

**Short-term accommodation of Hong
Kong English speakers towards
native English accents and the effect
of language attitudes**

Wenling Cao

PhD

University of York

Language and Linguistic Science

September 2018

Abstract

Accommodation, also known as convergence, refers to a process whereby a speaker changes the way he or she speaks to be more similar to another speaker. This dissertation focuses on two themes: language attitudes and short-term accommodation. A study using the matched-guise method is conducted to examine Hong Kong people's attitudes towards British English, American English and Hong Kong English (henceforth HKE). Results suggest that after the handover British English is still rated as the most prestigious English variety in Hong Kong. HKE is also found to have a high level of acceptance in terms of social attractiveness.

For short-term accommodation, two studies are conducted to investigate the phonetic convergence of HKE speakers towards native English accents, and the effect of language attitudes on convergence. Study 2 consists of a group of HKE speakers completing separate map tasks with a Received Pronunciation speaker and a General American English speaker. Their pronunciations of the THOUGHT vowel, the PATH vowel, rhoticity, fricative /z/ and fricative /θ/ are examined before, during and after the map tasks. The results suggest that the HKE speakers produce more fricative [z] and converge on rhoticity after exposure to the native accents. However, divergence is found on the PATH vowel and fricative /θ/, and maintenance is found on the THOUGHT vowel. These findings suggest that the HKE speakers tend to converge on the linguistic features which are more salient to them. Study 3 examines the effect of language attitudes on speech convergence, and no correlation is found between language attitudes and the HKE speakers' convergence on rhoticity.

Finally, the hybrid exemplar-based model is proposed to explain the complex results of the three studies. It provides a framework for speech accommodation which covers speech perception and production, and includes social factors as important elements in the model.

Table of Contents

Abstract.....	2
List of Figures.....	7
List of Tables	12
Acknowledgements	16
Author’s declaration	18
Chapter 1. Introduction	19
1.1 Aims of the research.....	19
1.2 Thesis structure	21
Chapter 2. Hong Kong English and language attitudes	25
2.1. Definition of attitude	25
2.2. Measurement of language attitudes	26
2.2.1. The matched-guise method.....	26
2.2.2. The verbal-guise method	27
2.2.3. Direct approaches	27
2.2.4. Implicit Association Test.....	28
2.3. Dimensions in the matched-guise and verbal-guise methods	28
2.4. Language attitudes and motivation in second language learning.....	29
2.5. Languages in Hong Kong.....	32
2.6. Hong Kong English	35
2.6.1. Phonological features of HKE	35
2.6.2. HKE as a new variety of world English?.....	37
2.7. Language attitudes in Hong Kong	41
2.7.1. Language attitudes in Hong Kong before and after 1997.....	41
2.7.2. Recent studies on language attitudes in Hong Kong	43
2.7.3. Recent studies of language attitudes and identity	47
2.8. Conclusion	49
Chapter 3. Study 1: Language attitudes towards British English, American English and Hong Kong English.....	51
3.1. Introduction	51
3.2. Research question and hypothesis	51
3.3. Experiment design.....	51
3.4. Results.....	54

3.4.1.	Results for accent evaluation	54
3.4.2.	Results for the questionnaire	66
3.5.	Discussion	71
3.5.1.	Status & Competence	71
3.5.2.	Social Attractiveness	72
3.5.3.	Factor 3	73
3.5.4.	Sex of speakers	73
3.5.5.	Attitudes and motivations	75
3.5.6.	Attitudes towards Hong Kong English	75
3.6.	Improvements	76
3.7.	Conclusion	76
Chapter 4.	Review of theories of short-term accommodation	77
4.1.	What is accommodation?	78
4.2.	Accommodation in sociolinguistics: Communication Accommodation Theory	79
4.3.	Accommodation in psycholinguistics: the interactive-alignment model	84
4.4.	Accommodation and exemplar-based theories	88
4.5.	Accommodation and L2 learning	97
4.6.	Summary	101
Chapter 5.	Review of studies of short-term accommodation and a pilot study	103
5.1.	Literature review of convergence studies	103
5.2.	Pilot study	112
5.2.1.	Experimental design	112
5.2.2.	Results	120
5.2.3.	Summary	127
5.3.	Improvements	129
5.3.1.	Variables	129
5.3.2.	Experimental design	130
5.3.3.	Normalization	131
5.4.	Conclusion	133
Chapter 6.	Study 2: Speech accommodation	135
6.1.	Research questions and hypotheses	135
6.2.	Experiment Design	138
6.2.1.	Participants	138
6.2.2.	Experiment design	139
6.2.3.	Materials	141

6.3. Data analysis	144
6.3.1. Data annotation and extraction	144
6.3.2. Vowel normalization	148
6.4. Results.....	148
6.4.1. Results of vowels.....	148
6.4.2. Results for consonants	186
6.4.3. The effect of talker sex	202
6.5. Chapter summary.....	207
6.6. Conclusions	213
Chapter 7. Study 3: Language attitudes and speech accommodation	214
7.1. Introduction	214
7.2. What is new in the present study?	214
7.3. Research questions and hypothesis.....	216
7.4. Experiment design	217
7.4.1. Participants	217
7.4.2. Recordings	217
7.4.3. Survey	217
7.4.4. Procedure	219
7.5. Results.....	219
7.5.1. Descriptive results of the survey.....	219
7.5.2. Linear mixed effects regressions	228
7.6. Speakers' profiles	232
7.7. Conclusion.....	245
Chapter 8. Discussion	246
8.1. Salience and short-term accommodation	246
8.2. The hybrid exemplar-based model for short-term accommodation.....	249
8.2.1. The existing models of speech accommodation	249
8.2.2. The hybrid exemplar-based model	250
Chapter 9. Conclusion	266
9.1. Scope of the research.....	266
9.2. Key findings	267
9.3. Contributions of the research to broader fields	269
Appendices.....	271
Appendix 1. Study 1: Accent evaluation form	271
Appendix 2. Study 1: Questionnaire of language attitudes	272

Appendix 3. Study 2: The maps used in the map tasks.....	274
Appendix 4. Study 2: Summary of rhoticity% for all participants	276
Appendix 5. Study 2: Summary of fricative [z] for all participants	277
Appendix 6. Study 2: Summary of fricative [θ] for all participants.....	278
References.....	279

List of Figures

Figure 2.1 The socio-educational model in Gardner (1985), copied from Gardner (1985, p.147).	31
Figure 2.2 The socio-educational model in Gardner (2006), copied from Gardner (2010, p.88).	32
Figure 2.3 Pie chart of languages spoken as mother tongue in Hong Kong (reproduced from Bacon-Shone et al., 2015, p.18).	33
Figure 2.4 Bar chart of languages spoken in Hong Kong (reproduced from Bacon-Shone et al., 2015, p.18).	34
Figure 2.5 The vowel space of HKE. Reproduced from Deterding et al. (2008, p.162).	35
Figure 3.1 Means of factor scores (from left to right: RP, GenAmE and HKE) for Status & Competence, Social Attractiveness and Factor 3.....	58
Figure 3.2 Means of ratings for Status & Competence for all the recordings.....	61
Figure 3.3 Means of ratings for Social Attractiveness for all the recordings.	62
Figure 3.4 Means of ratings for Factor 3 for all the recordings.	64
Figure 3.5 Proportions of English varieties spoken by teachers at three levels.....	66
Figure 3.6 Proportions of different English varieties used at and outside school.....	67
Figure 3.7 Proportions of different English varieties favoured in teaching and at work.	67
Figure 3.8 Proportions of judging which accent is more pleasant to speak or listen to..	69
Figure 4.1 The autonomous transmission account, taken from Pickering and Garrod (2004, p. 177).	85
Figure 4.2 The interactive-alignment model, taken from Pickering and Garrod (2004 p. 176).	86
Figure 4.3 A set of exemplars relating auditory/acoustic properties with category labels (taken from Johnson, 1997a, p.148).....	89
Figure 5.1 The left panel, reproduced from Deterding (1997, p.51), shows vowels produced by female RP speakers; the right panel, reproduced from Deterding et al. (2008, p.162), shows vowels produced by female HKE speakers.....	114
Figure 5.2 The LOT vowel and PATH vowel of RP and HKE based on Deterding (1997) and Deterding et al. (2008).....	115
Figure 5.3 The LOT and PATH vowels of GenAmE (data from Hagiwara, 1997), RP (data from Deterding, 1997) and HKE (data from Deterding et al., 2008).....	116

Figure 5.4 Procedure of the experiment design for the pilot study.....	117
Figure 5.5 Pictures used in the pre-task (top) and the post-task (bottom) for the LOT vowel.....	118
Figure 5.6 Maps used in the map tasks for the PATH vowel.	119
Figure 5.7 Participants' LOT vowel movements from the pre-task to the map task based on the transformed means of F1 and F2.....	121
Figure 5.8 Euclidean distance of the LOT vowel across the tasks for the RP and GenAmE group.....	123
Figure 5.9 Participants' PATH vowel movements from the pre-task to the map task based on the transformed means of F1 and F2.	124
Figure 5.10 Euclidean distance of the PATH vowel across the tasks for the RP and GenAmE group.	126
Figure 5.11 The THOUGHT vowel and PATH vowel of GenAmE (data from Hagiwara 1997), RP (data from Deterding, 1997) and HKE (data from Deterding et al., 2008).	129
Figure 5.12 Euclidean distances of eight HKE speakers across tasks and conditions..	133
Figure 6.1 The THOUGHT and PATH vowels of GenAmE (data from Hagiwara, 1997), RP (data from Deterding, 1997) and HKE (data from Deterding et al., 2001).	136
Figure 6.2 Experiment design of the formal study.	139
Figure 6.3 The environment of the recording booth of the experiments.	140
Figure 6.4 Map 1 for the native speakers.....	143
Figure 6.5 Map 1 for the participants.....	144
Figure 6.6 The actual locations of THOUGHT vowels for the HKE speakers, the RP speakers and the GenAmE speakers in Study 2, based on the results of Table 6.7.	149
Figure 6.7 F1 and F2 changes in THOUGHT vowels across the tasks in the two conditions.....	151
Figure 6.8 The shift of THOUGHT vowels across tasks and conditions.	152
Figure 6.9 Euclidean distance of the THOUGHT vowel across the tasks in the two conditions.....	155
Figure 6.10 Euclidean distance of the THOUGHT vowel from the pre-tasks to the map tasks in the two conditions.....	156

Figure 6.11 Euclidean distance of the THOUGHT vowel from the pre-tasks to the post-tasks in the two conditions.....	157
Figure 6.12 An example of judging convergence/divergence based on F1 and F2, and Euclidean distance.	159
Figure 6.13 The difference of pre-distance between the two exposure conditions for the THOUGHT vowel.	160
Figure 6.14 Euclidean distance of the THOUGHT vowel across tasks and conditions for each participant.....	163
Figure 6.15 Individual shifts of THOUGHT vowels from the pre-task to the map tasks.	164
Figure 6.16 The actual location of the PATH vowel for the HKE speakers, the RP speakers and the GenAmE speakers in Study 2.....	167
Figure 6.17 F1 and F2 changes in PATH vowels across the tasks in the two conditions..	168
Figure 6.18 Shift of PATH vowels across tasks and conditions.....	169
Figure 6.19 Euclidean distance of the PATH vowels across the tasks in the two conditions.....	172
Figure 6.20 Euclidean distance of the PATH vowel from the pre-tasks to the map tasks in the two conditions.....	174
Figure 6.21 Euclidean distance of the PATH vowel from the pre-tasks to the post-tasks in the two conditions.....	175
Figure 6.22 An example of over-convergence.....	177
Figure 6.23 The difference of pre-distances between the two exposure conditions for the PATH vowel.	178
Figure 6.24 Euclidean distance of the PATH vowel across tasks and conditions for each participant.	179
Figure 6.25 Individual shifts of PATH vowels from the pre-task to the map tasks.. ...	180
Figure 6.26 Euclidean distance of the THOUGHT/PATH vowels between the native interlocutors and the HKE participants from the early stage to the late stage of the map tasks.....	185
Figure 6.27 Percentage of the presence of rhoticity across the three tasks and two exposure conditions.	186
Figure 6.28 The participants' percentage of rhoticity in the pre-tasks..	189
Figure 6.29 The participants' percentage of rhoticity across tasks and conditions.....	191

Figure 6.30 Percentage of production of fricative [z] across the three tasks.....	192
Figure 6.31 The participants' percentage of fricative [z] in their pre-tasks.....	194
Figure 6.32 Individual percentages of the production of fricative [z] across the three tasks.....	196
Figure 6.33 Percentages of accurate pronunciation of fricative [θ] across the three tasks.....	197
Figure 6.34 The participants' percentage of fricative [θ] in the pre-tasks.....	199
Figure 6.35 Individual percentages of the production of fricative [θ] across the three tasks.....	201
Figure 6.36 Boxplot of F1 of the THOUGHT vowel across tasks for the females and the males.....	203
Figure 6.37 Boxplot of F2 of the PATH vowel across tasks for the females and the males.....	203
Figure 6.38 Percentage of the presence of rhoticity for the female and male participants across tasks and conditions.....	205
Figure 6.39 Percentage of fricative [z] for the female and male participants across tasks.....	206
Figure 6.40 Percentage of fricative [θ] for the female and male participants across tasks.....	206
Figure 7.1 Boxplots of factor scores for RP, GenAmE and HKE (from left to right) across the three dimensions.....	224
Figure 7.2 Boxplots of scores for integrative and instrumental orientation towards British English (left) and American English (right).....	225
Figure 7.3 The correlation between Linguistic Attractiveness and rhoticity (top figure), and between Integrative Orientation and rhoticity in two conditions (bottom figure).....	231
Figure 7.4 HK20 and the native interlocutors' production of rhoticity across the timeline in the map task.....	234
Figure 7.5 Spectrograms of the rhotic "spot" and the non-rhotic "spot" produced by HK20 in her conversation with a GenAmE speaker.....	236
Figure 7.6 HK23 and the native interlocutors' production of rhoticity across the timeline in the map tasks.....	238
Figure 7.7 Spectrogram of "Bar2" (top) and "Bar3" (bottom) produced by HK23.....	239

Figure 7.8 Formant contours of “Bar2” (more rhotic, in black) and “Bar3” (less rhotic, in red) produced by HK23.	240
Figure 7.9 HK8 and the native interlocutors’ production of rhoticity across the timeline in the map tasks.....	242
Figure 7.10 Spectrogram and formant contours for “Farm2” (left) and “Farm3” (right).	243
Figure 8.1 Illustration of Stage 1 in the HEM.	252
Figure 8.2 Illustration of Stage 2 and Stage 3 in the HEM.....	253
Figure 8.3 Illustration of how different types of attitudes represented in the HEM.....	262

List of Tables

Table 2.1 Language use of Cantonese, Mandarin and English in Hong Kong (reproduced from Bacon-Shone et al., 2015, p.18).....	34
Table 2.2 Developmental cycles of new varieties of English (adapted from Kirkpatrick, 2007, p.32 and Groves, 2009, p.68).....	39
Table 2.3 Summary of Hansen Edwards (2015)'s conclusion for five groups of speakers from Hong Kong and their stages in Kachru's and Schneider's model (edited from Hansen Edwards, 2015, p.203-204).....	40
Table 2.4 Summary of ratings of British English and American English for the three dimensions in Cheng (2013).	45
Table 3.1 Recordings used in Study 1.	52
Table 3.2 Factor loadings of the rotated component matrix of 20 semantic differential traits for Study 1.....	55
Table 3.3 Mean factor scores and standard deviation (SD, in brackets) of RP, GenAmE and HKE for the three dimensions.....	58
Table 3.4 Mean factor scores and SD (in brackets) for the three dimensions.	59
Table 3.5 Post-hoc tests of recordings for factor scores of Status & Competence. * indicates a significant comparison.	60
Table 3.6 Post-hoc tests of recordings for factor scores of Social Attractiveness.....	62
Table 3.7 Post-hoc tests between two HKE recordings and two guise recordings.....	63
Table 3.8 Post-hoc tests of recordings for factor scores of Factor 3.....	63
Table 3.9 The higher rated variety in the comparison between RP and GenAmE for the bi-dialectal speakers.....	65
Table 3.10 Ranking of all the recordings in the three dimensions.....	65
Table 3.11 Proportions of places preferred to live and to work.....	69
Table 4.1 Summaries of propositions of SAT/CAT (edited from Giles et al., 1987, p.36).	82
Table 4.2 Principal distinctions (edited from Giles et al., 1991, p.11).....	83
Table 4.3 An example of conversation between a native English speaker (NS) and a non-native English speaker (NNS).	100
Table 4.4 Four types of implicit feedback extracted from Gass and Mackey (2007, p.181-182).....	100

Table 5.1 Examples of utterances used in XAB perceptual similarity test in Kim et al. (2011).	110
Table 5.2 LOT vowels and PATH vowels in RP, GenAmE and HKE.	113
Table 5.3 Overall mean and standard deviation (in brackets) of F1 and F2 across tasks and conditions for LOT vowels.	121
Table 5.4 Euclidean distance of the LOT vowel across the three tasks and the two groups.	123
Table 5.5 Overall mean and standard deviation (in brackets) of F1 and F2 across tasks and conditions for PATH vowels.	125
Table 5.6 Euclidean distance of the PATH vowel across the three tasks and the two groups.	126
Table 6.1 Comparisons between the pilot study and the present study.	135
Table 6.2 Summary of hypotheses based on variables.	137
Table 6.3 Language profile of all the participants.	138
Table 6.4 Variables of the formal study.	142
Table 6.5 Details of data analysis of the variables.	145
Table 6.6 A summary of each participant's pronunciation of the five variables.	147
Table 6.7 Overall mean and standard deviation (in brackets) of F1 and F2 across tasks and conditions for THOUGHT vowels.	150
Table 6.8 Post-hoc tests for F1 and F2 of THOUGHT vowel across the tasks in both RP and GenAmE condition.	153
Table 6.9 Euclidean distance of the THOUGHT vowel across the three tasks and the two conditions.	154
Table 6.10 The post-hoc tests results of Euclidean distance for task*exposure of the THOUGHT vowel.	156
Table 6.11 The post-hoc tests results of Euclidean distance for task*exposure of the THOUGHT vowel.	157
Table 6.12 Summary of predictions and actual results for the THOUGHT vowel.	165
Table 6.13 Overall mean and standard deviation (in brackets) of F1 and F2 across tasks and conditions for PATH vowels.	168
Table 6.14 Post-hoc tests for F1 and F2 of PATH vowel across the tasks in both RP and GenAmE condition.	170
Table 6.15 Summary of the model comparisons for F1 of the PATH vowel.	171

Table 6.16 Euclidean distance of the PATH vowel across tasks and conditions. Euclidean distance of the PATH vowel across the three tasks and the two conditions.....	172
Table 6.17 The post-hoc tests results of Euclidean distance for task*exposure of the PATH vowel.....	174
Table 6.18 The post-hoc tests results of Euclidean distance for task*exposure of the PATH vowel.....	175
Table 6.19 Summary of the model comparisons for the two subsets of PATH vowel.	176
Table 6.20 Summary of predictions and actual results for the PATH vowel.....	182
Table 6.21 Euclidean distance of the native interlocutors' vowels at the early/late stage of the map tasks.....	184
Table 6.22 Summary of the logistic mixed effects models for rhoticity.....	187
Table 6.23 The post-hoc tests results for task of rhoticity with p values adjusted in the Bonferroni method.....	188
Table 6.24 Summary of the logistic mixed effects models for fricative [z].....	193
Table 6.25 Summary of the logistic mixed effects models for fricative [θ].....	198
Table 6.26 Summary of the model comparisons for task*sex.....	203
Table 6.27 Summary of the model comparisons for task*sex.....	206
Table 6.28 Individual changes from the pre-task to the map tasks in the THOUGHT vowel, the PATH vowel, rhoticity, fricative [z] and [θ] in the RP and GenAmE conditions.....	210
Table 6.29 The participants' accommodation patterns for the five variables at the individual level.....	211
Table 7.1 A summary of the differences in experiment design between Studies 1 and 3.	215
Table 7.2 The eight recordings used in Study 3.....	217
Table 7.3 The integrative and instrumental statements of British English and American English.....	218
Table 7.4 Accent identification rate for the eight recordings in Study 3.....	220
Table 7.5 Identification rates for all the recordings (overall identification%) and for native English recordings (native English identification%).	221
Table 7.6 Factor loadings of the rotated component matrix of the 19 semantic differential traits in Study 3..	222

Table 7.7 Mean factor scores and SD (in brackets) for RP, GenAmE and HKE in the three dimensions in Study 3.....	223
Table 7.8 Means of the integrative and instrumental statements for British English and American English.....	226
Table 7.9 Individual preferences of English variety/accent across the attitude tests....	227
Table 7.10 Summary of the model comparisons for rhoticity.	230
Table 7.11 HK20's profile of attitudes and accommodation of rhoticity.....	233
Table 7.12 An example of accommodation of rhoticity in the conversation between HK20 and the GenAmE interlocutor.	235
Table 7.13 HK23's profile of attitudes and accommodation of rhoticity.....	237
Table 7.14 An example of accommodation of rhoticity in the conversation between HK23 and the RP interlocutor.....	239
Table 7.15 HK8's profile of attitudes and accommodation of rhoticity.....	241
Table 7.16 An example of accommodation of rhoticity in the conversation between HK8 and the RP interlocutor.....	242
Table 8.1 The average number of tokens of input and output that the HKE speakers received and produced in the two map tasks for the five variables.	260

Acknowledgements

Ten years ago, back in 2008, I had my first phonetic lesson in Australia where I studied as an exchange student. It must be a miracle that I fell in love with phonetics at first sight! That was where my journey of linguistics began. In 2014, I continued my journey of linguistics in York. During these four years in York, I have received generous helps from numerous people. My gratitude to these people can hardly be expressed verbally, or phonetically.

I would like to thank my two supervisors, Professor Paul Foulkes and Dr Márton Sóskuthy. Paul is smart, nice and strict, he always has insightful opinions to the questions I was concerned about. Marci is clever, supportive and direct, he easily diagnoses the core issue of the data/problems and gives useful advice. Without their guidance and support, I could not have completed my PhD. From Paul and Marci, I have learnt how to become a decent and respectful scholar. I hope one day I will become a respectful scholar like them, who are passionate on their work, rigorous on research and acknowledged among colleagues and students.

I would like to thank my two examiners, Sam Hellmuth from York and Bronwen Evans from UCL. They gave me incredible feedback on my thesis and the discussion with them in the viva was very inspiring.

I would also like to thank the members of the Department of Language and Linguistic Science in York. Thanks go to Huw Llewelyn-Jones for technical assistance on the experiments, thanks to Tarmar Keren-Portnoy for the support as my Thesis Advisory Panel member. I am glad that I met so many great scholars in York: Sam Hellmuth, Richard Ogden, Dom Watts, Paul Kerswill, Marilyn Vihman and Amanda Cardoso, thank you for all your support during my PhD. I would also like to thank my PhD colleagues, Miyuki Kamiya, Sarah Kelly, Katherina Walper Gormaz, George Brown, James Tompkinson, Sarah Tasker, Grace Wood, Salina Cuddy, Jonathan Stevenson, as well as my Chinese PhD fellows Yiyun Ding, Niki Chen, Hongyan Zhao, Yuxi Gong, Wei-min Wu, Shanshan Lou and Theodora Lee. Thank you all for being by my side when I was overwhelmed by my PhD. In particular, I would like to thank Jiabo Li, Rebecca Dodgson

and my friends from WeThinker. Mr Li gave me enormous support during the darkest time of my PhD and Rebecca gave me lots of help and encouragement in my writing up year.

My gratitude goes to Professor Florian Jaeger and people from HLP lab at the University of Rochester. The spring I spent in Rochester in 2017 was one of the best memories during my PhD. I would also like to thank Professor Jette Hansen Edwards from CUHK who hosted me in summer 2017. My visit to CUHK was meaningful and fruitful, it has encouraged me to carry on my research on Hong Kong English.

I would also like to thank my participants who helped in my PhD project. Thank you, Adrianna Comeaux, Megan Jenkins, Hannah How, Allison Bennett and Salina Cuddy, without your time and efforts on the map tasks, the experiments would not have been completed. I also appreciate the helps from my Hong Kong participants, including those from York and the visiting students from CUHK.

A special thank you to Daniel Yee, who has witnessed both the angel and evil side of me, yet has not given up on me. From him I have learnt to become a better person, kind, humble and warm.

Finally, my sincerest gratitude goes to my dearest family, my mum Sammi Yu, my dad David Cao, my sister Lily Cao and my two nieces Phoebe Lin and Chloe Lin. Sammi and David, thank you for your unconditional love and support, I could not imagine better parents than you two. My family has given me endless support along my journey of linguistics, both emotionally and financially. I was standing on their shoulders, that's why I could become who I am today.

On 26th September 2014, my flight from Hong Kong landed in Manchester airport and I was ready to start my PhD in York. On the same day, the Umbrella Movement started in Hong Kong. As a researcher who works on Hong Kong topic, this event has a huge impact on my research career. I hope my research on Hong Kong English will go some way to pay it back to society and people in Hong Kong.

PhD is not the end of the journey. My journey of linguistics continues...

Author's declaration

I declare the content of this thesis to be my own work and that all other content from outside sources has been appropriately referenced. This work is original and has not been previously submitted for an award at any institution. Some parts of this thesis have been presented elsewhere at conferences.

The findings from chapter 3 have been presented at the following conference:

Cao, W. L. (2017). The Future of Hong Kong English: A Language Attitude Perspective. *UKLVC11*. Cardiff, 29-31 August.

Some of the findings from chapter 6 have been presented at the following conference:

Cao, W. L., Foulkes, P & Soskuthy, M. (2017). Short-term accommodation of non-native speakers: a perspective of second language acquisition. *50th Annual Meeting of the Societas Linguistica Europaea*. Zurich, 10-13 September.

Chapter 1. Introduction

Accommodation, also known as convergence, refers to a process whereby a speaker changes the way he or she speaks to be more similar to the other speaker. Accommodation can occur at the syntactic level (Pickering & Ferreira, 2008) or at the phonetic level (Babel, 2012; Pardo, 2006). This dissertation investigates the phonetic accommodation of Hong Kong English speakers towards native accents in a conversation, and the effect of language attitudes on accommodation. This chapter provides an overview of the dissertation, including the aims of the research and the thesis structure.

1.1 Aims of the research

Accommodation is a common phenomenon in our daily life. Many studies have been done on convergence from the perspectives of sociolinguistics, psycholinguistics and cognitive science. For example, Babel (2012) found that American college students converged towards two model talkers in a shadowing task, especially on the /æ/ vowel. Pardo (2006) also found that speakers in a conversation converged towards each other after 40 minutes' talking. While most of the convergence studies focus on native speakers, Kim, Holton and Bradlow (2011) compared the convergence between native speakers and between native and non-native speakers. Their results suggested that the pairs of talkers who shared the same first languages (henceforth L1) showed convergence, whereas the pairs consisting of native and non-native speakers did not converge. This finding leads to one question, do non-native speakers converge? Intuitively, accommodation should not only be limited to native speakers and theories of accommodation do not imply this either; however, a lack of studies of convergence between native and non-native speakers leaves this question unanswered. Therefore, **the first aim of the project is to fill this gap, to examine the convergence of non-native English speakers**, specifically, Hong Kong English (henceforth HKE) speakers. To achieve this goal, a study of speech accommodation is designed to examine the convergence between HKE speakers, Received Pronunciation (henceforth RP) speakers and General American English (henceforth GenAmE) speakers in a conversational setting (see Chapters 5 & 6). The implications of this study are threefold. Firstly, it will complement the research on convergence and examine the theories of accommodation with different subjects. Secondly, as a result of globalization, people do not only interact with native speakers of their own language, but also people from diverse linguistic

backgrounds. This study will help us to better understand the communication between people with different language backgrounds, which may further facilitate other fields of study, such as dialect change and human-machine interaction. Thirdly, this study will also contribute to the development of a new model for short-term accommodation (see Chapter 8), as the existing models of speech accommodation could not explain the complex results found in the study.

The second aim of the project is to explore the role of attitudes in convergence. Communication Accommodation Theory (henceforth CAT: Giles, Coupland & Coupland, 1991) suggests that people converge to shorten their social distance from the other speaker. There is some empirical evidence to support this hypothesis. For example, Babel (2012) found that the higher the female participants rated a white model speaker, the more they would converge to the model speaker. Similarly, Pardo, Gibbons, Suppers and Krauss (2010) found a significant correlation between the ratings of closeness between college roommates and their convergence after they had cohabitated for 3.5 months. However, the effect of attitude is not very robust, as in Babel (2012) a significant effect of attitude was not found for the other model speaker. Therefore, this project is also interested in the impact of attitudes on convergence. The present study investigates the convergence of non-native speakers, as the effect of attitudes on non-native speakers' convergence remains unknown. To achieve this, a study of language attitudes is conducted to investigate whether HKE speakers' attitudes towards British English and American English affect their convergence. This study will not only test the proposal of CAT but also help us to understand whether and how social factors interfere with convergence.

Besides these two studies, an additional study focusing on Hong Kong people's attitudes towards different English varieties is conducted. The study provides a broad background for the study of language attitudes and accommodation. It also contributes to the ongoing debate about the status of HKE in Hong Kong.

The third aim of the project is to develop a preliminary model for short-term accommodation based on a series of exemplar theories (Johnson 1997, Pierrehumbert 2001, 2003). CAT and the interactive-alignment model are the two prevalent models for speech accommodation. CAT explains why people accommodate while the interactive-

alignment model suggests how accommodation occurs, however, these two models only explain some parts of speech accommodation. The hybrid exemplar-based model (henceforth the HEM) is proposed, aiming to provide a framework for speech accommodation. The HEM proposes that speakers would update their distributions of sounds based on the input they receive from the interlocutors, and the selection for the production goal would be constrained by articulatory difficulty and would be interfered by social factors. The HEM explains how accommodation occurs from perception to production, and how social factors affect the process of accommodation.

1.2 Thesis structure

To achieve the above aims, three studies are conducted: (1) a study of language attitudes of Hong Kong people; (2) a study of speech accommodation of HKE speakers towards native English accents; (3) a study of the effect of language attitudes on speech accommodation.

There are nine chapters in this dissertation. **Study 1** is included in Chapters 2 and 3, and investigates Hong Kong people's language attitudes towards British English, American English and HKE. **Study 2** is reported in Chapters 4, 5 and 6, and examines the short-term convergence of HKE speakers towards RP and GenAmE in a conversational setting. **Study 3** is reported in Chapter 7 and investigates the correlation between the HKE speakers' language attitudes and their performance of convergence. Chapter 8 discusses the results of these studies and proposes a new model for speech accommodation. Chapter 9 draws some conclusions for the whole project.

1.1.1 Study 1: Language attitudes towards British English, American English and Hong Kong English

Chapter 2 is a literature review of language attitudes and HKE. Chapter 2 starts with a definition of language attitude and reviews a few research methods of language attitude studies, for example, the matched-guise method (Lambert, Hodgson, Gardner & Fillenbaum, 1960), the verbal-guise method (Cooper, 1975), and the Implicit Association Test (Greenwald, McGhee & Schwartz, 1998). Two motivations in the socio-educational model (Gardner, 1985), i.e. integrative orientation and instrumental orientation, are also discussed, as these two orientations were found to be relevant to learners' performance

in L2 learning. Besides the theoretical concepts of language attitude, Chapter 2 also reviews the language landscape in Hong Kong, and the phonological features and status of HKE. These contexts provide a background for Studies 1 and 2 in the following chapters. Finally, a few studies of language attitudes towards British English and American English in Hong Kong are reviewed.

Chapter 3 reports a study of language attitudes using the matched-guise method. Previous literature suggested that Hong Kong people preferred British English during the British colonial period, but later research has found a higher acceptance of American English (Cheng, 2013; Zhang, 2009). Therefore, Study 1 aims to provide an updated investigation of Hong Kong people's language attitudes towards British English, American English and HKE. 107 Hong Kong participants completed Study 1, giving judgements on 20 semantic differential traits (e.g. politeness, formality, intelligibility etc.) of eight recordings. Two bi-dialectal speakers read *The North Wind and the Sun* in both an RP accent and a GenAmE accent. By comparing the participants' ratings of their RP samples with the ratings of the GenAmE samples, the Hong Kong participants' preferences of English accent can be gained. Two recordings of HKE with different degree of accentedness were used to elicit the Hong Kong participants' attitudes towards HKE. Besides the accent evaluation task, the Hong Kong participants also answered 16 questions relating to their background and experience of English learning.

1.1.2 Study 2: Speech accommodation

Chapter 4 is a literature review of theories of short-term accommodation. It aims to provide a comprehensive review of theories relating to speech accommodation from different angles. Firstly, **CAT** (Giles et al., 1991) is reviewed. CAT is a sociolinguistics theory which suggests that convergence is a strategy that people adopt to lessen the social distance between interlocutors. Secondly, **the interactive-alignment model** proposed by Pickering and Garrod (2004) in the early 2000s is reviewed. It suggests that speakers and listeners build up the dialogue as a joint activity and align at different levels of linguistic representations. These alignment processes are automatic and direct. Thirdly, **the exemplar-based models** are reviewed, including the perception account (Johnson, 1997a), the production account (Pierrehumbert, 2001, 2003) and the sociophonetic variation account (Docherty & Foulkes, 2014; Foulkes & Docherty, 2006). The exemplar-based models provide a framework to explain how accommodation occurs. In addition to

these theories, **the Interaction Hypothesis** (Gass & Mackey, 2007, 2015) is also reviewed from the perspective of L2 learning.

Following Chapter 4, Chapter 5 reviews a few empirical studies of convergence from the past 20 years. This part of the literature review mainly focuses on the studies conducted by Molly Babel and Jennifer Pardo. Babel conducted a series of studies (Babel, 2010, 2012) on phonetic convergence using the shadowing paradigm (Goldinger, 1998), while Pardo investigated convergence in a conversational setting using the map task (Pardo, 2006; Pardo, Jay and Krauss, 2010). Chapter 5 also reports on a pilot study for Study 2 (see Chapter 6). The pilot study contained 6 HKE participants. Half of the participants completed a map task with an RP speaker, and the other half completed the same task with a GenAmE speaker. A pre-task and a post-task were also conducted before and after the map task. Acoustic measurements of the changes in the LOT vowel and the PATH vowel from the pre-task to the map task and from the pre-task to the post-task were taken.

Chapter 6 reports on a study of speech accommodation of HKE speakers. This chapter aims to answer one of the key research questions in the dissertation: **do Hong Kong English speakers accommodate towards native accents after a short-term exposure to the native accents?** The chapter starts with a comparison between the pilot study and Study 2, following with the experiment design and data analysis for Study 2. The experiment design of Study 2 follows the paradigm of a pre-task, the map task and a post-task used in the pilot study. 19 HKE speakers were recruited for Study 2. Each participant completed the experiment with an RP speaker, and after 3-4 weeks the participant repeated the experiment with a GenAmE speaker. Two vocalic variables (i.e. the THOUGHT vowel and the PATH vowel) and three consonantal variables (i.e. rhoticity, fricative [z] and fricative [θ]) were chosen as target sounds. In the results section, the HKE participants' performance on each variable are reported at both the individual and group levels.

1.1.3 Study 3: Language attitudes and speech accommodation

Chapter 7 investigates the correlation between language attitudes and speech accommodation. The same HKE participants from Study 2 were used in Study 3. They completed an accent evaluation task similar to the one used in Study 1 that aimed to elicit their attitudes towards British English and American English. In addition, the HKE

participants' integrative and instrumental orientations towards British English and American English were elicited. These different dimensions of attitudes were used to predict their performance of convergence in Study 2. Chapter 7 also includes an analysis of individual speaker. Three examples are chosen from 19 HKE speakers, which provide deep insights into the conversation between the HKE speakers and the native speakers. Three particular episodes of their conversation are extracted to demonstrate how convergence occurs.

Finally, Chapter 8 discusses a few key findings in Study 2 and Study 3. It suggests that people tend to converge on the linguistic features which are more salient to them. Evidence from Study are provided to support this argument. Secondly, a full explanation of the HEM is provided, including Stage 1 – perception, Stage 2 – update of distributions and Stage 3 - production. Findings from Study 2 and Study 3 are used to support the claims of the HEM. Finally, a few proposals are made for future studies.

Chapter 2. Hong Kong English and language attitudes

This chapter will review language attitudes in Hong Kong, aiming to provide an overview of the status of different varieties of English in Hong Kong from the 1990s to today. Firstly, theories and methods of language attitude research will be reviewed, covering everything from the matched-guise method to the Implicit Association Test. Secondly, the linguistic landscape in Hong Kong will be introduced, followed by a debate of whether HKE is a new variety of world English. Finally, a few studies on language attitudes in Hong Kong will be reviewed, dating from an early 1990s study by Bolton and Kwok (1990) to the most recent studies on identity and language attitudes by Hansen Edwards (2015, 2016a).

2.1. Definition of attitude

Social psychology has been studying attitude since the early twentieth century. Allport (1935, p.810) defined an attitude as “a mental and neural state of readiness, organised through experience, exerting a directive or dynamic influence upon the individual’s response to all objects and situations with which it is related”. Another definition of an attitude comes from Edwards, who suggested that “an attitude is the degree of positive or negative affect associated with some psychological object” (1957, p.2). Edwards’ definition only involved associated feelings, while Allport went one step further, suggesting that an attitude could also influence people’s behaviour towards the attitude object.

A more recent definition of an attitude is the ‘three-components’ model, where an attitude consists of **cognitive**, **affective** and **behavioural** components. The three components are derived from ancient philosophical paradigms. According to Eagly and Chaiken (1993, p.10), the cognitive component concerns thoughts and beliefs that people have about the attitude object; the affective component consists of feelings or emotions that people have in relation to the attitude object; and the behavioural component encompasses people’s actions or a behavioural intention with respect to the attitude object. It is worth noting that people do not necessarily respond to an attitude object in respect of all three components. For example, people who believe that HKE represents part of a Hong Kong

identity (cognitive component) do not necessarily speak HKE (behavioural component) or have a positive feeling towards HKE (affective component).

The attitude object can be a group of people like Asian immigrants, or an ideology like capitalism, or an event like learning French. When the attitude object is a language, we are able to learn about people's attitudes towards a specific language. If we adopt the three-component model above, a language attitude consists of the cognitive aspect of a language, for example, "I believe that British English is spoken by well-educated people in Hong Kong"; the affective aspect, "I like the Received Pronunciation accent"; and the behavioural aspect, "I would like to speak British English."

2.2. Measurement of language attitudes

Like identity, attitude is a hypothetical construct which is not directly observable but can be inferred from observable responses (Eagly & Chaiken, 1993, p.2). The following will review a few common methods of measuring language attitudes, including the matched-guise method, the verbal-guise method, direct approaches and the Implicit Association Test. Greater focus is given to the matched-guise method.

2.2.1. The matched-guise method

The matched-guise method introduced by Lambert, Hodgson, Gardner and Fillenbaum (1960) is the most commonly used indirect method. It consists of passages of text read aloud in two or more languages/accents by a bi-dialectal speaker. Without being told, respondents believe that they are listening to passages being read by two different individuals. Instead of Likert scales, the matched-guise method uses semantic differential traits that originate from Osgood, Suci and Tannenbaum (1957) to elicit quicker responses and to minimise mental processing. Semantic differential traits contain two opposing semantic labels such as friendly/unfriendly at each end, and a 7-point (sometimes 4-point or other options) trait in-between. Respondents are asked to judge the speaker's personality based on these labels and scales. In this way, the respondent's private attitudes are elicited and the effects of speaker voice and social desirability bias are minimised. Using bi-dialectal speakers is designed to control for features like intonation, speech rate and voice quality across accents. Respondents can then focus on the pronunciation of the accents themselves. The matched-guise method allows a fair

degree of cross-study comparisons, because many studies share similar dimensions such as *status* and *social attractiveness*.

The matched-guise method also has its limitations. First, the authenticity of varieties produced by bi-dialectal speakers is usually questioned. The bi-dialectal speaker is normally a native speaker of one variety and able to mimic another variety. The imitated variety may therefore be perceived as ‘odd’ or ‘inauthentic’ by speakers from that community. Second, repeating a reading passage may “exaggerate the language variations and make them much more salient than they would normally be outside the experimental environment” (Garrett, 2010, p.57). Third, respondents may not identify varieties as the one researchers want to present, especially for non-native speakers of that language. It would be even harder to identify regional dialects, which sometimes are perceived as ‘non-standard’ or ungrammatical.

2.2.2. The verbal-guise method

When a study involves more than two dialects, it is difficult to find multidialectal speakers who can produce native-like accents for all the varieties. In this case, the verbal-guise method can be used. The verbal-guise method is a variant of the matched-guise method (Cooper, 1975). Instead of using the same speakers, the verbal-guise method uses different speakers to represent the various accents. In this way, the mimicry effect (a criticism of the matched-guise method) is reduced. However, the verbal-guise method might introduce many other paralinguistic variables caused by speaker differences, such as pitch and voice quality. The verbal-guise method is also criticized for using repeated passages (Cooper, 1975, p.6).

2.2.3. Direct approaches

Direct approaches involve using questionnaires and/or interviews to elicit people’s overt attitudes towards language varieties. Questionnaires usually contain statements relating to the status or solidarity aspects of the language under investigation, where respondents report their degree of agreement from Likert scales. Interviews, on the other hand, ask respondents to describe their feelings and opinions towards the language under investigation and towards the people who speak that language. Besides ‘what’, researchers are also interested in the question of ‘why’; that is, “Why would you

like/dislike that language?” These qualitative methods enable us to have a thorough understanding of individuals’ attitudes towards the language under investigation. However, it is difficult to quantify the results based on qualitative data.

2.2.4. Implicit Association Test

The Implicit Association Test technique (IAT: Greenwald, McGhee & Schwartz, 1998) is used to examine implicit attitudes. Implicit attitudes are defined as “introspectively unidentified traces of past experience that mediate favourable or unfavourable feeling, thought, or action towards social objects” (Greenwald & Banaji, 1995. p.8). In other words, implicit attitudes are out of people’s conscious awareness and control, and they are shown through people’s different reactions to objects. One of the reflections of implicit attitudes is reaction time. The IAT measures people’s reaction time towards an associate attribute (e.g. pleasant or unpleasant) and target categories (e.g. British English or American English). The shorter the reaction time that people associate a positive attribute to one category (e.g. pleasant and British English), the more positive an attitude towards the category they have. Because implicit attitudes are out of people’s awareness, the IAT is ideal to investigate people’s inhibited attitudes. One argument against the IAT is whether the difference in reaction time is actually a reflection of people’s different attitudes. For example, Karpinski and Hilton (2001) suggested that no significant correlation was found between IAT and explicit attitudes.

2.3. Dimensions in the matched-guise and verbal-guise methods

Principal Component Analysis (PCA) is a statistical method for variable reduction, which is widely used in the matched-guise and verbal-guise methods (Rindal, 2014; Van der Haagen, 1998; Zahn & Hopper, 1985). It calculates correlations within all the semantic differential scales and puts highly correlated scales together into a single component. Using PCA, we can reduce 21 semantic traits into three or four higher order dimensions which represent different aspects of attitudes. In principle, studies can be different in the number and labels of the ‘components’ they investigate. Respondents from various populations may interpret dimensions differently based on their cultural and social settings. However, such inconsistency in measurements can also cause difficulty when comparing results across studies. To improve on this point, Zahn and Hopper (1985) developed a standardized instrument of scales called the Speech Evaluation Instrument.

They measured 56 semantic differential scales using a large sample of 572 people. Based on PCA, 30 of the scales were selected as valid scales and three dimensions were investigated – *superiority*, *attractiveness* and *dynamism*. *Superiority* denotes a speaker's social status, intellectual achievement, and the speech characteristics of advantaged and educated members of society; *attractiveness* denotes the qualities of speakers and their speech which reflect both social and aesthetic appeal; *dynamism* involves social power, activity level and the self-presentational aspects of speech. Since then, many studies have chosen scales from Zahn and Hopper (1985). Some of them investigate the same three dimensions, but some of them use different labels or investigate new dimensions. For instance, Ladegaard (1998) selected 15 out of 30 semantic traits from Zahn and Hopper (1985) and renamed the three dimensions as *status and competence*, *personal integrity and social attractiveness*, and *linguistic attractiveness*. These renamed categories were later adopted by Rindal in her language attitude studies of English accents among Norwegian students (Rindal, 2010, 2014). Other studies such as van der Haagen (1998) used four categories: *status*, *dynamism*, *affect*, and *standard* based on the PCA results of Dutch learners of English.

Using the same dimensions can permit comparison of results from different studies. However, it should be noted that most of these early studies were based on Western cultures; dimensions investigated in these studies may therefore not be applicable to respondents from different cultural backgrounds. Bayard and Green (2002) examined cross-cultural variation based on data from 15 countries and ethnicities. They found three clear-cut dimensions (*competence*, *power* and *solidarity*) when 682 samples from six Western nations were included, whereas only two, less clear-cut, dimensions (*solidarity* and *power*) emerged from samples from Hong Kong, Singapore and Japan. This indicates that dimensions in language attitudes are not universal, PCA might generate different numbers of dimensions based on subjects. Dimensions might also contain different scales which imply various meanings.

2.4. Language attitudes and motivation in second language learning

The role of language attitudes and motivation in second language learning has been well studied by Gardner and his colleagues. One of their most influential outcomes is the socio-educational model (Gardner, 1985, 2006, 2010). As language attitudes and

motivation in second language learning are not the main purpose of the present study, only a brief review is provided below.

There are many versions of the socio-educational model. Here we only discuss the version in Gardner (1985) and Gardner (2006).

The socio-educational model (Gardner, 1985) suggests that second language acquisition is affected by the beliefs in the community concerning the importance and meaningfulness of learning the language, the nature of skill development expected, and the particular role of various individual differences in the language learning process (see Figure 2.1 below; from Gardner 1985, p.146-147). There are four different types of individual differences: intelligence, language aptitude, motivation and situational anxiety. Other factors such as language attitudes and personality would not directly affect the language performance, but they may affect performance indirectly through one of these variables.

In the socio-educational model (Gardner, 1985), motivation is defined as “the combination of effort, desire to achieve the goal of learning the language and favourable attitudes towards learning the language” (1985, p.10). Another concept in the socio-educational model is orientation, which refers to a class of reasons for learning a second language. Two orientations in language learning are distinguished: integrative orientation and instrumental orientation. Integrative orientation refers to a class of reasons that suggest that the individual is learning a second language in order to learn about, interact with, or become closer to the second language community (1985, p.54). Instrumental orientation refers to a class of reasons that suggest that the individual is learning a language for pragmatic reasons and has a desire to gain social recognition or economic advantage through knowledge of a foreign language (1985, p.11).

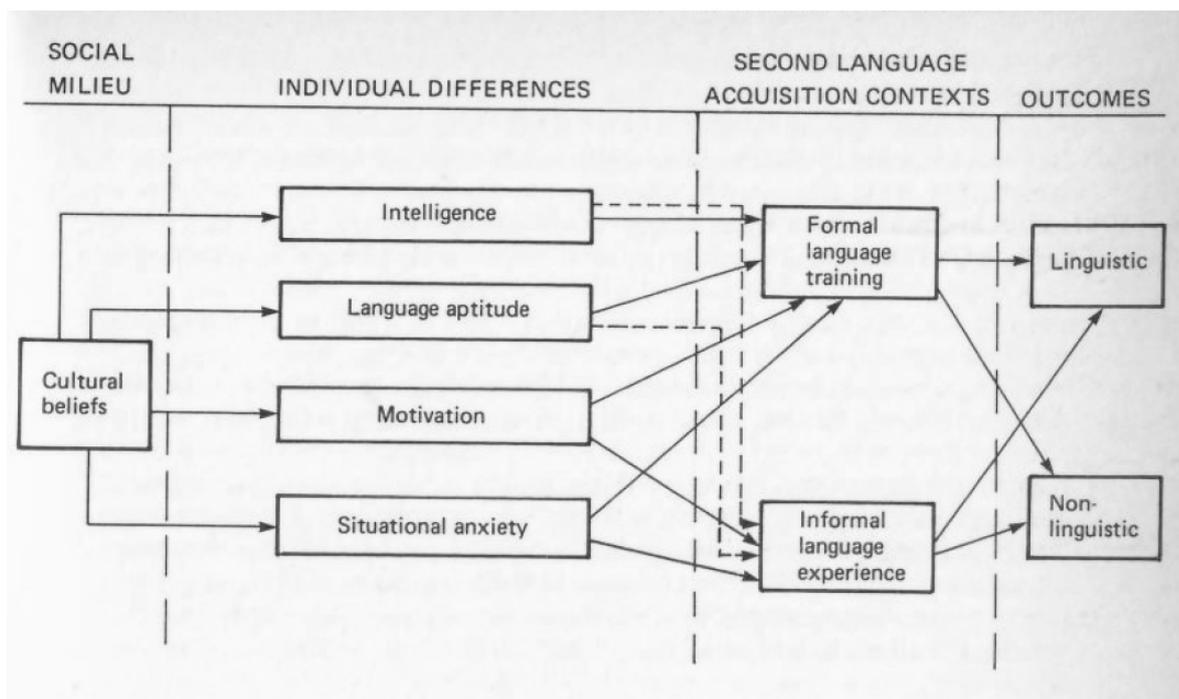


Figure 2.1 The socio-educational model in Gardner (1985), copied from Gardner (1985, p.147).

With more and more studies of language motivation in second language acquisition, Gardner updated the socio-educational model in 2006 (see Figure 2.2). For example, in Gardner (2006, 2010), *integrativeness* and *attitudes to learning situation* (not displayed in Figure 2.1) become foundations of motivation in second language learning. *Integrativeness* here consists of more than just integrative orientation. It refers to a genuine interest in learning the second language for the purpose of communicating with members of the other language community, and consists of integrative orientation (IO), interest in foreign languages (IFL), and attitudes towards the community speaking the target language (AFC).

Instrumentality, referred to as instrumental orientation in Gardner (1985), also supports motivation in Gardner (2006). However, its correlational effect is unstable, as some studies found that instrumental orientation did not influence a learner's performance whereas integrative orientation did.

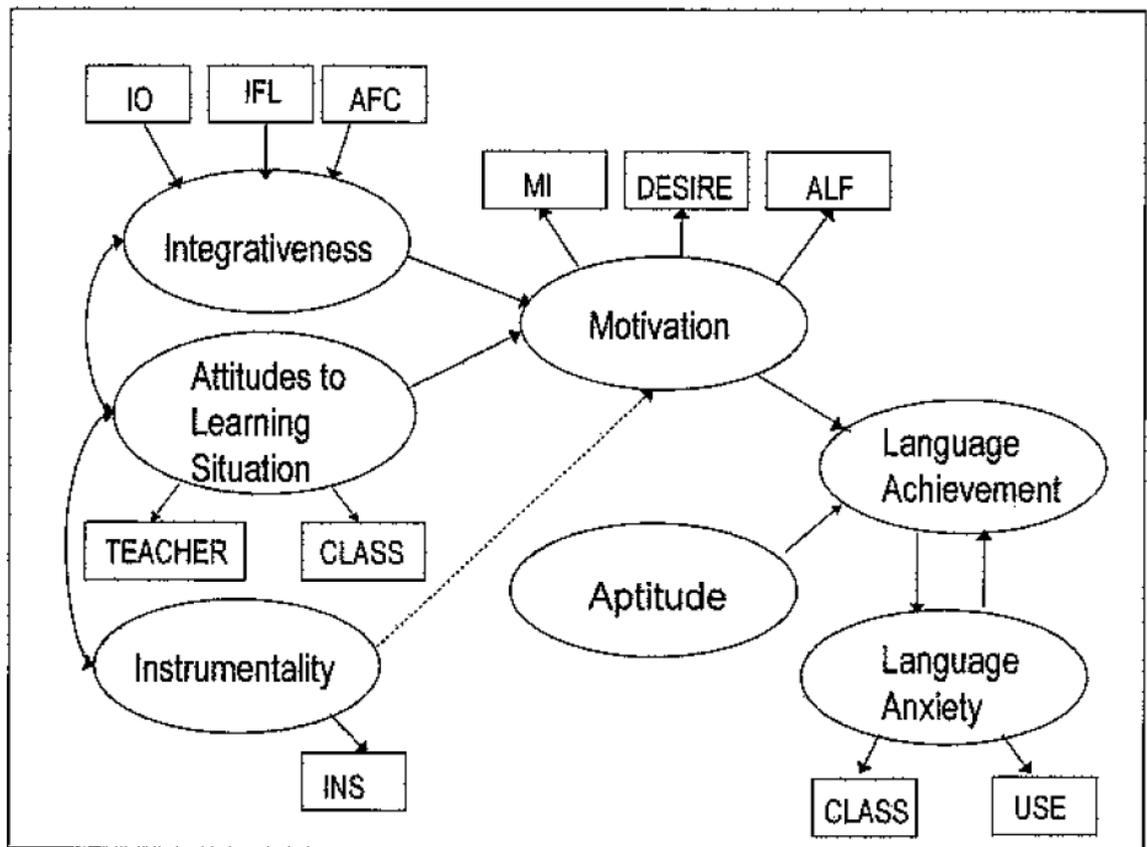


Figure 2.2 The socio-educational model in Gardner (2006), copied from Gardner (2010, p.88). IO represents integrative orientation, IFL represents interest in foreign languages, AFC represents attitudes towards the community speaking the target language, TEACHER refers to evaluation of the teacher while CLASS refers to evaluation of the course; MI represents motivational intensity, DESIRE represents desire to learn a language, ALF refers to attitudes towards the learning situation, CLASS at the right bottom corner refers to anxiety in language class while USE refers to anxiety in using the language.

2.5. Languages in Hong Kong

This section will present the linguistic landscape in Hong Kong, following by phonological features of HKE and a discussion of status of HKE.

Hong Kong, located in southern China, has been a Special Administrative Region (SAR) of China since the handover of sovereignty from the UK in 1997. The linguistic landscape of Hong Kong is very diverse. The three main languages used in daily life are Cantonese, English and Mandarin. Cantonese is the main language in Hong Kong, spoken as their usual language by 95.8% of the population (2011 Hong Kong Population Census: Census and Statistics Department, 2011, p.40), and widely used for communication in daily life. English is also an official language. It is spoken by 46.1% of the population and normally

used in government, business and tertiary education. As a result of the handover, Mandarin (or Putonghua), which is the official language in Mainland China, became a compulsory language course for primary and secondary students. In 1996, Mandarin was spoken as another language/dialect by 24.2% of the population. The figure increased to 46.5% in 2011 (1996-2011 Hong Kong Population Census: Census and Statistics Department, 1996, 2011).

Bacon-Shone, Bolton and Luke (2015) conducted a survey on language use, proficiency and attitudes in Hong Kong. They phone-interviewed 2049 citizens of Hong Kong, ranging from age 12 to over 80. In their survey, 89.1% of the interviewees claimed that Cantonese was their mother tongue, 4.7% claimed Mandarin and 0.6% claimed English, as shown in Figure 2.3. When they were asked what language(s) they knew, 99.6% of the interviewees said they knew Cantonese, 68% said Mandarin and 62.2% said English, as shown in Figure 2.4.

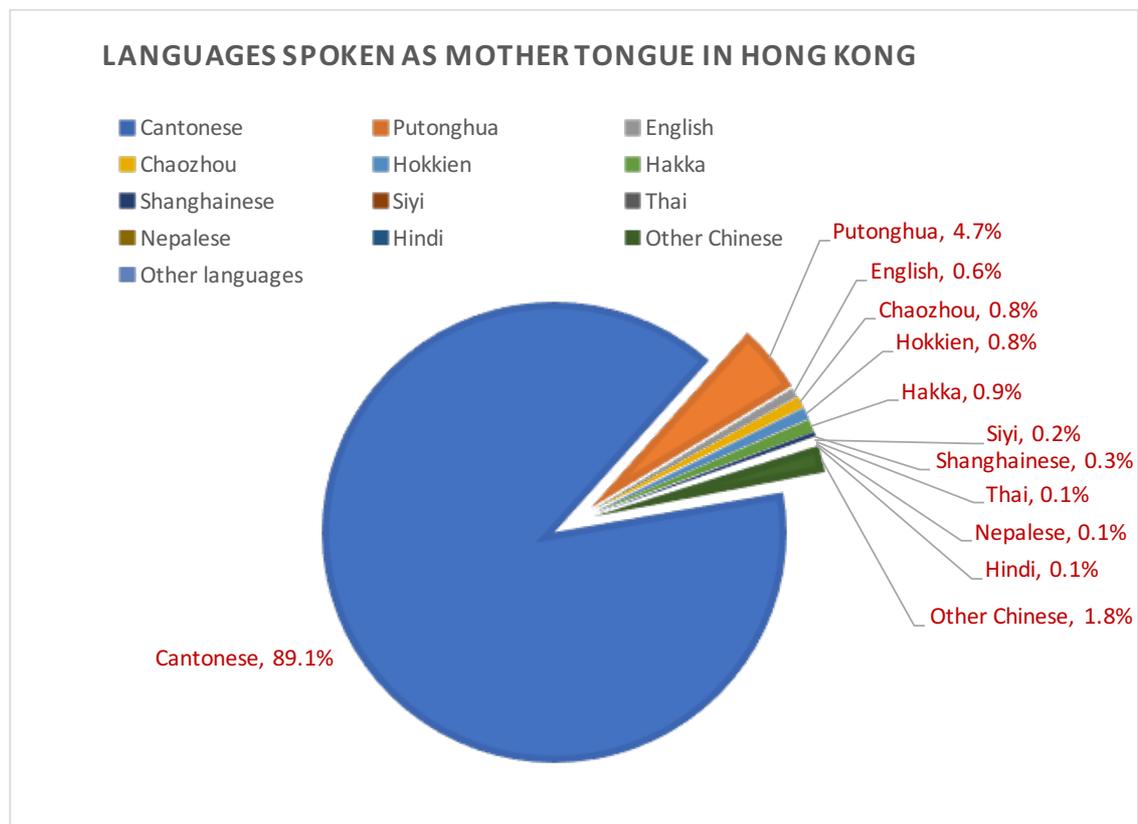


Figure 2.3 Pie chart of languages spoken as mother tongue in Hong Kong (reproduced from Bacon-Shone et al., 2015, p.18).

As shown in Table 2.1, Cantonese is the language that people use in their family, friends and work domains.

	Family members	Friends	Work colleagues	Work clients
Cantonese	97.4%	98.2%	97.3%	94.3%
Mandarin	6.6%	14.4%	15%	37.8%
English	10.9%	21.9%	33.2%	48.1%

Table 2.1 Language use of Cantonese, Mandarin and English in Hong Kong (reproduced from Bacon-Shone et al., 2015, p.18).

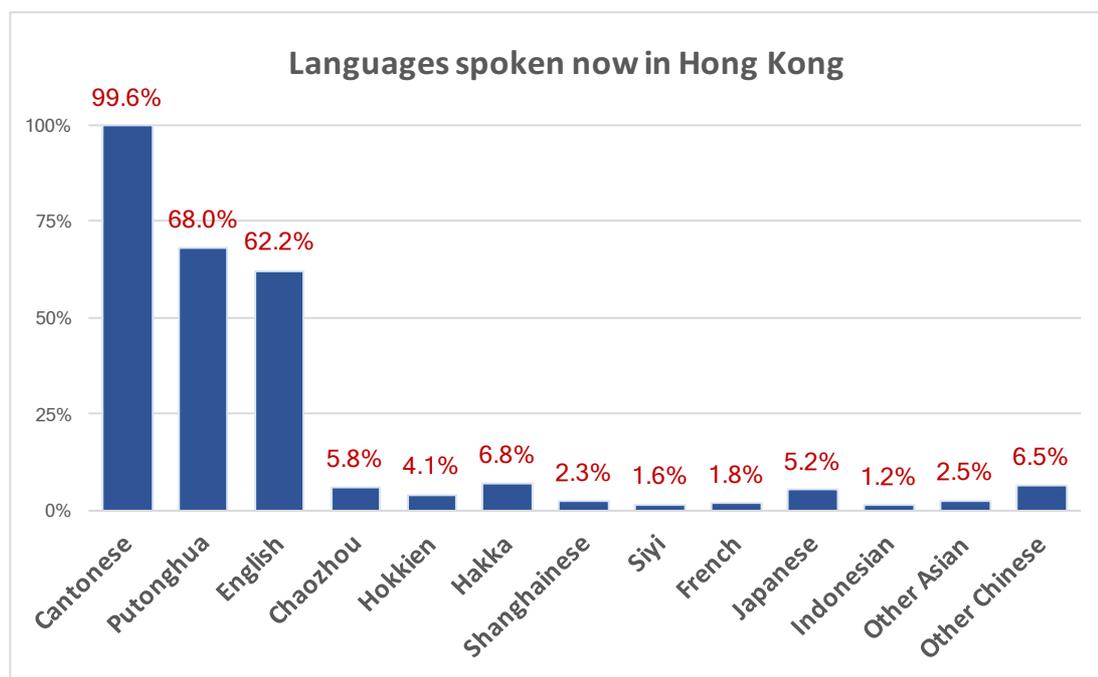


Figure 2.4 Bar chart of languages spoken in Hong Kong (reproduced from Bacon-Shone et al., 2015, p.18).

As suggested in Figure 2.3 and 2.4, though English is one of the official languages in Hong Kong and spoken by 62% of the interviewees, only 0.6% of them claimed English as their mother tongue. Therefore, English is a second language for most people in Hong Kong. Their English is very likely to be affected by their L1 (i.e. Cantonese), which will be discussed in the next section.

2.6. Hong Kong English

2.6.1. Phonological features of HKE

Due to its colonial history, HKE is traditionally closer to British English than American English in terms of phonological features. The following will briefly review some of the more prominent phonological features of HKE.

Vowels

Hung (2000) and Deterding, Wong and Kirkpatrick (2008) are two empirical studies of HKE. As the vowels in Hung (2000) were extracted from both male and female speakers but no normalization was reported, the monophthong report of Deterding et al. (2008) is adopted for the present study.

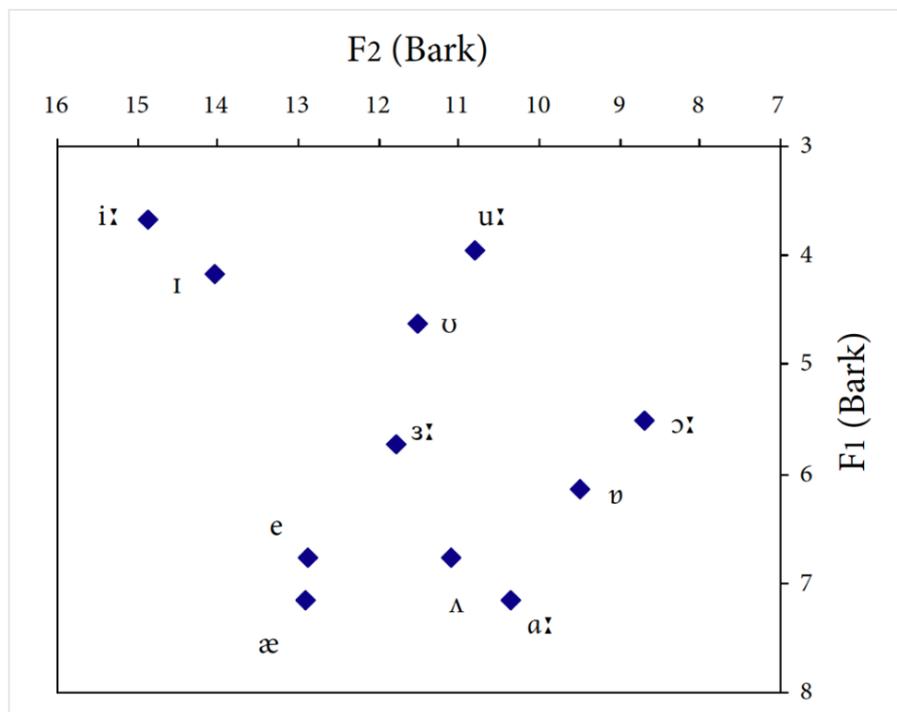


Figure 2.5 The vowel space of HKE. Reproduced from Deterding et al. (2008, p.162).

Deterding et al. (2008) measured a set of monophthongs based on the data of 15 female HKE speakers. Hung (2000) claims that four pairs of monophthongs (i.e. /i:/-/ɪ/, /e:/-/æ/, /u:/-/ʊ/ and /ɔ:/-/ɒ/) cannot be distinguished in HKE and should be treated as the same vowel phoneme. However, the mergers of these four pairs of vowels are not observed in Deterding et al. (2008).

Fricative /θ/

In HKE, the fricative /θ/ is sometimes pronounced as [f] in all word positions. For example, the word “think” is realised as [fɪŋk]. Deterding et al. (2008) also reported [t] as another variant.

Fricative /z/

In HKE, the voiced fricative /z/ is usually realised as [s] in all word positions. Hung (2008) indicates that HKE lacks a voiced/voiceless fricative contrast; therefore HKE speakers may have difficulty in distinguishing /f/-/v/, /s/-/z/, /θ/-/ð/ and /ʃ/-/ʒ/. This may be affected by Cantonese, which only has three voiceless fricatives: /f/, /s/ and /h/.

Final consonant cluster omission

Final consonant cluster omission refers to the phenomenon where the final stop of a cluster is usually omitted in HKE, for example, the word “land” is realised as [læn]. Note that a few studies show that deletion of /t,d/ is also found in native English varieties such as RP (Fabricius, 2002; Pavlík, 2017), American English (Bybee, 2002; Neu, 1980) and New Zealand English (Holmes & Bell, 1994). Hansen Edwards (2016b) found a 74% of /t,d/ deletion based on a HKE database (3009 tokens) of spontaneous conversations. Recent research tends to suggest that some degree of final stop deletion is a universal property for all varieties of English (Schreier, 2005).

L-vocalisation

L-vocalisation refers to the phenomenon that speakers use a vowel to replace the dark [ɨ] in a word, for example, *feel* is pronounced as [fi:u] instead of [fi:ɨ]. L-vocalisation is observed in HKE (Deterding et al., 2008; Hung, 2000). According to Hung (2000), the vocalisation of dark [ɨ] only occurs when the [ɨ] is preceded by a [-back] vowel and in the coda position. If the dark [ɨ] is not in the coda position, such as in *feeling* [fiɪŋ], no vocalisation is found.

When the dark [ɨ] is preceded by a [+back] vowel, deletion is observed. For example, *call* and *wall* are realised as [kɔ:] and [wɔ:].

2.6.2. HKE as a new variety of world English?

There has been a long debate on whether HKE is a newly emerged variety of world English or whether it is only the English that Hong Kong people speak with a Cantonese accent. The former suggests a recognised status of HKE, like Singapore English or Indian English, while the latter indicates that HKE is still in the process of being recognised and accepted by the local community. This section will briefly summarise the debates of the past two decades. Note that although the present study investigates HKE, the study has no intention to answer the question about the autonomy of HKE. Instead, the study focuses more on the language attitudes of Hong Kong people towards British English and American English. Reviewing the debates about the status of HKE should provide a background for the following section on language attitudes.

In early 1990s, linguists studying English in Hong Kong generally believed that the English spoken in Hong Kong was not an emerging new variety (Johnson, 1994; Luke & Richards, 1982; Tay, 1991), but rather a learner language. Debates and predictions on the future of HKE became prevalent before and after 1997, when the handover caused huge changes to education and language policy in Hong Kong. Bolton (2000) called for a re-examination of the issue. He argued that HKE is an emerging new variety, supported with solid evidence from HKE accent, vocabulary, history and literacy. Bolton's view was supported by Pang (2003).

To better understand the current status of HKE, the following section will review a few models of the development of new varieties of English. Sociolinguists and TESOL scholars have proposed different models of the developmental cycles of postcolonial English. Kachru (1992)'s development of non-native models, Moag (1992)'s constituent processes of the life cycle, Schneider (2007)'s dynamic models of the evolution of postcolonial English and Butler (1997)'s five characteristics are widely adopted in the studies of varieties of world English.

Kachru (1992, p.56) proposes a three-phase of development of non-native models. The initial stage is **non-recognition** of the local variety. At this stage, people tend to consciously imitate and identify with the native variety (i.e. British English in the case of Hong Kong) as it is associated with rulers from a higher social class. The second stage is the **co-existence of local and native varieties**. At this stage, the local variety still has a

lower status than the native one, but people would claim that the other person is using a localised English, for example 'Hong Kongized English'. The third stage is the **recognition** of the local variety. The local variety is slowly accepted as the norm.

Moag (1992, p.234) proposes five processes of the life cycle of a postcolonial variety of English: transportation, indigenization, expansion in use and function, institutionalization and restriction of use and function. **Transportation** means English is introduced and brought into a new environment with the intention of permanence, such as colonial administration. **Indigenization** indicates a process of language change where the local variety is distinct from the imported native variety and from other indigenized varieties elsewhere. For example, *beggar's chicken* in HKE refers to a dish in Chinese cuisine in which chicken is baked in lotus leaves and mud (Bolton, 2000, p.278). This phase is neither used nor understood by British English speakers or English speakers from Mainland China or Taiwan. **Expansion in use and function** refers to an extension of use of the local variety into education, the media and government services. **Institutionalization of the new variety** is represented by the localisation of English teachers, the widespread use of the local variety in the media and an emergence of creative writers using the local variety. **Restriction of use and function** is the last stage and only applies to some postcolonial areas such as Malaysia. In the case of Hong Kong, since 1998, Chinese has replaced English as the medium of instruction for most of the primary and secondary schools in Hong Kong. Note that Moag (1992)'s model is largely based on the case of Fiji. He also states that the processes are not fully consecutive; it is possible for overlaps between processes.

Butler (1997) proposes five principles of judging a variety of English. First, a standard and recognisable pattern of pronunciation passed from generation to generation. Second, specific vocabulary and phrases emerge to specify key features of the physical and social environment, and these words are regarded as peculiar to the variety. Third, a sense of history is developed, that due to the history of the language community, the variety becomes the way what it is now. Fourth, there is literature written in that variety of English without apology for 'mistakes'. Fifth, there are reference works such as dictionaries and style guides for that variety. This shows that instead of seeking for authority explanation from outside, people in that language community look to themselves for standardisation. Using this model, Bolton (2000) provided examples of

HKE accent, vocabulary, history and literacy, but he also admitted that at the time of his study there were no reference works on HKE.

Schneider (2007, p.33) suggests a five-stage dynamic model of the evolution of postcolonial varieties of English: foundation, exonormative stabilization, nativization, endonormative stabilization and differentiation. **Phase 1 (foundation)** refers to the initial stage of English being brought into a new territory where English was not used on a regular basis in the area before. **Phase 2 (exonormative stabilization)** refers to the stage where English is regularly spoken in the area and English is formally established as the language of administration, education and the legal system. **Phase 3 (nativization)** refers to a transition from accepting the imported native variety as the dominant language to increasing independence of the local variety, along with more contact and interaction between the two varieties. **Phase 4 (endonormative stabilization)** refers to the stage where the local variety is accepted as the norm in formal contexts. **Phase 5 (differentiation)** refers to the stage when the local variety becomes a new national language variety with a new national identity.

Groves (2009) tried to frame the development of HKE in the models of Kachru (1992), Moag (1992) and Schneider (2003, 2007), as shown in Table 2.2. She suggested that HKE is gaining more independence but is not yet a variety in its own right. Similarly, Sung (2015) also adopted Schneider’s model (2007) and concluded that HKE is in phase 3 (nativization) of Schneider’s model, supported by the fact that a large number of English L2 speakers were emerging in Hong Kong. Hansen Edwards (2015) suggested that HKE speakers are in different stages of acceptance of HKE, ranging from phase 2 (exonormative) to phase 4 (endonormative) in Schneider’s model.

<i>Scholar</i>	<i>Phases</i>				
<i>Kachru (1983)</i>	1 non-recognition	2 co-existence of local and imported varieties		3 recognition	
<i>Moag (1982)</i>	1 transportation	2 indigenisation			5 decline
		3 expansion in use	4 institutionalisation		
<i>Schneider (2003, 2007)</i>	1 foundation	2 exonormative stabilisation	3 nativisation	4 endonormative stabilisation	5 differentiation

Table 2.2 Developmental cycles of new varieties of English (adapted from Kirkpatrick, 2007, p.32 and Groves, 2009, p.68). The shaded areas indicate the possible stage(s) of the development of HKE.

	Group 1: Speak and accept HKE	Group 2: Speak HKE but only want to speak it in some contexts	Group 3: Speak HKE but do not want to speak it	Group 4: Unsure whether they speak HKE but do not want to speak it	Group 5: Do not speak HKE and do not want to speak HKE
Percentage	16.28%	7.81%	35.18%	24.75%	15.96%
Kachru's model	Stage 3: recognition	Stage 2: co-existence of local and imported variety	Stage 2: co-existence of local and imported variety	Stage 1: non-recognition	Stage 1: non-recognition
Schneider's model	Phase 4: Endonormative stabilization	Phase 3: Nativization	Phase 4: Nativization	Phase 2: Exonormative stabilization	Phase 2: Exonormative stabilization

Table 2.3 Summary of Hansen Edwards (2015)'s conclusion for five groups of speakers from Hong Kong and their stages in Kachru's and Schneider's model (edited from Hansen Edwards, 2015, p.203-204).

Hansen Edwards (2015) examined 307 tertiary institute students' language attitudes towards HKE and their cultural identity in October 2014. 30.61% of the respondents agreed that HKE was a real variety of English, 40.39% stated no and 28.99% stated they were unsure. For those who agreed that HKE was a real variety of English, they believed that HKE represented Hong Kong culture and their identity and it was a unique accent with its own unique vocabulary and identifiable features. For those who were against the autonomy of HKE, they believed that HKE was just English with a Cantonese accent with incorrect grammar and pronunciation, and that Hong Kong people were not 'native' speakers of English. 59.28% of the respondents stated that they spoke HKE, 15.96% stated they didn't speak HKE and 24.75% stated that they were unsure. Based on her data, Hansen Edwards (2015) summarised five different types of speakers of English and their corresponding stages using Kachru (1992)'s and Schneider (2007)'s models, as shown in Table 2.3.

Hansen Edwards (2015) conducted the research in September 2014. One year later, she repeated the study with 292 tertiary students. In the later study (Hansen Edwards, 2016a),

she found that the acceptance of HKE as a 'real' variety of English significantly increased from 30.61% to 42.80%.

In conclusion, though the debate about the development of HKE is still ongoing, most linguists agree that HKE is emerging as a new variety since more and more Hong Kong people are starting to accept and recognise HKE as a local variety which represents their identity.

2.7. Language attitudes in Hong Kong

2.7.1. Language attitudes in Hong Kong before and after 1997

There has not been much research on language attitudes towards different English accents in colonial Hong Kong. Bolton and Kwok (1990) found that over 65% of respondents chose British English as their model of English in a survey of attitudes towards Hong Kong accents, British accents and American accents. From this, it is reasonable to assume that British English was the dominant accent in Hong Kong before 1997. In education, around 10% of the schools at primary level and over 80% of the schools at secondary level used English as the medium of instruction (So, 1996). Local English teachers were either graduates from government-recognised universities or people who had a qualification in education. At that time they would have had opportunities to attend summer courses in Britain to improve their English proficiency. On the other hand, under the Expatriates English Teacher Scheme (EETS) many native-speaker teachers were recruited from Britain to support English teaching in Hong Kong. Boyle (1997, p.179) comments that:

“Whereas other places were concentrating on level of expertise rather than country of origin, Hong Kong was following a British-is-Best policy and through the British Council hiring all the teachers for the EETS from UK. [sic]”

Boyle explains that this phenomenon was consistent with the social trends in Hong Kong at that time: “the standard Hong Kong practice of hiring from the UK” could be observed in government offices and business sectors, which wanted British rather than American or Australian accents. British English, undoubtedly, was prevalent among Hong Kong people during the colonial period.

After the handover, the superiority of English diminished gradually. The most influential change was that Chinese became the medium of instruction in over 70% of the primary and secondary schools. From 1998, many primary and secondary schools changed their medium of instruction to Chinese. Only schools which satisfied the government's requirements for student ability, teacher capability and support strategies and programmes could continue to use English as their medium of instruction. Meanwhile, the Hong Kong government launched a 'biliteracy and trilingualism' educational policy. The 'trilingualism' policy aimed to ensure that students should be able to speak standard Cantonese, Mandarin and English, while the 'biliteracy' policy referred to students acquiring English and Chinese in standard writing. The influence of English, though still an official language in Hong Kong, has gradually reduced. To retain the quality of English teaching after the change of medium of instruction, a similar scheme called the Native-speaking English Teacher Scheme (NET) was launched in public primary and secondary schools. In the 2014/2015 school year, the NET Scheme recruited 455 native-speaking English teachers to local Hong Kong primary schools and 403 to secondary schools (Legislative Council Panel on Education, 2014).

Though these language policies have greatly affected the status of English in Hong Kong, English as an instrumental second language in Hong Kong still plays an important role in business and education. It is possible that British English still retains its prestige status since the influence of Britain did not fade away right after the handover. In the long term, the British accent may lose its advantage when students are exposed to other varieties of English accent through the NET, media and Internet. Moreover, American popular culture, spread through TV and movies, also overshadows the existing British colonial culture in Hong Kong. The people of Hong Kong's changing attitudes towards the American accent has been noticed by some members of the media. For example, the South China Morning Post (2013, October 29) reports that American English tutoring institutes in Hong Kong are popular among parents because they believe speaking with an American accent may give their children better job opportunities in the future. Another comment from South China Morning Post (Wordie, 2014) also states that the older generation mimics an upper-class British English accent, while young Hong Kong Chinese now adopt 'Valley Girl Speak', an accent from southern California in the USA.

All this evidence suggests that American English now affects Hong Kong more than it did before.

2.7.2. Recent studies on language attitudes in Hong Kong

Many recent studies of language attitudes in Hong Kong focus on the changes in young students' perception of English, Cantonese and Mandarin before and after the handover (Lai, 2001, 2005, 2011; Pennington & Yue, 1994; Poon, 2011) due to the change in medium of instruction. Lai (2005, 2011) investigated the first and second postcolonial generations' attitudes towards Cantonese, English and Mandarin using Gardner and Lambert's (1972) sociocultural model. The socio-educational model (Gardner, 1985) reviewed in section 2.5 is an extension of Gardner and Lambert's sociocultural model (1972). As mentioned above, integrative orientation refers to that class of reasons that suggest that the individual is learning a second language in order to learn about, interact with, or become closer to the second language community (1985, p.54); instrumental orientation refers to that class of reasons that suggest that the individual is learning a language for pragmatic reasons and has a desire to gain social recognition or economic advantage through knowledge of a foreign language (1985, p.11). Lai's studies (2005, 2011) conclude that the two generations all have the strongest instrumental orientation towards English, followed by Cantonese and Mandarin. In terms of integrative orientation, Cantonese and English received higher ratings than Mandarin. It is interesting to find that besides instrumental value, English also has a strong integrative value for postcolonial generations (2011, p.255). In addition, Lai (2005) observes that attitudes towards English have changed greatly from English being seen as a threat to Chinese identity in the 1980s (Pierson, Fu & Lee, 1980) to being an international language for communication right after the handover (Axler, Yang, Anson & Steven, 1998). It is now a marker of Hong Kong identity used to distinguish Hong Kongers from Mainland Chinese (Lai, 2005).

As already explained, British accents were prevalent in Hong Kong during the late colonial period, but not everyone agreed that Hong Kong people valued the British accent over other varieties of English. In the early 1990s, Bolton and Kwok (1990) argued that Hong Kong people had developed a separate 'Hong Kong identity' rather than a broader and more abstract 'Chinese' identity. Driven by the change in identity, Bolton and Kwok (1990) assumed that Hong Kong people would value Cantonese-accented English more

because the localised features of pronunciation would enable them to signal a ‘Hong Kong identity’. To test their hypothesis, they conducted a survey of attitudes towards HKE accents, two varieties of RP, and American accents on 131 undergraduate students using the verbal-guise method. Contrary to their hypothesis, 65.1% of the participants chose British native speakers when they were asked “When you speak English, who would you most like to sound like?”, whereas only 25.6% of the participants chose Hong Kong–accented English speakers and 6.2% of them chose American English speakers. This confirms that the British accent was still dominant at that time, but the results did not support Bolton and Kwok (1990)’s hypothesis. It seems too early to predict a ‘well-developed’ Hong Kong identity in 1990 when Hong Kong was still a colony of the United Kingdom. Also, unlike Singapore, where English is much more widely used, in Hong Kong, English is primarily used in education, government and business. Hong Kong people may choose Cantonese instead of Cantonese-accented English to signal their local identity. This explanation is also found in Candler (2001), where some secondary school students indicated that they would associate a Hong Kong accent in English with new immigrants, and they would choose Cantonese if they wanted to signal their identity.

A study by Luk (1998) provided stronger and more solid evidence for the prestigious status of British accents. Over 85% of the secondary school students in her study showed a preference for RP compared to a local HKE accent. Candler (2001) also found that over 50% of 81 secondary school students wanted to speak English with a British accent, while only around 22% of them wanted to speak with either American or HKE accents. From these results, we can conclude that Hong Kong students still favoured British accents over Hong Kong and American accents at that time.

Groves (2011) interviewed 140 undergraduate students from Hong Kong, asking them about the English variety they felt they should learn to speak and what English variety they spoke. 60% of the interviewees selected British English as the variety they should learn to speak, and 13.6% selected American English. In terms of what English variety they spoke, 74.3% of the interviewees selected HKE, 15.7% selected British English and 3.6% selected American English.

Results from a more recent study by Cheng (2013) also partly support the conclusion that the British accent is more favoured. In Cheng’s study, 21 young people from Hong Kong

completed a questionnaire which contained 22 statements with 5-point Likert scales. Cheng proposed three new categories of attitudes: the affective dimension, linguistic dimension and pragmatic dimension. The *affective dimension* “reflects and interacts with subjects’ preferences for the culture and other social aspects related to the linguistic variety”; the *linguistic dimension* relates to “whether the subjects believe that a certain linguistic variety is the standard, is more correct, more grammatical and purer”; and the *pragmatic dimension* deals with the instrumental value of a certain linguistic variety, such as if “the linguistic variety is useful for personal success in the modern world” (Cheng, 2013, p.3). Results showed that British English was preferred in the *linguistic* and the *pragmatic dimensions*, while American English was more favoured in the *affective dimension*. Based on the overall ratings, he claimed that young educated Hong Kong people might have changed their attitudes towards the two English varieties. However, one issue in his study is that seven out of his 21 respondents received their tertiary education in North America and the rest were educated in Hong Kong. If we look at the ratings of the local group, British English is rated overwhelmingly higher than American English for all three dimensions:

	Affective	Linguistic	Pragmatic	Overall
BrE	3.2857	4.0000	3.3036	3.5298
AmE	2.8929	2.2619	2.7321	2.6290
p	=.06	<.001	=.145	=.001

Table 2.4 Summary of ratings of British English and American English for the three dimensions in Cheng (2013). Data extracted from tables in Cheng (2013, p.12-23).

In other words, local young people from Hong Kong in his study actually favour British English. However, the very small sample size of his study can hardly said to be representative of the whole population.

Chan (2013) conducted a verbal-guise study with 71 young Hong Kong professionals, examining their attitudes towards native English varieties, such as British English, American English, Australian English, and non-native English varieties such as HKE, Chinese English, Indian English and Philippine English. Note that Chan (2013)’s division of native accents and non-native accents was based on Kachru’s (1985) three concentric circles. In his study, he found that over 55% of the participants could recognise HKE, 40.4% of them could recognise British English and 30.5% of them could recognise American English. British English was rated as the most appropriate accent in situations

such as teaching, news broadcasts, business meetings, job interviews, giving directions, discussions in class and chatting with friends who are non-native English speakers. American English was rated lower than British English in these situations. Moreover, British English was also rated highest in terms of perceived intelligibility.

Not all the studies agree that British English is preferred by people from Hong Kong. Zhang (2009) and Bayard and Green (2002) suggested that Hong Kong people prefer American accents.

Zhang (2009) investigated Hong Kong university students' attitudes towards eight English varieties using a verbal-guise method: Received Pronunciation (RP), Tyneside English, General American English, Australian English, Philippine English, Mandarin-accented English and two varieties of HKE. Each English variety was spoken by two female speakers. Participants rated each recording using a questionnaire of 21 semantic traits. The 21 semantic traits were categorised into two dimensions: *status* contained traits like *educated*, *intelligent* and *wealthy*, representing a dimension of social status and prestige in a linguistic variety; *solidarity* contained traits like *friendly*, *sincere* and *pleasant*, representing a dimension of social attractiveness. Results showed that overall ratings for General American English were higher than those for RP and for Tyneside English in both *status* and *solidarity*. However, when we look at the ratings for speakers of the same accent, huge intra-speaker differences were found. For instance, speaker RP 2 was rated second for both *status* and *solidarity*, but another speaker RP 1 was ranked 9th in *status* and 10th in *solidarity* out of 16. These ratings were even lower than one Mandarin-accented English speaker and one Hong Kong educated English speaker. For the American accent, speaker USA 2 was ranked top for both *status* and *solidarity*, whereas another American English speaker USA 1 was ranked 5th in *status* and 9th in *solidarity* out of 16 speakers. Such huge intra-speaker differences indicate that respondents' ratings in Zhang's study might be affected by factors other than the accent itself; for example, speakers' pitch, voice quality and speech rate.

Another verbal-guise study of attitudes towards different English accents is Bayard's project *Evaluating English Accents Worldwide* (Bayard & Green, 2002). 75 Hong Kong university students participated in a language attitudes survey of American, British, Australian and New Zealand English. Each variety was represented by one female and

one male voice. The survey contained three dimensions: *competence* which included the traits *reliable, hard-working, competent* and *ambitious*; *power* which included the traits *controlling, dominant, authoritative* and *assertive*; *solidarity* which included the traits *cheerful, friendly, warm* and *humorous*. Results showed that American accents regardless of speaker gender overwhelmed British accents for *status, power, solidarity* and *competence*. Greater differences between American accents and British accents were observed for *solidarity* and *competence*. Similar to Zhang (2009), Bayard also used the verbal-guise method to elicit listeners' attitudes and this may introduce interfering factors such as speaker differences. In particular, the recording text Bayard and Green (2002) used does not have a neutral content but is a personal letter from children to parents. Both speakers and listeners could be affected by the non-neutral content. My auditory impression is that the letter sounds more emotional when is read by the American speakers in Bayard and Green (2002)'s study. In addition, when participants were asked to identify the accent of the speakers, they could not identify the British accent but could successfully identify the American accent. This is contradictory to previous findings that Hong Kong participants had higher identification rates for British English than American English (Bolton & Kwok, 1990; Candler, 2001; Chan, 2013).

Though it may be too early to conclude that American English is favoured more than British English in Hong Kong, it is undoubtedly the case that some English speakers in Hong Kong speak English with an American accent. Deterding et al. (2008) found that six out of 15 English speakers they recorded in Hong Kong had clear American influences in their English accent. Hansen Edwards (2016c) also found that among the 68 English speakers of Hong Kong in her study, 63 had at least some American features in their pronunciation. The most prevalent American feature she found was rhoticity, where 52% of the elicited words were pronounced with rhoticity. More studies are needed to examine the influences of American English on HKE.

2.7.3. Recent studies of language attitudes and identity

A series of studies of language attitudes in Hong Kong were reviewed above. Though the review has focused more on the comparison of Hong Kong people's attitudes towards British English and American English, HKE has always been of interest to linguists in Hong Kong. They are especially interested in how much Hong Kong people accept HKE

and whether HKE is developing into a new variety of world English (section 2.7.2). One recent political event in particular has brought the topic of identity back to research in HKE: the Umbrella Movement of 2014. The Umbrella Movement is a pro-democracy movement that was partially student-led and lasted from 26th September to 15th December 2014. Tens of thousands of Hong Kong residents peacefully occupied major thoroughfares in Admiralty, Causeway Bay and Mong Kok, demanding universal suffrage for the 2017 election of Hong Kong's chief executive. A stronger Hong Kong identity was found among the younger generation after the Umbrella Movement. According to the Hong Kong University's Public Opinion Programme's surveys on People's Ethnic Identity (Public Opinion Programme, 2017), in a poll of late 2014, 59.8% of young respondents (ages 18-29) stated that they were a Hong Konger, and the number increased to 69.7% in a poll of late 2017. On the other hand, only 6.5% of the young respondents stated they were Chinese in the 2014 poll and this number decreased to 0.3% in 2017. The following will review recent studies by Hansen Edwards who tried to capture the changes in language attitudes and identity after 2014. As identity is not the main interest of the present study, this part of the review will be brief.

Hansen Edwards (2015) conducted a survey in September 2014 of Hong Kong students' language attitudes towards HKE and whether their cultural identity related to their language attitudes. 307 Hong Kong tertiary students were asked to select a cultural identity from *Hong Konger*, *Hong Kong Chinese* and *Chinese*. 61.24% of the students selected *Hong Konger*, 35.18% selected *Hong Kong Chinese* and 3.57% selected *Chinese*. The results for students' perceived cultural identity were consistent with the results of Hong Kong University's Public Opinion Programme. When these three groups of students were asked whether HKE is a real variety of English, 33% of the *Hong Konger*-selectors said yes, 25% of the *Hong Kong Chinese*-selectors said yes, and 45% of the *Chinese*-selectors said yes (note that there were only 11 students who selected their identity as *Chinese*). Hansen Edwards (2015) concluded that attitudes towards HKE were positive, as the majority of the students thought the "HKE accent is not a big deal as long as the speaker's English is understandable" (2015, p.201). On the other hand, they still preferred native varieties of English and thought HKE was not as good as the other native varieties.

Coincidentally, the survey which was used in Hansen Edwards (2015) was released on the same day that the Umbrella Movement started. This gave Hansen Edwards the chance to capture the change in people's attitudes towards HKE before and after the Umbrella Movement. She repeated the survey one year later with a different group of students who were matched for age and gender with the 2014 one. In Hansen Edwards (2016a), she found that for those who selected 'Hong Konger' as their cultural identity, 64% of them stated they spoke HKE, but in 2015 this number significantly increased to 79%. For those students who selected 'Hong Kong Chinese', in 2014 25% of them agreed that HKE was a real variety of English, but in 2015 the number significantly increased to 43%. As Hansen Edwards (2016a, p.163) admitted, it was not clear whether the change was due to the Umbrella Movement or as a result of a stronger Hong Kong identity among the younger generation, or a reflection of the changing linguistic and political landscapes in Hong Kong. What is certain is that the change itself is happening.

2.8. Conclusion

This chapter reviewed theories of language attitude, the phonological features and status of HKE, and studies of language attitudes in Hong Kong. Most of the studies agree that HKE is not yet an autonomous variety of English, as the acceptance of HKE is comparatively low among the people of Hong Kong. However, there is some evidence to indicate that acceptance is increasing (Hansen Edwards, 2015, 2016a), especially in recent years with more and more of the younger generation tending to mark their Hong Kong identity with HKE.

Previous studies have not reached an agreement on whether British English or American English is more prestigious or attractive in Hong Kong at this moment. Though most research seems to support the view that British English is still favoured by Hong Kong people (Bolton & Kwok, 1990; Candler, 2001; Chan, 2013; Cheng, 2013; Groves, 2011), there is also evidence to show that American English is spoken by some HKE speakers (Deterding et al., 2008; Hansen Edwards, 2015b) and American English is preferred to British English (Bayard & Green, 2002; Zhang, 2009). One reason could be that Hong Kong people's attitudes have changed over time and the function of English became different after the handover, as these studies were conducted at different times. Another reason could be the different methods these studies used. Except for Bolton and Kwok

(1990), studies supporting British accents used questionnaires, while those supporting American accents used the verbal-guise method. None of them use the matched-guise method (Lambert et al., 1960), the most common method in the study of attitudes.

The next chapter will present a study of Hong Kong people's attitudes towards British English, American English and HKE, using the matched-guise method.

Chapter 3. Study 1: Language attitudes towards British English, American English and Hong Kong English

3.1. Introduction

As mentioned in Chapter 2, before the change of sovereignty in 1997, Hong Kong people preferred British English due to the colonial history of Hong Kong. It has been more than 20 years since the handover, and Hong Kong has gone through a series of changes in its economics, political profile and linguistic landscape. Regarding the linguistic landscape, the use of Mandarin is undoubtedly increasing, while arguments about the status of HKE are also becoming more prevalent in academia.

This chapter presents a study of language attitudes, focusing on how Hong Kong people perceive British English, American English and HKE. The matched-guise method is used in the study. Since there have not been many studies of language attitudes in Hong Kong using the matched-guise method, the present study aims to fill this gap.

This chapter starts with the research question and then presents the experiment design, followed by the results for the matched-guise method. Besides the matched-guise results, the results of a questionnaire are also presented. Finally, the chapter ends with a brief discussion.

3.2. Research question and hypothesis

The present study is interested in one broad question: **What are Hong Kong people's attitudes towards British English, American English and HKE?**

Though no agreement has been reached on the question above, more studies have shown that British English is still preferred in Hong Kong, therefore, the hypothesis of the present study is that **British English is more prestigious and more attractive in Hong Kong than American English and HKE.**

3.3. Experiment design

3.3.1. Recordings

In total, eight recordings were used in this study. Four recordings were produced by two

bi-dialectal speakers: one male American speaker and one female British speaker. The male American speaker was 26 years old and came from California. He had been staying in the UK for 5 years, including 3 years of studying and 2 years of working, at the time of recording. The female British speaker was 20 years old and from southern England. She had received some training on RP and American accents before because of her interest in drama. The two bi-dialectal speakers were asked to read *The North Wind and the Sun* in both an RP accent and a General American accent. At least two trials for each accent were recorded, and four recordings were selected: one with the authentic RP accent and one with the guised American accent produced by the female bi-dialectal speaker; one with the guised RP accent and one with the authentic American accent produced by the male bi-dialectal speaker. Authenticity of their accents was examined by two experienced phoneticians from the Department of Language and Linguistic Science at the University of York.

Speaker	Recording	Sex
bi-dialectal RP speaker	RP	female
	GenAmE-guise	female
bi-dialectal GenAmE speaker	RP-guise	male
	GenAmE	male
HKE speaker with a strong accent	HKE-strong	male
HKE speaker with a mild accent	HKE-mild	male
RP speaker	RP (filler)	female
GenAmE speaker	GenAmE (filler)	female

Table 3.1 Recordings used in Study 1.

Two male Hong Kong students at the University of York read the same passage in their natural accent and produced two recordings for HKE. Note that one of the Hong Kong students spoke English with a strong Cantonese accent, while the other one had a mild Cantonese accent and his accent was strongly influenced by RP. As mentioned in Chapter 2, HKE is still a developing variety of English. It contains a wide range of speakers who produce English with different degrees of Cantonese accent. The two recordings here can only represent two types of HKE. Similarly, Zhang (2009) distinguished ‘broad HKE’ from ‘educated HKE’ in her study. For example, one of the broad HKE speakers she chose was a Chinese takeaway assistant who moved to the UK at the age of 30, whereas one of the educated HKE speakers had a master’s degree and was a manager in an interpreting department in the UK. However, the degree of accentedness is not necessarily related to a speaker’s educational background.

Two recordings of the same passage from a monolingual female American speaker and a female RP speaker were used as fillers. A summary of the recordings is shown in Table 3.1. All the recordings were used in an accent evaluation task in which listeners gave their judgements on 20 semantic differential traits.

3.3.2. Listeners and tasks

A total of 126 Hong Kong people participated in this study. They completed a survey containing an accent evaluation task after each of the eight recordings, followed by a questionnaire. Each recording lasted for about 1 minute, and the whole survey took around 20 minutes to complete. 19 respondents were excluded either because they spent less than 8 minutes on the survey or because they did not complete the survey.

Of the remaining 107 participants, 29 were secondary school students, 67 were college students and 11 were professionals. There were 76 (71%) male and 31 (29%) female participants. 25 of the participants (24%) had overseas experience of between 2 weeks and 20 years: 16 participants had visited the UK, 7 had visited North America, and 2 had visited Australia. Data of one participant who had spent 20 years in Australia was retained in the study because her attitudes were in the middle of the group.

3.3.3. Traits of accent evaluation

In total 20 semantic differential traits were used on the accent evaluation form. To avoid ambiguity, a 6-point scale was used for each semantic differential trait to exclude mid-points. Participants needed to complete the evaluation form in order to continue on to the next recording.

The traits were selected in three ways. First of all, a small preliminary survey of Hong Kong people's attitudes towards British English and American English was conducted before the language attitudes study. Five Hong Kong people were asked to describe British English and American English using their own words. Words mentioned by the Hong Kong people such as "intelligibility", "politeness" and "formality" were included as traits on the evaluation form. Second, discussions on the most popular online forum in

Hong Kong called Golden Forum¹ mentioned that people felt some Hong Kongers deliberately speak English “with an r-sound” to pretend they were American-born Chinese who can speak native English. This related to the perceived sincerity of a speaker, thus “sincerity” was included as one of the traits. Third, in order to compare the present study with other similar attitudes studies, some of the semantic differential traits in the evaluation form were selected from Zahn and Hopper (1985), Rindal (2010, 2014) and Ladegaard (1998). All the traits can be seen in Table 3.2 or in Appendix 1.

3.3.4. Questionnaire

After completing all the recordings, the participants were also asked to complete a questionnaire (see Appendix 2). The questionnaire contained questions about sex, age, English educational background and overseas experience. The questionnaire also included questions relating to the function of the two English varieties in Hong Kong; for instance, where and when the participants would use British English/American English. Finally, participants were asked about their explicit attitudes towards British English and American English in terms of speaking and listening.

3.3.5. Procedure

A link to the online survey, consisting of eight recordings, accent evaluation forms and the questionnaire, was put on the “Bragging Station” of Golden Forum, in which young people usually discuss gossips and university life. To avoid the language barrier, the survey was translated into traditional Chinese by the author and checked by another native Cantonese speaker from Hong Kong.

3.4. Results

3.4.1. Results for accent evaluation

Each participant gave judgements on eight recordings including two authentic samples, two guised samples, two HKE samples and two fillers. All of the ratings including the two filler recordings were used in statistical analysis; however, ratings of the two filler recordings will not be reported in the following section. In total, 17,120 judgements were used in the analysis (107 responses × 6 recordings × 20 semantic differential traits).

¹Golden Forum URL: <https://forum.hkgolden.com/>

3.4.1.1. Principal Component Analysis

In order to determine evaluative dimensions of these traits, PCA of factor analysis with varimax rotation was conducted using SPSS. (For more discussion on PCA see Chapter 2.4, and Field, 2009, p.628.) PCA calculates correlations within all traits and combines highly correlated traits in the same component. Varimax rotation is a type of orthogonal rotation, which is designed to minimise the number of factors that have high factor loadings on each of the factors. Three components with Eigenvalues higher than 1 were selected. Because of the groupings of factors, the following labels were used: the first component was named as *Status & Competence*, the second component as *Social Attractiveness* and the third component as *Factor 3*. Correlations (loadings) of each trait to the three factors are shown in Table 3.2.

Traits	Factor 1 (Status)	Factor 2 (Attractiveness)	Factor 3
Confidence	.815	.345	.038
Education	.782	.196	.208
Formality	.780	.162	.295
Social class	.766	.264	.254
Intelligence	.763	.277	.204
Leadership	.737	.428	.059
Model of pronunciation	.737	.121	.345
Ambition	.689	.467	-.205
Reliability	.688	.268	.424
Aesthetic quality	.618	.285	.453
Intelligibility	.543	.240	.475
Interestingness	.268	.843	.202
Humour	.162	.840	.202
Enthusiasm	.246	.789	.182
Fashion	.317	.772	.234
Attractiveness	.460	.708	.287
Pleasantness	.446	.529	.498
Sincerity	.067	.173	.757
Politeness	.524	.205	.633
Friendliness	.216	.460	.631
<i>Eigenvalue (Rotation)</i>	<i>6.761</i>	<i>4.656</i>	<i>2.897</i>
<i>% of variance explained (rotation)</i>	<i>33.804</i>	<i>23.278</i>	<i>14.487</i>

Table 3.2 Factor loadings of the rotated component matrix of 20 semantic differential traits for Study 1.

The loadings in Table 3.2 indicate the correlation of each trait to the three dimensions. The higher the value, the stronger the correlation it has. A shaded grey background means scales were highly correlated and categorised into the same factor. Eigenvalues were used to determine how factors were included in the analysis. Three factors had an Eigenvalue higher than 1, therefore only three factors were extracted from the data. Based on the Eigenvalue of each factor, the percentage of variance explained was calculated for each factor and shown under “% of variance explained”. For example, factor 1 explained 33.8% of total variance. In total, 71.56% of variance was explained by the three factors.

Based on the literature review in Chapter 2, the labels *Status & Competence* and *Social Attractiveness* were adopted from Rindal (2010, 2014) and Ladegaard (1998). Traits in factor 1 such as *social class*, *education*, *formality* and *leadership* indicate a speaker’s social status associated with his/her intellectual achievement and merits. Traits like *model of pronunciation*, *aesthetic quality* and *intelligibility* refer to the linguistic characteristics of a speaker’s speech. Therefore, the label *Status & Competence* was chosen to represent this dimension. Factor 1 also contains scales such as *confidence*, *intelligence* and *reliability* which seem to indicate a speaker’s attractiveness based on their literal meaning. However, the low loadings of these traits for factor 2 (*Social Attractiveness*) suggests that the Hong Kong participants may correlate a person’s confidence, reliability, ambition and intelligence more to his social status than to his or her social attractiveness. These merits are also qualities of successful members of society. It is reasonable that respondents correlate these qualities to a person’s social status. The *Status & Competence* dimension here should be defined as a broad concept, which correlates a speaker’s social status not only to his/her economic and intellectual achievement, but also to speech competence and to some qualities associated with success.

Factor 2 contained the traits *interestingness*, *humour*, *enthusiasm*, *attractiveness* and *fashion*, which indicate a dimension of the speaker’s social attractiveness. Traits in factor 2 were all associated with people’s appealing characteristics rather than social status. Traits here were also consistent with the traits of the social attractiveness factor in Rindal (2014). Therefore factor 2 was named *Social Attractiveness*.

The third factor was named as *Factor 3* because it was difficult to categorise the three traits in this factor into one single label. PCA grouped *sincerity* together with *politeness*

and *friendliness*. These three traits seemed to indicate a perceptual feeling of interpersonal interaction; for example, one can perceive his/her interlocutor as a friendly, polite but not sincere person.

3.4.1.2. Factor scores

Instead of taking the mean of ratings for each factor, factor scores were used in the analysis. A factor score is a composite score for each individual on a particular factor. It is calculated from factor loadings (numbers shown in Table 3.2) and correlation coefficients. Factor loading represents the correlation of each trait to the factors, while correlation coefficients represent the correlation between each pair of traits. In this way, both impacts of factors and other variables can be taken into account. For example, a respondent had three factor scores for the RP-guise recording: the factor score of *Status & Competence* = -0.304, the factor score of *Social Attractiveness* = 0.783, and the factor score of *Factor 3* = -0.253. The number -0.304 represented the respondent's attitude towards the RP-guise recording at the dimension of *Status & Competence*. The higher the factor score is, the more positive an attitude it indicates. Zero indicates a neutral attitude based on the sample of the current study.

Before PCA, each respondent had ratings on 20 semantic differential traits; after PCA, the ratings of 20 semantic differential traits were reduced into three factor scores which represented *Status & Competence*, *Social Attractiveness* and *Factor 3*. Therefore, 1,926 factor scores (107 respondents * 3 factors * 6 recordings) were used in the following analysis. Factor scores enable us to compare respondents' attitudes across different recordings in three factors.

3.4.1.3. Descriptive results for English varieties

To gain an overall picture of the data, the means of the three English varieties are presented first: namely RP, GenAmE and HKE. The ratings for the same type of accent were combined; for example, the means and standard deviation for RP were calculated based on the ratings of the authentic RP sample produced by the bi-dialectal RP speaker and the guised RP sample produced by the bi-dialectal GenAmE speaker. Means and standard deviations of RP, GenAmE and HKE are presented in Table 3.

	Status & Competence	Social Attractiveness	Factor 3
RP	0.490 (0.88)	0.071 (1.04)	-0.038 (1.06)
GenAmE	0.044 (0.81)	-0.306 (1.02)	0.130 (1.01)
HKE	-0.72 (1.05)	0.056 (0.99)	-0.195 (0.99)

Table 3.3 Mean factor scores and standard deviation (SD, in brackets) of RP, GenAmE and HKE for the three dimensions. The higher the value is, the more positive an attitude it indicates. The highest value of each dimension is bolded.

The descriptive results indicate that RP was the most prestigious variety of the three, followed by GenAmE and HKE. For *Social Attractiveness*, the means of RP and HKE were very close while GenAmE was rated the lowest, suggesting that Hong Kong respondents accepted their local variety and thought it was attractive. For *Factor 3*, there was not much difference between the three varieties. The ratings of the three varieties are shown in Figure 3.1.

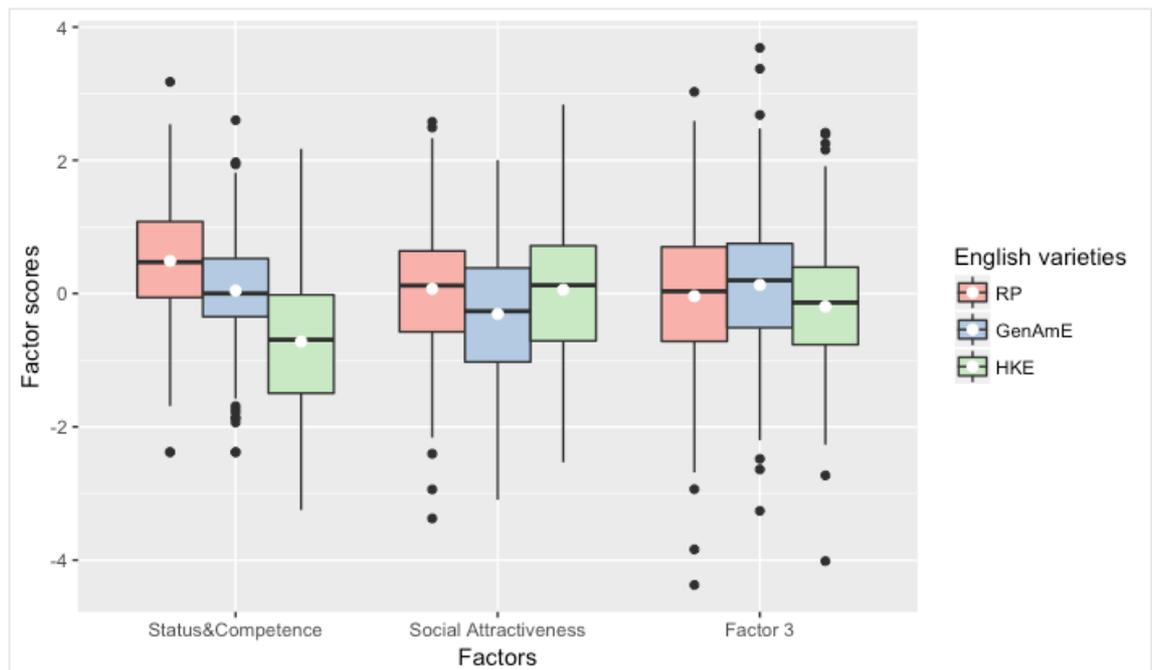


Figure 3.1 Means of factor scores (from left to right: RP, GenAmE and HKE) for Status & Competence, Social Attractiveness and Factor 3. The white dots in the bars represent the means, the black line in the bars represent the medians.

As the matched-guise method was used in the study, it is more interesting to compare the RP samples with the GenAmE samples produced by the bi-dialectal speakers. Therefore, only descriptive results are shown for English varieties. More advanced statistical

analysis including linear mixed effects regressions were used for the results of matched-guise method, and the results of the regressions are presented in the next section.

3.4.1.4. Results of the matched-guise method

This section firstly presents the descriptive results of ratings of six recordings, then follows the linear mixed effects regressions on the factor scores of *Status & Competence*, *Social Attractiveness* and *Factor 3*.

As shown in Table 3.4, RP was rated as more prestigious than GenAmE for both the speakers. For *Social Attractiveness*, RP was rated as more attractive than GenAmE for the female bi-dialectal speaker, but it was rated as less attractive than GenAmE for the male bi-dialectal speaker. An opposite pattern is observed for Factor 3, that RP was rated lower for the female bi-dialectal speaker but higher for the male bi-dialectal speaker.

Table 3.4. shows the means of the factor scores for each recording.

Speaker	Recording	Status & Competence	Social Attractiveness	Factor 3
Female bi-dialectal RP speaker	RP	0.618 (0.92)	0.521 (0.89)	-0.454 (1.10)
	GenAmE-guise	0.190 (0.81)	-0.558 (0.99)	0.156 (1.06)
	<i>p</i>	<.001	<.001	<.001
Male bi-dialectal GenAmE speaker	RP-guise	0.362 (0.83)	-0.378 (0.98)	0.378 (0.85)
	GenAmE	-0.103 (0.78)	-0.052 (0.99)	0.104 (0.96)
	<i>p</i>	<.001	=.009	= 0.09
HKE-mild speaker	HKE-mild	-0.135 (0.88)	0.062 (1.05)	-0.176 (0.96)
HKE-strong speaker	HKE-strong	-1.305 (0.88)	0.050 (0.92)	-0.212 (1.02)
	<i>p</i>	<.001	=1.000	=1.000

Table 3.4 Mean factor scores and SD (in brackets) for the three dimensions. P-values were based on results of post-hoc tests of linear mixed effects regression. The higher value in comparisons was bolded.

To further analyse the results, separate linear mixed effects regressions were used. The full model included **recording** and **sex** as the fixed effects, and included random intercepts by **participant**. **Recording** refers to the six recordings that the participants rated on. **Sex** refers to the participants' self-reported biological sex. **Participant** refers to the 107 participants who involved in the study.

The factor scores of *Status & Competence*, *Social Attractiveness* and *Factor 3* were used as dependent variables. A full model containing two fixed effects factors and one random effects factors was adopted for each of the regressions.

The formula of the full is shown below:

$$\text{Full model} = \text{Factor scores} \sim \text{recording} + \text{sex} + (1 \mid \text{participant})$$

Results of linear mixed effects regressions are presented according to the dimensions below.

Status & Competence

Two nested models with recording and sex removed respectively were run separately, and ANOVA was used to compare the nested models and the full model. The results of model comparison suggest that including **recording** as the fixed effect significantly increased the model fit (Chi-square = 311.05, DF = 5, $p < .001$). No significant effect was found for **sex** (Chi-square = 0.3011, DF = 1, $p = 0.58$).

Post-hoc tests of three specific contrasts of **recording** (RP vs GenAmE-guise, RP-guise vs GenAmE, HKE-strong vs HKE-mild) were run for factor scores of *Status & Competence*, with adjusted p-values in the Bonferroni method. Table 3.5 shows the results of post-hoc tests.

Post-hoc tests of recordings		
RP – GenAmE-guise = 0.4284	SE = 0.1035, z = 4.138	$p < .001^*$
RP-guise – GenAmE = 0.4650	SE = 0.1035, z = 4.491	$p < .001^*$
HKE-strong – HKE-mild = -1.1707	SE = 0.1035, z = -11.309	$p < .001^*$

Table 3.5 Post-hoc tests of recordings for factor scores of Status & Competence. * indicates a significant comparison.

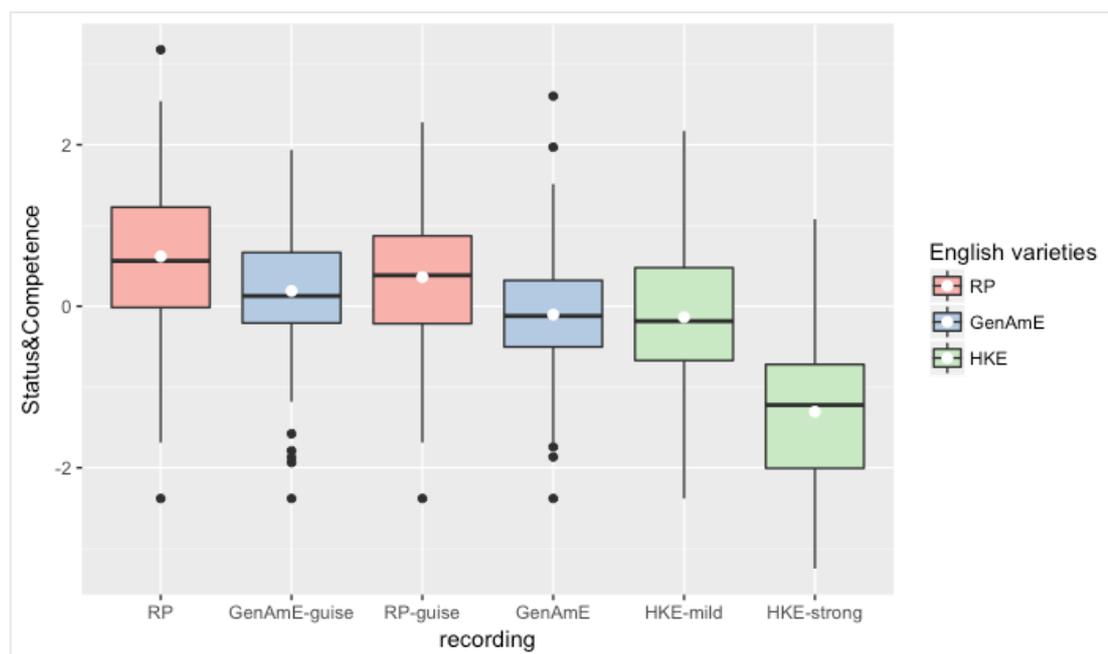


Figure 3.2 Means of ratings for Status & Competence for all the recordings. The white dots in the bars represent the means, the black line in the bars represent the medians.

Figure 3.2 shows the means of all the recordings for *Status & Competence*. The results suggest that for the two bi-dialectal speakers, their RP samples were rated significantly higher than their GenAmE samples. HKE-strong was significantly lower than all the other recordings in terms of *Status & Competence*, indicating that HKE-strong was the least prestigious accent of all. Similarly, HKE-mild was significantly lower than all the other native recordings except for GenAmE. HKE-mild was rated higher than HKE-strong.

Social Attractiveness

The same procedures of model comparison were run for the factor scores of Social Attractiveness. Results of model comparisons suggest that **recording** was the significant predictor for the model (Chi-square = 106.65, DF = 5, $p < .001$). Participants' **sex** was not significant for the model (Chi-square = 0.18, DF = 1, $p = 0.6714$).

Post-hoc tests of three specific contrasts of **recording** (RP vs GenAmE-guise, RP-guise vs GenAmE, HKE-strong vs HKE-mild) were run for factor scores of *Social Attractiveness*, with adjusted p-values in the Bonferroni method. Table 3.6 shows the results of post-hoc tests.

Post-hoc tests of recordings		
RP – GenAmE-guise = 1.0798	SE = 0.11, $z = 9.810$	$p < .001^*$
RP-guise – GenAmE = -0.326	SE = 0.11, $z = -2.962$	$p = .009^*$
HKE-strong – HKE-mild = -0.0114	SE = 0.11, $z = -0.104$	$p = 1.000$

Table 3.6 Post-hoc tests of recordings for factor scores of Social Attractiveness.

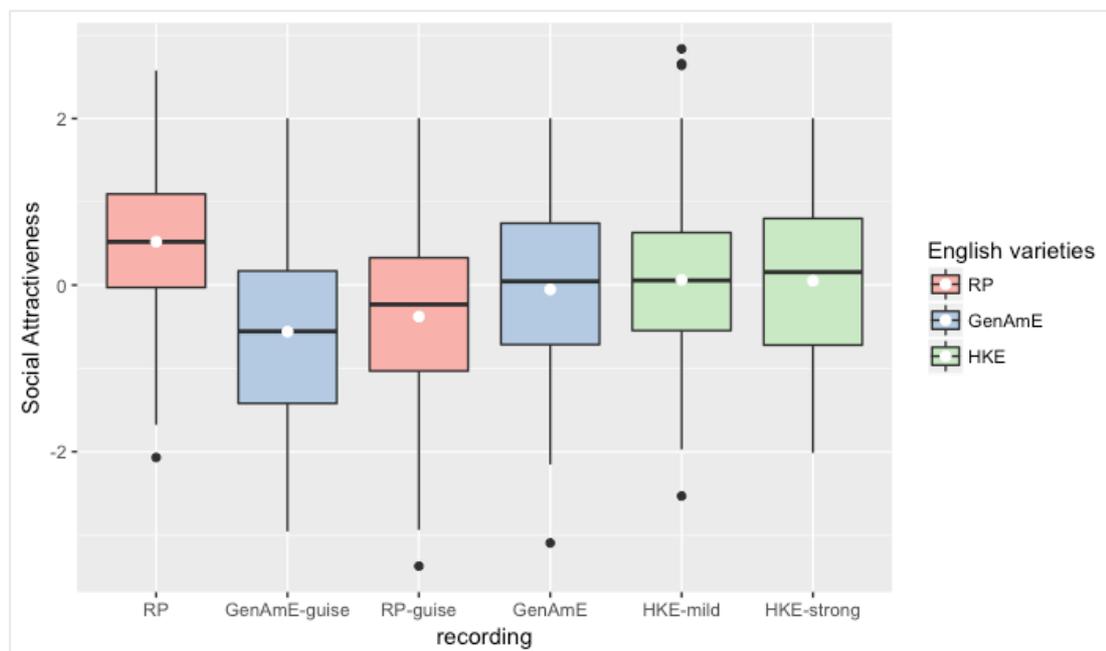


Figure 3.3 Means of ratings for Social Attractiveness for all the recordings. The white dots in the bars represent the means, the black line in the bars represent the medians.

Figure 3.3 shows the means of all the recordings for *Social Attractiveness*. Post-hoc tests of recording suggest the RP sample of the bi-dialectal RP speaker was rated higher than her GenAmE sample; however, the RP sample of the bi-dialectal GenAmE speaker was rated lower than his GenAmE sample.

The two bi-dialectal speakers elicited different attitudes towards RP and GenAmE, which seem to indicate that speaker's sex may have an impact on people's attitudes. As only one speaker of each sex was selected, further analysis of the effect of speaker's sex was not possible.

Post-hoc tests also suggest that the comparison between HKE-strong and HKE-mild was not significant, indicating that the HKE with a strong accent was viewed as equally attractive as the HKE with a mild accent. Interestingly, as shown in Table 3.7, HKE-

strong and HKE-mild were significantly higher than two of the native recordings: GenAmE-guise and RP-guise, which indicates that the respondents viewed the two types of HKE as more attractive than some of the native varieties.

Post-hoc tests between two HKE recordings and two guise recordings		
HKE-strong – GenAmE-guise = 0.609	SE = 0.11, z = 5.537	$p < .001^*$
HKE-mild – GenAmE-guise = 0.620	SE = 0.11, z = 5.641	$p < .001^*$
HKE-strong – RP-guise = 0.429	SE = 0.11, z = 3.899	$p < .001^*$
HKE-mild – RP-guise = 0.440	SE = 0.11, z = 4.003	$p < .001^*$

Table 3.7 Post-hoc tests between two HKE recordings and two guise recordings, * represents a significant comparison, p-value was adjusted with the Bonferroni method.

Factor 3

The same procedures of model comparison were run for the factor scores of *Factor 3*. Results of model comparisons suggest that **recording** was a significant predictor for the model (Chi-square = 52.803, DF= 5, $p < .001$) but not **sex** (Chi-square = 0.2549, DF= 1, $p = 0.6137$).

Post-hoc tests of three specific contrasts of **recording** (RP vs GenAmE-guise, RP-guise vs GenAmE, HKE-strong vs HKE-mild) were run for factor scores of *Factor 3*, with adjusted p-values in the Bonferroni method. Table 3.8 shows the results of post-hoc tests.

Post-hoc tests of recordings		
RP – GenAmE-guise = -0.6108	SE = 0.127, z = -4.773	$p < .001^*$
RP-guise – GenAmE = 0.274	SE = 0.127, z = 2.141	$p = .096$
HKE-strong – HKE-mild = -0.035	SE = 0.127, z = -0.278	$p = 1.000$

Table 3.8 Post-hoc tests of recordings for factor scores of Factor 3. * indicates a significant comparison.

Post-hoc tests of recording suggest that for the bi-dialectal RP speaker, her RP sample was rated lower than her GenAmE sample. For the bi-dialectal GenAmE speaker, his RP sample was not significantly different from his GenAmE sample. Again, the speaker's sex may have had an impact on the respondents' ratings. No significance was found between HKE-mild and HKE-strong in terms of *Factor 3*. Figure 3.4 shows the boxplots of ratings for all the recordings for *Factor 3*.

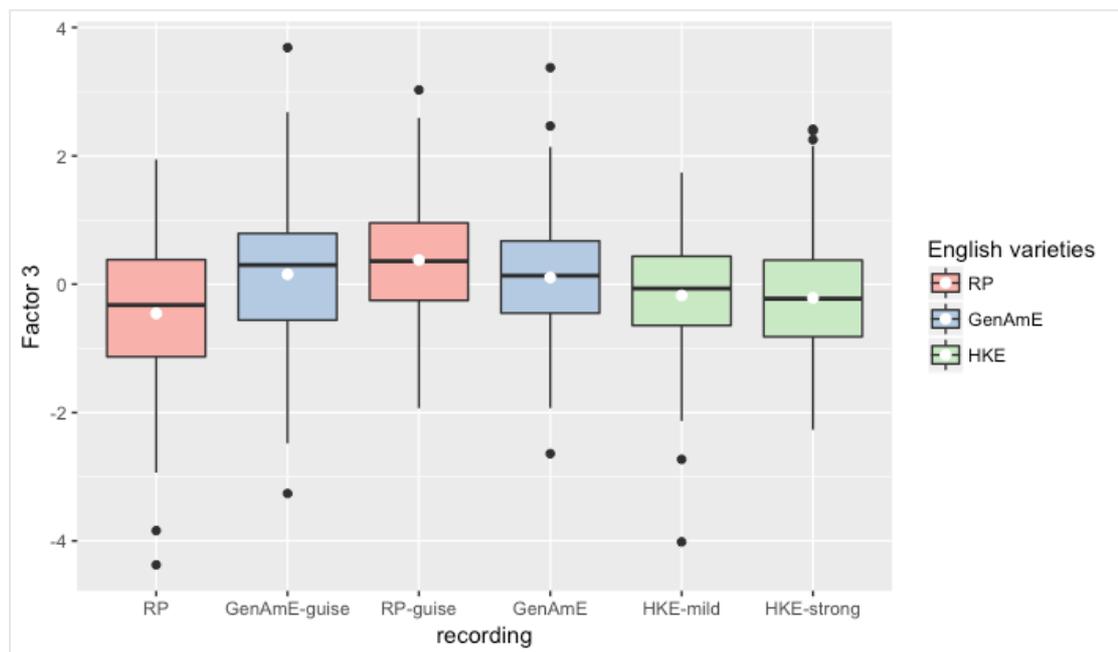


Figure 3.4 Means of ratings for Factor 3 for all the recordings. The white dots in the bars represent the means, the black line in the bars represent the medians.

3.4.1.5. Summary

This section reports the results for six recordings and the comparisons between the RP samples and the GenAmE samples from the two bi-dialectal speakers.

Table 3.9 summarises the higher rated variety in the comparison between the RP sample and the GenAmE sample for the two bi-dialectal speakers. For *Status & Competence*, RP was rated as more prestigious than GenAmE for both bi-dialectal speakers. For *Social Attractiveness*, RP was more attractive than GenAmE for the bi-dialectal RP speaker, whereas GenAmE was more attractive for the bi-dialectal GenAmE speaker. For *Factor 3*, GenAmE was higher for the bi-dialectal RP speaker but lower for the other bi-dialectal speaker.

Table 3.10 shows the ranking of all the recordings based on their means. Note that the ranking does not indicate a significant difference across varieties in different ranks. For example, for *Social Attractiveness*, even though HKE-mild and HKE-strong were ranked higher than GenAmE, the means of the two HKE varieties were not significantly different from the mean of GenAmE. Therefore, this table should be interpreted with caution.

	Status & Competence	Social Attractiveness	Factor 3
bi-dialectal RP speaker	RP	RP	GenAmE-guise
bi-dialectal GenAmE speaker	RP-guise	GenAmE	RP-guise

Table 3.9 The higher rated variety in the comparison between RP and GenAmE for the bi-dialectal speakers.

Rank	Status & Competence	Social Attractiveness	Factor 3
1	RP	RP	RP-guise
2	RP-guise	HKE-mild	GenAmE-guise
3	GenAmE-guise	HKE-strong	GenAmE
4	GenAmE	GenAmE	HKE-mild
5	HKE-mild	RP-guise	HKE-strong
6	HKE-strong	GenAmE-guise	RP

Table 3.10 Ranking of all the recordings in the three dimensions.

For the two HKE varieties, they had the lowest ratings on *Status & Competence*, suggesting that HKE was still seen as the least prestigious variety compared to the native varieties. For *Social Attractiveness*, the two HKE varieties were rated higher than the two native recordings (i.e. recording of GenAmE and recording of RP-guise), suggesting that the Hong Kong respondents had a high acceptance of HKE and found the two HKE varieties attractive. For *Factor 3*, the HKE varieties were ranked higher than the recording of RP.

The results of the linear mixed effects regressions also suggest that participant's **sex** was not a significant predictor for the models, indicating that the male and the female participants in the study did not differ in their attitudes towards these English varieties.

Note that for the two bi-dialectal speakers, the recordings of their authentic accent were rated higher on *Social Attractiveness*, while the recordings of their guised accent were rated higher on *Factor 3*. This suggests that the speakers' guised accents might have an effect on listeners' ratings.

In conclusion, the matched-guise results suggest that RP is still the most prestigious variety of the three. For *Social Attractiveness*, the results were different for the two bi-dialectal speakers. RP was rated as more attractive than GenAmE for the bi-dialectal RP speaker, whereas RP was rated as less attractive than GenAmE for the bi-dialectal

American speaker.

3.4.2. Results for the questionnaire

Besides the accent evaluation task, the respondents also answered 16 questions relating to their background and experience of English learning. The following section presents the results of this questionnaire, aiming to provide more information about people's attitudes towards RP, GenAmE and HKE.

3.4.2.1. Accents of English teachers in schools

Respondents were asked which accents their teachers had at the levels of primary school, secondary school and college. 'Hong Kong accent' was the dominant accent. More English teachers speak British English than American English at all levels, with the highest proportion of British English speakers among secondary school teachers. Results suggest that British English is adopted more than American English in the teaching of English in Hong Kong.

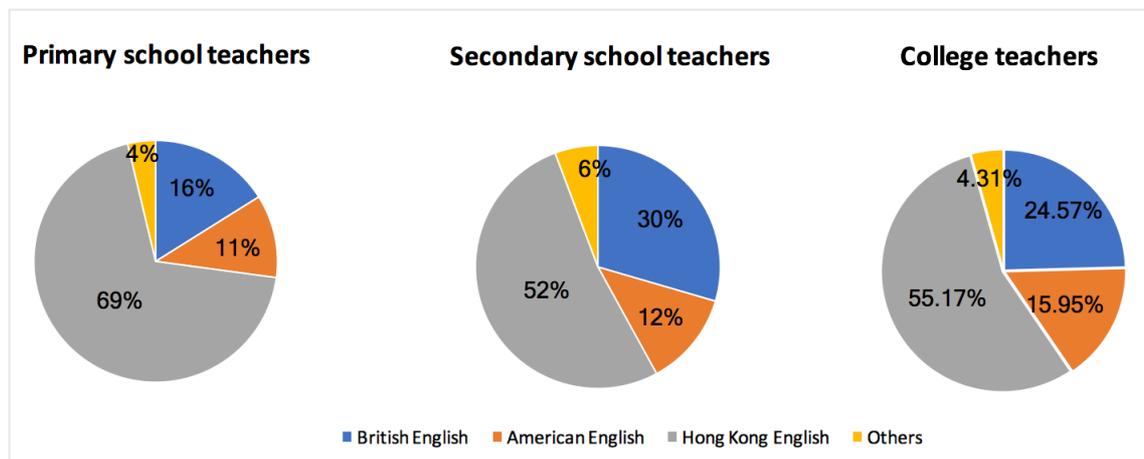


Figure 3.5 Proportions of English varieties spoken by teachers at three levels.

3.4.2.2. Target accents at or outside school

Respondents were asked which accent they speak at and outside school. Results in Figure 3.6 show that British English was used more than American English both at and outside school.

This is consistent with the results for the previous question. Since more teachers teach British English, more students use British English at school. Though American English is used less than British English in both situations, it should be noted that more people

speak American English outside school than at school. This indicates that American English may be used more outside education.

3.4.2.3. Accent that is favoured more in teaching and working

Respondents were asked which accent of English they prefer to be taught at school and to be spoken at work. As shown in Figure 3.7, it was not surprising to find that British English was the most preferred accent to be taught at school. In terms of work environments, both British English and American English were welcomed.

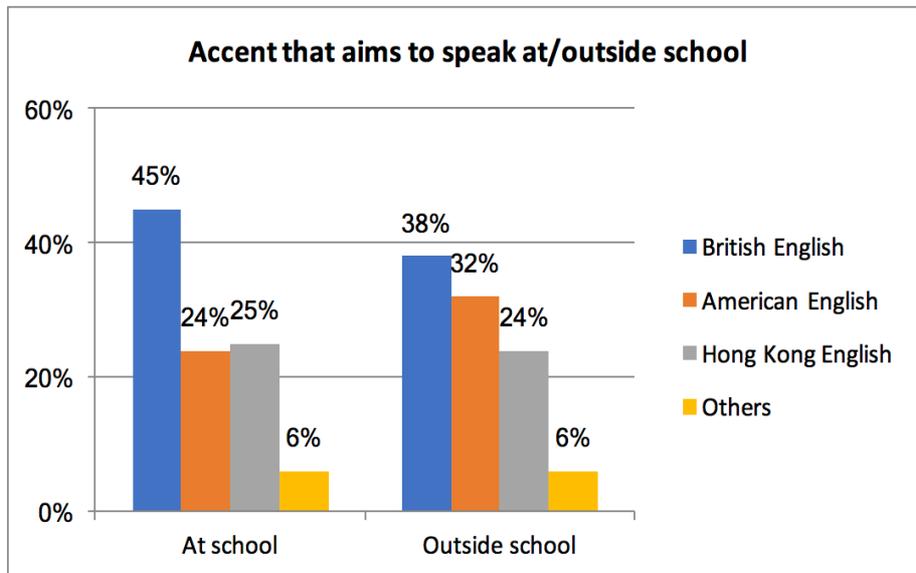


Figure 3.6 Proportions of different English varieties used at and outside school.

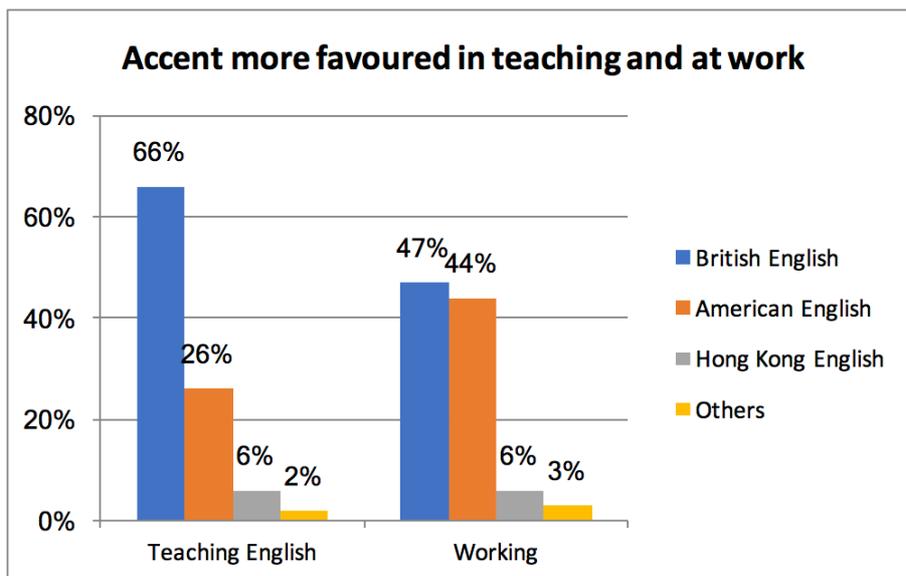


Figure 3.7 Proportions of different English varieties favoured in teaching and at work.

Bacon-Shone, Bolton and Luke (2015) had a similar question in their survey of language use and attitudes in Hong Kong. They interviewed 2046 Hong Kong citizens, asking them which variety should be taught in secondary schools. Their results suggested that 57% of interviewees selected British English as the model variety to be taught in secondary school, 7.8% selected American English, 3.4% selected HKE and 31.8% selected International English. As Bacon-Shone et al. (2015) did not define International English in their study, it is difficult to interpret what this variety meant for Hong Kong people.

Comparing Bacon-Shone et al. (2015) with the present study, both studies confirm that British English is preferred in the teaching of English in Hong Kong. Only 7.8% of the interviewees from Bacon-Shone et al. (2015) selected American English, while the present study shows a higher proportion of American English at 26%. HKE had a very low percentage in both the teaching of English and using English at work in the present study. Similarly, only 3.4% of the interviewees in Bacon-Shone et al. (2015) selected HKE. These results suggest that HKE is still not widely accepted in English teaching in Hong Kong.

3.4.2.4. Accent that is more pleasant to speak and listen to

Respondents were asked to judge which of British English, American English and HKE was the most “pleasant” in terms of speaking and listening. More respondents chose British English as the most pleasant accent to listen to. However, British English and American English have similar popularity in terms of how pleasant they are to speak.

Though over 50% respondents chose British English as the more pleasant accent to listen to, the ratings of RP in *Social Attractiveness* are not as clear as here. One explanation is that “pleasant to speak” relates to speech articulation. The respondents might think American English is easier to pronounce while British English is harder, based on their understanding of the articulatory characteristics of British English and American English.

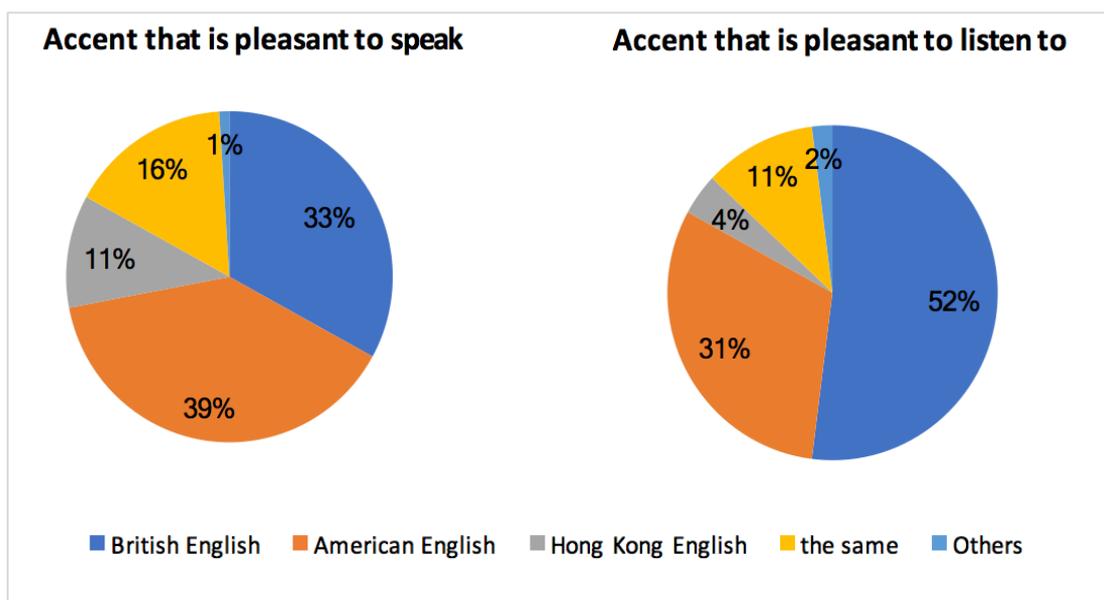


Figure 3.8 Proportions of judging which accent is more pleasant to speak or listen to. “The same” refers to an option that British English is the same as American English.

Bacon-Shone et al. (2015) also asked a similar question, i.e. which accent of English their interviewees preferred when speaking. 61.9% selected British English, 15.6% selected American English and 22.5% selected HKE. HKE had a higher percentage in Bacon-Shone et al. (2015) than in the present study, where only 11% of respondents selected HKE as the most pleasant accent to speak.

3.4.2.5. UK or US in terms of living and working

When the respondents were asked to choose between the Great Britain and the United States, more preferred to live and work in the UK than in the USA. This result in Table 3.11 suggests that Hong Kong people may have a stronger integrative orientation towards British English.

	United Kingdom	United States
Where would you rather live?	66.4%	33.6%
Where would you rather work?	58.9%	41.1%

Table 3.11 Proportions of places preferred to live and to work

3.4.2.6. Summary

This section reports the respondents’ direct attitudes towards British English, American English and HKE.

The results from the previous section suggest that RP was rated as the most prestigious variety. A few results from the questionnaire may provide an explanation for this. In the questionnaire, British English was selected as the favoured accent in teaching (66%) and the accent the respondents aimed to speak at school (45%). This suggests that the high status of RP in the accent evaluation may be due to the dominant role of British English in education in Hong Kong. *Status & Competence* from PCA contained traits like *formality*, *education* and *model of pronunciation*: it is not surprising that RP was rated highest in this dimension if British English is preferred in the teaching of English in Hong Kong.

In terms of *Social Attractiveness*, the results for accent evaluation also suggest that RP was rated as more attractive for the bi-dialectal RP speaker, and GenAmE rated higher for the other bi-dialectal speaker. A similar ambiguity was also found in the results of the questionnaire, i.e. that the respondents did not show a clear preference between British English and American English. In the questionnaire, 38% of the respondents preferred to speak British English outside school while 32% preferred American English. When they were asked which accent was more pleasant to speak, 33% of them selected British English and 39% selected American English.

HKE had a very high rating for *Social Attractiveness* in the accent evaluation. This is also reflected in the questionnaire, where around 25% of the respondents selected HKE as the accent they aimed to speak both in and outside school. However, HKE was still seen as the least favoured accent in teaching and working. Only 11% of the respondents chose HKE as the most pleasant accent to speak in, and 4% chose HKE as the most pleasant accent to listen to. This may be due to the different attitudes that were elicited from these tasks. The questionnaire elicited the respondents' direct attitudes towards HKE, while the accent evaluation elicited their indirect attitudes based on their ratings of 20 traits.

If people's attitudes towards one language/accent are very stable, then their direct attitudes and indirect attitudes are very likely to be the same. For example, in the current study the respondents showed a consistent attitude towards the status of British English/RP. However, for a developing English variety like HKE, it is not surprising that the respondents showed different attitudes in the accent evaluation and the questionnaire.

Depending on which dimension(s) of the attitudes are activated, the HKE speakers' production might be affected differently.

In conclusion, the questionnaire helped us to better understand the respondents' attitudes towards British English, American English and HKE.

3.5. Discussion

This chapter presents a study of language attitudes using the matched-guise method. The hypothesis of the present study is that British English is more prestigious and more attractive to the people of Hong Kong than American English and HKE. The results partially confirmed the hypothesis. For status, the results suggested that British English was more prestigious than American English and HKE. However, in terms of attractiveness, it was true that for the bi-dialectal RP speaker her British English was rated as more attractive than her American English, but for the bi-dialectal GenAmE speaker, his British English was rated as less attractive than his American English.

The following section will discuss each of the dimensions and compare the results with the previous studies reviewed in Chapter 2.

3.5.1. Status & Competence

A few studies suggested that British English is preferred in Hong Kong (Bolton & Kwok, 1990; Candler, 2001; Cheng, 2013; Chan, 2013; Groves, 2011). However, most of these studies used questionnaires to elicit people's attitudes. It is therefore difficult to directly compare the three dimensions of the present study to their results. Cheng (2013) was an exception. His questionnaire contained 22 statements which were grouped into three categories: the *affective dimension*, *linguistic dimension* and *pragmatic dimension*. These three categories allow us to compare his results with the three dimensions of the present study.

Cheng (2013, p.3) proposed a *linguistic dimension* which he defined as "a dimension [that] deals with the problem of correct and incorrect, i.e. whether the subject believes a certain variety is the *standard*". This is similar to the *Status & Competence* factor in the present study. His results suggested that British English was preferred in this dimension, which

was consistent with the result for *Status & Competence* in the present study. Furthermore, both the bi-dialectal speakers' RP samples were rated higher than their GenAmE samples for *Status & Competence* in the present study, which again suggests the superiority of British English in this dimension.

On the other hand, the British accent was found to have a lower status than the American accent in Zhang (2009) and Bayard and Green (2002). One reason might be that Zhang (2009) and Bayard and Green (2002) both used the verbal-guise method, which can be affected by individual differences. Another reason for gaining a different result in the present study might be that the present study had different respondent groups from those in Zhang (2009), and Bayard and Green (2002). Zhang (2009) and Bayard and Green (2002) only included university students in their studies, while 35% of the respondents in the present study were either secondary school students (25%) or professionals (10%). As the results of the questionnaire suggest that British English was used by more English teachers at school, it is possible that the secondary school respondents in the present study preferred British English to American English because British English is the standard form used in English learning at school.

3.5.2. Social Attractiveness

For *Social Attractiveness*, RP was rated as more attractive than GenAmE in the present study, which was also found among the 'local' group of Cheng (2013) for the *affective dimension*. Cheng (2013, p.3) defines it as "a dimension [that] deals with the problem of like and dislike, i.e. whether the subjects like to speak, listen, read and write in a certain linguistic variety, and this can be reflected and interacts with their preferences for the culture and other social aspects related to the linguistic variety". As the *Social Attractiveness* of the present study contains traits like *pleasantness*, *attractiveness* and *fashion*, it is possible to compare *Social Attractiveness* to Cheng's (2013) *affective dimension*. In this way, Cheng (2013) and the present study agree that British English is perceived as more attractive than American English by respondents from Hong Kong. Also, the results from the questionnaire indicated that over half of the respondents thought that British English was more pleasant to listen to.

If we look at the ratings of male and female speakers separately, in the present study,

GenAmE produced by the bi-dialectal GenAmE male speaker was rated as more attractive than his guised RP recording. This is also found in Bayard and Green (2002), where the male American speaker was rated much higher than the male British speaker in the *solidarity dimension*.

As for studies which involved female speakers, Zhang (2009) and Bayard and Green (2002) both found that the female American speakers in their studies were rated as more attractive than the female British speakers. In the present study, the RP recording from the bi-dialectal RP female speaker was rated as less attractive than her GenAmE recording.

3.5.3. Factor 3

Factor 3 was the third dimension of the present study, which contained three scales: *sincerity*, *friendliness* and *politeness*. This dimension seems to indicate a perceptual feeling of interpersonal interaction. The results for *Factor 3* were difficult to interpret. The female bi-dialectal RP speaker's RP sample was rated lower than her GenAmE sample for *Factor 3*; the male bi-dialectal GenAmE speaker's RP sample was rated higher than his GenAmE sample.

Bayard and Green (2002) found their female American speaker was rated higher on the scale of *friendliness* than the female British speaker, but it is worth pointing out that Bayard and Green (2002) used two different female speakers, while the present study used the same female bi-dialectal RP speaker. Similarly, Cheng (2013) found that "friendly" was chosen as the best description (1st out of 20 descriptions) of a typical American speaker, while it is only ranked 16th out of 20 descriptions of a typical British speaker.

3.5.4. Sex of speakers

The sex of the speakers seems to have an influence on ratings; however, it is worth noting that there were only one female and one male speaker in Study 1. *Social Attractiveness* and *Factor 3* demonstrated a different pattern for the male and female bi-dialectal speakers. The RP sample produced by the female speaker was rated less attractive than her American sample. Conversely, the RP recording from the male speaker was rated more attractive than his American recording.

A similar phenomenon is also found in van der Haagen (1998), based on Dutch learners of English. Four male and four female speakers were used in van der Haagen (1998). In van der Haagen (1998, p.57), four dimensions were extracted: *status* represented the social status a speaker appeared to have; *dynamism* referred to the perception of the degree of dynamism of a speaker; *affect* referred to the personal affect a subject felt for the speaker and *norm* indicated the extent to which a speaker was perceived to conform to the school norm.

For *status*, the male speakers' RP was rated as more prestigious than their GenAmE, while the female speakers' RP was rated as less prestigious than their GenAmE. This dimension is close to the *Status & Competence* of the current study, in which RP was rated as more prestigious than GenAmE for all speakers. For *dynamism*, GenAmE was rated higher than RP for all the speakers regardless of their sex. This dimension is close to the *Social Attractiveness* of the present study, where GenAmE was rated more attractive for the male bi-dialectal speaker and RP was rated more attractive for the female bi-dialectal speaker. For *affect*, the female speakers' GenAmE was rated higher than their RP, whereas no difference was found between the male speakers. This dimension is close to the *Factor 3* of the present study in which the female bi-dialectal speaker's GenAmE was rated lower than her RP, while the male bi-dialectal speaker's GenAmE was rated higher than his RP.

Another possibility is that the bi-dialectal speakers' guised accents co-vary with the effect of sex. The recordings of the speakers' authentic accent were rated higher than the recordings of their guised accent on *Social Attractiveness*. In contrast, the recordings of the speakers' guised accent were rated higher than the recordings of their authentic accent on *Factor 3*. Although the authenticity of the guised recordings was confirmed by one British phonetician and one Hungarian phonetician, the HK respondents might perceive the guised recordings differently.

As the present study only contained one male and one female speaker, there is not enough evidence to conclude that the findings were due to a 'sex effect'. Therefore, further research is needed on this aspect.

3.5.5. Attitudes and motivations

Gardner and Lambert (1972) proposed an ‘integrative motive’ and an ‘instrumental motive’ for learning a second language. Although Gardner and Lambert’s model (1972) is usually used in bilingual or multilingual studies, it is possible to apply to attitude studies of different English accents. In the present study, Hong Kong participants may favour British English or American English for different orientations. The results of the questionnaire show that British English was taught and spoken more than American English at school. On the other hand, American English has become as popular as British English at work. This suggests that Hong Kong people have a strong instrumental orientation towards British English and American English. When respondents were asked which country they would like to live in, 66% of them chose the UK. This indicated that the Hong Kong respondents may have a stronger integrative orientation for British English than American English. As the present study did not directly measure the respondents’ integrative and instrumental orientations, more studies are needed to examine these assumptions.

3.5.6. Attitudes towards Hong Kong English

In Chapter 2, the status of HKE was discussed. Most studies suggested that HKE is transiting from phase 2 (exonormative stabilization) to phase 3 (nativization) in Schneider’s model (2003, 2007). In the case of HKE, phase 2 (exonormative stabilization) refers to the stage where people in Hong Kong still use English in administration, education and the legal system and they value the imported native variety (i.e. British English) higher than the local variety (i.e. HKE). Phase 3 (nativization) refers to a transition from accepting the imported native variety (i.e. British English) as the dominant language to increasing independence of the local variety (i.e. HKE), along with more contact and interaction between the two varieties.

The results of the present study support the above view. For the two HKE varieties in the present study, they were rated the lowest in terms of *Status & Competence*, suggesting that HKE is still in phase 2 (exonormative stabilization) of Schneider’s model (2003, 2007). That is, the imported native variety (British English) is still regarded as superior to the local variety (HKE).

For *Social Attractiveness*, the two HKE recordings were rated significantly higher than the two native recordings (i.e. GenAmE-guise and RP-guise), suggesting that the Hong Kong respondents had a high acceptance of their own local variety. This is consistent with the argument that HKE is at the phase 3 (nativization) in Schneider's model (2003, 2007). They are transiting from regarding native varieties of English as authoritative to accepting HKE.

3.6. Improvements

There are a few things that could be improved in the current study. First of all, a task of accent identification is needed to ensure that respondents are able to identify the different varieties of English. Though previous studies have shown a high rate of accent identification for RP and GenAmE among Hong Kong people, it would be more precise if any further studies included an accent identification task.

Second, as discussed in section 5.5, it would be interesting to add measurements of integrative and instrumental orientation. As English is widely used in education and in the legal system in Hong Kong, introducing integrative and instrumental orientation in the questionnaire would help to better understand Hong Kong people's attitudes towards different varieties of English.

Third, since the male and female bi-dialectal speakers showed different results for *Social Attractiveness*, to avoid the interference of sex, same-sex bi-dialectal speakers should be used in further studies.

3.7. Conclusion

This chapter reported on a study of language attitudes. The results largely matched the hypothesis, i.e. that British English is more prestigious and attractive than American English to the people of Hong Kong. HKE was also found to be more attractive than the RP-guise and GenAmE-guise recording. The heterogeneous attitudes the HK respondents showed in the accent evaluation and in questionnaire might affect their production differently.

Chapter 4. Review of theories of short-term accommodation

In this chapter, theories of accommodation and some other related theories will be reviewed, including Communication Accommodation Theory (Giles et al., 1991), the interactive-alignment model (Pickering & Garrod, 2004), the exemplar-based models (Docherty & Foulkes, 2014; Foulkes & Docherty, 2006; Johnson, 1997a; Pierrehumbert, 2001) and the Interaction Hypothesis (Gass & Mackey 2007, 2015).

First of all, CAT and the interactive-alignment model will be reviewed. These two theories are directly related to speech accommodation. CAT has been a dominant theory of accommodation in sociolinguistics since the 1980s. It focuses on the motives of accommodation, suggesting that people accommodate to shorten social distance. The interactive-alignment model is a recent psycholinguistic account which proposes that accommodation is an automatic process, occurring at all levels of dialogue. These two models provide a foundation for studies of speech accommodation.

Secondly, some exemplar-based models will be reviewed. Though the exemplar-based models are not designed for speech accommodation, they provide a good paradigm to explain how speech accommodation occurs. For example, based on the exemplar-based perception model (Johnson, 1997a), the production model (Pierrehumbert, 2001) proposes a perception-production loop, suggesting that categories of sounds keep updating when people perceive new input, and production changes according to the updated categories. On the other hand, Docherty and Foulkes (2014) and Foulkes and Docherty (2006) use exemplar-based theories to explain sociophonetic variation. The present project aims to examine the role of language attitudes in accommodation. The exemplar-based sociophonetic account helps to explain how people's language attitudes affect their production.

Finally, the Interaction Hypothesis from L2 learning will be reviewed. As the present study is interested in convergence of non-native English speakers, it would be useful to review convergence from the perspective of L2 learning.

4.1. What is accommodation?

Accommodation, which is also known as convergence, has been widely studied in sociolinguistics, psycholinguistics and cognitive science. The definition of accommodation varies across different disciplines.

In sociolinguistics, convergence is defined as “a strategy whereby individuals adapt to each other’s communicative behaviours in terms of a wide range of linguistic/prosodic/non-verbal features including speech rate, pausal phenomena and utterance length, phonological variants, smiling, gaze, and so on” (Giles et al., 1991, p.7). This is opposed to divergence, which is defined as “the way in which speakers accentuate speech and non-verbal differences between themselves and others” (Giles et al., 1991, p.8).

If a speaker changes his pronunciation, speech rate, intonation or lexical choice to become more like his interlocutor in a conversation, he is performing convergence; in contrast, if he changes his speech to be less like the interlocutor’s, he is performing divergence. Convergence and divergence are not the only dimensions of linguistic change. Speech maintenance, where the speaker maintains his speech style, is also possible. Bourhis (1979, p.126) suggests that speech maintenance might be used by speakers in interethnic conversation as a strategy to maintain their ethnic identity and cultural distinctiveness.

In psycholinguistics, accommodation is also known as alignment or priming. When they are talking or writing, people tend to repeat an underlying basic structure that they have recently produced or have experienced others produce (Pickering & Ferreira, 2008, p.428). Different from sociolinguistic theories, which are mainly interested in the social motives of accommodation, cognitive psycholinguists are primarily interested in the underlying psychological or cognitive mechanisms of repetition, and how repetition/priming reflects learning and development.

Note that the terms *convergence* and *divergence* used in this project are slightly different from the convergence/divergence used in Giles et al. (1991). In Giles et al. (1991), “convergence” and “divergence” refer to speakers’ production changes towards their interlocutors *driven by social psychological motivations*. In this project, if speakers converge towards or diverge from their interlocutors, the changes might or might not be

driven by social motivations. The terms are neutral regarding speakers' underlying social motivations. For example, if a speaker diverges from his/her interlocutor, "divergence" would be used to describe the speaker's accent changes, no indication is made to the speaker's motivation of divergence.

4.2. Accommodation in sociolinguistics: Communication Accommodation

Theory

From the 1970s to the 1990s, accommodation attracted a great deal of attention from social psychologists. A representative scholar of accommodation is Howard Giles. Giles and his colleagues started working on accommodation in the 1970s, including Giles' first article on the "accent mobility" model in early 1973 (Giles, 1973), the speech accommodation theory (SAT: Giles, Mulac, Bradac & Johnson, 1987), and a broader version of CAT (Giles et al., 1991).

Since CAT was initiated by a group of social psychologists, the theory was heavily weighted towards explaining the social and psychological motives underlying accommodation. They were mostly interested in how speech was mediated by factors such as psychological motives, interpersonal perception and identity. In the following section, I will first briefly review the early versions of CAT: interpersonal speech accommodation (Giles, 1973) and SAT (Giles et al., 1987), and then introduce CAT (Giles et al., 1991) in more detail.

4.2.1. Accent Mobility

The paper Giles published in 1973 is regarded as the first empirical study of speech accommodation. In that paper, Giles used the term "accent mobility" to refer to the ability of an individual to modify his accent or pronunciation (Giles, 1973, p.89). Labov (1966) suggested that people use 'careful speech' in a formal context and use 'casual speech' in a less formal context. Giles (1973) argued that besides the contextual constraints suggested by Labov (1966), interpersonal constraints should also be considered. He proposed that individuals modify their speech depending on their social motives. If they wish to gain social approval from message receivers, they adapt their speech towards that of the receiver. On the other hand, if they wish to dissociate themselves from the receiver, their speech shifts away from that of the receiver.

To support his model, Giles recruited 13 males to complete an interview with (1) an RP speaker who was perceived as more prestigious in terms of age, educational level and accent, and (2) a Bristolian interviewer who shared the same accent as the 13 participants and was perceived as having equivalent prestige to them. They were interviewed about the same topic using the same procedure. After the interviews, the participants' recordings with both interviewees were extracted and paired for a perceptual judgement task. Giles (1973) hypothesized that the participants would use their idiolect accent with the Bristolian interviewer and would converge to the RP accent with the RP interviewer.

Two groups of listeners were recruited to judge which recording samples from the participants, those with the RP interviewer or those with the Bristolian interviewer, had either a 'broader' or a more formal accent. Results suggested that the majority of the listeners judged the samples produced with the Bristolian interviewer as broader and less formal than the samples produced with the RP interviewer. The empirical evidence thus supported Giles' model of accent mobility.

4.2.2. Speech Accommodation Theory

Inspired by Giles' (1973) article, an increasing number of studies were carried out on interpersonal speech accommodation. Supported by these empirical studies, SAT was developed, aiming to demonstrate the value and potential of social psychological concepts and processes for understanding the dynamics of speech diversity in social settings (Giles et al., 1991).

SAT is grounded on the hypothesis that during interaction individuals are motivated to adjust their speech styles as a strategy for gaining one or more social goals, such as evoking listeners' approval, attaining communication efficiency between speakers, and maintaining positive social identities (Giles et al., 1987, p.15). Speech convergence is often seen as an unconscious reflection of a speaker's desire for social integration.

One may ask, why are speakers motivated to adjust their speech style? Where do these motives come from? To answer these questions, knowing the basis of speech accommodation theories is vital.

Speech accommodation theories have their basis in some social psychological theories such as similarity attraction theory (Byrne, 1969), causal attribution theory (Heider, 1958) and social exchange theory (Homans, 1961).

Similarity attraction theory (Byrne, 1969) suggests that the more similarities two people share, the more likely it is that they will be attracted to each other. Convergence is taken as a strategy to attract the other interlocutor; in other words, to gain social approval.

Causal attribution theory (Heider, 1958) suggests that perceivers do not attribute others' behaviour only to the immediate outcome; they also take into account the actor's ability, effort, and external pressures. This provides theoretical support for how people evaluate convergence. For example, if a person converges towards his interlocutor because he wants to break down cultural barriers, listeners perceive this kind of convergence as a positive gesture; however, if the convergence occurs as a result of external pressures, this kind of convergence is perceived as a negative convergence.

Social exchange theory (Homans, 1961) suggests that people tend to engage in behaviour which is rewarded, and avoid behaviour that results in negative or unpleasant outcomes. We accommodate because we believe that by accommodating we will receive rewards (e.g. social approval). We accommodate when the costs of accommodating (e.g. changing one's identity) are lower than the benefits (e.g. receiving social approval).

4.2.3. Communication Accommodation Theory

As a developed and advanced version of SAT, CAT is better known in academia than SAT. In its early stages, SAT mainly focused on speech accommodation. Later studies in SAT did not limit themselves to speech, but were extended to non-verbal accommodation.

Giles et al. (1987, p.36) summarised the propositions of SAT/CAT in detail. The propositions hypothesise the conditions for convergence, maintenance or divergence. Table 4.1 shows Giles et al. (1987)'s summaries.

Giles et al. (1991) list the principal distinctions to be considered when characterising convergence and divergence, shown in Table 4.2.

Convergence	
People will attempt to converge towards the speech and non-verbal patterns believed to be characteristic of their message recipients, be the latter defined in individual, relational or group terms,	
<i>when speakers</i>	
1	desire recipients' social approval (when the costs are lower than the rewards);
2	desire a high level of communication efficiency;
3	desire a self-, couple-, or group- presentation shared by recipients;
4	desire appropriate situational or identity definitions;
<i>when the recipients'</i>	
5	actual speech in the situation matches the belief that the speakers have about the recipients' speech style;
6	speech is positively valued, non-stigmatized;
7	speech style is appropriate for the speakers as well as for the recipients.
Maintenance or Divergence	
Speakers will attempt to maintain their communication patterns, or even diverge away from their message recipients' speech and non-verbal behaviours,	
<i>when speakers</i>	
1	desire to communicate a contrastive self-image;
2	desire to dissociate personally from the recipients or the recipients' definition of the situation;
3	define the encounter in intergroup or relational terms, with communication style being a valued dimension of their situationally salient in-group or relational identities;
4	desire to change recipients' speech behaviour; for example, moving it to a more acceptable level;
<i>when recipients</i>	
5	exhibit a stigmatized form, a style that deviates from a valued norm, which is
6	consistent with speakers' expectations regarding recipient performance.

Table 4.1 Summaries of propositions of SAT/CAT (edited from Giles et al., 1987, p.36).

Upward/downward accommodation and symmetrical/asymmetrical accommodation indicate the direction of accommodation with respect to language varieties within a social hierarchy, e.g. with a standard variety at the top and regional dialects at the bottom. If speakers shift towards/away from a prestigious variety, it is an upward convergence/divergence; if speakers shift towards a more stigmatized or less socially valued variety, it is a downward convergence/divergence. Symmetrical and asymmetrical refer to whether the accommodation is one-directional or two-directional. For example, if a HKE speaker converges towards an RP speaker in a conversation, it should be counted as upward accommodation because British English is regarded as the most prestigious English variety in Hong Kong (see Chapter 3). If the accommodation does not only occur from the HKE speaker towards the RP speaker, but also from the RP speaker towards the

HKE speaker, it should be seen as symmetrical accommodation as the convergence is mutual. Large/moderate accommodation, unimodal/multimodal accommodation, and full/partial/hyper-/crossover accommodation indicate the scale of accommodation. For example, if the HKE speaker does not only verbally accommodate towards the RP speaker, but also accommodates in non-verbal expression (e.g. nodding head), this can be seen as multimodal accommodation.

Distinctions in characterising convergence and divergence

Upward versus downward
Symmetrical versus asymmetrical
Large versus moderate
Unimodal versus multimodal
Full versus partial versus hyper-/crossover
Subjective versus objective

Table 4.2 Principal distinctions (edited from Giles et al., 1991, p.11).

It is also worth mentioning subjective and objective accommodation. Subjective accommodation refers to speakers' beliefs regarding whether they or others are converging or diverging, while objective accommodation refers to speakers' shifts in speech measured independently as moving towards or away from others. Similar distinctions can be found in actual behavioural convergence and stereotypical convergence (Hewstone & Giles, 1986). The latter refers to speakers converging towards their stereotypical impression of the recipient group rather than the actual speech they hear. These distinctions of accommodation help us to distinguish how speakers actually converge linguistically from how speakers believe they are converging. Convergence is often cognitively mediated by speakers' stereotypes of how others will speak in their various social categories (Giles et al., 1987, p.18).

4.2.4. Consciously or non-consciously?

Whether accommodation is a conscious or non-conscious process has become an argument in recent years. Though many studies of CAT have considered convergence as an active communicative strategy, Giles et al. (1987) argue that:

“It should be noted, however, that use of terms such as strategies and intentions does not necessarily imply that these purposive behaviours are always performed or evaluated consciously with full awareness. Behaviour is organised by cognitive processes at many

levels simultaneously and it is unlikely that these ever are all monitored consciously.”
(p.26)

In other words, Giles et al. (1987) did not agree that convergence is an entirely conscious process. Furthermore, Giles et al. (1987, p.27) imply that in some cases convergence might be a scripted behaviour, that is “interactants may automatically apply a convergence scripts (or schema) to move towards more similar speech.”

Apparently, the argument of whether accommodation is conscious or non-conscious, automatic or non-automatic has gone beyond what CAT could answer, and this is not the focus of CAT either. Models from psycholinguistics such as the interactive-alignment account claim that accommodation is an automatic process. The next section reviews the interactive-alignment model from the perspective of psycholinguistics.

4.3. Accommodation in psycholinguistics: the interactive-alignment model

The interactive-alignment model was proposed by Pickering and Garrod (2004), aiming to explain the mechanism of language processing in a dialogue. It suggests that a successful dialogue involves the alignment of linguistic representations at all levels by the interlocutors.

The traditional view of dialogue suggests an isolation of the speaker and listener’s coding and decoding processes, as shown in Figure 4.1.

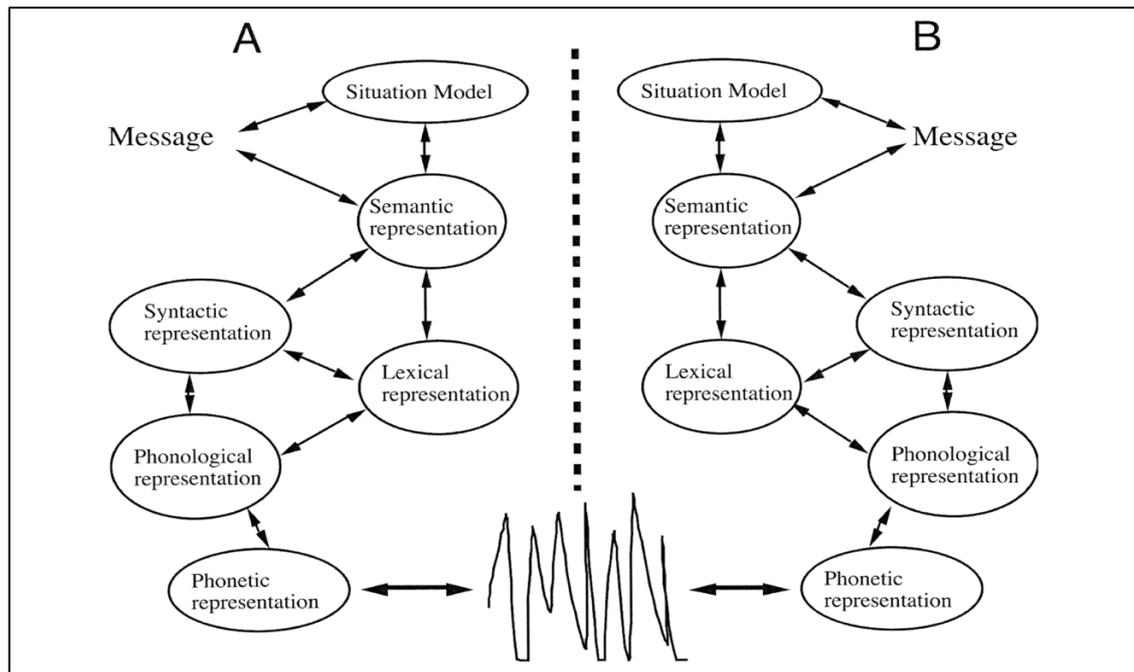


Figure 4.1 The autonomous transmission account, taken from Pickering and Garrod (2004, p. 177).

When two interlocutors start a dialogue, the speaker starts from a conceptualised idea or message that he wants to express. Then he formulates the message into different levels of linguistic representations. For example, an individual converts the message into syntactic and lexical representations, then into phonological and phonetic representations, and lastly into articulatory commands which generate sounds. In turn, the listener decodes the speaker's message by perceiving the sounds first. Then the sounds are decoded into higher levels of linguistic representations until the message is recovered.

In contrast, the interactive-alignment model argues that the speaker and the listener build up the dialogue as a joint activity, and align at different levels of linguistic representations (see Figure 4.2 for details). When the speaker generates syntactic representations, he takes into consideration the listener's syntactic representations; correspondingly, when it is the listener's turn to speak, he also considers the other person's representations and adjusts his speech to be aligned with the other person to minimize the collaborative effort. These alignment processes, according to Pickering and Garrod (2004), are automatic and direct.

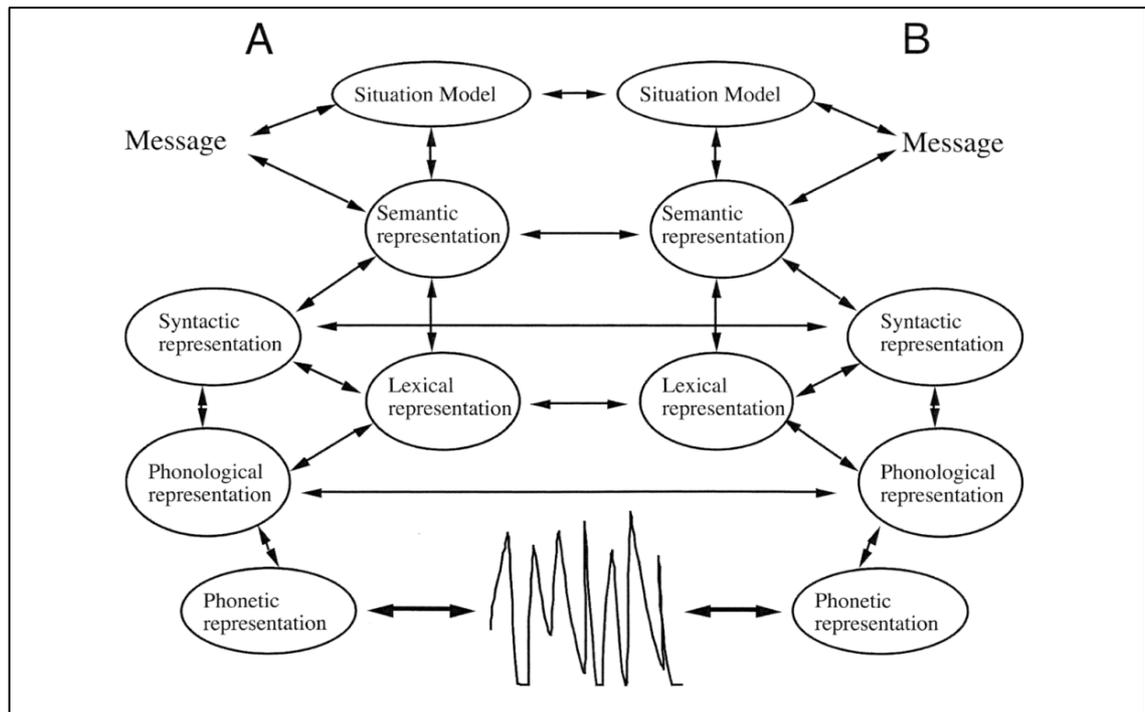


Figure 4.2 The interactive-alignment model, taken from Pickering and Garrod (2004 p. 176).

More specifically, the interactive-alignment model (Pickering & Garrod, 2004, p.172) contains six arguments:

- 4.1. alignment of situation models (Zwaan & Radvansky, 1998) is the basis of successful dialogue;
- 4.2. a primitive and resource-free priming mechanism contributes to the alignment of situation models;
- 4.3. alignment of other levels of representations including phonetic, phonological, lexical and syntactic levels is also achieved by the same resource-free priming mechanism;
- 4.4. interconnections between different levels suggest that alignment at one level leads to alignment at other levels;
- 4.5. another primitive mechanism allows interlocutors to repair misaligned representations interactively;
- 4.6. other explicit strategies may be applied only if the primitive mechanism fails to achieve alignment.

The resource-free priming mechanism refers to a process in which listeners hear an utterance that activates a particular representation, and are then more likely to produce an

utterance which uses the same representation. This mechanism is supported by Garrod and Anderson (1987)'s output/input coordination, which also suggests that speakers formulate their output according to the input received from the interlocutor.

The interactive-alignment model receives criticisms and suggestions from many scholars. For example, Cutting (2004) argues that the primitive and resource-free mechanism does not explain how the 'direct' links occur between the two interlocutors. In fact, Cutting challenges two fundamental issues of the model: what is the mechanism of alignment and what exactly are the interlocutors priming? Pickering and Garrod (2004) respond that the model does not commit to any specific mechanisms, as it may involve both transient activation and memory-based accounts. The former involves temporary activation of information in memory which is short-term and degrades rapidly, whereas the latter suggests some sort of implicit learning which is experience-based and long lasting. Specifically, Pickering and Garrod (2004) point out that their intention was to argue for the alignment of the structural aspect of the situation model rather than the content of the situation model. In other words, alignment in their model does not determine the content of production.

Another fundamental question raised by Krauss and Pardo (2004) is whether alignment is an automatic process. Pickering and Garrod (2004) answer that the automaticity of their model is derived from the automatic perception-behaviour effect as Dijksterhuis and Bargh (2001) propose, i.e. that humans are like other kinds of animals which act as or imitate what they perceive. There is evidence showing that perception and action share the same representational system in neurophysiology (Paus, Petrides, Evans & Meyer, 1993). Perceiving an action activates the mental representation of this action, which in turn will lead to performance of the action (Dijksterhuis & Bargh, 2001, p.8). What Pickering and Garrod (2004) do not really address are the inhibitors of the perceptual behaviour that Dijksterhuis and Bargh (2001) propose. Though Dijksterhuis and Bargh (2001) suggest a direct link between perception and behaviour, they also propose that the link may be moderated by "inhibitors" such as self-focused attention (i.e. attention to the self) and liking (e.g. feelings of empathy, attitudes towards other people). These inhibitors, however, are similar to the factors that CAT proposes (see section 2 in this chapter for details). This suggests that even for automatic accounts of accommodation like the

interactive-alignment model, social factors like preferences and attention should still play a role.

The interactive-alignment model is useful for understanding language processing in dialogue; however, as it does not explain how alignment occurs at each level of linguistic representation and as more focus is given to syntactic alignment, it may not be able to fully explain phonetic accommodation, which is of interest to the present study. Whether these implications derived from syntactic priming could also apply to phonetic convergence remains unknown.

4.4. Accommodation and exemplar-based theories

Unlike CAT and the interactive-alignment model, which were developed to explain accommodation, exemplar-based theories are cognitive models which focus on speech perception, production and language processing. As neither CAT nor the interactive-alignment model explains how convergence occurs between interlocutors, the exemplar-based theories provide a good paradigm for phonetic convergence. The exemplar-based approach began in the 19th century in psychology of memory and has been adopted in many fields; for example, Johnson (1997a; 2007) proposes an exemplar model to explain talker normalization in speech perception; Pierrehumbert (2001; 2003) explains phonological learning and pronunciation variation using an exemplar-based phonetic system; Foulkes and Docherty (Docherty & Foulkes, 2014; Foulkes & Docherty, 2006) use the exemplar-based model to account for sociophonetic variation. This section will review the above models.

4.4.1. Exemplar models in speech perception

Johnson (1997a)'s exemplar model was not the first exemplar model in speech perception. Built on a few previous studies of exemplar models of perception (Hintzman, 1986; Nosofsky, 1988), Johnson (1997a) proposes a model to explain talker normalization. He suggests that categorisation is achieved by comparing a new input item with each of the remembered instances/exemplars of each category. The similarity between the item and each exemplar determines the activation level of the exemplar. The greater the similarity, the higher the activation level. The new item should be categorised as an example of the category which has the highest activation level after comparing it to other categories.

An “exemplar” is used to refer to an association between a set of auditory/acoustic properties and a set of category labels (see Figure 4.3). Johnson (1997a) suggests that the phonetic details of each instance of the members of a perceptual category are stored. Along with the phonetic details, non-linguistic properties such as speaker’s gender, age, speaker’s name are also stored.

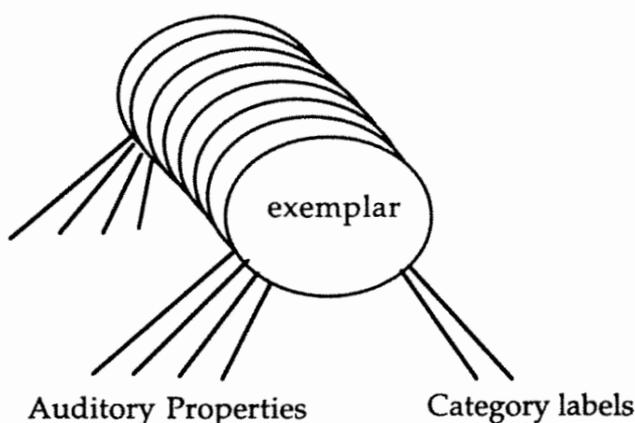


Figure 4.3 A set of exemplars relating auditory/acoustic properties with category labels (taken from Johnson, 1997a, p.148).

Let’s explain Johnson’s exemplar model with an example. When a British listener hears an /æ/ vowel (as in *path*) from a male American speaker, the listener firstly extracts acoustic properties and category information from the input: for example, acoustic properties like the F1 and F2 values of the input vowel and non-linguistic category information like *male*. Once the listener has the F1 and F2 values of the input vowel, the next step is to compare these values with the stored acoustic properties of each vowel category in his memory. For instance, he may compare the extracted F1 and F2 values with every instance of the British vowels /ɑ/ and /i/ that he has stored in his memory. The similarity between the input and /ɑ/-instances is greater than the similarity between the input and /i/-instances. Therefore, /ɑ/ beats /i/ and wins the competition between these two vowels.

The successful recognition of the input also depends on whether the listener has stored instances of /æ/ and whether a category is established for /æ/. If the listener has stored the acoustic properties and the category label for the /æ/ vowel, he should be able to recognise the input as /æ/ through comparisons. However, if the listener has no memory of the /æ/

vowel at all (this is very unlikely in reality, but let's make this assumption for the sake of explanation), it is more likely he would perceive the input as whichever category shares the greatest similarity in his memory, for example a British /a/ vowel. The top-down information *male* is expected to speed up the memory processing as it would help to filter out those exemplars which do not share the same gender label *male*. In other words, only those exemplars with a category label *male* are selected for the comparisons. Once the /æ/ category is activated successfully as a result of the acoustic comparisons, other labels which attach to these exemplars such as *American accent*, *tourists*, *Donald Trump* will be activated at the same time.

Frequency effects are also worth mentioning in the exemplar model. In Johnson (1997a), base activation level is used to adjust perception. High frequency words and recent exemplars tend to have a higher activation level than low frequency words and remote exemplars.

Attention weights are another parameter in Johnson (1997a). He argues that in categorisation, some auditory properties have a stronger effect than others. For those auditory properties which have higher degrees of sensitivity, larger attention weights are assigned. For example, in the example of perceiving the American /æ/ vowel, if other competitors are the vowels /a ε e i/, a higher attention weight will be given on the F1 dimension, as this is the main parameter to distinguish /æ/ from /a ε e i/.

4.4.2. Exemplar models in speech production

Pierrehumbert (2001, 2003) proposes an exemplar model of speech production, trying to model a complete perception-production loop based on exemplar representation. In perception, a stimulus passes through the auditory system and activates auditory/acoustic properties. The exemplar-based perceptual system compares the input with other stored properties and then determines which category is statistically most likely to underlie the current token. In production, the processing direction is the other way around. A cognitive idea/concept (e.g. I want to say "path") is created first in our brain, then the signal passes down to select the relevant label. A random sampling of the exemplar distribution is taken for the label. With the neighbourhood region of the selected exemplar activated, the average properties of the region form the production goal. According to the frequency effects inherent to exemplar theory, high frequency properties are more likely to be

selected because they constitute the statistical peak of distributions. Finally, the signal transforms into the articulation of the production goal.

Let's explain Pierrehumbert's model with an example. A Londoner wants to say "path". First of all, before the Londoner starts to plan for production, he might have heard and stored thousands of tokens of the word "path" in his lifetime. Among these tokens may be "path" with a Yorkshire accent, "path" with a foreign accent, "path" from his mother, and "path" from the policeman that just talked to him three minutes ago, etc. After he processes and stores all the input based on his previous experiences, distributions of the phonetic properties of "path" are developed; for example, glottal width for /p/, F1 and F2 for vowels etc. If we only focus on the vowel of "path" for the purpose of illustration, for a Londoner, F1/F2 values which represent the local accent, i.e. /ɑ:/, will be the most frequent tokens, compared to other F1/F2 values representing the northern representation /a/ or foreign accent representations such as /ɐ/. Therefore, the peak of distribution of F1 and F2 values of "path" should be close to the /ɑ:/ category. When he produces "path", the neighbourhood area of "path" is activated. Along with the label, the distributions of F1/F2 of the vowel are also activated. As the most frequent F1/F2 is more likely to be selected as a production goal, in the absence of any mediating factors such as attitudes and attention, the peak of F1/F2 distributions is selected, which is therefore realised as the local variant /ɑ:/ in production.

A great innovation of Pierrehumbert's production model of the exemplar-based account is that it extends Johnson (1997a)'s perception model to production, consisting of a complete perception-production loop. This is an important step for studies like speech accommodation. With the complete perception-production loop, it is now possible to make some predictions of how perceptual input from the interlocutor affects a speaker's pronunciation. A quote from Pierrehumbert (2003, p.133) explains this process:

"Each category is continually updated as the speaker perceives and encodes incoming examples; the updated distributions then provide the basis for productions by the speaker."

If we adapt this hypothesis to accommodation, in an ideal situation, a speaker starts with his own property distributions of each exemplar that he has developed based on his previous experience. When the speaker talks to his interlocutor, the distributions keep

updating. As the production goal is selected and calculated from the updated distributions, the speaker's pronunciation should shift towards the interlocutor's. However, this is only an assumption in an ideal situation. In a real conversation, a speaker faces more challenges than merely updating the distributions of his properties. For example, if the input the speaker receives is not sufficient to override the long-term experience, the speaker might still retain his previous pronunciation. Also, for non-native speakers, when talking to native speakers, what they perceive may not necessarily be identical to what the interlocutors actually pronounce. Other non-linguistic factors like the speaker's identity, the difficulty of the task and the speaker's English proficiency may also affect not only their perception but also their processing and selection of production goals.

Another issue with adopting Pierrehumbert's hypothesis in accommodation would be that if it is true that an average of the distributions is adopted as a production goal, what we should expect at the production end should be consistent. However, what we can usually observe in accommodation is that a speaker accommodates on one word but not another word, or a speaker converges at this particular time point but not 10 seconds later when he hears the same item. It seems accommodation does not behave as consistently as the hypothesis predicts. It could be that factors other than raw statistics can also affect the production outcome, such as attention and motor ability. Another reason could be that Pierrehumbert's model predicts a long-term result of updating the input, while accommodation focuses instead on short-term effects. For accommodation, maybe the salience of the new incoming item rather than its frequency plays a greater role. Pierrehumbert (2006, p.525) also claims that "*exemplar models are not sensitive to frequencies of ambient event per se, but rather to frequencies of memories.*"

As Pierrehumbert (2002) admits, the model does not predict "an exhaustive match between perception and production". She further points out that even though speakers perceive some tokens, if what they perceive is not committed to long-term memory, it does not have any influence on production. Pierrehumbert (2006, p.524), even suggests that the frequency distributions of one's perceptual system are not necessarily the same as those of one's production. However, this seems contradictory. If perception and production carry their own distribution information, how is the update of distributions possible? It is unclear to what extent perception and production share distribution information.

Johnson (1997a) also discusses the perception-production link. He suggests that people store not only auditory properties and category labels, but also articulatory properties in the exemplars. These articulatory properties could be used to direct speech production later, once they are activated. The gestural information generated while listening to others is based on one's ego exemplar. For instance, a HKE speaker's pronunciation of "three" might be [fri:]. When he hears a British person saying [θi:], what he stores for the word-initial consonant is not the gesture information of an interdental fricative [θ], but the gesture properties of his own pronunciation: a labiodental fricative [f]. When he tries to pronounce "three", the gesture properties of a labiodental fricative are activated, resulting in the HKE accent [fri:]. For Johnson (1997a)'s account, the connection between perception and production is achieved through the gesture properties that are stored in the perception processing. And more importantly, the gesture properties are based on one's own production.

It seems neither Pierrehumbert (2001)'s nor Johnson (1997a)'s model can fully explain accommodation, as they were not specially designed and developed for accommodation. However, as the perception-production loop is a key mechanism of accommodation, the models still provide a good basis for understanding accommodation.

4.4.3. Exemplar models in sociophonetics

When we discussed perception and production in the exemplar theories of the previous sections, we did not mention sociophonetic variation. In the real world, sociophonetic variation can be observed everywhere. For example, Docherty and Foulkes (2005) found that Newcastle males and females pronounce /t/ in word medial position (e.g. button) differently, with male speakers more likely to use laryngealized /t/ and female speakers to use the standard /t/. Speakers and listeners are able to associate the variants of a sound with different social categories. In the following section, how people develop sociophonetic variation in a community and how they produce sociophonetic variants will be discussed in the framework of exemplar theory.

According to exemplar-based models, when people perceive speech from other speakers, they do not merely store the phonetic properties that the speech signals carry, but also non-phonetic properties such as a speaker's identity, age, gender and accent. Social

categories associated with specific phonetic properties are stored as part of a general learning process. When the input of the same type of associations between specific phonetic variants and social categories accumulates, for example, laryngealized /t/ and Newcastle males, the strength of the association becomes stronger. With sufficient speech input, the association becomes established.

Foulkes and Docherty (2006) explain how children develop social-indexical variations in a community. They suggest that children start from clustering exemplars into three groups: adult males, adult females and children. The clustering is based on distinct phonetic differences, such as f0 and formant frequencies, between these three groups, driven by biological differences. Initially, children might only connect certain phonetic features with an individual talker, for example, associating high f0 with “mum”. With sufficient experience and exposure to more speakers, children might later develop a correlation between certain phonetic features and a group of speakers, for example, high f0 and “female”. Adults who move to a new community should follow a similar path to the one Foulkes and Docherty (2006) propose, but at a faster speed when learning new social-indexical associations.

As we have reviewed how social-indexical variations are developed and acquired, what is more relevant to accommodation is how people apply these social-indexical associations when they speak. Pierrehumbert (2001; 2003)’s production model can be used to explain this. Recalling from the previous section, Pierrehumbert (2001; 2002; 2003) suggests that production is achieved by activating a neighbourhood region of the selected label. The average properties of this region serve as the production goal. The selection of the neighbourhood can be biased by social factors. Foulkes and Docherty (2006) suggest that the ‘bias in selection’ proposed by Pierrehumbert provides the room for socio-indexical knowledge to shift the production target away from the raw statistics (p.430):

“If an individual has developed probabilistic associations between particular styles of speech and particular patterns of phonetic implementation, here too the bias could simply ensure a match for the situation of the utterance to be produced.”

As Docherty and Foulkes (2014, p.52) suggest, the sensitivity to the statistical properties would be modified by an individual’s pre-existing social-indexical knowledge, and it is

possible that distributions of phonetic properties could be overridden by social factors such as an individual's evaluation of the context of interaction.

Let's explain the model with an example. For Hong Kong people who grow up in a community where British English is valued as the most prestigious English variety, they gradually develop a social-indexical association between "RP" and "upper-class, well-educated elites". With this kind of association in mind, when they are in situations where they need to show their intelligence and educational background (e.g. a job interview or business negotiation), their English pronunciation may be more British-like compared to other situations. In the situation of a job interview, the activated neighbourhood region would include not only the area of HKE but also the area of British English. Exemplars that are associated with British English are activated. If we look at the F2 value for /ɑ:/ of "path", the most frequent F2 value should be close to the /ɑ:/ of British English, resulting in a British-like pronunciation of "path". However, this example only becomes established when the speaker's British English input is larger than his HKE input. For a case where the speaker's British English input is smaller than his HKE input, the F2 value closer to the [ɐ] of HKE would be selected as the production goal instead. Another possibility is that if the two sorts of input are roughly equal, then an average of British English F2 representations and HKE F2 representations would be taken as the production goal, resulting in a hybrid accent which sounds neither pure British nor pure HKE. Therefore, though social factors could affect production, there is still the possibility that speakers do not change their pronunciation, due to insufficient input.

The "variationist-embedded version" of the exemplar-based account consists of an ideal model for accommodation studies. On the one hand, social-indexical knowledge could affect speech perception and production through the "bias mechanism"; on the other, the completed perception-production loop explains how people accommodate and how perceptual learning contributes to production changes. As Docherty and Foulkes (2014, p.47) comment, one contribution of exemplar-based models is "*the integration of lexical and indexical dimensions without having to call upon any kind of specialized module or processing for handling social-indexical aspects of speech.*"

Though exemplar-based models have integrated speech perception, production, learning and processing, there are still many fundamental issues of the models that remain to be answered.

First of all, there is no consensus about whether acoustic properties or auditory spectrograms or other forms of speech should be used to represent exemplars. Johnson (1997b) suggests that auditory spectrograms be used in perception, while Johnson (1997a) uses formants instead. Docherty and Foulkes (2014) use waveforms in their examples, whereas Munson (2010) uses spectrograms. Another argument without consensus is whether the exemplar model is sound-based or word-based. Pierrehumbert (2001) suggests that the basic unit of exemplar representation is speech sound, whereas Johnson (2007) argues that in language people perceive words rather than sounds. In the examples above, I follow Pierrehumbert (2001)'s assumption that sounds are the unit of illustration in exemplar theories.

Both Pierrehumbert (2006) and Docherty and Foulkes (2014) mention “salience”. In sociolinguistics, “salience” is used to define different levels of sociophonetic variables. For example, “stereotypes” are more salient than “markers” according to Labov (2001, p.78)'s definitions. With the same phonetic feature, people may feel more “salience” in one speech community than in another community. This is especially true for bilingual/bi-dialectal speakers. On the other hand, what Pierrehumbert (2006) means by “salience” is more psychologically driven. For example, unfamiliar items are more easily detected and noticed than familiar items. You may not notice a change of window decorations in a local supermarket that you walk past every day, but you may notice someone dressed up in a giant dinosaur costume walking around because it is unusual. As Pierrehumbert (2006, p.525) suggests, exemplar models are sensitive to frequencies of memories but not to frequencies of ambient events. In other words, exemplar models do not explain how these salient sociolinguistic registers/experiences/events affect the model.

One comment from Docherty and Foulkes (2014) is that individual differences are usually omitted in studies of exemplar-based models, speech production and perception. Studies seem to consider their participants as ‘ideal speaker-hearers’ who would perform identically as a member of the community. In fact, individuals differ in their exposure to

phonetic features, their evaluation of interaction contexts and their experiences as learners of social-indexical associations. These factors will lead to personalised associations between certain phonetic parameters and social indexes for each individual. How individual variations interact with a broader shared communal knowledge remains a challenge for exemplar-based models.

It is worth mentioning that instead of simple exemplar theory, the hybrid exemplar-based theories are advocated here for accommodation. The hybrid exemplar-based theories allow categories and phonological coding in language processing. As Pierrhumbert (2006, p523) indicates, on one hand, “the hybrid exemplar-based theory imports the concept of levels of representation from generative models”; on the other hand, “it imports from exemplar theory the claim that probability distributions are acquired in great details through experience, that they continue to be updated in adult life, and that episodic factors can influence the way that these distributions are used in speech processing”.

4.5. Accommodation and L2 learning

The previous sections introduced some theories of accommodation from the perspectives of sociolinguistics and psycholinguistics. These theories such as CAT (Giles et al., 1991) and the interactive-alignment model (Pickering & Garrod, 2004) are directly relevant to speech accommodation, while the exemplar-based theories can be adopted to explain accommodation. However, these theories greatly lean towards accommodation studies of native speakers, and so they might not be able to fully explain the accommodation of non-native speakers. Therefore, this section aims to examine accommodation in the context of L2 learning, mainly focusing on the Interaction Hypothesis (Gass & Mackey, 2007, 2015).

Accommodation between a non-native speaker and a native speaker is different from the accommodation between two native speakers in several ways. Firstly, non-native speakers may not be able to perceive some sounds in their L2. For example, some HKE speakers might not be able to perceive the difference between fricative /z/ and /s/ because Cantonese lacks the voicing contrast for alveolar fricatives. Without perceiving the difference, accommodation is unlikely to occur for the non-native speakers. Discrimination of the pair of fricatives, in contrast, would not be a problem for native

speakers who have developed sensitivity to voicing in their mother tongue. Similar difficulty also exists in L2 learners' production. For example, due to lack of the phoneme fricative /θ/ in Cantonese, some HKE speakers might pronounce the word "thought" as [fɔt] instead of [θɔ:t]. For these HKE speakers, even if they can perceive the difference between fricative [f] and fricative [θ], they might not be able to pronounce fricative [θ], let alone converge towards native English speakers. To better understand the convergence of non-native speakers, it is important to put accommodation into a framework of L2 learning. The Interaction Hypothesis (Gass & Mackey, 2007, 2015; Long, 1996) is a suitable framework for this propose.

The Interaction Hypothesis (Gass & Mackey, 2007, 2015; Long, 1996) is developed from several L2 theories and models, including the Input Hypothesis (Krashen, 1985) and the Output Hypothesis (Swain, 2005). The Interaction Hypothesis suggests that L2 learning occurs in interactions when learners are exposed to the target language, produce the target language and receive feedback about their production from the native speakers. Initiated by Long in the 1990s, he proposes that:

"... environmental contributions to acquisition are mediated by selective attention and the learner's developing L2 processing capacity, and that these resources are brought together most usefully, although not exclusively, during negotiation for meaning." (Long, 1996, p.414)

Also, he suggests that

"...negotiation for meaning, and especially negotiation work that triggers interactional adjustments by the NS or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways." (p. 451- 452)

That is, the Interaction Hypothesis believes that negotiation in an interaction facilitates L2 learning. When there is a misunderstanding in the conversation between L2 learners and native speakers, they start to negotiate meaning. Throughout the negotiation, the L2 learners notice the differences between their output and the native speakers' input based on the feedback from the native speakers. The awareness of these differences trigger the L2 learners to adjust their production in order to match the native speakers' input. This

process can be seen as the L2 learners testing their interlanguage in negotiations, and in this way L2 learning can be eventually achieved.

This process is very similar to accommodation. Accommodation refers to the process where speakers change their pronunciation towards their interlocutors. If we put accommodation in the context of the Interaction Hypothesis and replace the speakers with L2 learners, convergence can be seen as a positive outcome of L2 learning, when L2 learners successfully adjust their pronunciation to be more native-like when they talk to native speakers in a conversation. Divergence and maintenance can also be regarded as two other types of outcome of L2 learners testing their interlanguage.

For example, HKE speakers believe that the word “three” is pronounced [fɹi:]. In their interlanguage, the fricative is labiodental fricative [f]. When they communicate with a native English speaker, if whenever they say [fɹi:] for the word “three”, the native speaker always repeats the word with [θɹi:], they may then notice the difference between their own pronunciation and the native pronunciation. At this point, they might try to adjust their pronunciation to be the same as the native speaker’s. If they successfully change the pronunciation to [θɹi:], this can be counted as **convergence**. However, they might not be able to change their pronunciation even though they notice the difference, due to lack of articulatory practice for the interdental fricative [θ]. They might still pronounce the word as [fɹi:], and this can be counted as **maintenance**. Another possibility is that they try to adjust their pronunciation but instead of producing the interdental fricative, they end with an alveolar fricative and pronounce it [sɹi:]. This can be counted as **divergence**. No matter whether convergence, divergence or maintenance, these processes are all relevant in L2 learning.

There are also some differences between accommodation and the process proposed by the Interaction Hypothesis. First of all, the Interaction Hypothesis emphasizes the process whereby L2 learners and native speakers negotiate meaning in the interaction, and the L2 learners learn through the native speakers’ feedback. However, convergence does not necessarily occur in a negotiation, and it may occur even if there is no mismatch of information in the conversation. One example is given below to illustrate this point.

Negotiation	1	NS: I have a car-[ɪ] park on my map.	} Convergence
	2	NNS: I have car-[ɪ] park too.	
	3	NS: You have a car-[ɪ] park?	} Convergence
	4	NNS: Yeah, I have a car-[ɪ] park.	

Table 4.3 An example of conversation between a native English speaker (NS) and a non-native English speaker (NNS).

Table 4.3 shows a conversation about car parks between a native English speaker and a non-native English speaker. According to the Interaction Hypothesis, learning is expected in a negotiation (i.e. from line 2 to line 4) when the non-native English speaker neglected the English particle “a” in his/her production (i.e. line 2) and the native speaker gave feedback by providing a correct form (i.e. line 3). The non-native speaker realised the difference and then repeated the native form (i.e. line 4). In this process, convergence is observed from the non-native speaker towards the native speaker in line 4.

Besides this, convergence is also possible in line 2 where the non-native speaker converges towards the native speaker on the rhoticity of the word “car”. In this part, there is no negotiation but convergence is still possible. The Interaction Hypothesis is interested in the process from line 2 to line 4 where there is a negotiation, while convergence can still occur even if there is no negotiation but merely repetition.

It is also worth mentioning the four types of implicit feedback proposed in the Interaction Hypothesis: confirmation checks, clarification requests, comprehension checks and recasts. Table 4.4 lists the definitions and an example for each type of implicit feedback.

Type	Explanation	Examples
Confirmation checks	expressions that are designed to elicit confirmation that an utterance has been correctly heard or understood;	<i>Is this what you mean?</i>
Clarification requests	expressions that are designed to elicit clarification of the interlocutor’s preceding utterances;	<i>What did you say?</i>
Comprehension checks	expressions that are used to verify that an interlocutor has understood;	<i>Did you understand?</i>
Recasts	a rephrasing of an incorrect utterance using a correct form while maintaining the original meaning.	<i>NNS: I have car park. NS: You have a car park? (from Table 4.3)</i>

Table 4.4 Four types of implicit feedback extracted from Gass and Mackey (2007, p.181-182).

Convergence may occur when L2 learners are giving or receiving these types of feedback. Again, convergence is not restricted to these four scenarios, and it might occur even if there is no negotiation.

Secondly, attention and awareness of the input are crucial in L2 learning and in the Interaction Hypothesis, but people do not necessarily need to be aware of when they converge. Schmidt (2001) suggests that learning is not possible without awareness because learners need to be aware of the input in order to internalize it. Attention can be seen as a mechanism that learners use to tune their focus into specific parts of the input they receive (Gass & Mackey, 2007). On the other hand, depending on the theories/models of accommodation adopted, convergence may involve some level of automaticity and unconsciousness. For example, the interactive-alignment model proposes that convergence is an automatic process, whereas Giles et al. (1987) argue that convergence should not be seen as an entirely conscious process.

To sum up, convergence can be regarded as a positive outcome of L2 learning in a conversation, but it is not only constrained by overt negotiation between native speakers and non-native speakers. Divergence and maintenance are two other possible outcomes of L2 learners testing their interlanguage.

4.6. Summary

In this chapter, theories and models relevant to speech accommodation have been reviewed. CAT (Giles et al., 1991), as a dominant sociolinguistic theory of accommodation, received a lot of attention until the 1990s. It accounts for why people accommodate from a social psychology perspective. With research growing in disciplines like psycholinguistics, speech perception and production, scholars started to consider how people accommodate, trying to discover the underlying mechanism of accommodation. The interactive-alignment model (Pickering & Garrod, 2004) is one of these attempts. The interactive-alignment model draws an ideal model for dialogue. Automatic alignments at all levels account for the success of a conversation. However, the model does not explain how signals align at each level, making it less applicable to actual cases.

Exemplar-based models have been attracting scholars' interests since the late 1990s. It suggests that no abstract representations of sounds are used in recognising sounds. Instead, exemplars associated with both linguistic properties and non-linguistic categories are stored in memory (Johnson, 1997a). A new stimulus activates the most similar exemplars by the calculation of distances. Items associated with the exemplars such as articulatory properties and social categories are also activated. Exemplar-based models also account for production (Pierrehumbert, 2001, 2002, 2003). Starting with a concept, the corresponding labels are activated. A random sampling of the exemplar distribution is selected for the labels. The neighbourhood region of the selected exemplar is activated, and the average properties of the region then make up the production goal. Finally, the signal transforms into the articulation of the production goal. Pierrehumbert's model extends Johnson (1997a)'s model from perception to production. It completes the perception and production loop, which makes it a possible model to explain accommodation.

The Interaction Hypothesis (Gass & Mackey, 2007, 2015) helps to understand convergence from the perspective of L2 learning. Convergence, divergence and maintenance can be seen as three possible outcomes of L2 learners testing their interlanguage. The Interaction Hypothesis proposes that, by receiving native speakers' feedback in the negotiations, L2 learners would be able to compare their pronunciation with the native speakers' pronunciation, and the awareness of the differences would facilitate L2 learning. This might be the case for convergence of non-native speakers.

The present project aims to investigate convergence by non-native English speakers towards native English accents, and the effect of language attitudes on convergence. If the above theories/models are applied to the present study, CAT predicts that people would converge more towards the accent they favour; the interactive-alignment model suggests that non-native English speakers would converge towards native speakers, not only at the phonological level but also at the syntactic level; the exemplar-based theories also predict convergence of non-native speakers, indicating that the input from native speakers would update the non-native speakers' distributions of categories which would provide a basis for the non-native speakers' production. In the next chapter, a pilot study of speech accommodation will be reported.

Chapter 5. Review of studies of short-term accommodation and a pilot study

In Chapter 4, a series of theories relating to speech accommodation were reviewed. This chapter will first review a few classic studies of convergence, followed by a pilot study of Study 2 (see Chapter 6), that is, convergence of non-native English speakers towards native English accents.

5.1. Literature review of convergence studies

In the 1980s and 1990s, convergence was widely studied when CAT (Giles et al., 1991) became a prevalent sociolinguistic theory. These studies aimed to test CAT and examined whether different social status or ethnic backgrounds affected speakers' accommodation. Later in the early 2000s, Pickering and Garrod (2004)'s proposal of an automatic account of alignment in dialogue elicited a few more studies of convergence, such as Babel (2010, 2012), Pardo (2006), Pardo, Jay and Krauss (2010), Pardo, Jordan, Mallari, Scanlon and Lewandowski (2013), Pardo, Gibbons, Suppes and Krauss (2012) and Pardo, Urmanche, Wilman and Wiener (2017). The following will mainly review studies of convergence conducted in the last 20 years, aiming to provide an updated picture of the study of convergence.

Babel (2010) examines dialect convergence and divergence in New Zealand English. 44 New Zealanders completed a shadowing task (Goldinger, 1998) which required them to repeat a list of words after an Australian model talker. The list consisted of words with the vowels in KIT, DRESS, TRAP, START, STRUT and THOUGHT. The participants' implicit attitudes towards Australia and New Zealand were also collected in an IAT. In addition to the implicit attitudes, the study also assigned the participants into two groups: one group was primed with a positive message of the Australian model talker and Australia, for example, participants were told the model talker was born in Auckland and would like to look for employment in New Zealand; the other group was primed with a negative message of the model talker and Australia, for example, participants were told the model talker thinks New Zealanders are stupid and lack culture. The results suggested that the explicit priming messages had no effect on the New Zealand participants' convergence towards the Australian model talker, but the participants' implicit attitudes

elicited by the IAT did. The more positive the participants felt towards Australia, the more they converged in the shadowing task. Another important finding from Babel (2010) is that not all the vowels underwent the same amount of convergence. The DRESS vowel was found to converge the most out of the five vowels. Babel (2010) explains that the DRESS vowel is one of the most distinct vowels between Australian English and New Zealand English. However, New Zealand English speakers are not particularly sensitive to DRESS vowels across dialects. Therefore, the sufficient phonetic space between the New Zealand participants' DRESS vowels and the Australian model talker's DRESS vowels had provided the room for the convergence, and the lack of sensitivity to DRESS vowels between dialects might further facilitate the accommodation.

Babel (2012), which is similar to Babel (2010), investigates the phonetic convergence of American college students and the role of attractiveness of model talkers in convergence. 111 American college students completed a shadowing task in which they were exposed to the model talkers' (either a black talker or a white talker) pronunciation of words with the five vowels /i æ a o u/. Half of the participants were exposed to the white talker while the other half were exposed to the black talker. In addition, within each group, half of the participants saw a picture of the talkers, while the other half did not. For those participants who could see the picture of the talkers, their attitudes (i.e. attractiveness ratings) towards the model talkers were collected. Some results from Babel (2012) echo Babel (2010). Firstly, the selectivity of vowel convergence was also found in Babel (2012). The participants did not converge the same on all five vowels: /æ/ significantly converged more than the other vowels. Secondly, the participants' attitudes towards the model talkers had some impact on their convergence. When the participants were exposed to the white model talker, the more attractive the females rated the white talker, the more they converged. In contrast, for the males, the more attractive they rated the white talker, the less they converged. No significant correlation was found between the participants' attitudes and the degree of convergence for the black talker. Thirdly, the male and the female participants converged differently on some of the vowels. For example, for the participants who did not see the picture of the talkers in the shadowing task, the females converged more on /æ/ than the males, while the males converged more on /ɑ/ than the females.

Similar to Babel (2010), where a greater effect of convergence was found for the DRESS vowel, one of the most distinct vowels between Australian English and New Zealand English, Babel (2012) also found the strongest convergence on /æ/ and /ɑ/ for the participants whose dialects had the most distinct features from GenAmE in these vowels. These findings lead to an indication that the acoustic-phonetic space of the vowels between participants and model talkers may affect the degree of convergence. Babel (2010) and Babel (2012) shows that the larger the acoustic-phonetic space a vowel has, the more likely it is to converge.

Another similarity between Babel (2010) and Babel (2012) is that both studies found a significant correlation between people's attitudes and their convergence. In Babel (2010), the attitudes were represented by the implicit attitudes towards Australia and New Zealand elicited by the IAT, while in Babel (2012) the attitudes refer to people's ratings towards the model talker's attractiveness. These findings suggest that not only implicit attitudes but also explicit attitudes could affect convergence, which provides a strong support for CAT.

Pardo and her colleagues also conducted a series of studies on convergence in a conversational setting (Pardo, 2006; Pardo et al., 2010) and in a shadowing task (Pardo et al., 2013; Pardo et al., 2017). Studies focusing on convergence in a conversational setting mainly investigate the effects of talker's sex and conversational role on convergence. Pardo (2006) investigated six same-sex pairs' (i.e. three male pairs and three female pairs) convergence in a map task. Before and after the map task, a pre-task and a post-task (which involved reading a word list) were also conducted. Different from Babel (2010, 2012), where acoustic measurements of vowels were used, Pardo (2006) used a AXB perceptual similarity test (Goldinger, 1998) to judge the degree of convergence. The AXB perceptual similarity test relies on naïve listeners' judgements on whether an item before shadowing (the A item) or an item after shadowing (the B item) sound more similar to the model talker's item (the X item). If the items after priming are selected more than the items before priming in the AXB similarity test, convergence is supported. For example, when comparing the items produced in the map task (i.e. the items after priming) and the items produced in the pre-task (i.e. the items before priming) in Pardo (2006), 65% of the judgements indicated that the map task items were more similar to the model talker's items, suggesting a convergence.

More importantly, Pardo (2006) found that a talker's sex and their role in the conversation affect their convergence. The results suggested that the female participants who gave instructions in the map task (henceforth givers) converged towards their female partner who received the instructions (henceforth receivers), but convergence was not found from the female receivers towards the female givers. For males, both receivers and givers showed convergence, but the male receivers demonstrated a stronger convergence than the male givers.

Similar effects of talker's sex and conversational role on convergence were also found in Pardo et al. (2010), where male participants converged more than female participants and givers converged to receivers but not the other way around. Pardo et al. (2010) also found an above-chance convergence in the map task and the post-task based on the AXB perceptual similarity judgements. What was new in Pardo et al. (2010) is that one member of each pair was instructed to imitate the other interlocutor. The results suggested that when the receivers were instructed to imitate, they showed a convergence; however, when the givers were instructed to imitate, no convergence was found. This finding suggests that the imitation instruction had an impact on how people converge. Another interesting finding in Pardo et al. (2010) is that acoustic measurements of vowel convergence showed no significant correlation with the results of the AXB perceptual similarity judgements. The acoustic measurements based on Euclidean distance suggested that no convergence was found among talkers, and givers who were instructed to imitate showed a divergence. Male and female pairs did not differentiate in their vowel convergence measurements either.

Note that the sample size in Pardo (2006) was small, as only six pairs of talkers were used in her experiment. Pardo et al. (2010) extended the scope of the sample size to 12 pairs of talkers, and a weaker convergence effect was found compared to Pardo (2006) overall. The percentage of convergence in Pardo (2006) was 65%, while in Pardo et al. (2010) it was only 53%. This indicates that the effect of convergence in a conversational setting is relatively small. It could be that completing a map task involves a higher cognitive load, which might block convergence. Even for the participants in Pardo (2006) and Pardo et al. (2010) who are native speakers of English, the effect of convergence was relatively

small. For non-native speakers of English, the effect of convergence might be even smaller.

A larger-scale study consisting of 108 talkers was conducted by Pardo and her colleagues (Pardo et al., 2017). Different from Pardo (2006) and Pardo et al. (2010), Pardo et al. (2017) investigated convergence in a shadowing task (Goldinger, 1998) which allows a strict control over word frequency and word type. 12 females were selected as model talkers, and the rest of the participants were divided into a same-sex female group, a same-sex male group and a mixed-sex group. The results suggested that participants of the same-sex groups did not converge differently from the participants of the mixed-sex group, and the interaction between the sex of the model talker and the sex of the shadowers was not significant. This finding challenges several previous studies in which the participants' sex was found to be a significant factor in convergence (Namy, Nygaard & Sauerteig, 2002; Pardo, 2006; Pardo et al., 2010). Pardo et al. (2017) explained that the significant effect of sex in the previous studies might be due to the low frequency words used in the experiments, as the female participants in Pardo et al. (2017) were found to converge more on low frequency words than on high frequency words. In addition, Pardo et al. (2017) also found that word frequency did not affect convergence, whether the frequency was treated as a continuous variable or as a categorical variable. Interestingly, word type was a significant predictor for convergence, in that bisyllabic words converged more than monosyllabic words.

Another novel finding of Pardo et al. (2017) is that among the five acoustic measurements of the vowels, significant convergence was found in duration only. Marginal convergence was found in Euclidean distance and F2, and no significant effect was found in F1 and f0. This aligns with the findings in Pardo et al. (2013); i.e. that talkers tended to converge and diverge in multiple dimensions of the sounds at the same time, and the patterns were chaotic. Furthermore, Pardo et al. (2017) found that all the acoustic attributes (i.e. duration, F1, F2, f0 and Euclidean distance) were significant predictors for the results of the AXB perceptual similarity test. In other words, listeners in the AXB perceptual test relied on all these acoustic parameters when they made their judgements. Duration and f0 were the strongest predictors of all.

Babel and Pardo mainly investigate convergence in a laboratory setting, either using the shadowing paradigm (e.g. Babel, 2012; Pardo et al., 2017) or eliciting conversational speech through the map task (e.g. Pardo, 2006; Pardo et al., 2010). There are other studies that focus on convergence in a more natural setting; for example, Pardo et al. (2012) and Evans and Iverson (2007) investigate convergence between college roommates.

Pardo et al. (2012) collected speech samples from four pairs of college roommates at the following times: late August, late October, December and January the year after. Students were asked to read five sets of American English vowels and two utterances of two sentences at each time point. Their attitudes towards their roommate (i.e. the closeness they felt towards their roommate) were also collected. The results of AXB perceptual similarity tests suggested that all the pairs except one showed significant convergence overall. The acoustic measurements of Euclidean distance also suggested a reduced distance of the vowels between roommates from August to October. However, great variations between pairs and the words that they converged on were found. Interestingly, a positive correlation between closeness ratings and convergence was found at the time of December, suggesting that the closer the students felt towards their roommate, the more they converged towards him/her after 3.5 months of cohabitation.

Evans and Iverson (2007) investigated 19 college students' long-term convergence towards standard southern British English after they had attended university for three months, for one year and for two years. These students all came from northern England and spoke English with a northern accent before they attended university. For example, for the word "bud" and "bath", they would pronounce them with [ʊ] and [a] in a northern English accent, whereas in standard southern British English these words are pronounced with [ʌ] and [ɑ]. Different from Pardo et al. (2012), Evans and Iverson (2007) examined not only the students' changes in production but also their changes in perception, and investigated the correlation between perception and production. The results of accent ratings suggested that the northern students changed their accents to be more southern over time. This also reflected in the acoustic measurements of vowels, in that they shifted their BATH vowel to be more back and higher, and centralised their BUD vowel. Regarding perception, the students did not seem to change their perception of the BUD and BATH vowels after living in a multidialectal environment for two years. Although the students performed better in recognising words in noise when the speech was

presented in standard southern British English, the recognition did not improve after two years' living in the south. Though no direct correlation between perception and production was found, there was some evidence to suggest that perception can be affected by the changes in production. For example, the students who produced more southern vowels overall performed better in identifying standard southern British English speech in noise compared to those who produced more northern vowels overall. These findings suggested that production changes can occur without changes in perception. Evans and Iverson (2007) explained that listeners might have a high perceptual tolerance of acoustic variations which allow them to change their accents without changing the underlying category representations. Another explanation was that perceptual adaptation might be talker-specific, and more long-term experience would be needed to adapt accent-general patterns.

It is worth noting that no matter in Pardo (2006), Pardo et al. (2010) or in Babel (2010, 2012), or in Evans and Iverson (2007), the participants mostly come from a homogenous dialect background in each of these studies. For example, the participants in Babel (2012) were mostly born and raised in California, while the northern English participants used in Evans and Iverson (2007) were from the same town in the Midlands, the UK. These participants presumably would speak a similar accent. When examining their convergence, the participants would start from a similar baseline. This might explain why an overall effect of accommodation is more likely to be found in these studies. In contrast, for a group of non-native English speakers who have huge variations on their accents, the homogenous pattern of convergence might not be found. Kim, Horton and Bradlow (2011) showed an example of this.

Kim et al. (2011) studied convergence in pairs who either shared the same language/dialects, or between native English speakers and Korean/Mandarin L2 English speakers. The design of mix-matching talkers from various language backgrounds allowed them to test the role of language distance in convergence. 32 pairs of talkers completed a diapix task (Van Engen, Baese-Berk, Baker, Choi, Kim & Bradlow, 2010) in which each participant was given one of two pictures and asked to find ten differences with his/her game partner. Similar to the map task used in Pardo (2006), the aim of the diapix task is also to elicit spontaneous conversation between the talkers. The 32 pairs consisted of three combinations: 8 pairs of talkers shared the same L1 and same dialect

(e.g. both were native speakers of English or Korean); 8 pairs of talkers shared the same L1 but with different dialects (e.g. an American English speaker from New York paired with an American English speaker from Florida) and 8 pairs of talkers had different L1s (e.g. a native speaker of English paired with a Korean learner of English). The results suggested that convergence of talkers with the same L1 and same dialect was significantly larger than the convergence of the other two groups. No significant difference was found between the convergence of talkers with the same L1 but different dialects and the convergence of talkers with different L1s. These findings indicated that closer language distance between talkers facilitates convergence.

As admitted in Kim et al. (2011), one limit of their study was the XAB perceptual similarity test they used, where naïve listeners were asked to compare three utterances rather than single words. One example is given below with three utterances copied from Kim et al. (2011, p. 149).

Position	Content	Speaker/Time point
X-utterance	<i>do you have a beehive</i>	Talker 1 / early stage
A-utterance	<i>is it at the top</i>	Talker 2 / early stage
B-utterance	<i>does the little boy have a visor</i>	Talker 2 / late stage

Table 5.1 Examples of utterances used in XAB perceptual similarity test in Kim et al. (2011). “early/late stage” in the column of “Time point” means that the utterance was produced in the early/late stage of the conversation.

As shown in Table 5.1, the three utterances were completely different. The listeners were asked to judge whether the A or B utterance sounded more similar to X. This might introduce a huge difficulty for listeners. It is worth noting that the overall percentages of convergence for the three groups were relatively low. Only the percentage for the same-L1/same-dialect pairs reached around 60%; the percentages for the group of same-L1/different-dialect and for the group of different-L1 were very close to 50%. It could be that the significant difference found between the same-L1/same-dialect group and the other two groups was due to the close language distance of the talkers helping the listeners to compare three completely different utterances in the XAB perceptual similarity test. When the utterances were from talkers with different dialects or different L1s, the listeners found it even more difficult to compare them.

Also, as Kim et al. (2011) only compared utterances, it is not clear how language distance affected the convergence of specific sounds. Interestingly, among the eight pairs of different-L1 speakers (e.g. a native English speaker paired with a Korean/Mandarin learner of English), non-native English speakers in three pairs were found to converge towards the native speakers, and in two pairs were found to diverge from the native speakers. In three pairs no changes were found. This reveals a complex picture of convergence for non-native English speakers.

This section has reviewed several studies of convergence from the past two decades, including laboratory-based studies such as Babel (2010, 2012), Pardo (2006) and Pardo et al. (2010), longitudinal studies such as Pardo et al. (2012) and Evans and Iverson (2007), and studies of convergence of speakers with different language backgrounds such as Kim et al. (2011). From these studies, several characteristics of convergence can be concluded. Firstly, convergence is not an entirely automatic process; social factors such as language attitudes and the conversational role of talkers affect the degree of convergence. This is shown in Babel (2010, 2012) where people converged more towards the talker/accent they favoured more. Pardo et al. (2012) also found that college students converged more if they had a higher rating of closeness towards their roommates. These findings reject the automatic account proposed by the interactive-alignment model (Pickering & Garrod, 2004). Secondly, convergence does not occur in all sounds and acoustic parameters for all speakers - selectivity of convergence is common. For example, Babel (2012) found that participants converged more on the /æ/ and /ɑ/ vowels, and Pardo et al. (2017) found that for the same vowels people converged on duration and F2 but not on F1 and f0. Kim et al. (2011) suggested that only 9 out of 32 pairs showed convergence and Pardo et al. (2017) found that 12 pairs of mixed-sex talkers showed distinct patterns of convergence, indicating that there might be huge individual differences in convergence. Thirdly, there are various ways of measuring convergence. Babel (2010, 2012) used the acoustic measurements of F1 and F2 of vowels, while Pardo (2006) used AXB perceptual similarity tests. Pardo et al. (2010) suggested a non-significant correlation between acoustic measurements and perceptual similarity tests, whereas a significant correlation was found later in Pardo et al. (2017).

Most of the studies of convergence focus on native speakers and convergence between non-native speakers and native speakers receives less attention. In the next section, a pilot

study of convergence between HKE speakers, British English speakers and American English speakers will be presented. Along with Study 2 in Chapter 5, these two studies aim to investigate the accommodation of non-native speakers (i.e. HKE speakers) towards native English accents.

5.2. Pilot study

In the previous chapters and in the last section, a series of theories relating to speech accommodation have been reviewed, as well as some classic studies of convergence. This section will shift from literature review to empirical study, aiming to answer a key research question: **do non-native English speakers accommodate to native accents during and after a short conversation?**

The interactive-alignment model, the exemplar-based theories and the Interaction Hypothesis reviewed in Chapter 4 all predict convergence. Therefore, we hypothesize that **non-native English speakers are able to accommodate to native accents after a short-term exposure.**

To test this hypothesis, a pilot study and a main study (see Chapter 6) were conducted. Both the pilot study and Study 2 chose HKE speakers as participants. The experimental design is similar for both studies: the HKE participants talked to a native English speaker (either RP or GenAmE) in a carefully designed experimental setting. A pre-task and a post-task were included to measure the changes in the participants' accent before and after the exposure to the native speaker. This section will focus on the pilot study.

5.2.1. Experimental design

5.2.1.1. Participants

Six Hong Kong students were recruited from the University of York. Four of them were visiting students from the Chinese University of Hong Kong, two were master's students from the University of York. All of them confirmed that they had had little overseas experience in an English-speaking country before completing the tasks. For the six participants, two are in the age group 18-25 years old and four are in the age group 26-30 years old. Their English proficiency was around IELTS 7 based on a background information questionnaire they completed before the experiment.

One RP speaker from southern England and one GenAmE speaker from California were recruited as the native interlocutors. They were both female master's students from the University of York, majoring in linguistics. They were informed about the aims of the study and the procedures of the experiments beforehand.

The six participants were divided into two groups according to the accents of their interlocutors. Three participants communicated with the RP speaker (henceforth the RP group) and three communicated with the American speaker (the GenAmE group). In the RP group, there were two female participants and one male participant (HK4). The same native female RP speaker was paired with all the members of the RP group in the map task. Similarly, the same GenAmE female speaker was paired with all members of the GenAmE group in the map task.

5.2.1.2. Materials

Two vowels were selected as variables in the experiment: the LOT vowel and the PATH vowel, as shown in Table 5.2. These two vowels were selected based on the phonological differences between HKE, RP and GenAmE.

	RP	GenAmE	HKE
LOT vowel	/ɒ/	/ɑ/	/ɒ/ or /ɔ/
PATH vowel	/ɑ:/	/æ/	/a/ or /ɑ:/

Table 5.2 LOT vowels and PATH vowels in RP, GenAmE and HKE.

RP vs HKE

Deterding conducted a series of studies on the formants of monophthong vowels in different varieties of English, including RP (Deterding, 1997) and HKE (Deterding et al., 2008). This allows us to compare RP with HKE, as similar methods were used in these studies. For example, they used the same formula of the Bark scale, they extracted vowels from connected speech (distinguishable from word list elicitation), and they had vowel formants from female speakers.

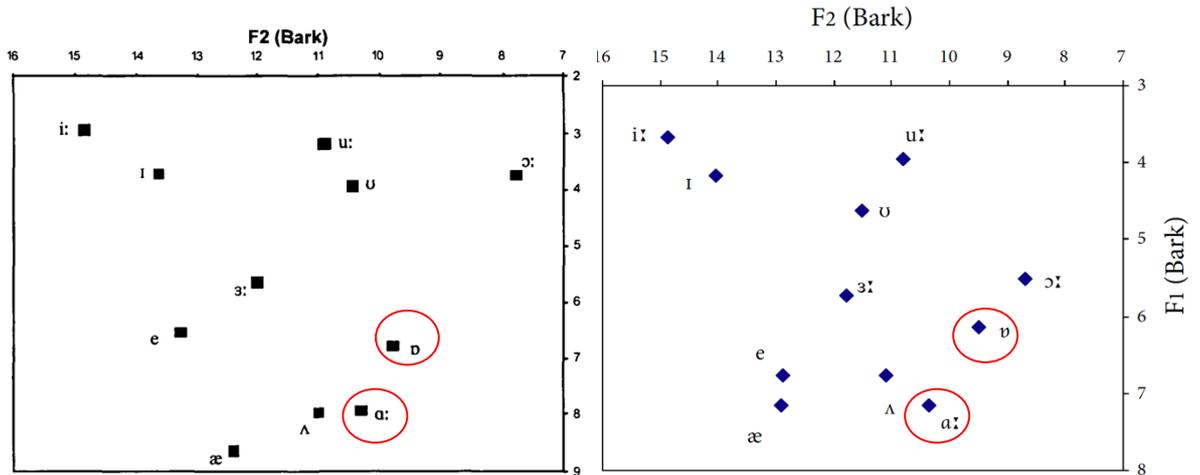


Figure 5.1 The left panel, reproduced from Deterding (1997, p.51), shows vowels produced by female RP speakers; the right panel, reproduced from Deterding et al. (2008, p.162), shows vowels produced by female HKE speakers. The LOT and PATH vowels in RP and HKE are highlighted in red circles. Note that y axes are different across the two figures.

To compare RP and HKE, the formant data provided in Deterding (1997) was used to plot the LOT and PATH vowels of RP in Figure 5.2. Regarding the HKE vowels, as Deterding et al. (2008) did not provide the exact formant data of the HKE vowels they plotted, the formant values of the LOT vowel and the PATH vowel for HKE were read and estimated from Figure 5.1(right).

According to Figure 5.2, the RP-*LOT* vowel is lower than the HKE-*LOT* vowel, and it is the same for the RP-*PATH* vowel. Note that the observed differences might still be biased due to physiological differences as the vowels were from two separate studies.

GenAmE vs HKE

Unfortunately, Deterding did not conduct a study on GenAmE which would have allowed us to compare GenAmE and HKE in the same way as the previous section did. The comparison between GenAmE and HKE is therefore less straightforward.

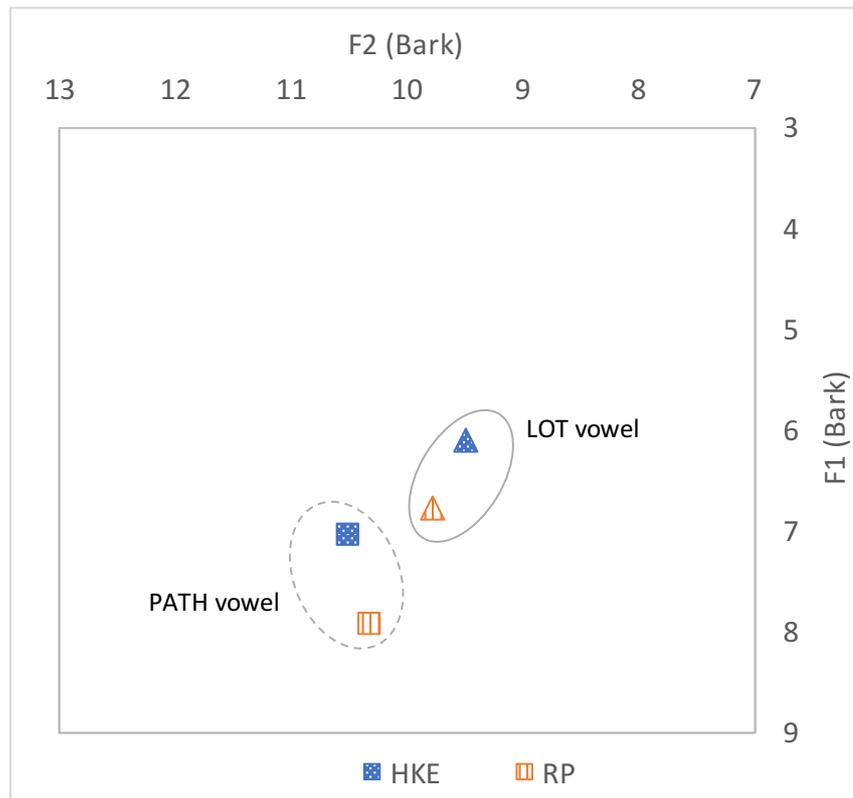


Figure 5.2 The LOT vowel and PATH vowel of RP and HKE based on Deterding (1997) and Deterding et al. (2008).

To compare with HKE, formant values of GenAmE vowels were selected from Hagiwara (1997). There are two reasons for using Hagiwara (1997) as a reference for GenAmE. First, the GenAmE speaker of the pilot study is from California, and Hagiwara (1997) reported formant values of female American English speakers from southern California. Second, the formant values reported in Hagiwara (1997) are similar to the formant values reported in Yang and Whalen (2015), who collected data from nine female American speakers, suggesting a high validity of the Hagiwara (1997) results.

Since only formant values in hertz were provided in Hagiwara (1997), the formants of the LOT vowel and PATH vowel in Hagiwara (1997) were transformed into Bark values using the formula from Deterding et al. (2008). The transformed Bark values of the LOT vowel and PATH vowel from Hagiwara (1997) are plotted alongside HKE and RP in Figure 5.3.

Figure 5.3 shows that the GenAmE-*LOT* vowel is lower and more front compared to the HKE-*LOT* vowel. The GenAmE-*PATH* vowel is also lower and more front.

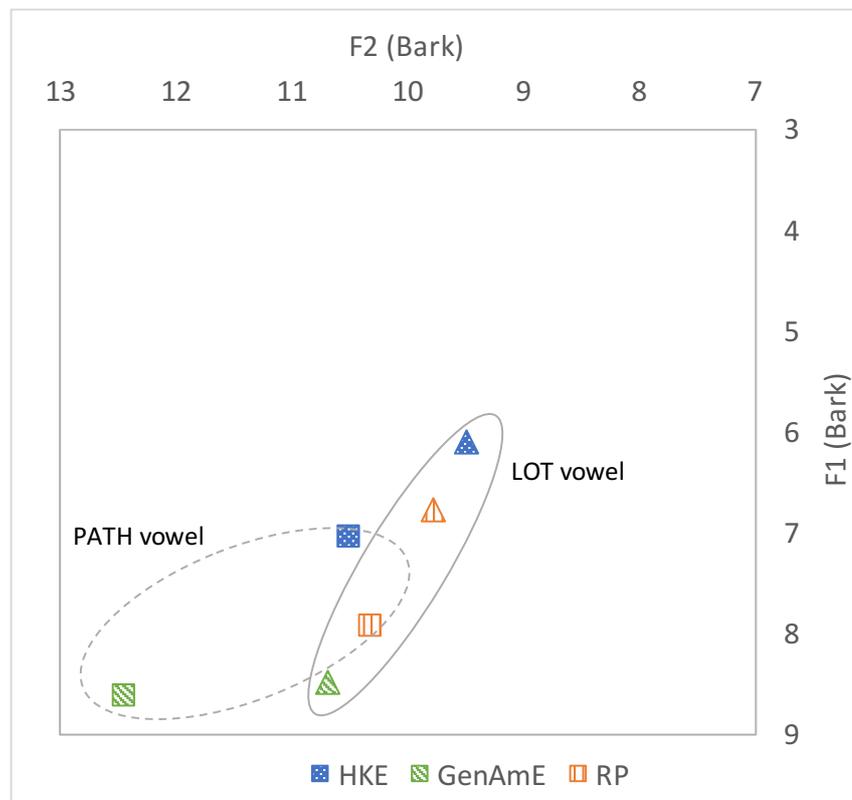


Figure 5.3 The LOT and PATH vowels of GenAmE (data from Hagiwara, 1997), RP (data from Deterding, 1997) and HKE (data from Deterding et al., 2008).

Though some preliminary predictions can be drawn from Figure 5.3, there are still some limitations. First, though similar methods were used in Deterding (1997), Deterding et al. (2008) and Hagiwara (1997), the three studies still differ in other ways. For example, Hagiwara (1997) used a word list method to elicit vowels, whereas Deterding (1997) and Deterding et al. (2008) extracted vowels from connected speech. Second, generally speaking HKE is closer to RP, but some HKE speakers may have strong American English accents. As Deterding et al. (2008) only included HKE speakers who were influenced by British English, the HKE in Deterding et al. (2008) may not be able to represent those HKE speakers who speak English with an American accent.

The purpose of estimating the vowel locations of HKE, RP and GenAmE is to provide a general overview of the three varieties. However, Figure 5.3 should be interpreted with caution.

5.2.1.3. Procedure

There were three tasks in the experiment: a pre-task, a map task and a post-task.

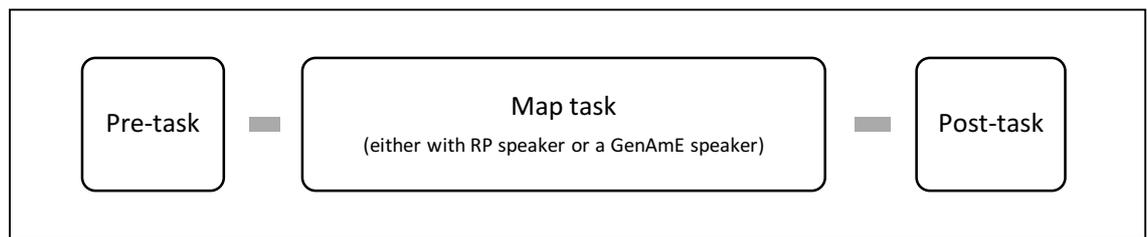


Figure 5.4 Procedure of the experiment design for the pilot study.

Each participants was seated on one side of a table. The native interlocutor who came into the recording studio for the map tasks was seated on the other side of the table. A paper board was placed in the centre of the table to prevent the interlocutor and participants from seeing one another. One lapel microphone was given to the participant and one to the interlocutor. The whole process was monitored and recorded by a technician. The procedure of the experiment was explained to the participants at the start. They were not told the purpose of the experiment. Training about the aims and the procedure of the experiment was given to the two native interlocutors beforehand. The training to the native interlocutors should not affect the experiment as the focus was the HKE speakers, not the native interlocutors.

Pre-task: The participants received two pictures containing images of dogs and passes (“pass” refers to a kind of ticket for museum or theme park; see Figure 5.5 for the pictures used for the LOT vowel). They were asked to describe what they could see on the pictures, including number of the animals, colour and location of the objects. The pre-task took them around two minutes to complete.

Map task: Each participant worked with an interlocutor who was either a native speaker of RP or a native speaker of GenAmE, as noted above. The Hong Kong participants were given the participant version of the maps while the native interlocutors were given the interlocutor version (see Figure 5.6). The Hong Kong participants were asked to draw a route on their maps with the help of the native speakers. The conversations were recorded. Each map task took them around 10 minutes to complete.



Figure 5.5 Pictures used in the pre-task (top) and the post-task (bottom) for the LOT vowel.

Post-task: The procedure of the post-task was the same as the pre-task. The participants received two pictures which were similar to the one they saw in the pre-task (see Figure 5.5). They were asked to describe the pictures using their own words. The post-task took them about 2 minutes to complete.

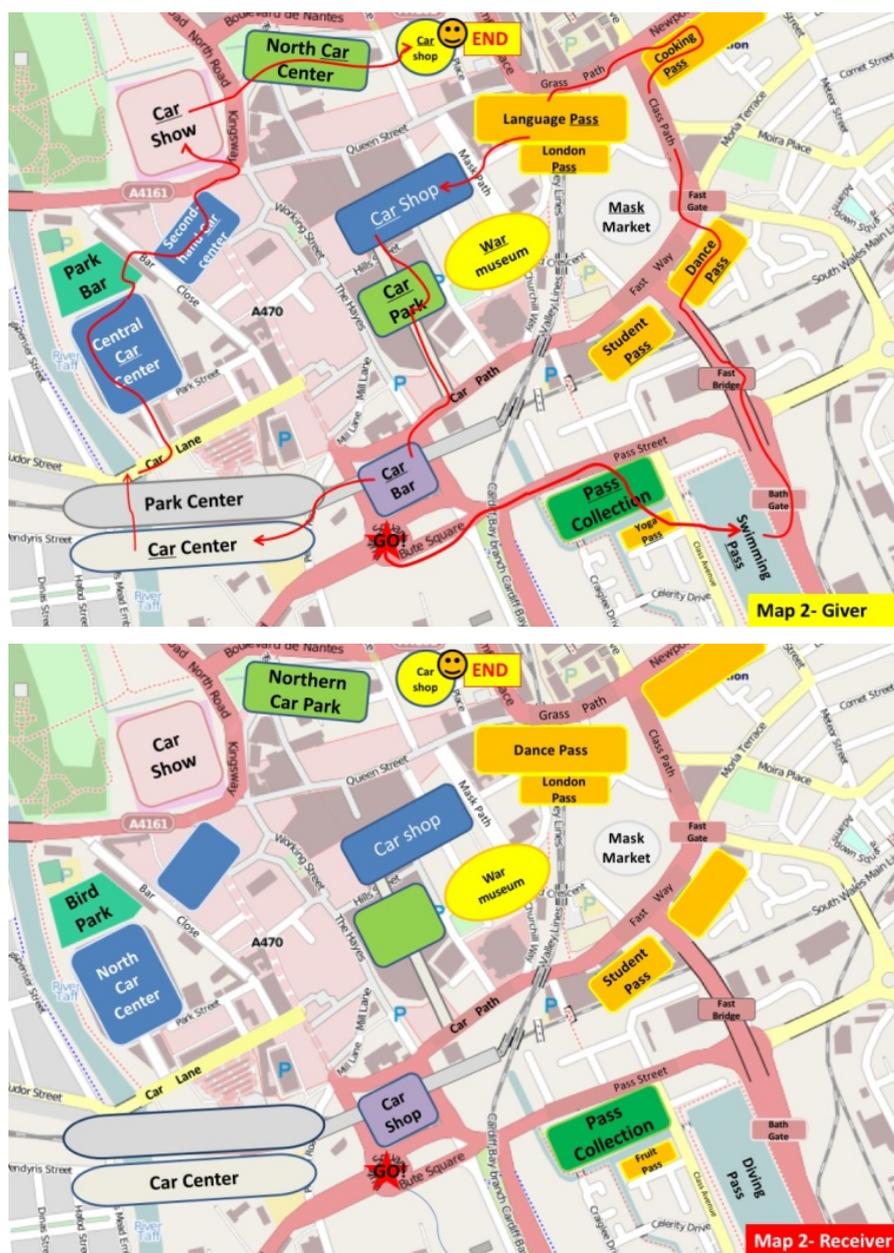


Figure 5.6 Maps used in the map tasks for the PATH vowel. On the top is the interlocutor version of the map, at the bottom is the participant version of the map. The participant version only contained a starting point and some of the landmarks, while the interlocutor version contained all of the landmarks and a route from a starting point to an end point.

5.2.1.4. Data analysis

The participants' and the native interlocutors' recordings were imported into Praat (version 6.0.28: Boersma & Weenink, 2017) for annotation. For all the recordings, all words containing the LOT vowel and PATH vowel were annotated. F1 and F2 values at the midpoint of the vowels were extracted using Praat scripts. On average, each

participant produced 10 tokens of LOT/PATH vowel in the pre-task and the post-task, and produced 35 tokens of LOT/PATH vowel in the map tasks.

As one of the participants was a male speaker, normalization was needed in order to compare speakers. The Watt and Fabricius method (Fabricius, Watt & Johnson, 2009; Watt & Fabricius, 2002) was used to normalize the target vowels. The Watt and Fabricius method uses a centroid point which is calculated from three corner vowels /i æ u/ to normalize all the vowel formants. A set of corner vowels were extracted from each speaker for normalization.

5.2.2. Results

Results will be shown in two sections. The first section will show the mean F1 and F2, and the Euclidean distance of the LOT vowel. The second section will show the results for the PATH vowel. Euclidean distance is a mathematical calculation which measures the distance between two points.

5.2.2.1. LOT vowel

Though Figure 5.3 showed that the RP-*LOT* vowel is lower than the HKE-*LOT* vowel, our data suggests that the LOT vowel of the RP speaker is higher than the LOT vowels of the participants, as shown in Figure 5.7.

Figure 5.7 shows that HK1 and HK3 from the RP group shifted towards the RP speaker mainly in the F2 dimension, while HK4 moved in the F1 dimension converging to the RP speaker. For the GenAmE group, HK7 did not change much, while HK5 and HK6 shifted towards the GenAmE speaker mainly in the F2 dimension. Note that changes in the GenAmE group were relatively small.

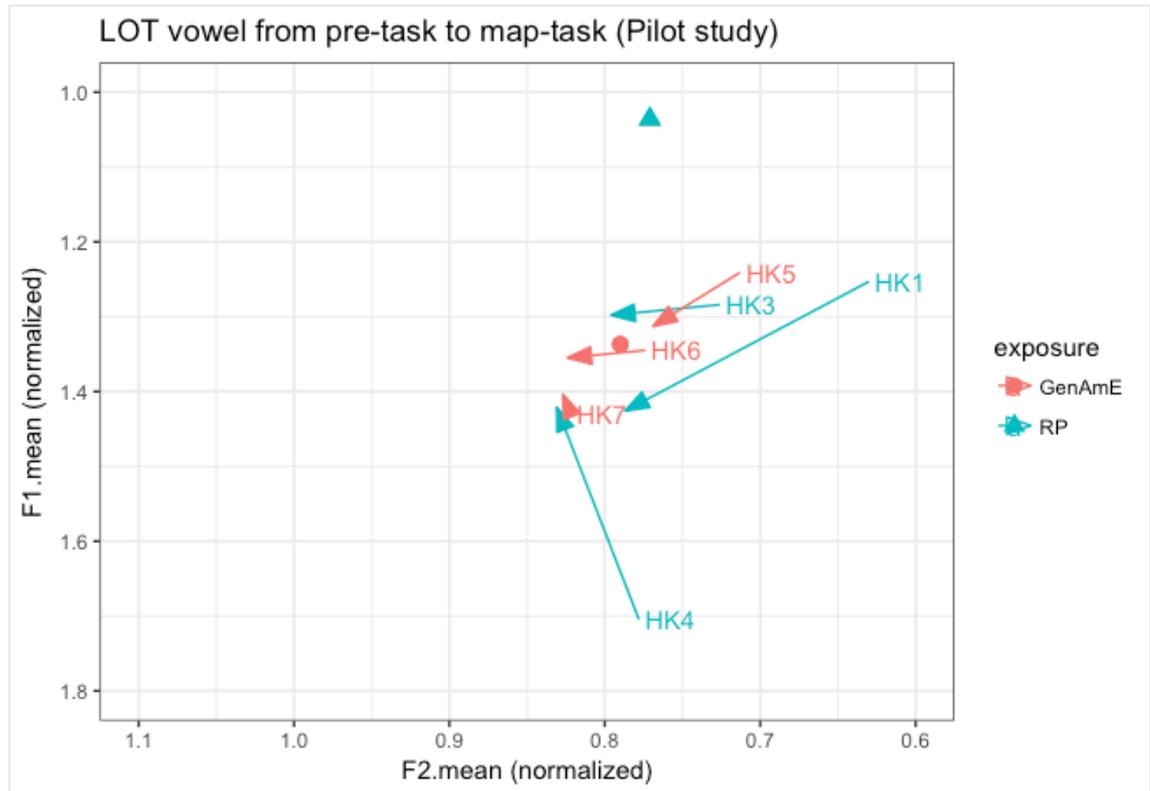


Figure 5.7 Participants’ LOT vowel movements from the pre-task to the map task based on the transformed means of F1 and F2. Each number represents a participant. Participants who were exposed to the RP interlocutor are in blue; participants who were exposed to the GenAmE speaker are in red. The starting point of an arrow represents the vowel in the pre-task (mean of F1 and F2), while the end point represents the vowel in the map task (mean of F1 and F2). The RP speaker and the GenAmE speaker are also marked for reference. It should be noted that the unit interval of the x-axis is 0.1 while the unit interval of the y-axis is 0.2.

LOT vowel (in transformed value)		pre	map	post	NS	pre→map change	pre→post change
RP group	F1	1.35 (0.23)	1.39 (0.16)	1.40 (0.16)	1.036 (0.076)	increase +0.04	increase +0.05
	F2*	0.70 (0.08)	0.81 (0.10)	0.72 (0.09)	0.771 (0.046)	increase +0.11	increase +0.02
GenAmE group	F1	1.31 (0.11)	1.34 (0.16)	1.32 (0.15)	1.337 (0.046)	increase +0.03	increase +0.01
	F2*	0.76 (0.08)	0.80 (0.11)	0.75 (0.08)	0.79 (0.04)	increase +0.04	decrease -0.01

Table 5.3 Overall mean and standard deviation (in brackets) of F1 and F2 across tasks and conditions for LOT vowels. NS represents overall means of the native interlocutors. The figures in the columns of “pre→map change” and “pre→post change” refer to the change from the pre-task to the map task/post-task. * indicates a significant effect based on one-way ANOVA.

Table 5.3 shows the means of the RP and GenAmE groups' F1 and F2 in the pre-task, the map task and the post-task, along with the native interlocutors' F1 and F2. A series of one-way ANOVA were used to examine the significant effect of **task** (pre vs map vs post). Dependent variables were F1 and F2 values of the two groups. Results of ANOVA suggested that a significant main effect was found for **F2 of both groups** (RP: $F = 3.889$, $DF = 2$, $p = .022$; GenAmE: $F = 3.965$, $DF = 2$, $p = .021$).

For the RP group, post-hoc tests of F2 indicated that the differences between **pre-task and map task** (pre-map = -0.11, $p < .001$) and **post-task and map task** (post-map = -0.01, $p < .001$) were significant. For the GenAmE group, post-hoc tests of F2 suggested a significant comparison between **post-task and map task** (post-map = -.05, $p = .04$).

These results suggested that the HKE participants in the RP group significantly converged towards the RP speaker in the F2 dimension from the pre-task to the map task. The main effect of **task** was found for the F2 of GenAmE group too, however, the difference between pre-task and map task was not significant.

Euclidean distance is used to calculate the distance between two points as the formula below indicates. Using Euclidean distance would be able to capture the changes in both F1 and F2 dimension. The present study calculated Euclidean distance between a participant's vowel and his/her interlocutor's across the three tasks.

$$\text{Pre-distance} = \sqrt{(\text{HKpre_F1} - \text{NSmean_F1})^2 + (\text{HKpre_F2} - \text{NSmean_F2})^2}$$

$$\text{Map-distance} = \sqrt{(\text{HKmap_F1} - \text{NSmean_F1})^2 + (\text{HKmap_F2} - \text{NSmean_F2})^2}$$

$$\text{Post-distance} = \sqrt{(\text{HKpost_F1} - \text{NSmean_F1})^2 + (\text{HKpost_F2} - \text{NSmean_F2})^2}$$

For all the tokens a participant produced in the pre-task, the distance between these tokens and the interlocutor's mean was calculated and named the pre-distance. Similarly, the distance between the tokens of the map task/post-task and the interlocutor's mean was calculated and named the map-distance or post-distance. For example, for HK1, the F1 and F2 values of each token of his vowels would be used in the above formulas; the mean F1 and F2 of the RP/GenAmE interlocutors' vowels would be used in the formulas too.

Two separate one-way ANOVA were run using **task** as the factor and the Euclidean distance of the RP group and the GenAmE group as dependent variables.

LOT vowel (in transformed value)		pre	map	post	pre→map change	pre→post change
RP group	distance	0.35 (0.21)	0.37 (0.15)	0.38 (0.15)	larger +0.02	larger +0.03
GenAmE group	distance	0.11 (0.08)	0.16 (0.11)	0.13 (0.11)	larger +0.05	larger +0.02

Table 5.4 Euclidean distance of the LOT vowel across the three tasks and the two groups. The figures in the columns of “pre→map change” and “pre→post change” refer to the change from the pre-distance to the map-distance/post-distance. No significant effect was found based on one-way ANOVA.

If the map-distance/post-distance is larger than the pre-distance, it suggests a divergence; if the map-distance/post-distance is smaller than the pre-distance, it suggests a convergence. Table 5.4 and Figure 5.8 show the means of Euclidean distance of LOT vowels across three tasks for both groups. Both the RP and GenAmE group shifted away from the native speakers in terms of Euclidean distance, as the map-distances and the post-distances were larger than the pre-distances for both groups.

The results of ANOVA indicated that no significant effect of **task** was found for the Euclidean distances of either group, suggesting that no convergence on interlocutor mean was observed for both groups.

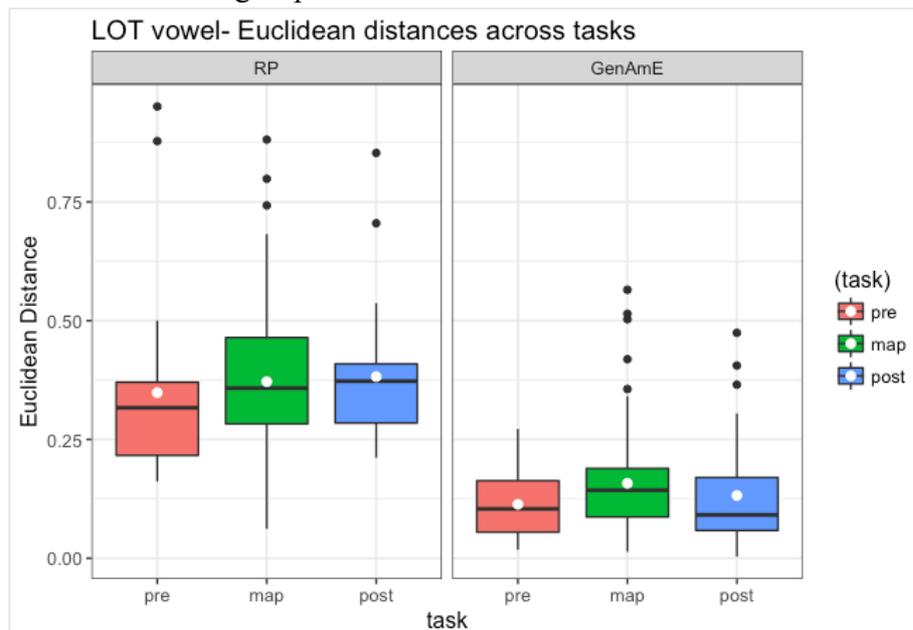


Figure 5.8 Euclidean distance of the LOT vowel across the tasks for the RP and GenAmE group. White dots in the boxplots represent the means of the distances. Black dots are outliers.

5.2.2.2. PATH vowel

The plotting of the PATH vowels of the participants and the two native speakers in Figure 5.9 is similar to Figure 5.3. The RP speaker's PATH vowel is slightly more back than that of the HKE participants of the RP group, and the GenAmE speaker's PATH vowel is lower and more front than that of the HKE participants of the GenAmE group. As shown in Figure 5.9, for the RP group, HK1 and HK4 shifted towards the RP speaker while HK3 moved away. The GenAmE group shows a more consistent pattern: they all shifted towards the GenAmE speaker.

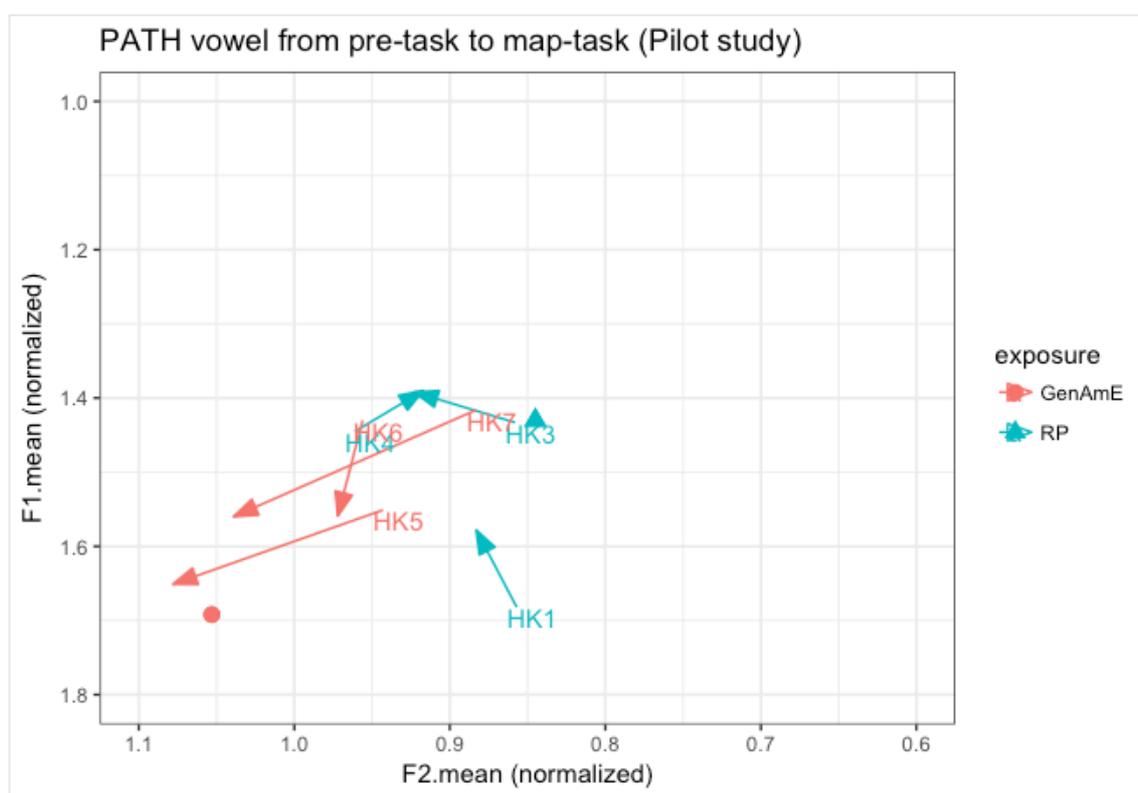


Figure 5.9 Participants' PATH vowel movements from the pre-task to the map task based on the transformed means of F1 and F2. Each number represents a participant. Participants who were exposed to the RP interlocutor are in blue; participants who were exposed to the GenAmE speaker are in red. The starting point of an arrow represents the vowel in the pre-task (mean of F1 and F2), while the end point represents the vowel in the map task (mean of F1 and F2). The RP speaker and the GenAmE speaker are also marked for reference.

PATH vowel (in transformed value)		pre	map	post	NS	pre→map change	pre→post change
RP group	F1	1.50 (0.14)	1.45 (0.25)	1.52 (0.16)	1.43 (0.061)	decrease -0.05	increase +0.02
	F2	0.89 (0.07)	0.92 (0.14)	0.92 (0.12)	0.845 (0.047)	increase +0.03	increase +0.03
GenAmE group	F1	1.48 (0.13)	1.56 (0.28)	1.58 (0.17)	1.692 (0.089)	increase +0.08	increase +0.1
	F2*	0.93 (0.15)	1.03 (0.16)	1.03 (0.17)	1.053 (0.032)	increase +0.1	increase +0.1

Table 5.5 Overall mean and standard deviation (in brackets) of F1 and F2 across tasks and conditions for PATH vowels. NS represents overall means of the native interlocutors. The figures in the columns of “pre→map change” and “pre→post change” refer to the change from the pre-task to the map task/post-task. * indicates a significant effect based on one-way ANOVA.

Table 5.5 shows the means of the RP and GenAmE groups’ F1 and F2 in the pre-task, the map task and the post-task. A series of one-way ANOVA were used to examine the significant effect of **task**. Dependent variables were F1 and F2 values of the two groups.

The results of ANOVA suggested that a significant main effect was found for the **F2 of the GenAmE group** ($F = 4.223$, $DF = 2$, $p = .017$). Post-hoc tests suggested a significant difference for the F2 of the GenAmE group between **the pre-task and the map task** (pre-map = -0.1, $p = .026$) and **the pre-task and the post-task** (pre-post = -0.1, $p = .04$). This indicates that the HKE participants in the GenAmE group significantly converged towards the GenAmE speaker in the dimension of F2.

Similar to the LOT vowel, the Euclidean distances between the HKE participants and the native interlocutors in the pre-task, the map tasks and the post-task were calculated. Table 5.6 and Figure 5.10 show the means of Euclidean distance across three tasks in the two groups. If the map-distance/post-distance is larger than the pre-distance, it suggests a divergence; if the map-distance/post-distance is smaller than the pre-distance, it suggests a convergence.

The results suggested that the RP group shifted away from the RP speaker as the map-distance and post-distance were larger than the pre-distance. On the other hand, the GenAmE group shifted towards the GenAmE speaker as the map-distance and the post-distance were smaller than the pre-distance.

PATH vowel (in transformed value)		pre	map	post	pre→map change	pre→post change
RP group	distance*	0.15 (0.09)	0.24 (0.18)	0.20 (0.10)	larger +0.09	larger +0.05
GenAmE group	distance	0.29 (0.14)	0.26 (0.23)	0.24 (0.10)	smaller -0.03	smaller -0.05

Table 5.6 Euclidean distance of the PATH vowel across the three tasks and the two groups. The figures in the columns of “pre→map change” and “pre→post change” refer to the change from the pre-distance to the map-distance/post-distance. * indicates a significant effect based on one-way ANOVA.

Two one-way ANOVA using **task** as the factor were run for the Euclidean distances of the two groups. A significant main effect of task was found for **the RP group** ($F = 3.577$, $DF = 2$, $p = .031$). Post-hoc tests suggested a significant difference between the **pre-task and the map task** (pre-map = $-.09$, $p = .025$). This indicates that though a trend of convergence was found for the GenAmE group, the changes across tasks were not significant; on the other hand, the divergence in the RP group was statistically significant.

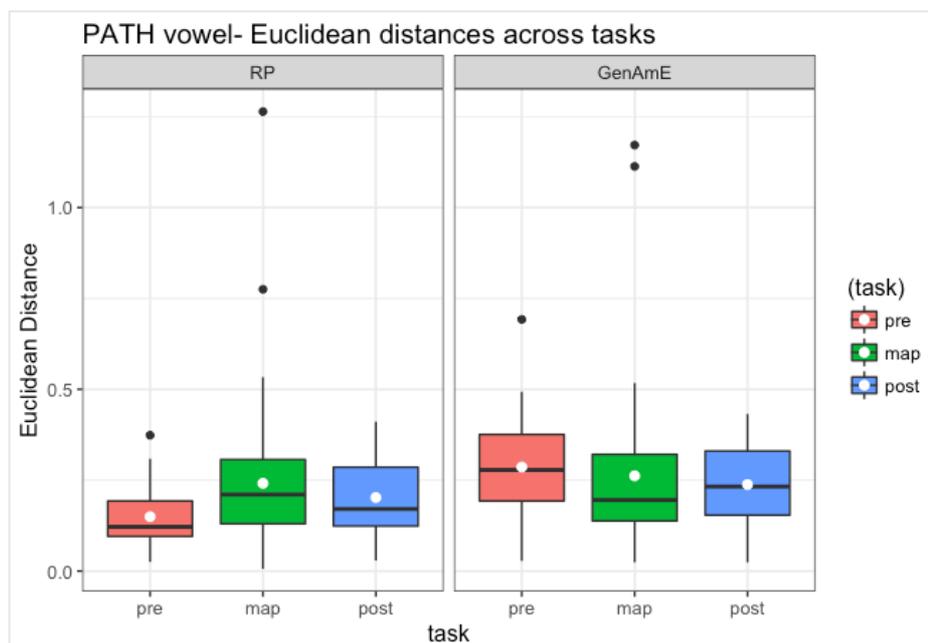


Figure 5.10 Euclidean distance of the PATH vowel across the tasks for the RP and GenAmE group. White dots in the boxplots represent the means of the distances. Black dots are outliers.

5.2.3. Summary

This section presented the design and results of the pilot study. Two groups of HKE speakers were recruited for the study. The participants completed three tasks in the experiment and talked to a native speaker in the map task. Their production of the LOT and PATH vowels were examined and compared with their native interlocutors' productions.

Results were reported as the mean of the RP and GenAmE groups' vowel formants and the Euclidean distance between the participants and interlocutors in the three tasks. A few key findings are summarised below:

- (1) For the LOT vowel, the HKE speakers in the RP group significantly converged their **F2** towards the RP speaker from the pre-task to the map task; the similar effect was not found for the HKE speakers in the GenAmE group.
- (2) For the LOT vowel, the results of Euclidean distance indicated a trend of divergence for the HKE speakers in both groups, however, no significant effect was found.
- (3) For the PATH vowel, the HKE speakers in the GenAmE group significantly converged their **F2** towards the GenAmE speaker from the pre-task to the map task, and from the pre-task to the post-task; no significant effect was found for the HKE speakers in the RP group.
- (4) For the PATH vowel, the results of Euclidean distance indicated a trend of convergence for the GenAmE group but the changes were not significant; the RP group was found significantly diverged from the native speaker in terms of Euclidean distance.

The reason for having a different story in the calculation of Euclidean distance may be imbalances in the participants' movements of F1 and F2. As shown in Figures 5.7 and 5.9, many participants only converged in one dimension but diverged in the other dimension. For instance, for the LOT vowel, HK1 only converged in F2 but diverged in F1; HK4 converged in F1 but diverged in F2. For these participants, the distance after exposure may be even larger than before exposure depending on the divergence dimension.

Another reason may be over-convergence. Over-convergence means that participants move towards the interlocutor but then move further from the interlocutor in the same direction. An example for this is HK6 as shown in Figure 5.7. HK6's LOT vowel started from further back than the GenAmE interlocutor's in the F2 dimension. When it moved towards the GenAmE interlocutor and became more front, it ended up being further away from the GenAmE interlocutor than it was at the beginning. In this case, do we count it as convergence or divergence? Based on the calculation of Euclidean distance, HK6 would be counted as diverging in the map task.

Vowel movement is different from driving a car where the driver can stop right by the parking sign. For L2 learners, it is common that they hypercorrect their pronunciation in order to achieve a native-like accent. In the case of the pilot study, some HKE speakers might try to converge towards the native speakers, however, due to hypercorrection, their pronunciation in the map tasks might be regarded as divergence compared to the pre-task based on the calculation of Euclidean distance.

In sum, the pilot study suggested that the HKE speakers changed some of their pronunciation of the LOT vowel and the PATH vowel during and after they talked to a native speaker. Though the effect of convergence was not very salient according to the results of Euclidean distance, the pilot study was still successful because it tested the experimental design and some changes on F2 were found. The next section will report several improvements that were made to Study 2.

5.3. Improvements

5.3.1. Variables

In the pilot study, the LOT vowel and PATH vowel were selected as variables. The results suggested that a trend of divergence was found for the LOT vowel in both groups. To elicit a stronger accommodation effect, the THOUGHT vowel was chosen to replace the LOT vowel in Study 2.

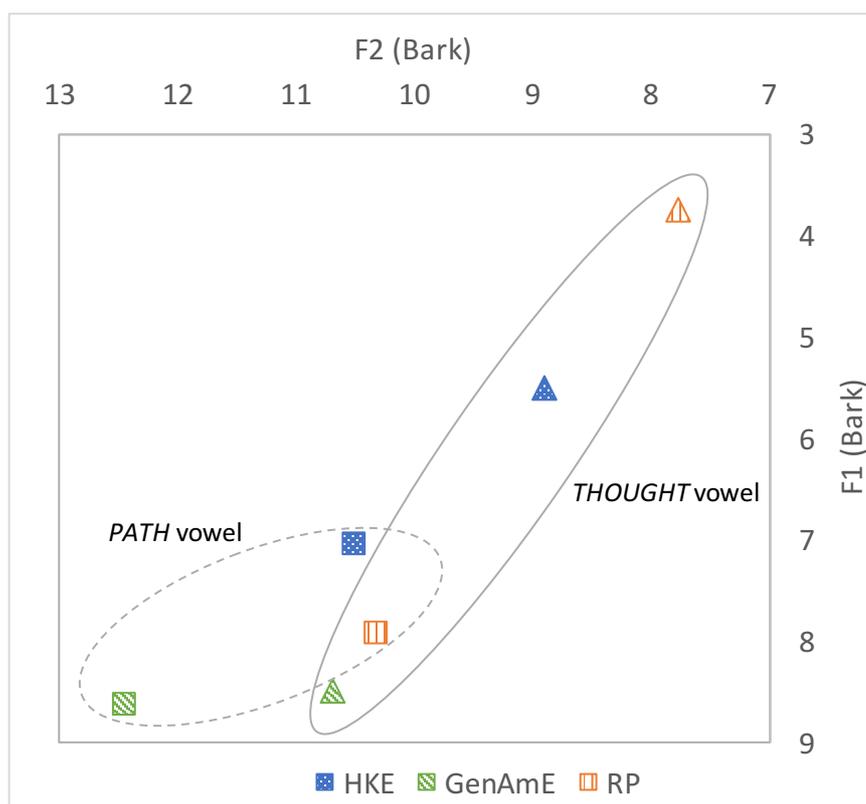


Figure 5.11 The THOUGHT vowel and PATH vowel of GenAmE (data from Hagiwara 1997), RP (data from Deterding, 1997) and HKE (data from Deterding et al., 2008).

If we plot the THOUGHT vowel and PATH vowel of RP, HKE and GenAmE in the same way as shown in Figure 5.11, the RP-*THOUGHT* vowel is a lot higher than the HKE-*THOUGHT* vowel, which may elicit more changes in the HKE participants. The PATH vowel is retained in Study 2.

Besides vowels, HKE also has some consonantal features which are distinct from RP and GenAmE; examples include the voiceless interdental fricative /θ/ (as in *thirty*) and the voiced fricative /z/ (as in *zoo*).

Fricative /θ/

Deterding et al. (2008) and Hung (2000) both mention that the voiceless fricative /θ/ is pronounced [f] in HKE. Deterding et al. (2008) found that the voiceless /θ/ occurred 42 times in word-initial position: 27 tokens (64.3%) were pronounced [θ], 14 tokens (33.3%) were pronounced [f] and 1 token (2.3%) was pronounced [t].

The fricative /θ/ was therefore added as a consonantal variable to Study 2 to observe whether the participants would produce more [θ] sounds after exposure to native accents.

Fricative /z/

Hung (2000) suggests that in HKE /s/ was the only alveolar fricative found in his data. No example of [z] was found in any position. For example, *seal* was pronounced as [sil], *razing* as [ɹeɪsɪŋ].

The voiced fricative /z/ was selected to see if the participants produce more voiced fricative /z/ after speaking to native speakers.

/ɹ/ (rhoticity)

Rhoticity is the last variable added to Study 2. Though HKE is closer to RP, some HKE speakers carry an American accent in their English. Rhoticity is chosen also because this is one of the most distinctive features between RP and GenAmE. Deterding et al. (2008) also found out that 6 out of 15 female HKE speakers in their study had clear American influences in their speech. It will be interesting to see whether exposing participants to RP and GenAmE would affect HKE speakers' production of rhotic words.

In summary, two vocalic variables are selected for Study 2: the THOUGHT vowel and the PATH vowel. Three consonantal features are selected: rhoticity, fricative /z/ and fricative /θ/.

5.3.2. Experimental design

In the pilot study, the participants were divided into two groups. One group talked to an RP speaker and the other group talked to a GenAmE speaker. This kind of experiment design would not allow us to directly compare the RP group with the GenAmE group as

different participants were used in each group. Therefore, the experiment design will change to a within-subject design, where the same group of HKE speakers are exposed to RP and GenAmE speakers in two separate experiments.

With the new design, we are able to examine which accent elicits stronger accommodation from the participants.

5.3.3. Normalization

Though the pilot study only had one male participant, Study 2 intends to collect data from both male and female HKE speakers. Therefore, normalization across speakers is essential to compare between speakers. Previous studies of accommodation show that Lobanov (1971) and Labov ANAE (Labov, Ash & Boberg, 2006) are the most common normalization methods. Lobanov (1971) is used in Babel (2010, 2012), while Labov ANAE (Labov et al. 2006) is widely used in Pardo's studies (Pardo et al., 2012; Pardo et al., 2013; Pardo et al., 2017). The following section reviews these two normalization methods, as well as Watt and Fabricius (2002).

Vowel normalization is usually classified as either a vowel-intrinsic method or a vowel-extrinsic method. According to Adank, Smits and van Hout (2004), a vowel-intrinsic method refers to a procedure that uses only acoustic information contained within a single vowel to normalize that vowel token, for example, Bark Difference Metric (Syrdal and Gopal, 1986). A vowel-extrinsic method refers to a procedure that requires acoustic information from more than one vowel of a talker to normalize other vowels, for example, Labov ANAE (Labov et al. 2006) and Lobanov (1971).

Watt and Fabricius (2002) is a vowel-extrinsic method that derives a centroid point from three corner points: the frontest vowel [i], the lowest vowel [a] and the backmost vowel [u]. An *S*-value was calculated from these three corner vowels, which is used to normalized all the vowels. The pilot study used the Watt and Fabricius (2002) method of normalization. However, as Watt and Fabricius (2002) uses the lowest vowel [a] to calculate the *S*-value, the PATH vowel in the pilot study happened to be the lowest vowel for many speakers. Therefore, to avoid using the target vowel in the normalization, Watt and Fabricius (2002) was not adopted in Study 2.

Lobanov is a vowel-extrinsic and a speaker-intrinsic method. It calculates means and standard deviation for all vowels, and uses them to normalize other vowels. To use Lobanov, one would extract formant values from all the vowels for each speaker.

Labov ANAE is a vowel-extrinsic and speaker-extrinsic method. Firstly, a log mean of all vowels for each speaker (S) is calculated. A grand log mean (G) is then extracted based on all the speaker's individual log means. After calculating G and S , a scaling factor F is taken from the anti-log of the difference ($G-S$). Each speaker's raw values are multiplied by the scaling factor F .

To decide which normalization method to use in Study 2, a small comparison study was conducted to compare Lobanov (1971) or Labov ANAE (Labov et al. 2006).

Eight HKE speakers including five male speakers and three female speakers from the participants of Study 2 were randomly selected for the comparison study. They first completed an experiment with a RP speaker, after 3-4 weeks, they repeated the experiment again with a GenAmE speaker (see Chapter 6). F1 and F2 formants of the THOUGHT vowel and the PATH vowel were extracted from all of their recordings. A set of vowels / i ε æ ɒ ʌ u/ were also extracted for normalization.

With the same vowel data, the methods Lobanov (1971) and Labov ANAE (Labov et al. 2006) were tested by using them to normalize the vowels. Figure 5.12 shows the results for the participants' Euclidean distances in the two different normalization methods. The results show a similar outcome, where a small increase in distances from the pre-task to the map task was found for the PATH vowel in both RP and GenAmE conditions. For the THOUGHT vowel, a small decrease from the pre-task to the map task was found in the RP condition.

The two methods did not differ much in terms of representing the results of the current study, Labov ANAE (Labov et al. 2006) was selected for two reasons. First, Labov ANAE (Labov et al. 2006) retains hertz for the data format, which gives a more straightforward view of the vowel spaces. Second, the experimental design of the present study is close to Pardo's studies, in which Labov ANAE (Labov et al. 2006) was used.

Using Labov ANAE (Labov et al. 2006) will allow comparisons among similar studies in the future.

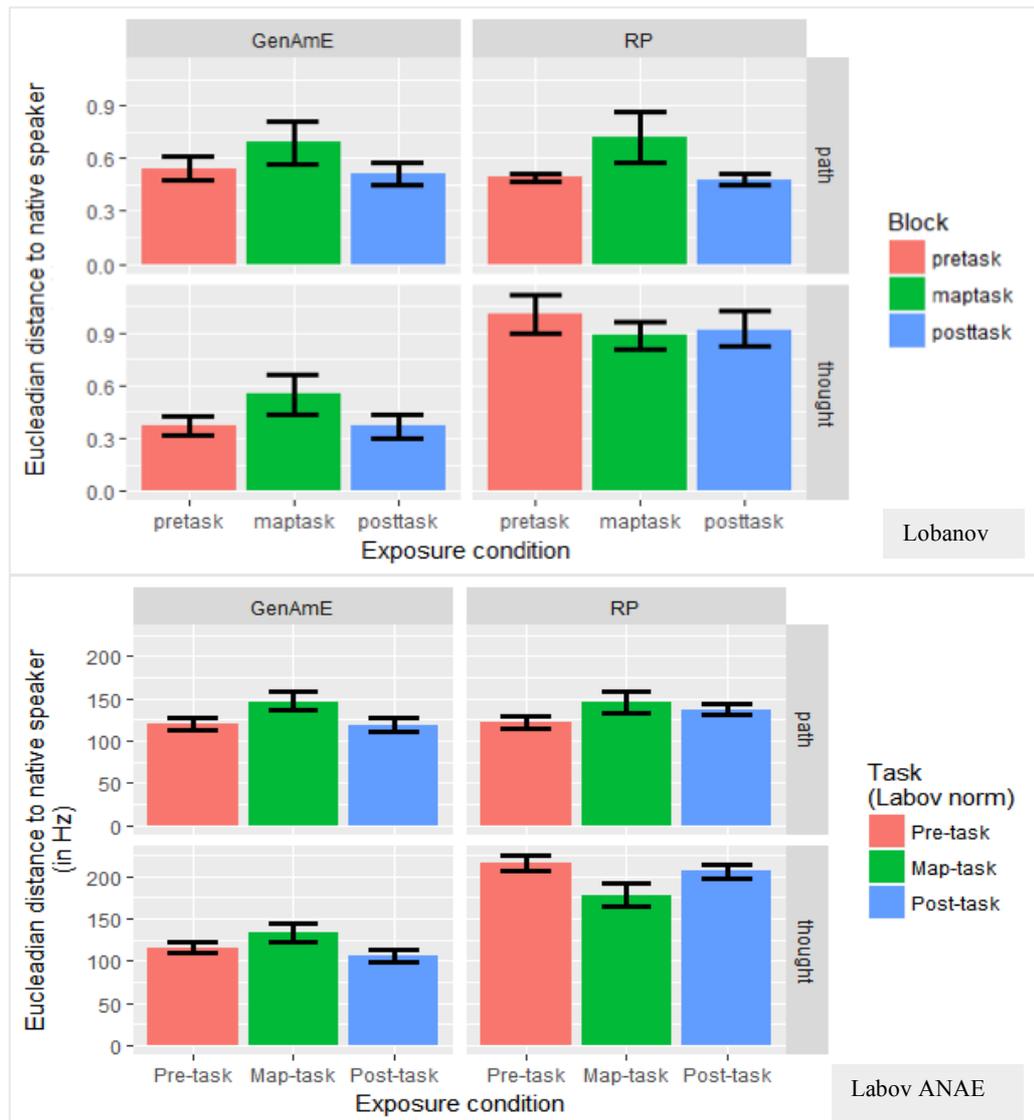


Figure 5.12 Euclidean distances of eight HKE speakers across tasks and conditions. On the top was the results normalized in Lobanov (1971), at the bottom was the results normalized in Labov ANAE (Labov et al. 2006).

5.4. Conclusion

This chapter reviewed a few studies of accommodation and the pilot study for the speech accommodation study. The results of the pilot study partially support our hypothesis that HKE speakers will be able to accommodate to native speakers, though the accommodation effects varied considerably on the dimensions of F1, F2 and Euclidean distance from interlocutor mean. A few improvements were made to the selection of

variables and normalization method, but not many changes were made to the experimental design and analysis.

Chapter 6. Study 2: Speech accommodation

This chapter presents the formal study of speech accommodation. The rationales and aims here are the same as those described in chapter 5.

The present study was revised from the pilot study described in chapter 5. Some changes were made (see Table 6.1 for details). First of all, the experimental design changed to a within-subject design: the participants were not only exposed to the RP accent, but also the GenAmE accent in a separate study. Second, four consonantal variables were added. Third, instead of the LOT vowel, the THOUGHT vowel was used in Study 2.

	pilot study	formal experiment
Participants	<ul style="list-style-type: none"> • six HK participants; • one native speaker from each accent was recruited. 	<ul style="list-style-type: none"> • nineteen HK participants; • two native speakers from each accent were recruited.
Experiment design	<ul style="list-style-type: none"> • each participant was exposed to one native accent, either GenAmE or RP; • maps for the pre-task/post-task were different from the maps for the map task; • two sets of maps were used 	<ul style="list-style-type: none"> • each participant was exposed to both accents in two separate experiments, with a 3-4 week interval in-between; • maps for the pre-task/post-task and the map task were the same; • three sets of maps were used
Variables	<ul style="list-style-type: none"> • two vowel variables: LOT and PATH; • each variable was represented by three words; e.g. LOT was represented by <i>dog</i>, <i>spot</i> and <i>shop</i> as labels in the map task; • each variable contained around 20 labels in the two maps in total. 	<ul style="list-style-type: none"> • two vowel variables: THOUGHT and PATH; • four consonant variables were added: rhoticity, fricative [z] and fricative [θ]; • same as the pilot study; e.g. fricative [z] was represented by <i>zoo</i>, <i>zone</i> and <i>zero</i> as labels in the map task; • each variable contained around 50 labels in the three maps in total.

Table 6.1 Comparisons between the pilot study and the present study.

6.1. Research questions and hypotheses

In this chapter, we mainly focus on the following research question:

Do Hong Kong English speakers accommodate to native accents after a short-term exposure to the native accents?

Both CAT and the interactive-alignment theory predict accommodation in a conversation, therefore, the main hypothesis of the present study is: **Hong Kong English speakers accommodate to native accents when they talk to native speakers.**

Accommodation may also occur from native English speakers to non-native English speakers, however, as the present project mainly focuses on non-native English speakers, the accommodation of native English speakers is not considered in this project.

Figure 6.1 (replicated from Figure 5.4.1) shows the hypothesized THOUGHT and PATH vowels of RP, GenAmE and HKE.

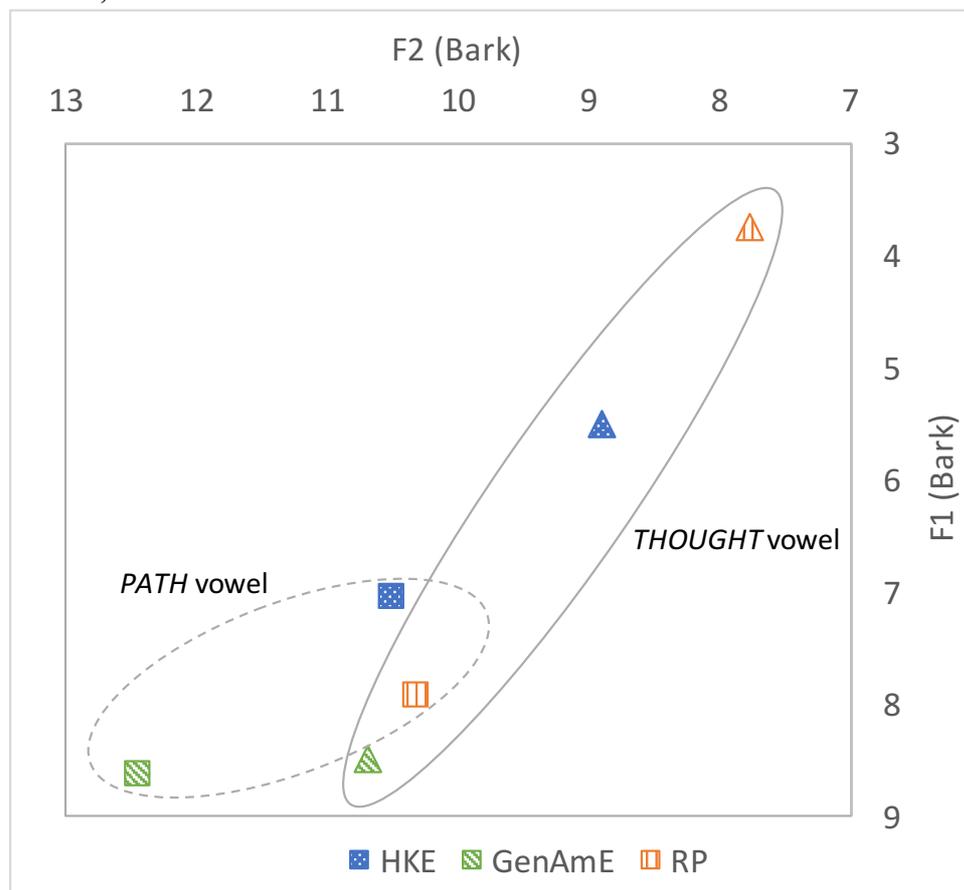


Figure 6.1 The THOUGHT and PATH vowels of GenAmE (data from Hagiwara, 1997), RP (data from Deterding, 1997) and HKE (data from Deterding et al., 2001).

As the locations of the THOUGHT vowels and PATH vowels are hypothesized, the hypotheses made below are very general. The actual locations of the vowels may be

different from the hypothesized locations, in that case, the hypotheses would be adjusted to a more defined detail.

H1 and H2 are predicted for the vowels:

H1. If the participants accommodate on the THOUGHT vowel, they are expected to shift their F1 and F2 towards the RP-*thought*/GenAmE-*thought* vowels when they talk to the RP/GenAmE speakers. Euclidean distance between the participants and the native speakers in the map tasks is expected to be shorter than the distance in the pre-tasks.

H2. If the participants accommodate on the PATH vowel, they are expected to shift their F1 and F2 towards the RP-*path*/GenAmE-*path* vowels when they talk to the RP/GenAmE speakers. Euclidean distance between the participants and the native speakers in the map tasks is expected to be shorter than the distance in the pre-tasks.

H3 and H4 are predicted for the consonantal variables:

H3. If the participants accommodate on rhoticity, they are expected to produce fewer rhotic words when they are exposed to the RP accent and they are expected to produce more rhotic words when they are exposed to the GenAmE accent.

H4. If the participants accommodate on the fricative [z], and fricative [θ], they are expected to produce more native-like items for these features.

A summary of the hypotheses is shown below.

<i>If the main hypothesis is true, the following is expected from the pre-tasks to the map tasks:</i>			
	Variables	RP condition	GenAmE condition
vowels	THOUGHT	F1/F2 shift towards NS Euclidean distance↓	F1/F2 shift towards NS Euclidean distance↓
	PATH	F1/F2 shift towards NS Euclidean distance↓	F1/F2 shift towards NS Euclidean distance↓
consonants	Rhoticity	rhoticity% ↓	rhoticity% ↑
	Fricative [z]	z% ↑	z% ↑
	Fricative [θ]	θ% ↑	θ%↑

Table 6.2 Summary of hypotheses based on variables. ↑ represents an increase of the value; ↓ represents a decrease of the value. NS stands for native speakers.

The same predictions as above are drawn from the pre-tasks to the post-tasks; however, larger differences are expected to be found between the pre-tasks and the map tasks

because the participants' accommodation may not be sustained all the way to the post-tasks.

6.2. Experiment Design

6.2.1. Participants

Twenty-three Hong Kong participants were recruited from the University of York. Four of the participants did not complete the experiments, therefore only nineteen participants' data was included in the analysis. Profiles of the participants' language proficiency and experience are shown in Table 6.3.

	Gender	Starting age of learning English (years old)	Age at arrival (years old)	Length of time since arrival	English proficiency overall (Listening/ Speaking)
HK1	Male	6	21	3 years	IELTS 6
HK2	Male	2	19	6 years	IELTS 6.5 (7/7)
HK3	Female	3	17	8 years	IELTS 7.5
HK4	Male	3	21	1.5 years	IELTS 7.5
HK5	Male	4	20	0.5 year	IELTS 7 (7.5/6)
HK6	Female	2	19	0.42 year	n.a
HK7	Female	3	17	3 years	IELTS 6.5 (6.5/6.5)
HK8	Male	3	21	3.17 years	HKDSE 5* (5*/5*)
HK9	Male	3	21	1.5 years	IELTS 8 (8.5/8)
HK11	Female	n.a	n.a	n.a	n.a
HK12	Male	3	17	9 years	IELTS 8 (9/7)
HK13	Female	1	18	4 years	IELTS 8 (9/7)
HK16	Female	3	20	0.5 year	IELTS 7.5 (8/6.5)
HK18	Female	3	28	0.5 year	IELTS 7.5 (8/6.5)
HK19	Female	4	17	3 years	IELTS 7 (8/6.5)
HK20	Female	4	20	0.42 year	IELTS 7
HK21	Female	3	17	4.5 years	IELTS 7
HK22	Female	3	22	1 year	IELTS 6.5 (7/6.5)
HK23	Female	3	20	0.33 year	IELTS 7.5

Table 6.3 Language profile of all the participants. "Starting age of learning English" refers to the initial age when the participants started to learn English; "Age at arrival" means their age when they first arrived in an English-speaking country; "Length of time since arrival" means how long they have lived in an English-speaking country. Note that HK11's data is missing owing to the absence of her background questionnaire, and HK6 refused to provide her English proficiency information. Not all the participants provided the subset scores of their IELTS Listening and Speaking. "HKDSE" is short for Hong Kong Diploma of Secondary Education: according to a study conducted by the Hong Kong Examinations and Assessment Authority, HK8's English proficiency is equal to IELTS 7.17-7.32 (HKEAA, 2012).

All of the HK participants except one were born and raised in Hong Kong, speaking Cantonese as their first language. One male participant was born in Canada but he moved to Hong Kong at the age of two and had grown up in Hong Kong since then. As he spoke native Cantonese and his English still retained some features of Cantonese, his data was included in the analysis. There were seven male and twelve female participants.

Four female native English speakers were recruited as interlocutors in the experiments. Two of them were native speakers of RP, while the other two were native speakers of GenAmE. The two RP speakers were both from southern England, speaking standard southern British English. One GenAmE speaker was from California, and the other one was from Massachusetts. All of the native speakers were Master's students in the Department of Linguistics when they participated in the experiments. They were informed about the experiment design and research aims and were trained beforehand.

6.2.2. Experiment design

The experiment consisted of three parts: a pre-task, a map task and a post-task, as shown in Figure 6.2. Each participant completed the experiment with an RP speaker, and after 3-4 weeks the participant repeated the experiment with a GenAmE speaker. The order of which native accent they were exposed to was counterbalanced: that is, nine participants talked to an RP speaker first and the rest talked to a GenAmE speaker first. The four native speakers (i.e. RP1, RP2, GenAmE1, GenAmE2) shared 38 map tasks (19 participants \times 2 map tasks) in total. RP1 completed 11 map tasks and RP2 completed 8 map tasks; GenAmE1 completed 7 map tasks and GenAmE2 completed 12 map tasks.

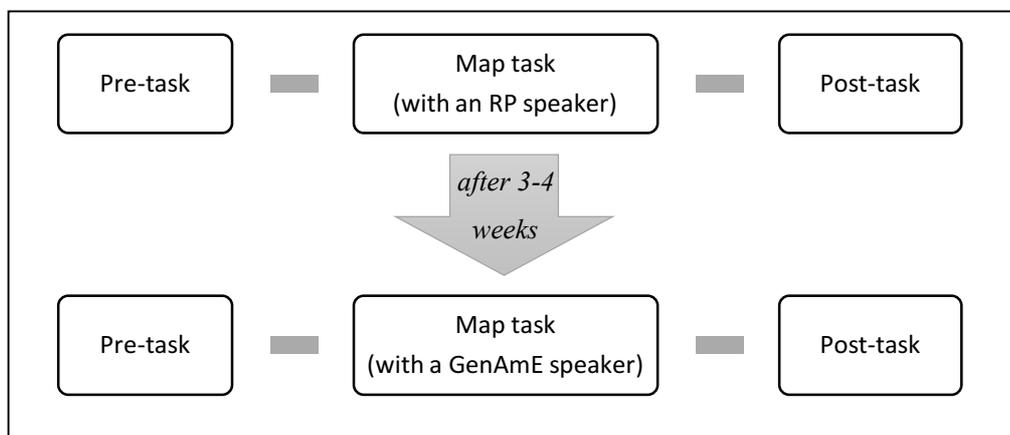


Figure 6.2 Experiment design of the formal study.

The native interlocutors were assigned one map task a day, except for two cases where GenAmE1 completed a map task at 2pm and another one at 4pm on the same day, and GenAmE2 completed a map task at 11am and another one at 4pm separately on a different day. As they did not complete the map tasks consecutively, their production should not be affected by the HKE participants and the repetition of the map tasks.

Apart from the within-subject design, the other parts of the experiments remained the same as for the pilot study. Figure 6.3 shows the environment of the recording booth where the experiments were conducted.



Figure 6.3 The recording booth where the experiments were conducted.

Each of the participants was seated on one side of a table, and put on a lapel microphone with the help of a technician. They received the instructions in Cantonese from the author in the recording booth. After the instructions, the author left the recording booth and the participants started the pre-task. When the participant completed the pre-task, the native interlocutor (either an RP interlocutor or a GenAmE interlocutor) came into the recording studio to complete the map tasks together with the participant. The native interlocutor was seated on the other side of the table, and put on another lapel microphone. A paper board was placed in the centre of the table to prevent the interlocutor and the participant from seeing one another. The whole process was monitored and recorded by a technician

and the author, who were sitting in another room next to the recording booth. After the map tasks, the native interlocutor left the recording booth, the participant then completed the post-task alone. The procedure of the experiment was explained to the participants before the pre-task in Cantonese. They were not told the purpose of the experiment.

6.2.2.1. Pre-task

Differing from the pilot study, the pre-task was a description task based on the maps used in the map task. The participants were given three maps. On the maps, there were some landmarks and a starting point. The participants were told to describe all the landmarks and items they could see on the maps. Their descriptions were recorded. The pre-task usually took 15 minutes to complete.

6.2.2.2. Map-task

The map task was similar to the pilot study. The participants kept the maps they received in the pre-task. The native interlocutor, on the other hand, used a different version of the maps. The maps the participants used only contained a starting point and some landmarks, whereas the maps for the native interlocutors had all the landmarks, a starting point, an end point and a route.

The participants were told to draw a route on the maps with the help of the native interlocutor. The native interlocutor gave instructions and directed the participants to the end point. The participants were also told to correct any wrong landmarks on their maps and to fill in the missing landmarks. Both speakers were recorded. The map task usually took 30 minutes to 40 minutes to complete.

6.2.2.3. Post-task

The post-task was a description task. The participants were given the maps that were used by the native interlocutor in the map task. They were told to describe the maps as they had done in the pre-task. With a full version of the maps, they were told to describe the maps following the route. Their descriptions were recorded. This task usually took 15 minutes to complete.

6.2.3. Materials

6.2.3.1. Target features

Five features were selected as variables. They were divided into two categories: vocalic features and consonantal features. The THOUGHT vowel and PATH vowel were grouped as vocalic features, and rhoticity, fricative [z] and fricative [θ] were grouped as consonantal features.

To ensure that the speakers would produce words containing these features in their speech, three words were selected for each feature and were used as the landmarks in the map task. All the selected words were controlled by word frequency and degree of familiarity from the MRC psycholinguistic database, as word frequency and familiarity have been proven to have effects on perceptual convergence (Goldinger, 1998).

Variables		HKE	RP	GenAmE	Words as landmarks
Vocalic features	THOUGHT	[ɔ]	[ɔ]	[ɑ]	<i>thought, cause, small</i>
	PATH	[a]	[ɑ:]	[æ]	<i>pass, path, bath</i>
Consonantal features	Rhoticity	rhotic/ non-rhotic	non-rhotic	rhotic	<i>car, bar, star</i>
	Fricative [z]	[s]	[z]	[z]	<i>zone, zero, zoo</i>
	Fricative [θ]	[f]/[s]	[θ]	[θ]	<i>three, thirty, thousand</i>

Table 6.4 Variables of the formal study.

Note that the participants do not necessarily have all the Hong Kong English features that are listed in Table 6.4. Some participants produced native-like vowels but they pronounced fricative [z] with [s]; some participants had rhotic-[ɹ], which suggests an American accent, but they used the RP-*path* vowel. In section 6.3, a summary of the participants' pronunciation of the five variables is listed in Table 6.6.

6.2.3.1. Maps

Different maps were used in this study. Two sets of three maps were designed for the study; one set for the participants and another one for the native interlocutors. The two sets of maps were very similar. The maps for the native speakers had all the landmarks and a route leading from a starting point to an end point. The maps for the participants did not have a route and some of the landmarks that appeared on the native speakers' maps were missing. The information gaps between the two maps were used to provoke conversation in the map task.

Below are examples of the first map (for the other two maps see the Appendix).

The landmarks representing the target features (see Table 6.4) were randomly placed on the three maps. For each feature, the representative landmarks appeared around 50 times in total on the three maps. For example, landmarks for the voiced fricative /z/, represented by *zoo*, *zone* and *zero*, appeared 50 times on the three maps.



Figure 6.4 Map 1 for the native speakers. The starting point is a “GO” indicated by a red star at the bottom of the map; the end point is a yellow smiley face on the left-hand side of the map. The arrows represent the direction of the route.



Figure 6.5 Map 1 for the participants. This map only contains the starting point, a “GO” indicated by a red star at the bottom of the map.

When the participants had to repeat the experiment with a different native speaker 3-4 weeks later, a revised set of maps were used. The revised maps only differed in the locations of the landmarks and the route directions. The purpose of using revised maps was to avoid a learning effect from their previous experiments.

6.3. Data analysis

6.3.1. Data annotation and extraction

The whole data set contained a total of 114 tasks (19 participants × 3 tasks × 2 experiments), in over 40 hours of recordings. All the recordings were imported into Praat (version 6.0.28, Boersma and Weenink, 2017) for acoustic and auditory analysis.

Firstly, all of the target words were manually annotated in Praat (see Table 6.5 for details). For all the vowels, boundaries began at the start of periodic waveforms and ended at the end of the periodic waveform. F1 and F2 values were extracted using Praat scripts. Note that for the THOUGHT vowel and PATH vowel, not only the words used as landmarks

were annotated, but any other words containing the target vowels were also annotated. For example, for the THOUGHT vowel words like *cause* were also annotated.

<i>Vocalic variables</i>	<i>Parameters</i>	<i>Details of data analysis</i>
THOUGHT vowel	Formants	F1 and F2 values were extracted from the midpoint of the annotated vowels;
	Euclidean distance	Euclidean distances were calculated with the mean F1/F2 of the corresponding native interlocutors.
PATH vowel	Formants	F1 and F2 values were extracted from the midpoint of the annotated vowels;
	Euclidean distance	Euclidean distances were calculated with the mean F1/F2 of the corresponding native interlocutors.
<i>Consonantal variables</i>	<i>Parameters</i>	<i>Details of data analysis</i>
Rhoticity	rhoticity%	All words that contain [ɹ] sounds were annotated (words with [ɹ] in word-initial position were excluded). An auditory judgement of rhoticity was made on each word. Percentages of rhoticity were calculated: rhoticity% = tokens of rhotic words/tokens of all annotated [-ɹ] words.
Fricative [z]	z%	Words with [z] in word-initial position were annotated and transcribed. Percentages of [z] were calculated: z% = tokens of words with [z] /tokens of all the annotated [z]-words.
Fricative [θ]	θ%	Words with fricative [θ] in word-initial position were annotated and transcribed. Percentages of [θ] were calculated: θ % = tokens of words with [θ] /tokens of all the annotated [θ]-words.

Table 6.5 Details of data analysis of the variables.

Euclidean distance was calculated to demonstrate the change in distance between the participants and the native interlocutor across the tasks, using the following formulas:

$$\text{Pre-distance} = \sqrt{(\text{HKpre_F1} - \text{NSmap_F1})^2 + (\text{HKpre_F2} - \text{NSmap_F2})^2}$$

$$\text{Map-distance} = \sqrt{(\text{HKmap_F1} - \text{NSmap_F1})^2 + (\text{HKmap_F2} - \text{NSmap_F2})^2}$$

$$\text{Post-distance} = \sqrt{(\text{HKpost_F1} - \text{NSmap_F1})^2 + (\text{HKpost_F2} - \text{NSmap_F2})^2}$$

Instead of the overall means, the means of the participants' corresponding native interlocutors were used to calculate Euclidean distance. These means were matched with

each data point of the participants' vowels. For example, HK1 talked to RP1 in his first experiment and talked to GenAmE1 in his second experiment. The means of RP1 and GenAmE1's THOUGHT/PATH vowels would be used in the above formulas. F1 and F2 value for each token of HK1's THOUGHT/PATH vowels in the three tasks would be used in the formulas too.

For all the consonantal features, the words which contained the target features were annotated. After that, each annotated word was transcribed using auditory judgements, and finally the percentages of target features were calculated. For rhoticity, tokens of linking-r were included because linking-r only appeared a few times in the data set of the native speakers. For fricative [z] and [θ], each token was annotated as either "yes" or "no". When the token matched the native pronunciation, e.g. a voiced fricative [z], a "yes" was given. Otherwise, a "no" was marked. Details are shown in Table 6.5. Eight minutes of recordings from each participant (in total 154 minutes) were extracted for an agreement test. A trained phonetician from the Department of Language and Linguistic Science at the University of York listened to each annotated token of rhoticity, fricative [z] and fricative [θ] in the recordings, and gave his judgements. For rhoticity, an 89% of agreement was reached, 93% was reached for fricative [z] and 78% was reached for fricative [θ].

Though all the participants spoke HKE, some of them might have a stronger influence by GenAmE while some of them might speak a more RP-like HKE. The participants' pronunciation of the five variables might be various. A summary of their pronunciation of the five variables is shown in Table 6.6. The summary is based on the data of the participants' two pre-tasks. For the THOUGHT and PATH vowels, a broad auditory transcription was conducted by the author. When the participants used more than one variant of the vowels in their production, all the variants are listed in the table. For example, HK3 pronounced some of the PATH vowels with a more RP-like variant [ɑ] and pronounced some others with a more HKE variant [a]. A convention of [ɔ̄] is used to represent a more RP-like THOUGHT vowel which is higher than Cardinal Vowel 6 and closer to Cardinal Vowel 7.

For rhoticity, the percentage of rhotic words in the pre-tasks was calculated for each participant. For the two fricatives, the percentages of the native variants (i.e. [z] and [θ]) were calculated.

As demonstrated in Table 6.6, the participants did not show a homogenous accent on the five variables.

	THOUGHT	PATH	Rhoticity	Fricative /z/	Fricative /θ/
HK1	[ɔ]	[a]	5%	0%	0%
HK2	[ɔ]	[a]	2%	0%	29%
HK3	[ɔ]	[ɑ]/[a]	20%	1%	87%
HK4	[ɔ]	[a]	7%	5%	96%
HK5	[ɔ]/[ɔ̃]	[ɑ]/[a]	58%	3%	99%
HK6	[ɑ]/[ɔ]	[a]	80%	0%	36%
HK7	[ɔ]	[a]	30%	0%	75%
HK8	[ɑ]/[ɔ]	[æ]/[a]	91%	46%	98%
HK9	[ɔ]/[ɔ̃]	[a]	3%	98%	99%
HK11	[ɔ]	[a]	3%	0%	99%
HK12	[ɔ]	[a]	7%	48%	23%
HK13	[ɑ]/[ɔ]	[æ]/[a]	65%	1%	93%
HK16	[ɔ]	[a]	2%	17%	95%
HK18	[ɔ]	[a]	32%	52%	97%
HK19	[ɔ]	[a]	32%	0%	41%
HK20	[ɔ]	[a]	25%	0%	40%
HK21	[ɔ]	[a]	6%	0%	3%
HK22	[ɔ̃]	[a]	1%	1%	98%
HK23	[ɑ]/[ɔ]	[æ]/[a]	63%	0%	70%

Table 6.6 A summary of each participant's pronunciation of the five variables.

It is likely that a participant contains both accents of RP and GenAmE in his/her production. For example, HK5 spoke English with rhoticity but he also pronounced the vowels with an RP-like accent. HK8 had American-like vowels and rhoticity but he still retained an HKE accent in his pronunciation of fricative /z/. The lack of homogeneity is also observed between participants. For example, some participants had a rhotic accent (e.g. HK5, HK6, HK8) while some of them did not. Some participants had acquired fricative /θ/ (e.g. HK4, HK5, HK8, HK13) whereas some participants might be still in the process of acquiring it (e.g. HK2, HK12).

The diversity of the participants' accents reflects the fact that HKE is still a developing variety where a homogenous representation of the variety has not established yet. The

participants might accommodate differently on each variable and a clear-cut pattern of accommodation might not be found.

6.3.2. Vowel normalization

For all the vowels, F1 and F2 were normalized using the Labov ANAE method (Labov et al, 2004; see discussion of normalisation methods in Chapter 5.4). A set of vowels /i, ε, æ, ɒ, ʌ, u/ was extracted from each speaker for each normalization task. Along with all the THOUGHT and PATH vowels, in total 23,493 data points (14,009 target vowels and 9,484 normalization vowels) were imported into the R program, and normalized using the *vowels* package (Erik & Kendall, 2007).

6.4. Results

The results are shown in two sections: vowels and consonants. The first section shows the results for the two vowels, the THOUGHT vowel and the PATH vowel. The second section shows the results for consonants: rhoticity [-ɹ], fricative [z] and fricative [θ].

6.4.1. Results of vowels

In this section, the results for the THOUGHT vowel and the PATH vowel are presented. For each vowel, the results for F1 and F2 are presented first, followed by the results for Euclidean distance and a summary.

6.4.1.1. THOUGHT vowel

6.4.1.1.1. Results for F1 and F2

Descriptive results

H1 predicts that if the participants accommodate towards the native speakers, the participants will shift their F1 and F2 towards the native speakers, and Euclidean distance for the map tasks will be shorter than the distance for the pre-tasks.

Figure 6.6 shows the actual locations of the THOUGHT vowels for the HKE speakers, the RP speakers and the GenAmE speakers in the present study. The hypothesized locations of the THOUGHT vowels are also shown in Figure 6.6. The scales of the x and y axis for the actual locations were transferred from Barks to hertz to facilitate the comparisons of the locations.

The actual locations of the speakers' THOUGHT vowels are not the same as the hypothesized locations. This may be due to three reasons. First, different groups of HKE, RP and GenAmE speakers were used in these studies. The locations of the RP-*thought* vowel and the GenAmE-*thought* vowel in the present study were based on two RP speakers and two GenAmE speakers, whereas the hypothesized locations were based on a group of RP and GenAmE speakers. Second, as mentioned before, HKE is still a developing variety of English, and so a huge variation within HKE speakers is expected. The HKE speakers in the present study can only represent some variations of HKE. Third, the hypothesized locations are based on the Bark scale, which is closer to people's perceptual system; while the actual locations are based on acoustic measurements of people's production. Some of the differences may be due to the different systems of scale used.

Figure 6.6 suggests that the RP-*thought* vowels are higher and more front than the HKE-*thought* vowels, while the GenAmE-*thought* vowels are lower and more front than the HKE-*thought* vowels.

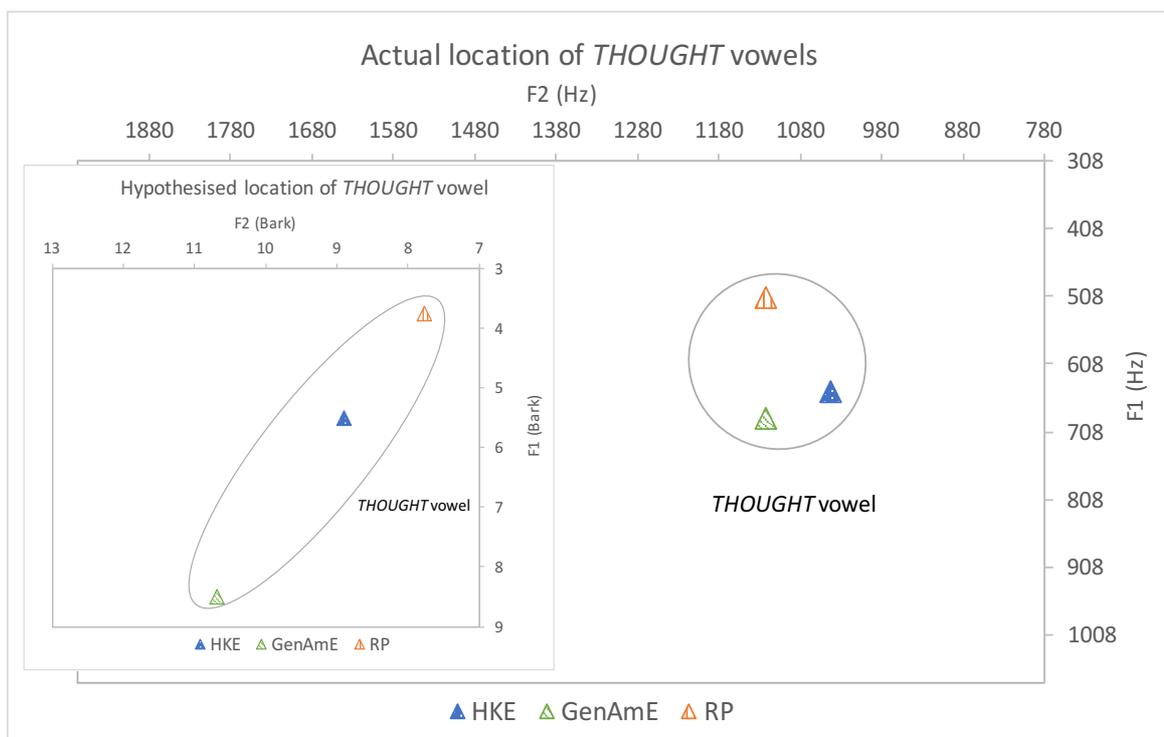


Figure 6.6 The actual locations of THOUGHT vowels for the HKE speakers, the RP speakers and the GenAmE speakers in Study 2, based on the results of Table 6.7. The small window on the left shows the hypothesized location of the THOUGHT vowel based on Hagiwara (1997), Deterding (1997) and Deterding et al. (2001).

Based on this, H1 is adjusted as follows:

If the participants accommodate towards the native speakers, their F1 is expected to shift higher (i.e. F1 value to decrease) and F2 is expected to shift to be more front (F2 value to increase) when they talk to the RP speakers. In the GenAmE condition, their F1 is expected to shift lower (or show no change as the difference is too small; F1 value to increase or no change), their F2 is expected to shift more front (F2 value to increase). Euclidean distances for the map task are expected to be shorter than the distances for the pre-tasks.

Table 6.7 shows the means of F1 and F2 for the HKE speakers in both conditions, the RP speakers and the GenAmE speakers. For the THOUGHT vowels, changes were mainly found from the pre-tasks to the map tasks. As predicted, the participants' F1 decreased and F2 increased in the RP condition, but the change in F2 was not significant. On the other hand, in the GenAmE condition, their F2 increased as expected, but their F1 also increased, which is not in line with the prediction. From the pre-tasks to the post-tasks, the changes were less obvious. The participants increased their F2 in both conditions as expected; however, they also increased their F1 which was not in line with the prediction. A series of linear mixed effects regressions were run to test the significance, which are reported in the next section.

THOUGHT vowel (in Hz)		pre	map	post	NS	pre→map change	pre→post change
RP condition	F1	649 (62)	599 (81)	653 (72)	510 (57)	decrease * -50	increase +4
	F2	1044 (116)	1053 (146)	1048 (107)	972 (147)	increase +9	increase +4
GenAmE condition	F1	648 (66)	627 (101)	650 (65)	689 (77)	decrease* -21	increase +2
	F2	1025 (103)	1103 (167)	1055 (114)	1122 (150)	increase* +78	increase* +30

Table 6.7 Overall mean and standard deviation (in brackets) of F1 and F2 across tasks and conditions for THOUGHT vowels. NS represents overall means of the native interlocutors. The figures in the columns of “pre→map change” and “pre→post change” refer to the change from the pre-distance to the map-distance/post-distance. * indicates a significant effect based on linear mixed effects regressions.

Figure 6.7 shows the changes in participants' means of F1 and F2 across the tasks. From the pre-tasks to the map tasks, in the RP condition, the participants shifted their

THOUGHT vowels higher but remained the same in the F2 dimension; in the GenAmE condition, they shifted their THOUGHT vowels slightly higher and moved more front.

The shifting can also be seen in Figure 6.8, where individuals are shown. Figure 6.8 plots the mean F1 and F2 of each participant's THOUGHT vowel and the corresponding native speakers' THOUGHT vowels across the three tasks in the same vowel space.

Figure 6.6 visualises the results of Table 6.7.

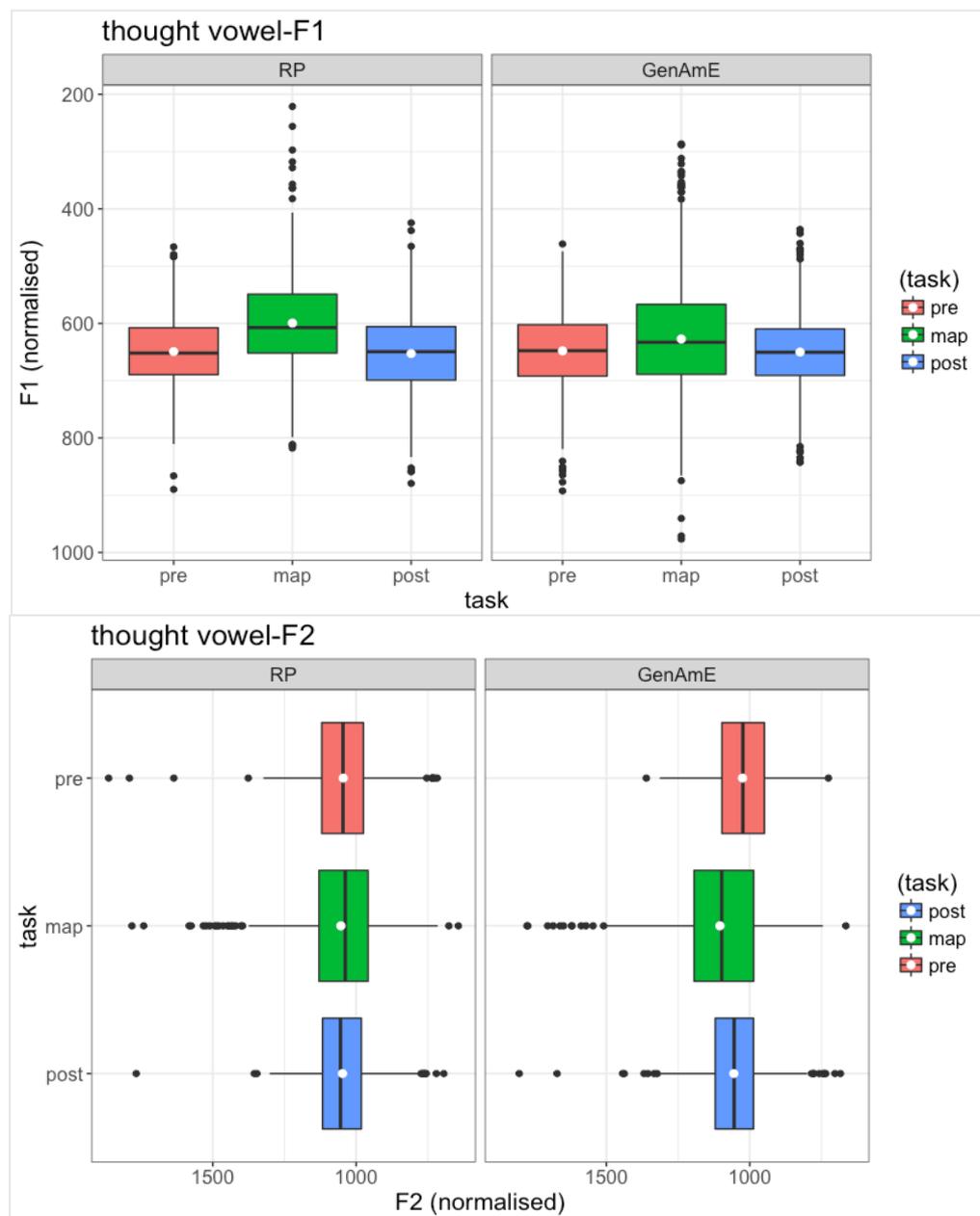


Figure 6.7 F1 and F2 changes in THOUGHT vowels across the tasks in the two conditions. White dots in the boxplots represent the means, while the black line in the middle represents the medians. Black dots are outliers. Note that the scales were inverted and the boxplots of F2 were flipped for the convenience of viewing.

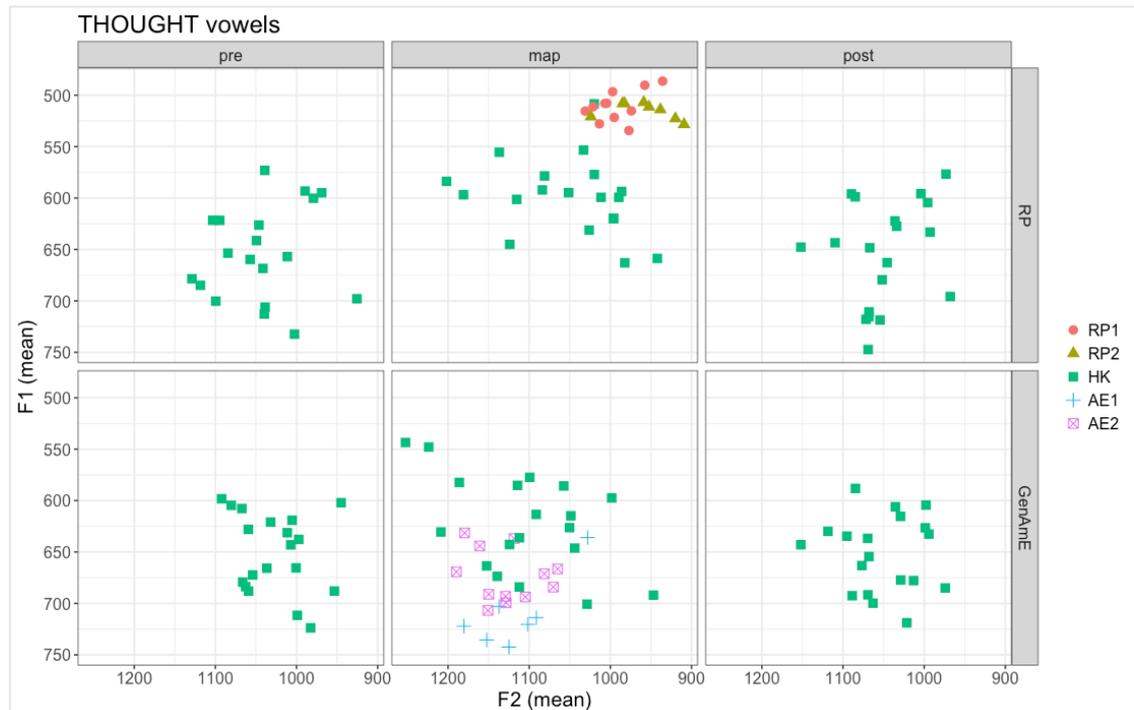


Figure 6.8 The shift of THOUGHT vowels across tasks and conditions. Each of the green plots represents one participant; the native interlocutors are shown using different coloured and shaped dots.

Linear mixed effects regressions

To examine the significance of the participants' changes in F1 and F2, a series of linear mixed effects regressions were run. The full model included the following fixed effects: task, exposure, the interaction between task and exposure, and participant sex. It included random intercepts by participant, by word and by interlocutor. It also included random slopes by participant over task, exposure and the interaction between task and exposure. The dependent variables were F1 and F2 values, which are continuous variables.

Interlocutor was included as a random effect instead of a fixed effect because the present study was not interested in the convergence caused by individual interlocutors. Also, **interlocutor** would be used as a fixed effect only if the convergence effect caused by these native interlocutors can be repeated with a different group of HKE participants, which apparently is not the assumption of the present study.

Task refers to the pre-tasks, map tasks and the post-tasks the participant completed in the experiments. **Exposure** refers to the accent they were exposed to in the map tasks, either RP or GenAmE. **Participant sex** refers to their self-reported biological sex. **Participant**

refers to the 19 participants who were involved in the study. **Word** refers to the different words that represent the target vowel. **Interlocutor** refers to the specific native interlocutors the participants talked to.

The formula of the full model is shown below:

Full model = Formant value ~ task*exposure + sex + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

To test the effect of task*exposure, a nested model with task*exposure removed was compared with the full model using ANOVA.

The formula of the nested model is:

Nested model = Formant value ~ task + exposure + sex + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

The comparison suggests that the interaction of **task*exposure** significantly improves the model fit for F1 (Chi-square = 8.8224, DF=2, $p = 0.012$) and F2 (Chi-square = 18.076, DF=2, $p < .001$). In other words, exposure to different English accents had an effect on the participants' changes in F1 and F2 across tasks.

Post-hoc tests of four specific contrasts (pre vs map, pre vs post in both RP and GenAmE conditions) were conducted separately for F1 and F2, with adjusted p values in the Bonferroni method. Table 6.8 shows the results of post-hoc tests.

Post-hoc tests of <i>task*exposure</i>		
RP condition, F1		
pre-map = 54.291	SE =8.335, z =6.514,	$p < .001^*$
pre-post = -1.639	SE =5.354, z =-0.306,	$p = 1.000$
RP condition, F2		
pre-map = -10.232	SE =14.619, z = -0.700,	$p = 1.000$
pre-post = -6.685	SE = 8.437, z = -0.792,	$p = 1.000$
GenAmE condition, F1		
pre-map = 27.229	SE =10.846, z =2.511,	$p = 0.048^*$
pre-post = -1.118	SE =5.452, z =-0.205,	$p = 1.000$
GenAmE condition, F2		
pre-map = -75.934	SE =16.331, z = -4.650,	$p < .001^*$
pre-post = -24.590	SE =7.780, z =-3.161,	$p = .006^*$

Table 6.8 Post-hoc tests for F1 and F2 of THOUGHT vowel across the tasks in both RP and GenAmE condition. * indicates a significant effect.

In summary, the results for F1 and F2 partially support H1. In the RP condition, the participants converged their F1 and F2 towards the RP speakers as H1 predicted, but a significant effect was only found in the F1 dimension. In the GenAmE condition, they significantly converged their F2 towards the GenAmE speakers as H1 predicted, but they significantly diverged from the GenAmE speakers in the F1 dimension.

6.4.1.1.2. Euclidean distance

Euclidean distances between the native speakers' THOUGHT vowels and the participants' THOUGHT vowels were calculated for the three tasks in the two conditions.

According to H1, if the participants converge towards the native accents, map-distance and post-distance are expected to be shorter than pre-distance. If they diverge from the native accents, map-distance and post-distance are expected to be larger than pre-distance.

Table 6.9 and Figure 6.9 show the results for Euclidean distance. As shown in these results, the map-distance is smaller than the pre-distance and the post-distance, suggesting a trend of convergence in the RP condition. On the other hand, the map-distance is larger than the pre-distance and the post-distance in the GenAmE condition, indicating a divergence. To examine the statistical significance, the results of linear mixed effects regressions are reported in the next section.

THOUGHT vowel (Hz)		pre	map	post	pre→map change	pre→post change
RP condition	distance	193 (88)	178 (102)	194 (83)	smaller -15	larger +1
GenAmE condition	distance	153 (85)	177 (113)	139 (84)	larger +24	smaller -14

Table 6.9 Euclidean distance of the THOUGHT vowel across the three tasks and the two conditions. The figures in the columns of “pre→map change” and “pre→post change” refer to the change from the pre-distance to the map-distance/post-distance. No significant effect was found based on linear mixed effect regressions. “.” indicates a marginal significance at 0.01.

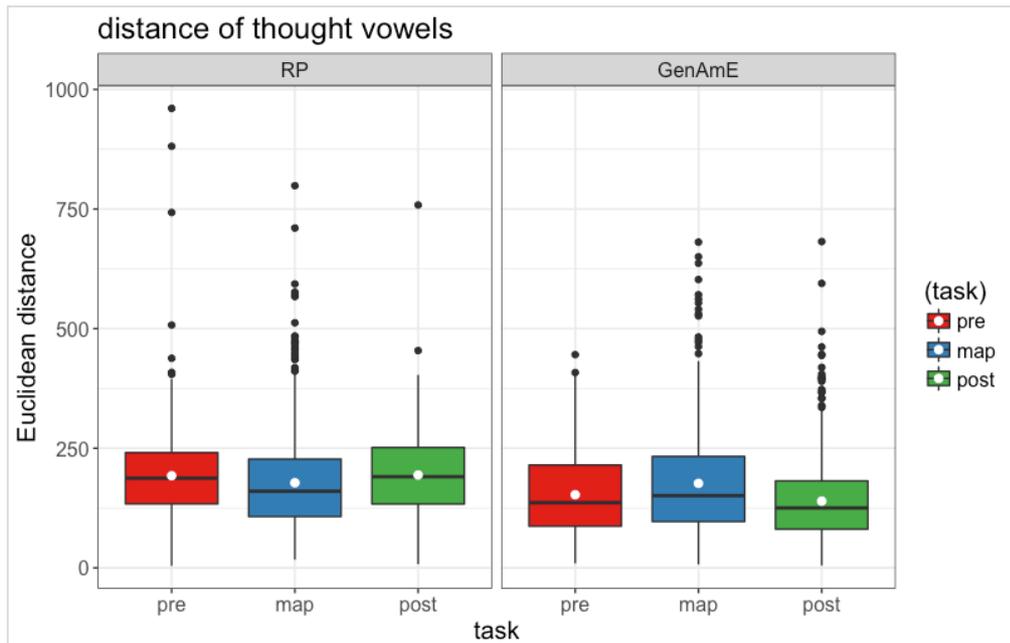


Figure 6.9 Euclidean distance of the THOUGHT vowel across the tasks in the two conditions. White dots in the boxplots represent the means of the distances. Black dots are outliers.

Linear mixed effects regressions

To perform a clearer picture of convergence, the data set was divided into two subsets. One subset only contained the data of the pre-tasks and the map tasks (henceforth Subset_premap), which tested the main effect of convergence in the map tasks. The other subset contained the data of the pre-tasks and the post-tasks (henceforth Subset_prepost), which tested whether the convergence effect would last till the post-tasks.

A series of linear mixed effects regressions were run for Euclidean distance of each subset. The full model used for Euclidean distance is the same as the one used for F1 and F2, except that the dependent variable was the Euclidean distances of the THOUGHT vowel. A nested model with task*exposure removed was compared with the full model using ANOVA.

The formulas of the full model and nested model for Euclidean distance are shown below:

Full model = distance ~ task*exposure + sex + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

Nested model = distance ~ task + exposure + sex + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

For **Subset_premap**, the comparison suggests that the interaction of **task*exposure** significantly improves the model fit (Chi-square = 6.6618, DF=1, $p = 0.009$). In other words, the participants' accommodation from the pre-tasks to the map tasks was significantly different in the two conditions. This is shown in Figure 6.10, where the participants converged in the RP condition but diverged in the GenAmE condition.

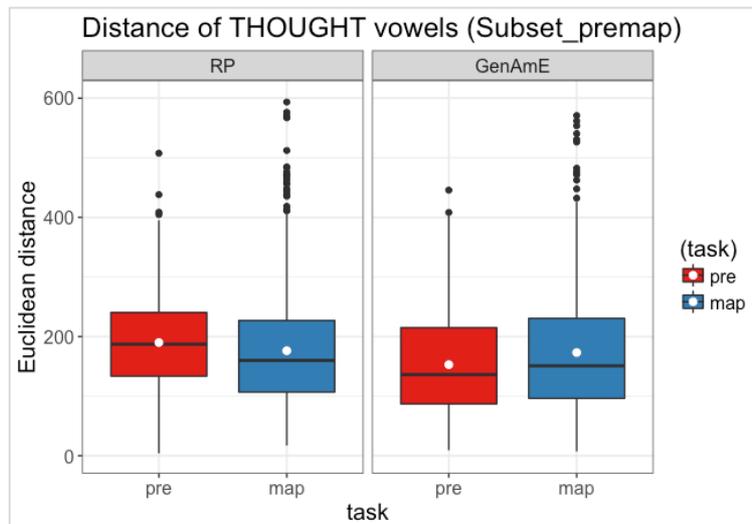


Figure 6.10 Euclidean distance of the THOUGHT vowel from the pre-tasks to the map tasks in the two conditions. To present a closer view of the changes, y axis was adjusted to 600 Hz. 9 outliers (black dots) are not shown in this figure.

To test whether the participants' convergence/divergence in each of the conditions was significant or not, post-hoc tests of **task*exposure** were conducted, with adjusted p values in the Bonferroni method. The results are shown in Table 6.10.

Post-hoc tests of <i>task (pre × map) * exposure (RP × GenAmE)</i>			
RP condition	pre-map = 16.59,	SE = 12.72, z = 1.304,	$p = 0.266$
GenAmE condition	pre-map = -23.27,	SE = 10.56, z = -2.204,	$p = 0.055$

Table 6.10 The post-hoc tests results of Euclidean distance for task*exposure of the THOUGHT vowel. · indicates a significant effect at 0.01 level.

The results of the post-hoc tests did not contradict with the significance of task*exposure. Post-hoc tests compared the pre-distance with the post-distance in each condition, whereas the significance of task*exposure indicates that the changes of Euclidean distance from the pre-tasks to the map tasks in the RP condition were significantly different from the changes in the GenAmE condition.

For **Subset_prepost**, the same model comparison was run. The interaction of **task*exposure** does NOT improve the model fit (Chi-square = 1.2541, DF=1, $p = 0.26$), indicating that the participants' accommodation from the pre-tasks to the post-tasks was not significantly different in the two conditions. Figure 6.11 shows that the changes of Euclidean distance from the pre-tasks to the post-tasks in the two conditions.

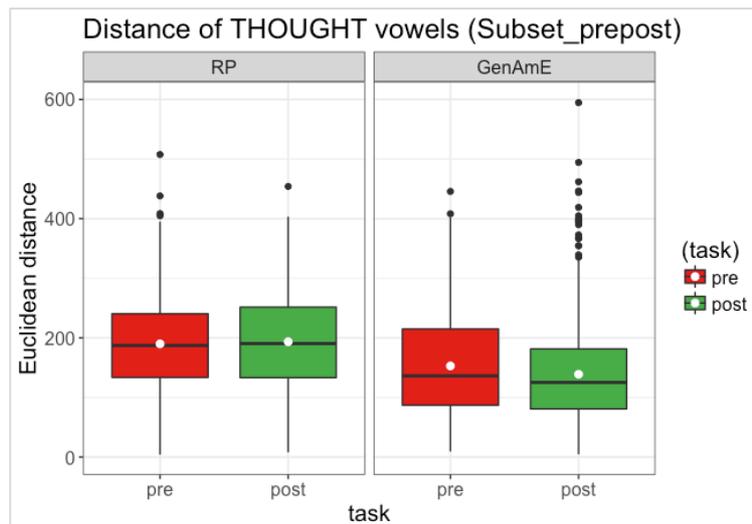


Figure 6.11 Euclidean distance of the THOUGHT vowel from the pre-tasks to the post-tasks in the two conditions. To present a closer view of the changes, y axis was adjusted to 600 Hz. 5 outliers (black dots) are not shown in this figure.

For the Subset_prepost, post-hoc tests with adjusted p values in the Bonferroni method were conducted. No significance was found between pre-distance and post-distance in both conditions, as shown in Table 6.11.

Post-hoc tests of task (pre × post) * exposure (RP × GenAmE)			
RP condition	pre-post = -2.254,	SE = 7.939, z = -0.284,	$p = 1.000$
GenAmE condition	pre-post = 12.038,	SE = 7.271, z = 1.656,	$p = 0.196$

Table 6.11 The post-hoc tests results of Euclidean distance for task*exposure of the THOUGHT vowel.

Note that the results for Euclidean distance here are not fully in line with the results for F1 and F2 in the previous section. The results for F1 and F2 suggested a convergence of F1 in the RP condition and a convergence of F2 in the GenAmE condition. Similar results were also found in Pardo et al. (2013) and Pardo et al. (2017) where participants converged and diverged in the multiple dimensions at the same time. Pardo et al. (2017) found a marginal convergence on Euclidean distance and F1, and no significant convergence was found in F2. In Pardo et al. (2013), vowels showed different patterns of

convergence in F0, duration and Euclidean distance. For instance, [ɑ] in high frequency words converged in duration and Euclidean distance but diverged in F0.

The inconsistent results found in the present study might be due to two reasons.

Firstly, F1 and F2 only reflected the change in a single dimension, whereas Euclidean distance considered the dimensions of F1 and F2 at the same time, and calculated the change as a whole unit. It is possible that a change may be considered convergence in the F1 and F2 dimensions but would be regarded as a divergence in Euclidean distance.

One example is shown in Figure 6.12. Figure 6.12 shows a HKE speaker's change of his/her THOUGHT vowel from the pre-task to the map task. A token of the participant's THOUGHT vowel in the pre-task is plotted with its F1 and F2, represented by the blue circle in the bottom left corner. A token of the THOUGHT vowel in the map task is plotted with its F1' and F2', represented by the blue circle in the top right corner. Similarly, a token of the GenAmE speaker's THOUGHT vowel is also plotted with its F1[^] and F2[^], shown as the green square. Judging from the F1 and F2 of these three points, the conclusion would be that the HKE speaker's F1 becomes higher and F2 becomes more back from the pre-task to the map task, therefore the HKE speaker converges towards the GenAmE speaker in the F1 and F2 dimension. However, if we calculate Euclidean distance and compare the pre-distance with the map-distance, because the map-distance is larger than the pre-distance, the conclusion would be that the HKE speaker diverges from the GenAmE speaker.

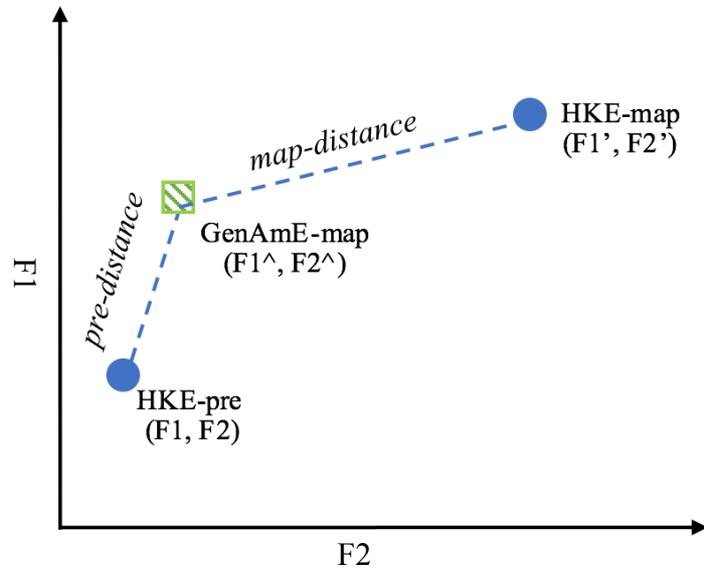


Figure 6.12 An example of judging convergence/divergence based on F1 and F2, and Euclidean distance. The green square represents the GenAmE speaker's THOUGHT vowel, the blue circle on the left bottom corner represents the HKE's THOUGHT vowel in the pre-task, the other blue circle on the right top corner represent his/her THOUGHT vowel in the map task.

Although the present study calculated Euclidean distances in a slightly different way from the above (it calculated the Euclidean distances between each THOUGHT token and the means of the native speaker's THOUGHT vowel), this example still shows how different conclusions can be drawn from F1 and F2, and Euclidean distance. Both methods reveal some parts of the whole picture. However, as Euclidean distance captures the changes in the F1 and F2 dimensions at the same time, it has an advantage in terms of determining the convergence of a sound as the whole unit. An alternative method to test convergence would be a AXB perceptual similarity test as used in Goldinger (1998) and Pardo (2006). However, due to the time constraints, the perceptual judgement task was not included in this PhD project.

Secondly, the results for F1 and F2 only found a significant convergence either in the F1 dimension (i.e. in the RP condition) or in the F2 dimension (i.e. in the GenAmE condition). The change in only one dimension might not be strong enough to elicit a significant effect on Euclidean distance. Therefore, it is not surprising to find a significant effect only on F1 and/or F2 but not on Euclidean distance.

In summary, the results for Euclidean distance suggest that the distances between the participants' THOUGHT vowels and the RP speakers' THOUGHT vowels were reduced

from the pre-tasks to the map tasks, but the differences were not significant. In the GenAmE condition, the distances between the participants and the GenAmE speakers were larger from the pre-tasks to the map tasks, however, the differences were not significant either. The linear mixed effects regressions suggest that **task*exposure** is a significant predictor for the model of Subset_premap: in other words, **the changes of Euclidean distance from the pre-tasks to the map tasks in the RP condition were significantly different from the changes of which in the GenAmE condition for the THOUGHT vowel.**

6.4.1.1.3. Accommodation at an individual level

Previous sections reported the overall effect of convergence on the THOUGHT vowel; however, individuals might not converge towards the native speakers from the same starting point. For example, Table 6.6 shows that some HKE participants spoke the THOUGHT vowels with an American accent, whereas some participants spoke the vowels with an RP accent. To understand whether the participants' accent of the THOUGHT vowel would affect their convergence, the difference of pre-distances in the two exposure conditions is subtracted (i.e. Difference of distance = Pre-distance_{GenAmE condition} – Pre-distance_{RP condition}) for each participant and plotted in Figure 6.13.

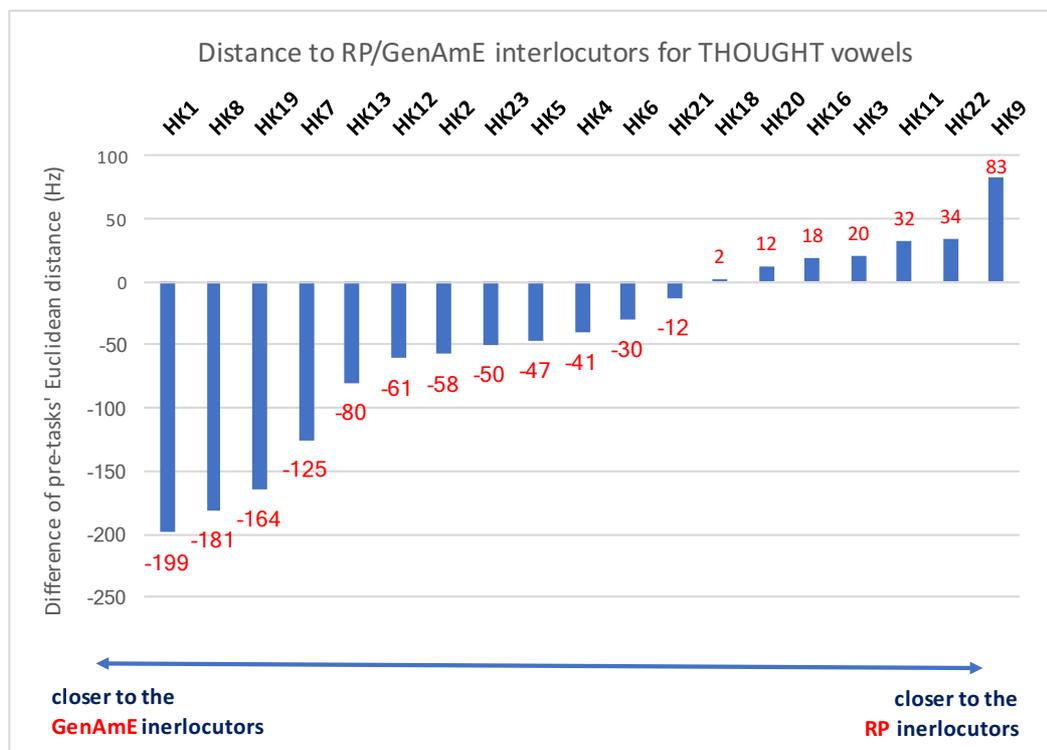


Figure 6.13 The difference of pre-distance between the two exposure conditions for the THOUGHT vowel. The smaller the value is, the closer the distance is to the GenAmE speakers.

If the participant's THOUGHT vowel was closer to the GenAmE interlocutor in comparison to the RP interlocutor, he/she would have a negative value in Figure 6.13. On the other hand, if the participant's THOUGHT vowel was closer to the RP interlocutor, he/she would have a positive value. That is, the smaller the value is, the closer the participant's THOUGHT vowel is to the GenAmE speaker.

Based on the order of the participants shown in Figure 6.13, the Euclidean distances across the tasks and two exposure conditions for each participant are shown in Figure 6.14. The participants were grouped into three coloured frames based on their patterns of accommodation in the two exposure conditions.

For the participants in the red frame, **it seems that they were more likely to converge towards the interlocutor who was further away from them.** For example, HK1, HK8, HK19, HK7, HK13 and HK23 whose THOUGHT vowels were further away from the RP interlocutors showed a convergence in the RP condition, but in the GenAmE condition only a small divergence was found. Specifically, the degree of convergence seems to correlate with the participant's relative distance of the THOUGHT vowel to the RP interlocutors. The further the participant's THOUGHT vowel was from the RP speaker (the more leftward they were placed in Figure 6.14), the more he/she converged in the RP condition. It could be that the large distances between the participants and the RP interlocutors provide spaces for them to converge toward the RP interlocutors. Also, they might be more likely to notice the differences of pronunciation when the acoustic distances are larger.

A similar pattern is also observed on HK9 who is in the orange frame in Figure 6.14. HK9 had the largest distance from the GenAmE interlocutor on the THOUGHT vowel; he then showed a trend of convergence in the GenAmE condition but not in the RP condition.

For the participants who are in the blue frame, their THOUGHT vowels were neither too close to the GenAmE interlocutors nor too close to the RP interlocutors. For these participants, **most of them had the same pattern of accommodation in the two exposure conditions.** If they converge in the RP condition, they would converge in the

GenAmE condition too, such as HK6 and HK18. On the other hand, if they diverge in the RP condition, the same pattern was found in the GenAmE condition, such as HK5, HK4 and HK20. In other words, the distances of the THOUGHT vowel between the participants and the native interlocutors did not seem to affect the participants' accommodation. Their accommodation patterns (e.g. convergence or divergence) might be affected by other factors such as individual's ability of imitation.

Figure 6.15 shows the shifting of the THOUGHT vowels from the pre-task to the map task for each participant of the two groups. In Figure 6.15, each participant has two sets of arrows, i.e. the red arrows represent the RP condition, and the blue arrows represent the GenAmE condition. The start of each arrow represents the location of the participant's THOUGHT vowel in the pre-tasks, plotted with the mean of the vowel in pre-tasks. The end of the arrow represents the location of the vowel in the map tasks, plotted with the mean of the vowel in the map tasks. The locations of the native speakers' THOUGHT vowels are also provided. Note that in Figure 6.15 the locations of the native speakers are different for each participant. Instead of using overall means, the plotting used the means of the native interlocutors that each participant was exposed to. For example, HK1 talked to RP1 and AE1 in the two map tasks; the two dots in the plot of HK1 represent the means of RP1 and AE1's THOUGHT vowels in their conversations with HK1. In this way, a more accurate shift between the participants and the native interlocutors can be shown.

THOUGHT vowel-Participants' Euclidean distances of across tasks

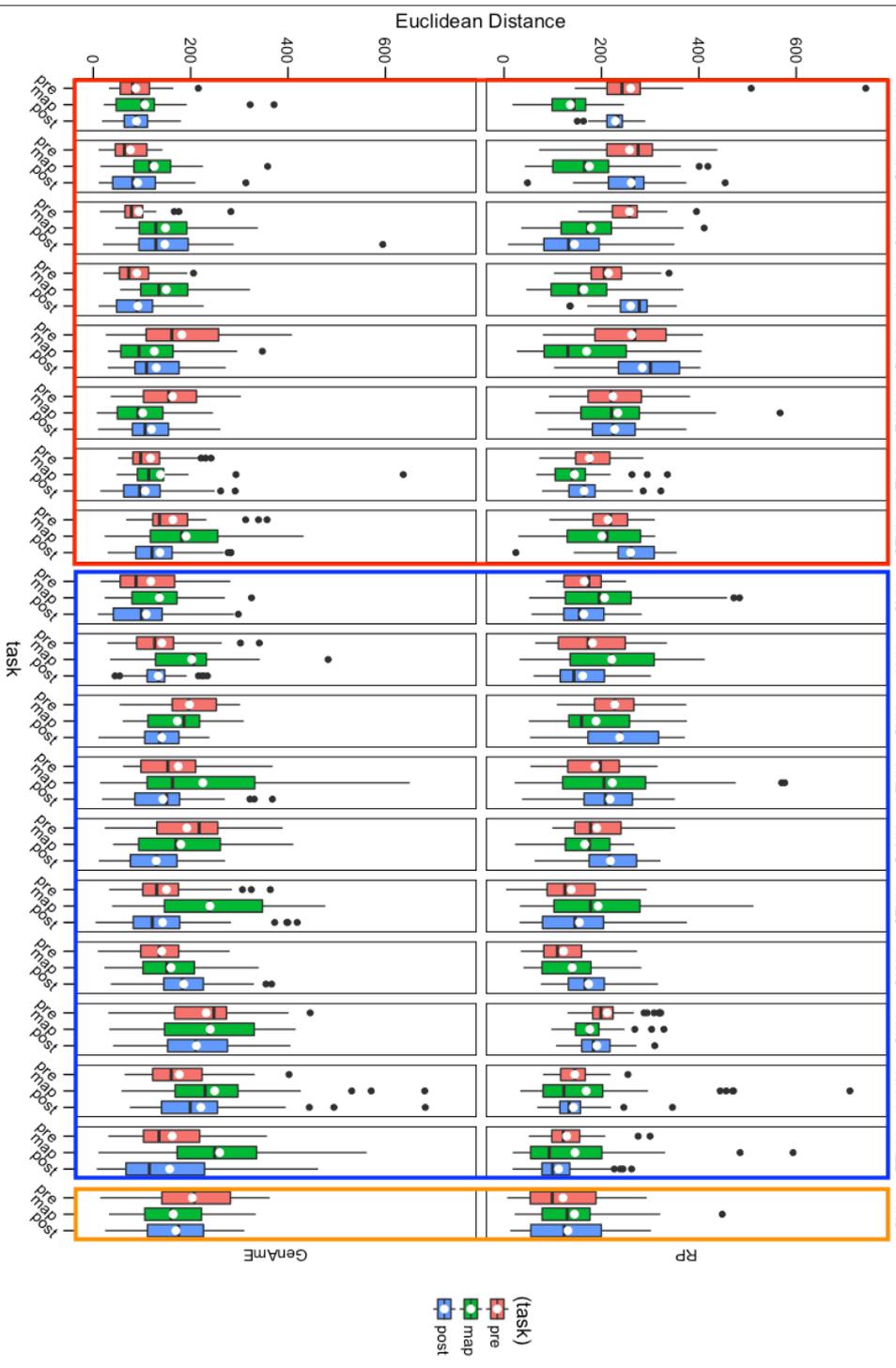


Figure 6.14 Euclidean distance of the THOUGHT vowel across tasks and conditions for each participant. The participants in the red frame converged in the RP condition but diverged in the GenAmE condition. The participant in the blue frame had the same pattern of accommodation in the two exposure conditions. The participant in the orange frame converged in the GenAmE condition but diverged in the RP condition. From left to right, the distance between the participant's THOUGHT vowel and the GenAmE interlocutor's THOUGHT vowel decreased.

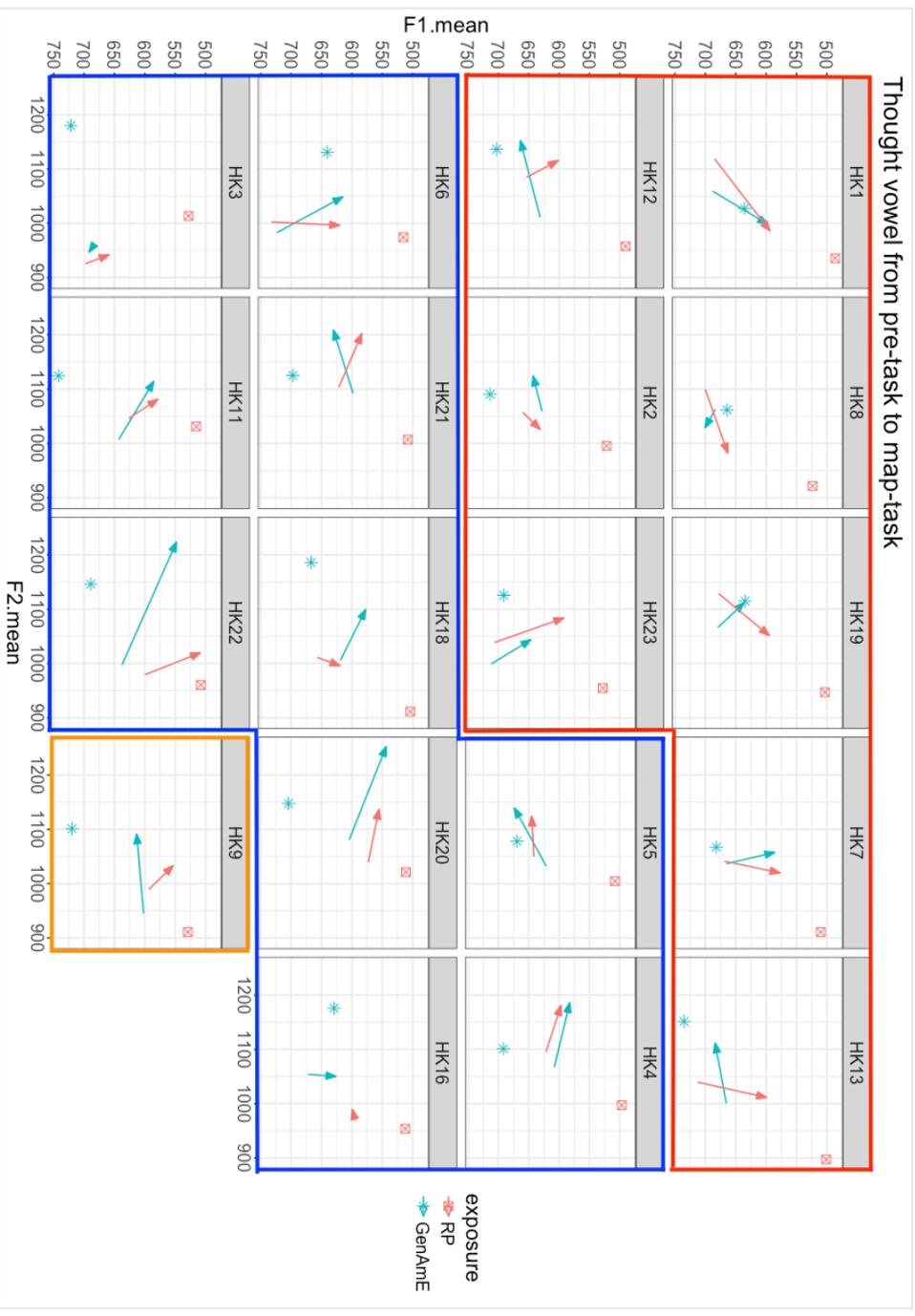


Figure 6.15 Individual shifts of THOUGHT vowels from the pre-task to the map tasks. The start point of an arrow represents the participant's THOUGHT vowel in the pre-task, the end of an arrow represents the vowel in the map tasks. The pink/blue dots represent the native interlocutors.

6.4.1.1.4. Summary

This section reported the results for the THOUGHT vowel. H1 predicts that if the participants accommodate towards the native speakers, (1) F1 values are expected to decrease and F2 values are expected to increase when they talk to the RP speakers; (2) in the GenAmE condition, F1 values are expected to increase or have no change and F2 values are expected to increase; and (3) Euclidean distances of the map tasks are expected to be shorter than the pre-tasks in both conditions.

The results are summarized in Table 6.12.

THOUGHT vowel		Pre → Map		Pre → Post	
		Results	Match?	Results	Match?
RP condition	F1	decrease (-50)*	YES	increase (+4)	-
	F2	increase (+9)	YES	increase (+4)	YES
	Euclidean Distance	smaller (-15)	YES	larger (+1)	-
		Pre → Map		Pre → Post	
GenAmE condition	F1	decrease (-21)*	-	increase (+2)	YES
	F2	increase (+78)*	YES	increase (+30)*	YES
	Euclidean Distance	larger (+24)	-	smaller (-14)	YES
Significant predictor					
Linear mixed effects regression	F1	task*exposure (Chi-square = 8.8224, DF=2, $p = 0.012$)			
	F2	task*exposure (Chi-square = 18.076, DF=2, $p < .001$)			
	Euclidean Distance (Subset_premap)	task*exposure (Chi-square = 5.3294, DF=1, $p = 0.02$)			
	Euclidean Distance (Subset_prepost)	task*exposure (<i>not significant</i>)			

Table 6.12 Summary of predictions and actual results for the THOUGHT vowel. Only when the result matches the H1 prediction is a capital “YES” marked in the “Match?” column. * indicates a significant effect based on the post-hoc tests of the linear mixed effects regressions.

In the RP condition, the participants behaved as H1 predicted. They significantly converged their F1 and F2 towards the RP speakers from the pre-task to the map task, and had a shorter Euclidean distance in the map task; however, only the change in F1 was significant. In the GenAmE condition, from the pre-task to the map task, the participants only matched the prediction in the F2 dimension. From the pre-tasks to the post-tasks, the participants behaved as H1 predicted; similarly, only F2 had a significant effect.

The results of the linear mixed effects models suggest a significant effect of the interaction of **task** and **exposure** for F1, F2 and Euclidean distance of Subset_premap. This suggests that the exposure to the native accents in 1 hour made a difference to the participants' pronunciation across tasks.

At the individual level, some HKE participants seem to be more likely to accommodate towards the native interlocutor who was more further away.

6.4.1.2. PATH vowel

6.4.1.2.1. Results for F1 and F2

Descriptive results

H2 predicts that if the participants shift towards the native speakers in F1 and F2, the Euclidean distance in the map tasks will be shorter than the distance in the pre-tasks.

Figure 6.16 below shows the actual locations of PATH vowels for the HKE participants, the RP speakers and the GenAmE speakers in the present study. The actual locations of PATH vowels were plotted using the means of F1 and F2 shown in Table 6.13. The hypothesized locations of the PATH vowel are also shown in Figure 6.16.

Similar to the THOUGHT vowel, where actual locations were different from the hypothesized locations, the actual locations of PATH vowel are also different from the hypothesized ones. As shown in Figure 6.16, the three PATH vowels are very close to each other. It could be that the GenAmE speakers of the present study had adopted some features of British English in their accent during their study in the U.K, therefore their PATH vowels are very close to the RP speakers' PATH vowels. Another possibility is that the native speakers accommodated their PATH vowels towards the HKE speakers in the map tasks, resulting in the small distances of the PATH vowels between HKE speakers, the RP speakers and the GenAmE speakers.

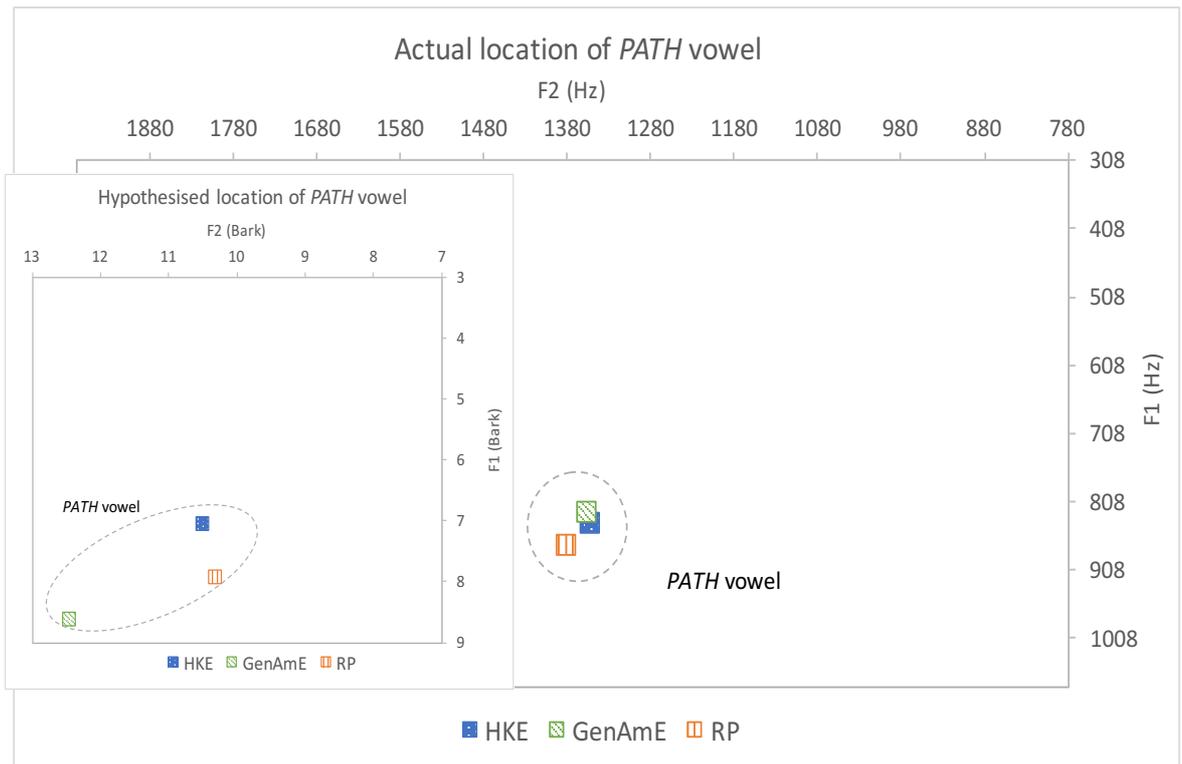


Figure 6.16 The actual location of the PATH vowel for the HKE speakers, the RP speakers and the GenAmE speakers in Study 2. The small window on the left side shows the hypothesized location of the PATH vowel based on Hagiwara (1997), Deterding (1997) and Deterding et al. (2001).

Based on the actual locations of the PATH vowel, H2 is adjusted as below:

If the participants accommodate towards the native speakers, (1) their F1 is expected to shift lower (i.e. F1 values to increase) and F2 is expected to shift to be more front (F2 values to increase) in the RP condition; (2) in the GenAmE condition, their F1 and F2 are expected to have little change because the locations of the HKE-*thought* vowel and GenAmE-*thought* vowel are too close; (3) Euclidean distance will be shorter in the map tasks compared to the pre-tasks and post-tasks in the RP condition but not in the GenAmE condition.

The results for F1 and F2 are shown in Table 6.13. The changes mainly occurred between the pre-tasks and the map tasks. Different from the predictions in H2, the participants significantly decreased their F1 in the RP condition. On the other hand, in the GenAmE condition, the participants decreased F1 and increased F2, and both the changes were significant. It is worth noting that the standard deviations of the GenAmE speakers were twice as large as those of the RP speakers, suggesting a larger variance in the GenAmE condition.

PATH vowel (in Hz)		pre	map	post	NS	pre→map change	pre→post change
RP condition	F1	841 (70)	807 (108)	841 (68)	873 (67)	decrease* -34	no change 0
	F2	1370 (134)	1373 (145)	1354 (128)	1381 (84)	increase +3	decrease -16
GenAmE condition	F1	839 (65)	799 (119)	839 (68)	825 (125)	decrease* -40	no change 0
	F2	1353 (130)	1434 (192)	1361 (146)	1357 (184)	increase* +81	increase +8

Table 6.13 Overall mean and standard deviation (in brackets) of F1 and F2 across tasks and conditions for PATH vowels. NS represents overall means of the native interlocutors. The figures in the columns of “pre→map change” and “pre→post change” refer to the change from the pre-distance to the map-distance/post-distance. * indicates a significant effect based on linear mixed effects regressions.

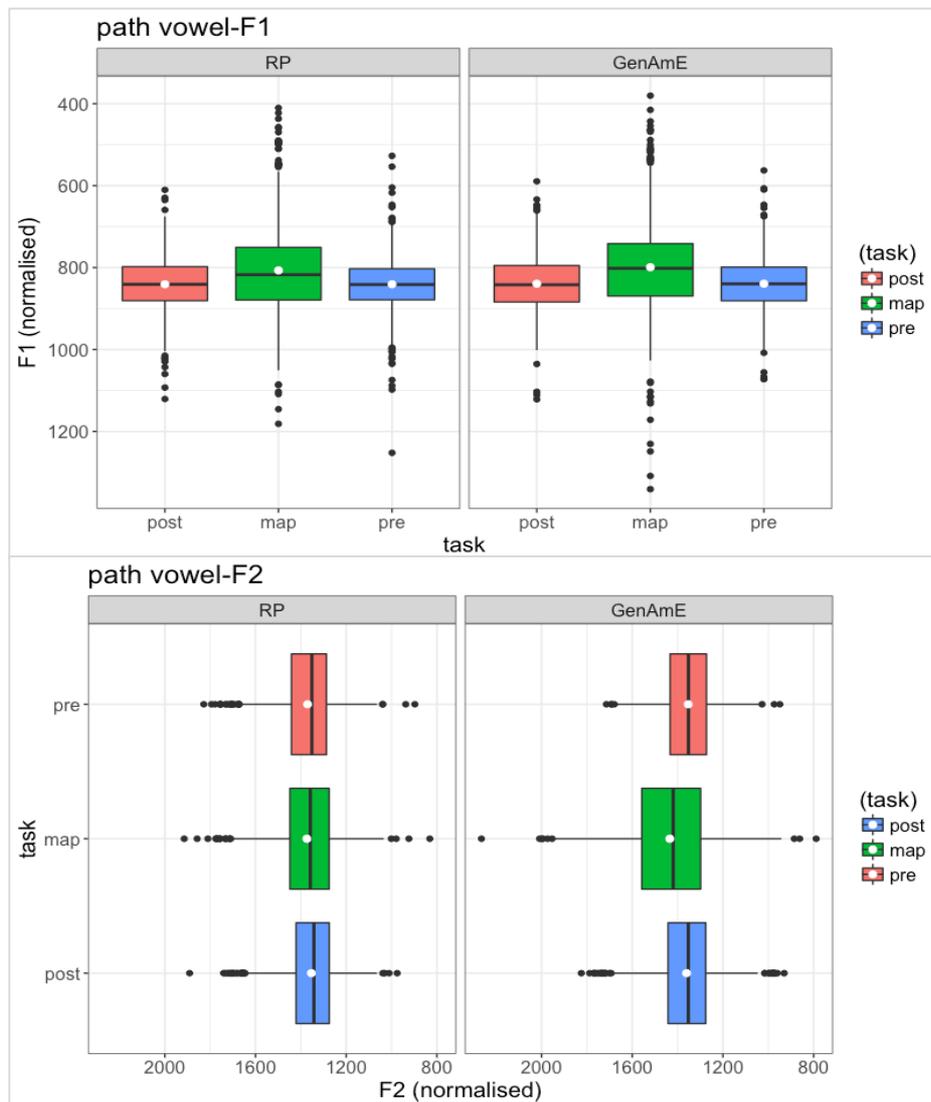


Figure 6.17 F1 and F2 changes in PATH vowels across the tasks in the two conditions. White dots in the boxplots represent the means, while the black line in the middle represents the medians. Black dots are outliers. Note that the scales were inverted and the boxplots of F2 were flipped for the convenience of viewing.

Figure 6.17 visualizes the results of Table 6.13. Figure 6.17 shows that the participants shifted their PATH vowels to be slightly higher from the pre-task to the map task in the RP condition; on the other hand, they shifted the vowels higher and more front from the pre-task to the map task in the GenAmE condition.

The shifting across the tasks is also shown in Figure 6.18. Figure 6.18 plots the mean F1 and F2 of each participant's PATH vowel and the corresponding native speakers' PATH vowels across three tasks in the same vowel space.

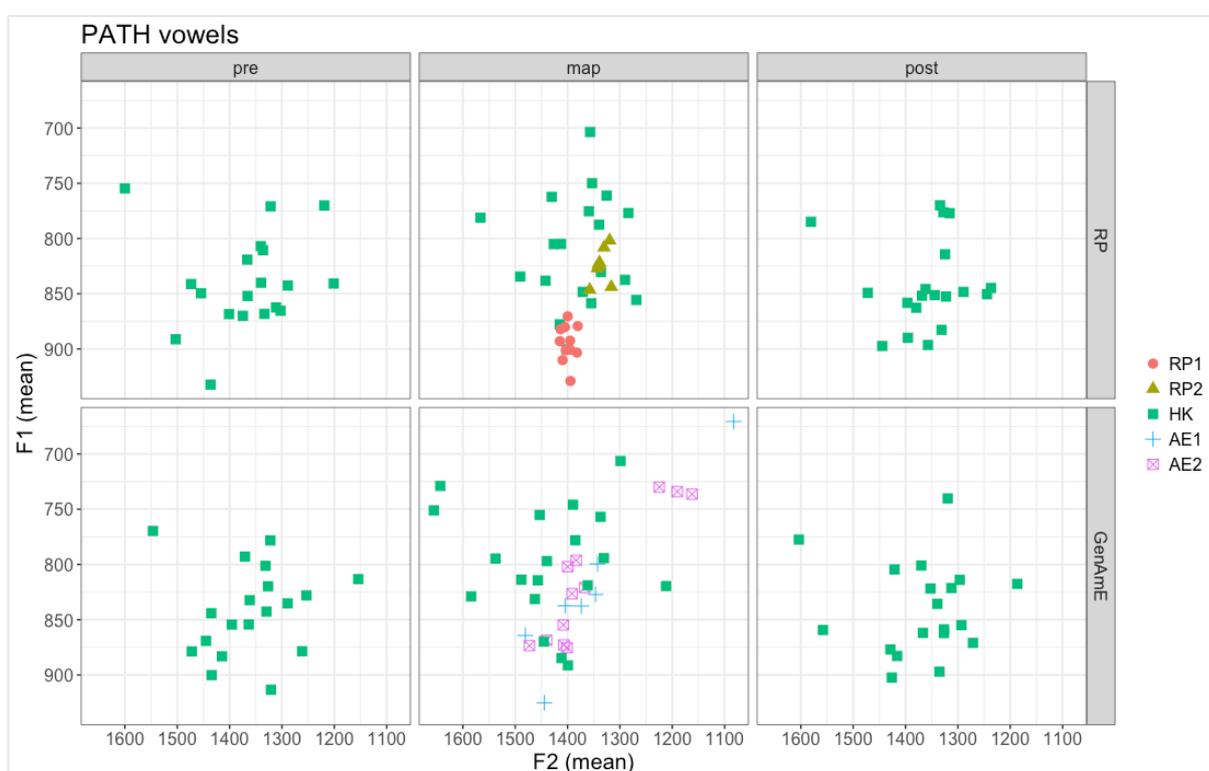


Figure 6.18 Shift of PATH vowels across tasks and conditions. Each of the green square plots represents one participant; the native interlocutors are shown in different coloured and shaped dots.

Linear mixed effects regressions

A series of linear mixed effects regressions were run for F1 and F2. The full model used here is the same as the one used for the THOUGHT vowel. The dependent variables were the F1 and F2 values of the PATH vowel. A nested model with task*exposure removed was compared with the full model using ANOVA.

The formulas of the full model and nested model for F1 and F2 are shown below:

Full model = Formant value ~ task*exposure + sex + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

Nested model = Formant value ~ task + exposure + sex + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

The comparison suggests that the interaction of **task*exposure** does not improve the model fit for F1 (Chi-square = 1.2617, DF=2, $p = 0.532$), but it significantly improves the model fit for **F2** (Chi-square = 14.585, DF=2, $p < .001$). In other words, exposure has no effect on the tasks for F1, but it has an effect on the tasks for F2.

Post-hoc tests of four specific contrasts (pre vs map, pre vs post in both the RP and GenAmE conditions) were conducted separately for F1 and F2, with adjusted p values in the Bonferroni method. Table 6.14 shows the results of post-hoc tests. The results suggest that in the RP condition, only the change in F1 was significant, while in the GenAmE condition, both the changes from the pre-tasks to the map tasks in F1 and F2 were significant.

Post-hoc tests of <i>task*exposure</i>			
RP condition, F1			
pre-map = 34.678,	SE = 8.645, z = 4.011,		$p < .001^*$
pre-post = 0.247,	SE = 5.978, z = 0.041,		$p = 1.000$
RP condition, F2			
pre-map = -4.480,	SE = 17.721, z = -0.253,		$p = 1.000$
pre-post = 13.862,	SE = 8.990, z = 1.542,		$p = 0.492$
GenAmE condition, F1			
pre-map = 43.939,	SE = 10.496, z = 4.186,		$p < .001^*$
pre-post = 1.228,	SE = 6.654, z = 0.185,		$p = 0.325$
GenAmE condition, F2			
pre-map = -74.940,	SE = 21.448, z = -3.494,		$p = .0019^*$
pre-post = -7.283,	SE = 12.234, z = -0.595,		$p = 1.000$

Table 6.14 Post-hoc tests for F1 and F2 of PATH vowel across the tasks in both RP and GenAmE condition. * indicates a significant effect.

As **task*exposure** was not a significant predictor for F1, it would be worth examining **task** and **exposure** as a separate fixed effect for F1. A different model called Model 1 was adopted for F1. Model 1 of the PATH vowel contained the same fixed effects (task, exposure and sex), the same random intercepts (by participant, by word and by interlocutor), and the same random slopes by participant over task and exposure.

Compared to the full model, Model 1 did not include **task*exposure** as the fixed effects factors and the random slopes. The dependent variables were the F1 values of the PATH vowels.

The formula of the Model 1 for the PATH vowel is shown below:

Model 1 = Formant value ~ task + exposure + sex + (task + exposure | participant) + (1 | word) + (1 | interlocutor)

Three nested models were run separately with **task**, **exposure** and **sex** removed from the fixed effects factors, and ANOVA was used to compare the nested models and the Model 1. The results of the ANOVA including the Chi-square, DF and *p* value are shown in Table 6.15. As each predictor was removed as a whole, only a single set of statistics is shown for **task**, even when it had more than a single corresponding estimate.

F1 value of the PATH vowel				
Term	Estimate	Chi-square	DF	<i>p</i> value
task (map)	-39.4839	17.061	2	<.001*
task (post)	-0.7103	-	-	-
exposure (RP)	4.5483	0.5709	1	0.449
sex (male)	-45.6370	7.8718	1	0.005*

Table 6.15 Summary of the model comparisons for F1 of the PATH vowel. The estimates represent comparisons against a reference value (pre-task for **task**, GenAmE for **exposure**, female for **sex**). The Chi-squares, degree of freedom and p-values are taken from ANOVA results of comparing the nested models with the first model. * indicates a significant effect.

Comparison of the models suggests that **task and sex were significant predictors but exposure was not a significant predictor for F1**. This confirms the results of the post-hoc tests. Exposing to different accents had no effect on the participants' F1, but completing the map tasks elicited a divergence from the native interlocutors on their F1.

Summary

In summary, **the results for F1 and F2 of the PATH vowel did not support H2**. In the RP condition, H2 predicted an increase on F1 and F2, however, the results suggest a significant decrease on F1 and no significant change on F2. In the GenAmE condition, H2 predicted that little change would be found in F1 and F2, however, the results suggest a significant decrease on F1 and a significant increase on F2.

6.4.1.2.2. Euclidean distance

Euclidean distances between the native speakers' PATH vowels and the participants' PATH vowels were calculated for the three tasks in the two conditions.

According to H2, map-distances and post-distances are expected to be shorter than pre-distances in the RP condition but not in the GenAmE condition. If they diverge from the native accents, map-distances and post-distances are expected to be larger than pre-distances.

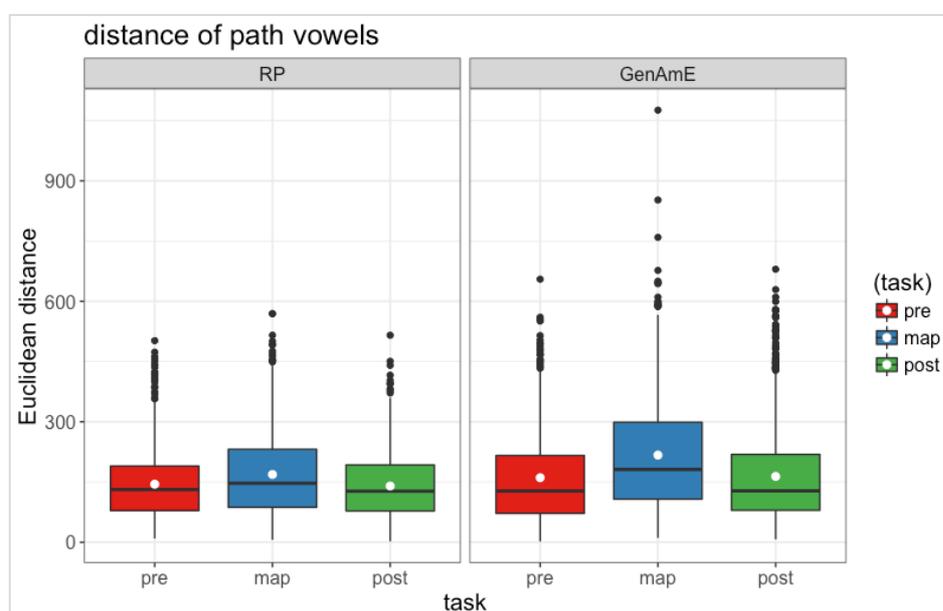


Figure 6.19 Euclidean distance of the PATH vowels across the tasks in the two conditions. White dots in the boxplots represent the means of the distances. Black dots are outliers.

PATH vowel		pre	map	post	pre→map change	pre→post change
RP condition	distance	145 (87)	169 (105)	140 (82)	larger* +24	no change -5
GenAmE condition	distance	161 (119)	217 (147)	164 (120)	larger* +56	no change +3

Table 6.16 Euclidean distance of the PATH vowel across tasks and conditions. Euclidean distance of the PATH vowel across the three tasks and the two conditions. The figures in the columns of “pre→map change” and “pre→post change” refer to the change from the pre-distance to the map-distance/post-distance. * indicates a significant effect based on the linear mixed effects regressions.

Descriptive results of Euclidean distance are shown in Table 6.16 and Figure 6.19. The results suggest that the map-distance is larger than the pre-distance and the post-distance

in both the RP and the GenAmE conditions. In other words, a trend of divergence is observed in both conditions.

To examine the statistical significance, results of linear mixed effects regressions are reported in the next section.

Linear mixed effects regression

The same procedure of linear mixed effects regressions was repeated for the Euclidean distance of the PATH vowel, except that the data set was split to two subsets: Subset_premap contained data of the pre-tasks and the map tasks, and Subset_prepost contained data of the pre-tasks and the post-tasks. For each subset, a full model was compared with a nested model with task*exposure removed.

The full model used for Euclidean distance is the same as the one used for F1 and F2, except that the dependent variable was the Euclidean distances of the PATH vowel. The full model contained the same fixed effects, the same random intercepts and random slopes as the one used for F1 and F2.

The formulas of the full model and the nested model of Euclidean distance are shown below:

Full model = distance ~ task*exposure + sex + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

Nested model = distance ~ task + exposure + sex + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

For **Subset_premap**, the results of the comparison suggested that adding **task*exposure** did NOT significantly improve the model fit (Chi-square = 2.056, DF=1, $p = 0.15$). In other words, the changes of the participants' Euclidean distance from the pre-tasks to the map tasks in the RP condition were not different from the changes of which in the GenAmE condition. Figure 6.20 shows the changes from the pre-tasks to the map tasks in the two conditions.

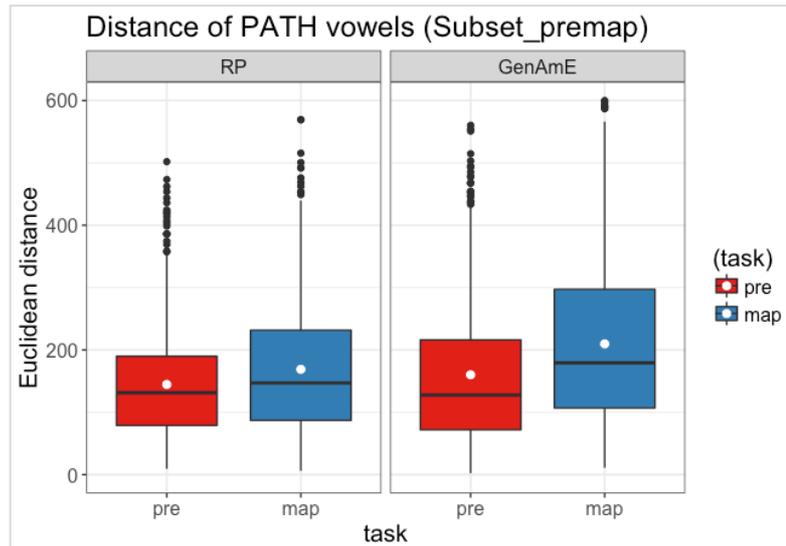


Figure 6.20 Euclidean distance of the PATH vowel from the pre-tasks to the map tasks in the two conditions. To present a closer view of the changes, y axis was adjusted to 600 Hz. 9 outliers (black dots) are not shown in this figure.

Post-hoc tests were conducted, with adjusted p value in the Bonferroni method. The results of post-hoc tests are shown in Table 6.17.

Post-hoc tests of <i>task (pre × map) * exposure (RP × GenAmE)</i>	
RP condition	pre-map = -27.318, SE = 9.331, $z = -2.928$, $p = 0.007^*$
GenAmE condition	pre-map = -54.694, SE = 15.838, $z = -3.453$, $p = 0.001^*$

Table 6.17 The post-hoc tests results of Euclidean distance for task*exposure of the PATH vowel. * indicates a significant effect.

As shown in Table 6.17, the results suggested that in both the RP and the GenAmE conditions, the map-distances were significantly larger than the pre-distances. In other words, **the participants' PATH vowels diverged from the native speakers from the pre-tasks to the map tasks in both exposure conditions. This does not support H2.**

The results of post-hoc tests do not contradict with the result of model comparison. The former tested the significance of changes from the pre-tasks to the map tasks IN each exposure condition, while the latter tested the significance of changes from the pre-tasks to the map tasks ACROSS the two exposure conditions.

For **Subset_prepost**, the full model was compared with the nested model, and the results suggested that **task*exposure** did NOT significantly improve the model fit (Chi-square = 1.271, DF=1, $p = 0.259$). That is, the participants' changes from the pre-tasks to the

post-tasks in the RP condition were similar to the changes of which in the GenAmE condition. This is shown in Figure 6.21.

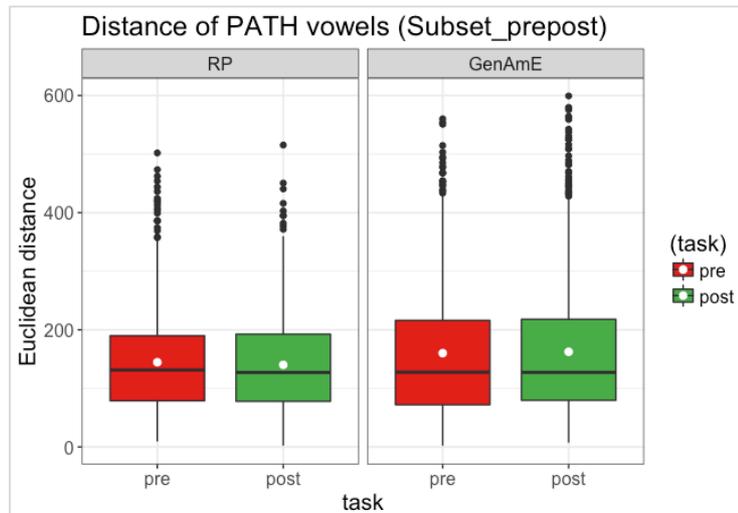


Figure 6.21 Euclidean distance of the PATH vowel from the pre-tasks to the post-tasks in the two conditions. To present a closer view of the changes, y axis was adjusted to 600 Hz. 4 outliers (black dots) are not shown in this figure.

Post-hoc tests were conducted, with adjusted p value in the Bonferroni method. The results of post-hoc tests are shown in Table 6.18.

Post-hoc tests of <i>task (pre × post) * exposure (RP × GenAmE)</i>	
RP condition	pre-post = 9.256, SE = 7.984, $z = 1.159$, $p = 0.493$
GenAmE condition	pre-post = -3.604, SE = 7.678, $z = -0.469$, $p = 1.000$

Table 6.18 The post-hoc tests results of Euclidean distance for task*exposure of the PATH vowel.

The results suggest that the changes from the pre-tasks to the post-tasks were not significant in both conditions.

Since **task*exposure** was not significant for the two subsets, the effect of **task** and **exposure** as an individual fixed effect was examined using a different set of models. A similar procedure was repeated for Euclidean distance. Model 2 contained the same fixed effects, random intercepts and random slopes as used in the Model 1; the dependent variable was Euclidean distance of PATH vowel.

The formula of the Model 2 for the PATH vowel is shown below:

Model 2 = distance ~ task + exposure + sex + (task + exposure | participant) + (1 | word) + (1 | interlocutor)

For each subset, three nested models were run separately with each of the fixed effect factors removed, and ANOVA was used to compare the nested models and the first model. The results of the ANOVA including the Chi-square, DF and *p* value are shown in Table 6.19.

As shown in Table 6.19, **task** was found to be significant for Subset_premap but not for Subset_prepost. Note that **exposure** was not significant in both subsets, suggesting that exposure to different English varieties had no effect on the participants' Euclidean distances. It is not surprising to find a non-significance of exposure as the PATH vowels of the RP speakers were very close to the PATH vowels of the GenAmE speakers.

Subset_premap				
Term	Estimate	Chi-square	DF	<i>p</i> value
task (map)	40.743	15.26	1	<0.001*
exposure (GenAmE)	29.157	3.37	1	0.066
sex (male)	11.037	0.2664	1	0.606
Subset_prepost				
task (post)	-2.547	0.2111	1	0.646
exposure (GenAmE)	22.189	1.613	1	0.204
sex (male)	1.004	0.0021	1	0.963

Table 6.19 Summary of the model comparisons for the two subsets of PATH vowel. The estimates represent comparisons against a reference value (pre-task for **task**, RP for **exposure**, female for **sex**). The Chi-squares, degree of freedom and p-values are taken from ANOVA results of comparing the nested models with the first model. * indicates a significant effect.

For the PATH vowel, the results for Euclidean distance are different from the results for F1 and F2. In previous sections, the participants were found to have shifted their F1 and F2 towards the GenAmE speakers, however, the results for Euclidean distance in the GenAmE condition suggested a divergence (see Table 6.17). A similar issue was discussed in the section about the THOUGHT vowel, in that Euclidean distance captures the changes in the F1 and F2 dimensions at the same time, whereas F1 and F2 alone only represent a single dimension. Besides this, it is worth noting that the standard deviation of the GenAmE speakers' PATH vowel was large, which is also reflected in Figure 6.18. The spreading of the GenAmE speakers' PATH vowel may be one of the reasons that the results for F1 and F2 were different from the results for Euclidean distance.

Another possibility is over-convergence. The participants might shift their PATH vowels towards the GenAmE speakers; however, they might over-converge so that the map-distances are larger than the pre-distances, resulting in a divergence based on the results of Euclidean distance. This is demonstrated in Figure 6.22. If we compare the F1 and F2 of HKE-pre and HKE-map, the changes indicate a convergence; however, if we compare the pre-distance and the map-distance, as the pre-distance is smaller than the map-distance, it indicates a divergence.

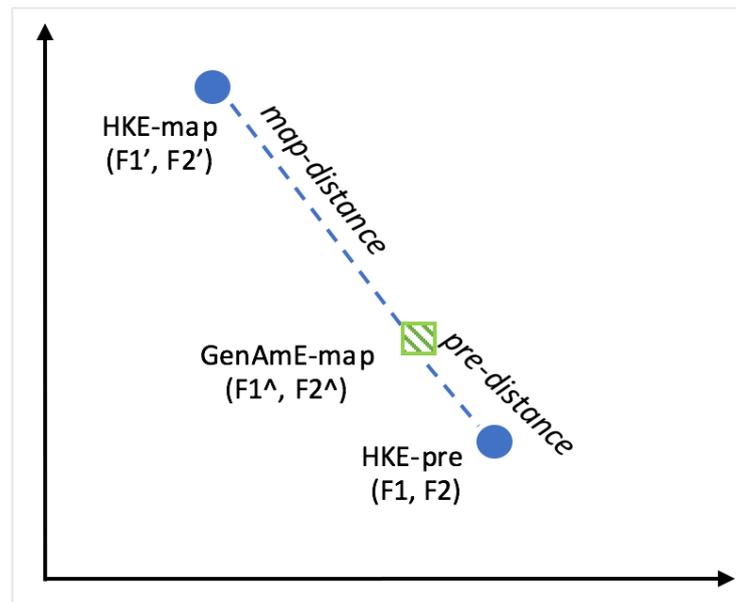


Figure 6.22 An example of over-convergence. The green square represents the GenAmE speaker's PATH vowel, the blue circle on the right bottom corner represents the HKE's PATH vowel in the pre-task, the other blue circle on the left top corner represent his/her PATH vowel in the map task.

Summary

In summary, the results for Euclidean distance for the PATH vowel did not support H2. The map-distances were larger than the pre-distance and the post-distance in the RP condition and the GenAmE condition, suggesting that the participants diverged from the native speakers on the PATH vowel.

Exposure was not a significant predictor for the linear mixed effects model, indicating that exposure to RP or GenAmE did not make a difference to the participants' PATH vowels. This may be due to the RP speakers and the GenAmE speakers of the present study having very similar PATH vowels.

6.4.1.2.3. Accommodation at an individual level

Though the results from previous sections suggest that no convergence was observed for the HKE participants as a group, individual participants might have different patterns of accommodation in the two exposure conditions. To measure the participants' the PATH vowels in relation to the RP/GenAmE speakers, the difference of the two pre-distances (Difference of distance = Pre-distance_{GenAmE condition} – Pre-distance_{RP condition}) is calculated for each participant and the results are plotted in Figure 6.23.

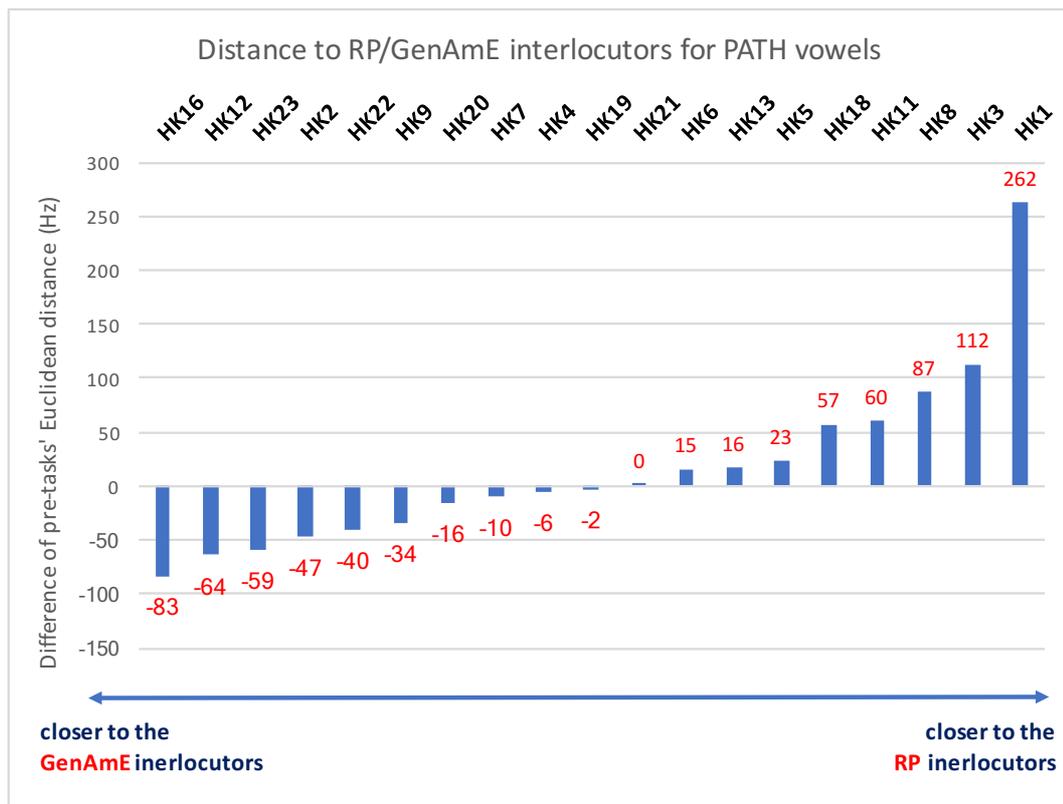


Figure 6.23 The difference of pre-distances between the two exposure conditions for the PATH vowel. The smaller the value is, the closer the distance is to the GenAmE speakers.

In Figure 6.23, a negative value indicates that the participant's PATH vowel is closer to the GenAmE interlocutors' PATH vowel. A positive value, in contrast, indicates that the participant's PATH vowel is closer to the RP interlocutors' PATH vowel.

Based on the order of the participants in Figure 6.23, the participants' Euclidean distances across the tasks and conditions are shown in Figure 6.24. The two patterns found in the THOUGHT vowel are also observed for the PATH vowel.

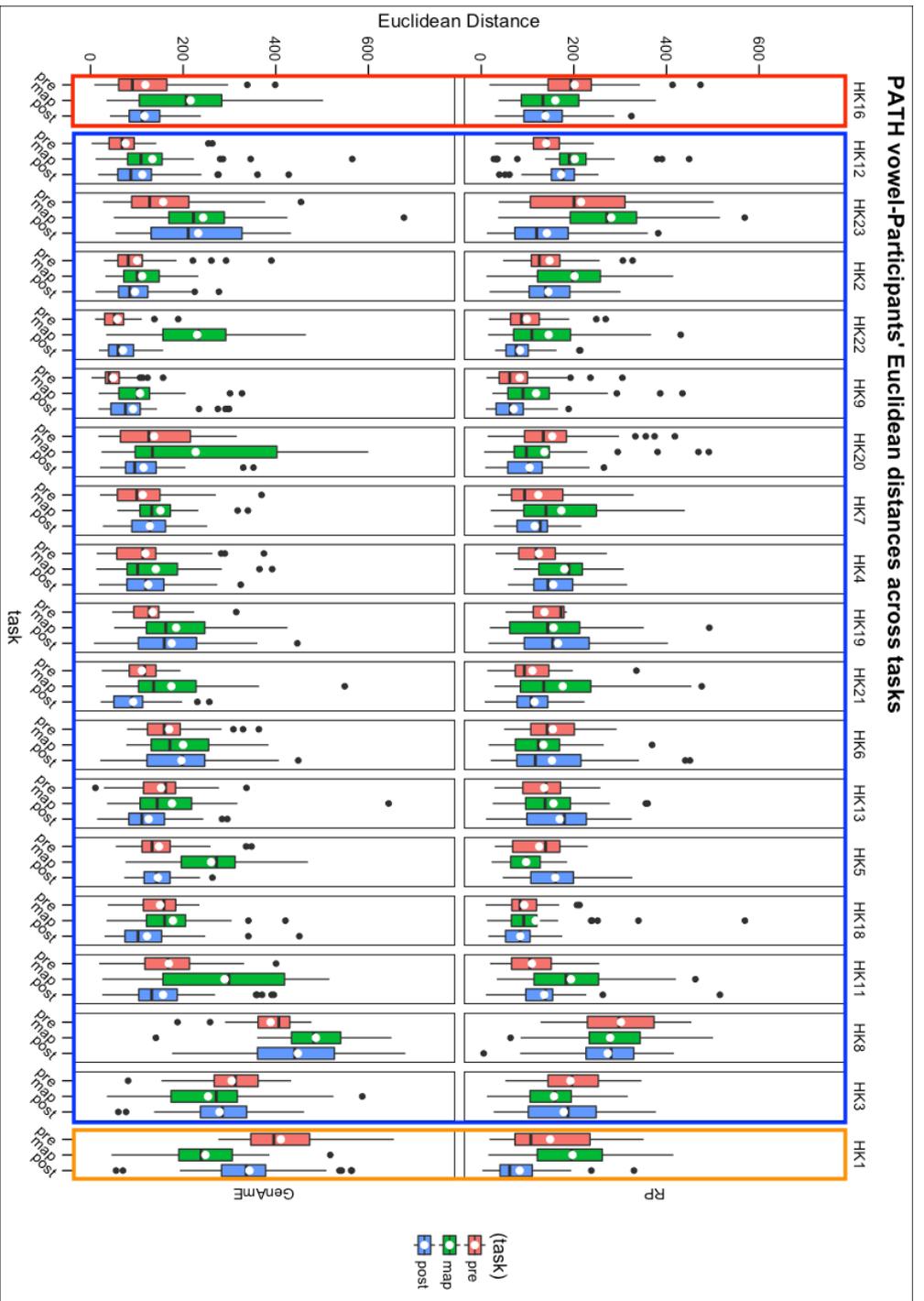


Figure 6.24 Euclidean distance of the PATH vowel across tasks and conditions for each participant. The participant in the red frame converged in the RP condition but diverged in the GenAmE condition; the participants in the blue frame accommodated at the same direction in the RP and GenAmE conditions; the participant in the orange frame converged in the GenAmE condition but diverged in the RP condition. From left to right, the distance between the participant's PATH vowel and the GenAmE interlocutor's PATH vowel decreased.

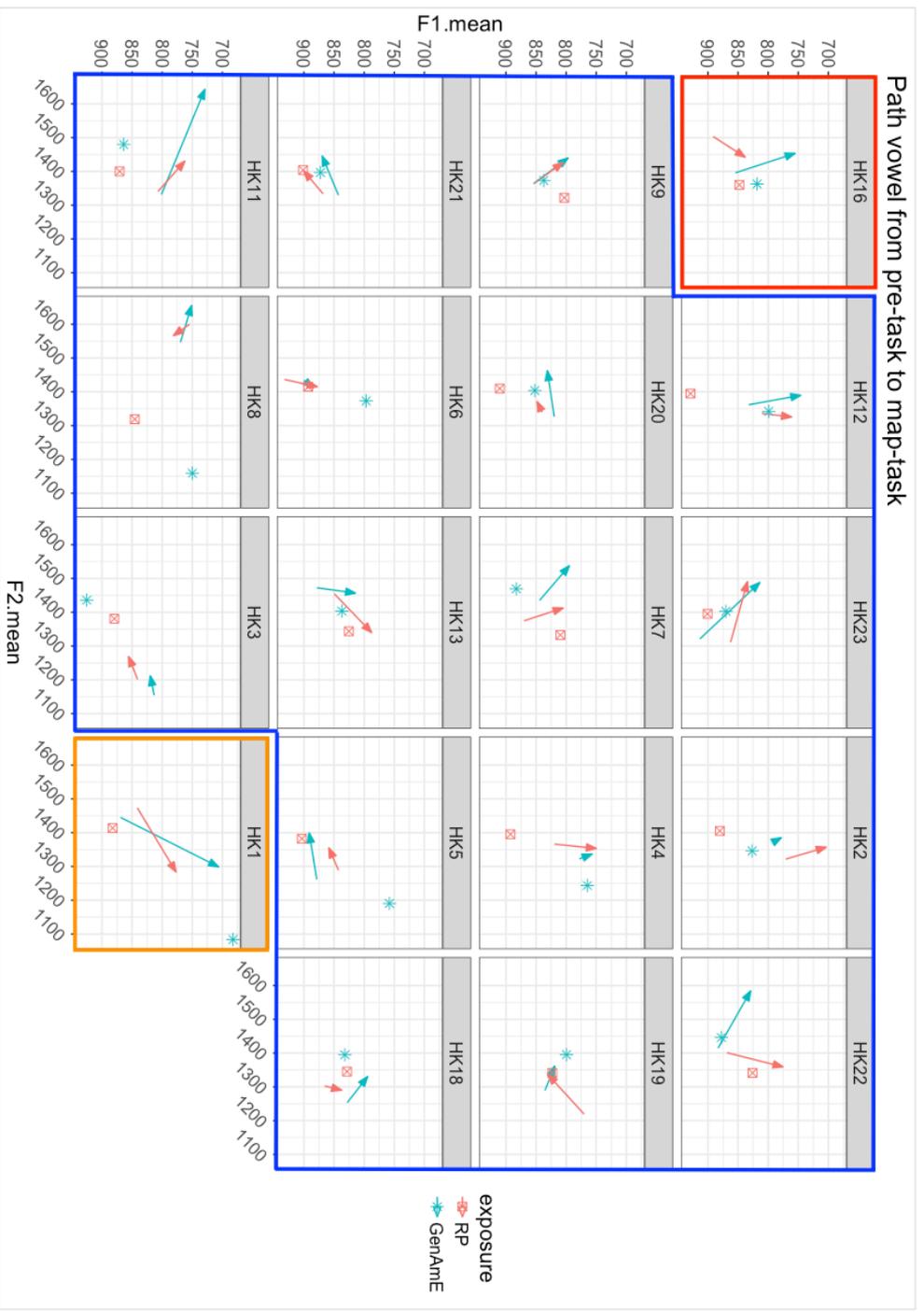


Figure 6.25 Individual shifts of PATH vowels from the pre-task to the map tasks. The start point of an arrow represents the participant's *PATH* vowel in the pre-task, the end of an arrow represents the vowel in the map tasks.

HK16 and HK1, who are placed in the red frame and orange frame in Figure 6.24, converged towards the interlocutors whose PATH vowels were more distant from them. For example, HK16's PATH vowel was further away from the RP interlocutor and close to the GenAmE interlocutor, he showed a convergence in the RP condition. On the other hand, HK1's PATH vowel was further away from the GenAmE interlocutor, he converged in the GenAmE condition.

Apart from these two participants, the rest of the participants did not seem to distinguish their accommodation patterns based on the interlocutors they talked to. Most of the participants in the blue frame (except for HK20, HK6, HK5 and HK8) accommodated at the same direction in the two exposure conditions. If they diverge in the RP condition, they would diverge in the GenAmE condition too, such as HK12, HK23 and HK2. Similarly, if they converge in the RP condition, they would converge in the GenAmE condition, such as HK3.

Both the THOUGHT vowel and the PATH vowel showed two patterns of accommodation at the individual level: (1) if the participant's vowel is distinctly closer to one interlocutor and further away from another, he/she would converge towards the one who is more further away from him/her; (2) if the participant's vowel is not distinctly further away from either of the native interlocutors, he/she would accommodate at the same direction towards the interlocutors. These two patterns indicate that the participants' accommodation might be relevant to the acoustic distances between their vowels and the native interlocutors' vowels.

Note that only 2 participants shown the first pattern on the PATH vowel whereas 9 participants shown the first pattern on the THOUGHT vowel. It could be that most of the HKE participants' PATH vowels were close to the native interlocutors' PATH vowels (as shown in Figure 6.16). The distances between the HKE participants' PATH vowels and the RP/GenAmE interlocutors' PATH vowels were not large enough to elicit a difference on accommodation.

Figure 6.25 shows the shifting of the PATH vowels from the pre-tasks to the map tasks for each participant. The order of the participants in Figure 6.25 is sorted based on the Figure 6.23. In Figure 6.25, each participant has two sets of arrows, i.e. the red arrows

represent the RP condition, and the blue arrows represent the GenAmE condition. The start of each arrow represents the location of the participant’s PATH vowel in the pre-tasks, plotted with the mean of the vowel in the pre-tasks. The end of the arrow represents the location of the vowel in the map tasks, plotted with the mean of the vowel in the map tasks. The locations of the native speakers’ PATH vowels are also provided. Note that a few participants’ shifts seem to support the assumption of over-convergence. For example, HK22’s shift in the GenAmE condition (the blue arrow) indicates an over-convergence, the same as HK11 and HK23’s shifts in the GenAmE condition.

6.4.1.2.4. Summary

This section reported the results for the PATH vowel. H2 predicts that if the participants accommodate towards the native speakers, (1) F1 and F2 values are expected to increase when they talk to the RP speakers; (2) in the GenAmE condition, F1 and F2 values are expected to have little change; and (3) Euclidean distances of the map tasks are expected to be shorter than the distances of the pre-tasks in the RP condition but not in the GenAmE condition. The results are summarized in Table 6.20.

PATH vowel		Pre-task → Map-task		Pre-task → Post-task	
		Results	Match?	Results	Match?
RP condition	F1	decrease (-34)*	-	no change	-
	F2	increase (+3)	YES	decrease (-16)	-
	Euclidean Distance	larger (+24)*	-	smaller (-5)	YES
		Pre-task → Map-task		Pre-task → Post-task	
		Results	Match?	Results	Match?
GenAmE condition	F1	decrease (-40)*	-	no change	YES
	F2	increase (+81)*	-	increase (+8)	-
	Euclidean Distance	larger (+56)*	-	larger (+3)	-
Significant predictor					
Linear mixed effects regressions	F1	task (Chi-square = 17.061, DF=2, $p < .001$)			
	F2	task*exposure (Chi-square = 14.585, DF=2, $p < .001$)			
	Euclidean Distance (Subset_premap)	task (Chi-square = 15.26, DF=1, $p < .001$)			
	Euclidean Distance (Subset_prepost)	task*exposure (<i>not significant</i>) task (<i>not significant</i>)			

Table 6.20 Summary of predictions and actual results for the PATH vowel. Only when the result matches the prediction is a capital “YES” marked in the “Match?” column.

In the RP condition, the participants did not behave as H2 predicted. They significantly diverged in F1 and in the Euclidean distance, suggesting a divergence from the pre-task to the map task. In the GenAmE condition, the participants shifted their F1 and F2 towards the GenAmE speakers, however, the results for Euclidean distance indicated a divergence.

The results of the linear mixed effects models suggest a significant effect of **task** for the Euclidean distance in Subset_premap. In other words, the participants diverged from the native interlocutors from the pre-tasks to the map tasks on the PATH vowels. No significant effect was found for **exposure**, suggesting that the participants did not differ in how they accommodated towards the RP speakers and the GenAmE speakers on the PATH vowels.

6.4.1.3. Native interlocutors' accommodation towards HKE participants

The previous sections report the accommodation of the THOUGHT and the PATH vowel from the HKE participants to the native interlocutors. This section will briefly discuss the accommodation of the two vowels from the native interlocutors towards the HKE participants.

As the native interlocutors only completed the map tasks in the experiments, it is not possible to compare their pronunciation changes between the pre-tasks, map tasks and the post-tasks. To examine whether the native speakers change their pronunciation due to the exposure to HKE, their production of the two vowels were divided into the “early production” and the “late production”. The early production included the vowels the native interlocutors produced during the initial 15 minutes of the map tasks, while the late production included the vowels they produced after 15 minutes of the map tasks. The reason for using 15 minutes as a division point because most of the participants completed the map tasks in about 30 minutes.

Euclidean distances between the native speakers' vowels and their corresponding HKE participants' vowels in the pre-tasks were calculated using the formulas:

$$\text{Early-distance} = \sqrt{(\text{NS early}_{F1} - \text{HKpre}_{F1\text{mean}})^2 + (\text{NS early}_{F2} - \text{HKpre}_{F2\text{mean}})^2}$$

$$\text{Late-distance} = \sqrt{(\text{NS late}_{F1} - \text{HKpre}_{F1\text{mean}})^2 + (\text{NS late}_{F2} - \text{HKpre}_{F2\text{mean}})^2}$$

For the native interlocutors, F1 and F2 value of each token for the vowels were used in the formulas. For the HKE participants, the means of F1 and F2 values of the vowels in the pre-tasks were used in the formulas.

By comparing the Early-distance and the Late-distance, the native speakers' changes during the map tasks can be learnt. If the Late-distance is smaller than the Early-distance, it indicates that a convergence of the native interlocutors towards the HKE interlocutors from the early stage to the late stage in the map tasks. If the Late-distance is larger than the Early-distance, it suggests that a divergence from the early stage to the late stage in the map tasks.

Table 6.21 shows the means and standard deviations of the Early-/Late-distance for each native interlocutor. The descriptive results suggest that RP1 shifted her THOUGHT vowels towards the HKE participants in the map tasks, and both GenAmE1 and GenAmE2 seem to accommodate their PATH vowels towards the HKE participants.

THOUGHT vowel							
NS	Early	Late	Early→Late change	NS	Early	Late	Early→Late change
RP1	205 (89)	194 (92)	smaller -11	GenAmE1	152 (79)	160 (83)	larger +8
RP2	163 (80)	168 (114)	larger +5	GenAmE2	185 (103)	199 (97)	larger +14
PATH vowel							
NS	Early	Late	Early→Late change	NS	Early	Late	Early→Late change
RP1	132 (75)	146 (76)	larger +14	GenAmE1	193 (126)	180 (133)	smaller -13
RP2	109 (77)	128 (100)	larger +19	GenAmE2	224 (176)	190 (154)	smaller -34

Table 6.21 Euclidean distance of the native interlocutors' vowels at the early/late stage of the map tasks. The figures in the columns of "Early→:ate change" refers to the change from the Early-distance to the Late-distance.

Figure 6.26 visualises the results of Table 6.21. Note that for the two GenAmE interlocutors, the standard deviations of the distances for the PATH vowels are large,

suggesting that the distances between the GenAmE interlocutors and the HKE participants are various.

The results here provide some descriptive evidence of accommodation from the native interlocutors towards the HKE participants. As the project's focus is on the accommodation of the HKE speakers, further analysis on the accommodation of the native interlocutors will not be reported here.

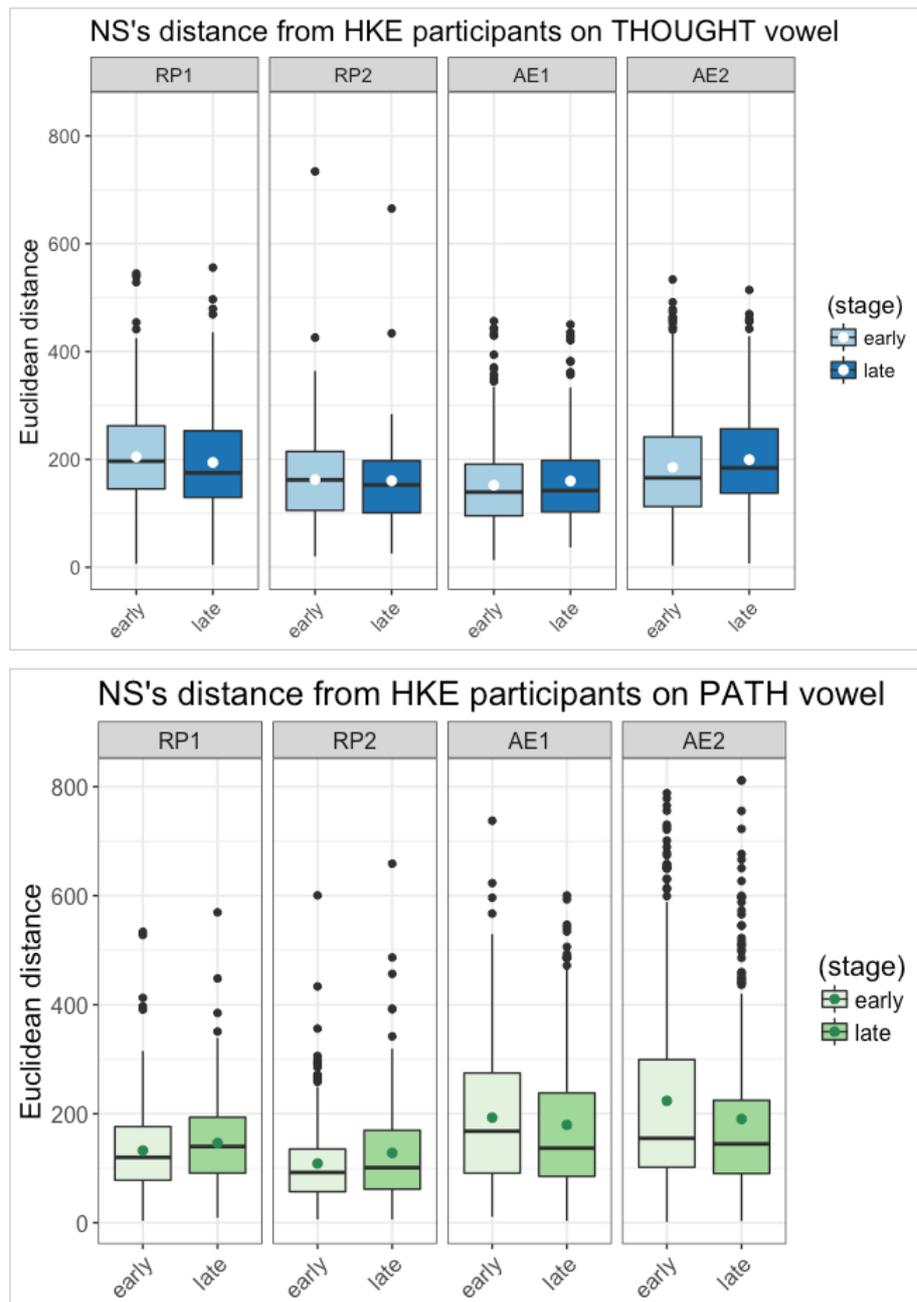


Figure 6.26 Euclidean distance of the THOUGHT/PATH vowels between the native interlocutors and the HKE participants from the early stage to the late stage of the map tasks. The dots inside the bars represent the means, the black lines represent the medians and the black dots represent outliers.

6.4.2. Results for consonants

6.4.2.1. Rhoticity

6.4.2.1.1. Descriptive results

For rhoticity, 9,568 tokens were extracted from the recordings. Percentages of rhotic tokens were calculated. The dependent variable was a binomial variable with two codes: the presence of rhoticity was coded as “yes” and the non-presence of rhoticity was coded as “no”. Therefore, bar charts are chosen to represent the percentage of rhotic tokens across the three tasks and two conditions in Figure 6.27.

H3 predicts that if the participants accommodate to the accents they are exposed to, rhoticity% is expected to be lower in the RP condition because RP is a non-rhotic accent. In the GenAmE condition, rhoticity% is expected to be higher because GenAmE is a rhotic accent.

A small (4.38%) decrease was observed from the pre-task to the map task in the RP condition, while a small (4.75%) increase was found from the pre-task to the map task in the GenAmE condition. The accommodation effects also lasted through to the post-tasks.

The results from Figure 6.27 suggest that H3 was supported.

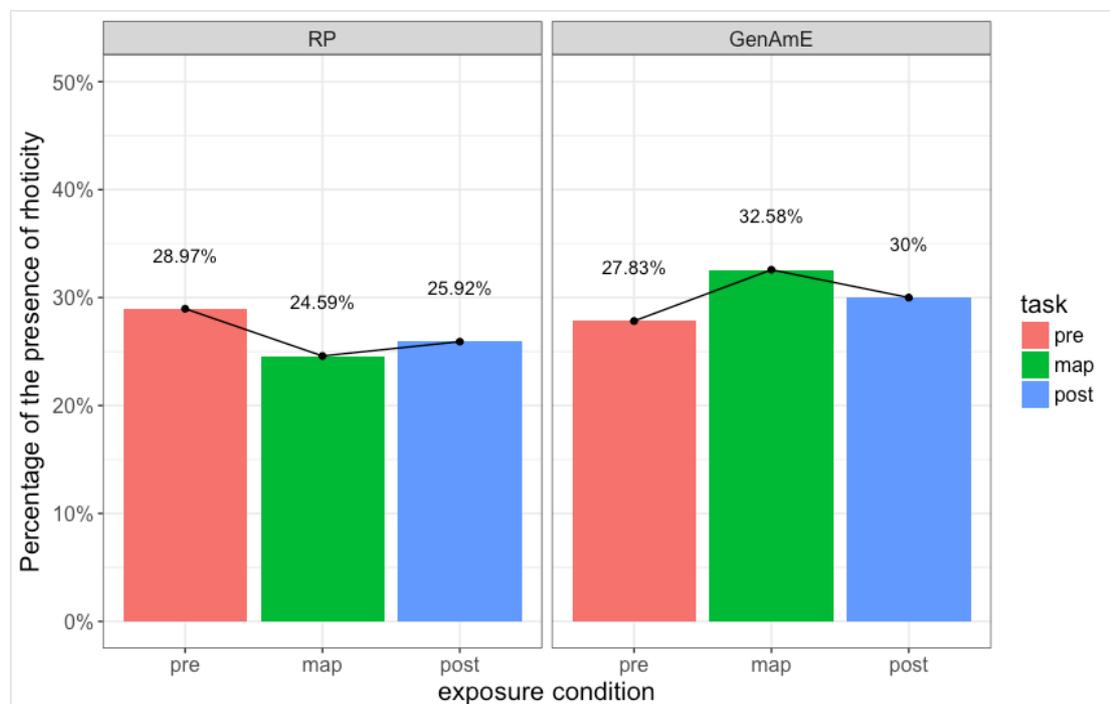


Figure 6.27 Percentage of the presence of rhoticity across the three tasks and two exposure conditions.

6.4.2.1.2. Logistic linear mixed effects regression

For the data of rhoticity, logistic mixed effects regressions were used. The full model contained the same fixed effects and random effects. It also included random intercepts by participant, by word and by interlocutor, and random slopes by participant over task, exposure and the interaction between task and exposure. The dependent variable was a binomial variable with two codes: the presence of rhoticity was coded as “yes” and the non-presence of rhoticity was coded as “no”.

The formula of the full model is:

Full model = rhoticity ~ task*exposure + sex + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

In Table 6.22, the estimate column shows the effect of each independent variable on the log odds of presence of rhoticity. A positive estimate value indicates that the participants in the tested condition produced more rhotic words compared to their performance in the reference condition. For example, when the males (the tested condition) were compared with the females (the reference condition) for **sex**, the positive estimate (1.5034) suggested that the male participants produced more rhotic words than the female participants, but because the *p* value for this comparison was not significant (*p* =.087), the difference of presence of rhoticity between male and female was not statistically significant. In contrast, a negative estimate value suggests that the participants in the tested condition produced less rhotic words compared to their performance in the reference condition.

Term	Estimate	SE	z value	p value
task (map)	0.8997	0.2846	3.166	0.002*
task (post)	0.6296	0.2364	2.664	0.008*
exposure (RP)	0.5634	0.3154	1.786	0.074
sex (male)	1.5034	0.8800	1.708	0.087
task map:exposure RP	-1.2639	0.4139	-3.055	0.002*
task post:exposure RP	-1.1151	0.3669	-3.039	0.002*

Table 6.22 Summary of the logistic mixed effects models for rhoticity. The estimates represent comparisons against a reference value (pre-task for **task**, GenAmE for **exposure**, female for **sex**). * indicates a significant effect.

Note that in Table 6.22, the interaction between the pre-task and the map task (the 5th row in the table) is significant. This indicates that the changes of the percentage of rhoticity

from the pre-tasks to the map tasks in the RP condition (i.e. -4.4%) was significantly different from the changes of which in the GenAmE condition (i.e. +4.8%).

Similarly, the interaction between the pre-task and the post-task (the 6th row in the table) is significant too. It suggests that the changes of the percentage from the pre-tasks to the post-tasks in the RP condition (i.e. -3%) was significantly different from the changes of which in the GenAmE condition (i.e. +2%).

To test the hypothesis of rhoticity, four post-hoc tests of specific contrasts (pre vs map, pre vs post in both the RP and GenAmE conditions) were conducted, with adjusted *p* values in the Bonferroni method. The results of post-hoc tests are shown in Table 6.23.

Post hoc tests for <i>task*exposure</i>			
RP			
pre-map = 0.3642,	SE = 0.2927,	<i>z</i> = 1.244,	<i>p</i> = 0.854
pre-post = 0.4855,	SE = 0.2892,	<i>z</i> = 1.679,	<i>p</i> = 0.372
GenAmE			
pre-map = -0.8997,	SE = 0.2844,	<i>z</i> = -3.164,	<i>p</i> = 0.006*
pre-post = -0.6296,	SE = 0.2367,	<i>z</i> = -2.660,	<i>p</i> = 0.031*

Table 6.23 The post-hoc tests results for task of rhoticity with *p* values adjusted in the Bonferroni method. * indicates a significant effect.

The results suggest that only in the GenAmE condition did the participants significantly produce more rhotic words from the pre-task to the map task, and from the pre-task to the post-task. In the RP condition, the participants produced less rhotic words from the pre-task to the map task, but the effect was not significant.

6.4.2.1.3. Accommodation at an individual level

The previous sections reported the participants' overall accommodation on rhoticity, this section will discuss the participants' accommodation at an individual level. Figure 6.28 shows the participants' percentages of rhoticity in the pre-tasks (also shown in Table 6.6).

Based on their degree of rhoticity, the participants are divided into three groups: (1) **the non-rhotic group** includes the participants whose percentage of rhoticity was less than 10% in their pre-tasks (HK22 to HK12 in Figure 6.28); (2) **the mild-rhotic group** includes the participants who had a percentage of rhoticity between 20% to 50% (HK3 to HK18); and (3) **the heavy-rhotic group** includes the participants whose percentage of

rhoticity was more than 50% (HK5 to HK8). The three groups of the participants are placed in different coloured frames in Figure 6.29.

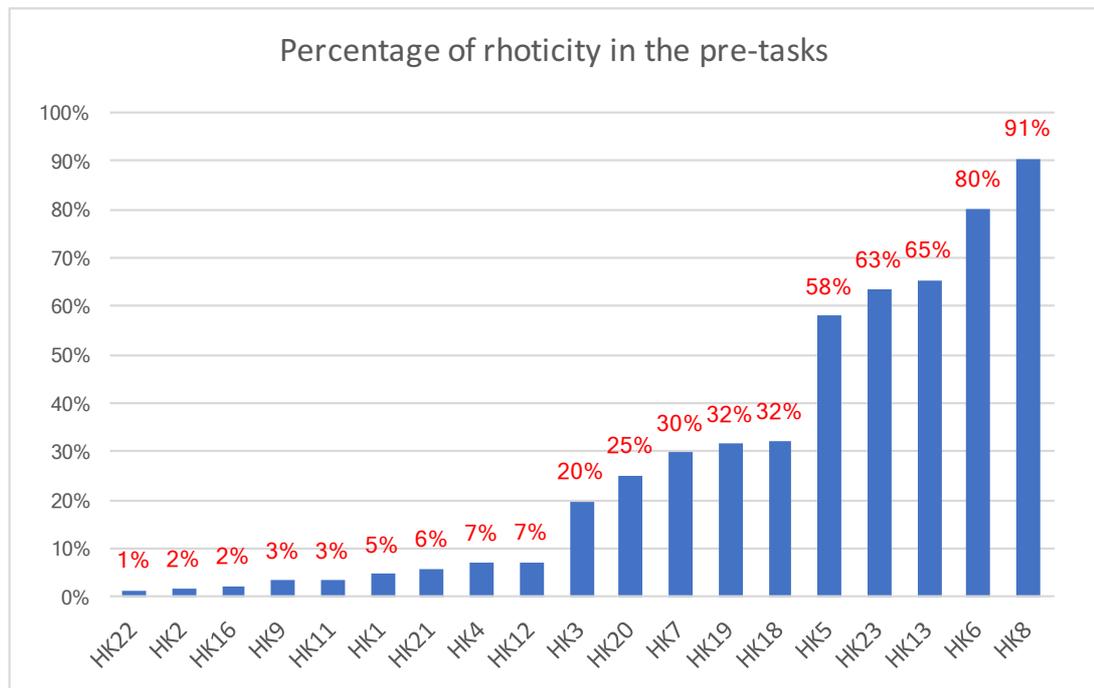


Figure 6.28 The participants' percentage of rhoticity in the pre-tasks. From left to right, the participant's degree of rhoticity increases.

Based on the order of the participants in Figure 6.28, Figure 6.29 shows the individual's accommodation of rhoticity across the tasks and conditions. In overall, almost all the participants converged towards the native interlocutors on rhoticity in both exposure conditions.

For the non-rhotic group in the red frame, most of the participants had a small change of percentage from the pre-tasks to the map tasks in both conditions. As these participants rarely produced any rhotic words in their pre-tasks, they would have little space to reduce their degree of rhoticity for even more when they were exposed to the RP speakers in the map tasks. However, when they were in the GenAmE condition where they have the space to increase the degree of rhoticity, they did not seem to change much. Only HK16 and HK12 increased over 20% of rhoticity in the GenAmE condition. It could be that most of the participants in the non-rhotic group might not have associated rhoticity with words in their production. Even though they receive input of rhotic words from the GenAmE speakers in the map tasks, the lack of the associations between rhoticity and the words might constrain their production.

For the mild-rhotic group in the blue frame, the participants might have started establishing the category of rhoticity, as they already produced some rhotic words in their pre-tasks. Unlike the non-rhotic group who can only converge from non-rhotic to rhotic, the mild-rhotic group can converge on both directions, yet, not all the participants showed a convergence in both conditions. For example, HK3 and HK19 only converged in the RP condition, and HK7 converged in the GenAmE condition. HK20 and HK18, on the other hand, converged in both conditions.

For the heavy-rhotic group in the green frame, all the participants showed a convergence in both conditions. Even for HK8 who had 91% of rhoticity in his pre-tasks, still changed his rhoticity slightly ($\pm 4\%$) according to the accents he was exposed to. As all these participants had over 50% of rhoticity in their pre-tasks, theoretically, they should have more space to change their rhoticity in the RP condition. HK5 and HK23 showed this pattern clearly. They decreased 16% and 33% of rhoticity in the RP condition, but they only increased 1% and 2% of rhoticity in the GenAmE condition. HK13 and HK6 also had a larger change of percentage in the RP condition than they had in the GenAmE condition, however, the differences of changes between the two conditions are not very huge.

6.4.2.1.4. Summary

The results suggest that the participants converged toward the RP speakers and the GenAmE speakers on rhoticity, though significant changes were only found in the GenAmE condition from the pre-tasks to the map tasks and from the pre-tasks to the post-tasks.

At the individual level, for the participants who had a heavy-rhotic accent, they converged in both conditions. For the participants who had a mild-rhotic accent, most of them showed a convergence in the RP condition. For the participants who rarely produced rhotic words, they did not change much in both conditions.

In sum, the results support H4 and indicate that HKE speakers show comparatively strong accommodation on rhoticity.

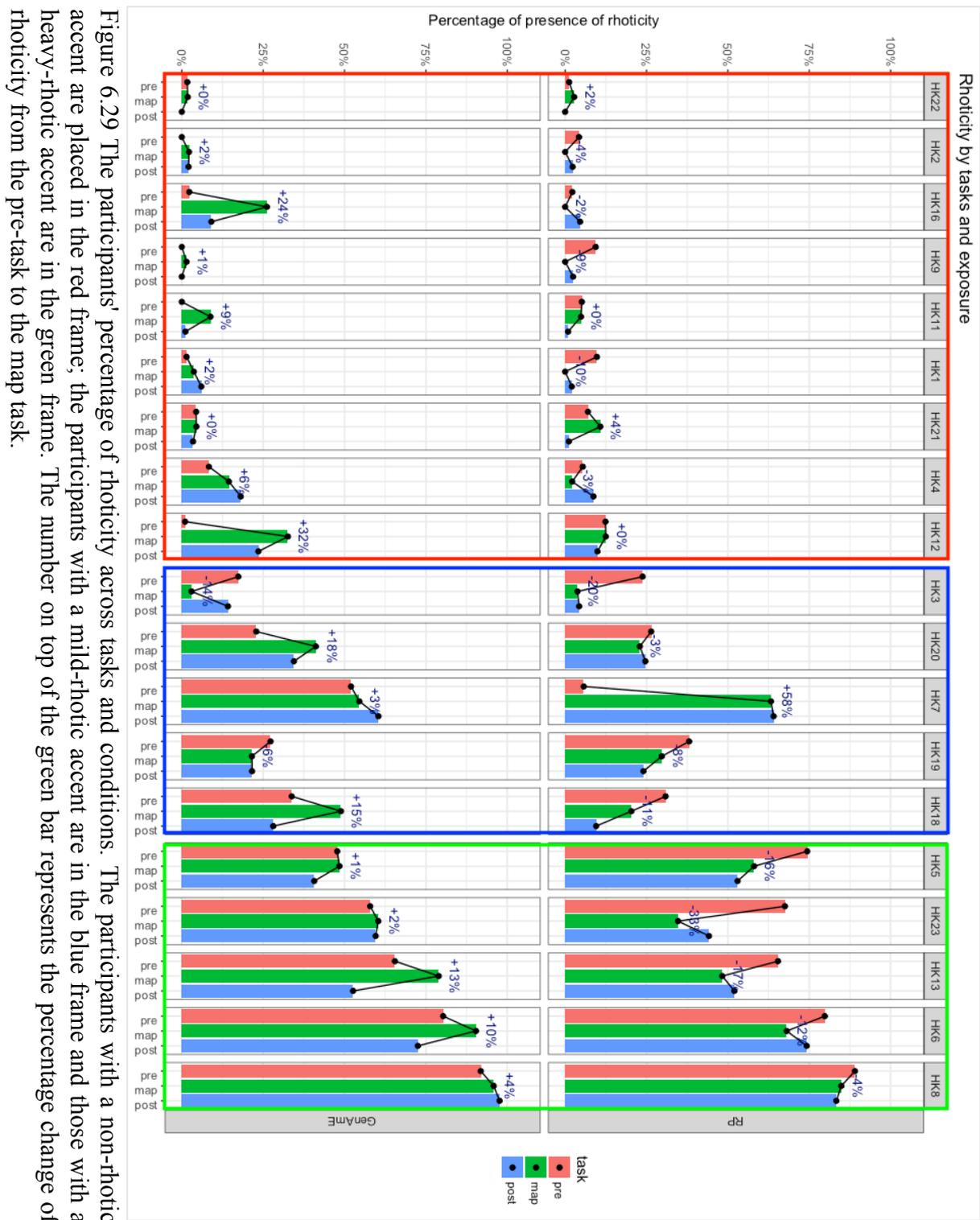


Figure 6.29 The participants' percentage of rhoticity across tasks and conditions. The participants with a non-rhotic accent are placed in the red frame; the participants with a mild-rhotic accent are in the blue frame and those with a heavy-rhotic accent are in the green frame. The number on top of the green bar represents the percentage change of rhoticity from the pre-task to the map task.

6.4.2.2. Fricative [z]

6.4.2.2.1. Descriptive results

For the voiced fricative /z/, 7,084 tokens were extracted from the recordings. The dependent variable was a binomial variable in which fricative [z] was coded as “yes” and other variants were coded as “no”. The percentages of the tokens of fricative [z] across three tasks are shown as the bar charts in Figure 6.30.

H4 predicts that if accommodation occurs, the participants will produce more fricative [z] in the map tasks and the post-tasks compared to the pre-tasks.

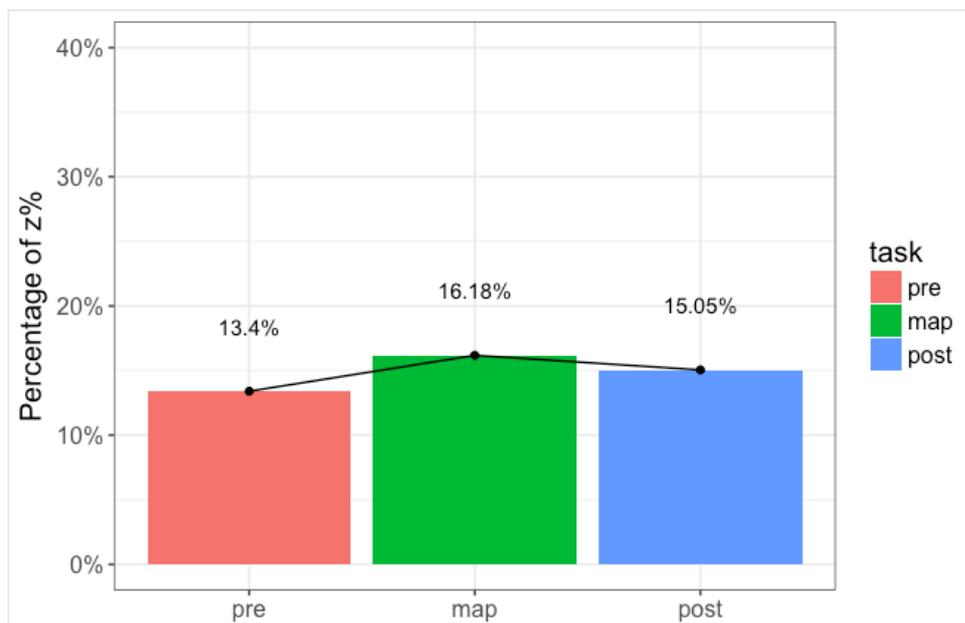


Figure 6.30 Percentage of production of fricative [z] across the three tasks.

The results suggest that only around 13% of the fricatives were pronounced as voiced [z]. [s] was a common variant, and a small number of [ʃ] tokens were also found. Figure 6.30 suggests a small increase of z% was found from the pre-tasks to the map tasks (2.78%), as well as from the pre-tasks to the post-tasks (1.65%). This result supports H4.

6.4.2.2.2. Logistic mixed effects regression

For the fricative [z] data, logistic mixed effects regressions were used. Different from the previous sections, the full model of fricative [z] only contained **task** and **sex** as the fixed effects; **exposure** was excluded because fricative [z] has the same realization in RP and GenAmE. The model included random intercepts by participant, by word and by

interlocutor, and random slopes by participant. The dependent variable was a binomial variable in which fricative [z] was coded as “yes” and other variants were coded as “no”.

The formula for the full model is shown below:

Full model = fricative [z] ~ task + sex + (task | participant) + (1 | word) + (1 | interlocutor)

In Table 6.24, the estimate column shows the effect of each independent variable on the log odds of the presence of fricative [z]. A positive estimate value indicates that the participants in the tested condition produced more fricative [z] compared to their performance in the reference condition. In contrast, a negative estimate value indicates that the participants in the tested condition produced less fricative [z] than in the reference condition.

The results suggested a significant effect between the pre-tasks and the map tasks. No significance was found between the pre-tasks and the post-tasks. **These results suggest that the participants produced significantly more fricative [z] in the map tasks compared to the pre-tasks.** This result supports H4.

Term	Estimate	SE	z value	p value
task (map)	1.8296	0.5472	3.343	<.001*
task (post)	0.2851	0.5213	0.547	0.584
sex (male)	1.3864	0.9724	1.503	0.133

Table 6.24 Summary of the logistic mixed effects models for fricative [z]. The estimates represent comparisons against a reference value (pre-task for **task**, female for **sex**). * indicates a significant effect.

6.4.2.2.3. Accommodation at an individual level

This section will discuss the participants’ accommodation on fricative [z] at an individual level. The participants’ percentages of fricative [z] in the pre-tasks are shown in Figure 6.31. The variation of the participants’ production of fricative [z] in the pre-tasks is very large. Some participants did not produce any fricative [z] in their pre-tasks, such as HK6 and HK19, while some participants produced around 50% of the fricative [z], such as HK12 and HK18.

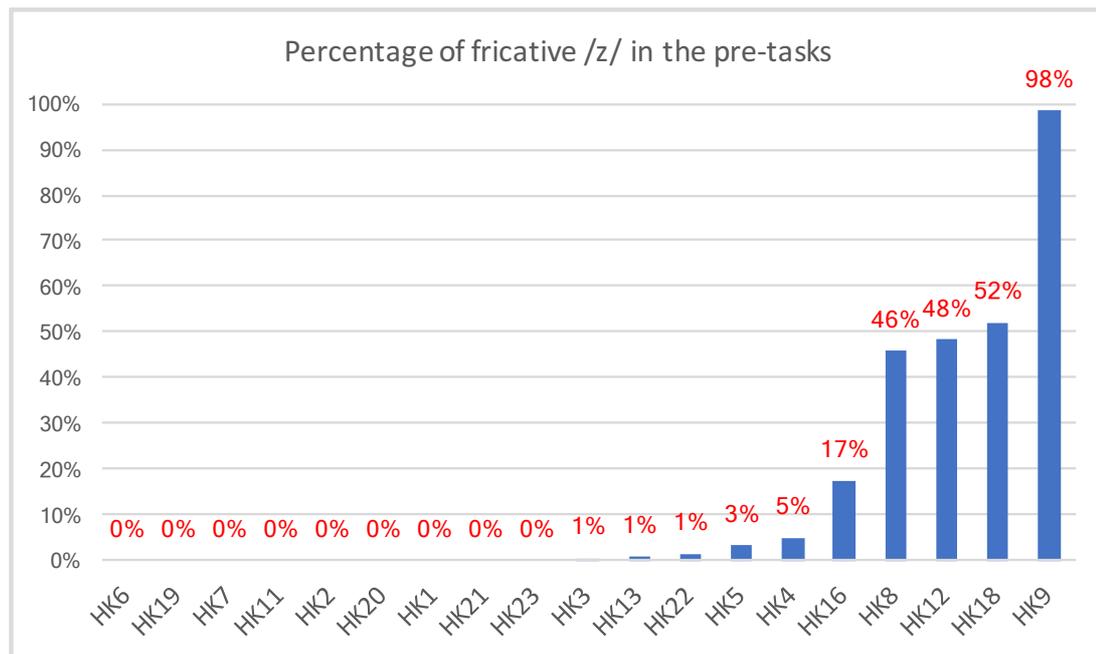


Figure 6.31 The participants' percentage of fricative [z] in their pre-tasks. From left to right, the percentage of fricative [z] increases.

Based on the participants' percentages of fricative [z], they are divided into three groups: (1) **the [s]-dominant group** includes the participants whose percentage of fricative [z] was less than 5% (HK6 to HK4); (2) **the mixed group** includes the participants whose production of the fricative contained both [z] and [s]; for this group, the participants' percentage of fricative [z] was between 15% and 55% (HK16 to HK18); and (3) **the [z]-dominant group** includes HK9 whose percentage of fricative [z] was almost 100%.

According to the order of the participants in Figure 6.31, Figure 6.32 shows the participants' accommodation of fricative [z] across the three tasks. The three groups are placed in different coloured frames in Figure 6.32.

For the [s]-dominant group, the participants rarely produced any fricative [z] in their pre-tasks. Some participants such as HK6, HK19 and HK7 retained 0% of fricative [z] across the three tasks. For these participants, they might not have established the [z]-category in their production at all. The amount of input they received from the native interlocutors in the map tasks might not be enough for them to establish a new category. Therefore, even though they were exposed to the native speakers in the map tasks, they still retained fricative [s] for the fricative. Other participants in the group showed a small increase of percentage from the pre-tasks to the map tasks. These participants might start acquiring

the category of fricative [z] through the input, however, as the dominant category was still fricative [s], their production did not change too much.

For the mixed group, most of the participants showed a convergence of fricative [z] from the pre-tasks to the map tasks, except for HK16. For these participants, they might have both the [s]-category and [z]-category in their production. When they were exposed to the native input in the map tasks, the probability of [z]-category being selected would be increased.

For the [z]-dominant group, HK9 showed a divergence from the pre-tasks to the map tasks. This might be due to the ceiling effect. HK9's percentage of fricative [z] in the pre-tasks was 98%, therefore he had little space to improve. In this case, fricative [s] might be a resource for him to express sociolinguistic meanings, such as using the fricative [s] as an L1-variant to express his Hong Kong identity.

6.4.2.2.4. Summary

This section reported the results for fricative [z]. The participants produced more fricative [z] when they talked to the native speakers in the map tasks, indicating a convergence. This result supports H4.

At the individual level, the participants who had fricative [s] as the dominant category for the fricative showed a small increase of percentage from the pre-tasks to the map tasks, suggesting that they might be learning the new category (fricative [z]) during the map tasks. The participants who had both the [z]-category and [s]-category showed a convergence, and the participant who had fricative [z] as the dominant category showed a ceiling effect.

Fricative [z]'s performance by tasks

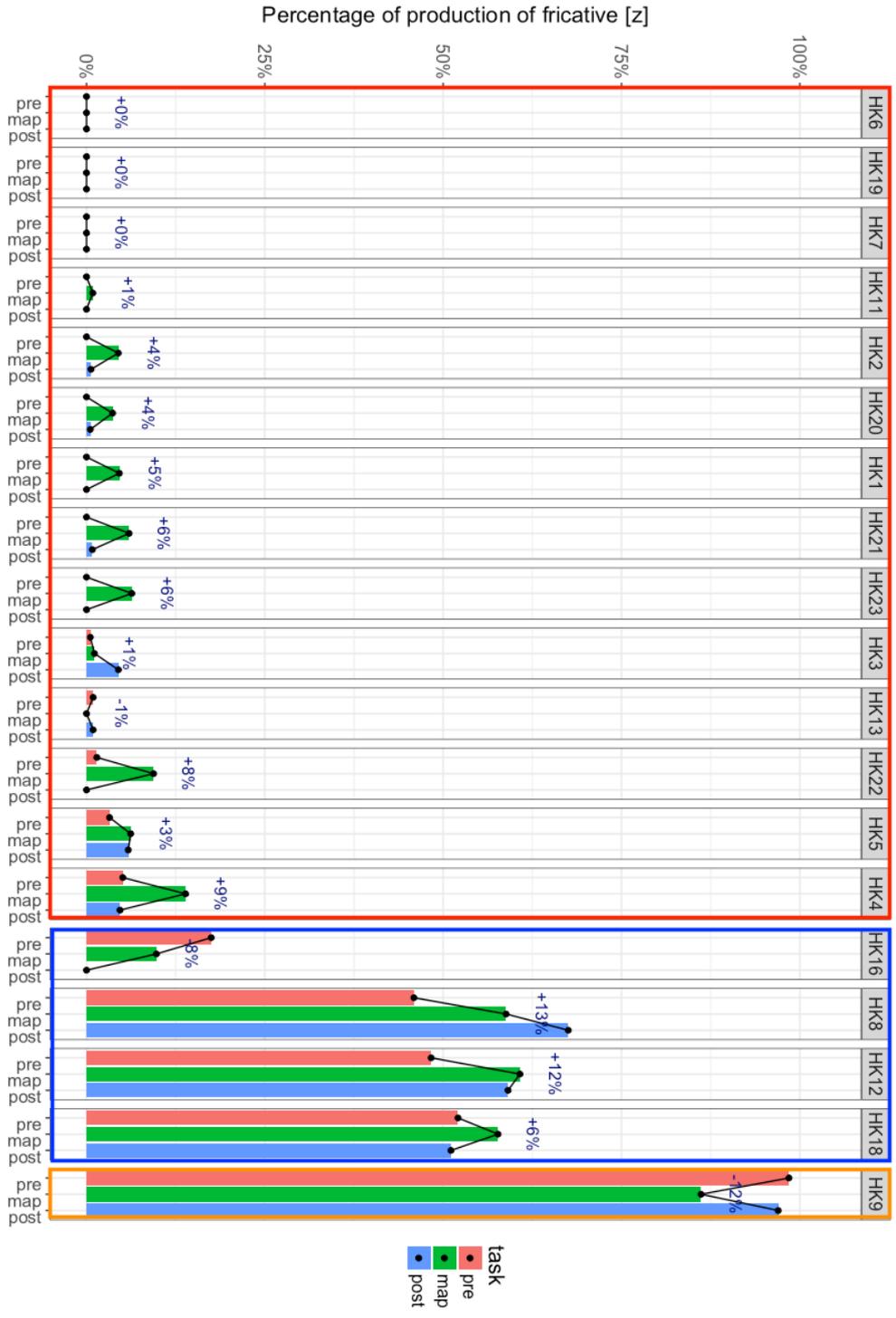


Figure 6.32 Individual percentages of the production of fricative [z] across the three tasks. Participants in the red frame have fricative [s] as the dominant category; those in the blue frame have both the [z]-category and [s]-category; and those in the orange frame have fricative [z] as the dominant category. The number on top of the green bar represents the percentage change from the pre-task to the map task.

6.4.2.3. Fricative [θ]

6.4.2.3.1. Descriptive results

For fricative [θ], a total of 6,730 tokens were extracted from the recordings. Only words with [θ] in word-initial position were included. There were two common variants of [θ] produced by the participants: [f] and [t]. [f] was usually seen in the words *three*, *thirty* and *thought*, while [t] was used in the word *thousand*. There were also a small number of [s] and [ʃ] tokens in the data.

H4 predicts that if accommodation occurs, the participants will produce more fricative [θ] in the map tasks and the post-tasks compared to the pre-tasks. The results shown in Figure 6.33 do not support this hypothesis.

Figure 6.33 shows the percentages of accurate pronunciation of fricative [θ] across the three tasks. Overall, the participants had good performance for [θ]-production, with around 66% in all tasks, which is almost identical to the 64% mentioned in Deterding et al. (2001). The changes across the tasks were very limited. A small decrease in production of fricative [θ] (-2.61%) was found from the pre-tasks to the map tasks. From the pre-tasks to the post-tasks, less than 1% increase was found.

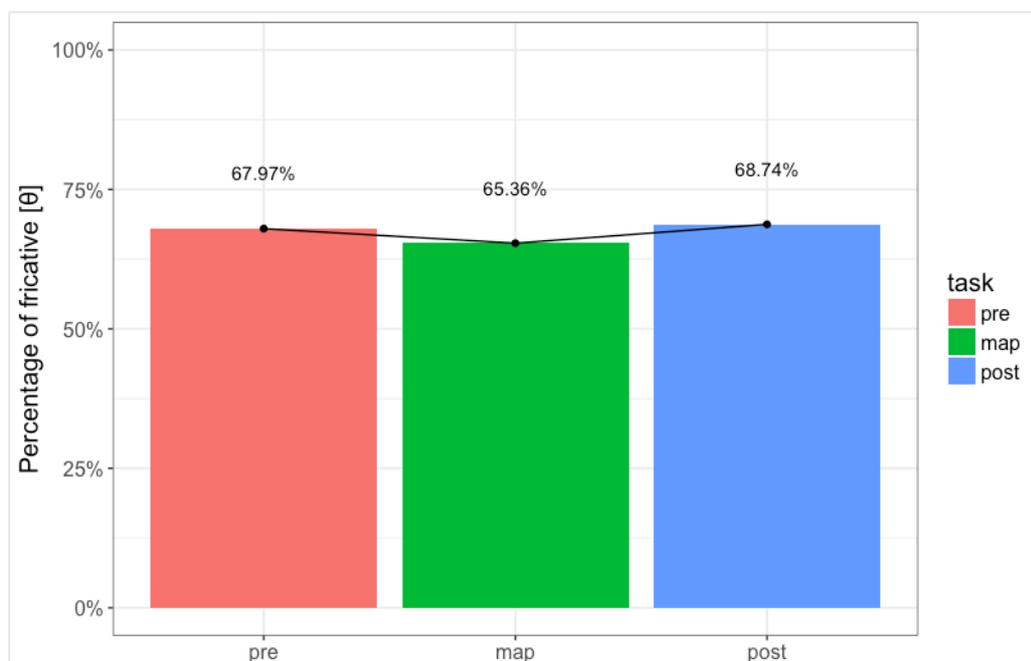


Figure 6.33 Percentages of accurate pronunciation of fricative [θ] across the three tasks.

6.4.2.3.2. Logistic mixed effects regression

For the fricative [θ] data, logistic mixed effects regressions were used. The full model of fricative [θ] contained **task** and **sex** as the fixed effects, and included random intercepts by participant, by word and by interlocutor, and random slopes by participant. The dependent variable was a binomial variable in which a realization of fricative [θ] was coded as “yes” and the other variants were coded as “no”.

The formula for the full model is shown below:

Full model = Fricative [θ] ~ task + sex + (task | participant) + (1 | word) + (1 | interlocutor)

In Table 6.25, the estimate column shows the effect of each independent variable on the log odds of presence of fricative [θ].

Term	Estimate	SE	z value	p value
task (map)	-0.3318	0.1726	-1.923	0.055 ·
task (post)	0.0029	0.1595	0.019	0.985
sex (male)	-1.5442	1.9501	-0.792	0.428

Table 6.25 Summary of the logistic mixed effects models for fricative [θ]. The estimates represent comparisons against a reference value (pre-task for **task**, female for **sex**). · indicates a marginally significant effect.

A positive estimate value indicates that the participants in the tested condition produced more fricative [θ] compared to their performance in the reference condition. In contrast, a negative estimate value indicates that the participants in the tested condition produced less fricative [θ] than in the reference condition.

As shown in Table 6.25, a marginally significant effect was found between the pre-tasks and map tasks, suggesting that **the participants produced less fricative [θ] in the map tasks compared to their performance in the pre-tasks**. Note that no significant effect was found between the pre-tasks and the post-tasks. These results do not support H4.

6.4.2.3.3. Accommodation at an individual level

This section will discuss the participants’ accommodation of fricative [θ] at the individual level. The participants’ percentages of fricative [θ] in the pre-tasks are shown in Figure 6.34. Like rhoticity and fricative [z], the participants are divided into three groups based on their production of fricative [θ] in the pre-tasks.

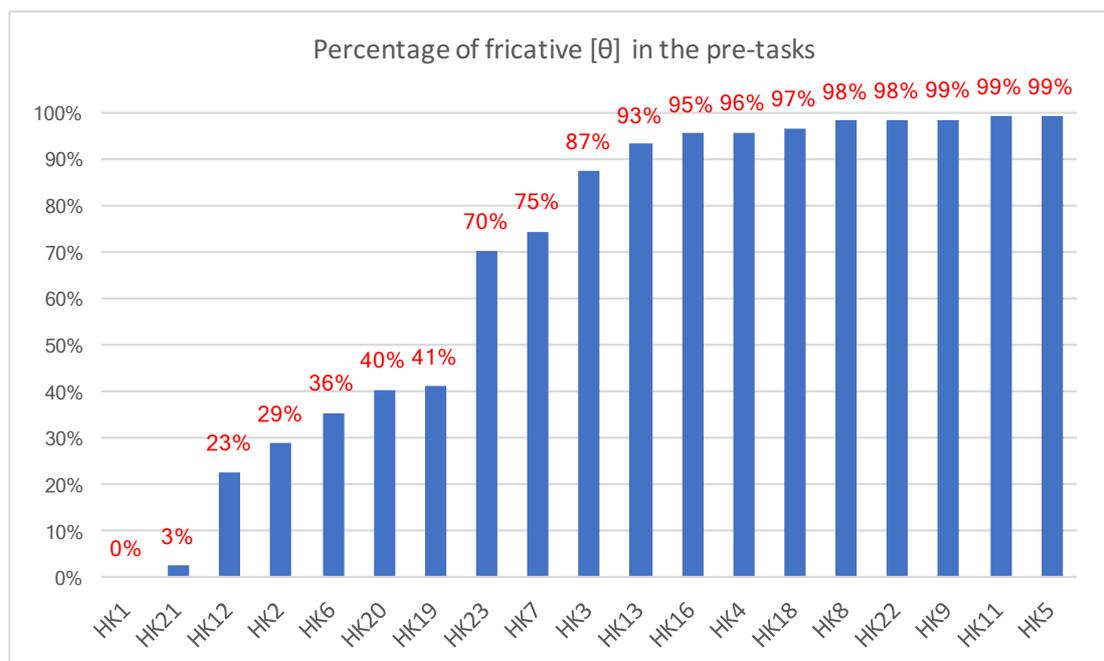


Figure 6.34 The participants' percentage of fricative [θ] in the pre-tasks.

The [f]-dominant group contains HK1 and HK21, whose percentages of fricative [θ] are below 5%. **The mixed group** contains the participants whose percentage was ranged from 20% to 50% (HK12 to HK19). **The [θ]-dominant group** includes the participants whose percentage was higher than 60% (HK23 to HK5).

Based on the order of the participants in Figure 6.34, Figure 6.35 shows their percentage of fricative [θ] across three tasks. The three groups are placed in different coloured frames in Figure 6.35.

As shown in Figure 6.35, most of the participants showed a divergence on fricative [θ] from the pre-tasks to the map tasks. For the [f]-dominant group, HK1 retained 0% of fricative [θ] across three tasks, suggesting that he might not have established the [θ]-category in his production. Although a convergence was found on HK21, the change was very small.

For the mixed group, the participants might have both the [f]-category and [θ]-category for the fricative. HK12 and HK2 showed a divergence while the rest of the group showed a convergence.

For the [θ]-dominant group, all the participants had over 60% of fricative [θ] in their pre-tasks. A few participants in this group showed a ceiling effect in the map tasks, as they produced over than 90% of fricative [θ] in their pre-tasks. In other words, they would have little space to increase their production of fricative [θ] in the map tasks. HK23, HK7 and HK3 who would still have space to improve, unexpectedly, produce less fricative [θ] in the map tasks.

It seems that most of the participants produced less fricative [θ] from the pre-tasks to the map tasks regardless of their baselines in the pre-tasks. Unlike rhoticity and fricative [z], not many participants showed a convergence on fricative [θ]. This might be due to the articulatory difficulty of fricative [θ] for HKE participants, whose L1-Cantonese does not have any interdental sound.

6.4.2.3.4. Summary

This section reported the results for fricative [θ]. The participants showed a divergence from the pre-tasks to the map tasks, producing less fricative [θ] when they talked to the native speakers. This result does not support H4. At the individual level, most of the participants showed a divergence on fricative, regardless of their baselines in the pre-tasks.

Fricative [θ]'s performance by tasks

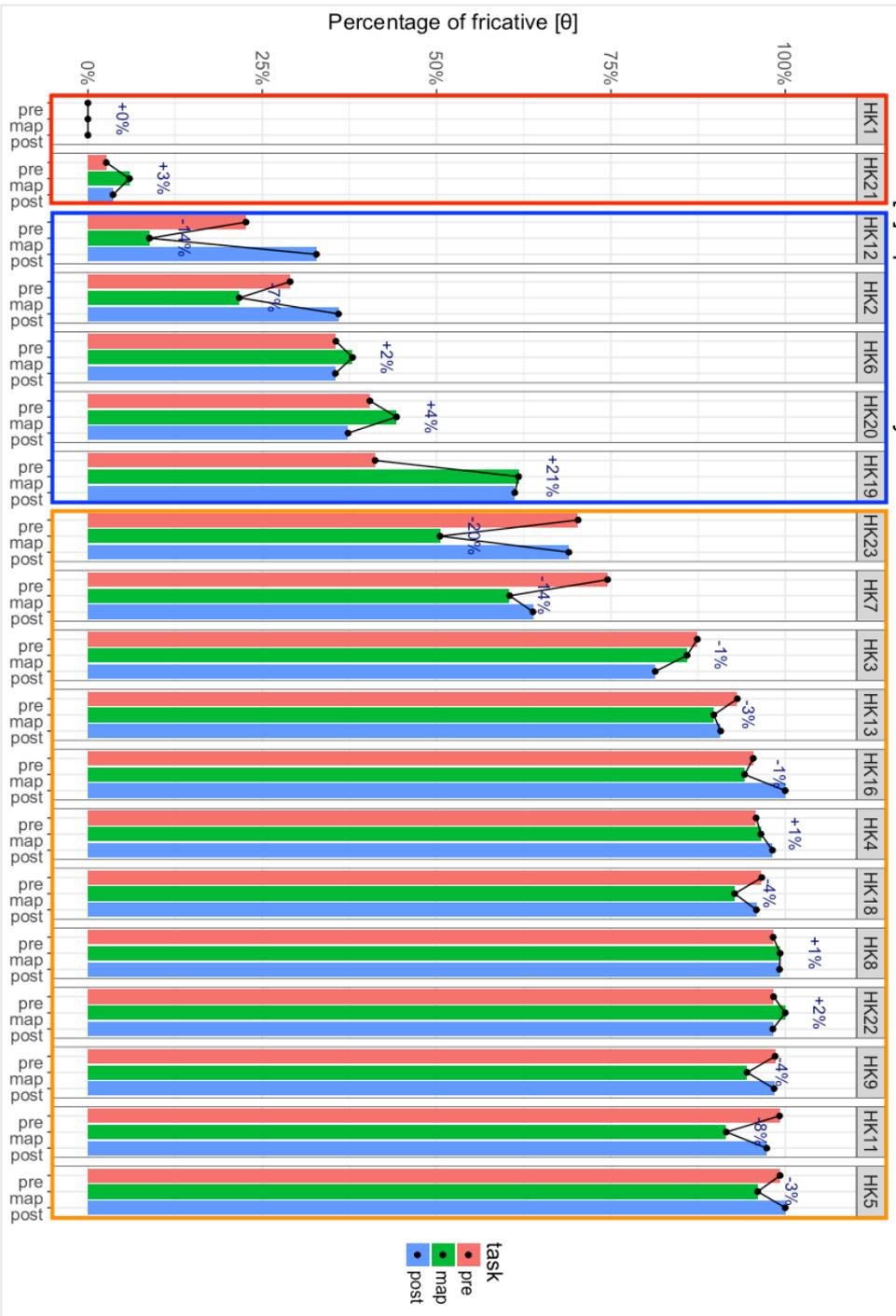


Figure 6.35 Individual percentages of the production of fricative [θ] across the three tasks. Participants who had fricative [f] as the dominant category are placed in the red frame; those who had both [f]-category and [θ]-category are placed in the blue frame; and the individual who had ceiling effect on fricative [θ] are grouped in the orange frame. The number on top of the green bar represents the percentage change from the pre-task to the map task.

6.4.3. The effect of talker sex

Since talker sex was found to be a significant factor in convergence in previous literature (Babel, 2012; Pardo, 2006), it is worth examining the effect of sex in the present study. Note that in the present study, there are 7 male participants and 12 female participants, and all the native interlocutors are females. In other words, there are 7 mixed-sex pairs and 12 same-sex pairs. Therefore, the effect of sex in the present study should be regarded as a comparison of convergence between mixed-sex pairs and same-sex pairs.

Though sex was one of the fixed factors in the linear/logit mixed effects models and some results of sex alone were reported above, the interaction of task and sex was not reported. This section mainly focuses on the interaction effect of task and sex, aiming to explore whether the HKE speakers' sex affects the degree of convergence in the map tasks.

6.4.3.1. Vowels

A series of linear mixed effects regressions were run for the THOUGHT and the PATH vowels. Dependent variables were F1, F2 values and Euclidean distance of the two vowels. The full model for sex contained **task**, **exposure**, **sex** and **interaction of task and sex** as the fixed factors. It also included random intercepts by participant, by word and by interlocutor, and random slopes by participant over task, exposure and the interaction between task and exposure. Comparison was carried out between the full model and a nested model for sex with **task*sex** removed. ANOVA was used to run the comparisons and in total six comparisons were run.

The formulas of the full model and nested model for sex are shown below:

Full model for sex = dependent variable ~ task*sex + exposure + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

Nested model for sex = dependent variable ~ task + sex + exposure + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

For the six dependent variables, the comparison results of the ANOVAs including the Chi-square, DF and p value are shown in Table 6.26. As shown in Table 6.26, a marginally significant effect of **task*sex** was found for F1 of the THOUGHT vowel and F2 of the PATH vowel only. No significance was found for the other dependent variables.

<i>THOUGHT</i>	Chi-square	DF	<i>p</i> value
F1	5.9062	2	0.052
F2	1.9956	2	0.369
Euclidean distance	1.1803	2	0.554
<i>PATH</i>	Chi-square	DF	<i>p</i> value
F1	1.0068	2	0.605
F2	5.7076	2	0.058
Euclidean distance	3.6585	2	0.160

Table 6.26 Summary of the model comparisons for task*sex. The Chi-squares, degree of freedom and p-values are taken from ANOVA results of comparing the nested models with the first model for sex. “.” indicates a marginal effect of **task*sex** for this variable.

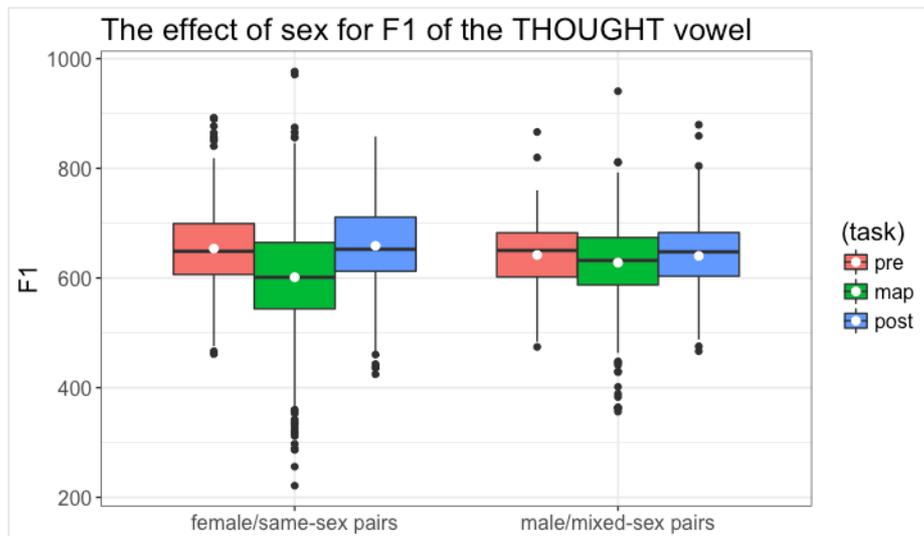


Figure 6.36 Boxplot of F1 of the THOUGHT vowel across tasks for the females and the males.

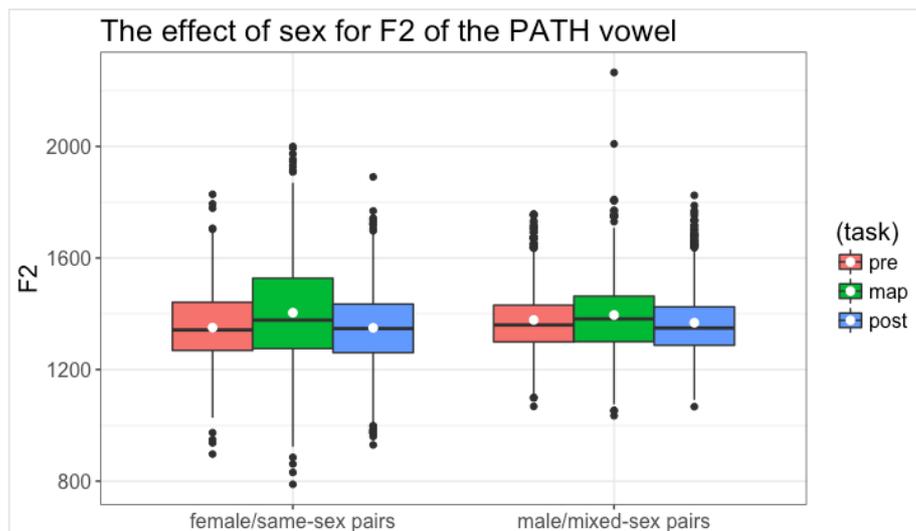


Figure 6.37 Boxplot of F2 of the PATH vowel across tasks for the females and the males.

For F1 of the THOUGHT vowel, when comparing the change from the pre-tasks to the map-tasks, the female participants (i.e. same-sex pairs) made more changes (female-male = 32.274 Hz) than the male participants (i.e. mixed-sex pairs).

For F2 of the PATH vowel, when comparing the change from the pre-tasks to the map-tasks, the female participants made more changes than the male participants (female – male = 77.894Hz).

6.4.3.2. Rhoticity

For rhoticity, logistic mixed effects regressions were used. The full model for sex contained **task, exposure, sex** and **three-way interaction of task, exposure and sex** as the fixed factors. It also included random intercepts by participant, by word and by interlocutor, and random slopes by participant over task, exposure and the interaction between task and exposure. The dependent variable was a binomial variable with two codes: the presence of rhoticity was coded as “yes” and the non-presence of rhoticity was coded as “no”.

Comparison was carried out between the full model and a nested model for sex with **the three-way interaction of task, exposure and sex** removed.

Full model for sex = rhoticity ~ task * exposure + sex + task: exposure: sex + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

Nested model for sex = rhoticity~ task*exposure + sex + (task*exposure | participant) + (1 | word) + (1 | interlocutor)

The results for comparison suggested a non-significant effect of **task: exposure: sex (Chi-square = 10.367, DF = 5, p =.065)**, indicating that the same-sex pairs (i.e. female participants) converged the same as the mixed-sex pairs (i.e. male participants) on rhoticity when they talked to the RP speakers and the GenAmE speakers.

Figure 6.38 shows the percentage of presence of rhoticity for the female and male participants across tasks and conditions. The female participants tended to converge more

than the male participants in the RP condition. When they talked to the GenAmE speakers, the male participants tended to converge more than the females.

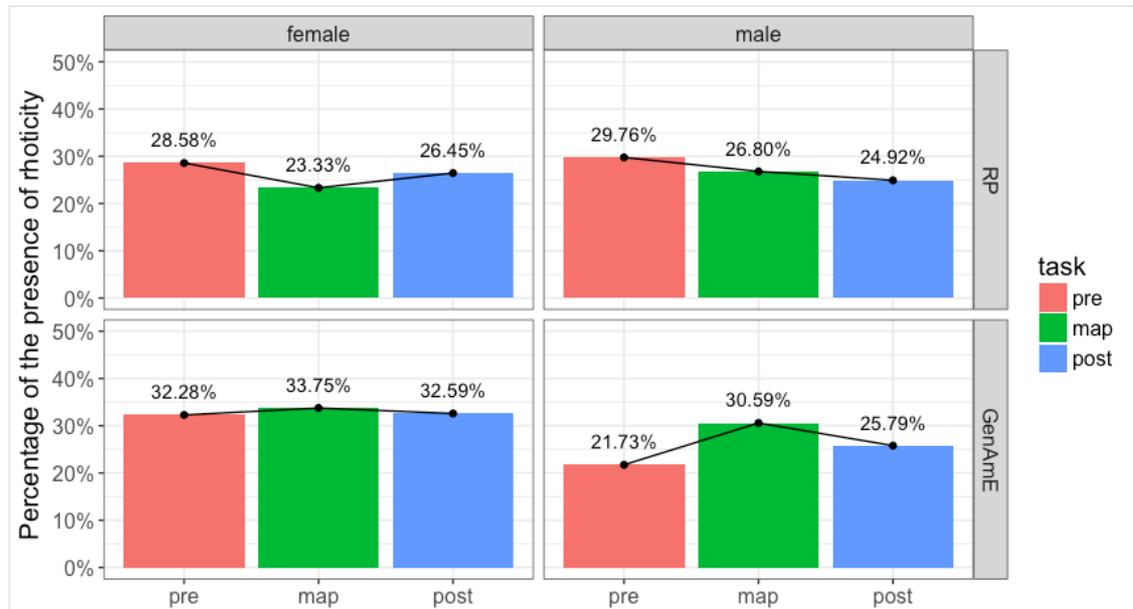


Figure 6.38 Percentage of the presence of rhoticity for the female and male participants across tasks and conditions.

6.4.3.3. Fricative [z] and fricative [θ]

For the two fricatives, logistic mixed effects regressions were used. Different from previous variables, **exposure** was not included as a fixed effect for the fricatives. Therefore, the full model for sex contained **task**, **sex** and **the interaction of task and sex** as the fixed factors. It also included random intercepts by participant, by word and by interlocutor, and random slopes by participant over task. The dependent variable was a binomial variable in which a realization of fricative [z]/fricative [θ] was coded as “yes” and the other variants were coded as “no”.

Comparison was carried out between the full model and a nested model for sex with **task*sex** removed.

Full model for sex = fricative ~ task*sex + (task | participant) + (1 | word) + (1 | interlocutor)

Nested model for sex = fricative ~ task + sex + (task | participant) + (1 | word) + (1 | interlocutor)

The results of comparisons including the Chi-square, DF and p value are shown in Table 6.27 and suggested a marginal effect of **task*sex** for fricative [z] and a significant effect for fricative [θ].

Fricatives	Chi-square	DF	p value
Fricative [z]	4.6506	2	0.097
Fricative [θ]	7.3864	2	0.025*

Table 6.27 Summary of the model comparisons for task*sex. The Chi-squares, degree of freedom and p-values are taken from ANOVA results of comparing the nested models with the first model for sex. * indicates a significant effect of **task*sex** for this variable.

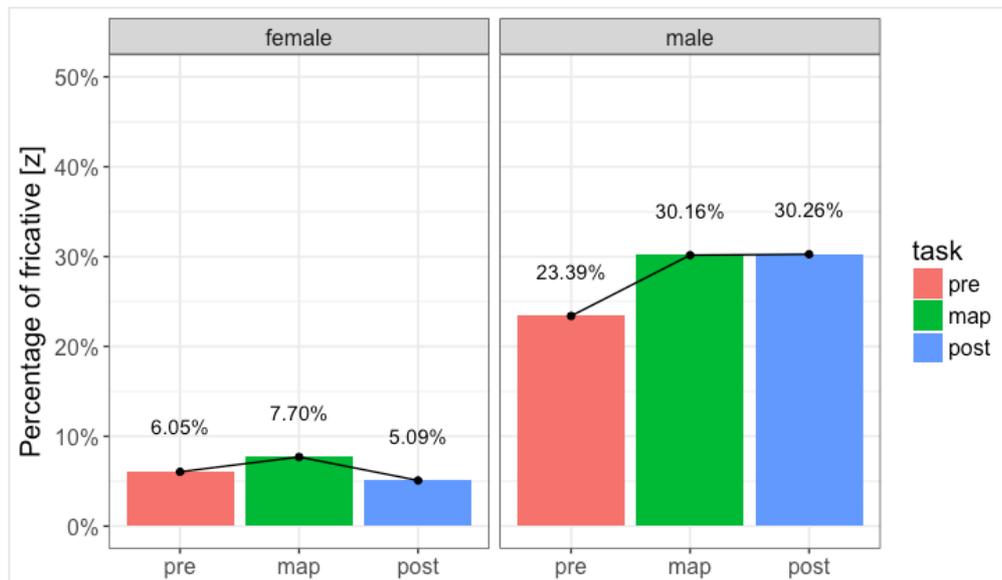


Figure 6.39 Percentage of fricative [z] for the female and male participants across tasks.

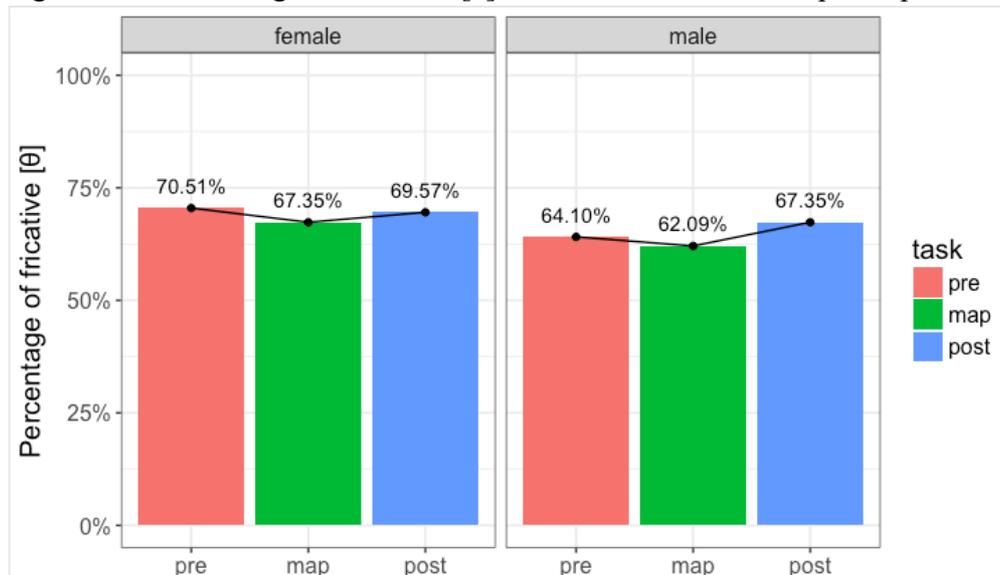


Figure 6.40 Percentage of fricative [θ] for the female and male participants across tasks.

For fricative [z], the full model containing the interaction of **task*sex** indicated a significant effect of **sex** (Estimate = 3.8069, SE = 1.66, $z = 2.29$, $p = .022$), suggesting that the male participants generally produced more fricative [z] than the female participants, regardless of the effect of task. This finding is also reflected in Figure 6.39. However, when the effect of task was considered, no significant effect was found for the interaction of **task*sex**. It could be that the increase in percentage found for the male participants was mainly contributed by one or two male participants (e.g. HK8 and HK12), and the huge individual differences reduced the effect of **task*sex**.

For fricative [θ], a significant effect was found for **task*sex**. The significance came from the comparison between the pre-tasks and the post-tasks when interacted with sex (SE= 0.25, $z = 2.52$, $p = .011$). That is, the male participants significantly produced more fricative [θ] than the female participants from the pre-tasks to the post-tasks.

6.4.3.4. Summary

This section reported the effect of **talker sex** for the five variables. The results suggested that for the Euclidean distance of the THOUGHT and the PATH vowel, rhoticity and fricative [z], no significant effect of **talker sex** was found, indicating that the female participants accommodated the same as the male participants did on these variables. For fricative [θ], the male participants significantly produced more fricative [θ] than the female participants from the pre-tasks to the post-tasks.

6.5. Chapter summary

This chapter presented a study of speech accommodation in HKE towards native English accents, namely RP and GenAmE. 19 participants completed two experiments, one with an RP speaker, the other with a GenAmE speaker. The participants' speech was collected before, during and after their conversations with the native speakers. The THOUGHT vowel and the PATH vowel were selected as variables, as well as rhoticity, fricative [z], and fricative [θ].

6.5.1. Summary of predictions and the results

Four predictions were made before the experiments. The following section summarises the accommodation from the pre-tasks and the map tasks, and responds to the four predictions.

H1. If the participants accommodate on the THOUGHT vowel, they are expected to shift their F1 and F2 towards the RP-*thought*/GenAmE-*thought* vowels when they talk to the RP/GenAmE speakers. Euclidean distance between the participants and the native speakers in the map tasks is expected to be shorter than the distance in the pre-tasks.

The results for F1 and F2 partially support H1, however, the results for Euclidean distance do not support this hypothesis.

In the RP condition, only the participants' **F1** showed a significant **convergence** towards the RP speakers. Though a small convergence was found for Euclidean distance, the change was not significant. In the GenAmE condition, only the participants' **F2** showed a significant **convergence** towards the GenAmE speakers. No significant change was found for Euclidean distance.

H2. If the participants accommodate on the PATH vowel, they are expected to shift their F1 and F2 towards the RP-*path*/GenAmE-*path* vowels when they talk to the RP/GenAmE speakers. Euclidean distance between the participants and the native speakers in the map tasks is expected to be shorter than the distance in the pre-tasks.

The results for F1 and F2 partially support H2. However, the results for Euclidean distance do not support H2.

In the RP condition, the participants' **F1** and **Euclidean distance** showed a **divergence** on PATH vowel.

In the GenAmE condition, the participants' F1 and F2 shifted towards the GenAmE speakers, but the results for Euclidean distance indicated a **divergence**.

H3. If the participants accommodate on rhoticity, they are expected to produce fewer rhotic words when they are exposed to the RP accent and they are expected to produce more rhotic words when they are exposed to the GenAmE accent.

The results support H3. In the RP condition, the participants showed a **convergence**, that is, they produced less rhotic words in the map tasks, though the differences were not significant. In the GenAmE condition, they showed a **convergence**, in that they significantly produced more rhotic words in the map tasks and in the post-tasks.

H4. If the participants accommodate on fricative [z] and fricative [θ], they are expected to produce more native-like items for these features.

The results partially support H4. For fricative [z], the participants significantly **converged** towards the native speakers in the map tasks. For fricative [θ], the participants showed a marginally significant effect of **divergence**; that is, they produced less fricative [θ] when they talked to the native speakers.

6.5.2. Summary of accommodation at an individual level

This chapter also reported the participants' accommodation on the five variables at the individual level. Table 6.28 shows the individual's convergence/divergence on the five variables. The cells which are shaded in an orange colour indicate a trend of convergence. Table 6.28 reveals a huge individual difference on the participants' accommodation across the five variables.

One reason might be that the HKE participants had different baselines when accommodating on these five variables, which is shown in Table 6.6. For example, HK4 had RP-like THOUGHT vowels and PATH vowels, produced words with non-rhoticity, rarely used fricative [z] (5%), but had no difficulty in pronouncing fricative [θ] (96%). On the other hand, HK8 had GenAmE-like THOUGHT vowels and PATH vowels, produced rhotic words (91%), used some fricative [z] (46%) and pronounced fricative [θ] (98%) most of the time. When the HKE participants had different starting points, it is unlikely that they would have the same pattern of accommodation on these variables.

RP condition	THOUGHT	PATH	rhoticity	[z]	[θ]
Prediction	–	–	–	+	+
HK1	-150.38	+48.30	-9.76%	+2.56%	0%
HK2	-30.60	+53.86	-4.26%	+2.99%	-19.85%
HK3	-35.65	-35.41	-20.01%	+2.44%	-4.42%
HK4	+39.51	+55.02	-3.24%	+12.2%	+1.43%
HK5	+41.66	-28.31	-16.33%	+2.36%	-2%
HK6	-38.64	-20.34	-11.69%	0%	+8.71%
HK7	-50.56	+49.68	+57.53%	0%	-39.99%
HK8	-82.59	-23.29	-4.19%	+12.3%	-1.49%
HK9	+23.58	+35.39	-9.38%	-14.18%	-2.18%
HK12	+9.90	+61.82	+0.11%	-1.04%	-16.86%
HK13	-92.57	+19.87	-17.1%	-1.69%	+2.77%
HK16	+17.77	-41.89	-2.22%	-21.66%	-2.18%
HK18	-24.75	+24.82	-10.51%	+18.11%	-4.36%
HK19	-78.48	+19.54	-8.41%	0%	+8.25%
HK20	+55.13	-16.84	-3.48%	+8.11%	-14.71%
HK21	+49.38	+65.34	+3.81%	+9.26%	+1.25%
HK22	+17.05	+47.01	+1.55%	+6.66%	+1.49%
HK23	-12.37	+65.31	-32.84%	+6.90%	-37.16%
GenAmE condition	THOUGHT	PATH	rhoticity	[z]	[θ]
Prediction	–	–	+	+	+
HK1	+18.54	-162.79	+2.23%	+6.25%	0%
HK2	+20.44	+10.89	+2.22%	+6.67%	+3%
HK3	+8.27	-50.92	-14.37%	-1.08%	+1%
HK4	+61.22	+58.23	+6.14%	+7.32%	+1.35%
HK5	+17.68	+112.53	+0.68%	+4.26%	-4.35%
HK6	-25.11	+30.01	+10.08%	0%	-3.18%
HK7	+60.44	+38.17	+2.55%	0%	+11.65%
HK8	+48.62	+108.71	+3.97%	+16.35%	+2.94%
HK9	-39.19	+57.55	+1.47%	-11.20%	-5.21%
HK12	-61.34	+57.79	+31.60%	+24.98%	-9.5%
HK13	-56.60	+23.50	+13.47%	0%	-11.76%
HK16	+19.08	+97.59	+23.85%	+5.22%	+0.07%
HK18	-12.30	+27.93	+15.11%	-9.36%	-2.59%
HK19	+54.71	+50.64	-5.73%	0%	+13.35%
HK20	+89.40	+89.62	+18.29%	0%	+29.11%
HK21	+50.39	+78.70	+0.08%	+3.75%	+5.02%
HK22	+97.47	+172.03	+0.13%	+9.68%	+1.89%
HK23	+27.04	+86.26	+2.49%	+5.88%	-4.1%

Table 6.28 Individual changes from the pre-task to the map tasks in the THOUGHT vowel, the PATH vowel, rhoticity, fricative [z] and [θ] in the RP and GenAmE conditions. For the THOUGHT and PATH vowels, the differences were based on Euclidean distance; for the consonants, the differences were based on the percentage change. The changes which indicate a trend of convergence are highlighted in orange. “Prediction” means in what way the changes would match the predictions.

Table 6.29 summarise the participants' accommodation patterns when they were divided into different groups based on their baselines of the sounds.

For the vowels, the HKE participants seem to be more likely to converge towards the interlocutors who were further away from them. This pattern is more obvious for the THOUGHT vowel. For the THOUGHT vowel, six HKE participants who had the largest distance from the RP interlocutors showed a convergence in the RP condition and a divergence in the GenAmE condition. Similarly, the HKE participant who had the largest distance from the GenAmE speaker showed a convergence in the GenAmE condition and a divergence in the RP condition.

Vowels			
	distance is closest/closer to the GenAmE speakers	distance is in between the RP and GenAmE speakers	distance is closest/closer to the RP speakers
THOUGHT	6 out of 8 HKE participants converged in the RP condition;	9 out of 10 HKE participants showed the same accommodation in the two conditions;	1 HKE participant converged in the GenAmE condition;
PATH	1 HKE participant converged in the RP condition;	12 out of 17 HKE participants showed the same accommodation in the two conditions;	1 HKE participants converged in the GenAmE condition;
Consonants			
	non-rhotic group (r% < 10%)	mild-rhotic group (20% < r% < 50%)	heavy-rhotic group (r% > 50%)
Rhoticity	5 out of 9 HKE participants converged in both conditions;	3 out of 5 HKE participants converged in one condition;	5 HKE participants converged in both conditions;
	[s]-dominant group (z% < 5%)	the mixed group (15% < z% < 55%)	[z]-dominant group (z% > 60%)
Fricative [z]	11 out of 14 HKE participants showed a small convergence;	3 out of 4 HKE participants showed a convergence;	1 HKE participant showed a ceiling effect, produced less rhotic words;
	[f]-dominant group (θ% < 5%)	the mixed group (20% < θ% < 50%)	[θ]-dominant group (θ% > 60%)
Fricative [θ]	1 out of 2 HKE participants showed a small convergence;	3 out of 5 HKE participants showed a divergence;	9 out of 12 HKE participants showed a ceiling effect.

Table 6.29 The participants' accommodation patterns for the five variables at the individual level.

For those participants who were neither close to the RP interlocutors nor closer to the GenAmE interlocutors, they showed the same accommodation in the two conditions. If they diverge in the RP condition, they would also diverge in the GenAmE condition. This pattern is common for both the THOUGHT vowel and the PATH vowel.

For the consonants, the patterns are less clear.

For the participants who produced less than 10% of the target consonants in the pre-tasks, they seem to show a small convergence on the sounds. This could be that they just started establishing the category of the target sounds; by receiving more input of the target sounds, they would be able to converge to a certain extent.

For the participants who produced about 20% - 50% of the target sounds in the pre-tasks, there is not a single pattern on their accommodation. Some participants showed a convergence while some others showed a divergence.

For the participants who produced more than 50% of the target sounds in the pre-tasks, they showed a convergence on rhoticity, but showed a divergence on the two fricatives. It could be that many participants who produced over 90% of fricative [z]/[θ] in the pre-tasks had reached a ceiling effect in the map tasks.

6.5.3. Other key findings

For the THOUGHT vowel (based on Euclidean distance) and rhoticity, **task*exposure** was found to be a significant predictor, indicating that the changes the HKE speakers had from the pre-tasks to the map tasks were significantly different from the changes they had in the GenAmE condition.

For the PATH vowel, only **task** was significant for the Euclidean distance, no significance was found for exposure and task*exposure. This suggests that exposing to different accents in the map tasks did not affect the HKE speakers' convergence on the PATH vowel.

For fricative [z] and fricative [θ], **task** was found to be a significant predictor, indicating that exposing to the native accents in the map tasks significantly changed their production on these two fricatives.

Talker sex was found to have no significant effect on the convergence of the THOUGHT vowel, the PATH vowel, rhoticity and fricative [z], indicating that the male and female HKE speakers accommodated similarly on these sounds. For fricative [θ], the male participants significantly converged more than the female participants did from the pre-tasks to the post-tasks.

For the native interlocutors, the two GenAmE speakers showed a trend of convergence towards the HKE participants on their PATH vowels.

6.6. Conclusions

Overall, the participants only showed a significant convergence on rhoticity and the fricative [z]. Divergence was found on the PATH vowel's Euclidean distance and fricative [θ]. For the THOUGHT vowel, no significant changes of Euclidean distance were found across tasks. At the individual level, the participants showed convergence towards the native speakers differently across the two vocalic variables and four consonantal variables.

Chapter 7. Study 3: Language attitudes and speech accommodation

7.1. Introduction

In the previous chapter, an accommodation study of HKE speakers towards native English accents was presented. This chapter aims to examine whether and how language attitudes affect HKE speakers' accommodation. The same participants from Study 2 (see Chapter 6) were used in the present study, which allows us to correlate the participants' language attitudes with their performance on accommodation. The experiment design of the attitude study in the present chapter is revised from Study 1 (see Chapter 3).

This chapter starts with the improvements made to the present study compared to Chapter 3, followed by the research question and hypothesis, experiment design, results and an analysis of individual speaker's profile.

7.2. What is new in the present study?

Study 1 surveyed 107 people from Hong Kong, investigating their attitudes towards British English, American English and HKE using the matched-guise method. The respondents listened to eight recordings and after each recording completed an accent evaluation form with 20 semantic differential traits. They also completed a questionnaire which aimed to elicit their direct attitudes towards the three English varieties. A female and a male bi-dialectal speaker were recruited to produce the recordings. Results from Study 1 suggested that British English was rated as more prestigious than American English for both bi-dialectal speakers; British English was rated as more attractive than American English for the female bi-dialectal speaker, whereas American English was rated as more attractive for the male bi-dialectal speaker. The results of the questionnaire suggested that British English was chosen as the more favoured English variety in teaching in Hong Kong. The findings of Study 1 were consistent with previous studies on language attitudes in Hong Kong.

Based on Study 1, a few changes were made to the design of the present study (see Table 7.1 for a summary). In Study 1, the sex of the bi-dialectal speakers seemed to influence

the respondents' attitudes; therefore in the present study, the male bi-dialectal speaker was replaced by another female bi-dialectal British speaker.

	Study 1	Present study
Respondents	107 Hong Kong respondents	18 Hong Kong respondents, the same participants from Study 2
Speakers & recordings	Speakers: <ul style="list-style-type: none"> • 1 female bi-dialectal British speaker • 1 male bi-dialectal American speaker • 2 male HKE speakers • 1 female British speaker and 1 female American speaker Recordings: <ul style="list-style-type: none"> • RP & GenAmE-guise • RP-guise & GenAmE • HKE-mild & HKE-strong • 2 fillers 	Speakers: <p>Same speakers & recordings except that the male bi-dialectal American speaker in Study 1 was replaced by another female bi-dialectal British speaker;</p> Recordings: <ul style="list-style-type: none"> • RP1 & GenAmE-guise1 (same) • RP2 & GenAmE-guise2 • HKE-mild & HKE-strong (same) • 2 fillers (same)
Accent identification	not included	Respondents were asked to identify the accent of the recording they heard;
Semantic differential traits	20 semantic differential traits; order of the traits was the same for all the recordings;	19 semantic differential traits (<i>pleasantness</i> was removed); order of the traits was randomised ;
Methods	the matched-guise method; PCA	same
Integrative & instrumental orientation	not included	included 20 statements which aimed to elicit respondents' integrative and instrumental orientation towards British English and American English

Table 7.1 A summary of the differences in experiment design between Studies 1 and 3.

In the traits on the accent evaluation form, *pleasantness* was removed from the present study. The result of PCA in Study 1 (see Table 3.2) suggested that *pleasantness* had similar factor loadings on the three dimensions (0.446 for *Status & Competence*, 0.529 for *Social Attractiveness* and 0.498 for *Factor 3*), indicating that *pleasantness* did not clearly belong to any of the three dimensions. Therefore, *pleasantness* was removed from the accent evaluation form in the present study.

Besides these changes, two sections were added in the present study. First of all, **an accent identification task** was added after each recording. This aimed to investigate whether the respondents were able to identify British English, American English and HKE. Secondly, 20 statements were added (see Table 7.3 for details) to elicit the respondents' **integrative and instrumental orientation** towards British English and American English. According to the social-educational model (Gardner, 1985, 2006, 2010), English learners have two types of motivation in language learning: integrative and instrumental orientation. Integrative orientation refers to "a class of reasons that suggest that the individual is learning a second language in order to learn about, interact with, or become closer to the second language community" (1985, p54). Instrumental orientation refers to "a class of reasons that suggest that the individual is learning a language for pragmatic reasons and has a desire to gain social recognition or economic advantage through knowledge of a foreign language" (1985, p11). For example, if a HKE speaker adopts RP accent because he/she likes British culture, in this case, he/she has a strong integrative orientation towards British English. On the other hand, if a HKE speaker adopts GenAmE accent because he/she believes that speaking American English could help him/her on career development, in this case, he/she has a strong integrative orientation towards American English. Respondents with different integrative and instrumental orientation may accommodate differently towards RP speakers and GenAmE speakers in the map tasks.

7.3. Research questions and hypothesis

There is one main research question for the present study: **do Hong Kong English speakers' language attitudes affect their accommodation towards native English accents?**

CAT (Giles et al., 1991) suggests that the purpose of accommodation is to shorten the social distance between interlocutors and to gain social approval from others. Therefore, the hypothesis of the present study is: **Hong Kong people accommodate more towards the English variety they favour.** In other words, in this study, the people who liked British English more would accommodate more towards the RP speakers in the map task and those who favoured American English over British English would accommodate more towards the GenAmE speakers in the map task.

7.4. Experiment design

7.4.1. Participants

19 respondents who participated in Study 2 completed a survey. HK11's data was removed from this study because she did not complete the survey. Therefore, in total 18 respondents participated in the study: seven male respondents and eleven female respondents.

7.4.2. Recordings

Eight speakers made the eight recordings used in the accent evaluation task. The same recordings from Study 1 were used in the present study, except that two recordings from the male bi-dialectal American speaker were removed. He was replaced by another female bi-dialectal British speaker. The new bi-dialectal British speaker was an actor who had received training in the American accent. She came from southern England and was 26 years old. She read *The North Wind and the Sun* in both an RP accent and a GenAmE accent. Two samples were selected from four trials and the authenticity of her accents was examined by two experienced phoneticians from the Department of Language and Linguistic Science at the University of York. The recordings used in the present study are shown in Table 7.2.

Speaker	Recording	gender
bi-dialectal RP speaker1	RP1	female
	GenAmE-guise2	female
bi-dialectal RP speaker2	RP2	female
	GenAmE-guise2	female
HKE speaker with a strong accent	HKE-strong	male
HKE speaker with a mild accent	HKE-mild	male
RP speaker	RP (filler)	female
GenAmE speaker	GenAmE (filler)	female

Table 7.2 The eight recordings used in Study 3.

7.4.3. Survey

The survey consisted of four parts: accent evaluation (PCA), accent identification, and an evaluation of integrative and instrumental orientation.

7.4.3.1. Accent identification

The respondents were asked to identify the origin of the speaker for each recording. Five regions (the UK, US, Australia, Canada and Hong Kong) were given. The respondents

were asked to choose one from the list of five regions. The reason for including Australia and Canada was that English accents from these regions are familiar to people from Hong Kong for historical reasons and due to immigration.

7.4.3.2. Accent evaluation in Principle Component Analysis

This part was identical to the matched-guise design in Study 1 (see Chapter 3.3 for details). The respondents completed an accent evaluation form right after they had listened to one of the eight recordings. The accent evaluation form was almost the same as the one used in Study 1, containing a 6-point scale and 19 semantic differential traits. The semantic differential trait *pleasantness* was removed and the order of the 19 semantic differential traits was randomised for each recording.

7.4.3.3. Evaluation of integrative and instrumental orientation

The evaluation contained 20 statements, shown in Table 7.3.

Item	Integrative statements of British English
1	I think people from Hong Kong should speak British English because Hong Kong is greatly influenced by the UK;
2	A person who speaks British English is usually arrogant, snobbish and a show-off;
3	I like British English (pronunciation);
4	I like people who speak British English and their way of life;
5	A person who speaks British English is usually educated, intelligent and well-off;
	Instrumental statements of British English
6	British English is less important in Hong Kong after the handover;
7	British English will help me greatly in getting better opportunities for further studies;
8	British English will help me greatly for better career opportunities in the 21st century;
9	I would like my children to speak British English;
10	British English is usually used in formal situations.
	Integrative statements of American English
11	I think people from Hong Kong should speak American English because it is so popular around the world now;
12	I like American English (pronunciation);
13	A person who speaks American English is usually educated, intelligent and well-off;
14	I like people who speak American English and their way of life;
15	A person who speaks American English is usually arrogant, snobbish and a show-off;
	Instrumental statements of American English
16	American English is going to replace the status and importance of British English after the handover;
17	American English will help me greatly in getting better opportunities for further studies;
18	American English will help me greatly for better career opportunities in the 21st century;
19	I would like my children to speak American English;
20	American English is usually used in informal situations.

Table 7.3 The integrative and instrumental statements of British English and American English.

Some of the statements were selected from Cheng (2013, p5): for example, “I like people who speak British English and their way of life.” Some statements were created by the author based on Gardner (2010). The respondents were asked to rate from 1 to 10, where “1” means “strongly disagree” and “10” means “strongly agree”. The statements were presented in a random order in the survey.

7.4.4. Procedure

The respondents were asked to complete the survey before they came to the recording booth for their first accommodation experiment. They were told that the purpose of the study was to investigate conversations between Hong Kong people and native English speakers.

7.5. Results

The results are presented in the following order: the descriptive results of the survey are presented first, followed by the results of linear mixed effects regressions which were used to investigate the correlation between attitudes and accommodation.

7.5.1. Descriptive results of the survey

7.5.1.1. Accent identification

In the accent identification task, the respondents were asked to identify the speaker’s origin by choosing from a list of the UK, the US, Australia, Canada and Hong Kong. Two types of identification rates are used: an accurate identification rate and a broad identification rate. For the **accurate identification rate**, if a respondent accurately identified the accent, one point would be given. The sum of the points divided by 18 (the total number of respondents) was the accurate identification rate. For example, the accurate identification rate for RP1 was 89%, which means 16 out of 18 respondents successfully identified that the speaker RP1 was from the UK. **The broad identification rate** adopted a criterion in which if an RP recording was identified as Australian English or a GenAmE recording identified as Canadian English, one point would also be given. For example, the broad identification rate of RP1 was 94%, suggesting that 17 out of 18 respondents identified that the speaker RP1 was from the UK or from Australia.

	British English			American English			HKE	
	RP1	RP2	RP-filler	GenAm E-guise1	GenAm E-guise2	GenAm E-filler	HKE-mild	HKE-strong
Accurate identification%	89% (16/18)	61% (11/18)	83% (15/18)	39% (7/18)	56% (10/18)	39% (7/18)	28% (5/18)	100% (18/18)
Broad identification%	94% (17/18)	67% (12/18)	94% (17/18)	61% (11/18)	83% (15/18)	78% (14/18)	-	-

Table 7.4 Accent identification rate for the eight recordings in Study 3. In the brackets, tokens of successful identification/total respondents are given.

Note that the accurate identification rate for HKE-mild was only 28%, which means only five respondents successfully identified the speaker of HKE-mild. For this recording, six respondents identified the HKE-mild recording as British English, four identified it as Australian English, two identified it as American English and one identified it as Canadian English. The reason for the low identification rate of the HKE-mild is that the speaker of HKE-mild had a strong RP accent in his English, so although his English still carried a degree of Cantonese accent, apparently most of the respondents were not able to identify these Cantonese features.

The results in Table 7.4 suggest that most of the respondents accurately identified the RP recordings as British English; around half of the respondents identified the GenAmE recordings as American English and all the respondents successfully identified the HKE-strong recording.

Table 7.5 shows the individual identification rate for all the recordings (overall identification%) and the identification rate for the six native English recordings (native English identification%). Each respondent rated eight recordings, including six native English recordings (3 RP recordings and 3 GenAmE recordings) and two HKE recordings. One point was given to one successful identification, and then the total number of points divided by eight was the overall identification rate, and divided by six was the native English identification rate.

	Overall identification%	Native English identification%	RP identification ratio	GenAmE identification ratio
HK16	100% (8/8)	100% (6/6)	3/3	3/3
HK9	100% (8/8)	100% (6/6)	3/3	3/3
HK8	88% (7/8)	100% (6/6)	3/3	3/3
HK5	75% (6/8)	83% (5/6)	3/3	2/3
HK12	75% (6/8)	83% (5/6)	3/3	2/3
HK18	75% (6/8)	67% (4/6)	2/3	2/3
HK4	75% (6/8)	50% (3/6)	3/3	1/3
HK6	63% (5/8)	67% (4/6)	2/3	2/3
HK19	63% (5/8)	67% (4/6)	3/3	1/3
HK22	63% (5/8)	67% (4/6)	3/3	1/3
HK23	63% (5/8)	67% (4/6)	3/3	1/3
HK13	63% (5/8)	50% (3/6)	2/3	1/3
HK1	50% (4/8)	50% (3/6)	2/3	1/3
HK20	50% (4/8)	50% (3/6)	2/3	1/3
HK2	38% (3/8)	33% (2/6)	2/3	0/3
HK3	38% (3/8)	33% (2/6)	2/3	0/3
HK7	25% (2/8)	17% (1/6)	1/3	0/3
HK21	13% (1/8)	0% (0/6)	0/3	0/3

Table 7.5 Identification rates for all the recordings (overall identification%) and for native English recordings (native English identification%). The ratio represents the token of successful identified recordings/tokens of the recordings of interest. Native English identification% represents the accuracy of identifying the RP and GenAmE recordings.

According to Table 7.5, most of the respondents had an overall identification rate of over 50%, except for HK2, HK3, HK7 and HK21, who performed poorly in both overall identification and native English identification.

7.5.1.2. Accent evaluation (PCA)

The respondents' ratings of the eight recordings were imported into SPSS for PCA analysis, following the same procedure as in Study 1 (see Chapter 3, section 4.1 for details). Three components were extracted from the data.

The first component was named *Linguistic Attractiveness*, indicating a dimension of how attractive the accent is based on its linguistic quality, such as pronunciation or intelligibility. The second component was named *Social Attractiveness*, indicating a dimension of the speaker's social attractiveness, such as humour or enthusiasm. The third component was named *Status*, indicating a dimension associated with a speaker's economic status and education background. Each of the components explained around

20% of the variance in the data, and together the three components explained 64.81% of the data. The traits for each component are shown in Table 7.6.

Traits	Factor 1 (Linguistic Attractiveness)	Factor 2 (Social Attractiveness)	Factor 3 (Status)
Aesthetic quality	.823	.239	.312
Intelligibility	.805	.309	.138
Model of pronunciation	.794	.121	.361
Sincerity	.667	.209	.238
Attractiveness	.591	.437	.377
Confidence	.570	.380	.421
Intelligence	.522	.163	.467
Enthusiasm	.165	.866	.125
Friendliness	.327	.800	.110
Humour	.126	.796	.239
Interestingness	.251	.774	.299
Fashion	.234	.533	.150
Formality	.250	.162	.777
Politeness	.204	.370	.759
Social class	.429	.143	.718
Education	.552	-.020	.610
Reliability	.434	.356	.601
Leadership	.214	.482	.587
Ambition	.069	.171	.078
<i>Eigenvalue(Rotation)</i>	<i>4.447</i>	<i>4.052</i>	<i>3.817</i>
<i>% of variance explained (rotation)</i>	<i>23.403%</i>	<i>21.326%</i>	<i>20.089%</i>

Table 7.6 Factor loadings of the rotated component matrix of the 19 semantic differential traits in Study 3. Traits which were highly correlated with the same factor are shaded in grey. Ambition did not correlate with any of the three factors, therefore it is not shaded.

If we compare the three factors here with the three factors from Study 1, what can be found is that the traits in *Status & Competence* of Study 1 largely overlap with the traits in *Linguistic Attractiveness* and *Status* here. And the traits in *Social Attractiveness* here are mostly the same as those in Study 1. In other words, the factor *Status & Competence* in Study 1 was split into two different factors (*Linguistic Attractiveness* and *Status*) in the current study, and factor *Social Attractiveness* in Study 1 remained the same here. Though the three factors in the two studies were different, they were highly consistent in how the traits were categorised.

To gain an overall picture of the data, the ratings for the same type of accent were combined. For example, the means and standard deviations for RP were calculated based on the ratings of the RP1 and RP2 produced by the two bi-dialectal speakers. Similarly, the ratings of GenAmE-guise1 and GenAmE-guise2 were combined. The means of RP, GenAmE and HKE are presented in Table 7.7.

	Linguistic Attractiveness	Social Attractiveness	Status
RP	0.442 (0.711)	0.321 (0.702)	0.346 (0.960)
GenAmE	-0.017 (0.806)	0.640 (0.898)	-0.261 (0.939)
HKE	-0.692 (1.296)	-0.433 (0.899)	-0.462 (0.960)

Table 7.7 Mean factor scores and SD (in brackets) for RP, GenAmE and HKE in the three dimensions in Study 3. The higher the value is, the more positive an attitude it indicates. The highest value among the three varieties is highlighted in bold.

The results of factor scores suggest that RP was rated highest in *Linguistic Attractiveness* and *Status*, while GenAmE was rated highest in *Social Attractiveness*. HKE was rated the lowest for all the three dimensions.

Figure 7.1 plots the ratings of each recordings for the three dimensions.

The results here are partially in line with the results of Study 1. In Study 1, RP was rated as the most prestigious and most attractive English variety in *Status & Competence* and *Social Attractiveness*, while in the present study, RP was still rated as the most prestigious and rated highest in *Linguistic Attractiveness*, but GenAmE was rated as the most socially attractive among all.

To explore the differences of results between Study 1 and Study 3, means of ratings on recordings produced by the two bi-dialectal speakers for *Social Attractiveness* were calculated. The results suggested that for the bi-dialectal speaker 1, her RP1 and GenAmE-guise 1 were rated approximately the same (RP1 = 0.26, GenAmE-guise1 = 0.23); for the bi-dialectal speaker 2, her RP2 was rated much lower than her GenAmE-guise 2 (RP2 = 0.38, GenAmE-guise 2 = 1.05). This suggests that the advantage of GenAmE in *Social Attractiveness* was mainly contributed by the recording of GenAmE-guise 2. The HKE participants also performed better in identifying the GenAmE-guise 2.

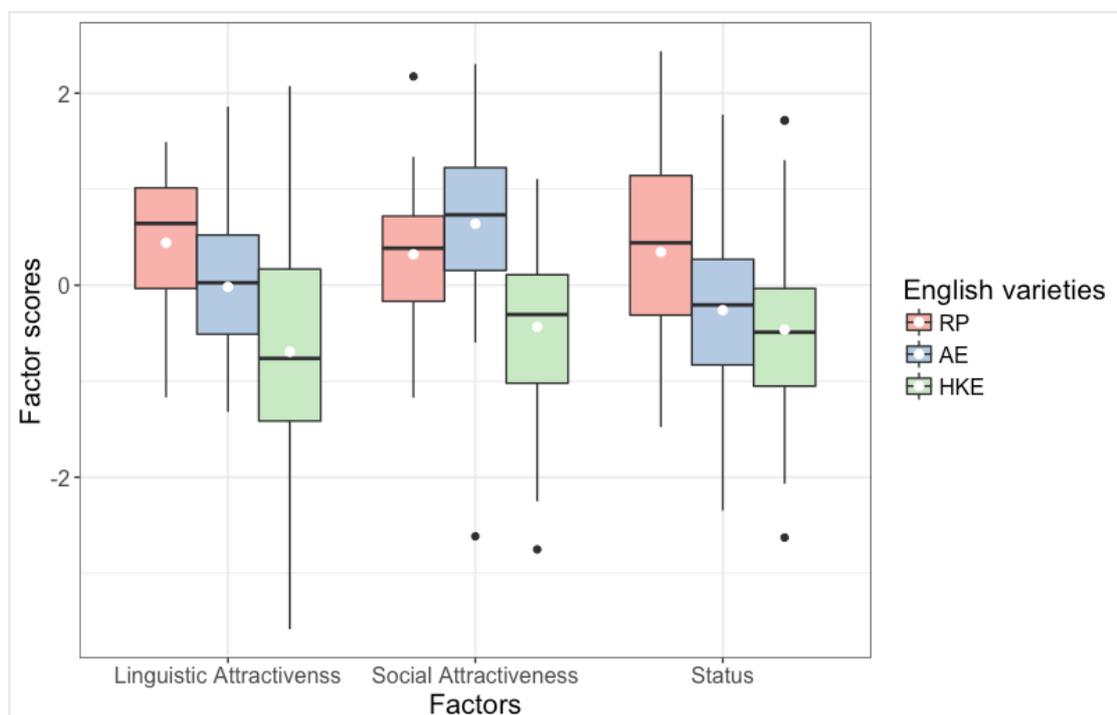


Figure 7.1 Boxplots of factor scores for RP, GenAmE and HKE (from left to right) across the three dimensions. The white dots represent the means, the black lines in the middle of the boxplots represent the medians. The higher the score, the more positive an attitude it indicates.

The inconsistency of ratings between the two bi-dialectal speakers was only found in *Social Attractiveness*. For the other two dimensions, the ratings for their RP samples and GenAmE-guise samples were highly consistent, in that both the RP samples were rated higher than the GenAmE-guise samples. It could be that the HKE participants' low identification rate on GenAmE-guise 1 (i.e. 39%) affect their attitudes of *Social Attractiveness* towards this sample.

Note that the purpose of the present study is different from Study 1. The present study has no intention to make a prediction of Hong Kong people's language attitudes from these 18 respondents. Instead, the present study focuses on individual attitudes.

7.5.1.3. Integrative and instrumental orientation

The respondents rated 20 statements relating to the integrative and instrumental orientation of British English and American English. The rating scale is 1 to 10. The higher the score, the more positive an attitude it indicates. Note that among all the statements, sixteen statements expressed a positive attitude while four statements

expressed a negative attitude. In order to maintain the consistency that a higher score indicates a more positive attitude, the ratings of the four negative statements were reversed; for example, “1” was reversed to “10”, “8” was reversed to “2” etc. The means of the respondents’ ratings of each statement and the overall means of integrative and instrumental orientation for British English and American English were also calculated and are shown in Table 7.8. Figure 7.2 shows the boxplots of the ratings.

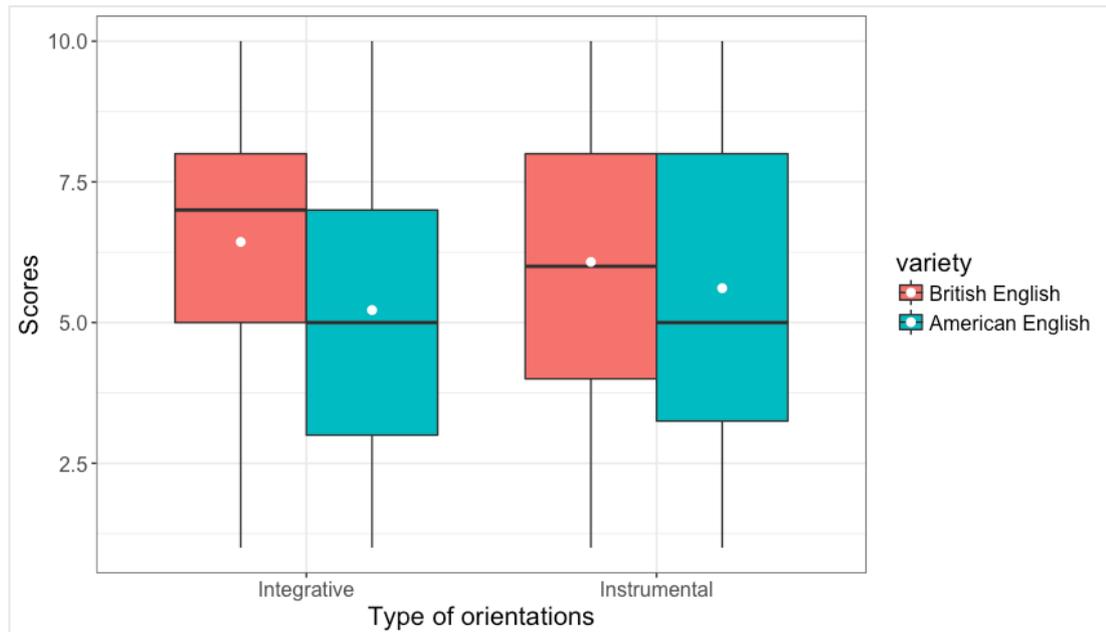


Figure 7.2 Boxplots of scores for integrative and instrumental orientation towards British English (left) and American English (right). The white dots represent the means, the black lines in the middle of the boxplots represent the medians. The higher the score, the more positive an attitude it indicates.

Independent t-tests were used to compare the overall means of integrative and instrumental statements between British English and American English. The results suggest that the ratings of **integrative statements** for British English were significantly higher than the ratings for American English (integrative: BE-AE = 1.11, $t=3.172$, $DF=178$, $p=0.001$). No significance was found between the ratings of **instrumental statements** for British English and the ratings for American English (instrumental: BE-AE = 0.46, $t=1.182$, $DF=178$, $p=0.23$).

Item	Integrative statements of British English	Mean
1	I think people from Hong Kong should speak British English because Hong Kong is greatly influenced by the UK;	6.44
2*	A person who speaks British English is usually arrogant, snobbish and a show-off;	7.44
3	I like British English (pronunciation);	7.50
4	I like people who speak British English and their way of life;	5.72
5	A person who speaks British English is usually educated, intelligent and well-off;	5.06
<i>overall</i>		6.43
Instrumental statements of British English		
6*	British English is less important in Hong Kong after the handover;	4.66
7	British English will help me greatly in getting better opportunities for further studies;	5.83
8	British English will help me greatly for better career opportunities in the 21st century;	6.17
9	I would like my children to speak British English;	6.67
10	British English is usually used in formal situations.	7.06
<i>overall</i>		6.07
Integrative statements of American English		
11	I think people from Hong Kong should speak American English because it is so popular around the world now;	4.83
12	I like American English (pronunciation);	4.67
13	A person who speaks American English is usually educated, intelligent and well-off;	5.17
14	I like people who speak American English and their way of life;	5.00
15*	A person who speaks American English is usually arrogant, snobbish and a show-off;	6.44
<i>overall</i>		5.22
Instrumental statements of American English		
16	American English is going to replace the status and importance of British English after the handover;	5.67
17	American English will help me greatly in getting better opportunities for further studies;	5.39
18	American English will help me greatly for better career opportunities in the 21st century;	5.83
19	I would like my children to speak American English;	4.33
20*	American English is usually used in informal situations.	6.83
<i>overall</i>		5.61

Table 7.8 Means of the integrative and instrumental statements for British English and American English. The statements which expressed a negative attitude are marked with *, for these statements, scores were reversed when calculating the means. The higher the score, the more positive an attitude it indicates.

7.5.1.4. Summary

The previous sections have reported the results of individuals' accent identification accuracy, attitudes of PCA, and attitudes of integrative and instrumental orientation. To gain an overall review of each individual's attitudes, Table 7.9 summarises each respondent's preference in all the attitude tests. Whichever the respondents rated higher

on each of the attitude tests is shown in the table. For example, HK12 had five “RP” labels, indicating that HK12 rated RP over GenAmE in all these attitude tests. Note that Table 7.9 only indicates a preference for the English variety, it does not indicate a significant difference between RP and GenAmE in the tests.

Participant	Attitudes from PCA			Explicit attitudes	
	<i>Linguistic Attractiveness</i>	<i>Social Attractiveness</i>	<i>Status</i>	<i>Integrative</i>	<i>Instrumental</i>
HK12	RP	RP	RP	RP	RP
HK18	RP	RP	RP	RP	RP
HK2	RP	RP	GenAmE	RP	RP
HK20	RP	RP	GenAmE	RP	same
HK9	RP	RP	GenAmE	RP	RP
HK4	RP	GenAmE	RP	RP	RP
HK13	RP	GenAmE	RP	RP	RP
HK16	RP	GenAmE	RP	RP	RP
HK3	RP	GenAmE	RP	RP	GenAmE
HK22	RP	GenAmE	RP	RP	GenAmE
HK23	RP	GenAmE	RP	RP	GenAmE
HK5	GenAmE	GenAmE	RP	RP	RP
HK19	GenAmE	GenAmE	RP	RP	same
HK7	RP	GenAmE	RP	GenAmE	GenAmE
HK1	RP	GenAmE	GenAmE	RP	GenAmE
HK21	GenAmE	GenAmE	RP	GenAmE	RP
HK6	GenAmE	GenAmE	RP	GenAmE	GenAmE
HK8	GenAmE	GenAmE	RP	GenAmE	GenAmE
Overall	RP	GenAmE	RP	RP	RP

Table 7.9 Individual preferences of English variety/accent across the attitude tests. Note that in the explicit attitude tests, the labels “British English/American English” were replaced by “RP/GenAmE” to make them consistent with the labels of attitudes from PCA. In the column *Instrumental*, “same” indicates that the respondent had the same ratings for RP and GenAmE.

According to Table 7.9, the majority of the respondents show a preference for RP in all the dimensions, apart from *Social Attractiveness* where GenAmE was favoured. Also, most of the respondents did not have a unique preference across the attitude tests, except for HK12 and HK18.

It is not surprising that individuals showed different attitudes in these tests. First of all, the two attitude tests elicited people’s attitudes from different dimensions. The accent evaluation of PCA recorded people’s immediate judgements on 19 traits, which revealed their indirect attitudes towards the speakers of the accents. The explicit attitudes test

focused on people's integrative and instrumental orientation towards an English variety. Secondly, we should not assume that people have a unique preference across all the aspects of attitudes. Attitudes are multi-dimensional and complex. Most of the time, people are consistent in what they like. For example, people who like oranges may tend to like orange-flavoured chocolates; however, they may not like an orange-scented shampoo or their hair dyed in an orange colour. Similarly, HK3 preferred GenAmE on the dimensions of *Social Attractiveness* and *Instrumental Orientation*, but she liked RP more on the dimensions of *Linguistic Attractiveness*, *Status* and *Integrative Orientation*.

7.5.2. Linear mixed effects regressions

In order to run linear mixed effects regressions between the results of attitudes and the results of accommodation, a few adjustments were made to the data.

Firstly, data from HK2, HK3, HK7 and HK21 were removed from all the linear mixed effects regressions due to their poor performances on identifying RP and GenAmE. HK21 did not identify any of the RP and GenAmE recordings (0% for the native English identification), HK7 only identified one (17%), and HK2 and HK3 only identified two recordings (33%). It is important to only include the data from the respondents who could identify RP and GenAmE. If a respondent could not identify RP, it is unlikely that his/her attitudes towards RP would influence his/her accommodation.

Secondly, only the percentage change of **rhoticity** from the pre-task to the map task was chosen as the dependent variable for the model, for two reasons. The first reason is that rhoticity and fricative [z] were the only two variables that showed consistent convergence in the experiment. As the model aimed to investigate the correlation between the respondents' attitudes and their accommodation, choosing the variables that showed convergence is vital. The second reason is that RP and GenAmE have a very clear distinction on rhoticity, but not on the fricative [z]. RP is non-rhotic and GenAmE is rhotic; however, RP and GenAmE do not differ in their use of the fricative [z]. Therefore, if attitudes were to affect accommodation as predicted, the effect of attitudes would reflect more on the respondents' production of rhoticity compared to their production of fricative [z]. Thus, only rhoticity was chosen as the dependent variable for the model.

Thirdly, only the scores of *Linguistic Attractiveness* and *Integrative Orientation* were included as the fixed effects factors for the regression models; other dimensions of attitudes like *Social Attractiveness*, *Status* and *Instrumental Orientation* were excluded. Due to the small size of the sample (14 respondents in two exposure conditions, 28 data points in total) and to avoid overfitting the regression models, it is safer to use a small number of fixed effects. *Linguistic Attractiveness* was chosen because it was the first component of PCA, representing the most powerful component among the three. It explained the largest proportion of variance (23.4%) compared to the other two components. *Integrative Orientation* was chosen because the integrative motive was found to have a clear and long-term impact on L2 learners' performance, whereas the role of instrumental motive was less direct and clear. Therefore, only *Integrative* was chosen to represent the respondents' explicit attitudes.

7.5.2.1. Attitudes and accommodation of vowels

A series of linear mixed effects regressions were run. The HKE participants were expected to produce less rhoticity in the RP condition, but in the GenAmE condition, they were expected to produce more rhoticity. Therefore, the interactions between exposure and attitudes were the interests for model comparisons. To explore the effect of *Linguistic Attractiveness* and *Integrative Orientation* separately, two full models were used. The first full model contained *exposure*, *Linguistic Attractiveness*, *Integrative Orientation* and interaction between *exposure* and *Linguistic Attractiveness* as the fixed effects, and contained a random intercept by participant. The second full model was similar to the first full model, except that interaction between *exposure* and *Integrative Orientation* was used as the fixed effect instead. The dependent variable was the respondents' percentage change of rhoticity from the pre-tasks to the map tasks.

The formula for the full models and the nested model are shown below:

Full model 1 = rhoticity ~ exposure*linguistic attractiveness + integrative orientation + (1 | participant)

Full model 2 = rhoticity ~ exposure*integrative orientation + linguistic attractiveness + (1 | participant)

Nested model = rhoticity ~ exposure + linguistic attractiveness + integrative orientation + (1 | participant)

The nested model contained no interaction between the fixed factors and was used to compare with the two full models. The results of the model comparisons including the Chi-square, DF and p value are shown in Table 7.10.

Term	Chi-square	DF	p value
exposure*Linguistic Attractiveness	0.0566	1	0.812
exposure*Integrative	0.0707	1	0.790

Table 7.10 Summary of the model comparisons for rhoticity. The Chi-squares, degree of freedom and p-values are taken from ANOVA results comparing the nested models with the full model.

The results suggest that *exposure*Linguistic Attractiveness* and *exposure*Integrative Orientation* did not improve the fit of the models. In other words, *Linguistic Attractiveness* and *Integrative Orientation* did not predict the respondents' accommodation on rhoticity. The correlations between rhoticity, Linguistic Attractiveness and Integrative Orientation were plotted in Figure 7.3.

The results of non-significant correlations between attitudes and accommodation may be due to three reasons. The most straightforward explanation is that attitudes might not have an impact on non-native speakers' convergence. Though most of the HKE speakers in Study 3 were advanced L2 learners, they might still have to place most of their attention in completing the map tasks. Little cognitive space might be left for attitudes to interfere speech production.

Secondly, the present study only examined the respondents' indirect attitudes (i.e. accent evaluation of PCA) and explicit attitudes (i.e. integrative and instrumental motives): it might be that these two types of attitudes do not correlate with people's accommodation while other types of attitudes do. For example, Babel (2010) elicited New Zealanders' attitudes towards Australia using IAT and found that the more positive attitudes the people had towards Australia, the more likely they converged towards Australian vowels.

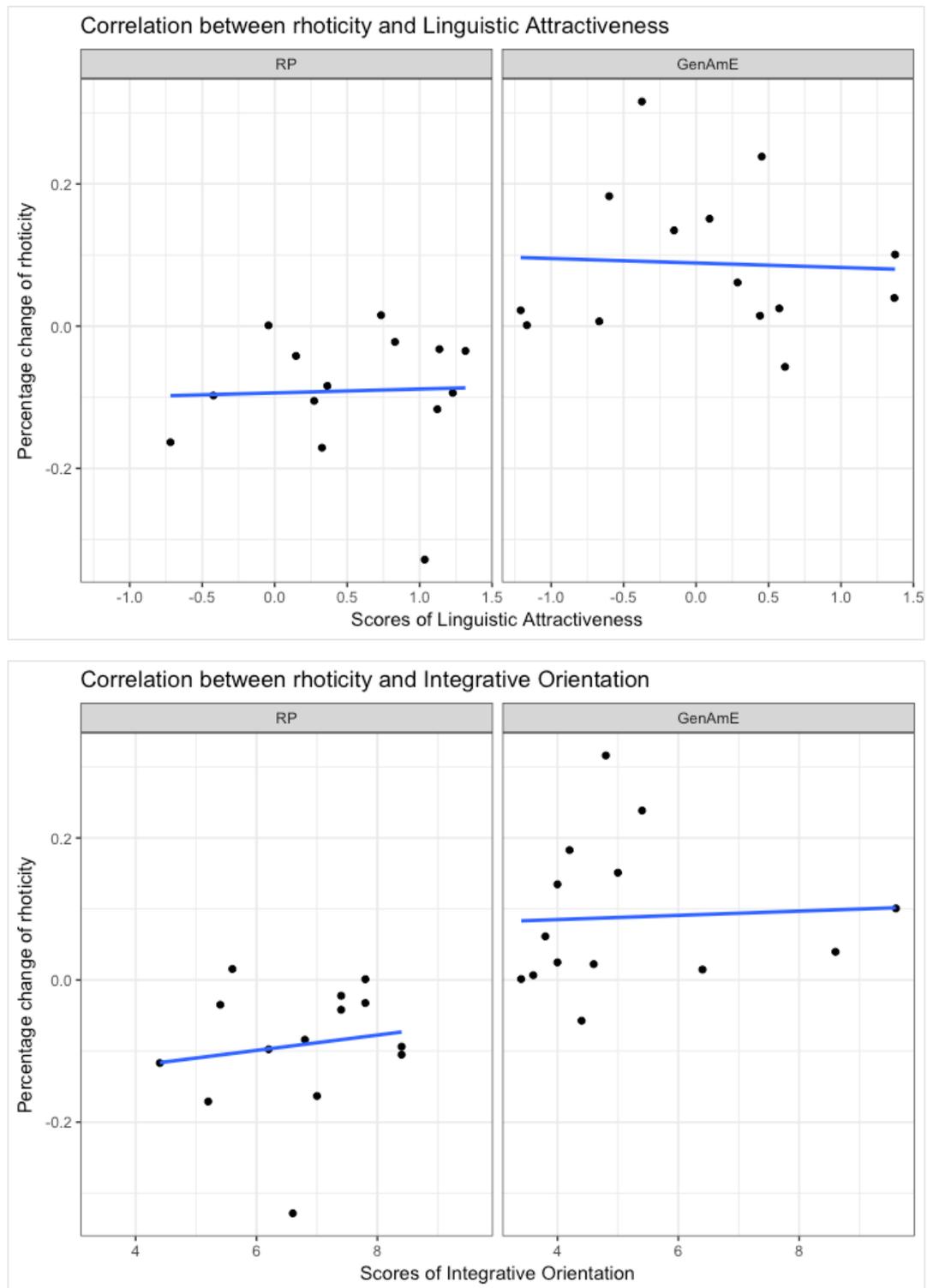


Figure 7.3 The correlation between Linguistic Attractiveness and rhoticity (top figure), and between Integrative Orientation and rhoticity in two conditions (bottom figure). Each black dot represents a respondent in one exposure condition, the blue line indicates the regression line.

For example, if *Status* is the most important attitude for respondent A, he/she might accommodate more towards the accent which he/she rates as more prestigious. However,

this might not be the case for respondent B, who cares about *Social Attractiveness* more. Thirdly, the respondents might be affected by different aspects of attitudes. As the respondents rarely had a unique preference for RP or GenAmE, it is likely that they responded to different dimensions of attitudes in their accommodation. Respondent B would not accommodate towards the accent that is more prestigious; instead, he/she would accommodate towards the accent that is more socially attractive. If this is the case, a consistent pattern between one aspect of attitudes and accommodation would not be found, because the respondents might not respond to the same aspect of attitudes in their accommodation. More examples are provided to support the third point in the next section.

Finally, it should be noted that the sample size of Study 3 is small. Studies with a larger data set are needed to explore the correlation between attitudes and non-native speakers' convergence.

7.5.2.2. Summary

This section examined the correlation between language attitudes and accommodation, using linear mixed effects regressions. The results suggest that *Linguistic Attractiveness* and *Integrative Orientation* did not correlate with the respondents' accommodation in rhoticity, indicating that the two types of attitudes examined in the present study had limited impact on people's overall degree of convergence.

Note that the results here were based on a small sample size: only 14 respondents were included in the data set. Studies with a larger sample size are needed to explore the correlation between attitudes and accommodation.

7.6. Speakers' profiles

In the previous sections, attitudes were found to have no impact on convergence of the HKE speakers in the statistical analysis. This result was not surprising because individuals showed a complex pattern on both accommodation and attitudes. They did not seem to converge on all sounds, neither did they showed a consistency on the attitudes towards British English and American English. Due to the complexity of convergence and attitudes, the statistical analysis might not be able to capture the whole picture.

Therefore, in this section, three HKE speakers' profiles were presented with examples extracted from the conversations between the HKE speakers and the native interlocutors.

The speakers represented three of the most common types of convergence observed in the present data set: (1) HK20 who converged more towards the GenAmE speaker; (2) HK23 who converged more towards the RP speaker; (3) HK8 who converged the same in both conditions. For each of these HKE speakers, a profile of their language attitudes and convergence on rhoticity was provided. In addition, an episode of the conversation from the map tasks was presented. Note that the patterns of convergence are not limited to these three types.

7.6.1. HK20: a respondent who converged more towards the GenAmE speaker

The first type of convergence is represented by HK20, a female participant who accommodated more in the GenAmE condition on rhoticity. Table 7.11 summarises HK20's preferences on different dimensions of attitudes and the statistical change in rhoticity between the pre-task and the map task.

Participant	HK20		Sex: female	
Attitudes	RP was preferred in: <i>Linguistic Attractiveness, Social Attractiveness, Integrative Orientation, Instrumental Orientation (same)</i>			
	GenAmE was preferred in: <i>Status, Instrumental Orientation (same)</i>			
Rhoticity		Pre	Map	rhoticity% change
	RP condition	26.43% (37/140)	22.95% (14/61)	-3.48%
	GenAmE condition	22.89% (19/83)	41.18% (35/85)	+18.29%

Table 7.11 HK20's profile of attitudes and accommodation of rhoticity. The columns of "Pre" and "Map" indicate the percentages of rhotic words in these tasks; the ratio in the brackets represents the token of rhotic words/all the r-words. The differences between these two tasks are shown in the column "rhoticity%". A negative value in the RP condition indicates a convergence, and a positive value indicates a divergence; a positive value in the GenAmE condition indicates a convergence, and a negative value indicates a convergence.

For attitudes, HK20 preferred RP on *Linguistic Attractiveness, Social Attractiveness* and *Integrative Orientation*, but preferred GenAmE on *Status*, and had no preference on *Instrumental Orientation*. Although HK20 produced most of the words in a non-rhotic manner, she still produced 20% of the words with rhoticity.

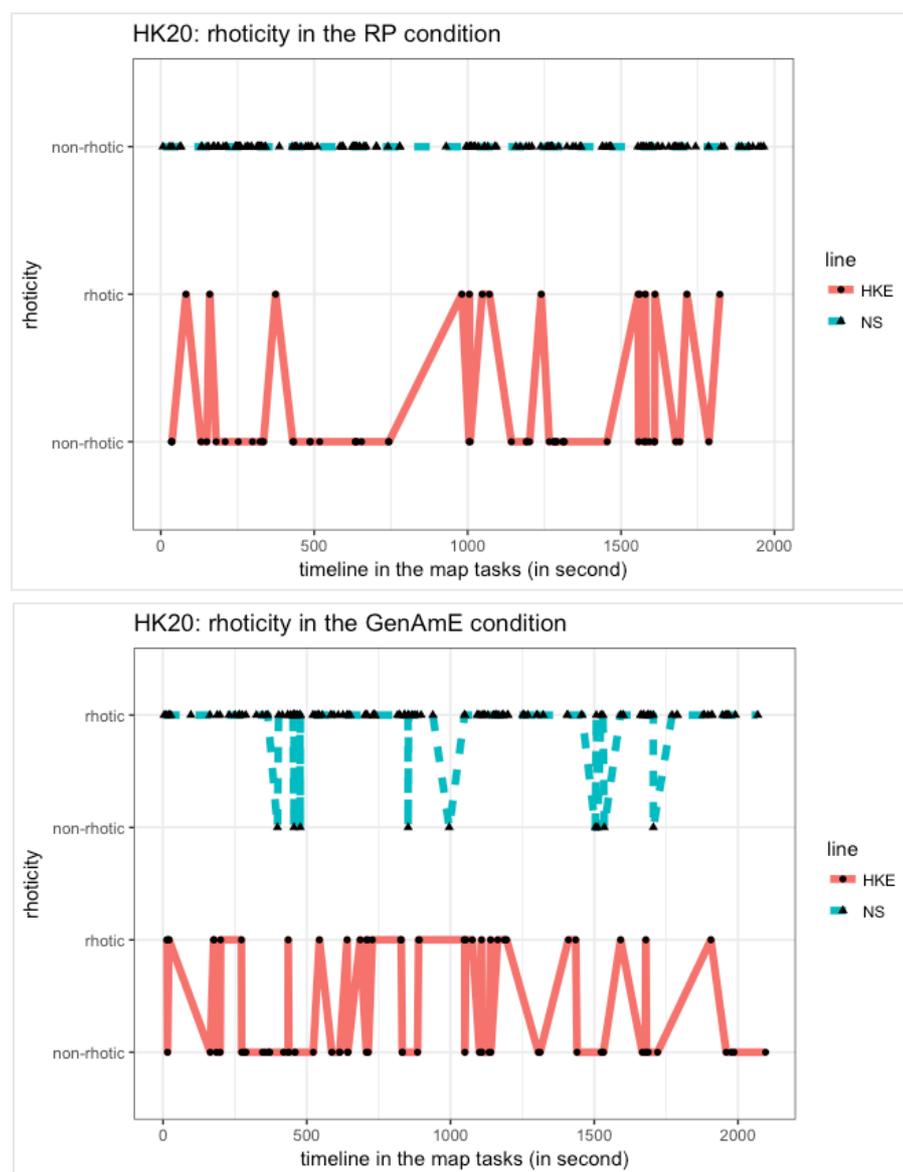


Figure 7.4 HK20 and the native interlocutors’ production of rhoticity across the timeline in the map task. The top picture represents the RP condition, the bottom picture represents the GenAmE condition. Each black circle represents a token of HK20’s r-words. If the word was produced with rhoticity, it was plotted in the “rhotic” level, if the word was produced without rhoticity, it was plotted in the “non-rhotic” level. The same was done for the native speakers’ r-words. The red lines represent the changes in rhoticity for HK20, and the blue lines represent the changes in rhoticity for the RP and the GenAmE speaker.

Figure 7.4 shows how HK20 accommodated chronologically in the map task in the two conditions. It also plots the native interlocutors’ production of rhotic words in the map task.

As shown in Figure 7.4, in the RP condition (top picture), the RP speaker did not produce any rhotic words in the map tasks. HK20 produced some rhotic words when she talked to the RP speaker; however, as Table 7.11 suggests, HK20's percentage of rhotic words in the map task (26.43%) did not change much compared to the percentage in the pre-task (22.95%). In the GenAmE condition (bottom picture), the GenAmE speaker occasionally produced some non-rhotic words: it might be that the GenAmE speaker converged towards HK20 on rhoticity. HK20 in return also produced more rhotic words in the map task (41.18%).

An episode of the conversation between them is shown below to demonstrate how accommodation occurred between HK20 and the GenAmE interlocutor.

-
- 1 **GenAmE:** and you should come to the green box on the left, called Car Wash Centre¹-[ɹ].
- 2 **HK20:** Car Wash Centre²-[ɹ]...not Spot³-[ɹ]...OK...hmm.
- 3 **GenAmE:** You have...you have what?
- 4 **HK20:** Car Wash Spot⁴-[ø].
- 5 **GenAmE:** Oh.
-

Table 7.12 An example of accommodation of rhoticity in the conversation between HK20 and the GenAmE interlocutor. Words with [ɹ] attached indicate a rhotic pronunciation, those with [ø] attached indicate a non-rhotic pronunciation.

In this conversation, HK20 and the GenAmE interlocutor were matching a landmark called “Car Wash Centre”. On the HK20’s map, the landmark was named as “Car Wash Spot”, while on the GenAmE interlocutor’s map it was marked as “Car Wash Centre”. The GenAmE interlocutor started by indicating the location of “Car Wash Centre¹-[ɹ]”. When HK20 repeated the name of the “Car Wash Centre¹”, she converged on the word “Centre²-[ɹ]”, producing it with rhoticity. Immediately after that, she added rhoticity to the word “Spot³” which is meant to be non-rhotic. The rhotic “Spot³-[ɹ]” caused confusion for the GenAmE speaker, therefore the GenAmE interlocutor questioned HK20. HK20 repeated “Car Wash Spot⁴-[ø]” to clarify, but this time she dropped the rhoticity on “Spot⁴-[ø]”. Spectrograms of “Spot³” and “Spot⁴” are provided below.

The word “Centre¹” and “Centre²” had a rhotic final vowel [ə̃], and the rhoticity of these two words cannot be observed in the spectrogram. Therefore only the spectrograms of “Spot³” and “Spot⁴” are provided in Figure 7.5. As shown in the top spectrogram, the falling F3 and the rising F2 at the end of the vowel in “Spot³” indicate a rhoticity.

This episode shows that HK20 converged towards the rhotic “centre” produced by the GenAmE interlocutor when the word was a key information in the conversation; meanwhile she over-generalised the rhoticity to the word “spot”. When the rhotic “spot” caused confusion, she dropped the rhoticity to clarify. From line 1 to line 2, HK20 seemed to adopt accommodation as a strategy to shorten the distance between her and the GenAmE interlocutor, by adding rhoticity to her production. From line 3 to line 4, when she realised that she had over-generalised the rhoticity, she then corrected her pronunciation of the rhotic “spot”.

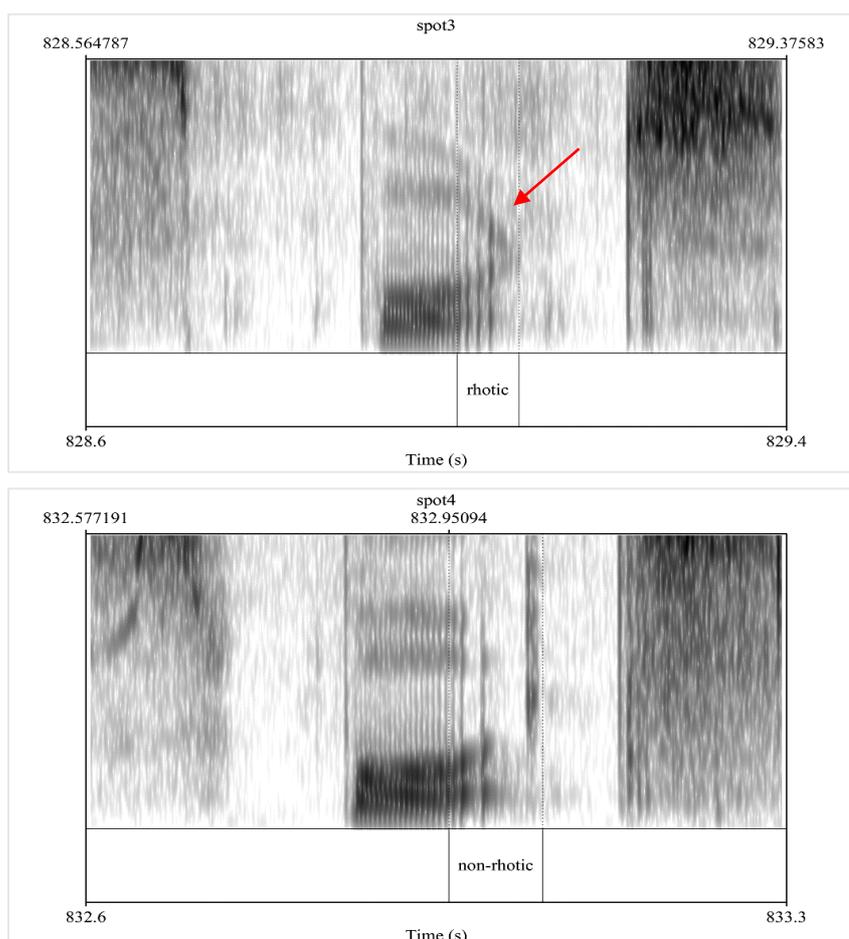


Figure 7.5 Spectrograms of the rhotic “spot” and the non-rhotic “spot” produced by HK20 in her conversation with a GenAmE speaker. The red arrow indicates the signs of rhoticity, i.e. the falling F3 and rising F2.

7.6.2. HK23: a respondent who converged more towards the RP speaker

The second type of convergence is represented by HK23. HK23 had diverse attitudes towards RP and GenAmE, as shown in Table 7.13. She preferred RP on *Linguistic Attractiveness*, *Status* and *Integrative Orientation*, while she rated GenAmE higher on *Social Attractiveness* and *Instrumental Orientation*. HK23 also had a mixture of rhoticity, i.e. she produced around 60% of the r-words with rhoticity.

Participant	HK23		Sex: female	
Attitudes	RP was preferred in: <i>Linguistic Attractiveness</i> , <i>Status</i> , <i>Integrative Orientation</i>			
	GenAmE was preferred in: <i>Social Attractiveness</i> , <i>Instrumental Orientation</i>			
Rhoticity		Pre	Map	rhoticity% change
	RP condition	67.53% (52/77)	34.69% (17/49)	-32.84%
	GenAmE condition	57.89% (33/57)	60.38% (32/53)	+2.49%

Table 7.13 HK23's profile of attitudes and accommodation of rhoticity. The columns of "Pre" and "Map" indicate the percentages of rhotic words in these tasks; the ratio in the brackets represents the token of rhotic words/all the r-words. The differences between these two tasks are shown in the column "rhoticity%". A negative value in the RP condition indicates a convergence, and a positive value indicates a divergence; a positive value in the GenAmE condition indicates a convergence, and a negative value indicates a convergence.

Figure 7.6 demonstrates HK23's accommodation of rhoticity chronologically in the two conditions. In the RP condition, the RP speaker did not produce any rhotic words, while HK23 produced more non-rhotic words in the map task (65.31% non-rhoticity) compared to the pre-task (32.67% non-rhoticity). The decrease of 30% of rhoticity from the pre-task to the map task can be seen as a convergence towards non-rhoticity.

In the GenAmE condition, most of the tokens produced by the GenAmE speakers were rhotic, except for a few words that were non-rhotic. HK23 did not change much of her production of rhoticity when she talked to the GenAmE speaker, in that she had around 60% of rhoticity in both the pre-task and the map task.

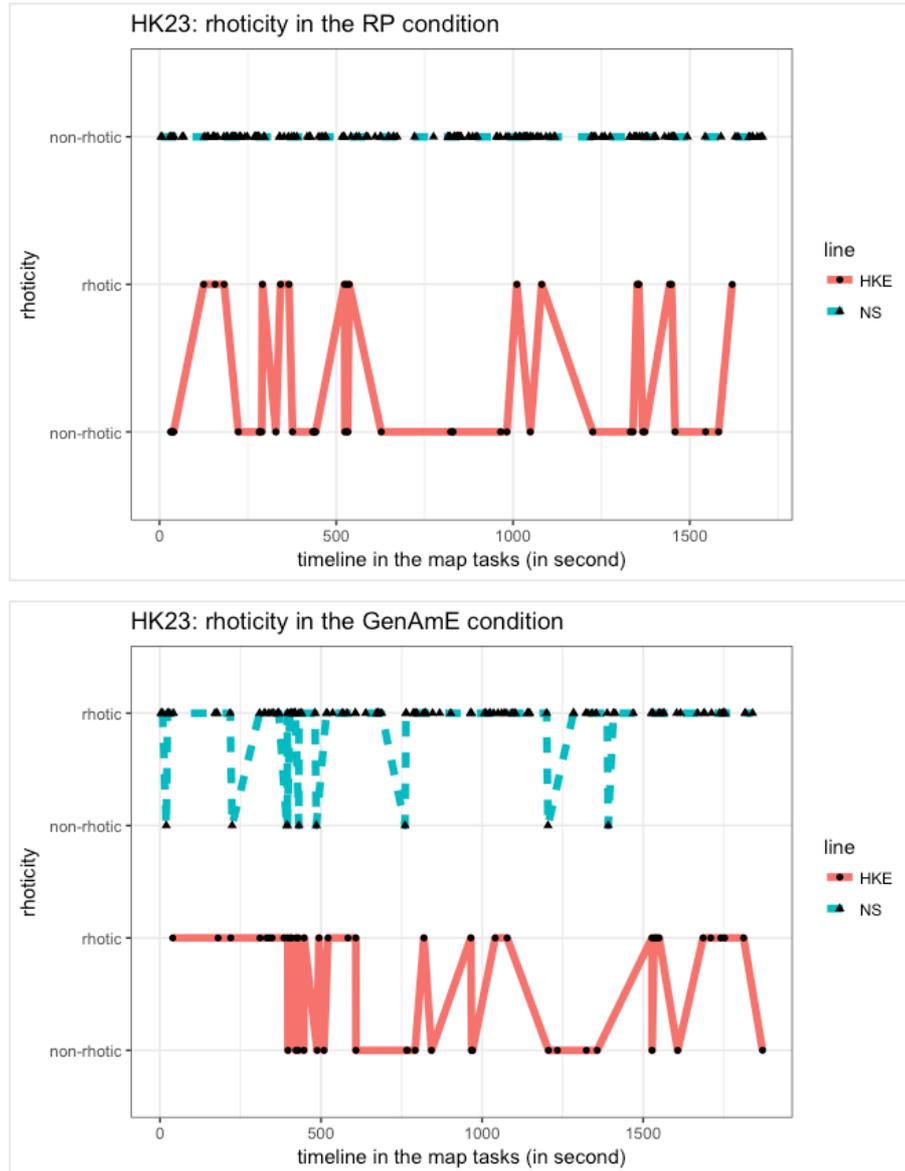


Figure 7.6 HK23 and the native interlocutors’ production of rhoticity across the timeline in the map tasks. The top picture represents the RP condition, the bottom picture represents the GenAmE condition. Each black circle represents a token of HK23’s r-words. If the word was produced with rhoticity, it was plotted in the “rhotic” level, if the word was produced without rhoticity, it was plotted in the “non-rhotic” level. The same was done for the native speakers’ r-words. The red lines represent the changes in rhoticity for HK23, and the blue lines represent the changes in rhoticity for the RP and the GenAmE speaker.

An episode of the conversation between HK23 and the RP interlocutor is shown in Table 7.14 to demonstrate how HK23 accommodated towards the RP speaker.

In this conversation, the RP interlocutor gave directions to the location of the “Thought Bar¹-[ø]”, pronouncing the “Bar¹” with non-rhoticity. HK23 responded immediately and repeated “Thought Bar²-[ɹ]”, indicating that she had found the place. At first, she

pronounced the “Bar²” with a strong rhoticity, later she repeated “Bar³-[ɹ]” with a less strong rhoticity and spelt out the word.

1	RP:	...sort of on the border between zone 11 and zone 20, to the Thought Bar¹-[ø].
2	HK23:	Thought Bar²-[ɹ] ...yep...same as... A Bar³-[ɹ] ...BR...B...A...R.
3	RP:	Yeah.

Table 7.14 An example of accommodation of rhoticity in the conversation between HK23 and the RP interlocutor. Words with [ɹ] attached indicate a rhotic pronunciation, those with [ø] attached indicate a non-rhotic pronunciation.

In line 2, after HK23 pronounced “Bar²-[ɹ]” with a strong rhoticity, she realised that her interlocutor was a non-rhotic English speaker. Therefore, to accommodate the RP interlocutor’s accent, she repeated “Bar³-[ɹ]” with a less rhoticity. Here, HK23 adopted a different strategy in accommodation. Instead of converging her rhotic “bar” to a non-rhotic “bar”, she adjusted the degree of rhoticity. In addition, she spelt out the word to further clarify her rhotic pronunciation of “bar”. This episode shows that accommodation on rhoticity is not limited to a shift between rhoticity and non-rhoticity; adjusting the degree of rhoticity is another approach.

Spectrograms and formant contours of “Bar²” and “Bar³” are shown in Figure 7.7 and Figure 7.8.

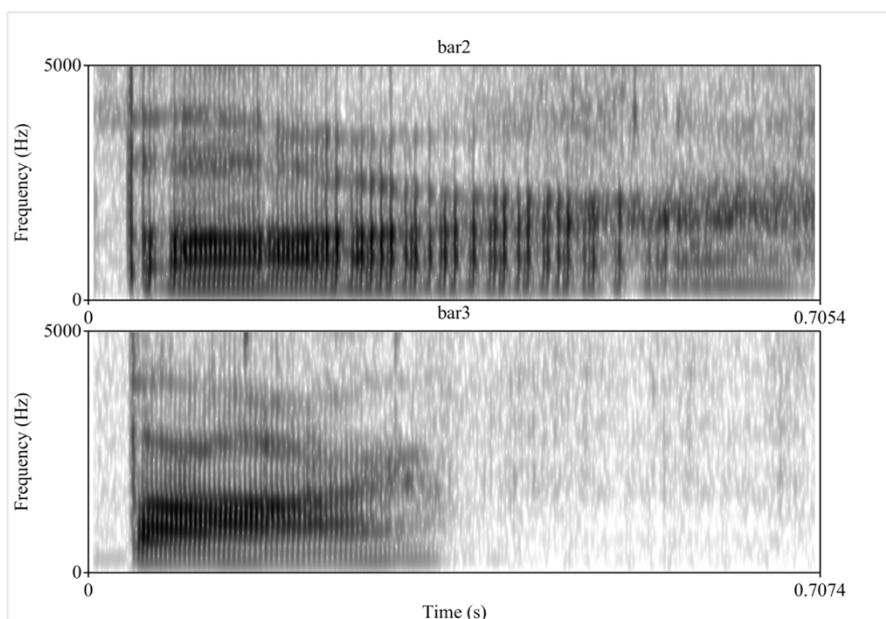


Figure 7.7 Spectrogram of “Bar²” (top) and “Bar³” (bottom) produced by HK23. The two words were extracted in a segment with the same duration.

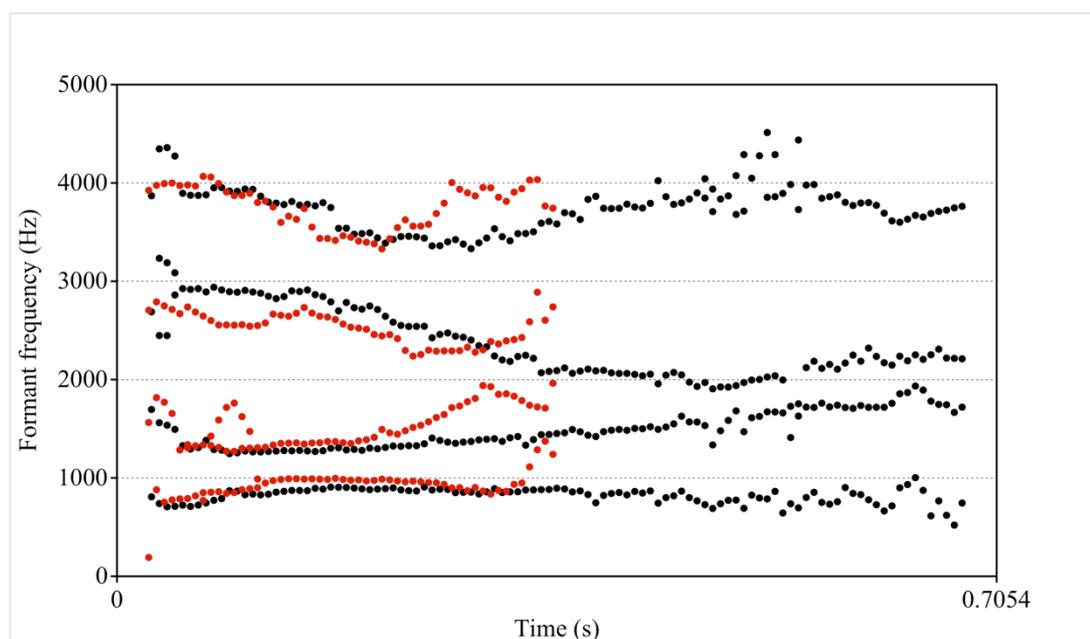


Figure 7.8 Formant contours of “Bar2” (more rhotic, in black) and “Bar3” (less rhotic, in red) produced by HK23.

As shown in Figure 7.7 and Figure 7.8, the duration of “Bar²-[ɹ]” is longer than the duration of “Bar³-[ɹ]”. In Figure 7.8, the black formant contours represent “Bar²-[ɹ]”, and the red formant contours represent “Bar³-[ɹ]”. The F3 of “Bar²-[ɹ]” started at around 2900Hz and it gradually decreased to 2000Hz at the end of the vowel (i.e. a decrease of 900 Hz). On the other hand, the F3 of “Bar³-[ɹ]” started at around 2700Hz and decreased to 2300Hz at the end of the vowel (i.e. a decrease of 400 Hz). The acoustic analysis shown in Figure 7.8 supports the auditory judgement that “Bar³-[ɹ]” was less rhotic than “Bar²-[ɹ]”.

7.6.3. HK8: a respondent who converged in both conditions

The third type of convergence is represented by HK8. HK8’s English accent showed a strong influence from American English. As shown in Table 7.15, HK8 produced around 90% of r-words with rhoticity in the pre-task. He also favoured GenAmE more than RP in most of the dimensions of attitudes, except that RP was rated more prestigious than GenAmE.

Participant	HK8		Sex: male	
Attitudes	RP was preferred in: <i>Status</i>			
	GenAmE was preferred in: <i>Linguistic Attractiveness, Social Attractiveness, Instrumental Orientation, Integrative Orientation</i>			
Rhoticity		Pre	Map	rhoticity% change
	RP condition	89.04% (65/73)	84.85% (56/66)	-4.19%
	GenAmE condition	91.86% (79/86)	95.83% (69/72)	+3.97%

Table 7.15 HK8’s profile of attitudes and accommodation of rhoticity. The columns of “Pre” and “Map” indicate the percentages of rhotic words in these tasks; the ratio in the brackets represents the token of rhotic words/all the words. The differences between the two tasks are shown in the column “rhoticity%”. A negative value in the RP condition indicates a convergence, and a positive value indicates a divergence; a positive value in the GenAmE condition indicates a convergence, and a negative value indicates a convergence.

Though HK8’s percentages of rhoticity reached a ceiling effect in the pre-task, he still showed some small changes in the map task. Figure 7.9 demonstrates HK8’s production of rhoticity in the two conditions. In the RP condition, the RP speaker produced two rhotic words as shown by the blue dotted line. The two rhotic tokens were of the same word, “car”, and they both appeared in a phrase like “the blue **car as** in zone 1”, suggesting a linking-r. Thus, these two tokens should not be taken as a convergence towards HK8 in this case. HK8 also produced some non-rhotic tokens when he talked to the RP speaker. In the GenAmE condition, HK8 only produced three tokens of non-rhotic words. This suggests that HK8 adjusted his rhoticity according to the speakers he was talking to.

HK8’s attitude preference for GenAmE did not seem to help him converge more towards the GenAmE speaker in the map tasks. This may be due to that fact that his rhoticity has reached a ceiling effect, so that he did not have the space to further increase his rhoticity.

An episode of the conversation between HK8 and the RP interlocutor is shown in Table 7.16 to demonstrate how HK8 accommodated towards the RP speaker.

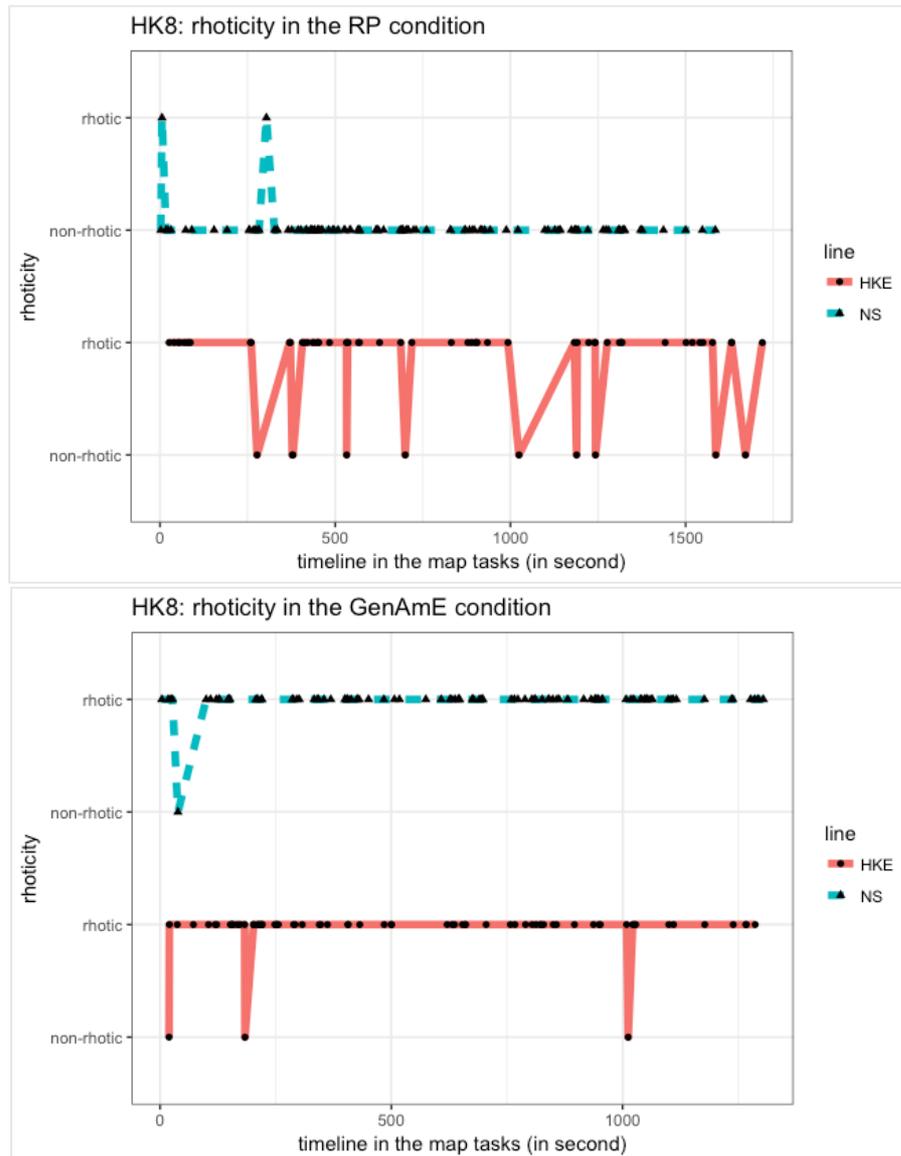


Figure 7.9 HK8 and the native interlocutors’ production of rhoticity across the timeline in the map tasks. The top picture represents the RP condition, the bottom picture represents the GenAmE condition. Each black circle represents a token of HK8’s r-words. If the word was produced with rhoticity, it was plotted in the “rhotic” level, if the word was produced without rhoticity, it was plotted in the “non-rhotic” level. The same was done for the native speakers’ r-words. The red lines represent the changes in rhoticity for HK8, and the blue lines represent the changes in rhoticity for the RP and the GenAmE speaker.

-
- 1 **RP:** And...and then you go to the **Farm¹-[ø]** Land.
 2 **HK8:** **Farm²-[ø]** Land...I have a **Farm³-[ø]** Zoo.
-

Table 7.16 An example of accommodation of rhoticity in the conversation between HK8 and the RP interlocutor. Words with [ø] attached indicate a non-rhotic pronunciation.

In this conversation, the RP interlocutor directed HK8 to the location “Farm¹-[ø] land”. HK8 repeated the name of the location “Farm²-[ø] land” with a non-rhotic accent, and told the RP interlocutor what he had on his map was a “Farm³-[ø] Zoo”, also with a non-rhotic accent on “Farm³”. At the beginning of the map tasks, HK8 pronounced the word “farm” with rhoticity, but after that, he then converged on his pronunciation of “farm” towards non-rhoticity, including the two tokens of “farm” shown in Table 7.16.

Note that HK8 pronounced “Farm²” and “Farm³” both with non-rhoticity, however, the vowel of “Farm²” was more back, close to an RP-*path* vowel; whereas the vowel of “Farm³” was more front, close to a GenAmE-*path* vowel. This example reveals the complexity of accommodation, that people might converge consistently on one dimension (e.g. HK8 converged on the rhoticity for “Farm²” and “Farm³”), but on another dimension they might behave differently.

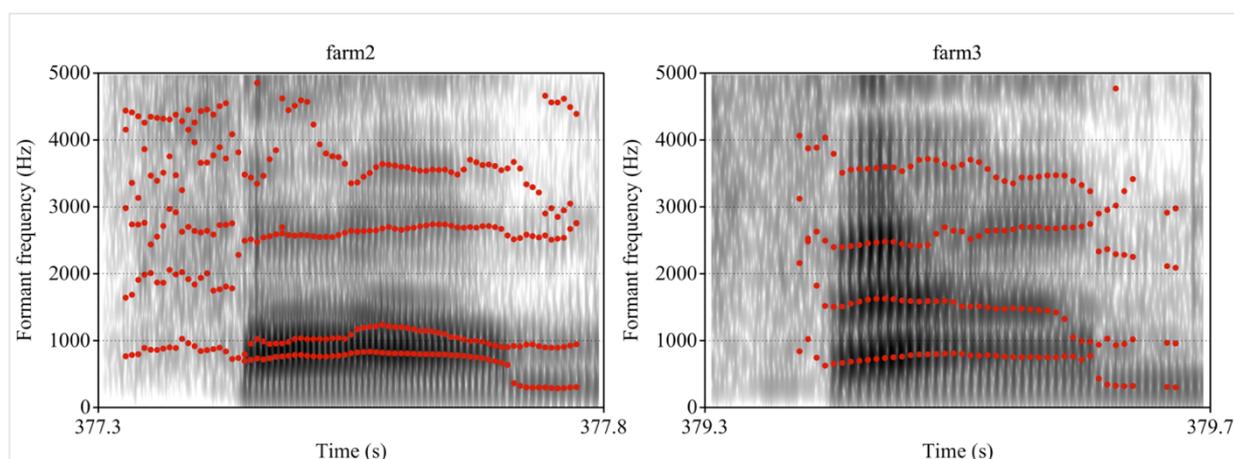


Figure 7.10 Spectrogram and formant contours for “Farm2” (left) and “Farm3” (right).

Spectrograms of “Farm²” and “Farm³” are shown in Figure 7.10. Figure 7.10 shows that the F3 values of “Farm²” and “Farm³” did not decrease at the end of the vowels, indicating that these two words were pronounced with non-rhoticity. Also, the F2 value of “Farm²” was lower than the F2 value of “Farm³”, suggesting that the vowel of “Farm²” was more back than the vowel of “Farm³”.

7.6.4. Summary

This section has reported on three examples of HKE’s accommodation on rhoticity.

The speakers' profiles suggest that the HKE speakers accommodated in their conversations with the native speakers. HK20 converged towards the GenAmE speaker by adding rhoticity to the word "centre". She also over-generalised the rhoticity to the word "spot", on which she corrected herself later. HK23 converged towards the RP speaker; however, she did not delete the rhoticity in her pronunciation of the word "bar", instead, she adjusted her degree of rhoticity of the word. HK8, who spoke English with a rhotic accent, converged towards the RP speaker by deleting the rhoticity in the word "farm". In addition, he also converged the vowel of "farm" towards the RP speaker in one token, but in another token of "farm" he only converged on rhoticity but not on the vowel.

The chronological plotting of rhoticity (see Figure 7.3, Figure 7.5 and Figure 7.8) also reveals the real-time interactions between the HKE speakers and the native interlocutors. It is interesting to find that the RP speakers rarely changed their production of rhoticity by producing all the words with non-rhoticity, while the GenAmE speakers seemed to perform differently according to the HKE speakers' levels of rhoticity. This might be that the GenAmE speakers recruited in the present study had spent some time in the UK when they completed the tasks and they might be in the process of a long-term accommodation towards British English. The RP speakers in the present study, however, might not have any motivation to accommodate towards American English in their life, nor in the experiments.

The three examples also suggest that people might respond to different aspects of language attitudes in their accommodation. HK20 might have responded to *Status*, in that she accommodated more towards the GenAmE speaker on rhoticity because she felt that GenAmE was a more prestigious accent. HK23, who rated RP as more linguistically attractive and prestigious, accommodated more towards the RP speaker on rhoticity. HK8 had a strong preference for GenAmE, however, he showed similar degrees of changes in the two conditions, suggesting that HK8 might not use language attitude as a modifier in his accommodation.

Based on the analysis above, the speakers' profiles complement the whole picture of convergence. Since the HKE speakers only talked to the native English speakers for less than one hour in the experiments, the convergence effect might not have been strong

enough to be captured using quantitative methods. The episodes of conversation between the HKE speakers and the native interlocutors reveal a qualitative side of accommodation.

7.7. Conclusion

This chapter presented the results of Study 3, trying to answer the question of whether Hong Kong people's language attitudes affect their accommodation towards native accents. Rhoticity was selected as the target variable in this chapter, because the HKE speakers showed the strongest effect of convergence on rhoticity. The quantitative results suggest that the HKE speakers' attitudes did not correlate with their accommodation of rhoticity. This might be due to the small sample size of the data (14 respondents), or was because people rely on different aspects of attitudes in their accommodation. The speakers' profiles explored the data from a different angle. Three examples were used to demonstrate the accommodation between the HKE speakers and the native interlocutors.

Chapter 8. Discussion

In the previous chapters, three studies of speech accommodation and language attitudes were reported. Although some significant results were found in these studies, it is not easy to interpret these results, because the HKE participants did not present a homogenous pattern for accommodation and language attitudes. This chapter consists of three parts. Firstly, I will argue the relationship between the salience of sounds and speech accommodation, with supporting evidence from Study 2. Secondly, I will propose a hybrid exemplar-based model (henceforth the HEM) for short-term accommodation, aiming to provide a framework for studies of speech accommodation. Finally, a few general proposals are made for future studies.

8.1. Salience and short-term accommodation

Study 2 investigated the HKE participants' accommodation on two vowels (i.e. the THOUGHT vowel and PATH vowel) and three consonants (i.e. rhoticity, fricative [z] and fricative [θ]). The results of Study 2 are complicated. For the THOUGHT vowel, the participants' changes in Euclidean distance from the pre-tasks to the map tasks in the RP condition were significantly different from the changes in the GenAmE condition. For the PATH vowel, the participants showed a divergence from the pre-tasks to the map tasks in both the RP and the GenAmE conditions. For rhoticity, convergence was found in the GenAmE condition. For fricative [z], the participants converged towards the native interlocutors from the pre-tasks to the map tasks. For fricative [θ], a marginal divergence was found from the pre-tasks to the map tasks.

Due to the HKE participants' diverse pronunciation of the five variables, the results for accommodation were different across the participants and the variables. Although it is difficult to summarise a unique pattern from the complex results, there is some evidence suggesting a correlation between the salience of sounds and accommodation. That is, **people tend to converge on the linguistic features which are more salient to them.**

Firstly, let's explain what determines the salience of linguistic features. In linguistics, there is not a consistent definition of salience and studies tend to define their own notion of salience. For example, salience refers to sociolinguistic sensitivity in Kerswill (1985),

where a stigmatized sound is more salient than a sound which does not carry any social meaning. Saliency can also mean high frequency, as in Pierrehumbert (2001, 2006), or preference, as in Mufwene (1991). In the present project, “saliency” is defined as **sounds which have a greater phonetic difference between the HKE speaker’s native repertoire and the native interlocutor’s repertoire, and sounds which carry specific social meanings.**

8.1.1 Phonetic difference

Evidence supporting this point can be found in the accommodation of the THOUGHT vowel and the PATH vowel. When calculating the relative distance of the participants’ vowels in comparison to the RP interlocutors’ and GenAmE interlocutors’ vowels, the participants who had the largest distance from the RP interlocutors tended to show a convergence in the RP condition but not in the GenAmE condition. This pattern was observed for HK1, HK8, HK19 and five other HKE participants for the THOUGHT vowel and for HK16 for the PATH vowel. Similarly, Babel (2010, 2012) suggests that people tend to converge on the vowels which have a larger acoustic-phonetic distance from the model talkers. Babel (2010) found that DRESS vowels elicited the strongest convergence effect in a study of convergence of New Zealand English, because DRESS is one of the most distinct vowels between Australian English and New Zealand English. On the other hand, Babel (2012) also found the strongest convergence on /æ/ and /ɑ/ for the participants whose dialects have most distinct /æ/ and /ɑ/ vowels from GenAmE.

Evidence can also be found from the consonantal variables. In Study 2, convergence was found on rhoticity and fricative [z], but not on fricative [θ]. This might be due to rhoticity and fricative [z] having a greater phonetic difference between HKE and RP/GenAmE than fricative [θ]. Firstly, rhoticity is a more salient feature compared to fricative [z] and fricative [θ] for HKE speakers, because convergence on rhoticity involves adding or deleting a phoneme /-ɹ/, while convergence on the fricatives only involves replacement. By converging towards the RP/GenAmE speaker on rhoticity, a HKE speaker would need to either add a phoneme /-ɹ/ (in the case of convergence towards GenAmE) or delete a phoneme /-ɹ/ (in the case of convergence towards RP). However, for the fricatives, the HKE speaker would only need to replace his/her HKE variants (i.e. fricative [s] as in *zoo*-[su:] and fricative [f] as in *three*-[fi:]) with the native variants.

Secondly, fricative [z] is more salient than fricative [θ] for HKE speakers because the phonetic differences between /z/-/s/ are larger than the differences between /θ/-/f/. For the pair of fricatives /z/-/s/, the most distinctive difference is **voicing** which is reflected in a few acoustic properties. Jongman, Wayland and Wong (2000) found that voiceless fricatives have longer fricative duration, larger amplitude and higher frequency of spectral peaks than voiced fricatives. On the other hand, the differences between the pair of fricatives /θ/-/f/ are less obvious. Jongman et al. (2000) show that fricative /f/ and fricative /θ/ are similar in fricative duration and spectral peaks, and the main acoustic difference between fricative /f/ and fricative /θ/ is in F2 transition information. Since fricatives /z/ and /s/ are significantly different in multiple acoustic cues whereas fricatives /θ/ and /f/ are only different in one cue, HKE speakers might be more likely to notice the differences between fricative /z/ and fricative /s/, which might in turn facilitate convergence.

Note that not all the studies agree that larger language distance between speakers facilitates convergence. Kim et al. (2011) argue the other way around; that is, closer language distance between talkers facilitates convergence. In their study, they found that the convergence between talkers who shared the same L1 and the same dialects was significantly larger than convergence between talkers who shared only the same L1 but spoke different dialects, and convergence between talkers who had different L1s. It is worth noting that the “language distance” in Kim et al. (2011) is different from the “acoustic-phonetic distance” in Babel (2010, 2012). The “language distance” in Kim et al. (2011) was a categorical variable which was defined by whether talkers shared the same L1 and/or same dialect, whereas the “acoustic-phonetic distance” in Babel (2010, 2012) was a continuous variable represented by the vowel distances between the talkers. The results of the present project tend to support Babel’s claim, as convergence of rhoticity and fricative [z] was found between the HKE speakers and the RP/GenAmE speakers who had different L1s.

8.1.2 Social factors

In the previous section, “salience” was defined as greater phonetic difference, which can be regarded as the language-internal explanations in Kerswill and Williams (2002)’s model of salience. In their model, they propose that language-internal explanations are

preconditions for salience, and extralinguistic factors are the centre and the cause of salience. Extralinguistic factors are defined as a combination of cognitive, social psychological or pragmatic factors in Kerswill and Williams (2002)'s model. For example, people's language attitudes towards a stigmatized sound, social relations between interlocutors, and talkers' cultural identities in the conversation can all be counted as extralinguistic factors.

In the present study, salience can also refer to sounds which carry social meanings, for example, rhoticity. In Study 2, the HKE speakers were found to converge towards the RP/GenAmE speakers on rhoticity in the map tasks. This might be also due to rhoticity/non-rhoticity being a stereotype for American English/British English. Instead of recognising that the [æ] vowel in the word "path" is a feature of GenAmE, it might be easier for HKE speakers to associate rhoticity with American English. This might explain why convergence occurred on rhoticity but not on the vowels in Study 2.

8.2. The hybrid exemplar-based model for short-term accommodation

8.2.1. The existing models of speech accommodation

The previous sections provide some explanations for Study 2. However, there are still some results that have not been explained. Can the existing models of speech accommodation such as Communication Accommodation Theory (CAT) and the interactive-alignment model explain the rest of the results?

Communication Accommodation Theory (CAT) proposes that speakers converge in order to shorten their social distance from their interlocutors. According to CAT, if a HKE speaker likes the RP accent, he/she would be more likely to converge towards the RP interlocutor in their conversation. However, this hypothesis is not supported by the results of Study 3. In Study 3, the HKE participants' attitudes towards the RP/GenAmE accent did not correlate with their convergence on rhoticity. In other words, the HKE participants who liked American English more did not seem to have a stronger convergence on rhoticity when they talked to the GenAmE speakers in the map tasks. Moreover, the HKE participants who preferred American English converged differently across the five variables. For example, HK8, who preferred American English the most out of all the participants, converged towards the GenAmE interlocutor on rhoticity,

fricative [z] and fricative [θ], but not on the THOUGHT vowel and the PATH vowel. If people converge towards the accent they like, why would they only converge on some sounds towards the GenAmE interlocutor but not on others? Apparently, CAT cannot explain the diverse patterns of accommodation observed across participants and variables.

On the other hand, the interactive-alignment model suggests that convergence occurs automatically between interlocutors. According to the interactive-alignment model, the HKE speakers would automatically converge towards the native interlocutors they talk to. The results of Study 2 did not support this hypothesis. At the group level, the HKE participants were found to diverge from the native interlocutors on the PATH vowel. At the individual level, a few HKE participants did not produce any fricative [z] or fricative [θ] across the pre-tasks, the map tasks and the post-tasks. These participants did not seem to automatically converge towards the native interlocutors on these sounds.

When CAT and the interactive-alignment model are used to explain the results of Study 2 and Study 3, it seems that neither of them can account for the complicated results found in these studies. One reason might be that these theories are too general to explain complicated speech phenomena in actual conversations. Another reason might be that individual differences between the HKE speakers in Study 2 and Study 3 are too large, and that these theories cannot explain the diverse findings at the individual level.

While CAT explains why accommodation occurs and the interactive-alignment model explains how accommodation occurs, these models only provide explanations for a very small part of accommodation. An ideal model of accommodation should include both perception and production. More importantly, it should allow social factors to affect the process. The next section will propose a hybrid exemplar-based model for accommodation.

8.2.2. The hybrid exemplar-based model

To fill this gap, the hybrid exemplar-based model (HEM) is proposed for short-term accommodation. The HEM is built on the perception and production models of exemplar-based theories (Johnson, 1997; Pierrehumbert, 2001, 2003). The HEM is a preliminary model which aims to provide a framework for speech accommodation. Although the

classic study of Goldinger (1998) was designed to examine the exemplar theory in speech imitation, the study focused on the frequency effect of the exemplar theory and the role of social factors was not considered. Babel (2012) mentioned the exemplar-based theories in discussion; however, Babel did not expand the theory in full to speech accommodation. Therefore, this section will explain the HEM with examples from Study 2 and Study 3.

The HEM consists of three stages: (1) the perception stage; (2) the stage of distributions update; and (3) the production stage. These three stages demonstrate the process of short-term accommodation from perception to production, and how social factors affect the process. Figure 8.1 and Figure 8.2 show the three stages of the HEM.

8.2.1.1. The three stages of the HEM

Stage 1: Perception

Stage 1 shows how a HKE speaker (for the sake of illustration, the HKE speaker is called *Kevin*) perceives the input from a GenAmE speaker (for the sake of illustration, the GenAmE speaker is called *Sophie*). When Kevin hears Sophie speaking, his perception is not exactly the same as Sophie's production. In order for Kevin to understand Sophie, her speech goes through Kevin's perceptual mechanism. This mechanism follows the perception model of Johnson (1997). Johnson (1997) proposes that when people perceive sounds, they compare the new input items with previously stored exemplars. The most similar exemplars are activated and the new items are then categorised into the activated exemplars. For instance, when Kevin hears [æ] from Sophie, he compares the acoustic cues of Sophie's [æ] with all the vowels he stores (e.g. [i], [u] and [æ]) based on his prior experience. As the input vowel shares the greatest similarity to his stored-[æ] compared to other vowels such as [i] and [u], the input vowel is recognized as [æ]. During this process, social labels associated with Sophie such as "the native interlocutor", "female", "young" are also attached to the input vowel and stored in Kevin's memory. This is shown in Figure 8.1.

Apart from recognizing sounds, the perceptual mechanism also adjusts the weight of the input based on the **salience of the sounds**. This connects to the discussion in section 8.1, that is, salience is determined by phonetic difference and social factors. For example, Kevin pronounces the THOUGHT vowel as [ɔ], which is very different from Sophie's

THOUGHT vowel (i.e. [ɑ]). In this case, Kevin is more likely to notice Sophie's THOUGHT vowel due to the large phonetic difference between [ɔ] and [ɑ]. The weight of the THOUGHT vowel is greater than other vowels that are less salient to Kevin.

On the other hand, social factors might also affect how people perceive the sounds. For example, rhoticity is a linguistic marker for American English. Kevin is more likely to notice the rhotic feature in Sophie's speech than other features of the American accent. If Kevin has a strong attitude towards the American accent (e.g. he likes the American accent very much or he hates the American accent), he would also be more likely to notice the rhoticity and other features of the American accent in Sophie's speech. These social factors increase the weight of rhoticity when Kevin perceives it. In Figure 8.1, the weight of each sound is represented by the size of the dots. The larger the dots are, the greater the weight of the sounds would be.

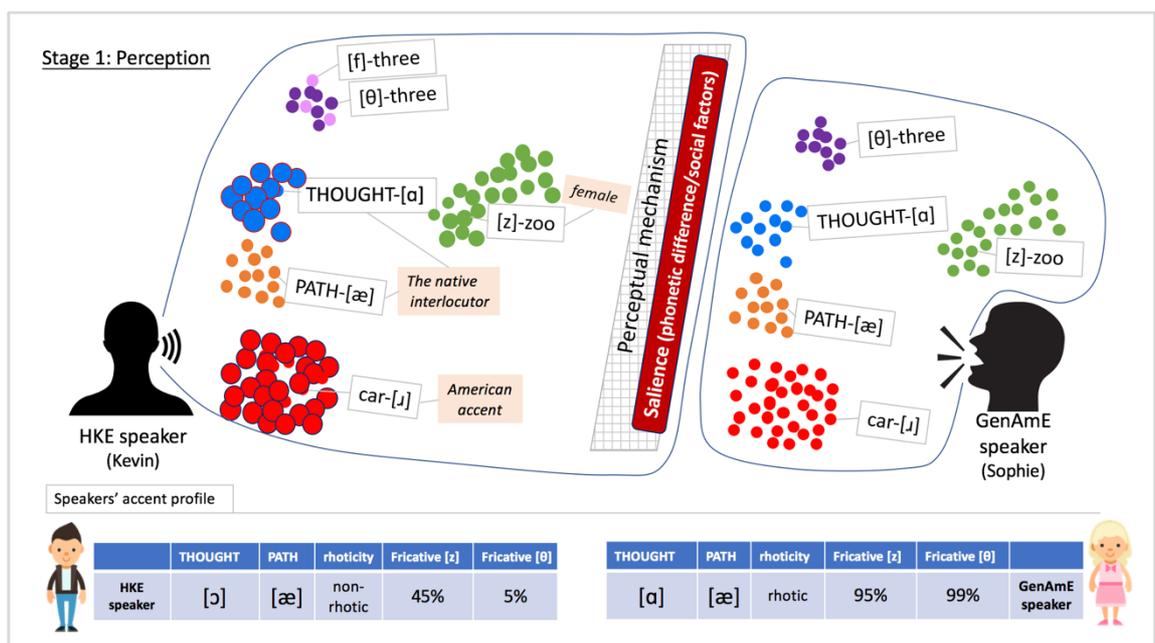


Figure 8.1 The illustration of Stage 1 in the HEM. On the left-hand side, the dots represent the target sounds perceived by Kevin; the sounds are associated with different labels based on the input. The larger the size of the dots, the greater weight of the sound in Kevin's memory. On the right-hand side, the dots represent Sophie's production.

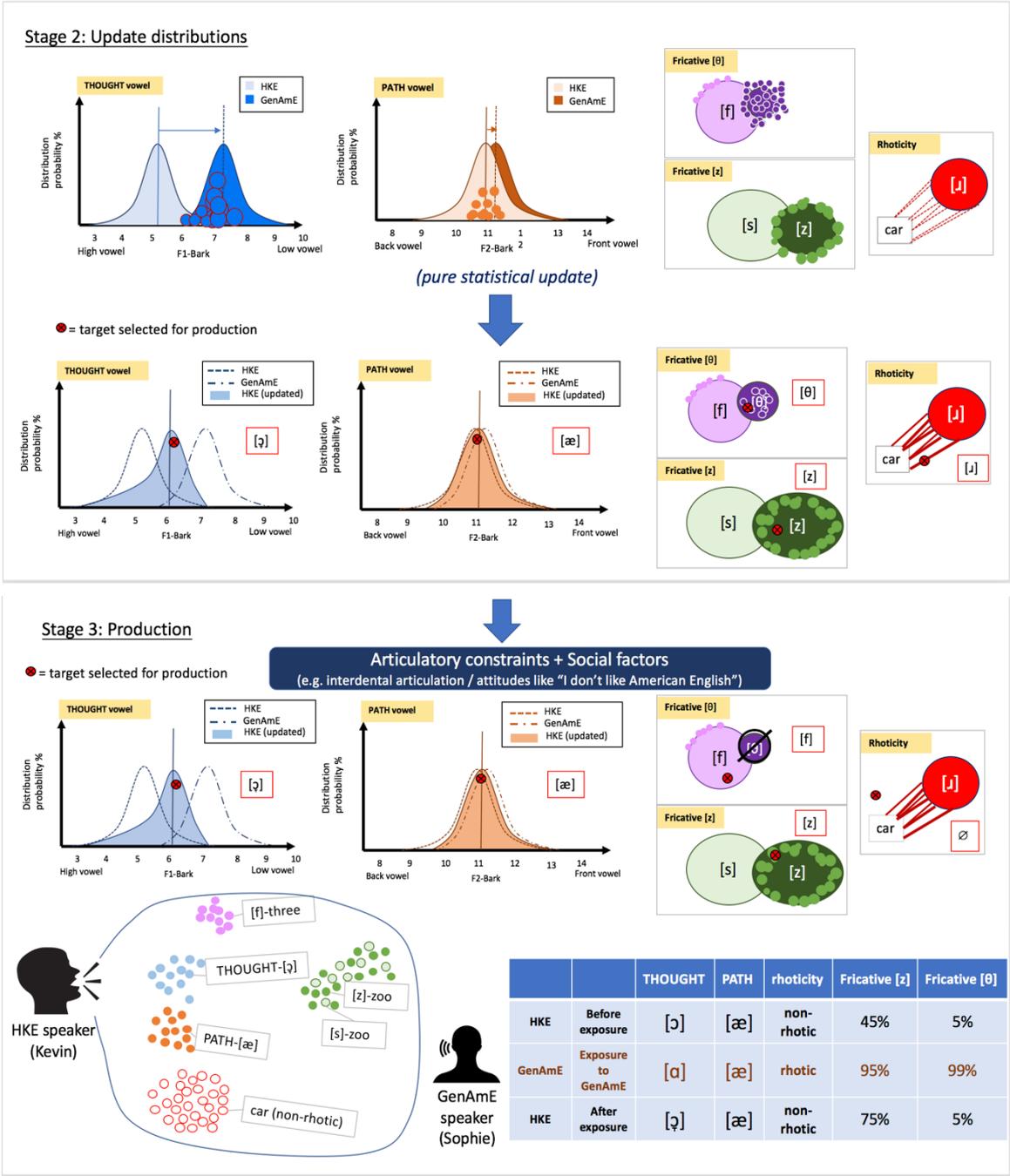


Figure 8.2 Illustration of Stage 2 and Stage 3 in the HEM. For the vowels, the distribution probability represents the probability of different scales of F1/F2 being selected as the production goal. For the consonants, the circles represent different categories of the sounds. The larger the circle, the higher probability of this category being selected.

Stage 2: Update distributions

Before Kevin talks to Sophie, he establishes the distributions for the THOUGHT vowel, the PATH vowel, and for rhoticity, fricative /z/ and /θ/ based on his prior experience. For instance, he had received thousands of input tokens such as [su:] for the word “zoo” and [sɪp] for the word “zip” from his surrounding environment, where people speak HKE. On the other hand, he had also received some amount of native input like [zɪɹʊθ] for the word “zero” and [zi:bɹə] for the word “zebra” from his English teachers or the media. The distributions of the z-sound are shaped by his prior experience of fricatives /z/ and /s/.

When Kevin perceives Sophie’s speech in Stage 1, the input he receives would update his prior distributions of sounds. As some sounds might be weighted more than others, the update of distributions would be affected by not only the quantity of the input, but also the quality (e.g. weight) of the input.

In Figure 8.2, the picture of Stage 2 shows how Kevin’s distributions of the five sounds update. For the THOUGHT vowel, the distribution of F1 is used to demonstrate the process of updating. The distribution probability of F1 presents the probability of different scales of F1 being selected in production. For example, the distribution probability of HKE’s F1 is ranged from 3 Barks to 7 Barks and the peak of the distribution is at 5 Barks. This suggests that HKE speakers’ THOUGHT vowels might have F1 values ranging from 3 to 7 Barks, but the probability of having a THOUGHT vowel with F1 in 3 Barks is relatively low, whereas the probability of having a THOUGHT vowel with F1 in 5 Barks is the highest.

Before exposure, the distribution of Kevin’s F1 for the THOUGHT vowel ranged from 3 Barks to 7 Barks, with the peak of distribution at 5 Barks. When Kevin receives the input from Sophie, whose THOUGHT vowel ranges from 6 Barks to 10 Barks, the peak of distribution of Kevin’s F1 moves from 5 Barks to 6 Barks. Similarly, for the PATH vowel, Kevin’s distribution of F2 is updated based on the input he receives. As Kevin’s PATH vowel is very close to Sophie’s PATH vowel, the distribution of F2 does not change too much even after the update.

The update of the consonants is more complicated than the update of the vowels. Figure 8.2 shows that before Kevin talks to Sophie, he has developed two categories for fricative

/z/ based on his prior experience: the [z]-category and the [s]-category. As a HKE speaker who receives more input of HKE from his surrounding environment, Kevin's [s]-category might be more dominant than his [z]-category, represented as a larger circle of [s] in Figure 8.2. In the case where a speaker has developed both a [z]-category and a [s]-category, the update of distribution for the z-sound would be similar to the update of the vowels. That is, Kevin receives input of [z] from Sophie and updates his distribution of the [z]-category. As a result, the possibility of [z] being selected is increased.

However, not all the HKE speakers would have developed both the [z]-category and the [s]-category for fricative /z/ based on their prior experience. For example, there are a few HKE participants in Study 2 who did not produce any fricative [z] across the three tasks. For these HKE participants, they might only have the [s]-category for fricative /z/. If this is the case, even though these participants receive input of fricative [z] from native speakers, they might still not be able to accommodate on this sound. Similarly, for the HKE participants who only have the [f]-category and did not establish the [θ]-category for fricative /θ/, they might not be able to accommodate on fricative /θ/ either.

The update of rhoticity is slightly different from the update of the fricatives. For HKE speakers, it is possible that some of them might not have the [z]-category or the [θ]-category for the fricatives. However, [ɹ]-category should not be a problem for HKE speakers. For example, they have no difficulty in pronouncing /ɹ/ in words like “really” and “right” (Hung 2000). In this case, rhoticity can be seen as establishing an association between a word and the [ɹ]-category in syllable final position. For those HKE speakers who speak English with some degree of rhoticity, they might have established the associations between some words (e.g. “car” [kɑɹ]) and the [ɹ]-category in syllable final position. For those who speak a non-rhotic accent, this kind of association might not be built.

In other words, what is updating in Stage 2 for rhoticity is not the [ɹ]-category itself (because HKE speakers already have acquired it), but the associations between [ɹ] and words. For example, in Figure 8.2, Kevin does not speak English with rhoticity before he talks to Sophie. Therefore, the word “car” does not associate with [ɹ] based on his prior experience. When Kevin hears Sophie pronounce “car” as [kɑɹ], the association between “car” and [ɹ] starts to establish. The more rhotic-“car”s Kevin receives from Sophie, the

stronger the association that is built. In contrast, the HKE participants who produced the word “car” with rhoticity in their pre-tasks would disassociate the connection between “car” and [ɹ] when they received the non-rhotic input from the RP speakers.

The top picture of Figure 8.2 shows Kevin’s statistical update of distributions when he receives input from Sophie. If accommodation is all about a pure statistical update of distributions based on the input of perception, then the HKE participants should demonstrate a convergence on all the five variables. However, this is not the case. In Study 2, even for rhoticity and fricative [z] (where convergence was found at the group level), not all the participants showed a convergence on these two variables. The correlation between perception and production in accommodation is not straightforward. For example, Kim et al. (2011) found that Korean learners of English did not accommodate towards American English speakers after a short exposure to American English. Evans and Iverson (2007) found that students from the north of England accommodated their BATH and BUD vowels to be more southern after they studied at a university in the south of England for two years. Their perception of these two vowels, however, did not change much.

If accommodation is not about a pure statistical update of distributions, what are the other factors that might be relevant? Stage 3 will try to address to this question.

Stage 3: Production

According to the production model of the exemplar-based theory, Pierrehumbert (2001, 2003) suggests that when people plan for production, the neighbouring regions of the exemplars of the production goal are activated. The average properties of the activated regions constitute the production goal. For example, if Kevin plans to pronounce the word “zoo”, the distribution of z-sounds is activated for the initial consonant of “zoo”. As Kevin has updated his distribution of z-sounds in Stage 2, his selection of the production goal is based on the updated distributions. As a result, [z] is more likely to be selected as the production goal after the update. At this point, the selection of the production goal is still based on statistical distributions.

However, the selection of the production goal might be affected by articulatory constraints and be overwritten by social factors. For example, if Kevin does not acquire

the articulatory cues of fricative /θ/ (e.g. move the tongue tip in between the upper and lower teeth), even if he has updated the distribution of fricative /θ/ in his perception, he still might not be able to pronounce fricative /θ/. The articulatory constraints might also apply to some HKE participants who could not voice fricative [z] in their production. Articulatory cues of some sounds are not likely to be acquired through perception alone.

Social factors, on the other hand, would also affect the selection of the production goal. For example, even if Kevin has established the association between “car” and rhoticity in Stage 2, if he has a strong attitude towards American English such as “I hate American English”, he might still pronounce the word “car” without rhoticity. That is, his attitude towards American English would overwrite his selection of the production goal based on the updated distributions.

At this point, the HEM is explained through the example of Kevin and Sophie in Figure 8.1 and Figure 8.2. For the sake of illustration, Kevin’s accommodation patterns for five variables and his language attitudes are manipulated. In the following sections, I will try to use the HEM to explain some of the results of Study 2 and Study 3.

8.2.1.2. Salience and the HEM

Salience is an important element in the perception mechanism of Stage 1. It decides how the input is weighted in the perceiver’s memory. In section 8.1, salience is defined as sounds which have a greater phonetic difference between the HKE speaker’s native repertoire and the native interlocutor’s repertoire, and sounds which carry social meanings. The reason why people tend to accommodate on the linguistic features which are more salient to them is that these salient sounds are weighted more in people’s memory, and in turn they contribute more to the update of distributions in Stage 2.

For instance, when calculating the relative distances of the THOUGHT vowel between each HKE speaker and the native interlocutors, those participants (e.g. HK1, HK8, HK19, HK7, HK13, HK2 and HK23) who had the largest distance from the RP speakers showed a convergence in the RP condition. For these HKE participants, when they perceived the THOUGHT vowel from the native interlocutors, due to the large phonetic difference, they might be more likely to notice the RP-*thought* vowel in the RP condition compared

to the GenAmE-*thought* vowel in the GenAmE condition. The input of the THOUGHT vowel from the RP speakers would be weighted more in their memories. The similar pattern was also observed on HK9 who had the largest distance from the GenAmE speaker on the THOUGHT vowel. He showed a convergence in the GenAmE condition but not in the RP condition.

The question remaining is what is the threshold of phonetic difference for accommodation? For example, for the THOUGHT vowel in Study 2, the pattern of converging towards the native speaker who was further away was only observed when the phonetic difference between the HKE speaker and the native interlocutor was larger than 50Hz. When the phonetic difference was below 50Hz, the HKE participants did not seem to differentiate their accommodation patterns based on the exposure conditions. The threshold of phonetic difference might not be the same for different vowels. For example, the threshold of phonetic difference for the PATH vowel was 80Hz. More studies would be needed in order to answer this question.

Another factor that might determine the salience of sounds is social factors. In Study 2, the HKE participants showed a convergence on rhoticity from the pre-tasks to the map tasks ($SE = 0.284, z = -3.164, p = .006$), and from the pre-tasks to the post-tasks ($SE = 0.237, z = -2.66, p = .031$) when they talked to the GenAmE interlocutors. This might be due to rhoticity being a linguistic marker of American English. When the HKE participants heard the GenAmE interlocutors speaking, instead of noticing the /æ/ vowel as one of the features of American English, they might be more likely to notice the rhoticity in the interlocutors' speech.

8.2.1.3. Distributions in the HEM

Stage 2 might be most challenging part in the HEM. Depending on whether the participants have established the categories of the target sounds, the update of distributions would be different. This is shown in the results for consonants in Study 2.

If the participants have not established the category of the target sound based on their prior experience, the distribution of the target sound is unlikely to be updated. As a result, accommodation might not occur. For example, in Study 2, HK6, HK19 and HK7 did not

produce any fricative [z] in their pre-tasks, and even after their exposure to the native interlocutors, they still did not produce any fricative [z] in the map tasks or in the post-tasks. The similar pattern was observed for HK1 as well, who did not produce any fricative [θ] across the tasks.

If the participants have established the category of the target sound, they will update the distribution of the target sound depending on the input they receive. In Study 2, the mixed group for fricative [z] (i.e. the HKE participants whose percentage of fricative [z] in the pre-tasks was between 15% to 55%) showed a convergence from the pre-tasks to the map tasks. Similarly, all the participants in the heavy-rhotic group for rhoticity (i.e. the participants whose percentage of rhoticity was over 50%) showed a convergence in both the RP and GenAmE condition.

Note that no test was run in Study 2 to examine the HKE participants' prior established categories of the five variables. The only cue to estimate their established category (categories) for each variable is their pronunciation in the pre-tasks. For example, if a HKE participant produced both fricative [s] and fricative [z] in his/her pre-task, it is very likely that he/she had both [s]-category and [z] category for fricative /z/. However, the estimation has its limits, as a few participants (e.g. HK2, HK20, HK1, HK21, HK23) who did not produce any fricative [z] in their pre-tasks also showed a convergence when they were exposed to the native speakers in the map tasks. How to estimate the speaker's prior established distributions of a sound is a challenge for the HEM.

Another piece of supporting evidence for the update of distributions is that the HKE participants tend to converge on the variables in which they received more input. In Study 2, each variable contained around 50 labels in the three maps. However, a strict control of the amount of input the HKE speakers received on each variable was not possible in the spontaneous conversations. Table 8.1 shows the average number of tokens of native words the HKE speakers received and the average output they produced for each variable in the two map tasks. For example, for the THOUGHT vowel, the HKE speakers on average received 126 tokens of native-*thought* vowels and they produced an average of 67 tokens of HKE-*thought* when the data of the two maps was combined. As shown in Table 8.1, the variables that were most frequent in the input are rhoticity and fricative [z], which are also the two consonantal variables that showed an effect of convergence. This

suggests that the greater the input the HKE speakers received in their exposure to the native speakers, the more likely it was that they would converge on the sounds.

	THOUGHT	PATH	Rhoticity	Fricative [z]	Fricative [θ]
Input from NS	126 (24)	130 (22)	266 (37)	187 (34)	140 (20)
Output of HKE	67 (13)	68 (15)	125 (19)	88 (22)	123 (19)

Table 8.1 The average number of tokens of input and output that the HKE speakers received and produced in the two map tasks for the five variables. The number represents the tokens of words. Standard deviations are shown in brackets. The shaded variables are the variables that showed convergence in the results section. “NS” means the native speakers.

It is also worth mentioning that the new distributions resulting from the updates might be temporary in short-term accommodation. In Study 2, most of the significant accommodation found from the pre-tasks to the map tasks did not last to the post-tasks. For example, a significant convergence was found on fricative [z] at the group level from the pre-tasks to the map tasks; however, no significant change was found from the pre-tasks to the post-tasks. It could be that the amount of input people receive in short-term accommodation might not be enough to elicit a long-lasting change in their accents. If speakers consistently update their distributions of target sounds through conversations, with a sufficient amount of input, a long-term accommodation might occur.

8.2.1.4. Language attitudes in the HEM

Previous studies of speech accommodation showed that social factors such as language attitudes (Babel, 2010, 2012) and the role of a speaker in a conversation (e.g. information giver or receiver) (Pardo, 2006) would affect speech convergence. This section will explain how social factors affect speech accommodation in the HEM and will try to explain the results of Study 3 in the HEM.

In the HEM, social factors could affect speech accommodation in Stage 1 and/or in Stage 3. In Stage 1, social factors are one of the elements which determine the salience of sounds. The more salient the sounds are that speakers perceive, the greater the weight given to the sounds in the input. In Stage 3, social factors might overwrite the updated distributions and change the selection of production goal. Let’s use Babel (2012) as an example. Babel (2012) found that the more attractive the white model talker was rated, the more likely the female American participants were to converge towards him. According to the HEM,

one possibility is that when the female American participants perceived sounds from the white model talker, those who liked the model talker more would add greater weight to the model talker's input. As a result, these female participants would accommodate more towards the white model talker. Another possibility is that the female American participants who liked the white model talker might change their selection of production goal to be more like him in Stage 3.

In Study 3, no correlation was found between language attitudes (i.e. Linguistic Attractiveness and Integrative Orientation) and the HKE speakers' convergence on rhoticity. This result does not seem to match the HEM's prediction. One explanation could be that what matters to convergence is the HKE speakers' attitudes towards the interlocutors (like the attitudes elicited in Babel (2012)) rather than their attitudes towards the accent the interlocutors speak (e.g. the attitudes elicited in Study 3). The attitudes towards a specific interlocutor might have a more direct and stronger impact on the HKE speakers' performance of convergence. Figure 8.3 demonstrates the differences between the attitudes towards the interlocutor and the attitudes towards the accent the interlocutor speaks using the example of Sophie and Kevin again.

Before talking to Sophie, Kevin does not know what accent his interlocutor speaks. He is only told that she is a native speaker of English. When Kevin and Sophie start to communicate, Kevin hears Sophie speaking English with rhoticity, pronouncing the word "path" with the [æ] and pronouncing the word "thought" with the [ɑ] vowel. These phonetic cues, primarily, would be connected to the label "the native interlocutor Sophie". That is, it is **Sophie** who speaks English like this, not someone else. During the conversation, Kevin might have some impressions about Sophie, for example, "she is clever but a little bit arrogant" or "She sounds formal and old-fashioned" etc. These can be regarded as Kevin's attitudes towards Sophie, the interlocutor. If Kevin has enough experience and knowledge in English accents, he might be able to tell that Sophie speaks the GenAmE accent. If Kevin has a specific attitude towards GenAmE accent, such as "people who speak the GenAmE accent are informal, friendly, intelligent and well-educated", only when he successfully associates "**the native interlocutor Sophie**" with the label "**GenAmE accent**", would his attitudes towards the GenAmE accent be activated. Otherwise, even if Kevin has a strong attitude towards the GenAmE accent, his production would not be affected by his attitudes towards the GenAmE accent. In that

case, Kevin’s attitudes towards the interlocutor might have a more direct impact on his production.

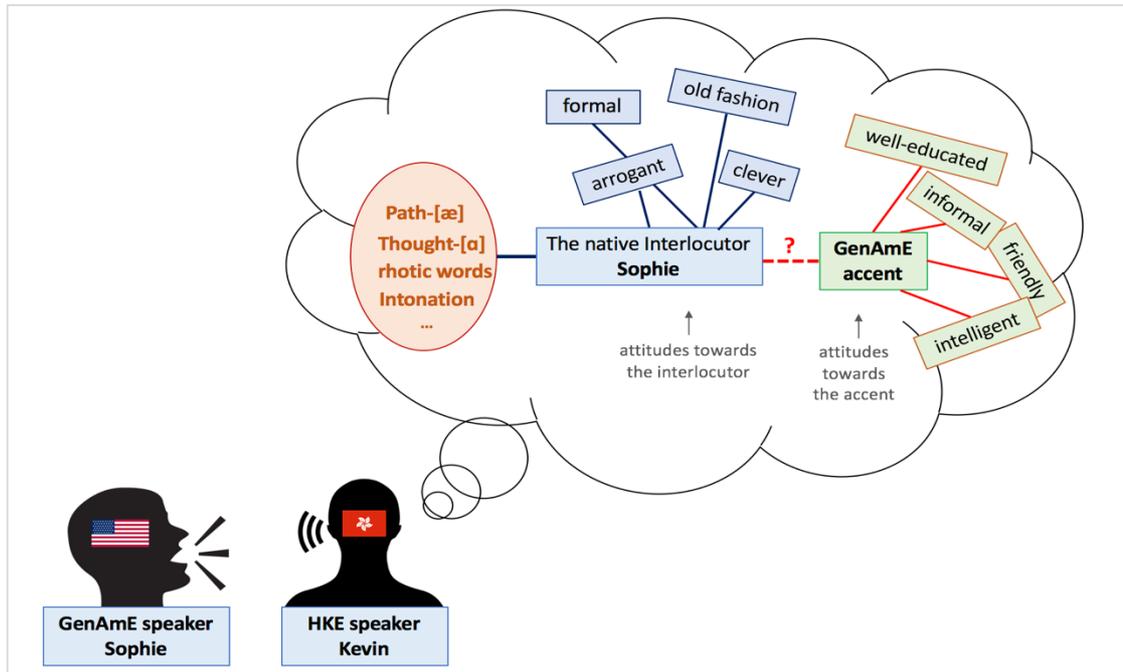


Figure 8.3 Illustration of how different types of attitudes represented in the HEM. The orange circle represents the phonetic cues the HKE speaker perceives from the GenAmE speaker; the blue box associating with a few blue labels of attitudes represent the HKE speaker’s attitudes towards the interlocutor; the green box associating with a few green labels of attitudes represents the HKE speaker’s attitudes towards the GenAmE accent. A solid line represents a strong association, whereas the dotted red line represents an uncertain association.

What remains unknown in Study 3 is whether or how well the HKE speakers associated the native interlocutors with their corresponding accents. Though the HKE completed an identification task of native English accents in Study 3 (see Chapter 7.5.1), the recordings used in the identification task were produced by a different group of RP/GenAmE speakers who did not participate in the map tasks. Therefore, it is possible that the HKE speakers in Study 3 might have associated the native interlocutors they talked to in the map tasks with the wrong accent, or they might not have built up the association between the interlocutor and a specific accent at all. This might explain why the HKE speakers’ language attitudes towards RP/GenAmE did not affect their convergence.

Although only language attitudes are discussed here, other social factors such as a speaker’s cultural identity, gender and social class might also affect how people accommodate. Speakers might also be affected by more than one social factor. It will be

interesting to explore the hierarchical effects of social factors (if there is any) on speech accommodation. For example, if speech convergence was affected by both cultural identity and language attitudes, would one factor affect the speech convergence more or earlier than the other?

8.2.1.5. Limits of the HEM

Compared to CAT and the interactive-alignment model, the HEM provides a framework of speech accommodation which covers both perception and production, and includes social factors as important elements in the model. However, as a preliminary model, there are still many questions that the HEM cannot answer.

Firstly, it is difficult to explain divergence in the HEM. In Study 2, according to the HEM, the HKE speakers would update their distributions of the PATH vowel and fricative [θ] in Stage 2, and converge towards the interlocutors on these sounds. However, the results of Study 2 suggested that the HKE participants diverged on the PATH vowel from the pre-tasks to the map tasks and a marginal divergence was also found on fricative [θ]. One explanation could be that the HKE participants who are not good at distinguishing /f/ and /θ/ might perceive some tokens of fricative [θ] as fricative [f]. As a result, the [f]-category would be updated instead of the [θ]-category. Since perceptual tests were not included in Study 2, it would be difficult to test how the HKE speakers perceive /f/ and /θ/ based on the present results.

Secondly, in order to assess what the distributions would be after the update, one would need to know the speaker's prior-established distributions and how the inputs are weighted in Stage 1. However, it is not easy to assess these two elements precisely as there are too many variables (e.g. different social factors) that might influence them. For instance, for two HKE speakers who both show a positive attitude towards American English, one might add the weight to rhoticity in the input while the other might add the weight to the THOUGHT vowel. Because of this, it is difficult to make a precise prediction based on the HEM.

The HEM does not aim to answer all the questions in speech accommodation. For sociolinguists, the HEM helps them to understand how speech accommodation occurs;

for psycholinguists, the HEM explains to them how social factors might affect the process of speech accommodation. With the present results of Study 2 and Study 3, some evidence has been found to support the claims of the HEM. However, there are still some parts of the HEM that cannot be explained by the present project. These will be left to future studies.

8.3. Future directions

Apart from the HEM, divergence and the correlation between language attitudes and accommodation also deserve more attentions in future studies.

Divergence is rarely discussed in many studies of speech accommodation. However, the results of Study 2 suggest that divergence was as common as convergence in the HKE speakers' conversations with native English speakers. In Study 2, the HKE speakers were found to diverge on the PATH vowel and fricative [θ]. Similarly, in Kim et al. (2011), non-native speakers in 2 pairs (out of 8) were found to have diverged from the native interlocutor, and in 3 pairs were found to have made no changes in the conversation. More studies are needed to address divergence. Sharma (2018) proposes that speakers might diverge from the interlocutor's speech style and shift to their most personal style to adopt a stance of honest frankness and to reveal the 'real me'. This might be true for the HKE speakers in Study 2 too. The next step of the present project will investigate the divergence of HKE speakers using quantitative discourse analysis.

Although previous studies found a significant effect of language attitudes on speech convergence, the results of Study 3 did not find the same effect. In section 8.2, I argued that what matters to convergence might be the attitudes towards the interlocutor rather than the attitudes towards the accent the interlocutor speaks. More studies are needed to explore the different effects of these two types of attitude on speech accommodation.

Many studies of speech accommodation (including Study 2 of this project) only collect production data from speakers and their interlocutors. The HEM has shown that to understand the whole process of accommodation, perceptual data is essential. Fundamentally, the perception-production link is the core for understanding the process of speech accommodation. Sociolinguistics and psycholinguistics have been working

together on speech accommodation for decades, in the future, more interdisciplinary collaborations between different disciplines such as cognitive science and neuroscience will help to reveal the whole picture of speech convergence.

Chapter 9. Conclusion

9.1. Scope of the research

This dissertation studied the speech accommodation of HKE speakers and the effect of language attitudes on accommodation. Three research questions were addressed in the project: (1) what are Hong Kong people's attitudes towards British English, American English and HKE? (2) do HKE speakers accommodate towards native accents after a short-term exposure to the native accents? (3) do HKE speakers' language attitudes affect their accommodation towards native English accents?

To answer these questions, three studies were conducted.

Study 1 elicited language attitudes from 107 Hong Kong respondents using the matched-guise method. The respondents listened to eight recordings spoken in different English accents (i.e. RP, GenAmE and HKE) and gave their judgements on 20 semantic traits for each recording. Three dimensions of language attitude, i.e. *Status & Competence*, *Social Attractiveness* and *Factor 3* were extracted using Principal Component Analysis. These three dimensions explained 72% of the variance in the data.

Study 2 collected about 40 hours of speech data from 19 HKE speakers. The HKE speakers completed a map task with a RP speaker and repeated the same task with a GenAmE speaker three to four weeks later. In the map tasks, the native interlocutors gave instructions to the HKE speakers so that the HKE speakers could draw a route on the map and correct the wrong landmarks. A pre-task and a post-task were also conducted. Realisations of the THOUGHT vowel, the PATH vowel, rhoticity, fricative [z] and fricative [θ] from both the native speakers and the HKE speakers were extracted from the map tasks, the pre-tasks and the post-tasks. By comparing the HKE speakers' pronunciation across the three tasks, their convergence on each variable was calculated.

Study 3 had the same HKE participants as Study 2. The participants completed a similar accent evaluation task to the one in Study 1. Their attitudes towards the RP accent and the GenAmE accent were collected and used as predictors for their convergence on rhoticity.

9.2. Key findings

Several key findings in the three studies are summarised below:

- (1) In Study 1, Hong Kong respondents rated RP as more prestigious and more attractive than GenAmE. HKE was rated as more attractive than two of the native recordings, suggesting a high level of acceptance of HKE in terms of social attractiveness;
- (2) In Study 2, the HKE speakers' THOUGHT vowels were found to have no significant changes after they talked to the native interlocutors but the PATH vowels were found to diverge from the native interlocutors from the pre-tasks to the map tasks;
- (3) In Study 2, the HKE speakers significantly converged towards the native interlocutors on rhoticity and fricative [z] but they were also found to have diverged from the native interlocutors on fricative [θ];
- (4) In Study 3, the HKE speakers' attitudes towards British English and American English did not predict their convergence on rhoticity.

The findings in Study 1 aligned with those in the previous literature on status; that is, British English is seen as more prestigious than American English in Hong Kong (Bolton & Kwok, 1990; Candler, 2001; Chan, 2013; Cheng, 2013; Groves, 2011). In terms of attractiveness, the results supported Cheng (2013), in which British English was rated as more attractive than American English, but the results were different from Zhang (2009) and Bayard and Green (2002), where American English was found to be more attractive. These results suggested that British English still maintains its prestigious status in Hong Kong 21 years after the handover from the UK. The different results for attractiveness found between Study 1 and those in the previous literature indicate that Hong Kong people's attitudes towards accent attractiveness might be changing over time, due to the influence of American popular culture and American businesses in Hong Kong.

Another finding in Study 1 is that HKE was rated as more attractive than two of the native recordings (i.e. GenAmE-guise and RP-guise), suggesting that acceptance of HKE is increasing in Hong Kong. This finding provides evidence for the ongoing debate about the status of HKE; that is, HKE is a developing new variety of English and it is in phase 3 (nativization) of Schneider's model (2003, 2007). This stage refers to the status where people in Hong Kong are transiting from accepting the imported native variety (i.e.

British English) as the dominant language to increasing independence of the local variety (i.e. HKE).

The results of Study 2 are complex. First of all, the HKE speakers were found to converge differently across the five variables. Convergence was only found on rhoticity and the fricative [z], divergence was found on the PATH vowel and fricative [θ], and maintenance was found on the THOUGHT vowel. Many previous studies have also shown the selectivity of convergence; for example, Babel (2012) found that American college students converged on the low vowels /æ/ and /ɑ/ out of the five vowels /i æ ɑ o u/. The present study suggested that people might tend to converge on the linguistic features which are more salient to them. Salience is defined as sounds with greater phonetic differences between speakers and interlocutors, and sounds carrying social meanings in the present study. Rhoticity is argued to be a more salient feature compared to fricative [z] and fricative [θ] for HKE speakers, because convergence on rhoticity involves adding or deleting a phoneme /-ɹ/ while convergence on the fricatives only involve replacement. Fricative [z] is also believed to be more salient than fricative [θ] for HKE speakers because the phonetic differences between /z/-/s/ are larger than the differences between /θ/-/f/. Extralinguistic factors might also be able to explain the convergence on rhoticity. Compared to the vowels, rhoticity/non-rhoticity is a stereotype feature of American English/British English. The HKE speakers might easily spot the difference in rhoticity between their own accent and the GenAmE accent. The significant convergence on rhoticity in the GenAmE condition also confirmed this.

Secondly, exposure to different English accents in the map tasks influenced the HKE speakers' accommodation on the THOUGHT vowel and rhoticity. For the THOUGHT vowel, a trend of convergence (statistically not significant) was found when the HKE speakers talked to the RP speakers, and a trend of divergence was found when they talked to the GenAmE speakers. For rhoticity, significant convergence was found in the GenAmE condition. These results indicate that non-native English speakers were able to adjust their pronunciation based on the accents/interlocutors they were exposed to.

Thirdly, the speaker's sex was found to have no effect on the HKE speakers' convergence on the five variables, except for fricative [θ]. This result echoed Pardo et al. (2017)'s finding that there was no difference in convergence between same-sex pairs and mixed-

sex pairs. Though previous literature found that females tend to converge more than males (Namy et al., 2002), in the present study, the male HKE participants converged more than the female HKE participants on fricative [θ] from the pre-tasks to the post-tasks.

Study 3 did not find a significant correlation between language attitudes and convergence. This result contradicted the findings in previous studies (Babel 2010, 2012; Pardo et al., 2010). Note that in the present study, due to the constraints of the small sample size (i.e. 14 HKE speakers) and to avoid overfitting, only the dimension *Linguistic Attractiveness* from PCA and *Integrative Orientation* from the explicit attitudes were chosen as predictors for the model. Also, the present study examined people's attitudes towards different English accents instead of their attitudes towards the interlocutors they were exposed to. It could be that the latter type of attitude has a more direct impact on people's convergence, as Babel (2012) found that the more attractive the model speaker was rated, the more likely the female participants were to converge. Another possibility is that the HKE speakers placed their attention on solving the map tasks which required high cognitive loads, so that little cognitive space was left for attitudes to interfere with people's production.

Apart from the three empirical studies, the hybrid exemplar-based model (the HEM) was proposed for short-term accommodation. The HEM suggested three stages: Stage 1 - perception, Stage 2 - update of distributions and Stage 3 - production. In Stage 1, when speakers hear speech from their interlocutors, speakers would add weights to the inputs of the interlocutors based on the salience of the sounds. The sounds which had larger phonetic differences from the interlocutors and/or sounds with social meanings would be weighted greater. In Stage 2, the speakers would update their prior established distributions of the sounds based on the input they received in Stage 1. The quantity and quality of the input would affect the updates. In Stage 3, the speakers would select a production goal from the new distributions, the selection would be affected by the speakers' articulatory constraints and social factors such as language attitudes.

9.3. Contributions of the research to broader fields

This dissertation has contributed to several disciplines in linguistics.

Firstly, Study 1 contributes to sociolinguistics by providing an updated investigation into Hong Kong people's attitudes towards different English varieties. It also contributes to the ongoing debate about the status of HKE, suggesting that Hong Kong people's acceptance of HKE is increasing.

Secondly, Study 2 and Study 3 complement research into phonetic convergence by extending the subjects of the research to Hong Kong English speakers. Previous studies of speech accommodation usually have a highly homogenous participant sample. As HKE is still a developing variety, the HKE participants in the Study 2 and Study 3 did not have a homogenous accent. Although complicated patterns were found in these studies, the results revealed the complicity of speech accommodation.

Thirdly, this project also contributes to the development of accommodation theories. For example, Study 2 did not support the automaticity proposed by the interactive-alignment model (Pickering & Garrod, 2004) as convergence was only found on some but not all sounds. Study 3 did not support CAT (Giles, et al., 1991) either, because the HKE speakers' attitudes did not predict their convergence on rhoticity. The HEM was proposed to explain the complex findings in Study 2 and Study 3. The HEM aimed to provide a framework for short-term accommodation which covers both the perception and the production, and includes social factors as important elements in the model.

Finally, the project explores convergence from the perspectives of sociolinguistics, psycholinguistics, cognitive science and second language learning. It connects a few disciplines in linguistics through a well-designed study on conversation between non-native English speakers and native English speakers. It also provides a good foundation for future studies which are interested in the convergence of non-native speakers.

Appendices

Appendix 1. Study 1: Accent evaluation form

The speaker appears to (be)							
Status & Competence							
unintelligent	1	2	3	4	5	6	intelligent
uneducated	1	2	3	4	5	6	educated
unambitious	1	2	3	4	5	6	ambitious
have low leadership quality	1	2	3	4	5	6	have high leadership quality
unconfident	1	2	3	4	5	6	confident
lower-class	1	2	3	4	5	6	upper-class
informal	1	2	3	4	5	6	formal
impolite	1	2	3	4	5	6	polite
Social Attractiveness							
unreliable	1	2	3	4	5	6	reliable
unpleasant	1	2	3	4	5	6	pleasant
hostile	1	2	3	4	5	6	nature
lack of sense of humour	1	2	3	4	5	6	have sense of humour
unattractive	1	2	3	4	5	6	attractive
old-fashioned	1	2	3	4	5	6	modern
uninteresting	1	2	3	4	5	6	interesting
unfriendly	1	2	3	4	5	6	friendly
unenthusiastic	1	2	3	4	5	6	enthusiastic
Linguistic Quality							
unintelligible	1	2	3	4	5	6	intelligible
not aesthetic	1	2	3	4	5	6	aesthetic
not a good model of pronunciation	1	2	3	4	5	6	a good model of pronunciation

Appendix 2. Study 1: Questionnaire of language attitudes

1. Your gender is: Male Female
2. Your group:
 Secondary school student University student
 Professional aged below 30 Professional aged over 30
3. How many years of formal English teaching have you had?
 Less than 6 years 6-12 years more than 12 years
4. Have you ever visited or lived in an English-speaking country? If yes, where and how long?
 Yes Where _____ Duration _____ No
5. Did you go to an English secondary school or Chinese secondary school?
 English secondary School Chinese secondary school

6. Which accent did/do your English teachers mostly have?

	AmE	BrE	Hong Kong English	Australian English	Canadian English	Other (please identify)	I don't know
Primary school English teacher							
Secondary school English teacher							
University English teacher							

7. In your opinion, which accent is favoured when teaching English in Hong Kong?
 American English British English Australian English
 Hong Kong English Others (Please identify)
8. In your opinion, which accent is favoured when working in Hong Kong?
 American English British English Australian English
 Hong Kong English Others (Please identify)
9. Which accent/pronunciation are you aiming at when you speak English at school?
 American English British English Australian English
 Hong Kong English Others (Please identify)
10. Which accent/pronunciation are you aiming at when you speak English outside school?
 American English British English Australian English
 Hong Kong English Others (Please identify)
11. Which pronunciation do you think is more pleasant to listen to?
 American English British English
 Hong Kong English No preference

12. Which pronunciation do you think is more pleasant to speak?

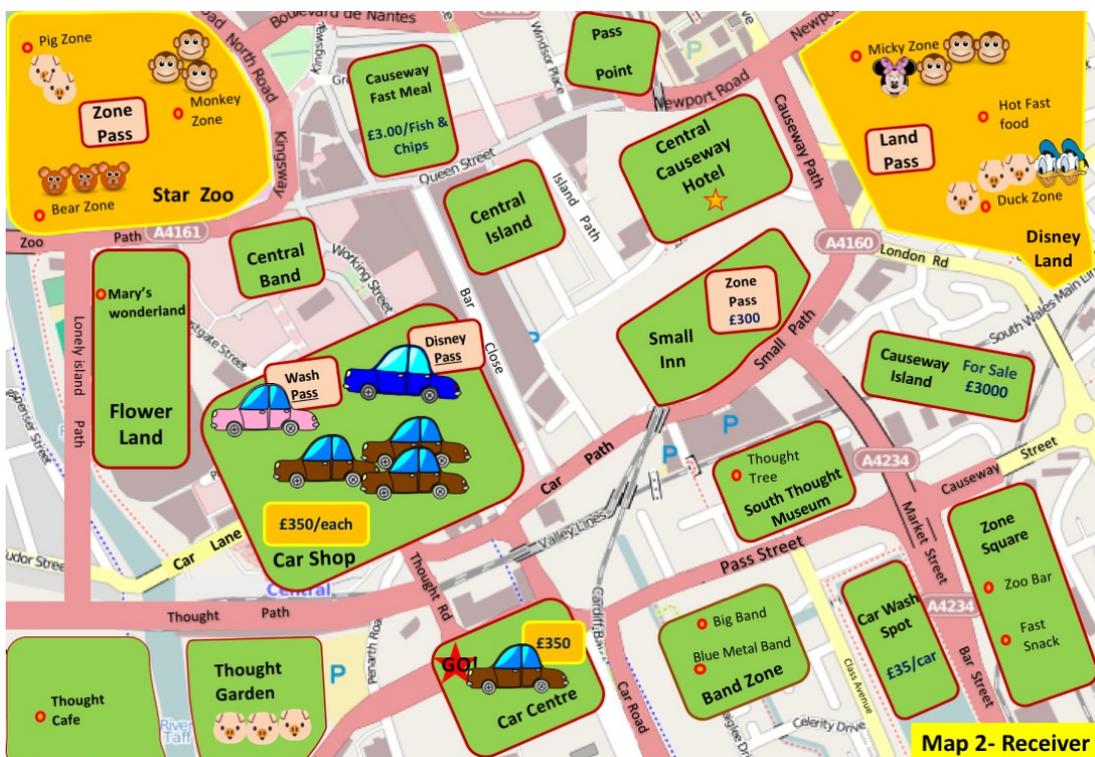
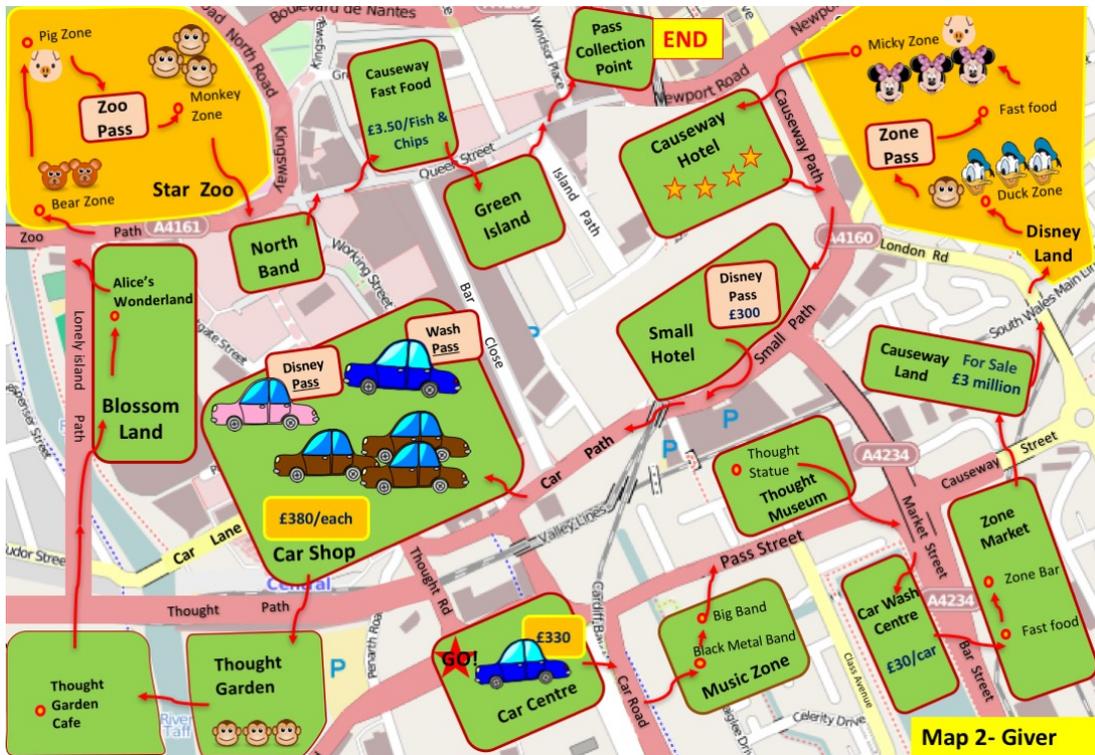
American English British English
Hong Kong English No preference

13. If you had to choose between Great Britain and the USA,

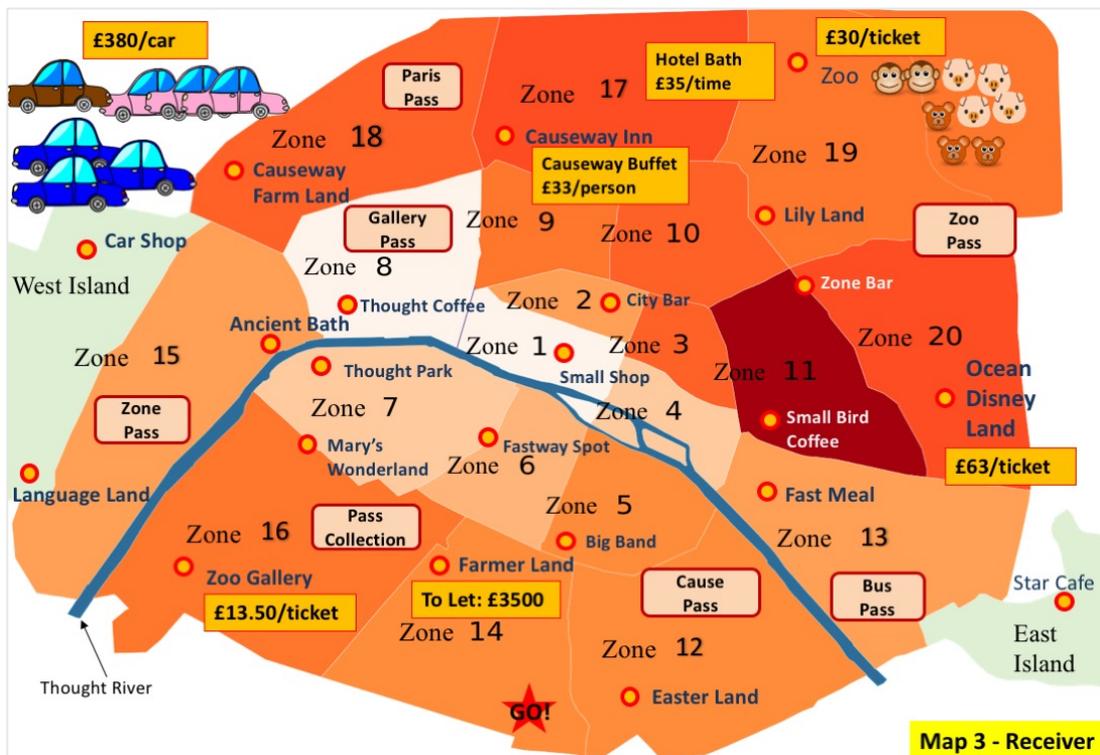
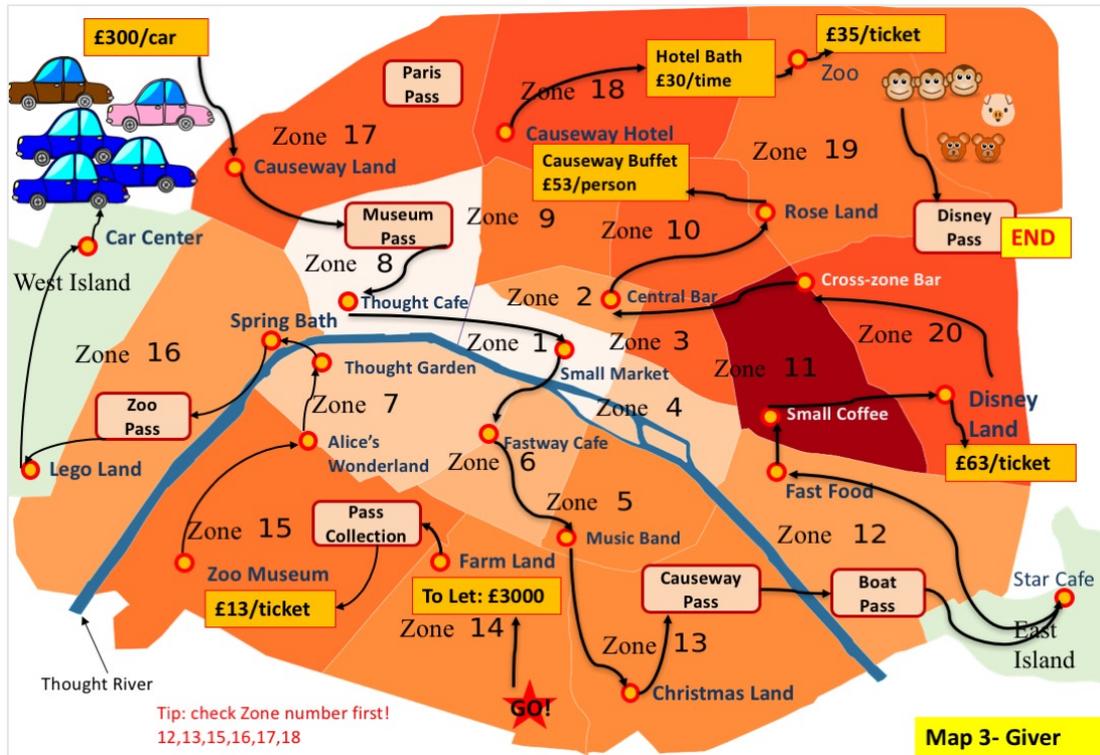
- where would you rather live? Great Britain the USA
- where would you rather work? Great Britain the USA
- with whom would you rather get married? British American

Appendix 3. Study 2: The maps used in the map tasks

Map 2



Map 3



Appendix 4. Study 2: Summary of rhoticity% for all participants

rhoticity%	Exposure to RP				Exposure to GenAmE			
	pre	map	post	NS-RP	pre	map	post	NS-GenAmE
HK1	9.76% (4/41)	0.00% (0/56)	2.04% (1/49)	0.00% (0/128)	1.47% (1/68)	3.70% (2/54)	6.00% (6/100)	99.12% (113/114)
HK2	4.26% (2/47)	0.00% (0/64)	2.30% (2/87)	0.93% (1/107)	0.00% (0/80)	2.22% (1/45)	2.08% (2/96)	99.42% (171/172)
HK3	23.81% (20/84)	3.80% (3/79)	4.31% (5/116)	0.00% (0/107)	17.36% (25/144)	2.99% (2/67)	14.17% (17/120)	98.74% (157/159)
HK4	5.41% (6/111)	2.17% (1/46)	8.70% (8/92)	0.00% (0/143)	8.33% (11/132)	14.47% (11/76)	18.10% (21/116)	98.95% (94/95)
HK5	74.39% (61/82)	58.06% (54/93)	52.99% (62/117)	0.00% (0/130)	47.76% (64/134)	48.44% (31/64)	40.71% (57/140)	98.35% (119/121)
HK6	79.78% (71/89)	68.09% (32/47)	74.21% (118/159)	0.00% (0/164)	80.30% (106/132)	90.38% (47/52)	72.53% (132/182)	98.84% (85/86)
HK7	5.71% (4/70)	63.24% (43/68)	64.06% (41/64)	3.03% (3/99)	52.00% (39/75)	54.55% (36/66)	60.44% (55/91)	100% (115/115)
HK8	89.04% (65/73)	84.85% (56/66)	83.33% (75/90)	1.92% (2/104)	91.86% (79/86)	95.83% (69/72)	97.65% (83/85)	99.01% (100/101)
HK9	9.38% (6/64)	0.00% (0/55)	2.44% (2/82)	0.00% (0/70)	0.00% (0/119)	1.47% (1/67)	0.00% (0/108)	99.46% (184/185)
HK11	5.13% (6/117)	4.88% (4/82)	0.92% (1/109)	0.00% (0/162)	0.00% (0/66)	8.82% (6/68)	1.12% (1/89)	95.17% (138/145)
HK12	12.39% (14/113)	12.50% (8/64)	9.93% (12/141)	0.00% (0/163)	1.01% (1/99)	32.61% (15/46)	23.53% (28/119)	96.37% (186/193)
HK13	65.38% (68/104)	48.28% (28/58)	52.00% (52/100)	0.00% (0/152)	65.48% (55/84)	78.95% (30/38)	52.63% (50/95)	97.80% (178/182)
HK16	2.22% (2/90)	0.00% (0/67)	4.60% (4/87)	0.00% (0/122)	2.30% (2/87)	26.15% (17/65)	9.09% (7/77)	94.57% (122/129)
HK18	30.85% (29/94)	20.34% (12/59)	9.52% (10/105)	0.78% (1/129)	33.78% (25/74)	48.89% (22/45)	28.09% (25/89)	96.58% (113/117)
HK19	38.10% (8/21)	29.69% (19/64)	24.05% (19/79)	0.00% (0/130)	27.27% (9/33)	21.54% (14/65)	21.59% (19/88)	92.00% (115/125)
HK20	26.43% (37/140)	22.95% (14/61)	24.64% (34/138)	0.00% (0/149)	22.89% (19/83)	41.18% (35/85)	34.45% (41/119)	93.57% (131/140)
HK21	7.00% (7/100)	10.81% (8/74)	1.14% (1/88)	0.00% (0/154)	4.40% (4/91)	4.48% (3/67)	3.45% (4/116)	89.71% (122/136)
HK22	1.23% (1/81)	2.78% (2/72)	0.00% (0/90)	0.00% (0/122)	1.69% (1/59)	1.82% (1/55)	0.00% (0/89)	93.10% (108/126)
HK23	67.53% (52/77)	34.69% (17/49)	44.19% (38/84)	0.00% (0/157)	57.89% (33/57)	60.38% (32/53)	59.57% (56/94)	92.50% (111/120)

Appendix 4. Summary of rhoticity% for all participants across the three tasks in the two conditions. The ratio in brackets shows the tokens of rhotic words / tokens of all words.

Appendix 5. Study 2: Summary of fricative [z] for all participants

z%	Exposure to RP				Exposure to GenAmE			
	pre	map	post	NS-RP	pre	map	post	NS-GenAmE
HK1	0% (0/86)	2.56% (1/39)	0% (0/78)	87.31% (117/134)	0% (0/98)	6.25% (3/48)	0% (0/68)	96.10% (74/77)
HK2	0% (0/75)	2.99% (2/67)	0% (0/78)	91.07% (102/112)	0% (0/85)	6.67% (3/45)	1.25% (1/80)	97.24% (141/145)
HK3	0% (0/93)	2.44% (1/41)	7.61% (7/92)	96.97% (96/99)	1.08% (1/93)	0% (0/48)	1.15% (1/87)	100% (125/125)
HK4	10.38% (11/106)	22.58% (7/31)	4.76% (4/84)	94.64% (106/112)	0% (0/110)	7.32% (3/41)	4.62% (3/65)	90.28% (65/72)
HK5	1.03% (1/97)	3.39% (2/59)	11.34% (11/97)	88.07% (96/109)	5.00% (6/120)	9.26% (5/54)	0% (0/91)	92.96% (66/71)
HK6	0% (0/72)	0% (0/27)	0% (0/71)	90.65% (97/107)	0% (0/78)	0% (0/20)	0% (0/72)	96.23% (51/53)
HK7	0% (0/72)	0% (0/60)	0% (0/57)	88.73% (63/71)	0% (0/57)	0% (0/40)	0% (0/65)	91.94% (57/62)
HK8	37.70% (23/61)	50% (29/58)	66.10% (39/59)	83.78% (62/74)	54.10% (33/61)	70.45% (31/44)	69.09% (38/55)	93.94% (62/66)
HK9	98.18% (54/55)	84.00% (21/25)	98.33% (59/60)	94.92% (56/59)	98.70% (76/77)	87.50% (35/40)	95.83% (69/72)	100% (130/130)
HK11	0% (0/78)	0% (0/68)	0% (0/62)	91.06% (112/123)	0% (0/69)	2.22% (1/45)	0% (0/57)	96.74% (89/92)
HK12	60.56% (43/71)	59.52% (25/42)	50% (39/78)	84.16% (85/101)	37.18% (29/78)	62.16% (23/37)	68.42% (52/76)	99.25% (132/133)
HK13	1.69% (1/59)	0% (0/33)	0% (0/58)	92.00% (69/75)	0% (0/49)	0% (0/24)	2.04% (1/49)	97.58% (121/124)
HK16	33.33% (17/51)	11.67% (7/60)	0% (0/63)	82.61% (57/69)	1.92% (1/52)	7.14% (3/42)	0% (0/72)	93.10% (81/87)
HK18	49.33% (37/75)	67.44% (29/43)	65.08% (41/63)	88.57% (62/70)	55.07% (38/69)	45.71% (16/35)	39.19% (29/74)	93.83% (76/81)
HK19	0% (0/14)	0% (0/43)	0% (0/58)	90.24% (74/82)	0% (0/35)	0% (0/46)	0% (0/62)	91.84% (90/98)
HK20	0% (0/116)	8.11% (3/37)	1.05% (1/95)	94.21% (114/121)	0% (0/77)	0% (0/45)	0% (0/54)	91.57% (76/83)
HK21	0% (0/85)	9.26% (5/54)	0% (0/36)	93.16% (109/117)	0% (0/58)	3.75% (3/80)	1.16% (1/86)	92.94% (79/85)
HK22	2.60% (2/77)	9.26% (5/54)	0% (0/64)	91.67% (66/72)	0% (0/61)	9.68% (3/31)	0% (0/57)	94.44% (68/72)
HK23	0% (0/54)	6.90% (2/29)	0% (0/55)	95.45% (105/110)	0% (0/59)	5.88% (2/34)	0% (0/42)	100% (77/77)

Appendix 5. Summary of z% for all participants across the three tasks in the two conditions. The ratio in brackets shows the tokens of voiced [z] / tokens of all z-words.

Appendix 6. Study 2: Summary of fricative [θ] for all participants

θ%	Exposure to RP				Exposure to GenAmE			
	pre	map	post	NS-RP	pre	map	post	NS-GenAmE
HK1	0% (0/54)	0% (0/58)	0% (0/41)	100% (89/89)	0% (0/78)	0% (0/57)	0% (0/61)	100% (66/66)
HK2	37.50% (15/40)	17.65% (12/68)	45.90% (28/61)	100% (81/81)	23.33% (14/60)	26.23% (16/61)	26.87% (18/67)	100% (108/108)
HK3	88.24% (45/51)	83.82% (57/68)	78.57% (44/56)	100% (64/64)	86.84% (66/76)	87.84% (65/74)	83.87% (52/62)	100% (89/89)
HK4	91.30% (42/46)	92.73% (51/55)	100% (49/49)	100% (72/72)	98.65% (73/74)	100% (60/60)	96.72% (59/61)	97.87% (46/47)
HK5	98.39% (61/62)	96.39% (80/83)	100% (47/47)	98.51% (66/67)	100% (74/74)	95.65% (66/69)	100% (68/68)	100% (63/63)
HK6	37.29% (22/59)	46.00% (23/50)	45.28% (24/53)	100% (68/68)	34.21% (26/76)	31.03% (18/58)	28.17% (20/71)	100% (58/58)
HK7	92.45% (49/53)	52.46% (32/61)	48.65% (18/37)	98.11% (52/53)	56.60% (30/53)	68.25% (43/63)	73.68% (42/57)	100% (66/66)
HK8	100% (47/47)	98.51% (66/67)	98.18% (54/55)	100% (53/53)	97.06% (66/68)	100% (70/70)	100% (69/69)	100% (62/62)
HK9	98.18% (54/55)	96.00% (48/50)	100% (54/54)	100% (43/43)	98.80% (82/83)	93.59% (73/78)	97.22% (70/72)	100% (102/102)
HK11	100% (72/72)	90% (81/90)	98.33% (59/60)	95.79% (91/95)	98.04% (50/51)	94.23% (49/52)	96.23% (51/53)	100% (65/65)
HK12	24.32% (18/74)	7.46% (5/67)	24.64% (17/69)	98.51% (66/67)	20.37% (11/54)	10.87% (5/46)	44% (22/50)	100% (84/84)
HK13	88.14% (52/59)	90.91% (40/44)	92.73% (51/55)	98.53% (67/68)	100% (43/43)	88.24% (30/34)	88.68% (47/53)	100% (82/82)
HK16	94.83% (55/58)	92.65% (63/68)	100% (73/73)	100% (60/60)	96.08% (49/51)	96.15% (50/52)	100% (52/52)	98.41% (62/63)
HK18	98.65% (73/74)	94.29% (66/70)	96.83% (61/63)	100% (52/52)	93.33% (42/45)	90.74% (49/54)	94.83% (55/58)	87.88% (58/66)
HK19	60% (3/5)	68.25% (43/63)	66.13% (41/62)	98.67% (74/75)	37.93% (11/29)	51.28% (20/39)	52.78% (19/36)	97.40% (75/77)
HK20	56.25% (45/80)	41.54% (27/65)	32.43% (24/74)	98.82% (84/85)	17.86% (10/56)	46.97% (31/66)	44.23% (23/52)	98.08% (51/52)
HK21	0% (0/66)	1.25% (1/60)	2.13% (1/47)	100% (95/95)	6.25% (3/48)	11.27% (8/71)	4.69% (3/64)	89.06% (57/64)
HK22	98.51% (66/67)	100% (73/73)	100% (61/61)	100% (57/57)	98.11% (52/53)	100% (61/61)	96.15% (50/52)	100% (63/63)
HK23	75.00% (48/64)	37.84% (14/37)	71.67% (43/60)	100% (77/77)	62.16% (23/37)	58.06% (36/62)	66.07% (37/56)	94.29% (66/70)

Appendix 6. Summary of θ% for all participants across the three tasks in the two conditions. The ratio in brackets shows the tokens of [θ] / tokens of all th-words.

References

- Adank, P., Smits, R., & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America*, 116(5), 3099–3107.
- Allport, G. W. (1935). Attitudes. In C. M. Murchison (Eds.), *Handbook of social psychology* (pp. 789-844). Worcester, MA: Clark University Press.
- Axler, M., Yang, A., & Steven, T. (1998). Current language attitudes of Hong Kong Chinese adolescents and young adults. In M.C. Pennington (Eds.), *Language in Hong Kong at century's end* (pp. 391-418). Hong Kong: Hong Kong University Press.
- Babel, M. E. (2010). Dialect divergence and convergence in New Zealand. *Language in Society*, 39(4), 437–456.
- Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, 40(1), 177–189.
- Bacon-Shone, J., Bolton, K., & Luke, K. (2015). *Language use, proficiency and attitudes in Hong Kong*. Social Science Research Centre, The University of Hong Kong.
- Bayard, D., & Green, J. (2002). *Evaluating English accents worldwide results: Hong Kong*. Retrieved from <http://www.otago.ac.nz/anthropology/Linguistic/Results/HK.html>.
- Boersma, Paul., & Weenink, David. (2017). Praat: doing phonetics by computer [Computer program]. version 6.0.28, retrieved 6 August 2017 from <http://www.praat.org/>.
- Bolton, K. (2000). The sociolinguistics of Hong Kong and the space for Hong Kong English. *World Englishes*, 19(3), 265–285.
- Bolton, K., & Kwok, H. (1990). The dynamics of the HK accent-social identity and sociolinguistic description. *Journal of Asian Pacific Communication*, 147–172.
- Bourhis, R. Y. (1979). Language in ethnic interaction: a social psychological approach. In *Language and Ethnic Relations* (pp. 117–140).
- Boyle, J. (1997). Native-speaker teachers of English in Hong Kong, *Language and Education*, 11(3), 163-181.
- Butler, S. (1997). Corpus of English in Southeast Asia: implications for a regional dictionary. In M. L. S. Bautista (Eds.), *English is an Asian language: the Philippine context* (pp.103-24). Manila: The Macquarie Library.
- Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14, 261–290.
- Byrne, D. (1969). Attitudes and attraction. In L. Berkowitz (Eds.), *Advances in experimental social psychology*. New York: Academic Press.

- Candler, R. (2001). *An investigation into Hong Kong non-native speakers' recognition of and attitudes towards difference accents of English*. MA thesis, the University of Hong Kong, Hong Kong.
- Census and Statistics Department. (1996). *1996 Populations By-Census*. Hong Kong: Hong Kong Special Administrative Region Government, People's Republic of China.
- Census and Statistics Department. (2011). *2011 Populations Census*. Hong Kong: Hong Kong Special Administrative Region Government, People's Republic of China.
- Chan, J. Y. H. (2013). Contextual variation and Hong Kong English. *World Englishes*, 32(1), 54–74.
- Cheng, K. S. K. (2013). Survey of young educated Hong Kong People's attitudes towards British English and American English : comparing differences by tertiary learning experiences. *British Journal of Arts and Social Sciences*, 12(I), 1–36.
- Cooper, R. (1975). Introduction to language attitudes II. *International Journal of the Sociology of Language*, 6, 5–9.
- Cutting, C. (2004). A call for more dialogue and more details (commentary in toward a mechanistic psychology of dialogue nonalignment). *The Behavioral and Brain Sciences*, 27(2), 203.
- Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English. *Journal of the International Phonetic Association*, 27(1–2), 47–55.
- Deterding, D., Wong, J., & Kirkpatrick, A. (2008). The pronunciation of Hong Kong English. *English World-Wide*, 29(2), 148–175.
- Dijksterhuis, A., & Bargh, J. (2001). The perception-behavior expressway: automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology*, 33, 1–38.
- Docherty, G. J., & Foulkes, P. (2014). An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua*, 142, 42–56.
- Eagly, A., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace College Publisher. New York.
- Edwards, A. L. (1957). *Techniques of attitude scale construction*. New York: Appleton-Century-Crofts.
- Evans, B. G., & Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *The Journal of the Acoustical Society of America*, 121(6), 3814.
- Fabricius, A. H. (2002). Ongoing change in modern RP: evidence for the disappearing stigma of t-glottalling. *English World-Wide*, 23(1), 115–136.

- Fabricius, A. H., Watt, D., & Johnson, D. E. (2009). A comparison of three speaker-intrinsic vowel formant frequency normalization algorithms for sociophonetics. *Language Variation and Change*, 21(3), 413–435.
- Field, A. (2009). *Discovering Statistics Using SPSS* (3rd ed.). London: Sage Publications.
- Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, 34(4), 409–438.
- Gardner, R. (1985). *Social psychology and second language learning: the role of attitudes and motivation*. London; Baltimore, MD: Edward Arnold.
- Gardner, R. (2006). The socio-educational model of second language acquisition: a research paradigm. In S.H. Foster-Cohen, M. Medved Krajnovic & J. Mihaljecic Djigunovic (Eds.), *Eurosla Yearbook (volume 6)*. Amsterdam: John Benjamins Publishing.
- Gardner, R. (2010). *Motivation and second language acquisition: the socio-educational model*. New York: Peter Language Publishing.
- Gardner, R. C., & Lambert, W E. (1972). *Attitudes and motivation in second language learning*. Rowley, MA: Newbury House.
- Garrett, P. (2010). *Attitudes to Language*. Cambridge: Cambridge University Press.
- Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition*, 27(2), 181–218.
- Gass, S., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction (1st edition)* (pp. 175-200). London: Lawrence Erlbaum Associates, Publishers.
- Gass, S., & Mackey, A. (2015). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction (2nd edition)* (pp. 180-207). London: Routledge.
- Giles, H. (1973). Accent mobility: a model and some data, 15(2), 87–105.
- Giles, H., Coupland, J., & Coupland, N. (1991). *Contexts of accommodation: studies in emotion & social interaction*. Cambridge: Cambridge University Press.
- Giles, H., Mulac, A., Bradac, J. J., & Johnson, P. (1987). Speech Accommodation Theory: the first decade and beyond. In *Communication Yearbook 10*(1), 13–48.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251–279.

- Greenwald, A. G., & Banaji, R. M. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4–27.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the Implicit Association Test. *The Implicit Association Test. Journal of Personality and Social Psychology*, 74(6), 1464–1480.
- Groves, J. (2009). Hong Kong English – Does it exist? *HKBU Papers in Applied Language Studies*, 13, 54–79.
- Groves, J. (2011). “Linguistic schizophrenia” in Hong Kong. *English Today*, 27(4), 33–42.
- Hagiwara, R. (1997). Dialect variation and formant frequency: the American English vowels revisited. *The Journal of the Acoustical Society of America*, 102(1), 655–658.
- Hansen Edwards, J. G. H. (2015). Hong Kong English: attitudes, identity, and use. *Asian Englishes*, 17(3), 184–208.
- Hansen Edwards, J. G. H. (2016a). The politics of language and identity: attitudes towards Hong Kong English pre and post the Umbrella Movement. *Asian Englishes*, 18(2), 157–164.
- Hansen Edwards, J. G. H. (2016b). The deletion of /t, d/ in Hong Kong English. *World Englishes*, 35(1), 60–77.
- Hansen Edwards, J. G. H. (2016c). Accent preferences and the use of American English features in Hong Kong: a preliminary study. *Asian Englishes*, 18(3), 197–215.
- Heider, P. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Hewstone, M., & Giles, H. (1986). Social groups and social stereotypes in intergroup communication: review and model of intergroup communication breakdown. In W. B. Gudykunst (Eds.), *Intergroup Communication* (pp. 10–26). London: Edward Arnold.
- Hintzman, D. L. (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review*, 93, 411–428.
- Holmes, J., & Bell, A. (1994). *Consonant cluster reduction in New Zealand English. Wellington Working Papers in Linguistics* 6, 56–82.
- Hong Kong Examinations and Assessment Authority. (2012). *Benchmarking Study between IELTS and HKDSE English Language Examination FAQ*. Retrieved from http://www.hkeaa.edu.hk/DocLibrary/IR/FAQs_IELTS_eng_Aug2013.pdf
- Homans, G. C. (1961). *Social behaviour*. New York: Harcourt, Brace and World.
- Hung, T. T. N. (2000). Towards a phonology of Hong Kong English. *World Englishes*, 19(3), 337–356.

- Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, 108(3), 1252.
- Johnson, K. (1997a). Speech perception without speaker normalization: an exemplar model. In K. Johnson and J.W. Mullennix (Eds.) *Talker Variability in Speech Processing* (pp. 145-166). San Diego: Academic Press.
- Johnson, K. (1997b). The auditory/perceptual basis for speech segmentation. *OSU Working Papers in Linguistics*, 50, 101–113.
- Johnson, K. (2007). Decisions and mechanisms in exemplar-based phonology. *Experimental Approaches to Phonology: In Honor of John Ohala*, (1968), 25–40.
- Johnson, R. K. (1994). Language Policy and Planning in Hong Kong. *Annual review of applied linguistics*, 14, 177–199.
- Kachru, B. B. (1985) Standards, codification and sociolinguistic realism: the English language in the outer circle. In R. Quirk and H.G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11-30). Cambridge: Cambridge University Press.
- Kachru, B. (1992). Models for non-native Englishes. In B. Kachru (Eds), *The other tongue: English across cultures (2nd edition)* (pp. 48–74). University of Illinois Press.
- Karpinski, A., & Hilton, J. L. (2001). Attitudes and the implicit association test. *Journal of Personality and Social Psychology*, 81(5), 774–788.
- Kerswill, P., & Williams, A. (2002). “Salience” as an explanatory factor in language change: Evidence from dialect levelling in urban England. In M. Jones and E. Esch (Eds.) *Language Change. The Interplay of Internal, External and Extra-Linguistic Factors*, (pp.1–32). Berlin: Mouton de Gruyter.
- Kerswill, P. (1985). A sociophonetic study of connected speech processes in Cambridge English: an outline and some results. *Cambridge Papers in Phonetics and Experimental Linguistics*, 4, 1-39.
- 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.
- Kirkpatrick, A. (2007). *World Englishes: Implications for international communication and English language teaching*. Cambridge: Cambridge University Press.
- Krashen, S. D. (1985). *The input hypothesis: issues and implications*. London: Longman.
- Krauss, R., & Pardo, J. (2004). Is alignment always the result of automatic priming? (Commentary in Toward a mechanistic psychology of dialogue nonalignment). *The Behavioral and Brain Sciences*, 27(2).

- Labov, W. (1966). *The social stratification of English in New York*. Washington DC: Centre of Applied Linguistics.
- Labov, W. (2001). *Principles of Linguistic Change: Social Factors (Volume 2)*. Oxford: Blackwell.
- Labov, W., Ash, S., & Boberg, C. (2006). *The Atlas of North American English*. Berlin: Mouton de Gruyter.
- Ladegaard, H. J. (1998). National stereotypes and language attitudes: the perception of British, American and Australian language and culture in Denmark. *Language & Communication, 18*(4), 251–274.
- Lai, M.L. (2001). Hong Kong students' attitudes towards Cantonese, Putonghua and English after the change of sovereignty. *Journal of Multilingual and Multicultural Development, 22*(2), 112–133.
- Lai, M.L. (2005). Language attitudes of the first postcolonial generation in Hong Kong secondary schools. *Language in Society, 34*(3), 363–388.
- Lai, M. L. (2011). Cultural identity and language attitudes – into the second decade of postcolonial Hong Kong. *Journal of Multilingual and Multicultural Development, 32*(3), 249–264.
- Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology, 60*(1), 44-51.
- Legislative Council Panel on Education. (2014). *The latest development of the Native-speaking English Teacher ("NET") Scheme*. Retrieved from <http://www.legco.gov.hk/yr14-15/english/panels/ed/papers/ed20141110cb4-116-1-e.pdf>
- Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America, 49*(2B), 606.
- Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching, 51* (3), 285-329.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-454). London: Academic Press.
- Luk, C. M. (1998). Hong Kong people's reactions of and attitudes towards accents differences. *Multilingua, 1*, 93–105.
- Luke, K. K., & Richards, J. C. (1982). English in Hong Kong: functions and status. *English World-Wide, 3*(1), 47–64.

- Moag, R. E. (1992). The life cycle of non-native Englishes: A case study. In B. Kachru (Eds.), *The other tongue: English across cultures* (pp. 233–252). University of Illinois Press.
- Mufwene, S. S. (1991). Pidgins, creoles, typology, and markedness. In F. Byrne and T. Huebner (Eds.) *Development and structures of creole languages: Essays in honor of Derek Bickerton* (pp.123-143). Amsterdam: John Benjamins.
- Munson, B. (2010). Levels of phonological abstraction and knowledge of socially motivated speech-sound variation: a review, a proposal, and a commentary on the papers by Clopper, Pierrehumbert, and Tamati, Drager, Foulkes, Mack, and Smith, Hall, and Munson. *Laboratory Phonology, 1*(1), 157–177.
- 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.
- Neu, H. (1980). Ranking of constraints on /t,d/ deletion in American English: A statistical analysis. In William Labov (Eds.), *Locating language in time and space*, 37-54. New York: Academic Press.
- Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 700-708.
- Osgood, C. E., Suci, G. J, & Tannenbaum, P. H. (1957). *Measurement of meaning*. Urbana: The University of Illinois Press.
- Pang, T. (2003). Hong Kong English: a stillborn variety? *English Today, 19*(2), 12–18.
- Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America, 119*(4), 2382–2393.
- Pardo, J., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics, 40*(1), 190–197.
- Pardo, J., Jay, I., Krauss, R. (2010). Conversational role influences speech imitation. *Attention, Perception and Psychophysics, 72*(8), 2254–2264.
- Pardo, J., 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.
- Pardo, J., Urmanche, A., Wilman, S., & Wiener, J. (2017). Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics, 79*(2), 637–659.
- Paus, T., Petrides, M., Evans, a C., & Meyer, E. (1993). Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study. *Journal of Neurophysiology, 70*(2), 453–469.

- Pavlik, R. (2017). Some new ways of modelling t/d deletion in English. *Journal of English Linguistics*, 45(3), 195–228.
- Pennington, M., & Yue, F. (1994). English and Chinese in Hong Kong: pre-1997 attitudes. *World Englishes*, 13, 1–20.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *The Behavioral and Brain Sciences*, 27(2), 169–190.
- Pickering, M., & Ferreira, V. S. (2008). Structural priming: a critical review. *Psychological Bulletin*, 134(3), 427–459.
- Pierrehumbert, J. B. (2001). Exemplar dynamics: word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency effects and the emergence of linguistic structure* (pp. 137–158). Amsterdam: John and Benjamins.
- Pierrehumbert, J. B. (2002). Word-specific phonetics. *Laboratory Phonology VII*, 101–140.
- Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, 46(2–3), 115–154.
- Pierrehumbert, J. B. (2006). The next toolkit. *Journal of Phonetics*, 34, 516–530.
- Pierson, H. D., Fu, G. S., & Lee, S. Y. (1980) An analysis of the relationship between language attitudes and English attainment of secondary school students in Hong Kong. *Language Learning*, 30, 289–316.
- Poon, A. Y. K. (2011). A survey on language attitudes: what are the implications for English language teaching? In *Proceedings of the 16th Conference of Pan-Pacific Association of Applied Linguistics* (pp. 243–246). Hong Kong.
- Public Opinion Programme, The University of Hong Kong (2017). People's Ethnic Identity. Retrieved 26th May, 2018 from <https://www.hkupop.hku.hk/chinese/popexpress/ethnic/>.
- Rindal, U. (2010). Constructing identity with L2: pronunciation and attitudes among Norwegian learners of English. *Journal of Sociolinguistics*, 14(2), 240–261.
- Rindal, U. (2014). Questioning English standards: learner attitudes and L2 choices in Norway. *Multilingua*, 33(3-4), 313–334.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158.
- Schmidt, R. (2001). Attention. In P. Robinson (Eds.), *Cognition and second language instruction*. Cambridge: Cambridge University Press, 3–32.
- Schneider, E. W. (2003). The dynamics of new Englishes: From identity construction to dialect birth. In K. Bolton & B. B. Kachru (Eds.), *World Englishes: Critical concepts in linguistics: Vol. 1* (pp. 125–185). London: Routledge.

- Schneider, E. W. (2007). *Postcolonial English: varieties around the world*. Cambridge University Press.
- Schreier, D. (2005). *Consonant change in English worldwide*. Basingstoke, UK: Palgrave Macmillan.
- Sharma, D. (2018). Style dominance: Attention, audience, and the ‘real me’. *Language in Society*, 47(1), 1-13.
- So, D. (1996). Profiles of language education in 25 countries: Hong Kong. In P. Dickson & A. Cumming (Eds.), *Profiles of Language Education in 25 countries* (pp. 41–46). Berkshire, UK: National Foundation for Educational Research.
- South China Morning Post. (2013, October 29). American accent schools flourishing in Hong Kong. *South China Morning Post*. Retrieved from <http://www.scmp.com/news/hong-kong/article/1342563/american-accent-schools-flourishing-hong-kong>.
- Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching*, 29, 73–87.
- Sung, C. C. M. (2013). Accent and identity: exploring the perceptions among bilingual speakers of English as a lingua franca in Hong Kong. *International Journal of Bilingual Education and Bilingualism*, 17(5), 544–557.
- Swain, M. (2005). The output hypothesis: theory and research. In E. Hinkel (Eds.), *Handbook on research in second language learning and teaching*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Syrdal, A. K. and Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America* 79:1086-100.
- Tay, M. W. J. (1991). Southeast Asia and Hong Kong. In J. Cheshire (Eds.) *English around the world: sociolinguistic perspectives* (pp. 319-32). Cambridge: Cambridge University Press.
- Thomas, E. R. and Kendall, T. 2007. NORM: The vowel normalization and plotting suite. [Online Resource: <http://lingtools.uoregon.edu/norm/>]
- Van der Haagen, M. (1998). *Caught between Norms: The English pronunciation of Dutch learners*. Holland Academic Graphics: Hague.
- Van Engen, K. J., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., & Bradlow, A. R. (2010). The wildcat corpus of native-and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech*, 53(4), 510–540.

- Watt, D., & Fabricius, A. (2002). Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the F1 ~ F2 plane. *Leeds Working Papers in Linguistics and Phonetics*, 9, 159-73.
- Wordie, J. (2014, March 1). Then & now: forked tongues. *South China Morning Post*. Retrieved from <http://www.scmp.com/magazines/post-magazine/article/1435747/then-now-forked-tongues>.
- Yang, B., & Whalen, D. H. (2015). Perception and production of English vowels by American males and females. *Australian Journal of Linguistics*, 35(2), 121–141.
- Zahn, C. J., & Hopper, R. (1985). Measuring language attitudes: the speech evaluation instrument. *Journal of Language and Social Psychology*, 4(2), 113–123.
- Zhang, Q. (2009). Hong Kong People's attitudes towards varieties of English, *Newcastle Working Papers in Linguistics*, 15, 151-173.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin* 123:162–85.