ALVEOLARIZATION IN HONG KONG CANTONESE: A SOCIOPHONETIC STUDY OF NEOGRAMMARIAN AND LEXICAL DIFFUSION MODELS OF SOUND CHANGE

May T. M. Chan
Worcester College



A dissertation submitted to the

Faculty of Linguistics, Philology & Phonetics
University of Oxford

in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in
General Linguistics and Comparative Philology



Trinity Term 2017

Esther Yewpick Lee Millennium Scholarships



Wilkinson Junior Research Fellowship
Worcester College, Oxford

Abstract

This thesis is a quantitative study of sociophonetic variation which focuses on the Hong Kong Cantonese velar coda consonants **-ŋ** and **-k**. These codas have, under linguistic contexts which will be described in further detail, become increasingly realized as alveolar nasal and oral stop consonants, [-n] and [-t], respectively (Bauer 1979). For the purpose of this thesis, the phenomenon of sound shift from **-ŋ** and **-k**, to [-n] and [-t] respectively will be termed 'alveolarization' (which should be kept distinct from 'alveolarization' as a descriptive term for the coarticulatory phenomenon whereby an alveolar constricture is engaged[1]). Insofar as the language of a speech community is a shared vehicle for communication and sound changes are constrained by the need for mutual intelligibility, the central aim of this thesis is to uncover the factors which contribute the most to driving this sound change. In describing the variation in these consonants, I examine the concomitant social and linguistic factors which might help explain it. While this study focuses on one specific set of linguistic variables, it aims to analyse a broad set of factors to obtain a picture of the complexity of this sound change.

The sound change is of theoretical interest as it provides an opportunity to evaluate the neogrammarian regularity hypothesis and lexical diffusion models of sound change. The neogrammarian regularity hypothesis states that sound changes are regular and admit no exceptions—they are purely driven by phonological context and show no lexical variation. On the other hand, the lexical diffusion model predicts that sound changes progress through the lexicon one lexeme at a time. By examining the effects of neighbouring phonological environment on the velar codas, and by performing a qualitative analysis of which lexemes might be leading the change, it is possible to test both models of sound change, which have often been set in opposition to one another.

Lexical frequency effects of velar coda realization are also explored, as an extension of the lexical diffusion model of sound change, but also for the reason that Cantonese is one of the few modern Sinitic languages remaining of which the coda consonants have not undergone extensive lenition from Middle Chinese. Usage-based models have argued that leniting sound changes display sensitivity to lexical frequency, and if the velar codas in Cantonese are shown to display lenition, then one would expect lenition to occur in frequent lexemes first. While alveolarization cannot be considered an instance of lenition, but instead, the substitution of one articulatory place feature for

[1] Evidently, due to the linguistic focus of this dissertation, it should also be kept distinct from the physiological process whereby alveoli are formed in lung development, also termed "alveolarization."

another, there is no prior motivation for hypothesizing that it will display lexical frequency effects. Yet, if lexical frequency effects can be found, it would be strong evidence for the lexical diffusion model of sound change.

In order to find an answer to the above strands of questioning, it was necessary to go beyond the binary categorical distinction adopted in many traditional sociolinguistic analyses of consonantal variation and embrace the use of statistical tools and research methods that help us analyse variation of consonants beyond this dichotomy. A three-way distinction between velar, alveolar, and reduced variants was adopted in order to capture the possibility that velar coda consonants (**ŋ**) and (**k**) (i) retain their velar feature, (ii) adopt the innovative alveolar feature, or (iii) undergo lenition. This necessitated the use of multinomial logistic regression models. Furthermore, I explore lexical frequency as a continuous measure based on spoken corpus data from the Hong Kong Cantonese Adult Language Corpus (HKCAC: Leung, Law & Fung 2001).

The results of this study indicate that there is evidence of sound change in velar codas (**ŋ**) and (**k**), and that there are two concurrent processes: (1) an increase in the use of alveolar variants [n] and [t], but also (2) an increase in lenition with decreasing age. When comparing the realization of (**ŋ**) and (**k**) across styles, it was found that, curiously enough, there was a cross-over effect, whereby speakers showed an unexpectedly higher rate of [alveolar] realization in the minimal pair readings than in the passage and word list reading tasks. This, together with a steady overall increase in [velar] realization with an increase in the formality of speech styles suggests that (1) speakers were aware of the social undesirability of the lack of realization of a coda, but also that (2) they did not necessarily associate the innovative [alveolar] feature with social undesirability—the interpretation altogether being that it is the *lenition of coda* which is the *least* socially desirable. These findings contrast the case of sound changes in Cantonese (as well as other Sinitic languages)—which has an orthographic system lacking phonetic encoding—with sound changes in, say, English, where the spelling is a better indicator of 'proper' pronunciation and thereby makes it easier for certain sound changes to be above the level of social awareness (for instance, 'g'-dropping in 'ing' endings).

Furthermore, it was found that neogrammarian factors are strong predictors of realization of the coda consonants, with 'alveolar', 'labial', and 'palatal' features of the following onset significantly explaining the variation in realization of (**ŋ**) and (**k**) as [alveolar] and [reduced], versus the reference category [velar]. A qualitative study of lexical differences also showed that certain lexemes are more progressively realized as [alveolar] or [reduced]. Finally, lexical frequency effects were significant in explaining realization of (**ŋ**) and (**k**) as [reduced] as opposed to [velar]. In other words, the more

frequent a lexical item, measured as mean-centered 'log(frequency)', the more likely it underwent lenition. This confirms usage-based predictions that more frequent lexical items are more likely to undergo leniting sound changes. However, lexical frequency effects were not found to operate on the realization of (**ŋ**) and (**k**) as [alveolar] as opposed to [velar], suggesting that alveolarization is a sound change that is unlikely to show any lexical frequency effects. Combined with the observation that neogrammarian effects are found to operate on the realization of (**ŋ**) and (**k**), a suggestion is put forward that alveolarization might be more likely for morphosyllables that frequently collocate with onsets that promote alveolar realization. This was proposed as a promising direction for future investigation.

Acknowledgments

There are many people who I would like to thank for their support throughout the long journey of doctoral research. This research would not have been possible without their generous support and friendships.

First and foremost, I would like to thank Prof. John Coleman for his support throughout not only the D.Phil. but also throughout the master's course. Supervision meetings with John always offered a refreshing perspective and new angles of exploration, and his words of motivation, especially his contagious interest in Chinese linguistics, kept me going at times when the going was tough. He made the Phonetics Laboratory a welcoming second home for me, and his support for the interdisciplinary seminar *Speech Lunch* has allowed us to invite interesting speakers from diverse backgrounds whose research and work relate to speech. Heartfelt thanks go to Dr Lauren Hall-Lew, my external examiner and unofficial mentor. Lauren has been a generous source of support and advice, always ready to put me in touch with people in relevant areas of research and informing me of the latest developments in the field of sociolinguistics, all of which broadened my research horizons immensely. Her patient, thorough reading of my thesis and her numerous suggestions have vastly improved this thesis (of course, any remaining gaps and insufficiencies are entirely my responsibility). I thank Dr Ian Watson for acting as my internal examiner and his comments on the thesis. I am very much indebted to my College Advisor, Prof. Andreas Willi, for his council and advice at crucial times in the research. Dr Dan Lunn has kindly offered his expertise on statistical analysis and taught me how to perform multinomial logistic regressions in MLWin, how to use R, and advised me on current best practices in statistical analysis.

I thank members of the Oxford Phonetics Laboratory for many lively discussions over coffee, tea (and cake!). Anastassia Loukina, Ladan Baghai-Ravary, Greg Kochanski, and Jon Edwards have variously offered me glimpses into the world of engineering in linguistic research and broadened my perspective on happenings in the world, and their friendships were a highlight in the—at times—unforgiving days of linguistic data annotation.

I would also like to thank all the Cantonese informants who kindly offered their time and responded to my request to collect data with enthusiasm and readiness. Their role in this research was indispensable and the research would truly not have been possible without their assistance. I also thank Claire Leung for helping to recruit student volunteers, and the young informants from Law Ting Pong Secondary School for taking part in this study.

There are numerous people who have enriched my understanding of language beyond my immediate doctoral research project. In my short stint doing research on genetics of specific language impairment, I have enjoyed working with Roni Nudel—who taught me almost everything I know about this field, and whose contagious enthusiasm has made it possible for a linguist to learn how to perform genetic research. In addition to Roni, with the generous support of Paloma Garcia-Bellido, Dianne Newbury, and Daniela Moralli, it was possible to produce journal papers from the research results. With Paloma I have had particularly enjoyable and enlightening discussions about language and the mysterious workings of the brain, and it is thanks to her that we solved the mystery of the genetic underpinnings of a curious case of specific language impairment in a young Spanish girl. I also enjoyed having Fabiola Ceroni, Rose Reader, and Laura Covill as my lab mates during this time at the Wellcome Trust Centre for Human Genetics.

Generous financial support was provided by the Esther Yewpick Lee Millennium Scholarship, the Clarendon Scholarship, and the Wilkinson Junior Research Fellowship of Worcester College. Without this financial support it would not have been possible to pursue my master's and doctoral studies, and gain exposure to so many diverse fields of research.

I think back to the Worcester College community with fondness. Worcester College was a place I could truly call my home throughout the years I spent there. I thank many people in the college for their camaraderie and mentorship. I had the great fortune of being part of the decanal team while at Worcester. The team was not only inspiring to work with, but constituted a group of scholars with diverse backgrounds from whom I learnt a great deal. Such is the case with the (at times draconian and fearsome) Deans Peter Darrah and Ben Morgan (albeit in reality they were the most supportive counsellors a student in trouble could wish for), and the rest of the Dean's team: Nirmali Fenn (most talented composer), Daniela Simone (most knowledgeable copyright lawyer), Hannah Ryley (most passionate English scholar), and Huw Davies (coolest game music composer)—I learnt from them how to cope with the ups and downs, the joys and tribulations that life brings. I also forged college friendships with other members of the SCR, such as Sugata Kaviraj (the dancing physicist), Michael Goldsmith (the all-round polymath classicist-computer scientist), Peta Fowler (the jolly classicist), Don Fraser (the soft-spoken geophysicist), Provost Jonathan Bate (the world-renowned playwright and Shakespearean scholar), John Parrington (the well-read geneticist), and Mike Searle (the globe-trotting geologist). The venerable gardeners of Worcester (headed by Simon Bagnall) have built a stunning oasis, and walks in the college gardens have offered a peaceful sanctuary for the mind—and their artistry

maintaining the impeccable lawn in the main quad made it much easier to convince undergraduates not to step on the grass! I also thank the Porters Mark Norman and Norman Blackman, with whom I have forged friendships while solving tricky problems, and with whom I had many enjoyable chats on night rounds. Fellow MCR mate Melissa Schuh provided me with nutritious as well as spiritual sustenance throughout the last phase of doctoral research, and I am lucky to enjoy continued status as her proud "college Mum".

Friendships outside of the college walls have further enriched my life in Oxford. The many meals and lively discussions I have shared with Kathryn Chung, Betsey Tsoi, Lynn Yu, Sunny Wong, Timothy Cheng, Karita Kan, Siron Ng, Samuel Ng, Sook Teng Vun, Cheng Cheng, Leo Goodstadt, Terry Wong, Jamie Wong, Cherry Xu, Benjamin Pope, Stephan Stiller, and Oliver Holton, to name but a few, have open my eyes to different disciplines and perspectives.

I thank my sister Evelyn Chan for her unwavering support. She has not only been my role model for as long as I remember, but my closest friend regardless of the geographical distance between us, always readily offering emotional support, encouragement, and guidance. I am lucky to count both Evelyn and Arjan as two of my closest friends. My parents have always allowed me the freedom to explore and pursue my interests, and I thank them for always being there for me. I thank my brother for bringing a lot of joy and good humour—I hope he will take us on more outdoor adventures!

I thank O.B., walking encyclopaedia and mentor, whose intellect, good nature, warmth and kindness have been an inspiration ever since we met. She taught us that small acts of kindness and a willingness to invest time in people can make a world of difference to another person's life. Her wisdom will always remain a source of moral guidance to me, and I hope to pass on however much I can to other young people in need.

I also thank Georg for his constant, steadfast support. He has stood by me through ups and downs, enriching my life in Oxford and beyond. I cannot wish for a better life companion and count myself very fortunate to be his partner.

Last but not least, my thanks go to David Parker, who is dearly missed, and whose belief in my abilities has made one of the most interesting chapters of my life possible.

I dedicate this thesis to the late Professor David Parker, who left behind a legacy of scholarship and wisdom to generations of students from Hong Kong.

Contents

Abstract i
Contents viii
List of Tables xii
List of Figures xv
List of Statistical Models xvii
Notation xviii
1 Introduction 1
1.1 The study of sound change 1
1.2 Theoretical models of sound changes 3
1.3 Neogrammarian sound change 3
1.3.1 The beginnings of the Neogrammarians 3
1.3.2 Homogeneity of the speech community and "exceptionless" sound change 6
1.4 Reactions against neogrammarian sound change and the earliest ideas on lexical diffusion 9
1.5 Lexical diffusion 13
1.5.1 Studies on lexical diffusion 16
1.5.1.1 The Chaozhou tone split 17
1.5.1.2 Homonym split from Middle Chinese to modern Mandarin (Peking dialect) 20
1.5.1.3 Lexical diffusion of Skikun velarization 22
1.5.1.4 Lexical diffusion evidence from other languages 24
1.5.1.5 Lexical diffusion of North American English short **a** tensing in Philadelphia 26
1.5.2 Lexically and phonetically gradual sound change 27
1.5.3 Summary of evidence for lexical diffusion 31
1.6 The neogrammarian vs lexical diffusionist "controversy" over regularity of sound change 31
1.7 Lexical frequency effects 42
1.8 Research questions 48
1.9 Quantitative sociophonetic variation 50
1.10 Aims of this thesis 53
1.11 Organisation of this thesis 55
2 Linguistic variables for this study: alveolarization in Hong Kong Cantonese 58
2.1 Alveolarization as a sound change in progress 59
2.2 Sound changes in Hong Kong Cantonese 62

2.3 Hong Kong Cantonese 64
2.3.1 Cantonese phonetics and phonology....67
2.3.2 The Cantonese syllable structure 68
2.3.3 Onsets....69
2.3.4 The rime: nucleus and coda 71
2.3.5 Tones....76
2.4 Phonological conditioning of **-ŋ** to [-n] variation 81
2.5 Phonological conditioning of **-k** to [-t] variation 87
2.6 Cross-dialectal motivation for the study of sound changes in Cantonese coda consonants....87
2.7 Historical development of Yue stop consonants from Middle Chinese....90
2.8 Summary 96
3 Methodology....98
3.1 Sampling....98
3.2 Informants 99
3.3 Recording set-up 101
3.4 Social variables 101
3.5 Speech styles 105
3.5.1 Casual speech elicitation....106
3.5.2 Formal tasks....107
3.5.2.1 Passage reading 107
3.5.2.2 Word list reading task....110
3.5.2.3 Minimal pairs reading task 112
3.6 Lexical frequency data 112
3.6.1 Lexical frequency as a continuous variable....118
3.6.2 The problem of zero frequency and Laplace smoothing 120
3.6.3 Limitations of using corpus-based frequency data....121
3.7 Data processing 123
3.8 Coding and quantifying tokens 124
3.8.1 Phonetic criteria used for coding (**ŋ**) 127
3.8.1.1 Tokens of (**ŋ**) coded as [velar]....127
3.8.1.2 Tokens of (**ŋ**) coded as [alveolar]....129
3.8.1.3 Tokens of (**ŋ**) coded as [reduced]....130
3.8.2 Phonetic criteria used for coding (**k**) 132
3.8.2.1 Tokens of (**k**) coded as [velar]....132
3.8.2.2 Tokens of (**k**) coded as [alveolar]....134
3.8.2.3 Tokens of (**k**) coded as [reduced]....135
3.8.3 Going beyond a binary distinction....137

3.9 Statistical analysis 137
3.9.1 Tools used for statistical analysis 138
3.9.2 Multilevel modelling 140
3.9.3 Multilevel modelling procedure 141
3.10 Summary 146
4 Full statistical model combining social, neogrammarian and lexical factors 148
4.1 The full modelling procedure 148
4.2 Full model for (**ŋ**) realization 150
4.3 Full model for (**k**) realization 157
4.4 Summary and discussion 163
5 Social dimensions of sound change 166
5.1 Age-based differences 166
5.1.1 Age-based effects on (**ŋ**) variation 166
5.1.2 Age-based effects on (**k**) variation 173
5.1.3 Summary of age-based variation of (**ŋ**) and (**k**) 177
5.2 Gender 178
5.3 Age and gender interaction 179
5.3.1 Age and gender interaction effects on (**ŋ**) to [reduced] variation 180
5.3.2 Age and gender interaction effects on (**k**) to [alveolar]/[reduced] variation 182
5.4 Speech style 184
5.4.1 Speech style-based effects on (**ŋ**) variation 184
5.4.2 Speech style-based effects on (**k**) variation 187
5.4.3 Interaction between speech style, age, and gender 190
5.5 Discussion 193
5.6 Summary 199
6 Neogrammarian regularity hypothesis 201
6.1 Preceding vowels 201
6.2 Following onsets 209
6.3 Tautosyllabic onset 215
6.4 Influence of tone 217
6.5 Suspension of phonological contrast 219
6.6 Discussion 220
6.7 Summary 222
7 Lexical diffusion 224
7.1 Lexical frequency effects 224
7.2 Which lexemes lead the change? 229
7.3 Homophonous lexemes 236
7.4 Inter- and intra-speaker lexical variation 240

7.5 Discussion 242
7.6 Summary 244
8 Conclusion 246
8.1 Brief review 246
8.2 Future directions 251
8.3 Concluding remarks 251
References 254
Appendices 268
Appendix I. Questions for casual speech elicitation 268
Appendix II. Passage reading 273
Appendix III. Word list 274
Appendix IV. Minimal pairs 277
Appendix V. Text files used for forced alignment 278
Appendix VI. Variables coding scheme 282
Appendix VII. Lexical distribution of the HKCAC 284
Appendix VIII. Raw data 288
Appendix IX. Mathematical modelling results 294

List of Tables

Table 1. Distribution of lexical occurrences of Middle Chinese tone III reflexes in modern Chaozhou lexical tone by MC voiced onsets 18
Table 2. Distribution of lexical occurrences of Middle Chinese tone III reflexes in modern Chaozhou lexical tone by MC sonorant onsets 19
Table 3. Mandarin -a/ -ai alternation 21
Table 4. Lexical diffusion pattern of word-final labial to velar shift in Skikun 23
Table 5. Dropping of voiceless uvular fricative in some Welsh dialects 25
Table 6. The *merry-marry* merger continuum 26
Table 7. Lexical variants among speakers in the transition zone between northern and southern English varieties for the variable (u) 29
Table 8. Variable rule analysis of Skikun labial-velar alternations 34
Table 9. Labov's prediction of the types of sound changes which are regular and those which are lexically diffused 37
Table 10. A typology of sound changes displaying lexical diffusion* 41
Table 11. Repetition of a common sentence by three speakers 44
Table 12. Repetition of an uncommon sentence by three speakers 44
Table 13. Rate of t/d-delection by word frequency 47
Table 14. Percentage of innovative alveolar realization of (ŋ) across age groups and speech styles, in Wong (2005) 60
Table 15. Cantonese consonants 70
Table 16. Aspirated and unaspirated onset minimal pairs 70
Table 17. Cantonese nuclear vowels 71
Table 18. Cantonese vowel space* 72
Table 19. Standard Cantonese rimes 73
Table 20. Cantonese phonemes and allophones* 75
Table 21. Numerical and IPA notation for tones in Cantonese 78
Table 22. Lexical items undergoing complete change from -aːŋ > [-aːn] 82
Table 23. Lexical items undergoing complete change from -ɐŋ > [-ɐn] 82
Table 24. Lexical items with -ɪŋ / -ɛːŋ finals displaying varying alveolarization 83
Table 25. Lexical items with -œːŋ finals displaying varying alveolarization 84
Table 26. Lexical items with -ɔːŋ finals displaying varying extent of alveolarization 85
Table 27. Alveolarization of -ŋ by preceding vowel 86
Table 28. Cross-dialectal comparison of nasal and stop coda attrition from Middle Chinese 89
Table 29. Historical development of Middle Chinese *-k to -t in Yue dialects 92
Table 30. Lechang dialect showing *-k to -t, in contrast with Hong Kong Cantonese 94

Table 31. Change from Middle Chinese *-t to -k in Yue dialects 95
Table 32. Token distribution in formal reading passage 110
Table 33. Token distribution in word list reading task 111
Table 34. Hong Kong Cantonese Adult Corpus (HKCAC) type frequency summary statistics 115
Table 35. Type frequency and rank minimum and maximum in this study 116
Table 36. Lexical frequency coding information 117
Table 37. Genres covered by the Hong Kong Cantonese Adult Corpus 122
Table 38. Factor weights and corresponding log-odds 146
Table 39. Mixed-effects multinomial logistic regression model results: (ŋ) > [n] variation (part 1) 152
Table 40. Mixed-effects multinomial logistic regression modelling results: (ŋ) > [reduced] variation (part 2) 154
Table 41. Mixed-effects multinomial logistic regression model results: (k) > [t] variation (part 1) 158
Table 42. Mixed-effects multinomial logistic regression modelling results: (k) > [reduced] variation (part 2) 161
Table 43. Multinomial logistic regression results for age-based variation of (ŋ) 166
Table 44. Types of sound changes at the level of the individual vs. community (Labov 1994:83) 170
Table 45. Multinomial logistic regression results for age-based variation of (k) 174
Table 46. Multinomial logistic regression results for gender-based variation of (ŋ) 178
Table 47. Multinomial logistic regression results for gender-based variation of (k) 178
Table 48. Multinomial logistic regression results of age and gender interaction in (ŋ) variation 179
Table 49. Multinomial logistic regression results of age and gender interaction in (k) variation 179
Table 50. Raw counts of (ŋ) realization by age and gender group 181
Table 51. Raw counts of (k) realization by age and gender group 184
Table 52. Multinomial logistic regression results for speech style-based variation of (ŋ) 185
Table 53. Raw counts of (ŋ) realization by speech style 187
Table 54. Multinomial logistic regression results for speech style-based variation of (k) 187
Table 55. Raw counts of (k) realization by speech style 189
Table 56. Multinomial logistic regression results of speech style and gender interaction in (ŋ) variation 190

Table 57. Multinomial logistic regression results of age and gender interaction in (k) variation 190
Table 58. Multinomial logistic regression results of speech style, age and gender interaction in (k) variation 190
Table 59. Multinomial logistic regression results for preceding vowel effects on (ŋ) variation 202
Table 60. Multinomial logistic regression results for preceding vowel effects on (k) variation 203
Table 61. Realization of (ŋ) by preceding vowel in all speech styles 205
Table 62. Realization of (k) by preceding vowel in all speech styles 206
Table 63. Multinomial logistic regression results for following onset effects on (ŋ) variation 209
Table 64. Multinomial logistic regression results for following onset effects on (k) variation 210
Table 65. Raw counts of (ŋ) realization by following onset, including percentage [-n] 213
Table 66. Raw counts of (k) realization by following onset, including percentage [-t] 214
Table 67. Multinomial logistic regression results for tone effects on (ŋ) variation 217
Table 68. Comparison of rate of [n] realization in rime sets with and without alveolar counterparts 220
Table 69. Multinomial logistic regression results for lexical frequency effects on (ŋ) variation 225
Table 70. Multinomial logistic regression results for lexical frequency effects on (k) variation 225
Table 71. Multinomial logistic regression results for lexical frequency * preceding vowel effects on (ŋ) variation 230
Table 72. Raw counts and percentages of (ŋ) realization by individual speaker 288
Table 73. Raw counts and percentages of (k) realization by individual speaker 289
Table 74. Realization of (ŋ) by frequency as a categorical variable (HKCAC frequency < 7: infrequent; HKCAC frequency ≥ 7 frequent) 290
Table 75. Realization of (k) by frequency as a categorical variable (HKCAC frequency < 7: infrequent; HKCAC frequency ≥ 7 frequent) 291

List of Figures

Figure 1. Development of modern Chaozhou tones from Middle Chinese tones from different dialectal and literary/colloquial substrata....20
Figure 2. Effects of initial consonant features on velar realization of coda consonants in Skikun....36
Figure 3. Average percentage of [-n] realization by age group, in Wong (2005)....61
Figure 4. Sinitic language families in China....65
Figure 5. Location of Hong Kong on a map of China, also showing Guangzhou, formerly 'Canton', from which Cantonese originated....66
Figure 6. Hong Kong bordering Shenzhen and the Pearl River Delta....66
Figure 7. Cantonese syllable structure....68
Figure 8. Schematic pitch diagrams of the Cantonese lexical tones....77
Figure 9. Fundamental frequency (F0) values of lexical tones in Hong Kong Cantonese....79
Figure 10. Development of final nasal and stop consonants in Chinese dialects....88
Figure 11. Informants' age distribution....103
Figure 12. Token of (ŋ) coded as [velar]; Speaker: MM35....128
Figure 13. Token of (ŋ) coded as [alveolar]; Speaker: MM33....130
Figure 14. Token of (ŋ) coded as [reduced]; Speaker: OM61....131
Figure 15. Token of (k) coded as [velar]; Speaker: OM53B....133
Figure 16. Token of (k) coded as [alveolar]; Speaker: MM33....134
Figure 17. Token of (k) coded as [reduced]; Speaker: OM61....136
Figure 18. Percentage (ŋ) realization by speaker ranked by age....168
Figure 19. Conditional density curve of (ŋ) realization by age....169
Figure 20. Percentage (k) realization by individual speaker....175
Figure 21. Probability density curve of (k) realization by age....176
Figure 22. Percentage (ŋ) realization by age and gender group....181
Figure 23. Percentage (k) realization by age and gender group....183
Figure 24. Percentage realization (k) by speech style....189
Figure 25. Percentage (ŋ) realization by age group across speech styles....192
Figure 26. Percentage (k) realization by age group across speech styles....192
Figure 27. Percentage (ŋ) realization by preceding vowel, all speech styles combined....205
Figure 28. Percentage (k) realization by preceding vowel, all speech styles combined....205
Figure 29. Comparison of percentage realization of (ŋ) as [n] by preceding vowel between Wong (2005) and current study....208
Figure 30. Percentage (ŋ) realization by following onset....213
Figure 31. Percentage (k) realization by following onset....214

Figure 32. Percentage (ŋ) realization by tautosyllabic onset....215
Figure 33. Percentage (k) realization by tautosyllabic onset....216
Figure 34. Percentage (ŋ) realization by tone....218
Figure 35. Percentage (k) realization by tone....219
Figure 36. Conditional probability density curve of (ŋ) realization by log(frequency)....228
Figure 37. Conditional probability density curve of (k) realization by log(frequency)....228
Figure 38. Realization of lexemes with -ɐŋ rime by frequency....231
Figure 39. Realization of lexemes with -aːŋ rime by frequency....232
Figure 40. Realization of lexemes with -ɛːŋ rime by frequency....233
Figure 41. Realization of lexemes with -œːŋ rime by frequency....233
Figure 42. Realization of lexemes with -ɔːŋ rime by frequency....234
Figure 43. Realization of lexemes with -ɪŋ rime by frequency....234
Figure 44. Realization of lexemes with -ʊŋ rime by frequency....235
Figure 45. Percentage of (ŋ) realization by homophonous lexemes....237
Figure 46. Percentage [ŋ] realization, most frequent characters in word list reading task....241
Figure 47. Percentage [ŋ] realization of least frequent characters in word list reading task....241
Figure 48. Scatter plot of frequency by rank in HKCAC data....284
Figure 49. Scatterplot of log(frequency) by log(rank) of HKCAC data with linear regression line....285
Figure 50. Scatter plot of log(frequency) by $1+\sqrt[3]{rank}$ in HKCAC data with linear regression line....286
Figure 51. Frequency by rank of words in the HKCAC....287
Figure 52. Realization of (ŋ) by frequency as a binary categorical variable (HKCAC frequency < 7: infrequent; HKCAC frequency ≥ 7 frequent)....290
Figure 53. Realization of (k) by frequency as a categorical variable (HKCAC frequency < 7: infrequent; HKCAC frequency ≥ 7: frequent)....291
Figure 54. Percentage realization of (ŋ) by log(freq)....292
Figure 55. Percentage realization of (k) by log(freq), raw data....293

List of Statistical Models

Model 1. “Best” mixed-effects multinomial logistic regression model for response variable “(ŋ) realization” obtained in MLwiN.....294
Model 2. “Best” mixed-effects multinomial logistic regression model for response variable “(k) realization” obtained in MLwiN.....297

Notation

In this thesis, phonological units will be in **bold** type[2], and IPA symbols are used to present morphosyllable subcomponents (onset, nucleus, and coda) along with tone, e.g. **kɔ**:**ŋ˥**. Wherever relevant, the Chinese character is included, along with a gloss, e.g. **kɔ**:**ŋ˥** 講 'to speak / tell'. Phonetic units will be within square brackets, e.g. [kɔːn].

The linguistic variables studied in this thesis are the velar nasal **-ŋ** and velar oral stop consonant **-k** in the coda position. As this is a sociophonetic study, and variation between speakers with respect to the use of these coda consonants will be explored, the sociolinguistic practice of denoting these linguistic variables using ( ) will be adopted here, such that I will be referring to variables (**ŋ**) and (**k**).

[2] Except when quoting directly from other texts which use solidi '/ /' for phonological units.

1 Introduction

1.1 The study of sound change

The wider aim of this thesis is to understand sound change in Hong Kong Cantonese. At a recent conference, when I presented some of my preliminary research results, there was a younger Cantonese speaker who was astonished to learn that 講 'to talk / tell' is in fact **kɔːŋ˨** in standard descriptions and used by older speakers of Cantonese, because in his idiolect this was firmly [kon], with an alveolar nasal coda. Going back two or three generations before him, [koŋ] would probably have been the norm. Somewhere along the way **-ŋ** became [-n], and **-k** became [-t] in the coda position. Examples like this beg the question, how widespread is alveolarization of **-ŋ** and **-k** in Hong Kong Cantonese today? Is it time to revise the accounts of the phonology of present-day Cantonese? But beyond finding out *what* happened, I aim to build upon the work of linguists who have tried to answer the theoretical and methodological question of how one can understand and explain such a sound change.

Research on sound change in progress, such as the study of centralization of **aw** and **ay** (in words such as 'mouse' and 'life') in Martha's Vineyard (Labov 1963), has shown that the changing social orientation of speakers through time affected their use of **aw** and **ay**, thereby using variation patterns within the speech community to explain how the sound change progressed. The repeated finding of ordered heterogeneity (Weinreich, Labov & Herzog 1968) within the speech community as fundamental to language change meant that any investigation into sound change could not comfortably ignore variation within the social dimension. To understand sound change, then, requires examining the social groupings that make up the relevant speech community.

This principle informs the basic methodological framework adopted in this dissertation, and broad social distinctions of gender and age will allow us to see who the main social drivers of the sound change might be. Note that such broad social categories are not without their problems. Eckert (1989:265), for instance, has criticised such approaches as problematic, for categories such as "gender" are treated as stable, immutable properties pertaining to speakers. Such criticism has generally been acknowledged by sociovariationists who employ such broad categories (e.g. Labov 2001:272). However, these social distinctions still remain instrumental as part of the researcher's elementary toolkit for obtaining a general picture of the extent of variation one might find within the speech community.

Furthermore, using age as a social category and employing the assumption that the rate of linguistic change over a speaker's life time is negligible[3], we can sample speakers from different age groups and compare the differences in their use of **-ŋ** and **-k**. Using age as an independent factor then, allows us to perform an apparent-time analysis of how swiftly a sound change has swept through a speech community.

Beyond social groupings, language-internal factors such as preceding and following phonological environments have been investigated in studies of sound change to identify how changes spread through the linguistic system. In this thesis, the focus will be on the immediate phonetic and phonological environments of **-ŋ** and **-k**: preceding vowel, tone, and following onset consonant. The phonological structure of Cantonese and the linguistic variables in this study will be described in more detail in Chapter 2. This chapter also briefly describes the Hong Kong Cantonese speech

[3] Although, refer to, for instance, Sankoff & Blondeau (2007), an investigation into the change of **r** from apical to dorsal in Montreal French, which found that some speakers did change their speech behaviour substantially throughout their life time. For a further discussion of using age as an independent factor for analysing sound change, refer to Section 3.3.

community (*who* are adopting the sound change), Cantonese phonology (*what* is the sound change), and what we know so far about how quickly the sound changes are progressing (*how* is the sound change transmitted through time). In the remainder of this introductory chapter, however, the focus is on the theoretical motivation behind this study.

1.2 Theoretical models of sound changes

Two widely discussed theoretical models of sound change are the neogrammarian regularity hypothesis and lexical diffusion. What follows is an introduction to the two models by guiding the reader through a brief historical exposition of the context against which these models were conceived and the predictions they make.

1.3 Neogrammarian sound change

1.3.1 The beginnings of the Neogrammarians

Until the 1860s, many comparative philologists studying sound change (in particular followers of Schleicher 1848) held the idea that the present state of language is a degenerate form of a perfect past state known as *Ursprache*, and as such, unworthy of investigation. The main interest for these philologists was to reconstruct Proto-Indo-European on the basis of the *Stammbaumtheorie* (or the "family-tree model") of language families. To this end, they used textual records to extract principles which could reconcile differences between cognates from related languages such that they would have a common form in Proto-Indo-European. The principles of sound change, then, were established with the aim of showing how one hypothetically unified language developed

into a number of observably different, related languages. Reconstructing the *Ursprache* was the objective, while contemporary language was hardly ever the focus of interest.

However, in the 1860s, a group of researchers, or the *Junggrammatiker* (Neogrammarians[4]) as they called themselves, started to question this approach. Initial resistance to the focus on *Ursprache* was offered amongst others by Scherer (1868), who suggested that a better understanding of sound change could be obtained if speakers and the mechanisms of natural speech were investigated. Influenced by the ideas put forward by Scherer, Osthoff and Brugmann's preface (1878) draws attention to the importance of understanding the mechanisms of speech in the study of sound change:

> The older linguistics, as no one can deny, approached its object of investigation, the Indo-European languages, without first having formed a clear idea of how human speech really lives and develops, which factors are active in speaking, and how these factors working together cause the progression and modification of the substance of speech. Languages were indeed investigated most eagerly, but the man who speaks, much too little.
>
> (Osthoff & Brugmann 1878, translation in Lehmann 1967:198)

Articulatory phonetics had already been of interest to linguists for decades by then, but had hardly been combined with the study of sound change. The *Junggrammatiker* wanted to revolutionise the study of sound change by calling for a systematic study of the physical (articulatory phonetic) and psychological aspects of speech. It is clear from the following excerpt just how much importance was placed by the Neogrammarians on understanding the present to explain the past.

---

[4] In this thesis, when referring to the original *Junggrammatiker*, the word "Neogrammarian(s)" is capitalized. Whenever the adjective "neogrammarian" is employed to describe sound changes or related ideas, the word is uncapitalized.

> We believe: as sure as we are that our Indo-European forefathers had need, just as we do, of their lips, tongue, teeth, etc. for the articulation of the sounds of their language, just as sure can we be of the fact that the entire psychological aspect of their speech activity (the emergence of sound images preserved in the memory from a subconscious state, and the development of concepts of sounds to words and sentences) was influenced by the association of ideas in the same way and in the same measure as today and as long as people are people.
>
> (Osthoff & Brugmann 1878, translation in Lehmann 1967:206)

Without an understanding of the phenomena which occur in speech communication *in situ*, they argued, any reconstruction of the original forms in Indo-European must necessarily be fruitlessly "[going] around in the most obvious circles" (ibid., 200). They add,

> Is not, after all, the authenticity, the scientific probability, of the original Indo-European forms, which are of course all purely hypothetical creations, totally dependent on whether they agree in general with the proper conception of the development of linguistic forms and on whether they are constructed according to correct methodological principles?
>
> (Osthoff & Brugmann 1878, translation in Lehmann 1967:200)

The reconstruction of Indo-European, they believe, should not arrive at conclusions which are implausible or which do not adhere to basic mechanisms by which sound changes can progress. This preface is often considered to be the Neogrammarian manifesto. By distinguishing their own approach from the practice of existing comparative philologists, the Neogrammarians called for a new direction in the study of sound change with a chief focus on understanding contemporary speech in order to explain the past, which was a phenomenally important change in direction for the history of linguistics.

1.3.2 Homogeneity of the speech community and "exceptionless" sound change

One key neogrammarian assumption in comparing diachronic states of dialects of sound change was synchronic homogeneity of the speech community. The Neogrammarians derived this idea from Winteler's (1876) dialectological work as evidence for this observation, with Osthoff and Brugmann asserting:

> In all living dialects the shapes of sounds peculiar to the dialect always appear much more consistently carried out throughout the entire linguistic material and maintained by the members of the linguistic community than one would expect from the study of the older languages accessible merely through the medium of writing; this consistency often extends into the finest shades of a sound.
>
> (Osthoff & Brugmann 1878, translation in Lehmann 1967:201–2)

The idea was that a speech community speaking the same dialect share sounds much more consistently alike than one might get an impression of in written texts. Furthermore, Leskien (1876; cf. Lehmann 1967:197), amongst others, had argued that phonetic changes operated in a regular, or "exceptionless" manner. Osthoff and Brugman adopted this idea, when they laid out a new foundation and objectives for studying sound changes, and suggested how historical reconstructions were to be made. In their own words, from an oft-quoted passage from Osthoff and Brugman (1878):

> First, every sound change, inasmuch as it occurs mechanically, takes place **according to laws that admit no exception**. That is, **the direction of the sound shift is always the same for all members of a linguistic community** except where a split into dialects occurs; and **all words in which the sound subjected to the change appears in the same relationship are affected by the change without exception**. Second, since it is clear that form association, that is, the creation of new linguistic forms by analogy, plays a very important role in the life of the more recent languages, this type of linguistic innovation is to be recognized without hesitation for older periods too, and even for the oldest. [My emphasis]
>
> (Osthoff & Brugman 1878, translation in Lehmann 1967:204)

As this quotation suggests, the Neogrammarians proposed that the study of sound change should follow the principle that:

> Mechanical sound change follows laws which allow no exceptions:
>
> a. it applies to all members of the linguistic community, unless there is a dialect split
> b. it applies to a given sound contained within the same phonological environment in *all words*

One has to note in particular this last point, which has important consequences for an explanation of the mechanism of sound change. If one can state with certainty that a given sound A changed to another sound B over time $t = t_0$ to t_n, and that *all* sounds A in the environment of C followed this course to sound B at the same time, and given that sound changes do not occur abruptly for all words overnight, then the only logical possibility would be that A advanced towards B in phonetically imperceptible increments, and that all tokens of A sounded the same (as long as they are in the environment of C) in all words in each of any given synchronic time frames from t_0 to t_n, i.e. when $t = t_1, t_2, t_3$, etc. Following this logic, neogrammarian sound change has been interpreted by subsequent scholars to mean a sound change that is *phonetically gradual* and *lexically abrupt*.

It is due to the above firm statement of inadmissibility of irregularities that the neogrammarian approach has earned the title "regularity hypothesis", a term used by Hockett (1965:186). Thus, sound changes strictly follow rules or laws, every speaker of the language will be equally affected by them, and a given sound change affects every word in the lexicon as long as the phonological environment is the same. In the neogrammarian approach, any exceptions to regular sound changes have been taken to be the result of analogy and paradigmatic levelling (Chen & Wang 1975: 256). In other words, if a given instance was found which did not accord to the regular sound laws, it was proposed that this was due to analogy or paradigmatic levelling. Another important point to remember is that Osthoff and Brugmann intended for the above laws to only apply to "mechanical sound change". In a footnote, they explain:

> We are of course speaking here only of mechanical sound change, not of certain phenomena of dissimilation and metathesis. The reason for the latter lies in the peculiarity of the words in which they appear; they are always the physical reflections of purely psychological stimuli and in no way nullify the concept of sound law.
>
> (Osthoff & Brugman 1878, translation in Lehmann 1967:209).

The Neogrammarians' intent to limit such sound laws to only those of a mechanical nature is often overlooked by subsequent scholars who discuss neogrammarian sound change—typically proponents of lexical diffusion, and who have assumed neogrammarians apply regular sound laws to *all* sound changes.

The neogrammarian model of sound change has remained deeply influential and a methodological basis on which comparative historical linguistic studies and reconstructions are conducted. For examples of reconstructions, see Hockett (1965:185),

and Song (2008) for an overview of Chinese reconstructions. Yet, while the comparative method proved fruitful in discovering generalizations of language change—and one of its major strengths was that it made it possible to "trace many linguistic changes over long periods of time" (Labov 1994:10)—this comparative method based on neogrammarian assumptions also had the major drawback that it assumed complete homogeneity of the speech community. The comparative method might be sufficient for those who are content with studying an artificially "normative dialect that was never any speaker's native language" (Labov 1994:11), but anyone who is interested in describing language change in the vernacular of speakers will find that this approach offers limited use for the study of language changes as they occur in real language settings, and an even more limited perspective on how such changes came about or were propagated within a speech community.

1.4 Reactions against neogrammarian sound change and the earliest ideas on lexical diffusion

The neogrammarian "manifesto" was influential in bringing about a shift in thinking about how to study sound changes, and especially instrumental in moving the focus solely on texts towards gaining a deeper understanding of the mechanisms of articulatory phonetics. But the idea that sound changes could only be exceptionless and affect all words in a given language at the same time was not uncontroversial, even at the time, which brings us to a discussion of the earliest instantiations of the lexical diffusion view of language change. The seeds for a lexical diffusion analysis of sound change were planted at least as far back as the late 19th century by Schuchardt (1885), amongst others. Schuchardt was a staunch opponent of the Neogrammarians, as he

disagreed with the fundamental assumptions in their approach to the study of sound change. He argued vehemently against the neogrammarian position, as he believed that postulating exceptionless neogrammarian sound laws, was "made untenable by the same set of facts that refutes the similarity of these laws to natural laws," as their postulation did not accord with empirical observations (Schuchardt 1885:10).

Schuchardt suggested the following as one example to illustrate why the lexical dimension of sound change should be investigated and that the researcher should not assume regularity of sound changes categorically:

> Modern French *e*—sometimes open, sometimes close, the difference is not essential here—in *chef*, *fève*, *pré*, *tel*, *mer*, *nez*, *ème*, *lène* = *caput*, *faba*, *prato*, *tale*, *mare*, *naso*, *amat*, *lana*, corresponds to Gallo-Vulgar Latin *ā* (classical *ā́* and *a̋* before a single consonant). The following consonant appears to be a matter of indifference. This was not so in Old French. This fact is reflected in present-day orthography: *chef*, etc., but *áime*, *láine*. If now *ā* before *m* and *n* has become *e* passing through the stage *ai*, could this not likewise have happened before the other consonants? And if *chaif*, *faive*, *tail*, *mair* were originally pronounced, it is conceivable that in a more recent period *chaif*, *faive*, *tel*, *mer* occurred.
>
> (Schuchardt 1885:24, translation in Vennemann & Wilbur 1972)

Based on the fact that the vowels in *áime* and *láine* are pronounced like *e* in *chef* in Schuchardt's time and, while the orthography dating back to Old French suggests that at some point *ai* was pronounced for *áime* and *láine*, Schuchardt proposed that, over time, the sound change from Gallo-Vulgar Latin *ā* to Modern French *e* progressively spread through the original set of lexical items containing *ā*. Had the sound change operated in an exceptionless manner, the sound change from *ā* to *e* should have affected all lexical items in the original lexical set containing *ā* at the same time and one should therefore not find any discrepancy in orthography. This was one of the examples that Schuchardt

used to illustrate that the way the Neogrammarians posited exceptionless sound laws was untenable.

Schuchardt proposed that sound changes could more aptly be described as occurring in successive waves affecting parts of the lexicon at a time. This idea was known as the *Wellentheorie*, or the "wave theory" of sound change. Although the theory is often attributed to Schmidt (1872), there are suggestions that Schuchardt had discussed this idea before Schmidt's publication (see Fought 1982:424; Campbell 2004:213). In this view, sound changes did not always affect all words in a lexicon at the same time, and the speech community was not assumed to be homogeneous. The influence of the sound changes was likened to ripples on water emanating from the location of impact, outward to more peripheral localities. In this model, a given word X sampled in one geographical location could be influenced by a sound change, with yet another word Y sampled from speakers in the same geographical location not influenced by the sound change (where X and Y both contain a segment in the same phonological environment that should otherwise be equally influenced by a sound change). Hence arose the idea of "*chaque mot a son histoire*" or "each word has its own history," an idea that can be traced back to Schuchardt although it is attributed to Gilliéron (1902). The wave model could apply to differences in geographical locality. It could be that word X containing the sound of interest is spoken one way in one geographical locality, while the same word X is spoken a different way in a different geographical locality.

One example that is often used to illustrate the "wave theory" model and the idea of each word having its own history is the change from Latin **k** > **ʃ** before **a** or front vowels in French, such that the Latin word *cantāre* 'to sing' became French *chanter*. However, in Normandy, it was found that the Latin **k** was retained in some French

words, such as *chaîne* (from Latin *catēna*, 'chain'), *chambre* (*camera*, 'room'), *champ* (*campus*, 'field'), *chandelle* (*candēla*, 'candle'), *chanson* (*cantiō*, 'song'), *chanter* (*cantāre*, 'to sing'), and *chat* (*cattu*, 'cat'). When isoglosses were plotted for the places in the Normandy region where these words were pronounced with a **k**, it was found that the isoglosses formed smaller regions which partially overlapped, providing evidence for the "wave" patterned phenomenon (for more details and a map, see Campbell 2004:214; Palmer 1972:273).

Schuchardt appeared intent to effect a change in the practice of linguistic study through relaxing the strict assumptions and confines imposed by the Neogrammarians as to how sound change should be studied. In Fought's interpretation of Schuchardt's position:

> [i]t is not so much the principle of regularity itself that he objects to […] as it is the procedures of phonological segmentation and the disregard of the integral word presupposed by it. It is difficult to take this part of his argument seriously nowadays, so accustomed are linguists to the segmentation and extraction of sounds from their contexts, but for Schuchardt it was an important issue, and so it must be understood.
>
> (Fought 1982:422)

His vision for a more lexically- and contextually informed approach to sound change, and better understanding of the mechanisms of sound change were largely overlooked at the time.[5] In all, the neogrammarian method prevailed within the study of sound change, especially in historical linguistic accounts.

[5] Not only has his idea of lexical diffusion not been accepted by Neogrammarians, his character was also called into question. For instance, (Markey 1979:xxii) has characterised Schuchardt as having an "almost incredible whimsy and irresponsible dilettantism" (Fought 1982:429).

1.5 Lexical diffusion

From roughly the 1960s, researchers started questioning how tenable the neogrammarian notions were in studying the mechanisms of sound change, and more effort was devoted to investigating the lexical dimension. One landmark in this movement was William S.-Y. Wang's seminal paper "Competing changes as a cause of residue" (1969), in which he argued for the need to re-examine prevalent assumptions about the nature of sound change. With Wang's call for the study of sound change from a lexical perspective, research efforts were more systematically directed towards investigating the lexical dimension, with the result that this research direction gradually expanded from the 70s and 80s (Chen & Hsieh 1971; Matthew Chen 1972; Chen & Wang 1975; JB Hooper 1976; Wang & Cheng 1977; Wang 1977; Wang 1979; Bauer 1979; Johnson 1983; Shen 1990; Chambers 1992:694-5; Milroy 1992:161-2; Krishnamurti 1998; Chambers & Trudgill 1998:159-163; Phillips 2001; Bybee 2002; Phillips 2006; Oliviera 2009; Baxter 2010; Bybee 2010).

Wang's original argument was based on observations that certain sound changes *could not* logically have progressed in the phonetically gradual and lexically abrupt manner which Neogrammarians proposed (Wang 1979:358). There are certain types of sound changes which Wang deemed as defying neogrammarian "laws". These include certain syntagmatic sound changes such as metathesis, where for instance **ks** becomes **sk**, and paradigmatic sound changes involving change in place of articulation such as **m** > **n** or **t** > **k** (Wang 1969:11). He argues:

> It simply makes no phonetic sense to say that a cluster like /ks/ gradually shifted to /sk/ through a series of imperceptible increments. Neither is it plausible to believe in a gradual continuum between members of well-attested changes, such as m > n or t > k, since altogether different articulators are involved.
>
> (Wang 1979:358)

As the quote suggests, Wang proposed that these sound changes cannot only proceed in a neogrammarian, phonetically gradual manner, for in the case of metathesis, there are no conceivable intervening phonetic increments which could lead to the transposition of adjacent phonemes made in a different articulatory position, and in the latter case, it is not clear whether a sound change involving the substitution of one phoneme for another produced in a different articulatory position can only proceed in phonetically imperceptible increments, as a discrete change from one to the other should also be considered a logical possibility, if not a more probable mechanism. A point to bear in mind is that, as pointed out here on p. 9, Osthoff and Brugmann's footnote clearly states that they did not consider metathesis to be a mechanical sound change, and that therefore it was not subject to regular sound laws. Thus, in fact, with respect to metathesis, Wang was in agreement with Osthoff & Brugmann. However, with respect to sound changes involving a change in place of articulation, Wang's prediction runs counter to that arising from Osthoff & Brugmann's neogrammarian manifesto in that Wang believed that it was implausible for such sound changes as **m** > **n** or **t** > **k** to progress in phonetically imperceptibly small steps, with all words in the lexicon affected at the same time.

Cheng & Wang (1977:86) further explain the lexical diffusionist position as follows:

> Although the view of phonological change that entails lexical abruptness, as inherited from the Neogrammarians, clearly is logically possible, we maintain that it has as yet received no systematic empirical confirmation. We feel, on the other hand, that there is a large variety of arguments and evidence which establish the existence of lexically gradual changes beyond any reasonable doubt. Until lexically abrupt changes are indeed verified, therefore, research on diachronic phonology must assume that lexical diffusion is the primary mechanism by means of which changes are implemented. One major topic in this research is the determination of the *chronological profile* of the diffusion process.
>
> (Cheng & Wang 1977:86)

They argued that sound changes must primarily transmit in a lexically *gradual* manner, in other words, they transmit through the lexicon incrementally, rather than that the sound change affects all words at the same time (or in a "lexically abrupt" manner). The way to determine this would be to trace the progression of the sound change through time. They suggested that the lexical split in Chaozhou (Chen & Wang 1975, discussed in detail in Section 1.5.1.1) was an instance of lexical diffusion.

Chambers and Trudgill (1998:159-60) support the notion that investigating the lexical dimension is essential—they argue on the basis of the methodological inadequacies of structuralist and generativist explanations of sound change that lexical variation of sound change across time should not be overlooked. To summarise, their critique of the structuralist position was the belief therein, traceable from Saussure, that the lexicon was irrelevant when studying sound change. Saussure drew upon the analogy that if phonemes are like keys in a musical instrument, then words are the melody—changing the keys of the instrument will result in a change in the melody but the fundamental change is in the phonemes themselves (Saussure 1959; see Phillips 2006:2 for a discussion). Similarly, those who worked in the structuralist tradition conceived of sound change as being equivalent to "phonemes change." More

specifically, one or more allophones associated with a phoneme can change category membership. If all allophones originally associated with a phoneme change membership, then this phoneme is obliterated. Allophones can also, instead of joining another phoneme category, "break off" and form a new phoneme category. Thus, structuralists viewed sound changes as shuffling of allophone membership with respect to phonemes. On the other hand, generativists describe sound changes as the result of a change in sets of ordered rules. If a sound change has happened, it is because a different rule set or different ordering of the same rules applies. However, while these theories can provide "static descriptions" (Chambers & Trudgill 1998) of stages of sound change, they fail to account for the intervening time and the mechanism of how one stage of the sound change progresses to the next. Chambers & Trudgill argue that neither the structuralist view that "phonemes change," nor the generativist rule ordering and rule addition analyses help explain patterns such as those found in lexical variants among speakers in the transition zone between northern and southern English varieties for the variable (u)[6], and therefore such accounts fall short of attaining descriptive or explanatory adequacy in accounting for sound change. In order to sufficiently describe the mechanism of sound changes, and especially the temporal dimension of sound change, Chambers & Trudgill argue, one must also consider whether the lexical diffusion model applies in the intervening stages of a sound change.

1.5.1 Studies on lexical diffusion

Further lexical diffusion analyses of sound changes in Chinese dialects include the Chaozhou tone split (Cheng & Wang 1977), alveolarization of Cantonese finals (Bauer

[6] The data they presented will be examined in detail in Section 1.5.2.

1979), delabialisation of velar onsets **k^{w}-** (Bauer 1983), as well as a study of the velarization of word-final stop consonants in Skikun (Li 1982). A selection of these studies is described in detail in the following sections. In doing so, one of the issues which will be discussed is what constitutes evidence of lexically-diffused change.

1.5.1.1 The Chaozhou tone split

Cheng and Wang (1977) described the split of the reflexes of Middle Chinese (MC) tone III into modern Chaozhou tones 2b and 3b as having progressed by means of lexical diffusion. Based on the distribution of modern Chaozhou tones 2b and tone 3b by MC onsets, Cheng & Wang argued that the even distribution of the split could not have arisen due to analogy, as there are no morphological analogies to draw in Chinese, nor could it have arisen from borrowing from other dialects, as a search through the Dictionary on Computer (DOC, a computerised dictionary of Chinese dialects developed under William S.-Y. Wang) found no other dialects with large numbers of similar forms and "borrowing is not phonetically selective" (Cheng & Wang 1977:91). This left them with two remaining explanations for the split: phonological conditioning, or a system-internal development that is caught in midstream. Based on the fact that tone splits in Chinese dialects have been found to be conditioned by Middle Chinese syllable onset consonant voicing, Cheng and Wang first examined whether the Chaozhou tone split was also conditioned by syllable-onset consonant voicing. Tables 1 and 2 illustrate how MC tone III is distributed across tones 2b and 3b in modern Chaozhou, in the contexts of MC voiced onsets and sonorants, respectively. The numbers in the table indicate occurrences of lexical items which originate from MC tone III by onset. The syllable onset consonants come from Dong's (1954)

reconstructions. As can be seen from the tables, no clear pattern of phonological conditioning could be found to operate on the tone split from MC tone III based on MC voiced (Table 1) or sonorant syllable-onset consonants (Table 2). While only two tables are reproduced here, Cheng & Wang searched for all the possible phonological distinctions that existed in MC which could have conditioned the tone split, including intial consonants, nuclear vowels and syllable-final consonants, and even sets of syllables which were homophonous in MC[7]. However, they did not find any phonological conditioning that could explain the split. The only way to explain this pattern, they suggested, was if the sound change spread through lexical diffusion.

Table 1. Distribution of lexical occurrences of Middle Chinese tone III reflexes in modern Chaozhou lexical tone by MC voiced onsets

		Chaozhou Tones	
		2b	**3b**
Middle Chinese syllable onsets	b	6	7
	v	1	3
	d	11	14
	dz	6	2
	z	3	3
	ɖ	3	4
	dʑ	1	3
	ʑ	3	5
	dj	2	1
	g	6	4
	ɣ	14	15
	Total	**56**	**61**

(table from Cheng & Wang 1977:94)

[7] For the full set of data, consult Cheng & Wang (1977).

Table 2. Distribution of lexical occurrences of Middle Chinese tone III reflexes in modern Chaozhou lexical tone by MC sonorant onsets

		Chaozhou Tones	
		2b	**3b**
Middle Chinese syllable onsets	m	12	10
	ɱ	2	5
	n	7	4
	l	18	9
	ɳ	3	3
	ŋ	11	8
	ɣj	8	4
	Ø	8	3
	Total	**69**	**46**

(table from Cheng & Wang 1977:95)

A subsequent study by Wang and Lien (1993), a more detailed investigation into the Chaozhou tone split, suggested that the mixture of literary and colloquial strata as the explanation for the split. The literary stratum was derived from Northern Luoyang and Chang'an dialects, while the colloquial stratum was derived mainly from Southern Min with a Yue substratum. They argued that modern tone 2b mainly developed from the literary stratum, in which Middle Chinese tone II with voiced obstruent initials was combined with MC tone III with voiced initials. This is shown diagrammatically in Figure 1. On the other hand, modern tone 3b was developed from the colloquial stratum MC tone III with voiced initials. Thus, while the overall tone distribution between tone 2b and tone 3b in modern Chaozhou could not be explained by means of phonetic conditioning, the tone split seemed distributed in a way that reflected the historical development from the literary and colloquial strata, which have each derived from different dialects. Thus, instead of being a case of lexical diffusion, the Chaozhou tone

split instead seemed to be the result of the literary and colloquial strata having been maintained over time but where interstylistic borrowing occurred, which resulted in MC tone III words to have divided more or less evenly between two modern tone categories. This interstylistic borrowing, Wang and Lien argued, are still not in violation of neogrammarian principles, which exempt cases of analogy and borrowing from showing law-like regularity.

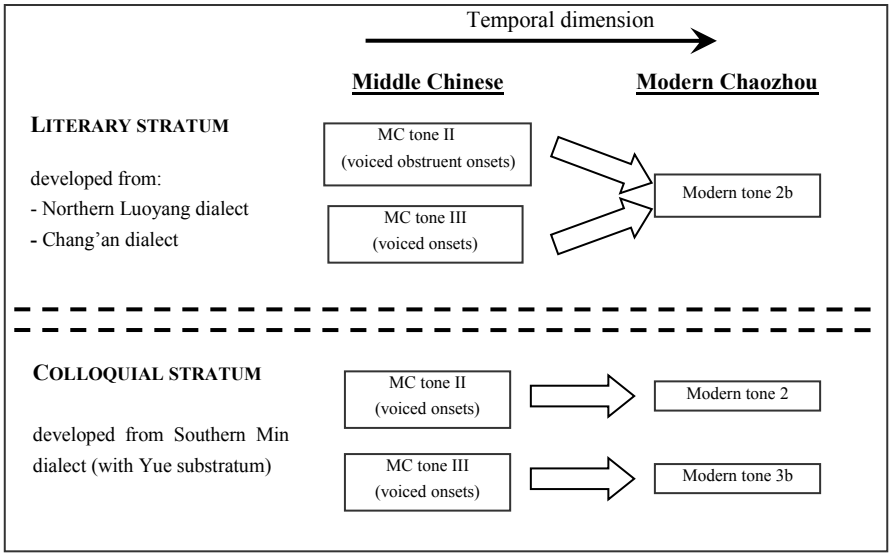


Figure 1. Development of modern Chaozhou tones from Middle Chinese tones from different dialectal and literary/colloquial substrata

1.5.1.2 Homonym split from Middle Chinese to modern Mandarin (Peking dialect)

The split from Middle Chinese to modern Mandarin is often described as displaying lexical diffusion. Chen & Wang (1975) explain why homonym splits are a good candidate for investigating the possibility of lexical diffusion:

> Obviously a short cut to factoring out non-obvious phonological conditions, and thereby establishing potential cases of lexical diffusion, is to produce evidence of divergent paths of historical development taken by words that were originally homophonous twins. […] The multiplicity of homonyms in Chinese provides an ideal ground for testing the hypothesis of lexical diffusion, since homonym splits are by definition spontaneous, or phonologically unconditioned.
>
> (Chen & Wang 1975:260)

The data presented by Chen and Hsieh (1971) shows the split of homonym pairs from Middle Chinese to modern Mandarin (Peking dialect), presented here in Table 3.

Table 3. Mandarin -a/ -ai alternation

Middle Chinese			**present-day Mandarin**		
牌	**b'ai**	→	**p'ai**		'signboard'
罷	**b'ai**	→	**pa**		'to cease'
釵	**tʂ'ai**	→	**tʂ'ai**		'hairpin'
叉	**tʂ'ai**	→	**tʂ'a**		'to cross'
捱	**ŋai**	→	**ai**		'to suffer'
涯	**ŋai**	→	**ia**		'shore'
街	**kai**	→	**kiai**	(→ **tɕie**)	'street'
佳	**kai**	→	**ka**	(→ **tɕia**)	'excellent'

(from Chen & Hsieh 1971:3, with Chinese characters added)

As Table 3 shows, the way in which the sets of rimes in Middle Chinese split into **-ai**/**-a** rimes in modern Mandarin was not phonetically conditioned, as each pair was homophonous to start with. As there was no morphological analogy for speakers to draw upon, Chen & Hsieh argued that the split showed lexical diffusion, as there was no positive evidence which could be found which pointed to dialect borrowing (Chen & Hsieh 1971:7). Altogether, the homonym split from Middle Chinese to modern (Peking)

Mandarin seemed to be an example of phonetically unconditioned, lexically diffused sound change.

1.5.1.3 Lexical diffusion of Skikun velarization

In a study of the shift from labial to velar of word-final stops in Skikun, an Atayalic language of Formosa, Li (1982) investigated the extent to which variation patterns associated with the sound change showed support for lexical diffusion. The data was from 10 speakers of the Skikun dialect, with male and female speakers whose age ranged from 32 to 84 years. Table 4 shows the distribution of variation patterns in the pronunciation of Skikun speakers by decreasing age.

As can be seen from the data, the oldest Skikun speakers were the most conservative, and showed no velarization of the word-final labial stops at all. The youngest speakers on the other hand, showed the most velarization, with the 32-year-old speaker 'W. B.' showing complete velarization. The speakers in between the oldest and the youngest speakers showed a gradual increase in velarization as the age of the speakers decreases. According to Li, this was evidence that the sound change progressed by means of lexical diffusion. Not only did lexical diffusion present itself *within* individual speakers, but lexical gradualness was also found *across* individual speakers. This data provides a prime example of a lexically diffused sound change.

Table 4. Lexical diffusion pattern of word-final labial to velar shift in Skikun

	S.T. f84	*S.P. f80*	*Y.K. f71*	*B.M. m65*	*P.S. m61*	*Y.S. m54*	*M.W. m50*	*Y.N. f55*	*H.Y. m46*	*Y.K. m36*	*W.B. m32s*	
qciyap	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	'opposite shore'
ʔiyup	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	'goshawk'
qatap-	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	'scissors'
tgtap	-p	-p	-p	-p	-p	-p	-p	-k	-p	-p	-k	'to fan'
ghap	-p	-p	-p	-p	-p	-p	-p	-p	-p/k	-p	-k	'seed'
qurip	-p	-p	-p	-p	-p	-p	-p	-p	-p/k	-p	-k	'ginger'
hmap	-p	-p	-p	-p	-p	-p	-p	-p	-k	-k	-k	'stab'
pshup	-p	-p	-p	-p	-p	-p	-p	-p	-k	-p	-k	'suck'
hmop	-p	-p	-p	-p	-p	-p	-p	-p	-k	-p	-k	'do magic'
talap	-p	-p	-p	-p	-p	-p	-p	-k	-p	-k	-k	'eaves'
tgiyup	-p	-p	-p	-p	-p	-p	-p	-k	-k	-p	-k	'sink'
miyup	-p	-p	-p	-p	-p	-p	-p	-k	-k	-k	-k	'enter'
qmalup	-p	-p	-p	-p	-p	-p	-p	-k	-k	-k	-k	'hunt'
mgop	-p	-p	-p	-p	-p	-p	-p	-k	-k	-k	-k	'share one cup'
qmuyup	-p	-p	-p	-p	-p	-p	-p	-k	-k	-p	-k	'fold'
kmiyap	-p	-p	-p	-p	-p	-p	-p	-k	-k	-k	-k	'catch'
mnep	-p	-p	-p	-p	-p	-k	-p	-k	-k	-k	-k	'to fish'
msuyap	-p	-p	-p	-p	-p/k	-k	-p	-k	-k	-k	-k	'yawn'
qom	-m	-m	-m	-m	-m	-m	-m	-m	-m	-ŋ	-ŋ	'anteater'
syam	-m	-m	-m/ŋ	-m	-m	-m	-m	-ŋ	-m	-ŋ	-ŋ	'pork'
qmtam	-m	-m	-m	-m	-m/ŋ	-m	-m/ŋ	-ŋ	-m	-ŋ	-ŋ	'swallow'
rom	-m	-m	-m	-m	-m	-m	-ŋ	-ŋ	-m	-ŋ	-ŋ	'needle'
qinam	-m	-m	-m	-m	-m	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	'peach'
hmham	-m	-m	-ŋ	-m	-m/ŋ	-m	-ŋ	-m	-ŋ	-ŋ	-ŋ	'grope'
yuhum	-m	-m	-ŋ	-m	-m	-m	-ŋ	-ŋ	-m	-ŋ	-ŋ	'gall'
prahum	-m	-m	-m	-m	-m/ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	'lips'
tmalam	-m	-m	-m	-ŋ	-m/ŋ	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	'taste'
mtlom	-m	-m	-m	-m	-m	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	'burn'
lmom	-m	-m	-m	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	'burn'
mktlium	-m	-m	-m	-m	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	'run'
cmom	-m	-m	-m	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	'wipe'
mnkum	-m	-m	-ŋ	-m	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	'dark'

(in Li 1982)

1.5.1.4 Lexical diffusion evidence from other languages

Languages other than Sinitic ones provide further evidence for lexical diffusion. Lexical diffusion was also found for the sound change **s** > **h** > **Ø** in Gondi dialects by Krishnamurti (1998). Not only that, but Krishnamurti found that the sound change progressed in a lexically, and possibly also phonetically, gradual way.[8]

Sommerfelt (1962:75; also discussed in Chen & Hsieh 1971:8; Chen 1972b:470–1; and Phillips 2006:3) suggested that the lenition of the initial voiceless fricative in Welsh also displayed lexical diffusion. Table 5 shows the data which Sommerfelt presented, namely the development of words such as **χwarę** 'to play', **χwanen** 'flea', and **χwa**:**ir** 'sister' across time (**t**$_1$, **t**$_2$, **t**$_3$, **t**$_4$). It was found that the loss of [χ] in such words diffused lexically in a progressive manner, rather than affecting all words with initial [χ] at the same time. This was obtained from observing that at stage **t**$_2$, **χwarę** was the only word which lost the initial [χ] for some time, with **χwanen** and **χwa**:**ir** unaffected. It was only afterwards, at stages **t**$_3$ and **t**$_4$, that **χwanen** and **χwa**:**ir** gradually joined the rank of [χ]-less words. This change also gradually spread geographically from South- to North Wales and across age groups, such that in Central Wales both [χ] and [ʍ] occurred at the time of Somerfelt's investigation, with in addition the innovative [w] spreading towards the north of Wales.

[8] See another example of lexically and phonetically gradual sound change by Chambers & Trudgill (1998) in Section 1.5.2.

Table 5. Dropping of voiceless uvular fricative in some Welsh dialects

	t_1	t_2	t_3	t_4
to play	**χwarę**	**ʁarę**	**ʁarę**	**ʁarę**
Flea	**χwanen**	**χwanen**	**ʁanen**	**ʁanen**
Sister	**χwa:ir**	**χwa:ir**	**χwa:ir**	**ʁa:ir**

(Sommerfelt 1962 in Chen 1972b:471)

In Baxter's (2010) investigation of the *merry-marry* merger in the Eastern Townships of North America (southeast of Montreal and along the east coast, including New Hampshire, Vermont and Maine), evidence was also found for lexical diffusion, despite such discontinuities in the progression of vocalic sound changes having remained elusive in previous studies, and thus formerly assumed to be neogrammarian in nature (Labov 1994:456, 543). What Baxter found was that, while certain lexical items, such as *arrow*, *variation*, *varicose*, and *various*, displayed the *merry-marry* merger, other lexical items, such as *carrot*, *narrow*, and *parish*, remained unmerged, with lexical items in between that only sometimes displaying mergers. This formed a 'merger continuum' with different lexical items scattered across it. This idea is summarised in Table 6 below.

Table 6. The *merry-marry* merger continuum

	(merged)	← merger continuum →		(unmerged)
Stage	**1**	**2**	**3**	**4**
Lexical items	*arrow* *variation* *varicose* *various*	*carry* *parents*	*Larry* *married*	*carrot* *narrow* *parish*

(from Baxter 2010:17)

The fact that different words in the *merry-marry* lexical sets are placed on different positions on the spectrum suggests that the sound change could not be neogrammarian, but displayed lexical diffusion instead.

1.5.1.5 Lexical diffusion of North American English short **a** tensing in Philadelphia

Labov (1994:429–37), summarising the work by Trager (1930; 1934; 1940), Cohen (1970), Ferguson (1975), Payne (1976; 1980), Halle and Mohanan (1985), Kiparsky (1989), Harris (1989), and Labov (1989), examined the tensing of short **a** in Philadelphia (as compared to New York). He investigated short **a** in words such as *ham*, *hand*, *hammer*, *hamming* (verb), *man*, *manner*, *manning* (verb), *can* (noun), *can* (aux.), *am*, *an*, *glad*, *bad*, etc. The speakers in the Middle Atlantic dialects (which include Philadelphia and New York) displayed variation in short **a** pronunciation as short *lax* low nonperipheral [æ], or mid-to-high, long *tense* vowel with a centering inglide [eːə]. Labov termed this the lax vs. tense opposition (Labov 1994:430).

First of all, Labov suggested that short **a** tensing in Philadelphia operated under neogrammarian conditioning: **a** was tensed before certain consonants (**m**, **n**, **f**, **θ**, **s**, **d**) when in a closed syllable. This meant that *ham* and *hand* contained the tense [eːə]

vowel, but *hammer* was pronounced with the lax [æ] vowel as **a** was in an open syllable. In addition, while a tense [eːə] in both *man* and *manning* (in the latter of which we would expect a lax [æ] due to it being in an open syllable) seemed to form an exception to the **a** tensing rule, Labov suggested that this could in fact be explained using the neogrammarian idea of analogy: *manning* as having been inflectionally derived from *man*, which has a tense [eːə].

On the other hand, there were words that seemed to defy the above rule which could not be explained by phonetic conditioning. For instance, the words *mad*, *bad*, *glad*, and *sad* showed patterns that could not be accounted for by neogrammarian type rules. While *mad*, *bad*, and *glad* are always tense, the word *sad* is lax, despite sharing the same **-d** ending, and also being an affective adjective. This suggested that the sound change was also progressing via lexical diffusion.

1.5.2 Lexically and phonetically gradual sound change

There are instances in the literature of lexically gradual *and* phonetically gradual sound changes (see Croft 2000:189 for a brief survey). For instance, Chambers and Trudgill (1998:106–109, 159–65) describe the variation of (**u**) in the transition areas between northern and southern varieties of English as being both lexically gradual and phonetically gradual. The data on which they based this analysis was from the Survey of English Dialects (SED) (Orton & Wakelin).[9] The change originated in the 17th century, in the London region, with the Middle English (ME) short **ŭ** gaining an unrounded variant [ʌ] (Chambers & Trudgill 1998). While speakers from the north always

[9] Note that this survey only collected data from the so-called NORM demographic (non-mobile rural males).

pronounced [ʊ] for (**u**), southern speakers pronounced [ʌ]. In between the north and the south, there was a wide range of variation. Chambers and Trudgill (1998) used the following words extracted from the SED interviews in order to analyse the variation of (**u**). Not every informant produced all of the following words, but each individual produced at least two-thirds of the following list:

> *brother*, *brush* (3 occurrences), *bump*, *bunch*, *butter*, *clump*, *come*, *cousins*, *cud*, *cutter*, *cutting* (2), *does*, *done*, *dove* (2), *dozen*, *drunk*, *duck* (2), *dull*, *dung* (2), *dust*, *dustpan*, *enough*, *funnel*, *gull*, *gun*, *gutter* (2), *hub*, *hundred*, *hungry*, *lump*, *mongrel*, *muck*, *mud*, *mushrooms*, *must*, *other*/*tother*, *pluck* (2), *puddles*, *puppies*/*pups*, *rubbish* (2), *sawdust*, *shovel*, *shut*, *slugs*, *stubble*, *stump*, *sun*, *thumb*, *thunder*, *tuft*, *tup (hog)*, *truss*, *tussock*, *uncle*, *up*, *us*.

The variation in 20 of these words which Chambers & Trudgill discussed, produced by 13 speakers from the different geographical regions described above, is shown in Table 7.

Table 7. Lexical variants among speakers in the transition zone between northern and southern English varieties for the variable (u)

Lexical items	**Speaker**												
	1	2	3	4	5	6	7	8	9	10	11	12	13
must	x	x	x	x	x	-	x	/	x	x	o	o	o
dozen	x	x	x	x	x	/	x	o	o	o	o	/	o
other	x	x	x	x	x	x	x	/	/	o	o	o	o
come	x	x	x	x	x	x	o	o	o	o	x	x	o
stubble	x	x	x	x	x	x	-	x	x	o	o	o	o
does	x	x	x	x	x	o	o	o	o	o	o	o	o
cousin	x	x	x	x	x	o	o	o	o	o	x	o	o
done	x	/	x	x	x	/	/	o	o	o	/	o	o
duck	x	x	/	x	/	/	x	o	o	/	/	o	/
thumb	x	x	x	x	o	o	o	o	o	o	/	x	o
shut	x	x	x	x	x	o	x	o	x	x	o	o	o
sun	x	x	x	x	x	o	o	o	x	x	o	o	o
hungry	x	x	x	x	x	o	o	o	o	o	o	o	o
up	x	x	x	x	x	o	o	o	o	o	o	o	-
hundred	x	x	x	x	o	o	o	o	o	o	o	o	o
pups	x	x	x	x	o	x	o	o	o	o	o	o	o
cud	x	x	x	o	x	x	o	o	o	o	o	o	o
butter	o	x	x	x	o	x	x	o	o	o	o	/	o
gull	x	x	o	x	o	o	x	o	o	o	-	o	-
uncle	x	x	x	x	o	o	o	o	x	x	o	o	o

x : [ʌ] / : variation between [ʊ] and [ʌ] within the same speaker

o : either [ʊ] or [ɤ] - : no data

(Chambers & Trudgill 1998:161, Table 10-3)

The data shows that speakers differed with respect to which vowel, [ʊ] or [ʌ], they used for (**u**) for a given word, while some speakers showed variation between the two pronunciations (indicated by '/') even for the same word. A phonetically intermediate value [ɤ] was also used for some words in certain intervening geographical regions between the north, where the [ʊ] variant was almost exclusively used, and the southern regions, where [ʊ] and [ʌ] were used. The coding of [ʊ] vs. [ɤ] vs. [ʌ] seemed to have been for practical reasons to make quantification easier, because a variety of phonetic

realisations such as raised and lowered variants [ʊ̝] and [ʊ̞], [ɤ̝] and [ɤ̞], as well as [ʌ̝] and [ʌ̞] were also observed (Chambers & Trudgill 1998). Chambers and Trudgill have termed varieties which are affected by a sound change in a lexically gradual *and* phonetically gradual manner 'fudged lects' (p.110), because there was an intermediate phonetic value [ɤ] used in the intervening areas (which affected some but not all words), as opposed to the phonetically more distant variant [ʌ]. In contrast, they considered varieties which are affected by sound change in a lexically gradual but phonetically abrupt way to be 'mixed lects' (p.110), where two distinct [ʌ] vs. [ʊ] variants were observed. See Chambers and Trudgill (1998:106–13) for the full description of the geographical spread of these variants. From this we can see that, while lexical diffusion models typically propose that sound changes progress in a lexically gradual but phonetically abrupt manner, there is evidence that lexically gradual *and* phonetically gradual sound changes can also occur. Chambers and Trudgill argue:

> Faced with data like this and a hypothesis that 'phonemes change', what conclusion can a dialectologist draw? Determining any phonemic analysis on the basis of data like this is a problem, to say the least. The hypothesis that 'new rules are added' is no more helpful. The variety in [Table 7] does not seem to be rule-governed, at least not in any sense that that term is used currently.
>
> (Chambers & Trudgill 1998:161)

They proposed that not only do we need to abandon the idea of strict categorical rules, but also that we should investigate the lexical dimension of a sound change, and how it develops through time.

1.5.3 Summary of evidence for lexical diffusion

From the above evidence, on the homonym split from Middle Chinese to modern Mandarin, velarization in Skikun, lenition of initial voiceless fricative in Welsh, the *merry-marry* merger in the Eastern Townships of North America, and even short **a** tensing in North American English—which Labov initially quoted as an example of neogrammarian change—it is clear that lexical diffusion can be found. These examples support Wang's suggestion that at least *certain* sound changes progress in a lexically gradual manner. However, as will be revealed in the following section, there have been other attempts at interpreting the data, to support the view of those who maintain that sound changes progress primarily through neogrammarian means (i.e. phonetically gradual and lexically abrupt).

1.6 The neogrammarian vs lexical diffusionist "controversy" over regularity of sound change

From the above discussion of the neogrammarian model and the lexical diffusion model of sound change, it has become apparent that the two views have been set in opposition to one another on the question of regularity of sound change. This has often been framed as a "paradox", "debate", or "controversy". While the "debate" over regularity of sound change has been in existence for a long time, studies up till the 90s have mainly sought to provide evidence to support either of the two stances. However, the question was brought to the fore in the literature again by Labov (1981; 1991; 1994). In particular, he raised questions regarding whether the seemingly lexically diffused variation patterns presented above are not in fact better explained by social and linguistic conditioning patterns which were overlooked. For instance, Labov (1991)

offers as an example of neogrammarian sound change the case of **æ**-raising in words such as *can*, and *black* in the largest cities involved in the Northern Cities Shift (occurring in major cities as Chicago, Detroit, Rochester, Buffalo, Syracuse and Cleveland). He claimed that **æ**-raising can be explained purely through phonetic conditioning by neighbouring segments. In terms of manner of articulation **æ**-raising was most advanced before nasal consonants, next-most advanced before voiceless fricatives, and least advanced when followed by voiced stops. In terms of place of articulation of the following segment, **æ**-raising was most advanced when followed by palatals and apicals, and least advanced before velars and labials. On the other hand, if **æ** was preceded by an initial liquid cluster, such as in the word *black*, it was found that **æ**-raising was least favoured (Labov 1991). However, Phillips (2006:16–7) countered Labov's analysis with:

> Unfortunately, Labov does not specify how many or which words were used as evidence of the fine-grained phonetic conditioning of the raising of /æ/. The problem with this is that **lexical diffusion is always found inside of phonetic conditioning**. So the fact that fine-grained phonetic conditioning is found does not exclude the possibility of lexical diffusion. [My emphasis]

In essence, one is unlikely to find lexical diffusion if one does not take the lexical dimension into account.

Similarly, Labov (1994) also offered a reassessment of the Skikun data in Li (1982), presented in Section 1.5.1.3 above, and raised the question of "whether it is their lexical identity, or some other factor, that is responsible" for the likelihood with which the substitution of word-final labial stops by velar stop took place (Labov 1994:445). While it was clear that word-final nasals shifted to velars more readily than word-final oral stops (Labov 1994:446), he examined the degree to which other factors,

including social factors and the place of articulation of word-initial stops, played a role in the Skikun shift to word-final velar stops. He used the data from the variable speakers (8 out of 10—the remaining speakers did not show any variation at all) as well as the following factors in a multivariate variable rule analysis: (i) sex, (ii) age, (iii) whether the final consonant is oral or nasal, and (iv) place of articulation of word-initial stops. These factors are presented in the left-most column in Table 8, which shows the results of the variable rule analysis by sex, age, final consonant, and initial consonant clusters, with velar final as the dependent variable.

Table 8. Variable rule analysis of Skikun labial-velar alternations

		%	Varbrul
SEX			
Male		41	.90
Female		38	.32
AGE			
70-		12	.01
60-69		15	.21
50-59		42	.55
30-49		74	.97
FINAL CONSONANT			
Oral		24	.29
Nasal		56	.76
INITIAL CONSONANTS			
Uvular, glottal	q-, qc-, ʔ-	5	.04
Labial+laryngeal	psh-	12	.22
Velar+apical	tgt-, tg-	12	.24
Velar+laryngeal	gh-	19	.35
No stop	r-, sy-, y-	42	.35
Apical	t-	25	.48
Uvular+nasal	qm-, qmt-, qin-	41	.57
Laryngeal+nasal	hm-, hmh-	38	.59
Apical/velar+nasal	tm-, km-, mg-, mtl-, mktl-,	56	.77
Labial, oral	pr-[#]	75	.87
Labial, nasal	m-, mn-[#], ms-[#], cm-, lm-[#]	60	.89

* Chi-squared per cell: 1.08 (indicating a good fit of prediction to observation)

[#] Note that these also contain an apical in addition to a labial.

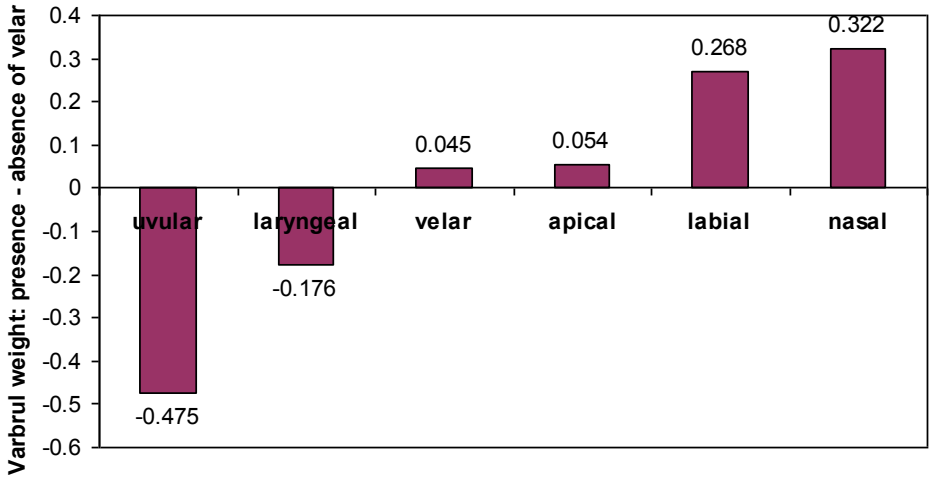
(from Labov 1994:447)

The next-to-last column shows the percentage with which a velar stop was used for word-final bilabial stops, with a higher percentage showing a higher degree of use of the innovative variant. The right-most column shows the Varbrul weight returned by the analysis, which is a measure of:

> the effects of multiple factors on a binary linguistic 'choice'—the presence or absence of an element, or any phenomenon treated as an alternation between two variants […] [t]he program identifies which factors significantly affect the response variable of interest, in what direction and to what degree.
>
> (Johnson 2009:359)

A Varbrul weight of 0.50 indicates that the factor has no effect on the outcome, with values of less than 0.50 indicating that the factor has a disfavouring effect and values of over 0.50 indicating that the factor has a favouring effect. Most notably, the "apical/velar + nasal", "labial + oral", and the "labial + nasal" onset consonants had Varbrul weights of 0.77, 0.87 and 0.89 respectively, suggesting that these initial consonants had a strong favouring effect on word-final stop consonant velarization, whereas other initial consonants had a largely disfavouring effect or no effect at all. Based on this analysis, Labov argued that there was a strong tendency for phonetic conditioning of word-final stop velarization by place of articulation of the word-initial consonant.

Labov observed that the variation between final bilabial vs. velar stop was modulated by the word-initial consonant: there is a "systematic form of dissimilation, where the place of articulation of the initials is differentiated as much as possible from the place of articulation of the final" (Labov 1994:446). To show the effects graphically, he separated the place features of the word-initial stops (e.g. for the Skikun word *prahum* the feature was 'labial', for the word *mktlium* the features were 'labial,' 'nasal,' 'velar,' and 'apical'). He then plotted the Varbrul weight differences (shown in Figure 2) for the presence vs. absence of these word-initial place features on word-final stop velarization, such that if the difference is zero, there was no overall effect of that feature. If the difference was below zero, the effect would be disfavouring.



(Labov 1994:449)

Figure 2. Effects of initial consonant features on velar realization of coda consonants in Skikun

On the basis of this tendency for labials and nasals to have a favouring effect and uvular and laryngeal features to have a disfavouring effect, Labov concluded that "a change that was initially presented as a clear case of lexical diffusion is instead a likely example of neogrammarian regularity" (Labov 1994:450). He therefore argued that the neogrammarian idea of complete regularity had to be revised to being of stochastic tendency rather than exceptionless rules:

> This re-examination of the Atayalic [Skikun] data indicates that a change that was initially presented as a clear case of lexical diffusion is instead a likely example of neogrammarian regularity. The neogrammarian viewpoint must of course be modified to accept stochastic regularities in place of absolute rules.
>
> (Labov 1994:450)

This revision of the neogrammarian argument seems puzzling. The fact that phonological rules can only predict where the innovative variants occur *stochastically*, implies that lexical diffusion *can* be the underlying mechanism. Further criticism of

Labov's neogrammarian position is offered by Phillips (2006:11–2), who has pointed out the lack of evidence and internal consistency for Labov's (1994:534) prediction that there is a dichotomous distinction between sound changes which are supposedly regular as opposed to those which are lexically diffused. Furthermore, Labov (1994:539) suggested that sound changes involving a shift in place of articulation are "necessarily more abstract than changes in manner," and unlike changes that only involve manner of articulation, lexical diffusion is more likely (540-41), although it is not entirely clear what he means by "necessarily more abstract". He suggested that vowel shifts in place of articulation, diphthongization of high vowels, consonant changes in manner of articulation, vocalization of liquids and deletion of glides and schwa are sound changes which are regular, while shortening and lengthening of segments, dipthongization of mid and low vowels, consonant changes in place of articulation, metathesis of liquids and stops and deletion of obstruents observe lexical diffusion. The full set of his predictions are summarised in Table 9.

Table 9. Labov's prediction of the types of sound changes which are regular and those which are lexically diffused

Regular sound change	**Lexical diffusion**
Vowel shifts in place of articulation	Shortening and lengthening of segments
Diphthongization of high vowels	Diphthongization of mid and low vowels
Consonant changes in manner of articulation	Consonant changes in place of articulation
Vocalization of liquids	Metathesis of liquids and stops
Deletion of glides and schwa	Deletion of obstruents

Phillips quoted examples in the literature which are precisely the types of sound changes which Labov claimed were regular, and yet for which evidence exists that they were lexically diffused. For instance, Phillips (1980) found that the raising of Old English [ɑ] to [ɔ] before nasals affected the most frequent words first, and thus was a case of lexical diffusion, providing a counter-example to Labov's prediction that vowel shifts involving place of articulation are necessarily regular. Section 1.1.1 in Phillips (2006) offers a more detailed discussion and counter-examples to Labov's proposed dichotomy, which will not be repeated here. Nonetheless, Phillips points out that this does not mean that we should therefore ignore the possibility of phonetic conditioning:

> determining the phonetic environments which impact the change does not solve the actuation problem of why this change occurred when it did nor does it fully explain the implementation problem - what drives this change to be diffused through similar phonetic environments, through the population of speakers, and through the lexicon. But it should shed some light on the paths through which a sound change is implemented in the lexicon.
>
> (Phillips 2006:58)

It is clear from the above discussions that researchers preoccupied with one approach tended to overlook the possibility that a given sound change can follow *both* (or even multiple) models, whether equally or to different extents. Krishnamurti (1998) has made a similar point, questioning the incompatibility of the lexical diffusion model with the neogrammarian regularity assumption that sound changes necessarily progress in small phonetically incremental steps. The argument put forward was that it is not inconceivable that a sound change progresses in a phonetically incremental manner and yet also progresses in a lexically gradual way, which is borne out not only by

Krishnamurti's own findings, but also by the data presented by Chambers & Trudgill (1998) above.

In fact, Phillips (2006:7) holds an even stronger position than Wang, and sides with Ogura,

> In lexical diffusion, diffusion from word to word in the language progresses from speaker to speaker in the community. …The three processes of diffusion from word to word, diffusion from speaker to speaker and diffusion from site to site progress side by side.
>
> (Ogura 1995:32)

and Oliviera,

> My position is more radical than Chen and Wang's, and I will say that *all* sound changes are lexically implemented, that is, there are *no* neogrammarian sound changes (although we can have neogrammarian long-term end results).
>
> (Oliviera 2009:103)

in suggesting that *all* sound changes progress in a lexically diffused manner. Phillips deems it as plausible that the temporal dimension was not sufficiently taken into account in most cases and that any instances of neogrammarian sound change were likely "very rapidly (lexically) diffused changes" (2006:6). She quotes studies which found lexical diffusion and classified these sound changes into at least six types (Phillips 2006:10). I have categorised these six types according to (1) whether frequent or infrequent words lead the sound change, (2) whether the sound change is phonetically gradual or abrupt, and (3) whether the end result of the sound change is neogrammarian, in other words, regular. These six types are summarised in Table 10. In terms of *which* words lead the sound change, it seems that instances can be found where frequent words lead the sound change, but there are also cases of infrequent words

leading the change. Bybee (1999; 2002b) found in her survey of modern English final **t**/**d**-deletion that it was frequent words which led the sound change. On the other hand, in Middle English /ö(:)/ unrounding, Phillips (1984) found that it was infrequent words which led the change. Similarly, there were changes which were phonetically gradual, others which were phonetically abrupt. The former type can be found in the French change from [ɑ̃] to [ɔ̃] (Hansen 2001), chain shift in Montreal French (Yaeger-Dror 1994), and glide-deletion in South American English (Phillips 1994). An illustration of a phonetically abrupt sound change would be stress shift in diatonic noun-verb pairs, such as *cónvict* (noun) and *convíct* (verb) in English (Phillips 1984). Last but not least, there are lexically diffused sound changes which result in neogrammarian regularity, and there are those which do not result in neogrammarian regularity. The shift from **s** > **h** > **Ø** in Gondi dialects (Krishnamurti 1998) is an example of the former, while Philadelphia **æ** tensing in North American English (Labov, Yaeger & Steiner 1972:47–95; Labov 1994:436–7) seems to be of the latter type. The following table summarizes these different types of sound changes and the studies which found instantiations of each type.

Table 10. A typology of sound changes displaying lexical diffusion*

1. Initiation	2. Phonetic Process	3. End Result
***Frequent* words lead sound change** e.g. Modern English final **t**/**d-**deletion, Bybee (2002b; 1999)	**Phonetically *gradual*** e.g - French [ɑ̃] > [ɔ̃], Hansen (Hansen 2001) - Chain shift in Montreal French, Yaeger-Dror (1994) - Glide deletion in Southern American English, Phillips (1994)	**Neogrammarian regularity results** e.g. **s** > **h** > **Ø** in Gondi dialects, Krishnamurti (1998)
***Infrequent* words lead sound change** e.g. ME **ö(:)** unrounding Phillips (1984) - Diatonic pairs in English	**Phonetically *abrupt*** e.g. Stress shift in diatonic noun-verb pairs in English, Phillips (1984)	**No neogrammarian regularity results** e.g. Philadelphia **æ** tensing, Labov et al. (1972:47–95) ; Labov (1994:436–7)

* Adapted from Phillips (2006:10)

Based on this (non-exhaustive) typology of how sound changes can progress, the view that lexical diffusion is the norm rather than the exception seems to be supported. In the words of Bybee (2002a:270):

> The evidence […] suggests that lexical diffusion is much more common than previously supposed. It may be that all sound change diffuses gradually through the lexicon. For this reason we must reject the dichotomy of regular versus lexical diffusion changes and look at the pattern of lexical diffusion for each change.

This brief survey of the types of changes in which phonetic regularity and/or lexical diffusion has been found indicates that the "controversy" over neogrammarian versus lexically diffused sound changes is most fruitfully addressed through a study of

both the influence of phonetic environment as well as the role of the lexicon in a given sound change.

1.7 Lexical frequency effects

As discussed above, proponents of the lexical diffusion hypothesis believe that sound changes transmit through the lexicon gradually. Comparing the lexicon at two consecutive points in time and showing that the sound change affected only a *subset* of words (with the same phonological environment where the sound change could apply) would constitute sufficient evidence that lexical diffusion had taken place. Some researchers have posed the additional question of *which* words would be more likely to undergo lexical diffusion: while lexical diffusion might provide a model for describing how certain sound changes affect lexical items one by one, in order to fully account for how sound changes are transmitted, a lexical diffusionist needs to answer which words are affected by a given sound change *first* (Oliviera 2009:94).

Given the fact that lexical items occur with different frequencies in usage, one of the logical consequences of adopting a lexical diffusion analysis of sound change involves examining whether a given sound change is conditioned by the frequency of occurrence of a lexical items. This possibility, in fact, was postulated by Schuchardt (1885:26):

> I will not deny that a word that has been spoken 10,000 times can favor the development of the sound x to x' in a word spoken only 8,000 times, etc. The greater or lesser frequency in the use of individual words that plays such a prominent role in analogical formation is also of great importance for their phonetic transformation. **Rarely used words drag behind; very frequently used ones hurry ahead**. Exceptions to the sound laws are formed in both groups.
>
> [My emphasis]
>
> (Schuchardt 1885:26, translation in Venneman & Wilbur 1972:58; see also Fought 1982:442).

Thus, Schuchardt asserts that frequency does not only play a role in analogy, but also in phonetic sound changes, emphasising that "[r]arely-used words drag behind; very frequently used ones hurry ahead" (1885:26). Schuchardt envisaged an experiment:

> Have somebody who does not know why you ask him to do it repeat a single word thirty, fifty, or eighty times in a row and you will find very marked variations in pronunciation. […] There is always talk about the principle of least effort whenever the causes of sound change are being debated. What is more natural than making things easier whenever over-frequency provides the strongest impulse for this and wherever the danger of misunderstanding is least?"
>
> (Schuchardt 1885:28).

In fact, Johnson (1983) has performed this very experiment, to test whether boosting frequency temporarily using repetitions has the effect of introduction of variation in the form of articulatory reduction. He asked his subjects to repeat two sentences—one containing commonly used, frequent words, the other containing uncommon, infrequent words—up to 50 times. He explained (1983:63):

> If the frequency rule were valid, the Common sentence should: first, from the very beginning show itself readily subject to free variation rules, and secondly, remain fairly consistent throughout the test. On the other hand, the Uncommon sentence should begin with each word precisely articulated and then after many repetitions begin to show the application of free variation rules.

The common sentence which Johnson asked his subjects (namely three family members) to repeat was “Jack will butter up the American finger food”. The uncommon sentence was “The maniac palled while converting tapioca to kelp”. Johnson then used narrow phonetic transcription to describe the 1st, 5th, 35th or 45th, and 50th instances of the sentences (1983:64–9). The results are summarised in Tables 11 and 12.

Table 11. Repetition of a common sentence by three speakers

“Jack will butter up the American finger food” (Johnson 1983:64–9)

	Speaker 1	**Speaker 2**	**Speaker 3**
1st	[dʒækʷ ɨl bʌɾ̩ əpˑ ðə əmerɪkʰn̩ fiŋgɾ̩ fud]	[dʒæk̚ wɪl bʌɾ̩ əp ði əmerɪgn̩ fiŋgɾ̩ fu̯d]	[dʒæk̚ wəl bʌɾ̩ əpˑ ði əmɛrəkˑn̩ fiŋgɾ̩ fud̚]
5th	[dʒæk wəl bərdɾ̩ əp ðəːmerɨkn̩ fiŋgɾ̩ fu̯d]	[dʒæk̚ wɪl bʌɾ̩ əp̚ ði əmerɨkn̩ fiŋgɾ̩ fud̚]	[dʒækʰ wɨl bʌɾ̩ əpˑ ðiyəmerɪkʰn̩ fiŋgɾ̩ fudˑ]
35th	[dʒæk̚ wəl bʌrdɾ̩ əp̚ðəːmerɨkn̩ fiŋgɾ̩ fu̯d]	[dʒæk̚ wɨl bʌɾ̩ əp̚ ðɪ əmɛrign̩ fiŋgɾ̩ fud]	[dʒækˑ wɨl bʌɾ̩ əp ði əmɛrɪkn̩ fiŋgɾ̩ fud̚]

Table 12. Repetition of an uncommon sentence by three speakers

“The Maniac palled while converting tapioca to kelp” (Johnson 1983:64–9)

	Speaker 1	**Speaker 2**	**Speaker 3**
1st	[ðə meniæk̚ pʰɑlɾ wɑyl kʰənvərrɪŋ tʰæpʰiokə tu kʰɛlp̚]	[ðə meniækʰ pʰɑld wɑyl kʰənvərdɪŋ ðə tʰæp̄iokʰɑ tu pʰɛlpʰ]	[ðə meniæk̚ pʰɑld̄ wɑyl kˑənvərrɪŋ ðə tʰæpʰiokʰɑ tu kʰɛlpʰ]
5th	[ðə meniæʔ̚ pʰɑlḛ ʰwɑyl kənvərrɪŋ tʰæpiokʰə tə kˣɛlpʰ]	[ðə meniæk̚ pʰɑlḛ wɑyl kʰənvərrɪŋ ð tæpiokʰə tʰu kɛlpʰ]	[ðə meniæk̚ pʰɑlḛ wɑyl kənvərrɪŋ ðə tʰæpˑiokˑə tə kˑɛlpʰ]
45th	[ðə meniækʔ pʰɑlɾ wɑyl kʰənvərrɪŋ tʰæpʰiokə tə kʰɛlp̚]	[ðə meniæk̚ pʰɑl wɑyl kʰənvərrɪŋ tʰæpiokə tə kʰɛlp̚]	[ðə meniæk̚ pʰɑld̚ wɑyl tʰərniŋ tʰæpʸoḳə ʔə kɛlpʰ]
50th	[ðə menyæk̚ pʰɑlt ʰwɑyl kənvərrɪŋ tʰæpiokə tə kʰɛlpɸ]	[ðə meniæk̚ pɑld wɑyl k̄ənvərrɪŋ tʰæpiok̄ə ɾə kɛl]	[ðə meniæ pˑɑld̚ wɑyl tʰərniŋ ðə tʰæpʸioɣə tə kʰɛlpɸ]

According to Johnson, the difference in degree of articulatory reduction between common and uncommon words in the first repetition was obvious. While common words showed more flapping and less aspiration, the uncommon words had stops which were in general more clearly aspirated. With further repetitions, while there was a slight increase in reduced variants in the common words, this increase in reduced forms tailed off very quickly (which is why he stopped transcribing beyond the 35th repetition). However, the features which were seen in the common words from the beginning also began to show in the uncommon words in later repetitions: flapping, deletion and non-released stops. Johnson's conclusion was that the relative consistency shown by the common words throughout the repetitions and the "dramatic change" in the uncommon words from the first repetition through multiple repetitions showed that certain free variation rules are "frequency bound" (1983:69). This demonstration that repetition has at least the short-term effect of phonetic reduction can *perhaps* be projected to the general relationship between lexical frequency and phonetic reduction over a longer period of time.

Lexical frequency has also been suggested as a factor in accounting for different rates of sound change between words in Cantonese. Bauer (1983) hypothesised that innovative use of the syllable-initial velar oral stop [k-] in Hong Kong Cantonese for labialised velar **kʷ-**, displayed sensitivity to word frequency and that the more frequent lexical items would lead the change. However, he also found some exceptions. While the word 果 **kʷɔ**:˩ 'fruit/result' was less frequent than 廣 **kʷɔ**:**ŋ**˩ 'broad/wide/Guang(dong)/Guang(zhou)', it displayed a far higher proportion of delabialization (Bauer 1983:155), a fact which he could not explain. Judging from the meaning of these words, however, it is likely that stylistic matters might play a role, and

that 果 ‘fruit’ might be used more often in casual speech settings, whereas 廣 ‘extensive’ might be used more frequently in formal speech settings. Whilst whether it was the higher frequency or the lower frequency words which led the change is still up for empirical validation in the case of **k^{w}-** delabialization in Cantonese, the hypothesis that frequency can explain which words lead a given sound change was at least also put forward by Bauer.

Further studies have repeatedly found lexical frequency effects in sound change and in speech production. Hooper (1976) found that sound change affected high frequency words first, and that it was analogical changes which affected low frequency words first. Jurafsky et al. (1998) showed that words which occurred in frequent collocations were more reduced. Bybee (2000a:70) found, in word-final **t**/**d**-deletion in American English, that high-frequency words were more likely to undergo deletion than low-frequency words. Examples where lexical frequency effects on speech production have been found seem to abound within the literature (Fidelholtz 1975; Johnson 1983; Phillips 1984; Phillips 2001; Bybee 2002a; Phillips 2006; Bybee 2010; Hooper 1976); yet how might lexical frequency effects be related to sound change?

The mechanism which Bybee (2002a:287) proposed for the influence of word frequency on sound change was as follows:

> Changes that affect high-frequency words first are a **result of the automation of production**, the normal overlap and reduction of articulatory gestures that comes with fluency (Browman & Goldstein 1992; Mowrey & Pagliuca 1995). The strong directionality of such changes indicates that they are not the result of random variation, but that they stem from reduction processes resulting from repetition and the normal automation of motor activity.

This idea echoes the findings by Johnson above, which found articulatory reduction when repetitions were performed within a short span of time. But Bybee's argument that such articulatory reduction associated with high frequency words can explain reductive sound changes as a long-term effect has also found support within the literature. As mentioned briefly above, Bybee (2000b) studied **t**/**d**-deletion to test whether there were any lexical frequency effects. Using a corpus of Chicano English in Los Angeles (Santa Ana 1991), Bybee studied the influence of lexical frequency on **t**/**d**-delection, and found that **t**/**d**-delection occurred more often in high frequency words than in low frequency ones. The frequency information was obtained from Francis & Kučera (1982) and Bybee used a cut-off point between high and low frequency of 35 per million occurrences, as this was found to be the median frequency for regular past tense verbs, which was the target of her investigation. That is to say, words occurring 35 or more times per million were considered high frequency, whereas words occurring less than 35 times per million were considered low frequency. Bybee's findings are summarised in Table 13. Out of a total of 2049 tokens, higher frequency words (54.5% out of 1650 tokens) were significantly more likely than lower frequency words (34.4% out of 399 words) to undergo deletion (χ^2 = 41.67, p < .001). This provides support that lexical frequency effects operate on **t**/**d**-deletion in English.

Table 13. Rate of t/d-delection by word frequency

	Number of words with High Frequency		Number of Words with Low Frequency
Retention	752 (45.6%)		262 (65.7%)
Deletion	898 (54.4%)		137 (34.3%)

*Bybee (2000b), $\chi^2 = 41.67$, $p < .001$

Phillips (2006; 1980), on the other hand, presented evidence that some sound changes affected low frequency words first. Phillips found that the unrounding of Middle English **ö(**:**)** to **e(**:**)** diffused lexically, and that the change started with the least frequent words. Phillips (2001) added that "changes that ignore the phonological integrity of the segments and the morphological composition of words affect the most frequent words first" (also quoted in De Schryver et al. 2008:164). Based on such findings, Bybee (2002a:269–71) proposed the Frequency Activation Hypothesis, which states that articulatorily reductive sound changes affect the most frequent words first, while analogical changes affect infrequent words first (Bybee 2002a:269–71; also in De Schryver et al. 2008:164). Examining lexical frequency effects then, allows us to examine whether quantitative effects operate in a sound-change-in-progress situation and *where* a lexically-diffused sound change might be activated first.

1.8 Research questions

One might ask, how is the above discussion of lexical frequency effects on reductive sound changes related to our focus of investigation, namely alveolarization in Hong Kong Cantonese? It is clear from the above discussion that there is a consensus that in casual, connected speech, consonants often undergo reduction due to influence from neighbouring phonetic environments and through what has been described as the principle of least effort. Consequently, we expect the final consonants **-ŋ** and **-k** to be produced with a significant amount of reduction in everyday connected speech—for it would be surprising if the coda consonants were fully produced all the time as if in very

careful speech. It is then possible that such reductions neutralize the **-ŋ**/**-n** and **-k**/**-t** contrasts in some contexts, and that a suspension of contrast can lead to reanalysis by the next generation. Models which aim to incorporate usage by speakers, such as exemplar theory, draw heavily upon the notion that frequency of both exposure in auditory perception and of articulatory production can result in such shifts in phonetic parameters and shuffling of phonetic categories (see e.g. Pierrehumbert 2001; Bybee 2002b; Hay, Drager & Warren 2010). Pierrehumbert (2001:140) proposed that:

> [i]f every encountered token of a category is stored as a separate exemplar, then frequent categories will obviously be represented by numerous tokens and infrequent categories will be represented by less numerous tokens. The difference in token count is one ingredient of the model's explanations of frequency effects.

In this model, it is through the gradual and repeated exposure of phonetic targets that patterns are learnt by a given speaker (Pierrehumbert 2001:137):

> Exact phonetic targets and patterns of variation must accordingly be learned during the course of language acquisition. The usage-based framework readily accommodates such findings by proposing that mental representations of phonological targets and patterns are gradually built up through experience with speech.

Pierrehumbert (2001:147) further argues on the basis of systematic production biases—namely hypo- or hyperarticulation—found by Lindblom (1990) that:

> hypo-articulation, or the tendency to undershoot articulatory targets in order to save effort and speed up communication […] is arguably the cause of leniting historical changes, such as schwa reduction and /t/-deletion.

In summary, findings within the literature support the notion that lexical frequency effects operate on leniting sound changes. Furthermore, a proposal has been put forward by proponents of the exemplar theory of speech perception and production that phonetic categories can and do shift through frequent exposure. Based on these ideas, this thesis will examine the extent of lenition and hypo-articulation in the production of velar stops and the possible role it might play in sound change. Given the fact that articulatory reduction in casual speech is a natural phenomenon and that lexical frequency effects have been associated with this reduction, to what extent is lenition actually observed and how might it be related to lexical frequency, both in casual speech as well as in more formal registers? Could lenition be the mechanism responsible for an erosion of velar coda stops, creating a precursory condition for alveolar realizations to take over? To what extent can a phonetically gradual progression be observed from velar to alveolar realization of the coda stops; or is alveolarization a case of phonetically abrupt substitution of phonetic realization in one place of articulation by that of another? These are questions which will be explored in this study of alveolarization in Hong Kong Cantonese.

1.9 Quantitative sociophonetic variation

The greatest difficulty with investigating the mechanisms of completed sound change has been the lack of evidence of the sound system in a prior state and the social conditions in which the language was spoken. There is an element of circularity when one creates a picture of linguistic structure in a prior state of the language using reconstructive methods which are based on neogrammarian regularity assumptions:

when one tries to uncover the mechanisms of sound change, what one finds can then *only* be neogrammarian sound change. Despite their proclaimed purpose to study the mechanisms of articulation and natural speech production, the Neogrammarians did not take this to also mean examining the extent of variation found in actual speech production. A systematic study of variation in everyday speech production appeared much later.

The seminal work by Weinreich, Labov and Herzog (1968), which explicitly recognized the importance of variation in the speech community, was set against the generative assumption that an idealised speaker-hearer's knowledge of language should be the main object of study. Believing that language is a communicative tool used between speakers in social settings, the authors suggested that linguists should instead concern themselves with real-life speech produced by speakers in communicative situations, and that a model of sound change should take into account the variability at a synchronic point in time. They therefore suggested that the aim of linguistic study should be to uncover the structured heterogeneity in sociolinguistic variation. Using recorded natural speech to study sound change in progress allows researchers to investigate the extent of variation occurring in natural speech to better understand the mechanisms of sound change. This thesis adopts this methodology, and will investigate alveolarization from the speech of a stratified sample of speakers in different stylistic settings to explore the structured heterogeneity that the sound change might display. Furthermore, I take issue with viewing language—and subsystems such as phonology—as a largely static system of rules, which is overwhelmingly how formal approaches to phonology seek to describe language. While viewing linguistic systems synchronically

as a system of largely immutable rules is problematic, this is even more so when one examines language diachronically, a necessity when describing sound change.

The study on the variation of (u) in British English (i.e. the STRUT lexical set in Wells 1982) described by Chambers and Trudgill (1998:159-160) and the study of Skikun velarization described in Sections 1.5.2 and 1.5.1.3 above are some of the few examples of studies looking into lexical diffusion which also take sociolinguistic factors into account. Many of the classic accounts of lexical diffusion (e.g. Wang 1969; Cheng & Wang 1977; Wang 1977) fail to take into account speaker-based variation. Yet, if we are to understand the mechanism by which sound changes progress, we need to understand, as far as possible, the variability within and amongst social carriers of the sound change. Thus, an exploration of the social dimension forms the basis of this investigation. In addition, it is the *usage* of language by speakers which requires our scrutiny. As described in earlier sections, the lexical dimension is one of the basic analytical layers which ties in with usage, and taking into account usage-based predictions of the effect of frequent exposure on speech production, and subsequently on a sound system, is crucial to an understanding of how sound change transmits.

1.10 Aims of this thesis

The descriptive aim of this thesis is to investigate the variation in production of the velar coda stops and thereby the progress of alveolarization in Hong Kong Cantonese: is it an ongoing sound change, nearing completion, or is it practically complete? By examining the extent of variation, and the rapidity of progression of this sound change, it is hoped that the factors which drive this sound change can be uncovered.

Going from the descriptive to the theoretical, chief emphasis will be placed on evaluating the neogrammarian hypothesis and the lexical diffusion model, two models of sound change often seen in competition with one another which together, as we have seen, formed a subject of debate in recent decades (e.g. Wang 1969; Labov 1994; Oliviera 2009). Lexical diffusionists have contended that assuming regularity obscures the mechanisms which underlie sound change, which may—and, some argue, can only—proceed in a lexically gradual manner. From a theoretical and historical-linguistic viewpoint, the development of the coda consonants in Cantonese presents an opportunity to study how sound change operates in a language which has "virtually no paradigms, either inflectional or derivational" (Chen & Wang 1975:256), such that the core tenets of both models can be evaluated, rather than for any exceptions found to be dismissed as another case of analogy.

Throughout this introductory discussion, a brief survey of the neogrammarian regularity hypothesis, lexical diffusion and lexical frequency effects in a usage-based model, it has become evident that the most productive way forward in studying alveolarization in Hong Kong Cantonese is through studying not only phonetic factors (such as neighbouring segmental environment, or tone of the sound segments undergoing change), but equally important to study social factors (gender, age, speaker-

based variation) and lexical factors (such as lexical frequency effects). Going back to the notion of ordered heterogeneity, with which we started this chapter, it is through examining the role of the speaker, the phonetic environment, but also the lexicon that we can fully understand the dynamic nature of sound change. As described above, this can be done by performing a quantitative sociophonetic variation study.

The focus of this thesis will be only on a limited set of linguistic variables—the velar coda consonants (**ŋ**) and (**k**) in Hong Kong Cantonese—but a comprehensive assessment of the predictions made by neogrammarian and lexical diffusion models of sound change will be performed. Alveolarization in Hong Kong Cantonese offers a classic case of a consonantal sound change involving shift in place of articulation, which—in accounts of historical linguistic change with a more formal phonological leaning—occurs when one phonological rule substitutes for another in certain phonological environments. Yet, lexical diffusionist accounts (e.g. Bauer 1979) have contended that it presents an archetypal case of lexical diffusion as lexical variation was found to exist in previous studies of Hong Kong Cantonese.

Given the above discussion, what predictions are made by the neogrammarian and lexical diffusion models, and how can we put them to the test in the case of alveolarization in Hong Kong Cantonese? The strongest possible case for lexical diffusion can be made if lexical diffusion can be attested (a) within a speaker's repertoire, and (b) within a given speech style, and (c) within the same phonetic environment. That is, a given sound change has affected some lexical items but not others that meet all of the above conditions. In addition, if the usage-based prediction that lexical frequency has an effect on sound change can be attested, and frequent (or infrequent) items can be shown to lead a particular sound change, this would constitute

strong evidence that the sound change transmits via lexical diffusion. On the other hand, the strongest case for neogrammarian regularity can be made if variation patterns can be solely explained by phonetic conditioning, but no (lexical) variation exists *within* phonetic conditions or stylistic factors. Furthermore, in a sound change in progress, it is generally expected that innovative variants are used more progressively by the younger speakers than by the older speakers. When examining age-based variation, the expectation would be that the younger speakers would show at least as much, if not a higher degree of alveolarization than the older speakers. These predictions will be put to the test within this dissertation.

1.11 Organisation of this thesis

The thesis is organised as follows. Chapter 2 describes the linguistic variables investigated here, namely alveolarization of the velar nasal (**ŋ**) and oral stop (**k**) codas. I start by presenting a phonology, which is followed by a review of sound changes occurring in Hong Kong Cantonese, with a focus on studies of coda consonants and alveolarization.

Chapter 3 describes the methodology adopted in this thesis. I describe the sampling methods employed, recording set-up, and the material used to elicit casual speech as well as reading data. In order to study lexical frequency effects, it was necessary to obtain lexical frequency information, which was extracted from the Hong Kong Cantonese Adult Corpus (Leung, Law & Fung 2001). The considerations which have gone into obtaining such information will be described in further detail in Chapter 3.

Chapter 4 describes the variation of velar codas in Hong Kong Cantonese using one statistical model which combines all the factors investigated in this thesis. The full model includes social factors such as age, gender, style, segmental environment such as

preceding vowel, tone, and following onset. It also includes lexical frequency effects. By doing so, it is hoped that an overall picture can be presented as to which factors seem the most significant in describing the variation.

Chapter 5 describes results from analyzing the social factors conditioning this sound change, which included age and gender. Age was treated as a continuous measure in the full statistical model, and this section will describe the interspeaker differences across age, as a way of projecting how the sound change might have progressed through the temporal dimension. However, the treatment of age in age groups is also briefly described, to offer a way of comparison with previous studies of the variable (**ŋ**). The possibility of age grading and methodological assumptions behind an apparent-time framework are also discussed.

Chapter 6 presents results from an analysis of neogrammarian factors which might operate on the sound change. The claim that it is primarily phonetic conditioning which drives sound change is examined by studying the phonetic environments which the velar codas can co-occur with. The preceding vowel, tone, and following onset are the immediate phonetic or phonological environments discussed here, and the results of a variation study of these factors will be presented.

Chapter 7 presents results of a lexical diffusion analysis. Similar to the age variable, lexical frequency can also be studied either as a categorical variable or a continuous variable, but here it is treated as a continuous variable. Treating lexical frequency as a continuum of the infrequent to the frequent prevents skewing the analysis by drawing an arbitrary distinction between frequent and infrequent items. Homophonous lexemes with different frequencies are also explored to advance our understanding of whether lexical frequency effects can be found in phonologically-

controlled environments where phonological conditioning supposedly does not operate and where the primary mode of transmission of sound change can logically only be through lexical diffusion.

Chapter 8 will summarise the findings of this thesis, and present implications that this study has on, not only the possible mechanisms of sound change of velar stops in Cantonese, but also on our understanding of how methodological considerations can influence what we can empirically find. It will also present future directions for further research.

2 Linguistic variables for this study: alveolarization in Hong Kong Cantonese

The linguistic variables in this thesis are the velar coda consonants **-ŋ** and **-k** in Hong Kong Cantonese. Since this is a sociovariationist study of sound change, this chapter will start with a description of recent sociolinguistic studies on alveolarization in Hong Kong Cantonese, with particular focus on age-related variation and evidence from previous studies suggesting that this is a sound change in progress. Alveolarization is also placed in the wider context of other sound changes in Hong Kong Cantonese in Section 2.2. The sociohistorical background of Hong Kong is furthermore elaborated on in Section 2.3, to allow readers to understand the social and demographic changes which have contributed to the formation of the modern Hong Kong Cantonese speech community. The remaining sections (Sections 2.4 and 2.5) will focus on language-internal factors, and describe the phonological contexts in which alveolarization of **-ŋ** and **-k** can occur.

Velar coda consonant alveolarization in Hong Kong Cantonese presents an interesting case for the study of sound change. As Bauer (1979:133-5) and Cheung (1986:213) describe, whether alveolarization occurred depended on which rimes the coda consonants were found in (reviewed in more detail in Section 2.3). Alveolarization of velar codas could involve a merger if a given rime with a velar coda has an alveolar counterpart. Therefore, it is a sound change that could create additional ambiguities to the already extensive homophony within the Cantonese lexicon. For instance, consider the standard Cantonese pronunciations for 行 'a row (of)' **hɔːŋ˨** and 百 'hundred' **baːk˧**. The alveolarized variants are [hɔːn˨] and [baːt˧], which would create confusion with 寒

'cold' **hɔ**:**n**˩ and 八 'eight' **ba**:**t**˧, respectively. To give readers without any background in Cantonese sufficient context to understand which phonological environments the velar codas **-ŋ** and **-k** can occur in, this chapter (Sections 2.3.1 to 2.3.5) also presents a phonological description of standard Hong Kong Cantonese.

A description of the historical linguistic development of the coda syllables from Middle Chinese to modern Cantonese is also provided in Section 2.7 and it is suggested that further significance of studying alveolarization of **-ŋ** and **-k** in Hong Kong Cantonese lies in obtaining a better understanding of the mechanisms which might have contributed to the development of coda syllables in other dialects.

2.1 Alveolarization as a sound change in progress

The first known account of alveolarization in Hong Kong Cantonese is the study of **-ŋ** variation by Bauer (1979). Bauer anecdotally suggested, through his own general observation of **-ŋ** produced by younger speakers as being different from the standard descriptions of Cantonese, that alveolarization was an ongoing sound change. This motivated him to investigate **-ŋ** in the speech of one young female speaker who was in her twenties, and to examine the phonological contexts in which alveolarization occurred. Since Bauer's (1979) study there have been a number of quantitative sociolinguistic studies on sound changes in Cantonese, particularly of the syllable onset consonants (Yeung 1980; Bauer 1983; Pan 1982; Bourgerie 1990; Ho 2004). Two studies, Shum (1993) and Wong (2005), have examined age-related variation in the alveolarization of **-ŋ** and found that the younger speakers made more use of the innovative alveolar coda than older speakers, suggesting that alveolarization of **-ŋ** was a

sound change in progress.

Shum (1993) looked at the velar nasal coda **-ŋ** in the speech of 40 subjects. The subjects were divided into two age groups: between 15 to 20 years of age and between 50 to 55 years of age. Shum found that the use of innovative [-n] was higher in the younger age group, with the sound change being phonologically conditioned and more likely to occur in rimes containing the central vowels **a**:, **ɐ**, the mid-open front vowel **ɛ**:, and back vowel **ɔ**:. The phonological contexts will be described in further detail in Sections 2.3 and 2.4.

A sociolinguistic study performed by Wong (2005) found an age difference, with younger speakers making more use of the innovative variants overall, although this was not tested for statistical significance. Wong investigated the variation of velar nasal codas in informants stratified according to sex[10] and four age groups, namely 11–20, 21–30, 31–40, and 41 years of age or above. She found stylistic as well as age-related variation. While she did not report how many tokens were analysed in total, her results are summarised in the Table 14. The overall average percentages by age group are displayed in Figure 3.

Table 14. Percentage of innovative alveolar realization of (ŋ) across age groups and speech styles, in Wong (2005)

Age groups	Formal	Conscious	Casual	Total
11–20	11.46%	33.33%	42.01%	24.14%
21–30	5.23%	8.70%	13.02%	8.19%
31–40	5.54%	13.04%	20.12%	11.20%
41 or above	2.08%	4.35%	10.06%	4.94%

[10] The "sex" variable will be discussed in further detail in Section 3.4. Here the focus is on describing the age variable and reviewing evidence which point to alveolarization as a sound change in progress.

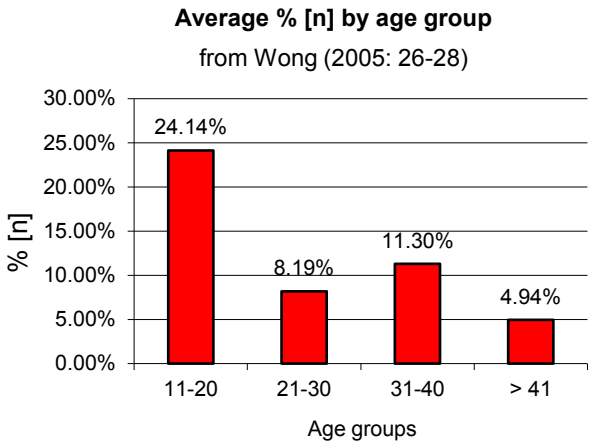


Figure 3. Average percentage of [-n] realization by age group, in Wong (2005)

Wong found that speakers produced the highest percentage of innovative alveolar realizations in casual speech, with progressively lower percentages as the formality of the speech setting increased. This trend also applied within each age group. One anomaly was that instead of there being a gradation of increasing percentages of [-n] realization down the age groups, the speakers aged between 21 and 30 showed a dip in percentage [-n] realization, suggesting that they made less use of the innovative variant than expected. These findings will be compared against the age-related findings in this study in Chapter 5, and any implications that this irregularity might have will be further explored.

Taken together, the age-based variation as found in the two studies by Shum (1993) and Wong (2005) suggest that sound change was progressing throughout the 90s and 00s. Assuming that the sound change was in progress and that speaker characteristics are relatively fixed in early adulthood, one would expect that the sound change would be well underway for speakers in their 50s and 60s (at the time of this study), and that if the sound change progressed even further, the younger speakers today

would make even more use of the innovative alveolar variants. This prediction will be revisited when the results of age-based variation will be discussed in detail in Section 5.1.

Compared to **-ŋ**, alveolarization from velar oral stop **-k** to [-t] has received little attention in the sociolinguistic literature on Cantonese. One of the few studies which has been performed on this variable is Zee's investigation of the careful reading by a native Cantonese-speaking 21-year-old male informant (1996). Other than describing that this one speaker displayed variation, including alveolarization, in his production of **-k**, there was no further examination of speech in other speakers, nor any information on age-based variation of **-k**. This dissertation hopes to address this gap in the literature by also examining variation in the production of **-k** by speakers of different age groups, to see whether there is any evidence of age-based variation or change in progress, and if yes, whether this change is more or less progressive than **-ŋ**.

2.2 Sound changes in Hong Kong Cantonese

Alveolarization of **-ŋ** and **-k** in coda position is only one type of a plethora of sound changes occurring in Hong Kong Cantonese. The ones that the general population is aware of are known as **la**:**n**˩ **jɐm**˥ 懶音, literally "lazy pronunciation", which, as the judgmental label probably already suggests, has engendered a lot of debate amongst educators. Some of these sound changes do not involve a loss of lexical distinction, while others do. For instance, the **ŋ-** in syllable-initial position is being lost and produced as a glottal or deletion [ʔ-/ø-], such that **ŋau**˩ 牛 'cow' is pronounced as [ʔau˩] or [au˩]. However, since there is no existing standard morphosyllable **au**˩ (with this particular combination of syllable structure and tone), this does not create any

confusion with another lexical item. There are sound changes, on the other hand, which involve a loss of lexical distinction through the merging of two consonants. The syllable-initials **n-** and **l-** have effectively merged in everyday speech, thereby eradicating the distinctions such as those between **nei**˩ 你 'you' and **lei**˩ 理 'reason/logic' so that both are pronounced [lei˩], except in broadcasting or formal register, where the distinction might still be maintained. The same holds for **k^{w}-** and **k-,** creating a loss of distinction between **k^{w}ɔːŋ**˥ 廣 'broad' and **kɔːŋ**˥ 講 'to talk/tell', with both pronounced as [kɔːŋ˦] (at least for those who do not display alveolization in addition). For more details of variation and sound changes in Hong Kong Cantonese, see e.g. Hashimoto (1972), Yeung (1980), Bauer (1983), Cheung (1986), Bourgerie (1990), Shum (1993), Zee (1996), Chen (1998), Law et al. (2001), Ho (2004), Wong (2005), Lee et al. (2006), Chan (2008), Ding (2010). Furthermore, there are tonal changes involving a merger of the high falling ˥˩ to high level ˥ tone and high rising ˧˥ to mid rising ˧˥ as discussed above in Section 2.3.5 (So 1996; e.g. Bauer & Matthews 2003; Bauer, Cheung & Cheung 2003). These sound changes are beyond the scope of this dissertation, but are mentioned for the reader to appreciate the wider set of sound changes in Hong Kong Cantonese which have been described in the literature. The next section describes the socio-historical context of Hong Kong Cantonese as well as provides a phonology for those readers who are unfamiliar with the language, and would like to understand the phonological contexts in which alveolarization of **-ŋ** and **-k** might occur.

2.3 Hong Kong Cantonese

Cantonese belongs to the Yue language family, one of seven major Sinitic language families (Zhan 1981; Baxter 1992:9; Yan 2006). As Figure 4 shows, the Yue language family belongs to the Southern branch of the Sino-Tibetan language family. The Yue language family is spoken primarily in the southeast of China, including amongst others Guangdong, Guangxi and Hong Kong. After British occupation of Hong Kong Island in 1842 and Kowloon in 1860, trading and commerce prospered, attracting a large influx of immigrants from the neighbouring Guangdong province. It was from this period onward through waves of immigration that Cantonese was brought into Hong Kong on a large scale. Cantonese derives its name from the city of Guangdong, which is the capital of Guangdong province and was formerly known as 'Canton'. This city is found within geographical proximity and is connected to Hong Kong via riverways in the Pearl River Delta (see Figures 5 and 6). According to a government census conducted in 2011, 89.6% of the population (totalling some 6.6 million people), spoke Cantonese (Census 2011). In addition, there is a widespread diaspora of Cantonese speakers in overseas communities across North America, Europe and Australia. However, this dissertation focuses only on Cantonese spoken within Hong Kong.

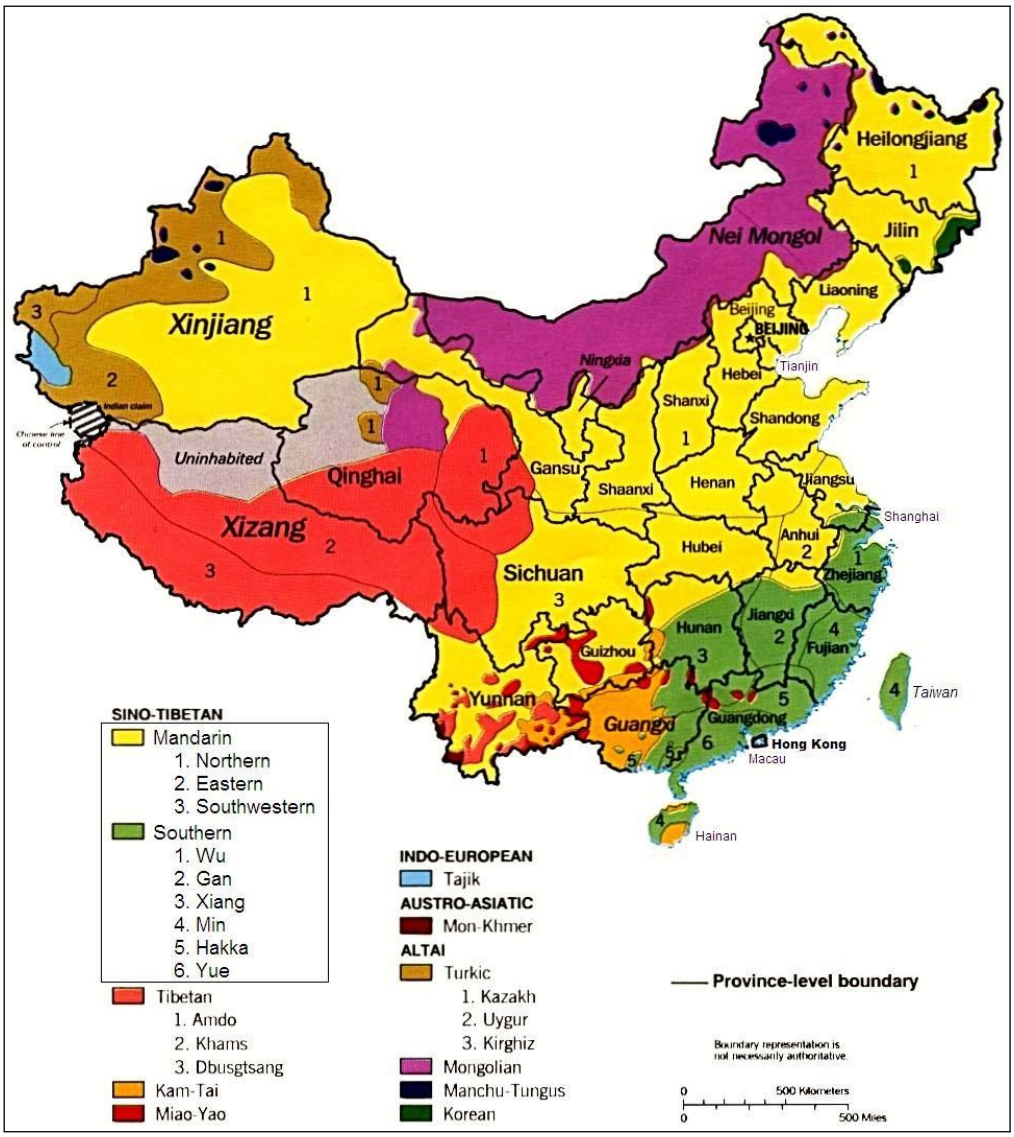


Figure 4. Sinitic language families in China[11]

[11] Source: http://www.lib.utexas.edu/maps/middle_east_and_asia/china_ling_90.jpg/ (Accessed 28 April 2012. Courtesy of the University of Texas Libraries, The University of Texas at Austin.)

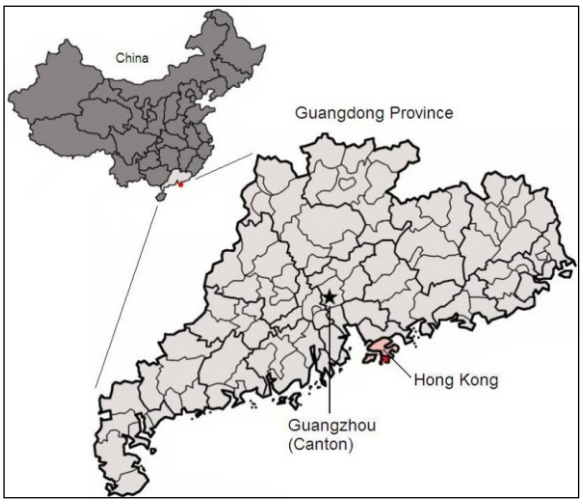


Figure 5. Location of Hong Kong on a map of China, also showing Guangzhou, formerly 'Canton', from which Cantonese originated [12]



Figure 6. Hong Kong bordering Shenzhen and the Pearl River Delta [13]

[12] Adapted from: http://commons.wikimedia.org/wiki/File:Location_of_Hongkong_by_Guangdong_(China).png/ (Accessed 29 April 2010).

[13] Source: http://commons.wikimedia.org/wiki/File:Pearl_River_Delta_Area.png/ (Accessed 29 April 2010).

2.3.1 Cantonese phonetics and phonology

Like most other Sinitic languages, Cantonese is a tone language, with a relatively simple syllable structure. In line with the long history of the Chinese writing system, where the basic orthographic component is the character and each character represents a morpheme, it is apt to describe the character as representing a morphosyllable: a single syllable with a distinct meaning.[14] Most words in modern Chinese languages consist of two or more morphosyllables. However, in this discussion of Cantonese phonology, the focus will be mostly on each individual morphosyllable, and leaves aside many of the questions associated with the combination of multiple syllables, such as tone change[15] and to a very limited extent, tone sandhi[16].

In the ensuing phonological description of Cantonese, I first describe the Cantonese syllable structure and the components which make up the Cantonese syllable in standard descriptions. This is followed by a discussion of phonotactic constraints on syllable structures in Cantonese and the related question of onset-coda co-occurrence restrictions. These will feed into an understanding of the syllables which can undergo alveolarization in Hong Kong Cantonese, the target of this investigation.

---

[14] A small number of Chinese characters which are ligatures, or héshū 合書, are exceptions to this, as they have the form of one character but combine multiple characters, and are thus typically pronounced using more than one syllable. An example would be "招財進寶" combined into a single ligature character . These characters, however, are not part of the standard writing system but are a special calligraphic / artistic device for very limited uses (e.g. Chinese New Year wall posters).

[15] In Cantonese, the phenomenon whereby tones change due to morphological and semantic factors (e.g. in compounding and reduplication), referred to as simply "tone change", should be kept distinct from tone sandhi, which occurs in Mandarin. Tone sandhi is a change in tone which happens when two tones are put adjacent to one another (for further details about Cantonese tone change, see Matthews & Yip 1994:23; Hashimoto 1972:93–100).

[16] See, e.g. Hashimoto (1972:112–3).

2.3.2 The Cantonese syllable structure

The basic syllable structure of Cantonese can be analysed as consisting of an obligatory nucleus with optional onset and coda consonants. This structure can be represented as follows (cf. Matthews & Yip 1994:19):

$$(C_o)\ V(:)\ (C_c)$$

Figure 7. Cantonese syllable structure

where C_o stands for onset consonant, V for nuclear vowel, C_c for coda consonant, and brackets ( ) for optional elements. The nucleus is represented by a vowel; however, the generalisation given in Figure 7 misses one possibility, namely the syllabic nasal consonants **m̩** and **ŋ̩** (e.g. **m̩˩** 唔 'negation', and **ŋ̩˧˥** 五'five', although the latter is more often pronounced [m̩] by the younger generation). As indicated by (:), the vowel occupying the nucleus position can be either short or long, but when not followed by a coda the nuclear vowel must be long.

In classical Chinese linguistic descriptions, the syllable is analysed into an 'initial' (or *shēngmǔ* 聲母) and a 'final' (*yùnmǔ* 韻母). The initial is equivalent to the onset consonant, while the final—for the purpose of analysing Cantonese—is equivalent to the rime (nucleus and coda) plus tone.[17] The method on which this initial-final analysis is based is called *fǎnqiè* (反切), which dates back to at least as early as A.D. 601, when

[17] In accounts of Old and Middle Chinese, the final is analysed as maximally containing a medial, a nuclear vowel and a coda, plus tone. Since Cantonese does not have the medial **j** which some other Chinese dialects have retained, and because this work is not a comparative study of Cantonese with respect to other Chinese dialects, we do not incorporate the medial in this analysis. The labialised velars **k^{w}** and **k^{hw}** are therefore analysed as a single segment, as shown in Table 15, rather than an onset + medial glide **w**.

Qièyùn (切韻) was published, even though the method probably already existed centuries before that (Baxter 1992:35). In *fǎnqiè*, the sound of a character is given through two other characters. For instance, the sound of 北 **bɐk**˥ 'north' can be given through '白黑切' in *fǎnqiè* notation. In this case, the onset **b** is obtained from the onset of the first character 白 **baak**˨ 'white' and the rime including tone of the second character 黑 **hɐk**˥ 'black'. While I recognize the importance of the onset + rime analysis in Chinese linguistics, the ensuing discussion will furthermore decompose the rime into the segmental sub-units, as it helps us isolate the segmental change from velar to alveolar coda consonants.

2.3.3 Onsets

There are 19 consonants in the phonological system of Cantonese,[18] summarised in Table 15. All of these consonants can appear in the onset position, where aspiration is contrastive. For instance, the unaspirated stop consonants **p-**, **t-**, **k-**, **ts-** are contrastive with the aspirated set **p^{h}-**, **t^{h}-**, **k^{h}-**, **ts^{h}-** respectively. In addition, the velar stops **k-** and **k^{h}-** also have contrastive labialised counterparts **k^{w}-** and **k^{hw}-**, respectively.

[18] Some analyses, however, assume that C_o is an obligatory element in the syllable structure represented in (1) and consider the zero consonant, Ø, as part of the inventory, bringing the number up to 20.

Table 15. Cantonese consonants [19]

	LABIAL	DENTAL/ALVEOLAR	PALATAL	VELAR	GLOTTAL
STOP					
Unaspirated	p	t		k kʷ	
Aspirated	pʰ	tʰ		kʰ kʰʷ	
NASAL	m	n		ŋ	
FRICATIVE	f	s			h
AFFRICATE					
Unaspirated		ts			
Aspirated		tsʰ			
APPROXIMANT	w	l	j		

Adapted from Bauer & Benedict (1997: 17)

The contrast between aspirated versus unaspirated onset stops is exemplified in Table 16. In each pair, the rimes and tones are the same, with the only difference in aspiration vs. non-aspiration in the onset consonants. Thus, the unaspirated voiceless bilabial stop in **pɐt˥** (畢, 'finish') contrasts with the aspirated voiceless bilabial in **pʰɐt˥** (匹, classifier for 'horse'), and so on.

Table 16. Aspirated and unaspirated onset minimal pairs

Unaspirated onset stops		Aspirated onset stops
畢 **pɐt˥** 'finish'	vs.	匹 **pʰɐt˥** 'classifier for 'horse'
打 **ta**:˧˥ 'dozen'	vs.	他 **tʰa**:˧˥ 'he'
加 **ka**:˥ 'add'	vs.	卡 **kʰa**:˥ 'stuck'
瓜 **kʷa**:˥ 'melon'	vs.	誇 **kʰʷa˥** 'exaggerate'
炸 **tsa**:˧ 'to fry'	vs.	詫 **tsʰa**:˧ 'surprise'

[19] According to Robert S. Bauer & Benedict (1997: 17), ʃ, tʃ, and tʃʰ exist as the palatalised allophones of s, ts, and tsʰ respectively.

2.3.4 The rime: nucleus and coda

The rime consists of a nucleus and an optional coda. Most analyses recognise a set of 13 phonetically distinct vowels which can appear in the nuclear position (e.g. Kao 1971; Cheung 1972; Hashimoto 1972; Chao 1947; and Zhan & Cheung 1987; see also discussion in Bauer & Benedict 1997:39). These vowels are summarised in Table 17. As the table shows, the vowels consist of a set of six short vowels, **ɐ**, **ɵ**, **ɪ**, **ʊ**, **o** and **e**, and a set of seven long vowels, **a**:, **ɛ**:, **i**:, **y**:, **œ**:, **ɔ**: and **u**:, respectively. Short vowels cannot appear without any coda consonant or offglide, with only long vowels able to appear on their own in the rime.

Table 17. Cantonese nuclear vowels

Short Vowels	Examples	**Long vowels**	Examples
ɐ	畢 **pɐt**˥ ‘finish’	**a**:	打 **ta**:˥ ‘dozen’
ɵ	出 **ts^{h}ɵt**˥ ‘to exit’	**ɛ**:	蛇 **sɛ**:˩ ‘snake’
ɪ	食 **sɪk**˩ ‘to eat’	**i**:	思 **si**:˥ ‘to think’
ʊ	六 **lʊk**˩ ‘six’	**y**:	書 **sy**:˥ ‘book’
e	你 **nei**˩ ‘you’	**œ**:	強 **k^{h}œ**:**ŋ** ˩ ‘strong’
o	高 **kou**˥ ‘tall’	**ɔ**:	過 **k^{w}ɔ**:˥ ‘to pass’
		u:	姑 **ku**:˥ ‘aunt’

A representation of the vowel space occupied by these phonetic categories is shown in Table 18. In the front there are the unrounded vowels [iː], [ɪ], [e] and [ɛː], and the rounded vowels [yː], [ɵ], and [œː]. There are two central unrounded vowels [ɐ] and [aː], and the back vowels consist of rounded [uː], [ʊ], [o] and [ɔː], respectively.

Table 18. Cantonese vowel space*

	Front		Central	Back
	Unrounded	**Rounded**	**Unrounded**	**Rounded**
High	[iː]	[yː]		[uː]
Mid-high	[ɪ]			[ʊ]
Mid	[e]	[ɵ]		[o]
Mid-low	[ɛː]	[œː]	[ɐ]	[ɔː]
Low			[aː]	

* based on (Cheung 1972:3; Bauer & Benedict 1997:40)

Note that while these vowels are phonetically distinct, these are not strictly phonemes. When considering how these vowels combine with coda consonants to form rimes, their distribution and phonemic contrasts can be more easily compared. The full set of coda consonants are the nasals **m**, **n**, **ŋ** and stop consonants **p**, **t**, **k**. Phonetically, the latter set of stop coda consonants are unreleased, and accounts have occasionally described them as glottalized (Hashimoto 1972:91; Karlgren 1926).

Table 19, adapted from Barrie (2003:5) and originally from Hashimoto (1972:90), summarises the set of possible rimes in Cantonese, i.e. those which can be found in canonical standard descriptions of Cantonese. In Table 19, rows display the nuclear vowels of Cantonese, while columns show the coda consonants. As mentioned before, and as captured by the first column of codas, only long vowels can appear without any coda consonant (the absence of coda consonants is indicated by the symbol for the empty set, ø). In other words, all short vowels need to be followed by a coda. The second, third and fourth columns show how the nuclear vowels can combine with offglides **i**, **y** and **u** to form diphthongs. The six remaining columns show nuclear vowels combining with the nasals **m**, **n**, **ŋ** and stop consonants **p**, **t**, **k**, respectively.

Table 19. Standard Cantonese rimes[20]

coda / V	Ø	ɪ	y	u	m	n	ŋ	p	t	k
iː	iː			iːu	iːm	iːn		iːp	iːt	
ɪ							**ɪŋ**			**ɪk**
e		ei								
ɛː	ɛː			ɛːu^{+}	ɛːm^{+}	ɛːn^{+}	**ɛːŋ**	ɛːp^{+}	ɛːt^{+}	**ɛːk**
uː	uː	uːi				uːn			uːt	
ʊ							**ʊŋ**			**ʊk**
o				ou						
ɔː	ɔː	ɔːi				**ɔːn**	**ɔːŋ**		**ɔːt**	**ɔːk**
yː	yː					yːn			yːt	
ɵ			ɵy			ɵn			ɵt	
œː	œː					œːn^{+}	**œːŋ**		œːt^{+}	**œːk**
aː	aː	aːi		aːu	aːm	**aːn**	**aːŋ**	aːp	**aːt**	**aːk**
ɐ		ɐi		ɐu	ɐm	**ɐn**	**ɐŋ**	ɐp	**ɐt**	**ɐk**

$^{+}$ Rimes which occur in colloquial or onomatopoeic morphosyllables that are generally not included in standard descriptions, but which some accounts recognise as well-formed (e.g. Zhan & Cheung 1990; Li et al. 1995; see also Bauer & Benedict 1997:44–5).

Note: cells which are *highlighted* contain rimes of interest in this study. The blue highlight indicates rimes which, when alveolarized would *not* lead to a loss of phonological distinction. On the other hand, rimes highlighted in yellow are ones which when alveolarized would lead to a loss of phonological distinction.

Furthermore, phonetic length of the nuclear vowel is dependent on whether there is a coda. In the absence of a coda (first column of Table 19), the nuclear vowel length is the longest. A long vowel with a coda consonant is medium long, while a short vowel with a coda consonant is the shortest (for a discussion and vowel length measurements, see Hashimoto 1972:90; Kao 1971; and Bauer & Benedict 1997:33–9).

[20] Analysis similar to Barrie (2003:5), but using IPA notation following Zee (1999:58-60). Nuclear vowels are arranged in rows; coda consonants are arranged in columns

In Table 19, rows of nuclear vowels (with the exception of **y**:) have been grouped into long and short pairs: **i**:/**ɪ**, **e**/**ɛ**:, **u**:/**ʊ**, **o**/**ɔ**:, and **ɵ**/**œ**:. As can be seen from the distribution of these vowel pairs with coda consonants, each vowel in a pair is in complementary distribution with the other vowel in the same pair, which is to say that one vowel in each of the pairs will not combine with the same coda consonants as the other vowel within the pair. The only exception to this is the **a**:/**ɐ** vowel pair, which is the only vowel pair which can appear with almost the full range of final nasal and stop coda consonants. Further exceptions are rimes marked with $^{+}$, particularly the rimes **œ**:**n** and **œ**:**t**, which are rimes recognized by some phonological accounts (see e.g. Kao 1971; Cheung 1972; Hashimoto 1972; Bauer & Benedict 1997:39–44) as existing in colloquial and onomatopoeic morphosyllables, but which standard descriptions of Cantonese tend not to explicitly recognise as well-formed rimes.[21]

The complimentarity of the long/short vowel pairs has allowed these pairs to be collapsed into single phonemes with allophonic distinctions amongst the 13 phonetically distinct vowels described earlier (Cheung 1972:3; see also Bauer & Benedict 1997:42). These are summarised in Table 20.

[21] While I am open to embracing colloquial morphosyllables as part of Cantonese, since this dissertation does not focus on Cantonese phonology, it does not matter much for our purposes whether rimes marked with $^{+}$ are included are not. If we are intent on studying the loss of phonological contrast, it is still possible to focus only on the rimes consisting of the nuclear vowels {**ɔ**:, **a**: or **ɐ**} + velar codas {**ŋ** or **k**}.

Table 20. Cantonese phonemes and allophones*

Front		Central	Back
/**i**ː/ /**y**ː/	[iː] [ɪ] [yː]		/**u**ː/ [uː]
/**e**/	[ɛː] [e]		/**ɔ**ː/ [ɔː]
/**œ**ː/	[œː] [ɵ]	/**a**/ [ɐ]	
		/**a**ː/ [aː]	

* description according to Cheung (1972:3), cf. Bauer & Benedict (1997:42)

As can also be seen from Table 19, the vowels **i**ː, **e**, **u**ː, **o**, as well as the rounded short vowel **ɵ** do not combine with velar coda consonants, with only the vowels **ɪ**, **ɛ**ː, **ʊ**, **ɔ**ː, **œ**ː, **a**ː and **ɐ** combining with velar coda consonants. In addition, **ɔ**ː, **a**ː and **ɐ**, can also co-occur with alveolar coda consonants **-n** and **-t**, and are the only vowels which can co-occur with both velar and alveolar coda consonants. The long vowels **a**ː, **œ**ː, **ɔ**ː, and **ɛ**ː, and short vowels **ɐ**, **ʊ**, and **ɪ** are thus the segmental environments which can immediately precede the velar coda consonants studied in this thesis.

The rimes which are relevant to our investigation of alveolarization are those which include velar coda consonants **-ŋ** and **-k**, highlighted in colour in Table 19. The rimes highlighted in light blue can undergo alveolarization, but will not create confusion, as no rime with alveolar counterpart exists in the standard description. The rimes highlighted in yellow are pairs where, if the velar coda were to be alveolarized, neutralization with the rime containing the alveolar counterpart would result, if the tones also matched. This has ramifications for our investigation of phonological conditioning of sound change, as it implies that for rimes involving the nuclear vowels **ɪ**, **ɛ**ː, **ʊ**, **œ**ː and a velar coda consonant, alveolarization does *not* lead to a loss in phonological contrast. Contrast this with rimes including nuclear vowels **a**ː or **ɐ** or **ɔ**ː

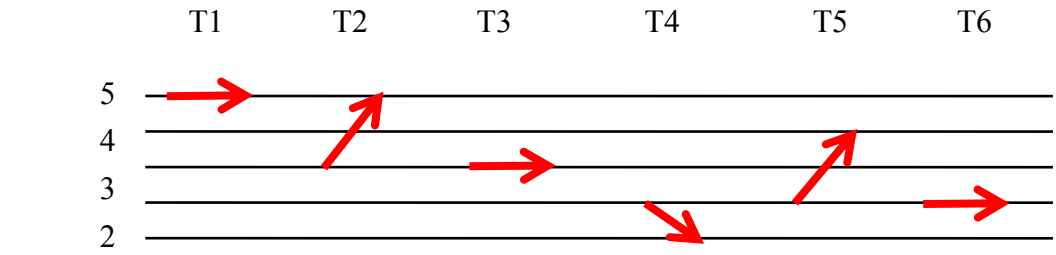
and a velar coda consonant, where alveolarization *could* lead to a loss of phonological contrast, if the syllables had the same onset and tone. In this thesis, asymmetries in coda consonant distribution are also explored to examine the effect of phonological organisation on sound change. In particular, a question to be addressed includes, would rimes such as **ɪŋ**, **ɪk**, **ɛ**:**ŋ**, **ɛ**:**k**, **ʊŋ**, **ʊk**, **œ**:**ŋ**, and **œ**:**k**, which do not have firmly established alveolar counterparts undergo alveolarization more readily than rimes such as **ɔ**:**ŋ**, **ɔ**:**k**, **a**:**ŋ**, **a**:**k**, **ɐŋ**, and **ɐk**, which do have alveolar counterparts, and for which confusion with existing morphosyllables would result when the coda is alveolarized? This question will be investigated by comparing how likely each of these sets of rimes undergoes alveolarization in Chapter 6.

2.3.5 Tones

Cantonese is a tonal language, which means that differences in pitch can create a lexical distinction. Contemporary accounts of standard Hong Kong Cantonese agree that it has six lexical tones (Matthews & Yip 1994:20; Cheung 1986), which include a combination of three level tones and three tonal contours. A schematic way of representing these lexical tones employs five relative pitch levels distributed across an individual's vocal range, from the lowest relative pitch value 1, to the highest, 5. A listener can adjust their expectation of the pitch range of a speaker and perceive pitch levels relative to the speaker's fundamental frequency average (Bauer & Benedict 1997:112).

Tone 1 is the high level tone, and can be represented as starting at the highest pitch level 5 and ending at 5. Using numerical notation, this pitch can be represented as 55 ('five-five'). Tone 2, the mid rising tone, starts at the middle pitch level, 3, and ends

at 5. This pitch contour can be represented numerically as 35. Similarly, Tone 3, the mid level tone, can be represented as 33; while Tone 4, the low falling tone, Tone 5, the low rising tone, and Tone 6, the low level tone, can be represented numerically as 21, 24 and as 22, respectively. These pitch levels and contours are represented in Figure 8.



Pitch levels:	11	35	33	21	24	22
IPA:	˥	˧˥	˧	˩˧	˨˦	˨

Figure 8. Schematic pitch diagrams of the Cantonese lexical tones

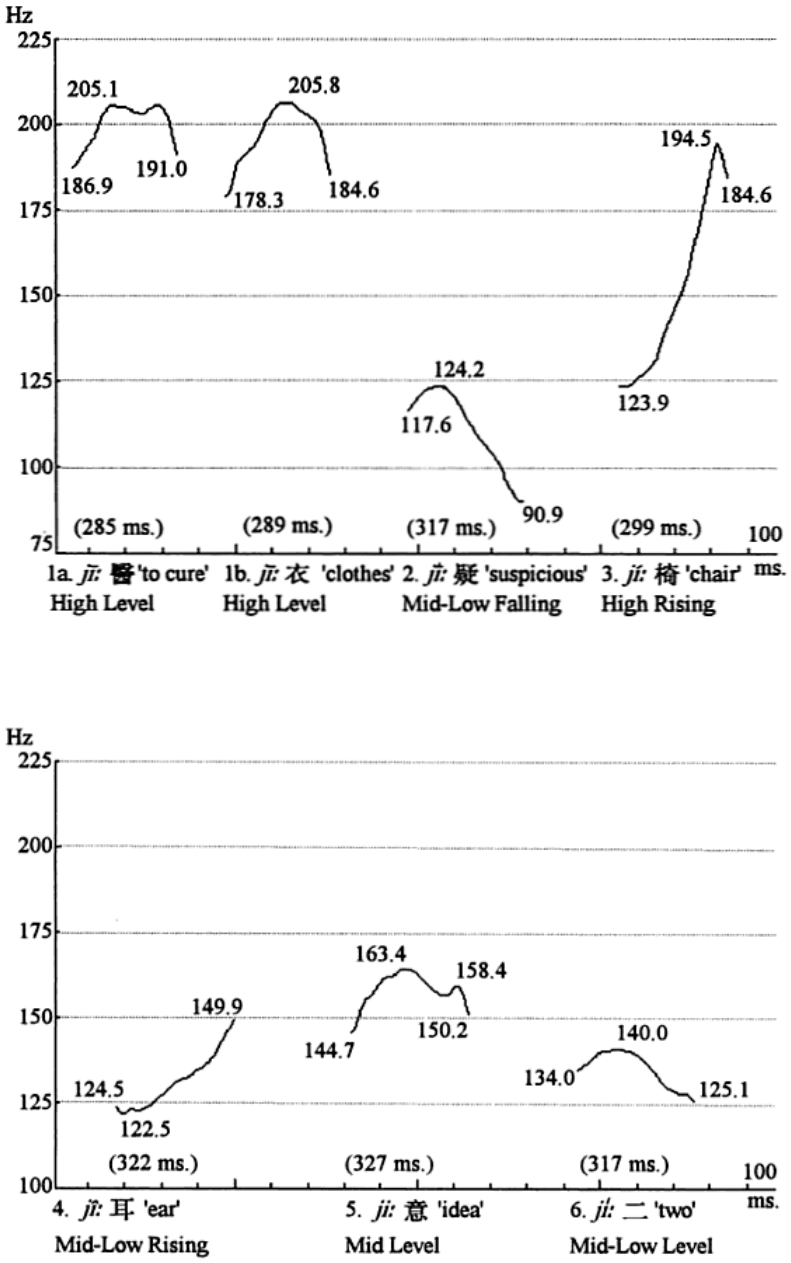
Y.R. Chao, a Chinese-American linguist, represented tones using time-pitch graphs not dissimilar to the schematic representation in Figure 8, to indicate the relative pitch change in the temporal dimension (Chao 1933). Subsequently, he also used them for the description of Cantonese tones (Chao 1947). These time-pitch graphs indicate the relative change in pitch level over time, with the end marked with a vertical bar. The time-pitch graphs for Cantonese tones are ˥, ˧˥, ˧, ˩˧, ˨˦ and ˨ for Tones 1-6 respectively. These time-pitch symbols were subsequently adopted by the International Phonetics Association (International Phonetic Association 1989:75-77), and are currently standard symbols for representing tone. Transcriptions in this thesis, as the reader might have noticed, include the Chinese character, phonological transcription and IPA tone notation

(i.e. ˥, ˧˥, ˧, ˨˩, ˧, or ˨) to indicate the tone which the syllable carries, as exemplified in the final column of Table 21.

Table 21. Numerical and IPA notation for tones in Cantonese

	Tone	**5-level numerical pitch notation**	**IPA notation**	**Example**
High Level	1	55	˥	思 **si**: ˥ 'to think'
High Rising	2	35	˧˥	史 **si**: ˧˥ 'history'
Mid Level	3	33	˧	試 **si**: ˧ 'to try'
Mid-Low Falling	4	21	˨˩	時 **si**: ˨˩ 'time'
Mid-Low Rising	5	24	˧˥	市 **si**: ˧˥ 'market'
Mid-Low Level	6	22	˨	事 **si**: ˨ 'a matter'

The above phonological description of Hong Kong Cantonese tones have to a large extent been validated by acoustic studies of fundamental frequency (F0). From the F0 traces from 28-year-old male who is a native speaker of Cantonese (Bauer & Benedict 1997:250), one can see that the F0 traces of each tone are impressionistically similar to the schematic descriptions of tone presented above.



From top to bottom, left to right: Tone 1 (1a. and 1b. high level tone), Tone 4 (2. mid-low falling), Tone 2 (3. high rising), Tone 5 (4. mid-low rising), Tone 3 (5. mid level), and Tone 6 (6. mid-low level). Figure from Bauer (1997:250) Figure 2.3, based on data from a 28-year-old male native speaker of Hong Kong Cantonese.

Figure 9. Fundamental frequency (F0) values of lexical tones in Hong Kong Cantonese

Traditionally, the tones carried by checked syllables, which have a short nuclear vowel and a final **-p**, **-k**, or **-t**, were considered 'clipped tones', which meant a further three types of tones were added to a total of nine. But in this study, since there is little phonetic motivation for considering them distinct from the three level tones, high level ˥, mid level ˧, and mid-low level ˨ tones, these 'clipped tones' will be considered shorter versions of the same level tones. The Guangdong variety of Cantonese, from which the Hong Kong variety has been derived, is typically considered to have seven lexical tones, the additional tone being the high falling one, ˥˩, as opposed to only six in Hong Kong Cantonese. There is also recent evidence that T2 (high rising, ˧˥) and T5 (mid-low rising, ˨˧) are merging into one (e.g. Fung & Wong 2011). However, tonal mergers as well as changed tone (tonal assimilation 變調 and morphological changed tone 變音) phenomena are beyond the scope of this thesis, and the six-way tonal distinction used in most descriptions of contemporary Hong Kong Cantonese will be adopted. In all, it is important to bear in mind—especially for those who are unfamiliar with lexical tones—that even when two syllables have the same onset, nucleus and coda, they are still not homophonous when they carry a different tone. In Section 7.3, where an analysis of homophonous lexemes are offered, this means that lexemes are identical with respect to the syllable substructures *as well as* tone.

2.4 Phonological conditioning of **-ŋ** to [-n] variation

As mentioned in Section 2.1, Robert Bauer's (1979) account of alveolarization of velar nasal coda **-ŋ** to [**-**n] [22], using data collected from one native Cantonese informant in a word list reading task, provided evidence that certain preceding vowels favoured alveolarization more than others. In particular, it was found that the rimes **-a**:**ŋ** and **-ɐŋ** led the change, practically always being alveolarized by this informant, followed by **-ɛ**:**ŋ**, which was almost always alveolarized, and **-ɔ**:**ŋ** and **-œ**:**ŋ**, which were only occasionally alveolarized. On the other hand, **-ɪŋ** and **-ʊŋ** never underwent alveolarization (Bauer 1979:133).

The data from Bauer (1979) of alveolarization of Cantonese **-Vŋ** finals are reproduced in Tables 22 to 26. As Tables 22 and 23 show, according to Bauer's account, all the lexical items investigated which contained the finals **-a**:**ŋ** and **-ɐŋ** seemed to have undergone alveolarization within this speaker.

[22] Bauer (1979:133) suggested that it was hard to ascertain the extent of final velar stop –**k** alveolarization due to extensive reduction to glottal [-ʔ] by his subject and thus left out the velar stops from his analysis.

Table 22. Lexical items undergoing complete change from -aːŋ > [-aːn]

	Tone + Final with –aːŋ > -aːn change					
Onset	**High Level** ˥	**Mid Rising** ˧˥	**Mid Level** ˧	**Low Level** ˩	**Low Falling** ˩˧	**Low Rising** ˨˩
p			蹦	蹦		
p^h					棚 膨 蟛	棒
k	耕					
k^h						
ts	掙 箏 爭 踭					
ts^h	撐 罉	橙	撐		倀	
m	鰻 蜢		孟		盲	猛
ŋ				硬		
h					行	
s	生 牲	省				
w					橫	
l						冷

(from Table I in Bauer 1979:137, Table I)

Table 23. Lexical items undergoing complete change from -ɐŋ > [-ɐn]

	Tone +Final with -ɐŋ > [-ɐn] change					
Onset	**High Level** ˥	**Mid Rising** ˧˥	**Mid Level** ˧	**Low Level** ˩	**Low Falling** ˩˧	**Low Rising** ˨˩
p	繃 崩			憑		
p^h					朋 憑[23]	
t	登 燈 瞪	等	凳	鄧		
t^h					騰 籐 謄	
k	更 羹 庚	梗 鯁 耿	更			
k^h		啃				
ts	增 曾 僧 憎			贈		
ts^h					層 曾	
m	猛				萌 盟	
n					能	
h	亨 莖	肯		幸 杏	行 衡 恆	
s	生 牲					
w					宏 閎 弘	
Ø	鶯 嚶					

(from Table II in Bauer 1979:137, Table II)

[23] Bauer also includes 貧, but this originally has a **-ɐn** final, not a **-ɐŋ** one. He also includes 凭, but this is merely an orthographic variant of 憑.

On the other hand, as shown in Table 24 (from Bauer 1979:138), lexical items containing the **-ɛːŋ** rime showed much more variation. The characters which are in bold and have greyed-out highlights are those which were pronounced with the innovative alveolar **-ɛːn** rime. The lexemes which are not highlighted showed variation between [-ɛːŋ] and [-ɛːn]. As can be seen, there is no clear pattern of phonological conditioning by the syllable-onset consonant or tone.

Table 24. Lexical items with -ɪŋ / -ɛːŋ finals displaying varying alveolarization

-ɪŋ / -ɛːŋ → [-ɪŋ] / [-ɛːn]

Bold and highlighted characters were pronounced only with alveolar [-ɛːn]. All others were pronounced with the standard [-ɪŋ] or [-ɛːŋ].

	Tone + Final -ɪŋ / -ɛːŋ → -ɛːn					
Onset	**High Level** ˥	**Mid Rising** ˧˥	**Mid Level** ˧	**Low Level** ˩	**Low Falling** ˩˧	**Low Rising** ˨˩
p		餅	柄	病		
p^h					平	
t	**釘**	頂		定 **掟**		
t^h	聽 廳					**艇**
k	驚	**頸**	**鏡**			
k^h						
ts	精	**井**	正	**阱 鄭** 淨		
ts^h	青	請				
m		**名**		命		
h	輕					
s	聲 腥	醒			成	
j		影			贏	
l		**零**	**靚**		零 **靈 鯪**	領 嶺

(from Bauer 1979:138, Table III)

The same seemed to be true for lexical items with **-œːŋ** and **-ɔːŋ** finals, as can be seen from Tables 25 and 26.

Table 25. Lexical items with -œːŋ finals displaying varying alveolarization

In casual conversation the informant pronounced the bold and highlighted characters with [-œːn].

	Tone + Final -œːŋ → œːn					
Onset	**High Level ˥**	**Mid Rising ˧˥**	**Mid Level ˧**	**Low Level ˨**	**Low Falling ˩˧**	**Low Rising ˨˧**
k	**疆 姜 薑**					
k^h					強	
ts	張 將 彰 樟 章	長 掌 獎	障 仗 漲	丈 杖		
ts^h	槍 窗 猖 昌 鎗	**搶**	唱 倡 暢		**塲** 長 詳 牆	
h	**香 鄉**	**響 享 向**	**向 嚮**			
s	相 箱 霜 傷 商	想 賞	相	上 尚	嘗 常 裳	上
j	央	快 鞅		樣 **讓** 嚷	羊 洋 楊 陽	仰 養 氧
l	倆			量 亮 諒	量 良 娘 涼 糧 梁	兩 輛

(from Bauer 1979:138, Table IV)

Table 26. Lexical items with -ɔːŋ finals displaying varying extent of alveolarization

In casual conversation the informant pronounced the bold and highlighted characters with [-ɔːn].

	Tone + Final -ɔːŋ →-ɔːn					
Onset	**High Level** ˥	**Mid Rising** ˧˥	**Mid Level** ˧	**Low Level** ˩	**Low Falling** ˩˧	**Low Rising** ˨˩
p	幫 邦	榜		磅 傍		
p^h					旁	
t	當	黨	檔	蕩	盪	
t^h	湯	躺	趟 燙		堂 糖 唐	
k	缸 綱 江 剛	港 **講**	鋼			
k^h			抗			
ts	莊 裝		壯	狀 撞		
ts^h	鶬	闖	創		藏 **牀**	
m				**望**	忘 亡 芒 荒	
ŋ					昂	
h	腔			項 巷	**航 行**	
s	桑	爽				
f	**方** 荒	**房** 訪	**放**		妨 防	
w	汪	枉		旺	**王 黃**	**往**
l					琅 郎 狼 浪	朗

(from in Bauer 1979:139, Table V)

The data shows no clear distribution within each of the final categories (**-a**ːŋ, **-ɐŋ**, **-ɛ**ːŋ, **-œ**ːŋ, **-ɔ**ːŋ) that points to phonological conditioning by the onset consonant or tone. In addition, with some onset consonant and tone combinations, there were lexical items which were alveolarized while other lexical items were not. This led Bauer to suggest that alveolarization progressed in this speaker via lexical diffusion. In summary, by placing the rimes on a continuum from 'never' to 'occasionally', 'almost always' and 'always' alveolarized, Bauer's findings can be represented as follows:

Table 27. Alveolarization of -ŋ by preceding vowel

Occurrence of -ŋ → [-n]	Never	Occasionally	Almost always	Always
	←			→
Finals undergoing alveolarization	**-ɪŋ** **-ʊŋ**	**-ɔːŋ** **-œːŋ**	**-ɛːŋ**	**-aːŋ** **-ɐŋ**

(based on description by Bauer 1979:133)

Bauer argued that alveolarization started with the **-a**:**ŋ** and **-ɐŋ** rimes, since there was no variation and these rimes were always alveolarized in his informant's speech. On the other hand, variation was still occurring in the other finals **-ɛ**:**ŋ**, **-ɔ**:**ŋ**, and **-œ**:**ŋ**, an indication that change might still be underway (Bauer, 1979: 134). He therefore postulated that alveolarization of final **-ŋ** spread forwards and upwards in preceding vocalic space and was transmitted through lexical diffusion. While this study was one of the first to document that the velar consonants were becoming alveolarized in Hong Kong Cantonese, it was not clear from this study how widespread the phenomenon was, or indeed, how this change was spreading though the speech community, as the study was based on just one informant reading out a list of words.

Further studies have described alveolarization within the literature on Cantonese sound change. In Zee (1996), the phonetic variation by a 21-year-old male speaker reading all the monosyllables with syllable-final oral and nasal stops in Hashimoto's (1972) syllabary were investigated using electro-palatography. Similar to Bauer's study, it was found that alveolarization from **-ŋ** to [-n] occurred in **-a**:**ŋ** and **-ɐŋ**, although Zee reported that it occurred only some of the time.

Wong (2005) found that overall (averaged across different speech styles, including formal speech, conscious speech, and casual speech) 34.19% of occurrences

of **-a**:**ŋ** were produced with the innovative alveolar variant [-a:n]; 40.83% of **-ɐŋ** as [-ɐn]; 16% **-ɛ**:**ŋ** as [-ɛ:n]; 14.17% **-œ**:**ŋ** as [-œ:n]; and 11.75% of **-ɔ**:**ŋ** as [-ɔ:n]. The use of [-ɪn] for **-ɪŋ** was very low, at 1.2%, while no production of [-ʊn] for **-ʊŋ** was found.

2.5 Phonological conditioning of **-k** to [-t] variation

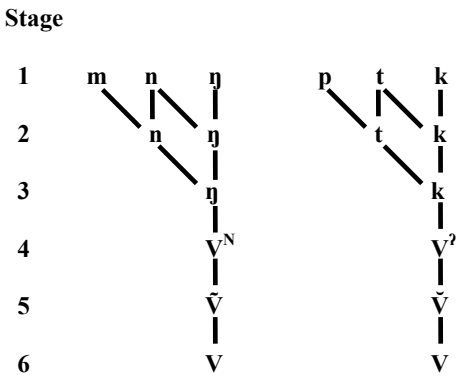
In Zee's (1996) account of phonological conditioning of **-k** to [-t], alveolarization was found in the rimes -**a**:**k** and **-ɐk**. On the other hand, he also found instances of *velarization* in the rimes -**ɐt**, **-ɔ**:**n** and **-ɔ**:**t**, with the latter rime having undergone complete velarization to [-ɔ:k]. The use of [-t] for **-k** seemed to be present only when the preceding phonological environment was **a**: or **ɐ**. In this data and similar to results for the velar nasals, Zee also found instances of velarization of **-a**:**t** and **-ɔ**:**t** to [-a:k] and [-ɔ:k]. He described that the **-a**:**t** rime was split into [-a:t] and [-a:k], while the **-a**:**k** rime split into [-a:k] and [-a:t] (1996:197). Such "confusion" (p. 197) of the words with **-a**:**k** and **-a**:**t** seemed to Zee to point towards the distinction between the coda stop consonants **-k** and **-t** having been lost altogether in this speaker.

2.6 Cross-dialectal motivation for the study of sound changes in Cantonese coda consonants

Being one of the very few Chinese dialects which still preserves the full set of Middle Chinese syllable-final consonants, namely the nasals ***-m**, ***-n**, ***-ŋ** and oral stops ***-p**, ***-t**, ***-k**, Hong Kong Cantonese offers the opportunity to study conditions under which final consonants might change and for these findings to be extended to other Chinese dialects, to obtain a better understanding of how the Middle Chinese coda consonants might have changed or undergone attrition over time. Among the southern Chinese

dialects, there is a great variety in the extent to which the Middle Chinese coda consonants have been preserved. Similar to Cantonese (or the Yue dialect group as a whole), Hakka and Southern Min have also preserved the full set of stop endings ***-p**, ***-t**, and ***-k** from Middle Chinese (Yue 2002:218). The Gan dialects only have two of those stops remaining, while the Wu dialects, the Jin dialects, and the Jianghai dialects of Mandarin only have a single glottal stop; on the other hand, Northern dialects have altogether lost the final stops and instead only have tonal features in their place (*ibid.*). For a full survey of comparisons of stops between dialects, see Yue (2002).

Comparing the finals in Middle Chinese and a selection of modern Chinese dialects, Chen (1972a) found that the development of final consonants in modern Chinese dialects can be generalised to consist of six stages. The diagram in Figure 10 shows the proposed stages of development.



from Chen & Wang (1975: 267)

Figure 10. Development of final nasal and stop consonants in Chinese dialects

Chen suggested that different modern Chinese dialects were at different stages of development, with stops in general having progressed at a much faster rate than the

nasals. In Stage 1, all Middle Chinese nasal consonants ***-m**, ***-n**, ***-ŋ** and stop consonants ***-p**, ***-t**, ***-k** were still present, but throughout Stages 2 and 3, consonantal endings merged, reducing the overall number of final consonants. In Stage 4, the phonetic features of the consonants reduced from Stage 3 left a phonetic trace, while in Stage 5 nasality and vowel length contrast were phonemicized after the final consonants were effaced. Finally, in Stage 6, even these features were lost, with the nuclear vowel being all that remained. Table 28 illustrates how these developments progressed in Peking Mandarin (the official *lingua franca* in mainland China) and in Guangzhou Standard Cantonese, along with other modern dialects.

Table 28. Cross-dialectal comparison of nasal and stop coda attrition from Middle Chinese

	Nasal codas at stage	**Stop codas at stage**	**Chinese Dialects**
A	1	1	Guǎngzhōu, Méixiàn, Xiàmén (lit.)
	1,5	4,6	Xiàmén (col.)
B	2	2	Nánchāng (Cháozhōu, lit.)
	2,5	4	Yángzhōu (Cháozhōu, col.)
	2,6	4	Sūzhōu
	2	6	Peking, Hànkǒu, Chéngdū
	2,5	6	Shuāngfēng, Chángshā
C	3	4	Fúzhōu
	3,5,6	4	Tàiyuán, Shànghǎi
	3-6	5	Chángshú
	3,5	6	Jìnán
	3,6	6	Wēnzhōu
D	5	6	Jǐnggǔ, Shuāngjiāng
	6	6	Fěngyí, Lìjiāng

(Chen & Wang 1975: 267, after Chen 1972)

Mandarin is in Stage 2 with respect to the nasals, as it has lost the final ***-m** ending. Mandarin stops, on the other hand, have progressed as far as Stage 6, since all

final stops have been lost. Compare this result with Standard Cantonese, which is in Stage 1 because it has still preserved all nasal as well as stop consonants from Middle Chinese. The Lijiang dialect in Band D, is at the most progressive end of the spectrum and has lost all final consonants as well as nasalization and length contrasts. As can be seen, the overall cross-dialectal trend is such that the final consonants disappeared over time. Furthermore, Chen & Wang 1975 believed that the different rates at which the nasal and stop consonants underwent attrition can be explained in terms of perceptual differences between the two types of finals. However, pertinent questions to be asked are: in so far as the above generalization holds, and given that standard Cantonese is at Stage 1 now, will it follow in the footsteps of the other Chinese dialects and display mergers of all its final consonants and ultimately lose them all? It is hoped that studying variation patterns in the current use of consonantal finals in Hong Kong Cantonese will at least offer insight into the status of merger and attrition processes which have hitherto taken place.

2.7 Historical development of Yue stop consonants from Middle Chinese

In her description of the parallel development of stop endings ***-p**, ***-t**, ***-k** in Yue dialects, including Hong Kong Cantonese, Yue[24] (2002) has found patterns which suggested that the direction of change of final stop consonants is determined by the

[24] Yue writing on Yue. Now, this is confusing... But this hapchance of orthographic identity is only present in the *Romanization/Pinyin spelling* of the author's name and the name for the southern Chinese language family, whereas in Chinese orthography they are represented by two completely different characters as well as pronounced differently in Cantonese. The author's name is 余(靄芹) **jy**:˩ (**ɔ:i˥ kʰɐn** ˩), whereas the dialect is 粵 **jy**:**t**˩. Two utterly different things.

height and tenseness of the preceding vowel and tone on the syllable.[25] She proposed that the historical change across dialects from bilabial ***-p** or velar ***-k** to alveolar **-t** seemed to be due to assimilation of tongue height after a close vowel (2002:241), although Cantonese and Lechang dialects are obvious exceptions. For an illustration of the differences across dialects, refer to Table 29.

[25] Moreover, she also suggested that the rate at which these changes transmitted through the various dialects seemed different and that the changes were transmitted from one lexical item to another, a finding which will be explored further in the ensuing discussion on lexical diffusion.

Table 29. Historical development of Middle Chinese *-k to -t in Yue dialects

	Hong Kong	**Pan-yu**	**Shun-de**	**He-chu**	**Tai-sham**	**Xin-xing**	**Lian-zhou**	**Hua-zhou**	**Wu-yang**	**Le-chang**
逼	pɪk	pet	pet	pet		pɐt	pɐt	bet		pek
力	lɪk	let	let	let		lɐt	lɐt	let	lɛt	lek
息	sɪk	set	set	*sek*		sɐt	sɐt	*łek*	łɛt	sek
直	tʃɪk	tset	tset	*tsɛk*		tsɐt	ts'ɐt	tʃ'et	ts'ɛt	tsek
色	sɪk	set	set	*sek*		sɐt		ʃet		sek
食	sɪk	set	set	*sek*		sɐt	*sɐt	ʃet	sɛt	sek
極	kɪk	ket	ket	kɛt		kɐt	k'ɐt	*k'ek*		kek
憶	jɪk	(j)iet	iet	*jɛk*		ɐt				iek
億	jɪk	(j)iet	iet			ɐt	*jɐt			iek
碧	pɪk	pet	pet	*pɛk*		pɐt		*bek*		pek
逆	jɪk	(j)iet	iet		ŋgɛt	ŋɐt	nɐt	*ŋek*		iek
璧	pɪk	pet	pet	pɛt		pɐt		*bek*		pek
席	tʃɪk	tset	tset			tsɐt	ts'ɐt	*t'ek*	t'ɛt	tsek
夕	tʃɪk	tset	tset	*tsek*		tsɐt	łɐt	*t'ek*	ts'ɛt	tsek
尺	tʃ'ɛk	*ts'iak*	*ts'ɛk*	*ts'iak*		ts'ɐt	ts'ɐt	tʃ'et		ts'ek,t'ɛt
石	sɛk	set	set			sɐt	sɐt	ʃet	sɛt	sek
亦	jɪk	(j)iet	iet	*jɛk*	jɛt	(j)iɐt		jet		iek

(Yue 2002:239)

The first column of Table 29 displays some Chinese characters with ***-k** endings in Middle Chinese. The remaining columns are transcriptions of contemporary

pronunciations in a range of Yue dialects, from Hong Kong Cantonese, to Panyu, Shunde and Lechang dialects, amongst others. With the exception of Hong Kong Cantonese (second column) and Lechang (final column), Middle Chinese ***-k** displayed a tendency to develop into [-t] in the context of high preceding vowels across Yue dialects. Contrary to this trend, Hong Kong Cantonese and Lechang dialects have preserved the ***-k** finals completely in the context of preceding high vowel.

The Lechang dialect displayed another peculiarity: lower back vowel **ɔ**: in addition to the high **ɪ** vowel seemed to favour the change from ***-k** to **-t**, a tendency not found in other Yue dialects (see Table 30). In contrast, in standard descriptions, the velar coda stop **-k** is still considered to have been maintained within Hong Kong Cantonese after the low back **ɔ**: vowel. But recall from Section 2.2 above, that this very rime, **-ɔ**:**k**, is suggested to be involved in ongoing alveolarization in Hong Kong Cantonese. Therefore, studying the extent of variation and degree of change of the **-ɔ**:**k** rime in present-day Hong Kong Cantonese might reveal to us the kinds of linguistic factors which might have influenced the historical change ***-k** to **-t** within the Lechang dialect.

Table 30. Lechang dialect showing *-k to -t, in contrast with Hong Kong Cantonese

	Hong Kong	Lechang
博	**pɔ:k**	**pɔt^{3}**
霍	**fɔ:k**	**fɔt^{3}**
莫	**mɔ:k**	**mɔt^{2}**
剝	**mɔ:k**	**mɔt^{5}**
托	**t'ɔ:k**	**t'ɔt^{3}**
雹	**pɔ:k**	**pɔt^{2}**
索	**sɔ:k**	**sɔt^{3}**
學	**hɔ:k**	**hɔt^{2}**
鶴	**hɔ:k**	**hɔt^{3}**

Yue (2002:238)

On the other hand, Yue suggested that the reverse change from alveolar ***-t** to bilabial **-p** and velar **-k** across dialects was motivated by assimilation of the coda to low vowels. This tendency is illustrated in the examples from various Yue dialects below.

Table 31. Change from Middle Chinese *-t to -k in Yue dialects

	Proto-Yue	**Hong Kong**	陽春 **Yangchun**	**Dongguan 1987**	增城 **Zengcheng**	**Xinjie**	吳川 **Wuchuan**	吳陽 **Wuyang**
	*at	ɐt	ɐk	ɐk	ɐk	ək	ak	ɐk
	*uat,*yat	uɐt,øt	ɐk	uɐk/ɐk(ɐt)	uɐk/ɐk	uək	?	?
Examples:								
'pen'	筆	pɐt	pɐk	(see below)	pɐk	pək	bak	bɐk
'dense'	密	mɐt	mɐk		mɐk	mək		mɐk
'chestnut'	栗	lɐt	lɐk		lɐk	lək		
'seven'	七	tʃ‿ɐt	ts'ɐk		ts'ɐk	ts'ək	t'ak	t'ɐk
'nephew'	侄	tʃɐt	tsɐk		tsɐk	tsək		
'louse'	蝨	sɐt	sɐk		sɐk	sək		
'solid'	實	sɐt	sɐk		sɐk	sək		
'day'	日	jɐt	jɐk		(j)iɐk	(j)iək	ʒak	ŋiɐk
'auspicious'	吉	kɐt	kɐk		kɐk	kək		kɐk
'one'	一	jɐt	jɐk		(j)iɐk	(j)iək	jak,*jat*	
'beg'	乞	hɐt	hɐk		hɐk	hæk		
'pawn'	卒	tʃøt			tsɐk	tsək		tɐk
'bone'	骨	kuɐt	kɐk		kuɐk	kuək		
'cave'	窟	kuɐt,fɐt	fɐk		k'uɐk	k'uək		

(Yue 2002: 229)

However, as Table 31 shows, Hong Kong Cantonese presents another anomalous case. In contrast to other dialects, which seemed to develop into **-k** due to assimilation to a preceding low vowel, Hong Kong Cantonese has preserved the Middle Chinese alveolar coda stop ***-t**, even within the same context of a low vowel. What is particularly interesting here is that these are exactly the low vowels which seem to condition contemporary alveolarization in Hong Kong Cantonese. So it seems that Hong Kong Cantonese has not only preserved the alveolar ***-t** from Middle Chinese, but that, furthermore, there is a progressive move towards further alveolarization within the context of the low preceding vowel **ɐ**. One has to bear in mind, however, that the above

data gathered by Anne Yue are from different sources and different time periods spanning from the early 1900s to the present day (e.g. Chao's field notes dating back to 1929, as well as Yue's own, and more contemporary accounts by Zhan & Cheung 1987; for details of all data sources, see Yue 2002:241–2). Nonetheless, if we are to assume that the above description is a representative synchronic description of final endings in contemporary dialects, the deviation of Hong Kong Cantonese and Lechang from other dialects constitutes a conundrum which needs to be explained.

What concerns us most in this discussion is that a preceding high vowel seems to have historically favoured the alveolarization of Middle Chinese ***-k** to **-t** across Yue dialects, whereas the same change from **-k** to [**-**t] in contemporary Cantonese seems to operate not on syllables with a high nuclear vowel, but rather on the low and mid ones. Lechang, however, seems to be an exception, and showed a closer conformity to the contemporary trend of **-k** alveolarization in Cantonese being conditioned by lower rather than higher vowels. Based on the above findings from a cross-dialectal analysis of the historical development of stop codas, this thesis will examine the proposal that tongue height could be a determining factor for alveolarization of velar codas in Hong Kong Cantonese, the results of which will be presented in Chapter 6.

2.8 Summary

This chapter described previous studies which have suggested that alveolarization is a sound change in progress in Hong Kong Cantonese and examined the age-based variation found in Shum (1993) and Wong (2005). This was placed within the wider context of sound changes that have been found to be occurring in Hong Kong

Cantonese. A basic description of Cantonese phonology was also offered: phonological patterning of initials and finals were examined, and co-occurrence restrictions were described, such that a more detailed description of the phonological conditioning operating on the alveolarization of **-ŋ** and **-k** could be offered. A cross-dialectal, historical linguistic motivation for studying the development of Hong Kong Cantonese velar coda stops was also presented. The following chapter will describe the methodology that has been adopted in order to investigate the questions posed within the preceding chapters.

3 Methodology

The data for this thesis was collected through sociolinguistic interviews conducted in 2010, for which ethics approval was obtained in advance from the Central University Research Ethics Committee of the University of Oxford. This chapter details the methodological considerations which went into the design of this study.

3.1 Sampling

Strictly speaking, in most—if not all—linguistic studies, external validity (whether results from an experiment can be generalised to the whole population that the study makes a claim about) can be called into question. The reasons are that sampling is hardly ever performed on a fully random basis, sample sizes are typically very small compared to the whole population of speakers of a language under investigation, and that there is a wide range of variation which can typically be found within a speech community. For instance, the snowballing technique (Milroy & Gordon 2003:32), where a friend of a friend is introduced to take part in the study, is fairly commonly employed within sociolinguistics, and it is obvious how this introduces a systematic bias in sampling. However, the other possibility, a fully randomised study, might cause the researcher to run into difficulty when aiming to obtain naturalistic data: involvement in a sociolinguistic interview requires substantial time commitment, and without direct or indirect introductions to the researcher, informants are typically reluctant to get involved in elaborate sociolinguistic studies; if they do respond to an approach by a stranger, it is questionable whether they will trust this person enough to feel at ease when conducting the interview and produce sufficiently naturalistic speech. Given the

aim to elicit speech which is as natural as possible, using the 'friend of a friend' sampling approach is still the lesser of the evils and is a more practical solution. Furthermore, the typically substantially larger sample sizes in social sciences are often unattainable due to the labour-intensive analyses that are needed in linguistic research. In the ideal world, and if the researcher had an unlimited amount of time to perform the sampling and location of subjects for the sociolinguistic interview, a fully random sampling method with large sample sizes would be preferable, but what a researcher is faced with is the real world.

That said, in typical sociolinguistic studies, ecological validity is higher than that obtained in more formal linguistic studies. Given that it is *speakers* who use language on a daily basis and that sound change occurs to this vehicle for communication throughout such daily use, the very aim of variationist sociolinguists to study naturalistic speech by speakers in their environment proper—social situations where speakers interact with one another—means that sociolinguistic studies are in a better position to make claims about the mechanisms of sound changes. This is a primary advantage to using a sociovariationist approach, and is perhaps the only way in which one can assess neogrammarian and lexical diffusion models of sound change.

3.2 Informants

This study used data from 30 speakers. All informants for this study were speakers of Hong Kong Cantonese who were born in Hong Kong or who have lived in Hong Kong since they were below 10 years of age. There is ample evidence that before the age of 10, speakers have no difficulty acquiring local features of a language, but that beyond the age of 10, the success with which a speaker acquires native-speaker proficiency

drops rapidly (e.g. Payne 1976; Kerswill & Williams 2000). While the influence that adult immigrants might have on sound change in Hong Kong Cantonese is an interesting question in itself, in order to avoid introducing a further variable, such speakers were excluded from the sample. The speakers were anonymised according to the following identity (ID) coding practice:

Age Group[26]:	Y(oung) | M(iddle) | O(ld)
Gender:	M(ale) | F(emale)
Age:	[Numeral]
Disambiguating code:	A, B

such that 'YF14A' means 'young female aged 14', with an 'A' at the end used to disambiguate between two informants in the data collection with the same age and gender. The informations will be referred to only by their ID code to preserve their anonymity.

The utmost care was taken to uphold the University's research ethics standards. Procedures for data collection and safeguarding complied with the University of Oxford's policy on research involving human participants and personal data, and ethics approval was obtained in advance from the Central University Research Ethics Committee (CUREC). All informants were informed of the purpose of the research and were encouraged to ask questions before and during the data collection and were free to withdraw from the research project at any time. They were asked to sign a consent form

[26] Age groups are merely coded to offer a comparison with age-based results from previous studies (Shum 1993 and Wong 2005). The age factor is included as a continuous measure in the multinomial logistic regression model presented in Chapter 4. See discussion in Sections 3.4 and 5.1 for further details on the methodology employed for and interpretation of age-based variation.

in advance to confirm they understood these details. Informants below the age of 18 also required parental consent in addition before they could take part.

3.3 Recording set-up

The data was recorded using lavalier microphones (Audio Technica AT803 omnidirectional condenser type, operating on phantom power provided by the recorder) and Marantz PMD660 solid state recorder at 44.1 kHz using uncompressed 16-bit PCM format. The use of clip-on lavalier type microphones meant that the equipment set-up was more likely to promote casual speech and minimize distraction from holding a conversation, when compared to recording set-ups which have more prominent microphone placements and require informants to be within a certain distance range to the microphone.

The casual speech interviews were conducted in pairs, and the informants were grouped into pairs based on acquaintance, in order to elicit as naturalistic speech as possible. Insofar as it was possible, the speakers were paired up with someone in their age group. All school children were paired with a friend in their age group. Only in a couple of instances did the pairing involve informants from different age groups, but they were all paired with people who were familiar to them to promote naturalness of their speech within the interview. The data collection was conducted in a school classroom for the 14-18 age group, while for other age groups, the interviews were conducted in a home setting.

3.4 Social variables

Age, gender and speech style are the social variables which are investigated in this work

to explore how they correlate with variation patterns in the linguistic variables. The age variable merits more in-depth discussion here, especially the considerations to be made when sampling informants across age groups, and whether to treat age as a categorical variable or as a continuous variable in the statistical analysis.

As reviewed in Section 2.1, sociovariationist studies of alveolarization such as Shum (1993) and Wong (2005) have used different age groupings to investigate age-based variation. Shum (1993) studied 40 speakers across two gender and two age groups: younger (aged 15-20); and older (aged 50-55), whereas Wong (2005) studied 40 speakers balanced across gender and four age groups: 11-20, 21-30, 31-40, and 41 years of age or above. The motivation for using these age groupings was not made clear in these studies; however, they concur in that speakers in the youngest age group were up to 20 years of age.

In this work, an important sampling criterion was that informants should be relatively balanced and distributed over a range of different ages to allow us to observe any age-based variation. As Eckert (1997) suggests, "age has significance because the individual's place in society, the community, and the family changes through time" (Eckert 1997:155; as quoted in Milroy & Gordon 2003:39). When selecting informants, an effort was made to include speakers who were 18 years of age or below. This is typically the age where high school students progress to either a tertiary institution or enter into the labour force, and it was deemed important to include enough speakers in the sample that fall within the school age range. Further sampling efforts were made to include enough speakers within the "middle" age range beyond the school age range, from aged 19 to 45, as well as within the "older" age range of 45 or beyond. The resulting age distribution of the sampled informants can be seen in Figure 11.

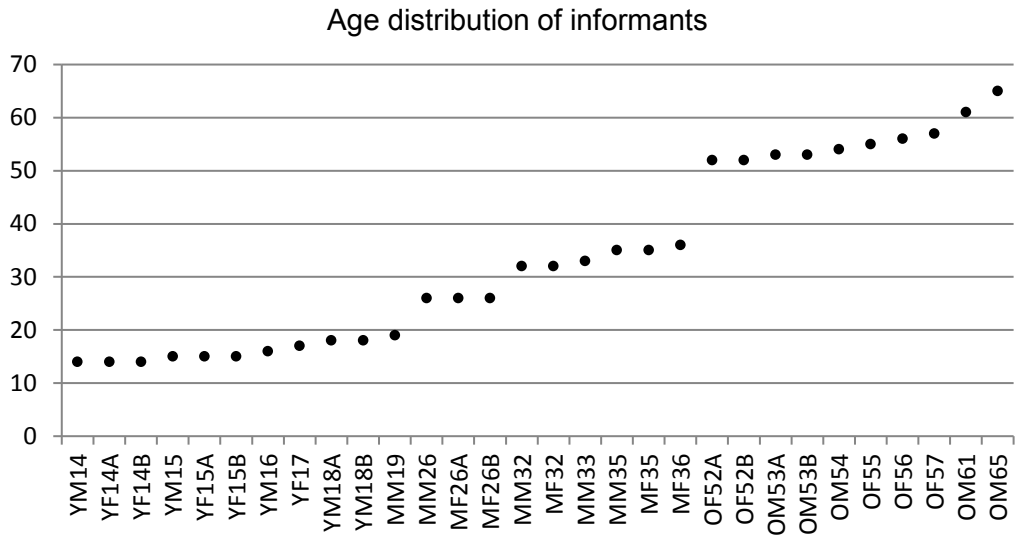


Figure 11. Informants' age distribution

Despite our foregoing discussion in Section 2.2 of the possibility that some of the sound changes in Hong Kong Cantonese have engendered strong opinions that they are a sign of "lazy pronunciation", without any empirical evidence there is no *a priori* reason to treat alveolarization as a sound change that is either above or below the level of speakers' conscious awareness. It has been suggested (e.g. Labov 1994) that sound changes that operate above the level of conscious awareness are closely linked to life stage (for instance, a teenager going to a high school might be more likely to use the most innovative variants in a sound change, whereas someone starting a professional career might make more use of conservative variants). This phenomenon is also termed "age-grading". On the other hand, sound changes which operate below the level of conscious awareness are less likely to be linked to life stages of the speakers. The reader is referred to Chapter 5 for a more detailed discussion of results with respect to the age variable.

Thus the "rule of thumb" as described above with respect to recruiting informants across different age ranges, was mostly to ensure a relatively balanced distribution of the data across age, rather than to set up any *a priori* expectation that the above rough age groupings used for sampling would necessarily correlate with variation in alveolarization.

For statistical analysis purposes, there is a clear advantage of using age as a continuous, rather than a categorical variable. As will be described in more detail in Section 3.9, using a continuous measure whenever possible is more preferable over using a categorical variable, as the latter introduces more parameters to be estimated and reduces the statistical power of the analysis. However, to offer a comparison with previous studies such as Shum (1993) and Wong (2005), which have stratified their sample of speakers according to age groups in their analysis of the age variable and have only present the aggregated statistics for each age group, the above-mentioned age groupings employed for recruiting informants will be used mainly for exploratory purposes.

As gender has classically been found to be a main social factor determining linguistic variation patterns in previous sociolinguistic studies, this factor was therefore included in this investigation of Hong Kong Cantonese variation. Socioeconomic class has also been shown to be a significant factor in predicting who the innovators are of sound change, especially in studies performed in North America. More specifically, plotting Varbrul factor weight against socioeconomic class has yielded findings which conform to the curvilinear hypothesis suggesting that innovations do not start at the peripheral socioeconomic classes, but rather from the ones in the middle, in particular, upper working class and lower middle class (Labov 2001:32–33; Cedergren 1973).

Within the context of Hong Kong, however, socioeconomic class (with the exception of education level) has been found to be relatively insignificant in accounting for use of innovative variants (e.g. Bauer 1983). Due to the limited resources which could be allocated to this doctoral dissertation and on the basis of findings in previous studies that speech patterns in Hong Kong Cantonese show less class-based differences, socioeconomic class was not included as a social factor in this work. However, all subjects within this study had at least completed primary school education. It was infeasible to have a higher threshold for educational background, as many speakers within the older generation in Hong Kong typically did not complete all of their secondary school education, and excluding older speakers on the basis of education would mean the picture of older speakers' speech characteristics would be less representative of what is the norm within the speech community.

3.5 Speech styles

Speech style has consistently been shown to be an important indicator of how the speech behaviour within a speech community develops in a situation of sound change in progress. In particular, speech style shifting is a good indication of how much social awareness speakers display towards innovative variants (Labov 2001:196). In this study, the sociolinguistic interview consisted of the following components, to extract different speech styles: casual speech, reading passages, word lists and minimal pair reading. These are described in more detail in the sections below. The whole interview lasted from about 45 minutes to 1 hour. About 20-40 minutes were devoted to the casual speech interview, and another five to ten minutes for each of the formal reading tasks.

3.5.1 Casual speech elicitation

As described briefly above, the sociolinguistic interviews were conducted in friendship or family pairs to facilitate more natural speech. The sociolinguistic interview was driven by a set of questions adapted from Tagliamonte (2006) designed to promote casual speech interaction. Questions about basic information (e.g. their age, their relationship to one another, where they lived, how long they have lived in Hong Kong, and what language(s) they spoke at home) were first asked. For some pairs, such elementary questions were sufficient to keep the conversation going, with informants bringing up more stories and anecdotes, but in other cases, it was necessary to prompt informant pairs using further questions adapted from Tagliamonte (2006), in Appendix I, to keep the conversation alive. The questions had to be adapted for important cultural reasons. While Labov and other sociolinguists quote the near-death experience as a topic which typically very successfully elicited casual and unmonitored speech in previous sociolinguistic studies, death is a social taboo which Hong Kongers generally feel very uncomfortable mentioning in daily talk, and asking them about a near-death experience might be construed by some to be a personal curse. Therefore, adaption of the near-death questions was needed to avoid inauspicious talk of death. Instead, it was found in this study that—as it is not uncommon for Hong Kong people to subscribe to superstitious ideas and to believe in the afterlife—talking about ghosts and whether informants believed in them or had seen any, achieved similar success. The adapted list of elicitation questions is included in Appendix I. Questions were based on various topics in daily life ranging from hobbies, relationship with parents, festivals and traditions, to any unusual experiences they may have had, such as witnessing an accident. Examples of questions included:

- *What are your favourite hobbies, and how much time do you spend on them?*
- *What is your favourite festival or holiday, and why? What do you do with your family or friends during this festival?*
- *Have you ever witnessed any accident?*
- *Do you believe in ghosts? Have you ever seen any?*

These questions were in general quite successful at engaging speakers to speak casually. Roughly 20-40 minutes of casual speech was elicited from each speaker pair.

3.5.2 Formal tasks

The formal tasks, which included passage-reading, word list reading, and minimal pair reading tasks, were designed to elicit tokens of interest in more formal speech settings. They were administered in this order so that the speech styles would be progressively more formal. After the informal tasks ended, informant pairs were split up and asked to perform the formal tasks separately. The informant who was not engaged in the formal tasks was asked to relax and read some magazines in a spot out of ear-shot from the informant who underwent the formal task.

3.5.2.1 Passage reading

The overarching aim of the passage reading task was to elicit a more formal speech style from each participant to study speech style effects but it was also set up to elicit tokens in order to test our hypotheses on the effects of phonological and lexical frequency effects. Overall, the passage contained 399 words (refer to Appendix II for the full passage). The aim was to elicit tokens of each of the final codas (**ŋ**) and (**k**) which met the following conditions: 7 preceding phonological contexts, 2 phonological contrastiveness conditions, and 2 lexical frequency conditions. Overall, 7 x 2 x 2

conditions. As mentioned in Section 2.3.4, there are seven nuclear vowels which can precede the velar codas: the long vowels **a**:, **œ**:, **ɔ**:, and **ɛ**:, and short vowels **ɐ**, **ʊ**, and **ɪ**. Recall also from Section 2.3.4 that certain rimes, when having undergone alveolarization, would create a loss of phonological contrast, while alveolarization of other rimes would not have the same effect. These constitute the 2 phonological contrastiveness conditions. The methodological design of the lexical frequency conditions will be described in detail in Section 3.6 on lexical frequency, although it will be mentioned here briefly, that there were two conditions: 'frequent' and 'infrequent', based on data from the Hong Kong Cantonese Adult Corpus (Leung, Law & Fung 2001). This was to ensure that not only frequent items were included in the formal tasks, but also a sufficient number of infrequent lexical items could be included, to allow the effects of infrequent lexemes to be analysed[27]. The median lexical frequency cut-off point of 7 was used, such that lexemes which appeared 7 times or more in the HKCAC were considered 'frequent', and those which appeared less than 7 times were considered 'infrequent'. Thus, within each preceding phonological vowel condition, there were four categories: phonological contrast exists, frequent (PHONCON: YES, FREQUENT), phonological contrast exists, infrequent (PHONCON: YES, INFREQUENT), phonological contrast exists, frequent (PHONCON: NO, FREQUENT), and finally, no phonological contrast exists, infrequent (PHONCON: NO, INFREQUENT).

Table 32 summarises these categories. Where 'n/a' is indicated, no morphosyllable exists in standard Cantonese with a corresponding alveolar rime. For instance, the alveolar rimes **-œ**:**n** and **-ɛ**:**n**, except in a very limited set of colloquial or

[27] Note: this is categorization of lexemes into 'frequent' vs. 'infrequent' was only used for deciding which lexemes to include in the formal tasks; in the final statistical analysis, lexical frequency was used as a continuous measure.

onomatopoeic morphosyllables, do not occur in descriptions of Standard Cantonese. The cells indicate the number of tokens present in the reading passage. For the **-ɐk** rime, a character corresponding to the 'PHONCON: NO, FREQUENT' condition could not be found so this cell was also empty.

Table 32. Token distribution in formal reading passage

	PHONCON*: YES		PHONCON: NO	
	FREQUENT^	INFREQUENT#	FREQUENT	INFREQUENT
-aːŋ	1	1	1	1
-œːŋ	n/a+	n/a	7	1
-ɔːŋ	2	1	2	2
-ɛːŋ	n/a	n/a	2	1
-ɐŋ	2	1	1	1
-ʊŋ	n/a	n/a	1	2
-ɪŋ	n/a	n/a	2	1
-aːk	5	1	1	1
-œːk	n/a	n/a	1	1
-ɔːk	2	2	1	1
-ɛːk	n/a	n/a	0	1
-ɐk	2	1	0	1
-ʊk	n/a	n/a	1	1
-ɪk	n/a	n/a	1	1

* PHONCON: YES: a phonologically contrastive lexeme exists, i.e. homophony to another morphosyllable occurs after alveolarization of velar stop
PHONCON: NO: a phonologically contrastive lexeme does not exist

^ FREQUENT: lexical frequency of 7 or more

\# INFREQUENT: lexical frequency of less than 7

\+ n/a: indicates that no homophonous pair exists due to systematic gap

3.5.2.2 Word list reading task

The word list contained 196 items, each of which included one or more tokens of target characters. The full list is presented in Appendix III. The token numbers for each category are included in brackets in Table 33, in a similar way of presentation as above.

Table 33. Token distribution in word list reading task

	PHONCON*: YES		PHONCON: NO	
	FREQUENT^	INFREQUENT#	FREQUENT^	INFREQUENT#
-aːŋ	8	9	5	6
-œːŋ	n/a+	n/a	11	9
-ɔːŋ	9	6	11	7
-ɛːŋ	n/a	n/a	11	5
-ɐŋ	9	8	6	2
-ʊŋ	n/a	n/a	11	7
-ɪŋ	n/a	n/a	9	7
-aːk	7	6	8	6
-œːk	n/a	n/a	6	6
-ɔːk	6	7	7	7
-ɛːk	n/a	n/a	5	6
-ɐk	6	6	1	6
-ʊk	n/a	n/a	8	7
-ɪk	n/a	n/a	6	6

* PHONCON: YES: a phonologically contrastive lexeme exists, i.e. homophony to another morphosyllable occurs after alveolarization of velar stop
PHONCON: NO: a phonologically contrastive lexeme does not exist

^ FREQUENT: lexical frequency of 7 or more

\# INFREQUENT: lexical frequency of less than 7

\+ n/a: indicates that no homophonous pair exists due to systematic gap

Morphosyllables were put in carrier phrases, which were either a lexical combination of two or more characters, or a short phrase. The key was to find carriers which contained characters which were fairly common, and which would likely come up in daily life, so that 1) the subjects could read it out loud, and 2) the linguistic variables under investigation could be investigated within the kinds of environments that speakers can encounter in their daily lives as these words were not too far removed from their daily experience of the language. The morphosyllable being investigated was put in different phonological environments, and particularly, was either followed by a word boundary, or by a non-assimilating onset consonant, such as bilabials **b-**, **m-**, **p-**, labiodentals **f-**, **w-**

or glottal **h-**. That said, there were some items which were followed by a velar and alveolar, to allow the effects of a following onset which has either a velar or alveolar feature to be estimated.

3.5.2.3 Minimal pairs reading task

The minimal pair reading tasks contained pairs which were minimally contrastive, differing only in the place of articulation of the coda consonant (velar or alveolar), such as in 冷 **la**:**ŋ˧** 'cold' vs. 懶 **la**:**n˧** 'lazy', and 生 **sɐŋ˥** 'birth' vs 新 **sɐn˥** 'new' for (**ŋ**) and 北 **bɐk˥** vs. 不 **bɐt˥** for (**k**). Filler pairs with similar phonological structure but which did not contain the experimental phonemes were added and the items were randomised. There was a total of 96 items, half of which were minimal pairs of interest, with the other half being fillers. The fillers contained minimal pairs of *onset* (as opposed to coda) phonemes, minimal pairs of non-(**ŋ**) or **-**(**k**) coda phonemes, phonologically identical pairs, and phonologically unrelated but (semantically related) pairs. The full list is presented in 0.

3.6 Lexical frequency data

This section describes how lexical frequency data was extracted from a corpus of spoken data. Due to the scarcity of corpora of spoken data in previous times, older studies investigating lexical frequency have used text-based corpora to obtain frequency information. For instance, Hooper (1976) used written frequency data from Dolby and Resnikoff, "The English Word Speculum", which, while including texts from a wide spectrum, still only included word frequency count from written text. Hooper (*ibid*, 98)

argues that while the exact frequency counts might not reflect those of spoken data, the overall trend should still hold, with frequent words in written texts being even more frequent and infrequent words in written text being even more infrequent (in geometric proportions was her guess). Another work on lexical frequency effects, Johnson (1983), also used frequency data from written texts, which was based on a combination of three works: Thorndike and Lorge (1944), Buckingham and Dolch (1936), and Horn (1925) (Johnson 1983:54–5). The first study by Thorndike and Lorge used children's literature to extract frequency information. Johnson took a list of 1000 of the most common words from Thorndike and Lorge (Johnson 1983:71). Buckingham and Dolch (1936) listed the most common words produced by free association, and these were divided by age of the speaker. From this work, Johnson used the word list for completion of fifth grade education in the US (Johnson 1983:71). Horn (1925), on the other hand, contained data from the most common words spoken by children up to six years old. Johnson then determined the frequency as follows: if a given word was absent from all these three sources, they would be considered "uncommon". If a given word was present in all three sources, then it would be labelled "common". Function words were excluded from Johnson's analysis.

Considering that the object of investigation is sound changes occurring in everyday speech contexts, the most relevant lexical frequency data would be those obtained from naturally-occurring everyday speech. With the wider availability of spoken corpora, their use in the examination of lexical frequency effects has become more commonplace. For instance, studies have made use of the Origins of New Zealand English (ONZE) spoken corpus, developed by the Linguistics Department of the University of Canterbury (Gordon, Maclagan & Hay 2007) to investigate lexical

frequency effects in the short front vowel shift in New Zealand English (Hay et al. 2015) and in word-medial intervocalic /t/ in New Zealand English. Pluymaekers, Ernustus & Baayen (2005) have also made use of the Spoken Dutch Corpus (Oostdijk 2000) to investigate lexical frequency effects in acoustic reduction in Dutch.

The lexical frequency data for this study were obtained empirically from the Hong Kong Cantonese Adult Corpus (Leung, Law & Fung 2001, referred to as "HKCAC" hereafter), which is a corpus of natural spoken Cantonese obtained from radio phone-in programmes. The use of word frequency information from the HKCAC approximates frequency counts from real, naturalistic speech.

Methodologically, reading passages and word list tasks were set up so that they contained a fairly balanced number of frequent and infrequent words. The HKCAC contains some 142,295 characters[28], also referred to as tokens in corpus analysis. The frequency data from this corpus was extracted by listing all the unique lexemes, also called *types*, in this corpus and counting the instances of occurrence for each type (the words lexemes and type will be used interchangeably in this section as they amount to the same thing). These counts are then frequency estimates[29] for each lexeme. Suppose one were to make a list of all the individual lexemes, and performed a count of how many times each of those appeared. The resulting list of lexemes by frequency of occurrence within the corpus can be given a *rank* of 1 for the most frequently occurring lexeme, a rank of 2 for the second-most frequent one, and so on. As with any corpus of natural language data, in the HKCAC the most common words showed a high number

[28] The description on the HKCAC webpage states that the corpus contains over 170,000 characters, but the data file available for download only contained 142,295 characters. In this work, the latter figure is taken to be the corpus size.

[29] Technically, the frequency obtained would be *per* 142,295, the total number of tokens in the HKCAC, but this is implicitly understood.

of occurrences, or frequency, with the number of occurrences tailing off rapidly as one moves down towards the list of less common words, with a long tail of words that have only a very small number of occurrences. This lopsided Zipfian distribution of frequency against rank can be gleaned from the following statistical summary of the HKCAC in Table 34. As frequency goes from the maximum down to the 3rd quartile, median and 1st quartile, its value drops rapidly from 6625, to 29, to 7, to 2, with a long tail of characters with a frequency of 1 following behind. Table 35 summarises information about the highest frequency of all (**ŋ**) and (**k**) tokens found in this study (1035 and 2206, respectively).

Table 34. Hong Kong Cantonese Adult Corpus (HKCAC) type frequency summary statistics

	Frequency
Total token count (i.e., total no. of characters)	142 295
maximum	6 625
3rd quartile	29
median	7
1st quartile)	2
Minimum	1
Type count (i.e. total no. of unique characters)	2 150
Lowest rank	1741

Table 35. Type frequency and rank minimum and maximum in this study

Highest rank in the data of this study	1
Lowest rank	1 742*
Highest HKCAC frequency out of all (ŋ) tokens: **kɔːŋ˧** 講 'to talk, speak'	1035
Highest HKCAC frequency out of all (k) tokens: **zɪk˥** 即 'immediate, that is to say'	2206

* As will be explained in Section 3.6.2, adding 1 to the lowest rank (1741) was a way to cope with characters which were found in this study, but which did not occur in the HKCAC corpus.

For a further description of the lexical distribution of the HKCAC, the reader is referred to Appendix VII.

As one of the aims of this dissertation is to investigate the effect of lexical frequency on the likelihood of lenition and alveolarization, it was necessary to elicit tokens of (**ŋ**) and (**k**) (in the formal tasks: passage reading, word list, and minimal pairs reading) which were frequent as well as those which were infrequent, as briefly mentioned in Section 3.5.2 above. This was in order to make sure that enough tokens were elicited which had a frequency higher than the median frequency of all words in the HKCAC[30], which was 7, but also enough tokens were elicited which had a frequency lower than the median frequency of 7. Table 36 demonstrates some examples of words, along with glosses, and corresponding frequency counts within the HKCAC (labelled f_{HKCAC}).

[30] That is, at the median frequency point of 7, there are as many lexemes which appear on the 'more frequent' end as there are on the 'less frequent' end. With the HKCAC containing 142,295 characters, this translates to a relative frequency of 49 per million[30].

Table 36. Lexical frequency coding information

Character		Gloss	HKCAC occurrences, f_{HKCAC}	Laplace-transformed $f_{HKCAC} = f_{HKCAC} + 1$ (LTf_{HKCAC})	$Log_{10}(LTf_{HKCAC})$	Rank	Log(rank)
騰	tʰɐŋ	'soar'	0	1	0	1742	3.241048
猛	maːŋ	'fierce'	1	2	0.30103	1741	3.240799
頸	kɛːŋ	'neck'	2	3	0.477121	1515	3.180413
症	tsɪŋ	'disease'	3	4	0.60206	1384	3.141136
廳	tʰɛːŋ	'hall'	10	11	1.041393	901	2.954725
從	tsʰʊŋ	'from'	50	51	1.70757	378	2.577492
性	sɪŋ	'nature / property'	100	101	2.004321	214	2.330414
生	sɐŋ	'life / Mr'	252	253	2.403121	103	2.012837
同	tʰʊŋ	'similar / with'	731	732	2.864511	36	1.556303
上	sœːŋ	'up / ascend'	331	332	2.521138	76	1.880814
講	kɔːŋ	'to speak'	1034	1035	3.01494	23	1.361728

There is a slight caveat which we should bear in mind, however, namely that the lexical frequency distribution of tokens in the data of this study containing our variables of interest, (**ŋ**) and (**k**), shows deviation from the lexical frequency distribution found in a natural language corpus. By eliciting more instances of certain tokens of interest, and by virtue of a different frequency distribution of types containing **-ŋ** and **-k** codas (as opposed to other codas, say, bilabial nasal or no coda), a study designed to examine frequency is effectively creating a 'local frequency' milieu by selectively focusing on particular phonemes, which is distinctly different in distribution from a 'global frequency' distribution. Going back to Section 1.7, where we discussed the increased likelihood of articulatory reduction with increased locally frequent exposure, it is not entirely clear how such 'local frequency boosts' might impact (**ŋ**) and (**k**) realization. The act of studying infrequent lexical items by eliciting them to appear more frequently in a sociolinguistic interview than they would appear in natural speech settings is a clear

case of the observer's paradox, whereby the very object under investigation is changed through the means of observation, which is a perennial problem for linguists studying natural language phenomena from collected data. However, while the frequency distributions of our collected data might be different from naturally-occurring speech, the overall trend whereby frequent lexemes were more frequent than the infrequent lexemes still held within the collected speech, with the only difference in by *how much* the frequent items were more frequent than the infrequent items.

3.6.1 Lexical frequency as a continuous variable

Treating lexical frequency as a binary categorical variable has traditionally been the norm in studies on word frequency and sound change. This involves using a frequency cut-off point to divide the lexicon up into two categories: 'frequent' vs. 'infrequent' lexical items. As mentioned in Section 1.7, Bybee (2002), used the relative frequency figure of 35 per million as a cut-off point, which was the median obtained from Francis & Kučera (1982) for regular past tense verbs, the target for her investigation of lexical frequency effects on **t**/**d**-delection. Walker (2012) used a categorical frequency measure obtained by arbitrarily dividing up corpus frequencies in neat groups with round figures, but this treatment did not seem statistically-motivated, and in fact, he admitted the arbitrary nature of this treatment.

There are clear-cut advantages to using lexical frequency as a categorical independent variable. It avoids fitting one linear curve through data that might potentially be structured distinctly differently within different intervals. It is also easy to manipulate data by binning lexical items into only two categories. The main problem pertains to where to draw the line: which frequency point to use as a cut-off. If we were

to employ the median, which many studies seem to do, then roughly as many tokens would be on one side as the other. Essentially this would mean dividing the total number of tokens in the corpus roughly by half and assigning the label "frequent" to the portion with higher frequencies and "infrequent" to the portion with lower frequencies. However, even using the median would be an arbitrary cut-off point, as it is not clear that the median reflects any empirically validated division between the frequent and the infrequent lexical items.

The problems associated with "binning" the frequency data and the arbitrariness of where to place the cut-off point can be avoided by using lexical frequency as a continuous measure. The fact that frequency and rank data from big corpora follow Zipf's law[31] has useful applications for linguistic analysis, the most important being that knowing a given word's position on a linear scale makes it easier for a linguist to interpret raw frequency numbers. Take the character 講 **goŋ˨˩** 'to speak / tell / talk,' for instance. This character occurs 1034 times in the HKCAC, and ranks as the 23rd most frequent character in the corpus (the most frequent, highly-ranked character in the corpus is 係 **hɐi˧** 'to be, yes/indeed' at 6625 occurrences). Compare this with the character 想 **sœ:ŋ˨˩** 'to think/want,' which has a frequency of 299 and ranks as the 89th most frequent character in the corpus. In other words, from the raw frequency values, 講 **goŋ˨˩** appears 735 times more than 想 **sœ:ŋ˨˩**, which might lead one to infer that 講 **goŋ˨˩** is ranked much higher than 想 **sœ:ŋ˨˩**. Whereas in reality, the gradation of frequency from the highest ranked lexemes drops very rapidly, and the gradation of frequency in the lowest ranked lexemes hardly drops, in fact tailing off almost horizontally. Using a

[31] See Appendix VII for an exposition.

log-log relationship makes it much easier to intuitively infer how lexemes rank compared to one another within all the lexemes in the corpus. The log(frequency) values corresponding to 講 **goŋ˨** and 想 **sœːŋ˨** are 3.015 and 2.476, respectively, out of the maximum log(frequency) of 3.821 for the most frequent character in the HKCAC, 係 **hɐi˧**.

The fact that log(frequency) is a relatively easier to interpret parameter for comparing frequency difference between two lexemes—due to the almost linear relationship of log(frequency) against log(rank)—as well as the fact that using log(frequency) as a continuous measure avoids the introduction of multiple parameters in the mathematical model mean that it has been adopted by many studies investigating lexical frequency and phonetic variation (e.g. Docherty, Hay & Walker 2006; Guy, Hay & Walker 2008; Dinkin 2008; Bullock, Toribio & Amengual 2014). Following these studies, log(frequency) will be used when performing statistical analyses of lexical frequency.

3.6.2 The problem of zero frequency and Laplace smoothing

While the HKCAC has roughly 142,295 characters, which is a fair size for a spoken corpus which contains phonetically transcribed data, there were occurrences of lexemes elicited in the structured tasks or which occurred in the casual speech style portion of the sociolinguistic interview which did not occur in this corpus. That is to say, these were infrequent, but had an f_{HKCAC} of zero. Yet, these lexemes were significant in this study because they contain the consonantal codas of interest, **-ŋ** and **-k**, and were in the 'infrequent' category. An example would be **ŋaːŋ˧** 硬 'hard', which occurred in the dataset but did not appear in the HKCAC, so the frequency for **ŋaːŋ˧** 硬 would be zero.

This in itself would not be a problem, but if we were to utilise the inverse linear log-log relationship between frequency and rank described above, we would need a work-around method, as the logarithmic function is undefined at zero. The work-around solution for this was to apply arithmetic Laplace smoothing by adding 1 to all of the HKCAC frequency data, such that characters which did not occur in the HKCAC could be assigned a frequency count of 1. This way, the inverse linear doubly logarithmic function was still defined at frequency = 1, and we have placed our tokens of interest on the log(frequency) scale, to allow a comparison against other lexemes within the data. This new transformed variable is displayed in the 5th column of Table 36. Lexemes which do not occur in the HKCAC were assigned the lowest rank of 1742.

3.6.3 Limitations of using corpus-based frequency data

It has to be added, however, that even obtaining frequency data from a spoken corpus is not without its problems. Certainly the assumption behind it, that the corpus from which the frequency data was obtained, in this case, a series of phone-in radio programmes, is representative of the overall speech community, is not unquestionable. The radio programmes covered a range of genres (see Table 37), but the unavoidable consequence is still that certain themes such as late-night ghost stories of the supernatural and the afterlife (even granted Hong Kong people's general propensity to take interest in such topics), as well as politics and economics, are overrepresented, whereas many words used in daily life, such as those that refer to, for instance, relatives and food, are underrepresented in this corpus.

Table 37. Genres covered by the Hong Kong Cantonese Adult Corpus

Themes	Program name	Time of broadcast	No. of files	No. of callers / participants	Recording time
Forums					
Politics	政黨論壇 [Political parties forum]	morning	2	11	100 mins
Economics	特區年代財經學人 [Economists of the SAR era]	morning	1	4	42 mins
Phone-in programs					
Current affairs	平息你的風波 [To appease your mind]	morning	15	15	69 mins
Personal	有冇心情顏聯武 [Are you in the mood?]	late night	6	6	67 mins
Personal	星空奇遇鐵達尼 [Startrek & Titanic]	evening	15	15	73 mins
Personal	海琪的天空 [Hoi-Kei's sky]	late night	9	10	69 mins
Ghost stories	恐怖熱線 [Ghost hotline]	late night	8	8	67 mins

Furthermore, this practice assumes that frequency data obtained from a spoken corpus from a prior state of the language is relatively static, and does not take into account the rise and fall of ideas and current topics that might propel certain lexical items from being used very infrequently to very frequently in a given period of time. To give a (made-up) example, an otherwise very low frequency character might shoot up in frequency when it is in the name of a prominent politician who just rose to power. Nonetheless, given that the spoken corpus was collected fairly recently and close to the point of data collection as well as reasonably-sized, plus the fact that there are no comparable corpora of Cantonese adult speech which allow the extraction of frequency information, I will accept the inadequacies of this method and assume that frequency data obtained from the HKCAC are representative enough for our current purpose.

3.7 Data processing

The recordings obtained from the interviews were annotated using the language archiving and annotation tool ELAN Version 4.1.2 (Wittenburg et al. 2006). Different tiers were created on which to make annotations, corresponding to all the variables investigated in this thesis, including 'subject ID', 'age', 'age group', 'gender', 'speech style', 'tautosyllabic onset', 'preceding vowel', 'tone', 'following onset', 'lexeme', 'lexical frequency', 'lexical rank', and 'auditory coding decision'. For data from reading tasks, the sound files for individual speakers were split into individual sound files for each task (reading passage, word list, and minimal pairs), and forced alignment was performed on these individual sound recordings. The input text for forced alignment was based on a phonemic *Jyutping* transcription (developed by the Linguistics Society of Hong Kong) without tone information and is provided in Appendix V for reference. The online forced alignment tool WebMAUS (Kisler, Schiel & Sloetjes 2012), based on an earlier tool MAUS (Schiel 1999), was used. The option for processing multiple files was selected and input parameters were set as follows:

```
Language:        English
Canonly:         True
Minpauslen:      5
Startword:       0
Endword:         999999
Outformat:       TextGrid
Mausshift:       10.0
Insprob:         0.0
Inskantextgrid:  True
Insorttextgrid:  True
Usetrn:          False
```

The MAUS forced aligner supports the export of alignment times in TextGrid format. These steps aligned the text annotations with the correct positions in the sound files

with fairly high accuracy and eliminated the need for manually segmenting the sound files for token extraction, given that these formal texts were structured in the same way across all speakers. Even though the language parameter was set to 'English', and the tool was optimised for English data, it performed well on Cantonese too. The TextGrid files were then imported into ELAN for phonetic annotation. The tokens of interest, (**ŋ**) and (**k**), where extracted using the "Filter Tier" tool (Tier > Filter Tier) in ELAN. The coding procedure outlined in Section 3.8 was then followed.

3.8 Coding and quantifying tokens

Each individual token of (**ŋ**) and (**k**) was annotated with the following information: Chinese character, and auditory code (i.e. [velar], [alveolar], or [reduced]). Due to the fact that some Chinese characters have multiple meanings and pronunciations (e.g. 行 can be **ha**:**ŋ**˩˩, 'to walk'; **hɐŋ**˩˩, 'possible' as in 可行; **hɔ**:**ŋ**˩˩, 'a row'; or **hɔ**:**ŋ**˧˥ 'shop or company') any polyphonous characters were disambiguated in the annotations using roman numerals (e.g. '行 1', '行 2', etc.). Using a lexicon which was compiled specifically for coding purposes, phonological information was inserted automatically based on the specific morphosyllable. For instance, the annotation '行 2' would automatically have the phonological information 'h + aa + ng + tone 4' inserted. Other information, such as HKCAC frequency, was also automatically inserted into the datasheet this way.

The tokens of standard velar nasal codas were coded using auditory judgment supplemented with acoustic spectrographic information obtained via *Praat* (Boersma & Weenink 2013). Previous work has suggested that relying on auditory judgment alone is

prone to errors in identifying the range of consonantal variation present in phonetic realization, due to existing lexical knowledge and categorical perception (Lavoie 2002:39) and that to reduce any such errors, spectrographic information should be used to correct any auditory judgments (Barry 1996:15). Especially in the case of unreleased stops, such as the syllable-final stops in Hong Kong Cantonese, spectrographic information would help by showing formant transitions that aid in identifying the unreleased consonant. This auditory judgment-driven approach assumes that the perceptual machinery of the brain can normalise the incoming speech stream from different speakers and form a fairly good judgment of the relative articulatory configuration of the vocal tract in forming the acoustic speech stream. The additional spectrographic information from the acoustic signal serves as a means to double-check this judgment and narrow down the margin of error, but is especially useful for maintaining consistency and providing an additional source of information for making auditory judgments.

Traditionally, in sociovariationist studies coding for binary alternations (presence vs. absence of a feature) has been the norm, such that binomial logistic regressions can be performed. Standard statistical packages developed for analysing linguistic variation typically only support binomial logistic regression. However, in recent decades, studies have typically moved beyond binary distinctions and considered consonant variation as having multiple possible outcomes (e.g. Stuart 2007 on the Glaswegian /r/; Docherty, Hay & Walker 2006 on the phrase-final /t/ in New Zealand English) or even gradient outcomes (e.g. Schmid, Moosmueller & Kasess 2015 on the velarized lateral in Viennese dialect; see also Hall-Lew & Fix 2012:794–5 for a literature overview on gradience in (L) vocalization). In this study, we are also interested in the rate of lenition

in the velar coda consonants, and whether this might be related to, and inform us of a possible mechanism for, alveolarization. Through a pilot study, it was also confirmed that in casual speech, there was a whole range of reduced variants of (**ŋ**) and (**k**) production within natural speech, and these are not merely divided into conservative vs. innovative variants (i.e. [-ŋ] vs. [-n]; [-k] vs. [-t]: for the full transcription and coding scheme used in the pilot phase, please refer to Appendix VI). Instead of discarding the tokens of (**ŋ**) and (**k**) where there was reduced articulation or deviation from velar or alveolar realizations, these instances were also coded for and subsumed under an umbrella category named [reduced]. This was a compromise made to balance efficiency of coding, and the relative importance of each of the coded variants in contributing to an understanding of the sound change mechanisms. However, the choice to not code phonetic realization into multiple minutely distinct categories but instead to code using one single "reduced / nondeterminate" category also arose from a methodological consideration, which is that auditory perception has its limits and that it is often impossible to identify the exact place of articulation of a consonant when it is highly fused with contiguous segments in connected speech.

When phonetic coding took place, no reference was made to the actual Chinese character annotation (e.g. whether a token was 北 **bɐk˥** or 不 **bɐt˥** within the minimal pairs task), thereby reducing the chance of bias and improving objectivity of the coding by focusing all attention on the physical properties of the consonant realizations. This, of course, is not possible for tokens in the casual speech task, where the meaning of the utterance already dictates which lexeme and (corresponding morphosyllable) it is. That said, all effort was made to be as objective as possible when performing auditory coding.

Tokens which coincided with bursts of noise or tokens from reading tasks which were misread were discarded from the analysis.

3.8.1 Phonetic criteria used for coding (**ŋ**)

The next sections will discuss the phonetic characteristics of tokens of (**ŋ**) and describe in more detail how coding decisions were made. In general terms, the following criteria were used for auditory categorization for (**ŋ**): whenever there was a clear velar nasal, the token was coded as a [velar]. If there was a clear alveolar nasal, the token was coded as an [alveolar]. For these coding decisions to be applied, it was a necessary criterion for a word to be segmented such that it could be understood in isolation to contain a velar or an alveolar nasal. So, for a phrase like 行快啲 **ha**:**ŋ˩ fa**:**i˧ di**:˥, the word **ha**:**ŋ˩** is only coded as a [velar] if [ha:ŋ] can be isolated from the sound segment and understood as the word 行 with a velar nasal closure, with little regard to the length of this closure.

3.8.1.1 Tokens of (**ŋ**) coded as [velar]

Phonetically, the velar nasal [ŋ] combines characteristics of consonants which have a constriction in the velar position and those that are produced with nasalization. While consonants with a velar closure are described in the literature as spectrographically having a characteristic 'velar pinch' (in other words, having second and third formants moving towards each other and coming together), in the case of the velar nasal [ŋ], the velar pinch is harder to discern, with the second and third formant typically moving towards one another, but not quite forming a clear 'pinch', which could be due to nasal resonance and anti-resonance formants. This is shown in Figure 12, a spectrogram

extracted from a casual speech segment, produced by speaker MM35, of the word 聽 'listen' **tʰɛːŋ˥** which was produced with a clear velar nasal [ŋ].

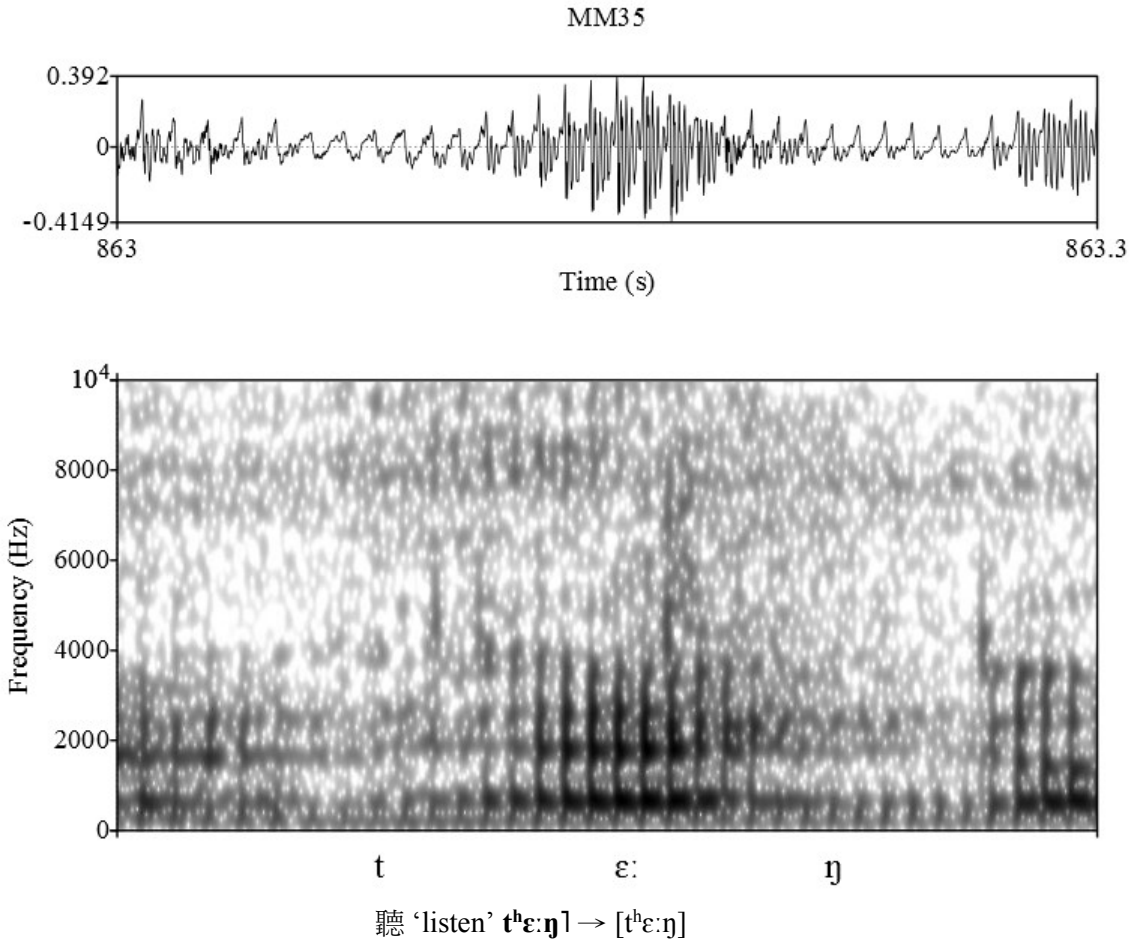


Figure 12. Token of (ŋ) coded as [velar]; Speaker: MM35

In the spectrogram, it can be seen that in the vowel [ɛː] the second formant, with a frequency at around 1800 Hz, and the third formant, with a frequency at around 2300 Hz, are edging closer to one another in the lead-up to [ŋ], but that a 'pinch' is not discernible. Throughout the duration of closure in [ŋ], nasal resonance and anti-rosonance patterns can be seen.

In this case, since the segment [ŋ] could be isolated and the velar closure was discernible auditorily, the token was coded as having a [velar] realization. These coding criteria were applied to all tokens of (**ŋ**) in order to be coded as [ŋ].

3.8.1.2 Tokens of (**ŋ**) coded as [alveolar]

The alveolar nasal [n] combines characteristics of consonants which have a constriction in the alveolar position and those that are produced with a nasal release. Spectrographically, the former trait, namely, alveolar constriction, is characterized by evenly-spaced first, second and third formants, while the latter trait, a nasal release, is characterized by nasal formants and anti-formants. This is shown in Figure 13, a spectrogram extracted from a casual speech segment, produced by speaker MM33, of the word (外來)生 'student' **sɐŋ˥** which was produced with an alveolar nasal [n].

In the spectrogram, it can be seen that in the alveolar nasal [n] the first formant, with a frequency of around 500 Hz, a second formant, with a frequency of around 1600 Hz, and the third formant, with a frequency of around 2700 Hz, are evenly spaced, which is characteristic of alveolars. Throughout the duration of closure in [n], nasal resonance and anti-resonance patterns can be seen.

In cases such as this, since the segment [n] could be isolated and the alveolar closure was discernible spectrographically as well as auditorily, the token was coded as

having an [alveolar] realization. These coding criteria were applied to all tokens of (**ŋ**) to be coded as [n].

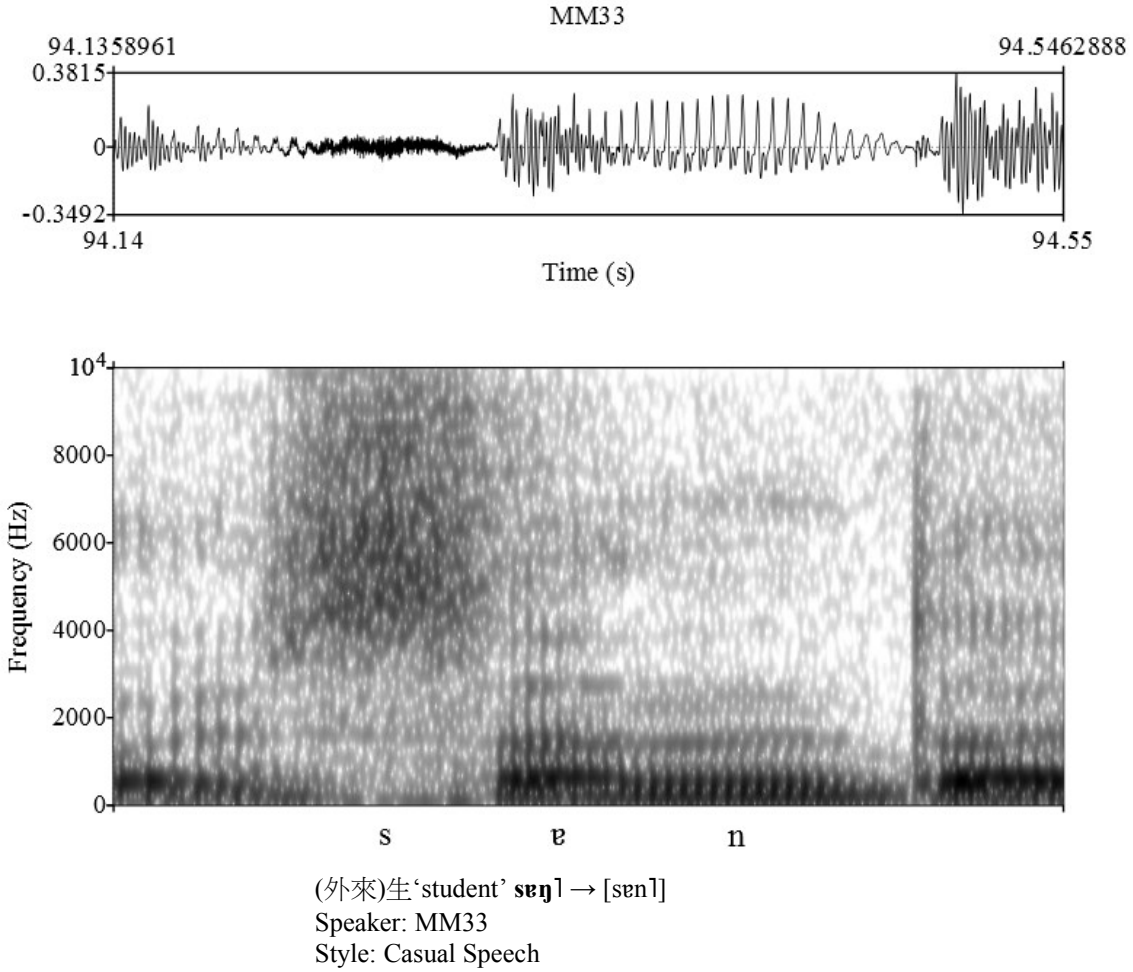


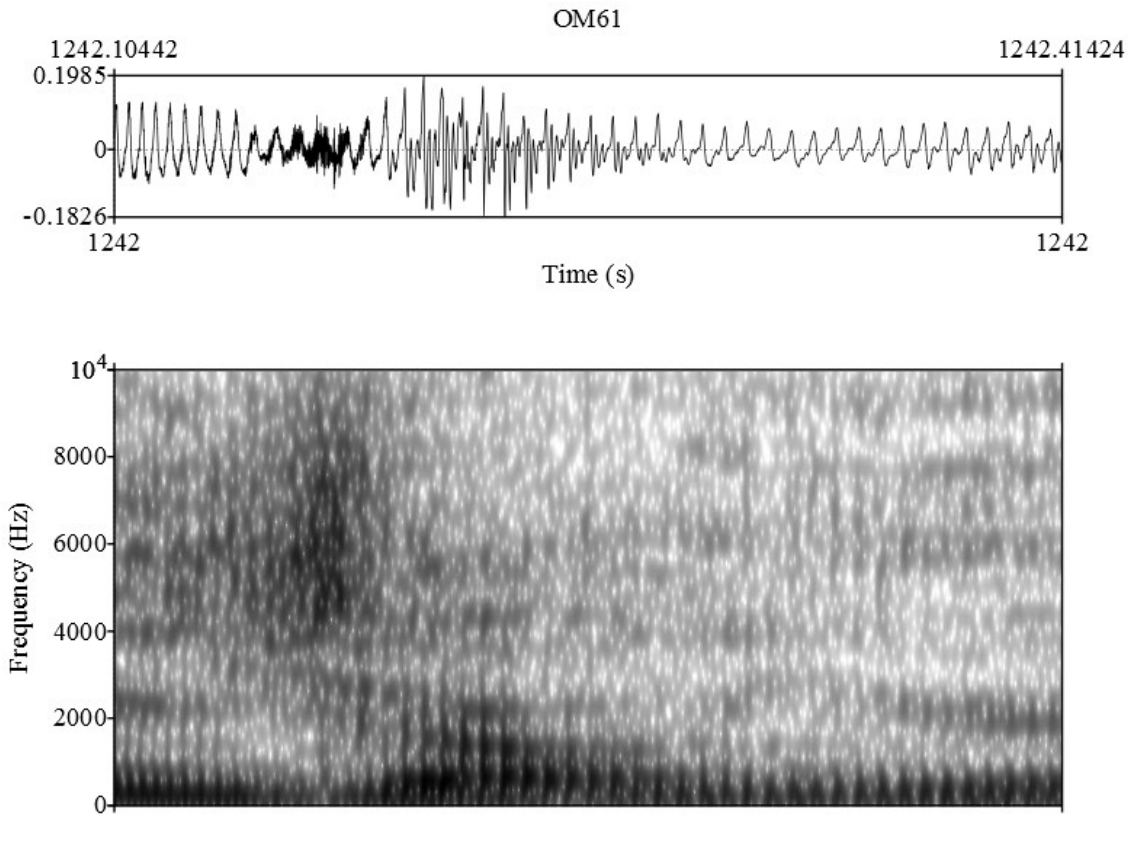
Figure 13. Token of (ŋ) coded as [alveolar]; Speaker: MM33

3.8.1.3 Tokens of (**ŋ**) coded as [reduced]

In all other cases, where neither a [velar] nor [alveolar] could be perceived, the token was coded as [reduced]. In other words, in the case of overlapping articulation without any auditorily perceivable way of pinpointing the place of articulation, the token was coded as [reduced]. Other cases included a complete dropping of the final velar nasal

coda or absence of a closure, as well as instances of assimilation where the velar or alveolar target was not reached, for instance when **-ŋ** combines with a following segment **j-** to form a [ɲ], or instances of uncertainty, when it could not be determined whether there was a velar or alveolar nasal closure.

As an illustration, Figure 14 shows a spectrogram from a casual speech segment, produced by speaker OM61, of (機)場(遠) '(air)port/field' **ts^{h}œːŋ˩** which had no discernible velar or alveolar closure and was followed by a following onset [j].



(機)場(遠) 'field/court' **ts^{h}œːŋ˩→** [ts^{h}œː]
Speaker: OM61
Style: Casual Speech

Figure 14. Token of (ŋ) coded as [reduced]; Speaker: OM61

In the spectrogram, it can be seen that the formants of the vowel [œ:] are clearly discernible in the nuclear portion of the vowel, but the intensity then drops and the formants become non-discernible, to be followed by the onset of the following syllable, [j]. In cases such as this, the token of (**ŋ**) was coded as [reduced].

3.8.2 Phonetic criteria used for coding (**k**)

The following criteria were used for auditory categorization for (**k**): whenever there was a clear velar stop, the token was coded as a [velar]. If there was a clear alveolar stop, the token was coded as an [alveolar]. As with (**ŋ**), for these coding decisions to be applied, it was a necessary criterion for a word to be segmented such that it could be understood in isolation to contain a velar or an alveolar nasal.

To illustrate, let us consider different cases and how phonetic coding decisions were made in each case.

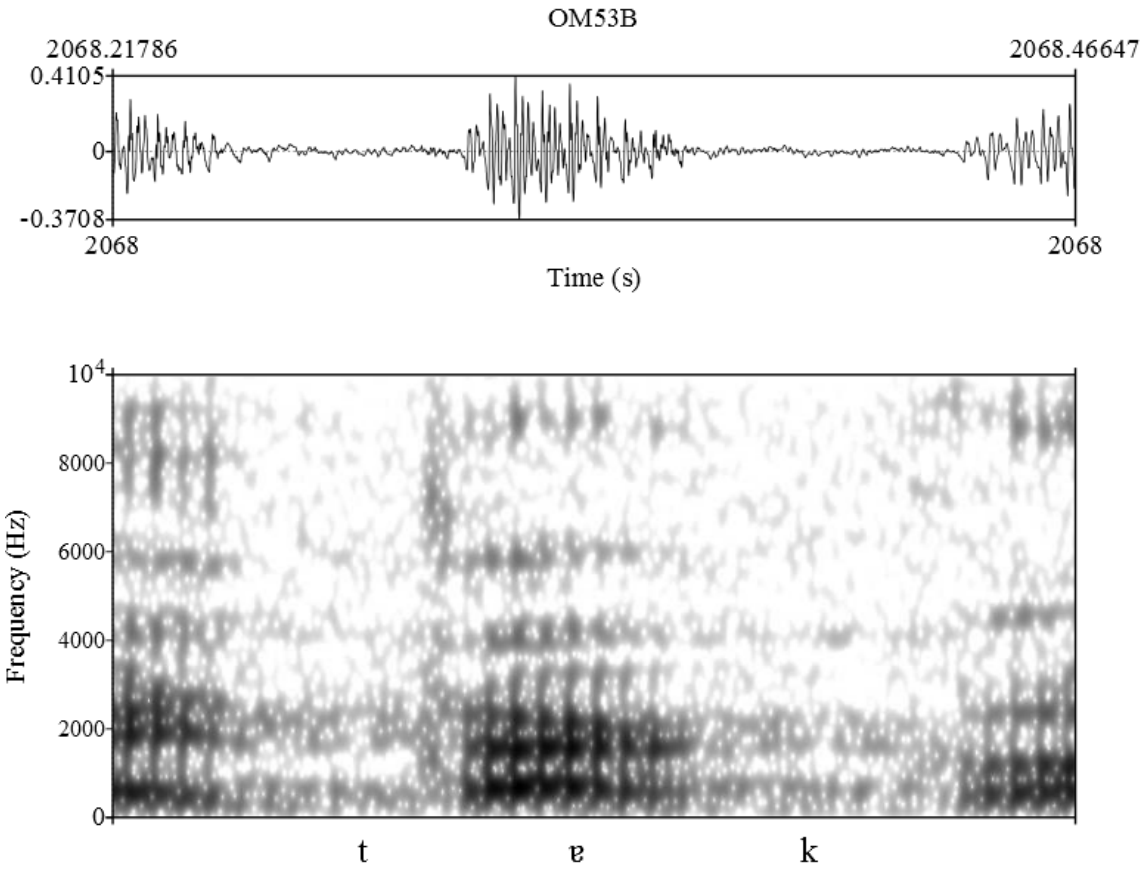
3.8.2.1 Tokens of (**k**) coded as [velar]

Phonetically, the velar stop [k] has a constriction in the velar position. In the case of Cantonese, the [k] is an unreleased stop. Spectrographically, the velar stop has a characteristic 'velar pinch', in other words, having second and third formants moving towards each other and coming together.

To illustrate this, let us consider the spectrograph from a sound segment produced by speaker OM53B, of (啟)德 '(Kai) Tak (Airport)' **tɐk˥**, which was produced with a velar stop [k]. The formants of the vowel [ɐ] transition from an evenly-spaced configuration, characteristic of the preceding alveolar [t], to the nuclear portion of the vowel, with frequencies of around 800 Hz for the first formant, 1700 Hz for the second

formant, and 2300 for the third formant. This configuration then changes at the end of the vowel portion in such a way that the second and third formants edge closer to one another, coming together during the constriction part of [k]. As an aside, it can also be seen that the constriction portion for [k] is temporally longer than the vowel duration, which echoes the finding in previous studies that the coda consonant following a short nuclear vowel is temporally longer. The fact that [k] is unreleased in Cantonese, is also spectrographically evidenced by the absence of a plosive burst in the higher frequencies.

In such clear cases of a presence of a [k], the token of (k) was coded as [velar].



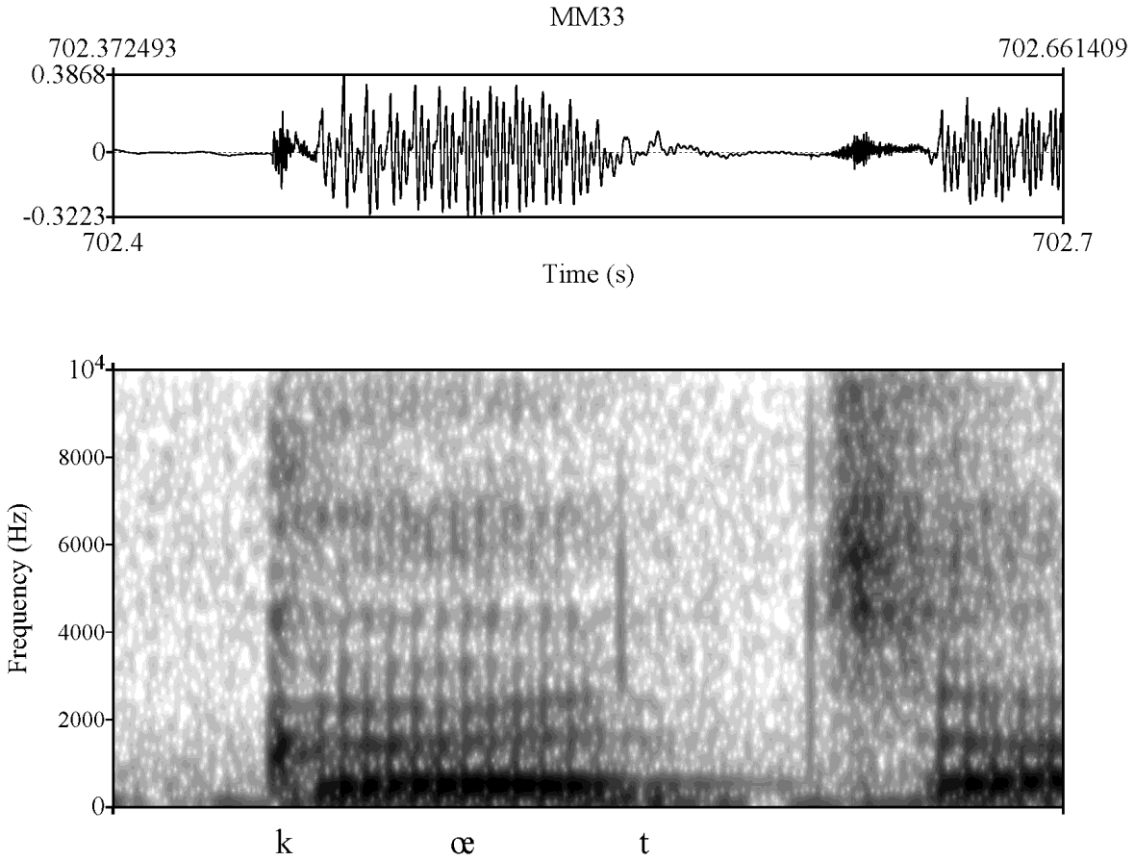
(啟)德 '(Kai) Tak (Airport)' **tɐk˥** → [tɐk]
Speaker: OM53B
Style: Casual Speech

Figure 15. Token of (k) coded as [velar]; Speaker: OM53B

3.8.2.2 Tokens of (**k**) coded as [alveolar]

Phonetically, the alveolar stop [t] is characterized by a full constriction at the alveolar ridge, and spectrographically it is similar to the [n], except that there are no nasal formants and anti-formants present. In the case of Cantonese, [t] is also unreleased.

As illustrated in the spectrogram below (Figure 16), of a sound segment produced by speaker MM33, (隻)腳(就) 'leg' **kœ**:**k**˦ was realized with an alveolar stop [t] at the end of the syllable.



(隻)腳(就) 'leg' **kœ**:**k**˦ → [kœ:t]
Speaker: MM33
Style: Casual Speech

Figure 16. Token of (k) coded as [alveolar]; Speaker: MM33

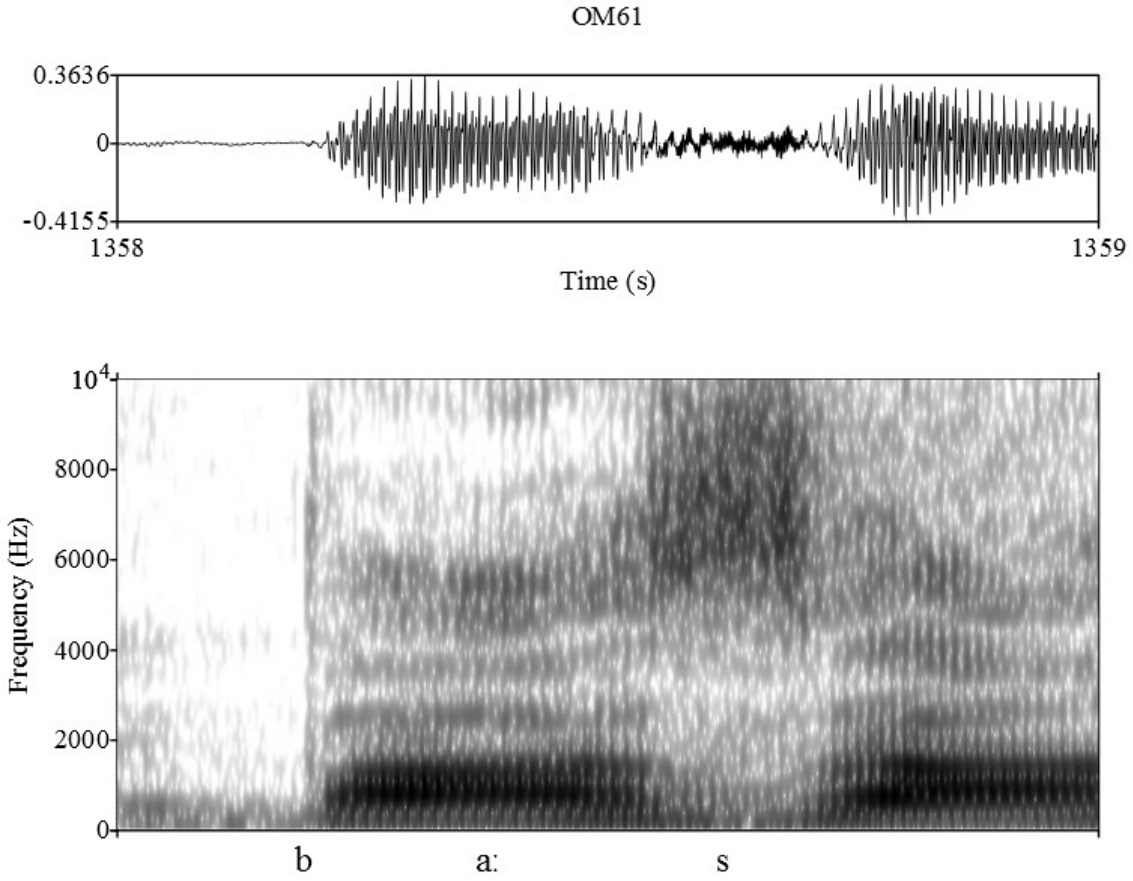
As can be seen, the vowel formants of [œ:] transit into the formants of [t] in the constriction part of the consonants, with the characteristically evenly-spaced formants of alveolar segments, the first formant at roughly 500 Hz, second formant at 1700 Hz, and the third formant at roughly 2200 Hz. In similarly clear cases of the presence of an alveolar stop [t], the tokens of (**k**) were coded as [alveolar].

3.8.2.3 Tokens of (**k**) coded as [reduced]

In all other cases, where no [velar] or [alveolar] could be perceived, the token was coded as [reduced]. This included cases of overlapping articulation without any auditorily perceivable way of pinpointing the place of articulation. Other cases included a complete dropping of the final velar coda or absence of a closure, instances of assimilation where the velar or alveolar target was not reached, for instance when **-k** was followed by an onset vowel segment and was realized as a glottal stop [ʔ]—or indeed any instance of glottalling—as well as instances of uncertainty (i.e. when it could not be determined whether there was a velar or alveolar nasal closure).

Figure 17 illustrates a case where the realization of a token of (**k**) was non-discernible as either [velar] or [alveolar]. In this case, the speaker, OM61, was producing the word 百(三) 'hundred' **pa**:**k**ɨ as [pa:]. As can be seen in the spectrogram, the vowel [a:] was directly followed by the following onset [s]. While the first, second and third formants of [a:] warp towards the formant configuration of the following [s], there was no spectrographic evidence of any velar or alveolar closure before the onset of

this following [s]. In cases such as this, and in cases where it was uncertain what the place of articulation was, the token of (**k**) was coded as [reduced].



百(三) 'hundred' **pa**:**k**ɨ → [pa:]
Speaker: OM61
Style: Casual Speech

Figure 17. Token of (k) coded as [reduced]; Speaker: OM61

3.8.3 Going beyond a binary distinction

The reason I adopted this coding scheme as opposed to a binary velar vs. alveolar one is because in casual speech there were lots of ways in which phonological environment interacted with the (**ŋ**) and (**k**) variables, and rather than treating those cases where there was no clear velar or alveolar realization as extraneous noise, this remaining category could inform us the extent to which nasal coda consonants are dropped or reduced, which might contribute to the percept of **la**:**nʴjɐm**˥ or 'lazy pronunciation'. In previous studies, the coding decisions were not very clearly spelled out: the percentage of alveolar realization was reported in Wong (2005) although she did not mention what the remaining portion from which the percentage was derived consisted of. Shum (1993) just referred to "percentage sound change", it being unclear whether only strictly alveolar realizations of (**ŋ**) were taken into account, or whether this was also combined with velar realizations of (**n**) in coda position. The current coding system provides more information on lenition and reduction in speech, with the option of leaving out the [reduced] category for comparison with other studies.

3.9 Statistical analysis

The choice of which type of statistical analysis to perform is dictated both by the nature of the data as well as the research questions for which an answer is to be obtained. This section describes the statistical tools used and compares some of the advantages and disadvantages of each.

3.9.1 Tools used for statistical analysis

Due to the above combination of categorical and continuous factors, a combination of the R statistical package (R Core Team and contributors 2013) and MLwiN (Rasbash et al. 2005) was used for statistical analysis. While Goldvarb was in the past considered the standard tool for statistical analysis in sociolinguistic variation studies, its lack of flexibility and overestimation of significance, compared to multilevel models (see discussion in Johnson 2009), meant that it was more preferable to opt for other widely-available and used statistical packages which report results in a way recognised in other (social) sciences. This also allowed the analysis to go beyond Goldvarb's restriction that all independent variables (or predictors) be categorical. How to divide up variables such as lexical frequency into discrete categories is sometimes an arbitrary decision, as we saw in the discussion in Section 3.6. Goldvarb also only supports binomial logistic regression, which means one can only analyse binary dependent (or response) variables. These limitations of Goldvarb highly constrain the questions we can ask as well as restrict our understanding of the factors under investigation, which we shall see in the results of using lexical frequency as a continous variable, and in our exploration of the [reduced] variants, in comparison with [alveolar] and [velar] realizations. In variationist sociolinguist studies, it has in recent years been considered standard practice to use multilevel modelling, also known as random-effects modelling (see, e.g. Tagliamonte & Baayen 2012 for a comparison of different models in analyzing was/were variation; and Agresti et al. 2000 for an walk-through of how to use random-effects models to analyze hierarchically-structured data). For a more detailed discussion, see Section 3.9.2.

The choice of MLwiN was driven by the fact that it is one of the few readily available statistical packages which supports both multinomial logistic regression and

multilevel (or mixed effects) modelling. It is developed by the Centre for Multilevel Modelling at the University of Bristol, and is a fairly well-documented tool with instructive manuals available, which is free for academic research use. The choice of R statistical package, on the other hand, was driven by the fact that it is open source and has an extensive range of libraries and powerful plotting functions (e.g. `ggplot2`).

In the following chapters, the full statistical models are presented for (**ŋ**) and (**k**), and where significant, the relationship between the independent factors (also known as predictors) and dependent factors (or responses), expressed as counts or percentages of (**ŋ**) and (**k**) realization and presented in graphs. Where needed, univariate statistics are also explored, which are essentially "naïve" explorations performed to examine the broad relationships between independent factors and dependent factors. This step is also important for making sure that statistical assumptions have not been violated. The independent factors in this study include external factors, such as speakers' age, gender, and speech style, but also internal factors, such as segmental environment (preceding vowel and following onset), tone, and lexical frequency. However, univariate statistics are merely a means to getting an impressionistic overview of how the response variable changes as univariate factors are varied. The full multinomial logistic regression model included as many factors as possible, and shows the factors which contribute significantly to progressive variants in this sound change, and allows us to test our original hypotheses presented in Chapter 1. It is only by performing a logistic regression on all the factors in one single model that one can offer an accurate picture of how the response variable correlates with a given independent factor by keeping all other independent factors constant. The statistical principle that effects of independent factors should be estimated *ceteris paribus* is thereby adhered to in this work as much as

possible. In other words, the statistical effect of any given independent factor on the dependent factor is estimated with *all else being equal*.

Overall, for the (**ŋ**) variable, 8392 tokens were coded, whereas for the (**k**) variable, 5532 tokens were coded. This difference in numbers reflects the fact that there are fewer lexical items in Cantonese with a stop coda **-k**, even though the occurrence of **-k** has already been artificially boosted in the elicitation tasks. A further breakdown of token counts will be presented in each subsection on individual variables below. Univariate descriptive statistics were obtained and graphs plotted in the R Language and Environment for Statistical Computing (Version 3.0.2, R Core Team 2013). Simple line and bar charts were plotted using OpenOffice Calc.

3.9.2 Multilevel modelling

This section will explore the benefits and disadvantages of using mixed models and explain the modelling procedure. The use of multilevel modelling has become commonplace in sociolinguistic variation analysis. Multilevel models take into account the hierarchical structure that is typically present in data from the social sciences. These models are also referred to as "mixed models", because they include "fixed" and "random" parts. The use of multilevel models in repeated-measures linguistic data is important because each speaker might systematically differ from another speaker with respect to their speech behaviour. General logistic regression models which do not incorporate multiple levels, if the data is structured as such, tend to over-estimate the level of significance, and underestimate the amount of variance, because these models assume that each response, or token in the case of our data, is an independent observation from the next token. This assumption of independence between individual

speech segments produced by the same speaker simply does not hold. Multilevel models can compare the data to the models while taking into account the fact that these speech segments are produced by the same speaker and might therefore systematically affect the outcome variable. Fixed effects models (supported by tools such as Goldvarb), do not sufficiently respect the heterogeneity of speakers, as such statistical models are built upon computed averages across all tokens produced by speakers grouped together (according to e.g. age or gender). It is now considered standard practice to use mixed effects models for sociovariation studies.

While there are obvious analytical advantages to using multilevel models, there are also some practical disadvantages. Since this is a relatively new tool, the amount of support available is relatively limited, compared to more established tools used within the field of sociolinguistic variation. Also, due to the estimation procedures used, it is not always the case that the models will converge and be able to provide estimates successfully when including all factors of interest in the model while including speakers as a random effect. As will be discussed below, when the full model is presented, the best model is the 'best available' model for which convergence of estimations can be reached using the given statistical algorithm.

3.9.3 Multilevel modelling procedure

For obtaining the best predictive model of the response variable, a backward selection procedure was used (Twisk 2006:80). In such procedure, the first step was to obtain the fullest model with all the possible predictor variables. Then a random intercept was allowed to explore whether a new full model with a random intercept is a better model than the previous one. If this was the case, the new model with random intercept was used. This new model was then further evaluated to see whether random slopes for each

of the predictor variables was necessary. In turn, a random slope was allowed for the predictor variables, one variable at a time, to see whether adding a random slope for a given predictor was necessary. If it was necessary to add a random slope for a given predictor variable, this was allowed in the final model. With all predictor variables assessed in this manner for the necessity of adding a random slope, insofar as the maximum likelihood estimation algorithms could converge, the full final model could then be obtained. The predictor variable(s) with the lowest significance (as measured by *p*-value) was excluded.

For technical reasons (see Baayen 2008:166; and Agresti 2007:139), lower-order terms (or main effects) were kept in the model when there was a higher-order interaction with another term, which is common practice in statistical modelling. Multinomial logistic regressions [32] are of a more generalized type of regression compared to binomial logistic regression. Whereas binomial logistic regressions are applied to categorical response variables with two outcomes, multinomial logistic regression can be applied to categorical response variables with multiple outcomes, say $j = 1, 2, \ldots, m$.

Essentially, a multinomial regression model estimates the probability π_{ij} that each token i is in each jth class of the m response classes, $j = 1, 2, \ldots m$. In our case, $m = 3$, with the category response values being [velar], [alveolar], and [reduced]. The category response [velar] ($j = 1$) was taken as the base class, which was computed as the residual probability after the other classes π_{i2} [alveolar] and π_{i3} [reduced] had been modelled. For an accessible description of how to fit such a model in the R statistical environment, refer to Rossiter & Loza (2012:47–56). This resulted in a model with $k + 1$ coefficients

[32] Multinomial logistic regression models can also be interchangeably called polytomous logistic regressions.

(β_0, β_1, β_2, β_3, …, β_k) for each of the $j = m - 1$ classes, which in our case are the two [alveolar] and [reduced] classes. The equation for this model is as follows:

$$\pi_{ij} = \frac{e^{(\alpha_j + \beta_{1j} x_{i1} + \ldots \beta_{kj} x_{ik})}}{1 + \sum_{l=2}^{m} e^{(\alpha_j + \beta_{1l} x_{i1} + \ldots \beta_{kl} x_{ik})}}, j = 2,...., m \quad (1)$$

$$\pi_{i1} = 1 - \sum_{j=2}^{m} \pi_{ij} \quad (2)$$

Fitting of these sets of equations was performed using a maximum likelihood estimation procedure (1st order MQL), which obtained fitted values for α and β, used to obtain the odds ratio of each token *i* being in response category [alveolar] or [reduced], relative to the base class [velar]. This translates to answering the question, how much more likely is it that a token *i* is in category [alveolar] as opposed to in category [velar]? And, how much more likely is it that a token *i* is in category [reduced] as opposed to in the category [velar]? The log-odds can then be computed by means of taking natural logarithm on both sides of the equation:

$$\ln \frac{\pi_{ij}}{\pi_{i1}} = \alpha_j + \beta_{1j} x_{i1} + \mathrm{K} + \beta_{kj} x_{ik} \ , j = 2, \mathrm{K} \ , m \quad (3)$$

where $m = 3$, in our case. The model which is estimated here can be expressed by the following:

$$\ln \frac{\pi_{Alveolar}}{\pi_{Velar}} = \beta_0 + \beta_1 age + \beta_2 gender + \beta_3 style + \beta_4 age \times gender + \beta_5 gender \times style + \beta_6 age \times gender \times style + \beta_7 vowel + \beta_8 tone + \beta_9 following_onset + \beta_{10} lexical_frequency \quad (4)$$

$$\ln \frac{\pi_{Reduced}}{\pi_{Velar}} = \beta_{11} + \beta_{12} age + \beta_{13} gender + \beta_{14} style + \beta_{15} age \times gender + \beta_{16} gender \times style + \beta_{17} age \times gender \times style + \beta_{18} vowel + \beta_{19} tone + \beta_{20} following_onset + \beta_{21} lexical_frequency \quad (5)$$

In other words, we are trying to estimate, given the odds ratio of a given token falling into the [alveolar] category vs. the [velar] category, how great of an effect (measured by the coefficients $\beta_{1, 2, 3, ...}$) each of the predictors 'age', 'gender', 'style', 'vowel', 'tone', 'following onset' and 'lexical frequency' has on this odds ratio. At the same time, we are also performing the equivalent estimation of the effects of the same set of predictors, given the odds ratio of [reduced] vs. [velar]. Once estimates for the coefficients $\beta_{1, 2, 3, ...}$ are obtained, the relative effects of each of the predictors can be interpreted by means of taking the exponent of β, $\exp(\beta)$, which gives us the estimated increase in odds due to that given factor (Agresti 2007:104–5). For instance, if using [velar] as the reference category and [alveolar] as the treatment category, we find that the estimation for the coefficient for 'age', β_1, is -0.042, then the decrease/increase in odds of a token being in category [alveolar] as opposed to in category [velar], for every 1-unit increase in age would be $\exp(-0.042) = 0.959$. That is, for each 1-year increase in age, there is a 4.11% *de*crease in odds of a token being of category [alveolar] as opposed to [velar]. This percentage decrease/increase in odds ratio is simply computed by $(\exp \beta_1 - 1) \times 100 = -4.11$, with the negative sign indicating a decrease.

The choice of reference category, essentially the baseline treatment level, was the conservative variant as far as this was possible. For instance, for the response variable, the category [velar] was chosen as the reference, such that estimates for the other categories ([alveolar] and [reduced]) could be compared against the conservative reference category [velar].

The output of results from a multinomial logistic regression model is similar to that of a binomial logistic regression model but differs somewhat from variable rule analysis results. These differences are explained in Johnson (2009) but are summarised here for ease of reference. The main difference between the output of a Goldvarb analysis and any other statistical package is whether the coefficients are expressed as sum contrasts (an example of which is factor weights) or as differences from the reference levels, also known as treatment contrasts (Johnson 2009:361). In the case of multinomial logistic regression, difference from the reference levels is obtained. As mentioned in Section 1.6, a factor weight of 0.50 indicates that the factor has no effect on the outcome, with values of less than 0.50 indicating that the factor has a disfavouring effect and values of over 0.50 indicating that the factor has a favouring effect. Table 38 (from Johnson 2009:361) shows a comparison between factor weights and log-odds. Statistical results in this thesis will be presented using log-odds and differences in odds ratios between the treatment and reference levels.

Table 38. Factor weights and corresponding log-odds

Factor weight	Log-odds
.000	−∞
.100	−2.197
.200	−1.386
.300	−0.847
.400	−0.405
.500	0
.600	+0.405
.700	+0.847
.800	+1.386
.900	+2.197
1.000	+∞

(Johnson 2009:361)

3.10 Summary

This section described the methodological design for this study. The social variables employed, namely age, gender and speech style, are fairly standard variables explored in sociolinguistic variation studies, and the methodology used in collecting data in a sociolinguistic interview, with casual speech, passage reading, word list reading and minimal pairs reading tasks were described. With the focus of this thesis on usage-based lexical frequency, it was also described how lexical frequency information was extracted from the Hong Kong Cantonese Adult Corpus (HKCAC) and what the implications were of using lexical frequency as a binary categorical vs. as a continuous measure, and why the latter is more preferable, due to the arbitrariness of where to draw the boundaries between 'frequent' vs. 'infrequent'. A description of statistical tools used in this thesis was also offered. The next chapters describe the full models which were

obtained from performing statistical analyses using the afore-mentioned methodology, and the results of statistically significant factors are discussed in detail.

4 Full statistical model combining social, neogrammarian and lexical factors

This chapter presents a full statistical model which aims to capture all the above-mentioned factors—the social, neogrammarian, and lexical—to estimate how large the effect each has on (**ŋ**) and (**k**) variation as [velar], [alveolar] and [reduced] realizations. The shorthand descriptions "[velar]", "[alveolar]" and "[reduced]" will be used throughout the discussion of statistical results for the categorical coding decisions that were made with respect to outcomes of (**ŋ**) and (**k**) realization (please refer to Sections 3.8.1 and 3.8.2 again for descriptions as to how these categorization decisions were made). As mentioned in Section 3.9, it is hoped that, in doing so, independent factors can be estimated *ceteris paribus*, such that the statistical effect of any given independent factor on the dependent factor is estimated with all else being equal.

4.1 The full modelling procedure

The multinomial logistic regression was performed in MLwiN, using a quasi-likelihood estimation method for linearization and subsequently applying the iterative generalised least squares (IGLS) estimation procedure (Rasbash et al. 2012:128). The modelling procedure described in Section 3.9.3 was followed. Firstly, fixed factors were included in the model. Social factors included in the model were: 'age', 'gender', and 'style', as well as interactions between them: 'age × gender', 'gender × style', 'age × gender × style'. Age was included as a continuous variable, the motivations for which were described in Section 3.4. Linguistic factors included phonological or phonetic contexts: 'preceding vowel', 'following onset', and 'tone'. In terms of the lexical dimension,

'lexical frequency' was included as a continuous explanatory variable instead of as a binary categorical variable ('frequent' vs. 'infrequent'), in line with the considerations discussed in Section 3.6.1, as was its interaction with 'preceding vowel' and 'following onset'. The full model thus obtained for (**ŋ**) and (**k**) realization estimated coefficients of the above factors given log-odds of [alveolar] vs. [velar], and the log-odds of [reduced] vs. [velar], according to Equations (6) and (7) below.

$$\log \frac{\pi_{Alveolar}}{\pi_{Velar}} = \beta_0 + \beta_1 age + \beta_2 gender + \beta_3 age + \beta_4 style +$$
$$\beta_5 age \times gender + \beta_6 gender \times style + \beta_7 age \times gender \times style +$$
$$\beta_8 vowel + \beta_9 tone + \beta_{10} following_onset + \beta_{11} lexical_frequency +$$
$$\beta_{12} lexical_frequency \times vowel + \beta_{13} lexical_frequency \times following_onset \quad (6)$$

$$\log \frac{\pi_{Reduced}}{\pi_{Velar}} = \beta_{14} + \beta_{15} age + \beta_{16} gender + \beta_{17} age + \beta_{18} style +$$
$$\beta_{19} age \times gender + \beta_{20} gender \times style + \beta_{21} age \times gender \times style +$$
$$\beta_{22} vowel + \beta_{23} tone + \beta_{24} following_onset + \beta_{25} lexical_frequency +$$
$$\beta_{26} lexical_frequency \times vowel + \beta_{27} lexical_frequency \times following_onset \quad (7)$$

Next, fixed factors were examined to remove any insignificant factors. Whenever interactions between factors were significant, the lower order terms were retained in the model (following Baayen 2008:166; and Agresti 2007:139). Lastly, a random intercept

was included where this was possible insofar as the model estimation could still converge.

4.2 Full model for (ŋ) realization

This section presents the full model for (ŋ) realization. A summary of the significant factors is provided in Tables 39 and 40, showing (ŋ) > [alveolar] variation, and (ŋ) > [reduced] variation respectively. A screenshot of the "best" model output from MLwiN is shown in Model 1 in 0. All the factors which were significant were included in the tables. The single-factor terms which were insignificant but where higher order interactions with other factors were significant (see discussion in Sections 3.9.3 and 4.1 above on higher order interactions) were also included in the table of results, as they merit further discussion.

First, let us examine Table 39, which shows the model portion for (ŋ) > [alveolar] variation, with the category [velar] taken as the reference response category. From the estimation results, it can be seen that age was estimated to be highly significant ($p = 2.56 \times 10^{-12}$) in predicting (ŋ) > [alveolar] variation. With every 1-year increase in age, it was 4.1% *less* likely for (ŋ) to be realized with [alveolar].

Gender was found not to be significant in predicting this variation ($p = 0.377$), there being no significant difference between males and females in the odds with which (ŋ) is alveolarized.

Style was significant, with more formal speech styles disfavouring [alveolar] realization. The odds with which [alveolar] was produced decreased across styles from passage reading (odds-ratio with respect to 'casual' speech style: 0.473, $p=5.425 \times 10^{-3}$), to word list (odds-ratio with respect to 'casual' speech style: 0.348,

$p = 1.050 \times 10^{-4}$), but increased again in the minimal pairs task (odds-ratio with respect to 'casual' speech style: 0.511, $p = 3.8707 \times 10^{-2}$). In other words, it was, respectively, 52.7%, 65.2%, and 48.9% *less likely* for (ŋ) to be produced as [alveolar], in the passage reading, word list reading, and minimal pairs reading tasks, than in casual speech style.

Table 39. Mixed-effects multinomial logistic regression model results: (ŋ) > [n] variation (part 1)

sr1 **Response Ref. Cat.: [velar]**	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
Fixed Part[$]						
- Intercept#	-1.131	0.366	-3.090	2.000×10^{-3}	0.323	-67.7%
External factors						
- Age	-0.042	0.006	-7.000	2.560×10^{-12}	0.959	-4.1%
- Gender (Female)	0.313	0.354	0.884	0.377	1.368	36.8%
- Style (ref.: Casual Speech)						
Passage Reading	-0.748	0.269	-2.781	5.425×10^{-3}	0.473	-52.7%
Word List	-1.055	0.272	-3.879	1.050×10^{-4}	0.348	-65.2%
Minimal Pairs	-0.672	0.324	-2.074	3.807×10^{-2}	0.511	-48.9%
Internal factors						
- Vowel (Ref.: ɐ)						
ɛː	-0.858	0.269	-3.190	1.425×10^{-3}	0.424	-57.6%
ɪ	-4.351	0.354	-12.291	1.013×10^{-34}	0.013	-98.7%
ʊ	-5.253	0.457	-11.495	1.405×10^{-30}	0.005	-99.5%
ɔː	-1.024	0.209	-4.900	9.607×10^{-7}	0.359	-64.1%
- Tone (Ref.: T1)						
T4	-0.274	0.104	-2.635	8.423×10^{-3}	0.760	-24.0%
T6	0.311	0.142	2.190	2.851×10^{-2}	1.36 5	36.5%
- Following Onset (Ref.: Velar)						
Labial	1.852	0.255	7.263	3.793×10^{-13}	6.373	537.3%
Palatal	3.937	0.237	16.612	5.724×10^{-62}	51.265	5026.5%
Alveolar	3.371	0.284	11.870	1.700×10^{-32}	29.108	2810.7%
Other	2.567	0.242	10.607	2.752×10^{-26}	13.027	1202.7%
- Lexical Frequency (LexFreq)						
log(freq): mean-centered	-0.306	0.264	-1.159	0.246	0.736	-26.4%
- LexFreq * Vowel(ɛː vs. ɐ)	-0.651	0.220	-2.960	3.085×10^{-3}	0.522	-47.8%
- LexFreq * Vowel(ɪ vs. ɐ)	-0.884	0.298	-2.966	3.013×10^{-3}	0.413	-58.7%
- LexFreq * Vowel(ɔː vs. ɐ)	-0.314	0.125	-2.512	1.200×10^{-2}	0.731	-26.9%
- LexFreq * Vowel(œː vs. ɐ)	-0.284	0.142	-2.000	4.550×10^{-2}	0.753	-24.7%
- LexFreq * Following Onset(Alveolar vs. Velar)	0.968	0.256	3.781	1.560×10^{-4}	2.633	163.3%
Random Part						
- Intercept variance	1.729	0.159	10.874	1.530×10^{-27}	5.635	463.5%

[$] Mostly those factors and interactions of factors which were significant are reported here to draw attention to the important factors. For all results, please refer to 0.

Intercept was included in the random part of the model.

It was found that preceding vowels were significant in predicting (**ŋ**) to [alveolar] realization. In this model, there was no statistically significant difference between, respectively, the preceding vowels **a**ː and **œ**ː, compared to the reference vowel **ɐ**, in predicting (**ŋ**) to [alveolar] realization, but overall, the vowels **a**ː, **ɛ**ː, **œ**ː and **ɔ**ː, were

significantly different from the vowels **ɪ** and **ʊ**, which highly disfavoured [alveolar] realization, with respect to the reference category **ɐ**. The model confirmed that the preceding vowels **a**:, **œ**: and **ɐ** are the most progressive in [alveolar] realization. The following onset was found to be highly statistically significant in explaining [alveolar] realization.

When (**ŋ**) was followed by an onset segment with the feature 'alveolar', the odds with which (**ŋ**) was alveolarized was 29.1 times ($p = 1.70 \times 10^{-32}$) that when the feature of the following onset is 'velar'. Other following onset features, namely 'labial', 'palatal' and 'other' (which includes pauses, such as phrase boundaries; and zero following onsets, where following syllables do not have an onset consonant—see Section 2.3.2 on the Cantonese syllable structure), were also more likely (at 6.4, 51.3, and 13.0 times, and $p = 3.793 \times 10^{-13}$, 5.724×10^{-62}, 2.752×10^{-26}, respectively) to favour [alveolar] realization than a following 'velar' onset.

As for lexical frequency, there was no significant effect of log(frequency) in predicting (**ŋ**) to [alveolar] realization.

Next, let us explore the second portion of the model, which estimates the effects of included factors on (**ŋ**) > [reduced] realization. The results are summarised in Table 40. From the estimation results, it can be seen that age was estimated to be highly significant ($p = 7.86 \times 10^{-17}$) in predicting (**ŋ**) > [reduced] variation. With every 1-year increase in age, it was 2.5% *less* likely for (**ŋ**) to be realized with the [reduced] variant. Gender was this time found to be significant in predicting (**ŋ**) > [reduced] variation ($p = 2.268 \times 10^{-3}$), with the odds of females overall showing [reduced] (as opposed to [velar]) variation 1.89 times that for males.

Table 40. Mixed-effects multinomial logistic regression modelling results: (ŋ) > [reduced] variation (part 2)

sr2 **Response Ref. Cat.: [velar]**	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
Fixed Part[$]						
- Intercept#	-1.625	0.223	-7.287	3.090×10^{-13}	0.197	-80.3%
External factors						
- Age	-0.025	0.003	-8.333	7.860×10^{-17}	0.975	-2.5%
- Gender (Female)	0.638	0.209	3.053	2.268×10^{-3}	1.893	89.3%
- Style (ref.: Casual Speech)						
Passage Reading	-1.676	0.159	-10.541	5.597×10^{-26}	0.187	-81.3%
Word List	-2.272	0.159	-14.289	2.551×10^{-46}	0.103	-89.7%
Minimal Pairs	-3.702	0.295	-12.549	4.018×10^{-36}	0.025	-97.5%
- Age * Gender(Female)	-0.025	0.005	-5.000	5.733×10^{-7}	0.975	-2.5%
- Gender(Female) * Style(PR)	0.503	0.212	2.373	1.766×10^{-2}	1.654	65.4%
- Gender(Female) * Style(Word List)	0.808	0.202	4.000	6.334×10^{-5}	2.243	125.0%
Internal factors						
- Vowel (Ref.: ɐ)						
aː	1.421	0.167	8.509	1.755×10^{-17}	4.141	314.1%
ɛː	1.134	0.175	6.480	9.172×10^{-11}	3.108	210.8%
ɪ	-1.614	0.194	-8.320	8.827×10^{-17}	0.199	-80.1%
ɔː	1.317	0.151	8.722	2.737×10^{-18}	3.732	273.2%
œː	1.544	0.158	9.772	1.483×10^{-22}	4.683	368.3%
ʊ	-0.767	0.175	-4.383	1.171×10^{-5}	0.464	-53.6%
- Tone (Ref.: T1)						
T4	0.313	0.076	4.118	3.815×10^{-5}	1.368	36.8%
- Following Onset (Ref.: Velar)						
Labial	3.080	0.143	21.538	6.793×10^{-103}	21.758	2075.8%
Palatal	2.954	0.177	16.689	1.569×10^{-62}	19.183	1818.3%
Alveolar	1.905	0.142	13.415	4.907×10^{-41}	6.719	571.9%
Other	1.811	0.143	12.664	9.321×10^{-37}	6.117	511.7%
- Lexical Frequency (LexFreq)						
log(freq): mean-centered	0.526	0.174	3.023	2.501×10^{-3}	1.692	69.2%
- LexFreq * Vowel(ɛː vs. ɐ)	-0.615	0.166	-3.705	2.115×10^{-4}	0.541	-45.9%
- LexFreq * Vowel(ɔː vs. ɐ)	-0.591	0.130	-4.546	5.464×10^{-6}	0.554	-44.6%
- LexFreq * Following Onset(Other vs. Velar)	-0.354	0.149	2.376	1.751×10^{-2}	0.702	-29.8%
Random Part						
- Intercept variance	0.459	0.059	7.780	7.252×10^{-15}	1.582	58.2%

[$] Only those factors and interactions of factors which were significant are reported here to draw attention to the important factors. For all results, please refer to 0.
Intercept was included in the random part of the model

Style was also significant, with more formal speech styles disfavouring [reduced] realization. The odds with which [reduced] was produced decreased across styles from passage reading (odds-ratio with respect to 'casual' speech style: 0.187, $p = 5.597 \times 10^{-26}$),

to word list (odds-ratio with respect to 'casual' speech style: 0.103, p 2.551 = × 10^{-46}, respectively), and finally to the minimal pairs task (odds-ratio with respect to 'casual' speech style: 0.025, p = 4.018 × 10^{-36}). In other words, it was, respectively, 81.3%, 89.7%, and 97.5% less likely for (**ŋ**) to be produced as [reduced] than in casual speech style, in the passage reading, word list reading, and minimal pairs reading tasks.

It was found that preceding vowels were significant in predicting (**ŋ**) to [reduced] realization. There was a statistically significant difference between the preceding vowels **a**:, **ɛ**:, **ɪ**, **ɔ**:, **œ**:, and **ʊ**, with respect to the reference vowel **ɐ**, in predicting (**ŋ**) to [reduced] realization (p = 1.755 × 10^{-17}, 9.172 × 10^{-11}, 8.827 × 10^{-17}, 2.737 × 10^{-18}, and 1.483 × 10^{-22}, respectively). The vowels **ɪ** and **ʊ** highly disfavoured [reduced] realization, with respect to the reference category **ɐ**, while the vowels **a**:, **ɛ**:, **ɔ**: and **œ**: were, respectively, 4.13, 3.11, 3.73, and 4.68 times more likely to favour [reduced] realization over the reference vowel **ɐ**. For a further discussion on the effects of preceding vowels discussion, refer to Section 6.1.

Following onset was also found to be highly statistically significant. When (**ŋ**) was followed by an onset segment with the feature 'labial' and 'palatal', the odds with which (**ŋ**) was [reduced] were 21.8 (p = 6.793 × 10^{-103}) and 19.2 times (p = 1.569 × 10^{-62}) the odds compared to when the feature of the following onset is 'velar'. Other following onset features, namely 'alveolar' and 'other' (which includes pauses and zero following onsets), were also more likely (at 6.7, and 6.1 times, p = 4.91 × 10^{-41}, 9.32 × 10^{-37}, respectively) to favour [reduced] realization than a following 'velar' onset.

As for lexical frequency, measured as mean-centered 'log(frequency)', it was significant for predicting (**ŋ**) > [reduced] variation (p = 2.501 × 10^{-3}), with an odds-ratio

of 1.69. In other words, all else being equal, with every 1-unit increase in 'log(frequency)', (**ŋ**) was 1.69 times *more likely* to be realized as [reduced].

The implications of social factors, neogrammarian factors, and lexical frequency effects on (**ŋ**) realization will be further explored in Chapters 5, 6 and 7. In the next section, the full model for (**k**) realization is presented and discussed.

4.3 Full model for (**k**) realization

Compared with the model for (**ŋ**) above, fewer terms could be included in the model for (**k**), as the model estimation could not complete when random factors and interaction between lexical frequency and other factors were added. A screenshot of the "best" model output from MLwiN is shown in Model 2. in 0. A summary of the significant factors is provided in Tables 41 and 42, showing (**k**) > [alveolar] variation, and (**k**) > [reduced] variation respectively. All the factors which were significant were included in these tables. As before, those factors which were insignificant but of interest were still included in the tables, if they merit discussion. First, let us examine Table 41, which summarises the model portion for (**k**) > [alveolar] variation, with the category [velar] taken as the reference response category.

From the estimation results, it can be seen that age was estimated to be highly significant in predicting (**k**) > [alveolar] variation ($p = 7.062 \times 10^{-32}$). With every 1-year increase in age, it was 4.6% *less* likely for (**k**) to be realized with [alveolar].

Gender was also found to be significant in predicting this variation ($p = 3.848 \times 10^{-29}$), there being a highly significant difference between males and females in the odds with which (**k**) is alveolarized, with the odds for females being 13.8 times that for males.

Style was also significant, with more formal speech styles disfavouring [alveolar] realization. The odds with which [alveolar] was produced decreased across styles from passage reading (odds-ratio with respect to 'casual' speech style: 0.419, $p = 4.45 \times 10^{-4}$), to word list (odds-ratio with respect to 'casual' speech style: 0.180, $p = 2.152 \times 10^{-11}$), but increased again in the minimal pairs task (odds-ratio with respect to 'casual' speech style: 0.399, $p = 4.775 \times 10^{-5}$). In other words, it was, respectively,

58.1%, 82.0%, and 60.1% less likely for (**k**) to be produced as [alveolar] in the passage reading, word list reading, and minimal pairs reading tasks, than in casual speech style.

Table 41. Mixed-effects multinomial logistic regression model results: (k) > [t] variation (part 1)

sr1 **Response Ref. Cat.: [velar]** [$]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
Fixed Part						
- Intercept	-1.854	0.411	-4.511	6.454×10^{-6}	0.157	-84.3%
External factors						
- Age	-0.047	0.004	-11.750	7.062×10^{-32}	0.954	-4.6%
- Gender (Female)	2.622	0.234	11.205	3.848×10^{-29}	13.763	1276.3%
- Style (ref.: Casual Speech)						
Passage Reading	-0.871	0.248	-3.512	4.45×10^{-4}	0.419	-58.1%
Word List	-1.741	0.256	-6. 695	2.152×10^{-11}	0.180	-82.0%
Minimal Pairs	-0.919	0.226	-4.066	4.775×10^{-5}	0.399	-60.1%
- Age * Gender(Female)	-0.045	0.007	-6.429	1.288×10^{-10}	0.956	-4.4%
- Gender(Female) * Style(MP)	-3.255	0.419	-7.768	7.942×10^{-15}	0.039	-96.1%
- Age * Gender(Female) * Style(MP)	0.080	0.010	8.000	1.244×10^{-15}	1.083	8.3%
Internal factors						
- Vowel (Ref.: ɐ)						
ʊ	-7.596	1.252	-6.067	1.302×10^{-9}	5.025×10^{-4}	-99.9%
ɪ	-7.706	0.782	-9.854	6.573×10^{-23}	4.501×10^{-4}	-99.9%
ɔː	-1.300	0.113	-11.504	1.253×10^{-30}	0.273	-72.7%
œː	-0.761	0.239	-3.184	1.453×10^{-3}	0.467	-53.3%
- Following Onset (Ref.: Velar)						
Labial	2.963	0.397	7.463	8.426×10^{-14}	19.356	1835.6%
Palatal	5.915	0.443	13.352	1.151×10^{-40}	370.554	36955.4%
Alveolar	5.301	0.386	13.733	6.427×10^{-43}	200.537	19953.7%
Other	3.756	0.389	9.656	4.458×10^{-22}	42.777	4177.7%
- Lexical Frequency (LexFreq)						
log(freq): mean-centered	-1.311	0.612	-2.142	0.032	0.270	-73.0%
- LexFreq(Frequent) * Following Onset(Labial vs. Velar)	2.413	0.627	3.848	1.189×10^{-4}	11.167	1016.7%
- LexFreq(Frequent) * Following Onset(Palatal vs. Velar)	2.232	0.656	3.402	6.679×10^{-4}	9.318	831.848
- LexFreq(Frequent) * Following Onset(Alveolar vs. Velar)	1.857	0.619	3.000	2.700×10^{-3}	6.404	540.4%
- LexFreq(Frequent) * Following Onset(Other vs. Velar)	1.363	0.615	2.216	2.667×10^{-2}	3.908	290.8%

[$] Only those factors and interactions of factors which were significant are reported here to draw attention to the important factors. For all results, please refer to 0.

It was also found that preceding vowels were significant in predicting (**k**) to [alveolar] realization. In this model, there was a statistically significant difference

between the preceding vowels **ʊ**, **ɪ**, **ɔ**: and **œ**: (p = 1.302 × 10^{-9}, 6.573 × 10^{-23}, 1.253 × 10^{-30}, and 1.453 × 10^{-3}, respectively) with respect to the reference vowel **ɐ**, in predicting (**k**) to [alveolar] realization. The vowels **ʊ**, **ɪ**, **ɔ**: and **œ**: were far less likely to favour [alveolar] realization than **ɐ** (99.9%, 99.9%, 72.7%, and 53.3% less likely, respectively).

Following onset feature was also found to be highly statistically significant. When (**k**) was followed by an onset segment with the feature 'alveolar', the odds with which (**k**) was alveolarized was 200.5 times the odds than when the feature of the following onset is 'velar' (p = 6.427 × 10^{-43}). Other following onset features, namely 'labial', 'palatal' and 'other' (which includes pauses and zero following onsets), were also more likely (at 19.4, 370.6, and 42.8 times, p = 8.426 × 10^{-14}, 1.151 = × 10^{-40}, and 4.458 × 10^{-12}, respectively) to favour [alveolar] realization than a following 'velar' onset.

As for lexical frequency, measured as mean-centered 'log(frequency)', it was significant for predicting (**k**) > [alveolar] variation (p = 3.2 × 10^{-2}), with an odds-ratio of 0.27. In other words, all else being equal, with every 1 unit increase in 'log(frequency)', (**k**) in lexemes were 0.27 times as likely, in other words, 73% *less* likely, to be realized as [alveolar] than as [velar].

Next, let us explore the second portion of the model, which estimates the effects of included factors on (**k**) > [reduced] realization. The results are summarised in Table 42. From the estimation results, it can be seen that age was estimated to be non-significant in predicting (**k**) > [reduced] variation. Gender was this time estimated to be highly significant in predicting (**k**) > [reduced] variation (p = 1.749 × 10^{-17}), with the odds of

females overall showing [reduced] (as opposed to [velar]) variation 5.9 times that for males.

Style was also significant, with more formal speech styles disfavouring [reduced] realization. The odds with which [reduced] was produced changed across styles from passage reading (odds-ratio with respect to 'casual' speech style: 0.203, $p = 3.433 \times 10^{-23}$), to word list (odds-ratio with respect to 'casual' speech style: 0.212, $p = 2.424 \times 10^{-27}$), and finally to the minimal pairs task (odds-ratio with respect to 'casual' speech style: 0.068, $p = 2.834 \times 10^{-56}$). In other words, it was, respectively, 79.7%, 78.8%, and 93.2% *less* likely for (**k**) to be produced as [reduced] in the passage reading, word list reading, and minimal pairs reading tasks, than in casual speech style.

Table 42. Mixed-effects multinomial logistic regression modelling results: (k) > [reduced] variation (part 2)

sr2 **Response Ref. Cat.: [velar]** [$]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
Fixed Part						
- Intercept	-1.219	0.223	-5.466	4.594×10^{-8}	0.296	-70.4%
External factors						
- Age	0.001	0.003	0.333	0.739	1.001	0.1%
- Gender (Female)	1.770	0.208	8.510	1.745×10^{-17}	5.871	487.1%
- Style (ref.: Casual Speech)						
Passage Reading	-1.597	0.161	-9.919	3.433×10^{-23}	0.203	-79.7%
Word List	-1.549	0.143	-10.832	2.424×10^{-27}	0.212	-78.8%
Minimal Pairs	-2.687	0.170	-15.806	2.834×10^{-56}	0.068	-93.2%
- Age * Gender(Female)	-0.034	0.005	-6.800	1.046×10^{-11}	0.967	-3.3%
- Gender(Female) * Style(PR)	0.853	0.338	2.524	1.161×10^{-2}	2.347	134.7%
- Gender(Female) * Style(WL)	0.675	0.282	-2.394	1.668×10^{-2}	1.964	96.4%
- Age*Gender(Female)*Style(RP)	-0.036	0.008	-4.500	6.795×10^{-6}	0.965	-3.5%
- Age*Gender(Female)*Style(Word	-0.028	0.007	-4.000	6.334×10^{-5}	0.972	-2.7%
Internal factors						
- Vowel (Ref.: ɐ)						
ɛː	0.682	0.184	3.707	2.101×10^{-4}	1.978	97.8%
ʊ	-1.254	0.118	-10.627	2.229×10^{-26}	0.285	-71.5%
ɪ	-2.094	0.117	-17.897	1.235×10^{-71}	0.123	-87.7%
aː	0.924	0.106	8.717	2.857×10^{-18}	2.519	151.9%
ɔː	0.554	0.100	5.540	3.025×10^{-8}	1.740	74.0%
œː	0.465	0.192	2.422	1.544×10^{-2}	1.592	59.2%
- Following Onset (Ref.: Velar)						
Labial	3.952	0.185	21.362	3.006×10^{-101}	52.039	5103.9%
Palatal	3.750	0.271	13.838	1.511×10^{-43}	42.521	4152.1%
Alveolar	2.434	0.176	13.829	1.691×10^{-43}	11.404	1040.4%
Other	3.853	0.177	21.768	4.629×10^{-105}	47.134	4613.4%
- Lexical Frequency (Ref.: Infreq.)						
log(freq): mean-centered	0.492	0.162	3.037	2.389×10^{-3}	1.635	63.6%
- LexFreq(Frequent) * Following Onset(Labial vs. Velar)	-0.389	0.181	-2.149	3.162×10^{-2}	0.678	-32.2%
- LexFreq(Frequent) * Following Onset(Palatal vs. Velar)	0.870	0.285	3.053	2.268×10^{-3}	2.387	138.7%
LexFreq(Frequent) * Following Onset(Other vs. Velar)	-0.360	0.168	-2.143	3.212×10^{-2}	0.698	-30.2%

[$] Only the those factors and interactions of factors which were significant are reported here to draw attention to the important factors. For all results, please refer to 0.

It was found that preceding vowels were significant in predicting (**k**) to [reduced] realization. There was a statistically significant difference between the preceding vowels **ɛː**, **ʊ**, **ɪ**, **aː**, **ɔː** and **œː** with respect to the reference vowel **ɐ**, in predicting (**k**) to [reduced] realization. The vowels **ɪ** and **ʊ** highly disfavoured [reduced] realization, with respect to the reference category **ɐ** (at 2.0, 0.3, 0.1, 2.5, 1.7, and 1.6 times odds-ratio,

$p = 2.101 \times 10^{-4}$, 2.229×10^{-26}, 1.235×10^{-71}, 2.857×10^{-18}, 3.025×10^{-8}, and 1.544×10^{-2}, respectively), while the vowels **a**:, **ɛ**:, **œ**: and **ɔ**: were more likely to favour [reduced] realization over the reference vowel **ɐ**.

Following onset was found to be highly statistically significant in explaining [reduced] variation of (**k**). When (**k**) was followed by an onset segment with the feature 'labial', 'palatal', 'alveolar' and 'other' the odds with which (**k**) was reduced were 52.0, 42.5, 11.4 and 47.1 times the odds when the feature of the following onset is 'velar' ($p = 3.006 \times 10^{-101}$, 1.511×10^{-43}, 1.691×10^{-43} and 4.629×10^{-105}, respectively).

As for lexical frequency, measured as mean-centered 'log(frequency)', it was significant for predicting (**k**) > [reduced] variation ($p = 2.389 \times 10^{-3}$), with an odds-ratio of 1.64. In other words, all else being equal, with every 1-unit increase in 'log(frequency)', (**k**) in lexemes were 1.64 times as likely, in other words, 63.6% more likely, to be realized as [reduced].

Finally, it is worth noting that the effects and significance values obtained in this model for (**k**) realization are noticeably higher than those for the model for (**ŋ**) realization. This is due to the fact that this model did not contain any random parts, and as such, did not control for random variation within the speaker level. Typically, models which include only fixed effects will return estimates which are numerically higher than mixed-effects models where group variation is taken into account.

As with (**ŋ**), the implications of social factors, neogrammarian factors, and lexical frequency effects for (**k**) realization will be further explored in Chapters 5, 6 and 7.

4.4 Summary and discussion

In this chapter, full statistical models were presented that compared the effects of all the factors discussed in this thesis on variation in realization of (**ŋ**) and (**k**) as [velar], [alveolar], and [reduced], respectively. These factors included social factors of 'age', 'gender', and 'style', phonological and phonetic context factors of 'preceding vowel', 'tone' and 'following onset', as well the usage-based factor of 'lexical frequency'. Including all these factors into the same model allowed their relative effects with respect to one another to be compared one at a time, with all other factors being held equal at the same time.

Starting with the factors which were consistently found to be significant in predicting (**ŋ**) and (**k**) realization, these were preceding vowel categories and following onset features. For (**ŋ**), the vowel categories **ɪ** and **ʊ** were found to consistently favour [velar] retention, while the vowel categories **ɐ**, **a**:, **ɛ**:, **œ**: and **ɔ**: favoured [alveolar] and [reduced] realizations to varying extents, but were far more likely to do so than the vowels **ɪ** and **ʊ**. The likelihood with which vowel category can predict (**k**) realization as [alveolar] or [reduced] is slightly less clear-cut, with a large amount of variation between vowels. Following onset was in the case of both (**ŋ**) and (**k**) found to be good at predicting [alveolar] and [reduced] realizations. The following onset feature 'alveolar' was found to be highly significant in predicting (**ŋ**) and (**k**) realization as [alveolar], compared to the following onset feature 'velar'. For (**ŋ**), it was found that the following onset features 'labial', 'palatal' and 'other' highly significantly favoured [alveolar] and [reduced] realizations, compared to the following onset feature 'velar'. For (**k**), the following onset features 'labial', 'alveolar', 'palatal', and 'other', were all

found to significantly more likely to favour [alveolar] and [reduced] realizations than the baseline feature 'velar' did.

As for social factors, age was found to be significant in explaining (**ŋ**) to [alveolar] and [reduced] variation, as well as in (**k**) to [alveolar] realization, but not (**k**) to [reduced] realization. Where age was significant, the effect was such that for every increase of one year in age, the odds of [alveolar] or [reduced] realization was between 2.5% to 4.6% less. Gender was not found to be significant in terms of predicting (**ŋ**) alveolarization, but was found to be significant in the case of predicting (**ŋ**) to [reduced], and highly significant in predicting (**k**) to [alveolar] and [reduced] realizations.

Style was also found to be highly significant in predicting (**ŋ**) and (**k**) realization. All formal tasks (passage, word list, and minimal pairs reading) were found to highly significantly disfavour [alveolar] and [reduced] realizations, compared to casual speech style. However, in the case of predicting (**ŋ**) to [alveolar] realization and (**k**) to [alveolar] realization, the minimal pairs task was found to be anomalous, in the sense that the odds with which it disfavoured [alveolar] realizations was lower than expected, if one were to assume that the minimal pairs task was the most formal speech style and would draw the most attention from speakers to their own speech.

Last but not least, lexical frequency, measured as mean-centered 'log(frequency)', was found to significantly account for (**ŋ**) to [reduced] realizations, and (**k**) to [alveolar] and [reduced] realizations. Overall, it was found that for (**ŋ**), frequent lexical items were more likely than infrequent lexical items to be realized as [reduced], in other words, to undergo lenition, at a rate of 1.69 times for every unit increase in 'log(frequency)'. For (**k**), frequent lexical items were 0.27 times as likely (i.e. 73% *less* likely) to be

realized as [alveolar], for every unit increase in 'log(frequency)'. Tokens containing (**k**) were also 1.64 times more likely to have [reduced] realization for every unit increase in 'log(frequency)'. These results support the predictions put forward by usage-based approaches that more frequent lexical items are more likely to undergo leniting sound change. The rather curious result that (**k**) being less likely realized as [alveolar] with an increase in 'log(frequency)', and the implications of this and other results on lexical factors will be discussed in Chapter 7.

5 Social dimensions of sound change

In this chapter, the social factors which were shown in the full statistical model to be significant in explaining the outcome of (**ŋ**) and (**k**) in the preceding chapter are described in further detail. The social, or external, factors which will be discussed include age, gender, and speech style. A summary is offered of the significant findings presented in Chapter 4 above with respect to these social variables, and interpretations of these results are offered, with reference to the predictions presented in the literature review in Chapters 1 and 2, as well as methodological considerations discussed in Chapter 3. Based on these results, univariate statistics and graphs are also presented to obtain some insight into how the social factors contribute to the variation in velar coda realization.

5.1 Age-based differences

5.1.1 Age-based effects on (**ŋ**) variation

The results of the full mathematical model presented in Chapter 4 show that age was a significant factor when it comes to describing the variation in realization of (**ŋ**). Table 43 shows only those results in (Tables 39 and 40) which are related to the "age" factor.

Table 43. Multinomial logistic regression results for age-based variation of (ŋ)

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
External factors						
(ŋ) > [n] variation						
- Age	-0.042	0.006	-7.000	2.560×10^{-12}	0.959	-4.1%
(ŋ) > [reduced] variation						
- Age	-0.025	0.003	-8.333	7.860×10^{-17}	0.975	-2.5%

It can be seen from the results that age was estimated to be highly significant for predicting both (**ŋ**) to [alveolar] variation ($p = 2.56 \times 10^{-12}$) as well as (**ŋ**) to [reduced] variation ($p = 7.86 \times 10^{-17}$). With every one-year increase in age, it was 4.1% *less* likely for (**ŋ**) to be realized with [alveolar] and 2.5% *less* likely for (**ŋ**) to be realized with the [reduced] variant.

It was suggested in studies reviewed in Chapter 2 that the innovative alveolar nasal was increasingly adopted by younger speakers of the Hong Kong speech community. As discussed in Section 2.4, Wong (2005) showed a clear age-related effect, whereby alveolar realization of velar nasals was displayed much more by the youngest group of speakers aged 11-20 (at 24.14%) than the older speakers aged 31-40 (11.30%) and aged over 41 (4.94%), notwithstanding an anomaly whereby the group of speakers 21-30 years of age showed an unexpectedly low rate of alveolarization (6.19%).

If alveolarization is a change in progress as the age-related variation patterns found by Wong (2005) suggests, we would expect the sound change to have progressed even further a few years down the line, with speakers in their 20s and 30s (at the time of this study) showing equally progressive use of innovative variants as the speakers in the age group 11-20 in Wong's study. To offer a comparison, the extent of interspeaker variation in this study is shown in Figure 18, which displays the percentage realization of the velar nasal coda (**ŋ**) as [velar], [alveolar] or [reduced], by individual speaker ranked by increasing age. The percentage for each type of realization is indicated by trend lines: green line with diamonds for [velar], red line with squares for [alveolar], and yellow line with inverted triangles for [reduced].

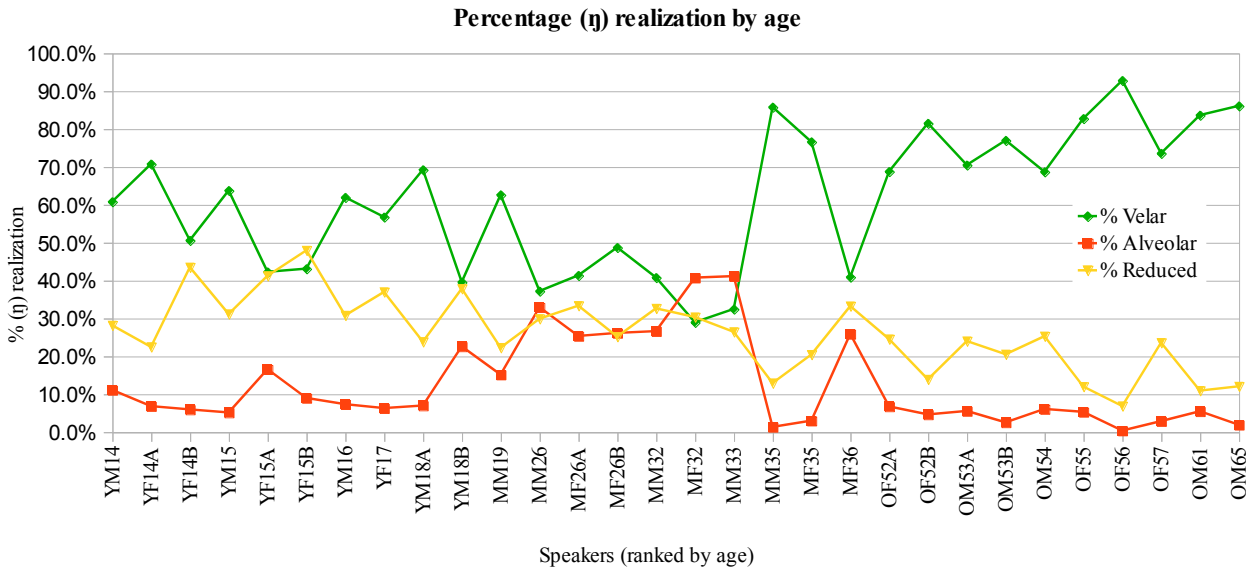


Figure 18. Percentage (ŋ) realization by speaker ranked by age

N= 8392

As can be seen, there is a discernible amount of interspeaker variation. One particularly noticeable variation pattern stands out, which is that speakers in their 20s and 30s display a markedly lower percentage of [velar] realization as well as the highest percentage of [alveolar] realization, with the exception of informants MM35 and MF35[33]. This group of speakers in their 20s and 30s would be in the same cohort as Wong's (2005) speakers who are aged 11-20. When one furthermore takes into account (**ŋ**) realization as reduced variants, it seems overall that the decline in use of [ŋ] is compensated for by increased percentages of either alveolar [n] or reduced variants. Where the younger age group showed a dip in their use of [n], this was partly compensated for by their increased use of reduced variants. In all, the younger speakers

[33] We can see that these two outliers behave distinctly differently from the other speakers in their 20s and 30s. These two speakers both display patterns very similar to the older speakers, with low rates of [alveolar] and [reduced] realizations, and high rates of [velar] realizations. The departure in their speech behaviour will be discussed in more detail in the discussion below.

did show innovation, not in use of [n] as expected, but by using an increased proportion of reduced variants. The overall trend is summarised in Figure 19 as a conditional probability density curve, showing how percentage use of (**ŋ**) on the y-axis changes conditionally upon the 'age' variable as a continuous measure. The graph was plotted using the `cdplot` function (written by Achim Zeileis) in the R package `ggplot2` (Wickham & Chang 2013). Function smoothing was obtained by `cdplot` through kernel density estimation in the `density` function (R Core Team and contributors 2013).

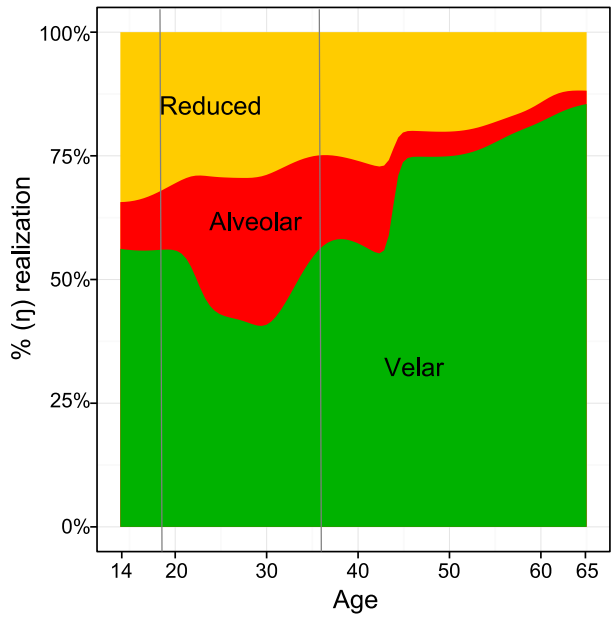


Figure 19. Conditional density curve of (ŋ) realization by age

Data from 30 speakers; total number of tokens, N = 8392

It can be seen from this univariate plot that the most innovative speech characteristics can be found in the middle age group speakers, who showed the highest

proportion of alveolar realization. While the younger speakers retain more velars than the middle age group speakers, they also showed the highest proportion of reduced variants.

Let us turn our attention to (**ŋ**) realization as [alveolar]. Assuming that any differences in methodology are insignificant, the finding here seem remarkably consistent with Wong's (2005) finding that this particular cohort is leading the change in terms of use of the innovative alveolar variant for (**ŋ**). In interpreting these age-related differences, it would be useful to appeal to Labov's (1994:83) categorization of sound changes at the level of the individual and at the level of the community into four possible scenarios, summarised in Table 44.

Table 44. Types of sound changes at the level of the individual vs. community (Labov 1994:83)

	Change across time within:	
	Individual	**Community**
1. Stability	×	×
2. Age-grading	√	×
3. Generational change	×	√
4. Communal change	√	√

In the first scenario, there is stability at both the individual and community level, with neither showing any change. In the second scenario, there is age-grading, which means that, as an individual progresses through life, they change their speech behaviour to "reflect" the generation of speakers they are joining at a given stage. However, the community as a whole remains stable. In the third case, there is generational change, meaning the individual, as they go through life, remains unchanged, maintaining the same speech behaviour they displayed when they were children or adolescents; yet the

community as a whole changes. As each generation of the youngest speakers develops new speech behaviour, the whole community consists of a collection of generational speech behaviours that marks each given generation. This third scenario is perhaps more commonly known as the "apparent-time" interpretation (Sankoff 2004). Finally, there is communal change, where both the individual as well as the community change. As each speaker goes through life, they change their speech behaviour throughout. Yet new generations of speakers also develop new speech patterns not seen before, such that there is also change on the community level. While these are logically possible combinations of changes that can happen on the individual and community level, age-grading and the apparent-time interpretation have both been considered as the most plausible when variation across age groups is found in a synchronic study.

> In making the choice between apparent time and age grading in the absence of reliable temporal benchmarks, sociolinguists displayed appropriate caution. Studies in which an apparent time interpretation was invoked usually focused on those aspects of language least subject to conscious manipulation or metalinguistic attention on the part of speakers—phonology rather than lexicon, for example. Researchers were also careful to point out that there might be an effect of age grading combined with change in progress. Within the domain of phonology, and with all these caveats, most studies tended to take apparent time as the default interpretation.
>
> (Sankoff 2004)

Going back to the comparison between Wong (2005) and this study, the fact that Wong's speakers aged 11-20 approximately belong to the same cohort as speakers in their 20s and 30s in this study, and that the speech behaviour of this cohort showed the most progressive use of alveolarization of (ŋ) within the speech community suggests that the apparent-time interpretation seems most likely for this sound change. In other

words, there is evidence for generational change, and each individual within this cohort as they go through life, remains unchanged, maintaining the same speech behaviour they displayed when they were children or adolescents. However, the community as a whole changes, and as described above, the generation of the youngest speakers develops new speech behaviour, the whole community consisting of a collection of generational speech patterns that marks each different generation.

Without a panel study tracing speakers throughout their lifetimes, it is not possible to ascertain whether age-related variation patterns found at a synchronic point in time is a case of age grading or apparent-time generational change. But it is generally believed that once speakers reach adulthood, linguistic features which are not socially marked will remain fairly stable, such that one can indeed appeal to the apparent-time interpretation. Furthermore, there are studies (e.g. Bailey et al. 1991, Trudgill 1988) which have re-examined the same speech community 10 to 15 years afterwards that have confirmed the apparent-time interpretation of the original age-varied data. However, Sankoff's (2007) study of the change of apical to dorsal **r** in Montreal French found that while most speakers did not change through adulthood, there was a substantial (but still minority) portion of speakers who changed to dorsal **r** in their adulthood, suggesting that if a phonological variable is above speakers' social awareness, it is still possible that the feature is subject to change throughout adulthood.

As for (**ŋ**) to [reduced] variation, Figure 18 shows that there is a steady increase in [reduced] realization (yellow line with inverted triangles) when moving from older to younger speakers, without the anomaly observed in speakers in their 20s and 30s for [alveolar] realization. Since there is no previous study with comparable coding criteria for phonetic realization of (**ŋ**) as [reduced], it is difficult to infer conclusively whether

this progressive increase in [reduced] realization down the age is an indication of sound change in progress through age-grading, whereby each successive generation of speakers makes less and less use of [reduced] realizations as they enter a particular life phase, to more reflect the behaviour of the age group that they belong to. However, if this is indeed age-grading, this would mean that (**ŋ**) as [reduced] realization might be above the level of social awareness, and that speakers are aware that [reduced] realizations are socially less acceptable. This idea will be further explored in Section 5.4, in our discussion of speech style.

Taken together, the age-based variation for alveolarization of (**ŋ**) seems to show generational change, with speakers in their 20s and 30s (at the time of this study) leading the innovation in terms of alveolarization. Overall, when taking the speech community as a whole, the younger the speakers, the more likely they were to make use of the alveolarized [n], with an overall 4.1% increase in likelihood with every one-year decrease in age. As for reduced realizations of (**ŋ**), the younger speakers displayed significantly more lenition, there being a 2.5% increase in likelihood for (**ŋ**) to be realized with the [reduced] variant with every decrease in age.

5.1.2 Age-based effects on (**k**) variation

The results of the full mathematical model presented in Chapter 4 show that age was a significant factor in describing the variation in realization of (**k**). Table 45 displays only those results in (Tables 39 and 40) which are related to age as a lower order factor.

Table 45. Multinomial logistic regression results for age-based variation of (k)

Response Ref. Cat.: [velar]	β (regress. coeff.)	Standard Error (S.E.)	Z-score (β / S.E.)	p-value (2-tailed)	Odds ratios (Exp(β))	% diff. in odds (cf. ref. cat.)
External factors						
(k) > [t] variation						
- Age	-0.047	0.004	-11.750	7.062×10^{-32}	0.954	-4.6%
(k) > [reduced] variation						
- Age	0.001	0.003	0.333	0.739	1.001	0.1%

The results show that while (**k**) to [alveolar] realization shows significant age-based variation, (**k**) to [reduced] realization did not. For (**k**) to [alveolar] realization, there is an overall 4.6% increase in the use of [alveolar] realization with every one year decrease in age, but any such significant effects could not be found for (**k**) to [reduced] realization.

When exploring the interspeaker variation which contributes to overall variation of (**k**) realization, it might be useful to refer to a plot of percentage (**k**) variation by speaker ranked by age. This is displayed in Figure 20. Raw token counts and percentages of (**k**) realization which went into plotting these graphs are not presented here, but are provided in 0 (Table 73), for reference. The overall increase in [alveolar] realization when moving down speaker age can be seen from the red line with square symbols. For (**k**) to [reduced] realization, indicated by the yellow line with inverted triangles, there is too much interspeaker variance for any significant age-based effect to be obtained. It can also be seen that the individual MM35 is a clear outlier not only for speakers of comparable age, but for all speakers. He displayed a very high rate of [velar] retention and amongst the lowest for [alveolar] production for (**k**). His speech patterns are in fact even more conservative than that of speakers from the older age group. As mentioned in Section 5.1.1, he is also one of the two outliers for (**ŋ**) realization. The speech characteristics of outliers and possible reasons for distinctness

from their age group will be examined in more detail further down in the discussion portion of this chapter.

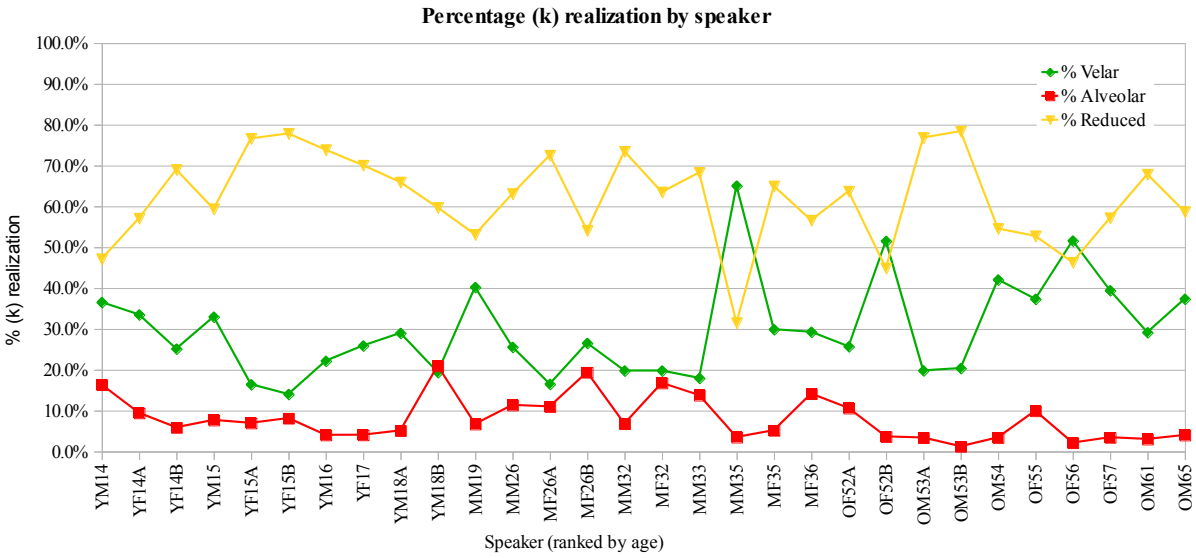


Figure 20. Percentage (k) realization by individual speaker

N=5532

Using the same plotting method as for Figure 19, the conditional probability density curve for (**k**) was also obtained, shown in Figure 21. This plot smoothes out the extreme interspeaker variation and could be an approximate picture of (**k**) realization on a speech community level. It can be seen that for (**k**) the proportion of [reduced] realization is much higher than for (**ŋ**), and that in fact, the proportion of [alveolar] realization is comparatively small.

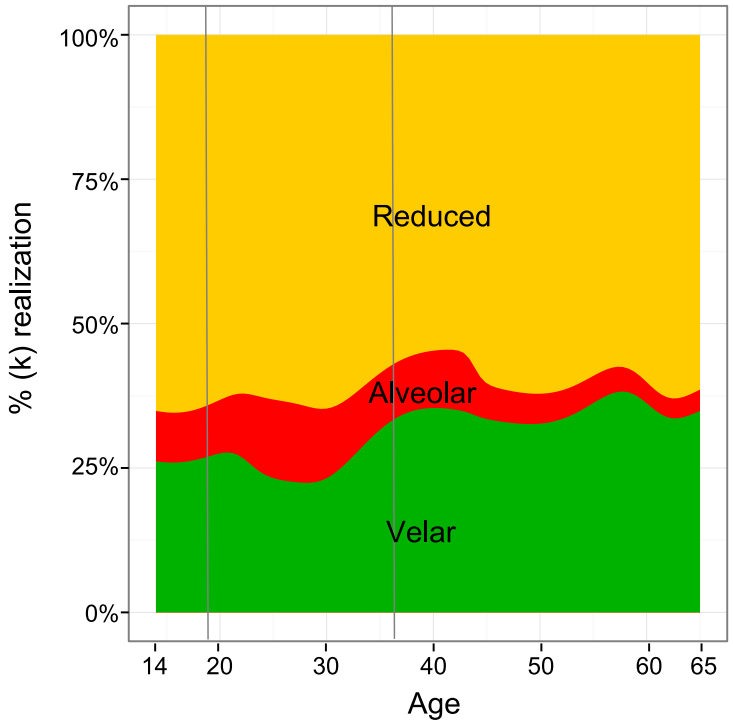


Figure 21. Probability density curve of (k) realization by age

From Figures 20 and 21, it can be seen that overall, the older speakers make more use of conservative [velar] realization, with younger speakers making more use of [alveolar] realization, but that there is a lot of variance of [reduced] realization across speakers of all ages, mirroring results obtained in the full statistical model.

These results are suggestive of a sound change in progress situation, with a progressive increase in the innovative [alveolar] realization by younger speakers of the speech community. Two crucial differences from (**ŋ**), however, are, first, that the amount of difference in [alveolar] realization between the age groups is less distinct, and second, that there is an overall much higher rate of reduced realizations, which shows a lot of variance across speakers without the kind of "age group"-based variation patterns as was observed for (**ŋ**). Altogether, these results suggest that alveolarization of

(**k**) is progressing, with younger speakers making significantly more use of the innovative [alveolar] variant than older speakers.

5.1.3 Summary of age-based variation of (**ŋ**) and (**k**)

Comparing (**ŋ**) with (**k**) variation patterns, two main differences between (**ŋ**) and (**k**) realization were discernible. First, the overall percentage of [velar] retention was much higher for (**ŋ**) than it is for (**k**). Second, while (**ŋ**) shows distinct age group-associated variation, particularly in terms of [alveolar] realization, there is far less distinct age-based variation for (**k**).

Taken together, the overall patterns in use of (**ŋ**) and (**k**) across age groups confirm that there is a reduction in [velar] realization for the younger and middle age groups, when compared to the older age group. However, rather than showing a steady decline in percentage of [velar] realization across age groups, the (**ŋ**) variable shows that the most progressive attrition of [velar] realization occurs in the middle age group, rather than the younger age group, and that this attrition is accounted for by a higher rate of [alveolar] realization. Evidence was found that sound change is occurring throughout the generations and velars give way to alveolars and reduced realizations. The result whereby the speakers in their 20s and 30s were the most progressive in (**ŋ**) to [alveolar] realization tied in with results found in the same cohort of speakers in Wong (2005), and the apparent-time interpretation was invoked to account for this anomalous generation, whereby this particular generation of speakers maintain their speech behaviour (of using a high amount of [alveolar] realization) while contributing to change in the overall community as a whole.

The next sections will look at gender-based variation patterns, followed by interactions between age and gender groups.

5.2 Gender

Sociolinguistic studies have repeatedly found gender-based differences in sound changes in progress. The same seems to hold for alveolarization in Hong Kong Cantonese. Gender-related results from the full statistical model are summarised here in Tables 46 and 47 for ease of reference.

Table 46. Multinomial logistic regression results for gender-based variation of (ŋ)

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
External factors						
(ŋ) > [n] variation						
- Gender (Female)	0.313	0.354	0.884	0.377	1.368	36.8%
(ŋ) > [reduced] variation						
- Gender (Female)	0.638	0.209	3.053	2.268×10^{-3}	1.893	89.3%

Table 47. Multinomial logistic regression results for gender-based variation of (k)

Response Ref. Cat.: [velar] [$]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
External factors						
(k) > [t] variation						
- Gender (Female)	2.622	0.234	11.205	3.848×10^{-29}	13.763	1276.3%
(k) > [reduced] variation						
- Gender (Female)	1.770	0.208	8.510	1.745×10^{-17}	5.871	487.1%

Gender was found to be significant in explaining (**ŋ**) to [reduced] variation ($p = 2.268 \times 10^{-3}$) as well as (**k**) to [alveolar] ($p = 3.848 \times 10^{-29}$) and (**k**) to [reduced] variation ($p = 1.745 \times 10^{-17}$), when all other factors were kept equal. However, gender was not found to be significant in explaining (**ŋ**) to [alveolar] variation. Overall, females were 89.3% more likely than men to use [reduced] realization for (**ŋ**), as well as 1276.3% and 487.1% more likely than men to use [alveolar] and [reduced] realizations for (**k**), respectively.

Since higher order interactions were also found between age and gender for (**ŋ**) to [reduced] variation, and (**k**) to [alveolar] and [reduced] variation, the above estimates for lower-order effects of age and gender respectively are merely provided for reference. Since the higher order interaction between age and gender was found to be significant, the exploration of the effects of gender is considered in more detail in the next section, by taking these interaction effects into consideration at the same time.

5.3 Age and gender interaction

The results of the interaction between age and gender on (**ŋ**) and (**k**) variation is repeated in Tables 48 and 49 for ease of reference.

Table 48. Multinomial logistic regression results of age and gender interaction in (ŋ) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
External factors						
(ŋ) > [alveolar] variation						
- Age * Gender(Female)	-0.008	0.008	-1	0.317	0.9920	-0.8%
(ŋ) > [reduced] variation						
- Age * Gender(Female)	-0.025	0.005	-5.000	5.733×10^{-7}	0.975	-2.5%

Table 49. Multinomial logistic regression results of age and gender interaction in (k) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
External factors						
(k) > [t] variation						
- Age * Gender(Female)	-0.045	0.007	-6.429	1.288×10^{-10}	0.956	-4.4%
(k) > [reduced] variation						
- Age * Gender(Female)	-0.034	0.005	-6.800	1.046×10^{-11}	0.967	-3.3%

As the results show, age and gender were found to interact in their effects on "(**ŋ**) to [reduced]", "(**k**) to [alveolar]" and "(**k**) to [reduced]" variation, in the full statistical models. This means that the effect of age on "(**ŋ**) to [reduced]"/"(**k**) to [alveolar]"/"(**k**)

to [reduced]" variation is not constant across the genders "male" and "female", and vice versa, namely the effect of gender on "(**ŋ**) to [reduced]"/"(**k**) to [alveolar]"/"(**k**) to [reduced]" variation is also not constant across age. The interpretation of interaction effects is more complex. Let us consider the coefficients separately in turn for each of the types of variation below.

5.3.1 Age and gender interaction effects on (**ŋ**) to [reduced] variation

Considering "(**ŋ**) to [reduced]" first, the coefficient for the lower order "age" factor for this variation is -0.025 (see Table 43). This coefficient estimates the effect of a one-year increase in age within the reference group for "gender", which is "male", so the corresponding odds ratio equals exp(-0.025) = 0.975, which is a change in odds ratio of -2.47%. Amongst females, the effect of increasing age on "(**ŋ**) to [reduced]" variation is obtained by adding the main effect -0.025 to the interaction effect -0.025 (i.e. -0.025 - 0.025) which is -0.05. Thus, the ratio by which the odds ratio changes amongst females when increasing age by one year is exp(-0.05) = 0.951, which is a change of -4.9%. In other words, for every one-year increase in age, the odds ratio with which females make use of [reduced] variation decreases *twice* as much (-4.9% vs -2.47%) as the odds ratio for men.

When exploring the interaction between age and gender graphically, it can be seen from Figure 22 (corresponding raw numbers and percentages presented in Table 50) that when the percentages of (**ŋ**) realizations are broken down across age and broad gender group categories used for sampling, interaction of effects between age and gender as described above can be roughly observed. Focusing only on [reduced] realization (marked by a yellow line with inverted triangles), and drawing an imaginary line

through the points for "younger males", "middle age males", and "older males", and comparing the slope of this line with that going through the points for "younger females", "middle age females" and "older females", one can intuit visually that the slope for males is less steep than that for females, and from the above calculations of effects obtained from the full mathematical model, the overall rate of decrease in odds ratio is about half as steep as that for females (-4.9% vs -2.47%).

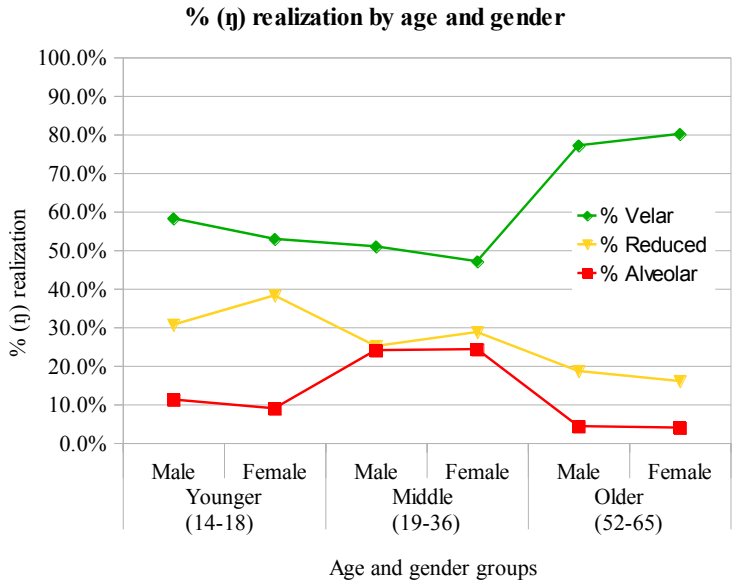


Figure 22. Percentage (ŋ) realization by age and gender group

N= 8392

Table 50. Raw counts of (ŋ) realization by age and gender group

	Younger (14-18)				Middle age (19-36)				Older (52-65)					
	Males		Females		Males		Females		Males		Females		Total	
Velar [ŋ]	862	58.2%	753	52.9%	707	50.9%	636	47.0%	1057	77.1%	1102	80.1%	**5117**	**60.9%**
Alveolar [n]	166	11.2%	127	8.9%	333	24.0%	328	24.3%	59	4.3%	54	3.9%	**1067**	**12.7%**
Other	453	30.6%	544	38.2%	348	25.1%	388	28.7%	255	18.6%	220	16.0%	**2208**	**26.3%**
Total	**1481**	**100.0%**	**1424**	**100.0%**	**1388**	**100.0%**	**1352**	**100.0%**	**1371**	**100.0%**	**1376**	**100.0%**	**8392**	**100.0%**

5.3.2 Age and gender interaction effects on (**k**) to [alveolar]/[reduced] variation

Analogously, the effects of interaction between age and gender on (**k**) to [alveolar] and [reduced] variation respectively, can be obtained in the same manner as described in Section 5.3.1 above, using results from Tables 45 and 49 above.

For (**k**) to [alveolar] variation, the rate with which the odds ratio changes amongst females when increasing age by one year is exp(-0.047-0.045) = exp(-0.092) = 0.912, which is a change in odds ratio of -8.8%. For men, this figure is exp(-0.047) = 0.954 (change in odds ratio of -4.6%). In other words, for every one-year increase in age, the odds ratio with which females make use of [alveolar] variation decreases *almost twice* as much (-8.8% vs -4.6%) as the decrease in odds ratio for men.

Similarly, for (**k**) to [reduced] variation, the rate with which the odds ratio changes amongst females when increasing age by one year is exp(0.001 -0.034) = exp(-0.033) = 0.968, which is a change in odds ratio of -3.2%. For men, this figure is exp(0.001) = 1.001 (change in odds ratio of 1%). In other words, for every one-year increase in age, the odds ratio with which females make use of [alveolar] variation decreases by -3.2% whereas the odds ratio for men increases by 1%.

When exploring the interaction between age and gender graphically for (k) variation, it can be seen from Figure 23 (corresponding raw numbers and percentages presented in Table 51) that when the percentages of (**k**) realizations are broken down across broad age group and gender categories used for sampling, interaction of effects between age and gender as described above can be roughly observed for (**k**) to [reduced] variation.

Focusing only on (**k**) to [reduced] variation (marked by yellow line with inverted triangles), and drawing an imaginary line through the points for “younger males”,

"middle age males", and "older males", and comparing the slope of this line with that going through the points for "younger females", "middle age females" and "older females", one can intuit visually that the slope for males is positive (cf. with calculation from full model: 1% increase in odds ratio) whereas the slope for females is negative (calculation from full model: -3.2% decrease in odds ratio).

Note that for the interaction between age and gender for (k) to [alveolar] variation, the calculations of changes in odds ratio when increasing age by one year across males vs females (-8.8% vs -4.6%) is not mirrored by the graphical trend obtained by age groupings in Figure 23. This discrepancy between results obtained from the full mathematical model and naïve univariate exploration of the effects of age and gender (broad groupings) highlights the fact that trends observed in univariate statistics might not remain when other factors and interactions are taken into account.

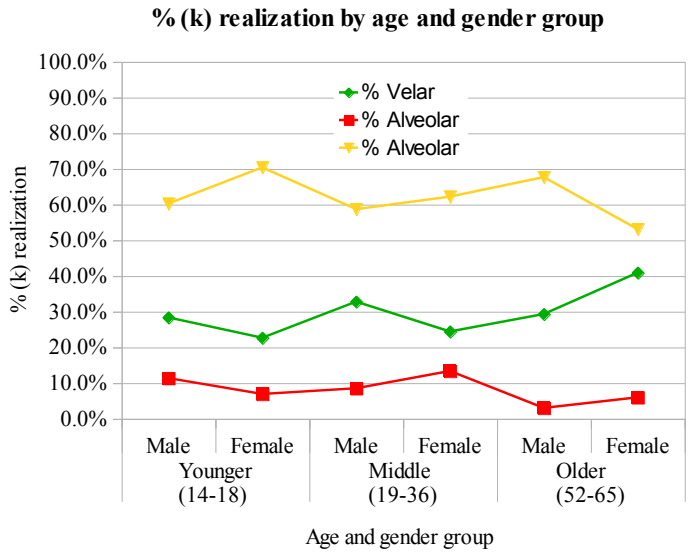


Figure 23. Percentage (k) realization by age and gender group

Table 51. Raw counts of (k) realization by age and gender group

	Younger (14-18)				Middle age (19-36)				Older (52-65)				Total	
	Males		Females		Males		Females		Males		Females			
Velar [ŋ]	257	28.4%	226	22.7%	281	32.8%	230	24.4%	262	29.3%	383	40.9%	**1639**	**29.6%**
Alveolar [n]	103	60.3%	69	70.4%	73	58.7%	126	62.2%	27	67.7%	56	53.1%	**454**	**8.2%**
Other	546	11.4%	702	6.9%	503	8.5%	586	13.4%	605	3.0%	497	6.0%	**3439**	**62.2%**
Total	**906**	**100.0%**	**997**	**100.0%**	**857**	**100.0%**	**942**	**100.0%**	**894**	**100.0%**	**936**	**100.0%**	**5532**	**100.0%**

5.4 Speech style

In terms of the effect of speech style on velar coda realization, it was found, from the full statistical models, that speech style was highly significant in explaining the variation in (**ŋ**) and (**k**) realization. Speech style-related results from the full statistical model are summarised for easier reference in Tables 52 and 53, and the results are discussed in more detail, with interpretations of the results where significant.

In the following sections, the effect of speech style will be discussed in turn for (**ŋ**) and (**k**) variation, respectively, in Sections 5.4.1 and 5.4.2. The effects of interaction of speech style and gender, as well as of speech style with age and gender, on (**ŋ**) and (**k**) variation will be summarized in Section 5.4.3. Age and speech style interaction was not found to be significant and will be left out of this discussion.

5.4.1 Speech style-based effects on (**ŋ**) variation

Speech style was found to be significant in explaining (**ŋ**) to [n] variation as well as [reduced] variation, with more formal speech styles disfavouring [n] realization. Table 52 lists the results from the full mathematical model.

Table 52. Multinomial logistic regression results for speech style-based variation of (ŋ)

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
External factors						
(ŋ) > [n] variation						
- Style (ref.: Casual Speech)						
Passage Reading	-0.748	0.269	-2.781	5.425×10^{-3}	0.473	-52.7%
Word List	-1.055	0.272	-3.879	1.050×10^{-4}	0.348	-65.2%
Minimal Pairs	-0.672	0.324	-2.074	3.807×10^{-2}	0.511	-48.9%
(ŋ) > [reduced] variation						
- Style (ref.: Casual Speech)						
Passage Reading	-1.676	0.159	-10.541	5.597×10^{-26}	0.187	-81.3%
Word List	-2.272	0.159	-14.289	2.551×10^{-46}	0.103	-89.7%
Minimal Pairs	-3.702	0.295	-12.549	4.018×10^{-36}	0.025	-97.5%

In terms of (**ŋ**) to [n] variation, the results show that the odds with which [n] was produced decreased across styles from passage reading (odds ratio with reference to 'casual' speech style: 0.473, $p = 5.425 \times 10^{-3}$), to word list (odds ratio with reference to 'casual' speech style: 0.348, $p = 1.050 \times 10^{-4}$), but anomalously increased again in the minimal pairs task (odds ratio with respect to 'casual' speech style: 0.511, $p = 3.8707 \times 10^{-2}$). In other words, it was, respectively, 52.7%, 65.2%, and 48.9% *less likely* for (**ŋ**) to be produced as [alveolar], in the passage reading, word list reading, and minimal pairs reading tasks, than in casual speech style, with all else being equal.

In terms of (**ŋ**) to [reduced] variation, more formal speech styles disfavoured [reduced] realization. The odds with which [reduced] (as opposed to [velar]) was produced gradually decreased across styles from passage reading (odds ratio with respect to 'casual' speech style: 0.187, $p = 5.597 \times 10^{-26}$), to word list (odds ratio with respect to 'casual' speech style: 0.103, p 2.551 = $\times 10^{-46}$), and finally to the minimal pairs task (odds ratio with respect to 'casual' speech style: 0.025, $p = 4.018 \times 10^{-36}$). In other words, it was, respectively, 81.3%, 89.7%, and 97.5% less likely for (**ŋ**) to be produced as [reduced] than in casual speech style, in the passage reading, word list reading, and minimal pairs reading tasks, with all else being equal. This conforms to our finding,

discussed in Section 5.1.1 on age-based variation, that the [reduced] variant for (**ŋ**) likely operates above the level of conscious awareness for speakers of all ages, with an increase in speech style leading to a decrease in the use of [reduced] variants of (**ŋ**).

Figure 24 graphically shows the percentages of variation in (**ŋ**) realization by speech style, although other factors were *not* controlled for (the raw data and percentages are presented in Table 53). The anomaly found in the full model whereby the rate at which (**ŋ**) was realized as the innovate [n] was unexpectedly high for the most formal speech style—minimal pair reading—is mirrored in this univariate plot. In simple terms, percentage of [alveolar] realization overall went up markedly from 10.3% in the word list reading task to 27.2% in the minimal pairs reading task, with a concomitant decrease from 25.0% to 6.8% in [reduced] realizations. This anomaly will be discussed in further detail in Section 5.5. Furthermore, there is an overall increase in [velar] realization (indicated by the green line with diamond symbols) when going from casual speech style to passage reading and word list style, although there was little difference between the two reading styles.

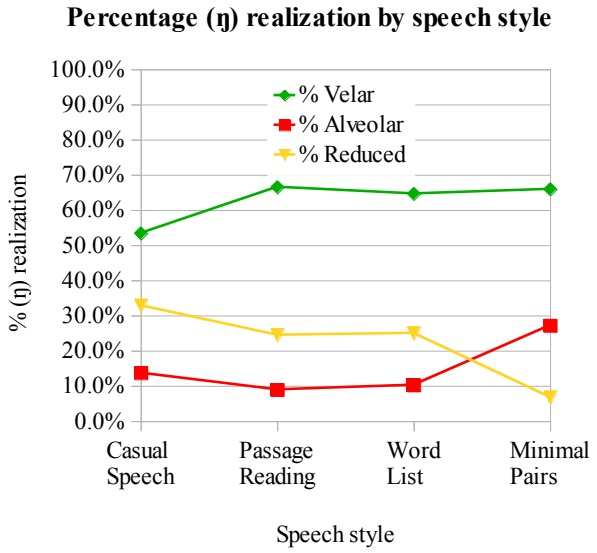


Figure 24. Percentage (ŋ) by speech style

N= 8392

Table 53. Raw counts of (ŋ) realization by speech style

	Casual Speech		Passage Reading		Word List		Minimal Pairs		Total	
Velar [ŋ]	1696	53.4%	1337	66.6%	1609	64.7%	475	66.0%	**5117**	**61.0%**
Alveolar [n]	435	13.7%	180	9.0%	256	10.3%	196	27.2%	**1067**	**12.7%**
Reduced	1044	32.9%	492	24.5%	623	25.0%	49	6.8%	**2208**	**26.3%**
Total	**3175**		**2009**		**2488**		**720**		**8392**	

5.4.2 Speech style-based effects on (**k**) variation

Speech style was also found to be significant in explaining (**k**) to [t] variation as well as [reduced] variation, with more formal speech styles disfavouring [t] realization. Table 54 lists the results from the full mathematical model.

Table 54. Multinomial logistic regression results for speech style-based variation of (k)

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
External factors						
(k) > [t] variation						
- Style (ref.: Casual Speech)						
Passage Reading	-0.871	0.248	-3.512	4.45×10^{-4}	0.419	-58.1%
Word List	-1.741	0.256	-6. 695	2.152×10^{-11}	0.180	-82.0%
Minimal Pairs	-0.919	0.226	-4.066	4.775×10^{-5}	0.399	-60.1%
(k) > [reduced] variation						
- Style (ref.: Casual Speech)						
Passage Reading	-1.597	0.161	-9.919	3.433×10^{-23}	0.203	-79.7%
Word List	-1.549	0.143	-10.832	2.424×10^{-27}	0.212	-78.8%
Minimal Pairs	-2.687	0.170	-15.806	2.834×10^{-56}	0.068	-93.2%

The results for (**k**) to [t] variation show the same anomaly as has been found for (**ŋ**) to [n] variation described in Section 5.4.1 above, namely that the minimal pairs condition shows a surprisingly small effect on rate of [t] realization, defying our expectation that an increase in speech style formality should show a steady decrease in the odds with which [t] is produced. As the results show, the odds with which [alveolar] was produced decreased across styles from passage reading (odds ratio with respect to 'casual' speech

style: 0.419, $p = 4.45 \times 10^{-4}$), to word list (odds ratio with respect to 'casual' speech style: 0.180, $p = 2.152 \times 10^{-11}$), but increased again in the minimal pairs task (odds ratio with respect to 'casual' speech style: 0.399, $p = 4.775 \times 10^{-5}$). In other words, it was, respectively, 58.1%, 82.0%, and 60.1% less likely for (**k**) to be produced as [alveolar] in the passage reading, word list reading, and minimal pairs reading tasks, than in casual speech style. Similar to (**ŋ**), this cross-over effect for style will be discussed in more detail in Section 5.5 below.

As for (**k**) to [reduced] realization, speech style was also found to be significant, with more formal speech styles disfavouring [reduced] realization. The odds with which [reduced] was produced changed across styles from passage reading (odds-ratio with respect to 'casual' speech style: 0.203, $p = 3.433 \times 10^{-23}$), to word list (odds-ratio with respect to 'casual' speech style: 0.212, $p = 2.424 \times 10^{-27}$), and finally to the minimal pairs task (odds-ratio with respect to 'casual' speech style: 0.068, $p = 2.834 \times 10^{-56}$). In other words, it was, respectively, 79.7%, 78.8%, and 93.2% *less* likely for (**k**) to be produced as [reduced] in the passage reading, word list reading, and minimal pairs reading tasks, than in casual speech style, with all else being equal. While there seems to be little difference between the odds ratios with which (**k**) to [reduced] was produced in the passage reading and word list reading tasks, it can be seen that minimal pairs did prompt speakers to drastically decrease their use of the [reduced] variant. This echoes the results obtained from (**ŋ**) to [reduced] variation, where the odds ratio with which [reduced] variants were produced was increasingly lower with an increase in formality in speech style, being lowest for the minimal pair reading task. These results suggest that the [reduced] variants for both (**ŋ**) and (**k**) are above the level of social awareness,

leading speakers to avoid their use in the task that demands the most conscious attention. This will also be discussed further in Section 5.5.

For reference purposes, Figure 24 graphically shows (**k**) variation. The cross-over in [alveolar] realization for the minimal pairs reading task which can be seen in this graph mirrors the result obtained from the full statistical model discussed above, whereby there is an unexpectedly high rate of innovative [alveolar] realization for (**k**) in the minimal pairs task. The raw data and percentages of (**k**) are also provided in Table 55 for reference.

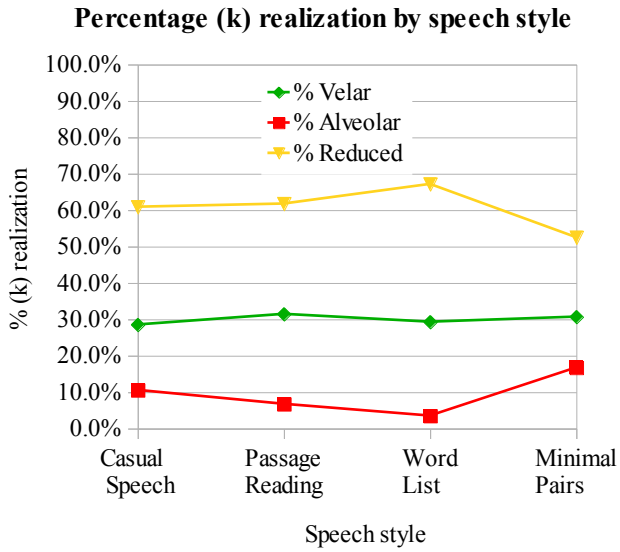


Figure 24. Percentage realization (k) by speech style

N=5532

Table 55. Raw counts of (k) realization by speech style

	Casual Speech		Passage Reading		Word List		Minimal Pairs		Total	
Velar [k]	535	28.5%	321	31.5%	565	29.3%	218	30.7%	**1639**	**29.6%**
Alveolar [t]	198	10.6%	69	6.8%	68	3.5%	119	16.8%	**454**	**8.2%**
Reduced	1141	60.9%	630	61.8%	1295	67.1%	373	52.5%	**3439**	**62.2%**
Total	**1874**		**1020**		**1928**		**710**		**5532**	

Next, the interaction between speech style, age, and gender will be discussed.

5.4.3 Interaction between speech style, age, and gender

The results of the interaction between speech style and gender on (**ŋ**) and (**k**) variation as obtained from the full mathematical model are summarized in Tables 56 and 57, respectively. Significant effects on (**k**) variation were also found for the interaction between age, gender and speech style, and Table 58 summarizes the results of the effects of this interaction from the full mathematical model for reference.

Table 56. Multinomial logistic regression results of speech style and gender interaction in (ŋ) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
External factors						
(ŋ) > [n] variation						
Not significant						
(ŋ) > [reduced] variation						
- Gender(Female) * Style(PR)	0.503	0.212	2.373	1.766×10^{-2}	1.654	65.4%
- Gender(Female) * Style(Word	0.808	0.202	4.000	6.334×10^{-5}	2.243	125.0%

Table 57. Multinomial logistic regression results of age and gender interaction in (k) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
External factors						
(k) > [t] variation						
- Gender(Female) * Style(MP)	-3.255	0.419	-7.768	7.942×10^{-15}	0.039	-96.1%
(k) > [reduced] variation						
- Gender(Female) * Style(PR)	0.853	0.338	2.524	1.161×10^{-2}	2.347	134.7%
- Gender(Female) * Style(WL)	0.675	0.282	-2.394	1.668×10^{-2}	1.964	96.4%

Table 58. Multinomial logistic regression results of speech style, age and gender interaction in (k) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
External factors						
(k) > [t] variation						
- Age*Gender(Female)*Style(MP)	0.080	0.010	8.000	1.244×10^{-15}	1.083	8.3%
(k) > [reduced] variation						
- Age*Gender(Female)*Style(RP)	-0.036	0.008	-4.500	6.795×10^{-6}	0.965	-3.5%
- Age*Gender(Female)*Style(Word	-0.028	0.007	-4.000	6.334×10^{-5}	0.972	-2.7%

While the above results show significant interactions between the social factors age, gender and speech style, the complexity of interpreting these interactions is deemed beyond the scope of this dissertation, and the value of describing the highly complex interactions is limited. Therefore, in the remainder of this section, univariate statistics will be explored for a "naïve" overview of how age, gender and speech broadly interact, when other linguistic factors are not accounted for.

When studying the "naïve" interaction between age group and speech style, it can be seen that speakers in the three age groups seem to respond differently to different speech style settings. Figures 25 and 26 show percentage of realization of (**ŋ**) and (**k**) respectively, by different age groups and across speech styles. Within each figure, each graph shows the results for each age group, and the percentages of realization on the y-axis and the four speech styles on the x-axis. Again, percentage of [velar] realization is displayed using a green line and a diamond (◆) symbol, [alveolar] realization using a red line and a square (■) symbol, and [reduced] realization using a yellow line and an inverted triangle (▼) symbol. As can be seen from the graphs, the age groups respond quite differently across speech styles, especially for (**ŋ**) realization.

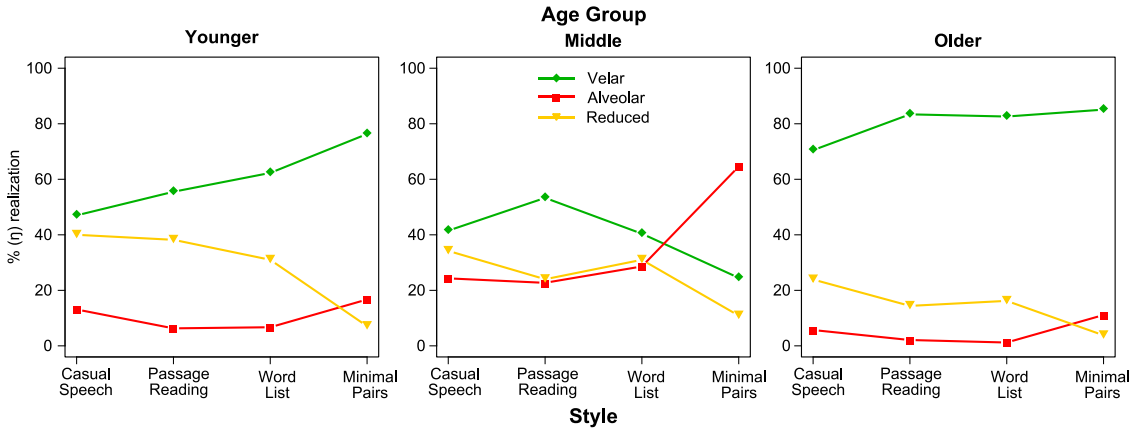


Figure 25. Percentage (ŋ) realization by age group across speech styles

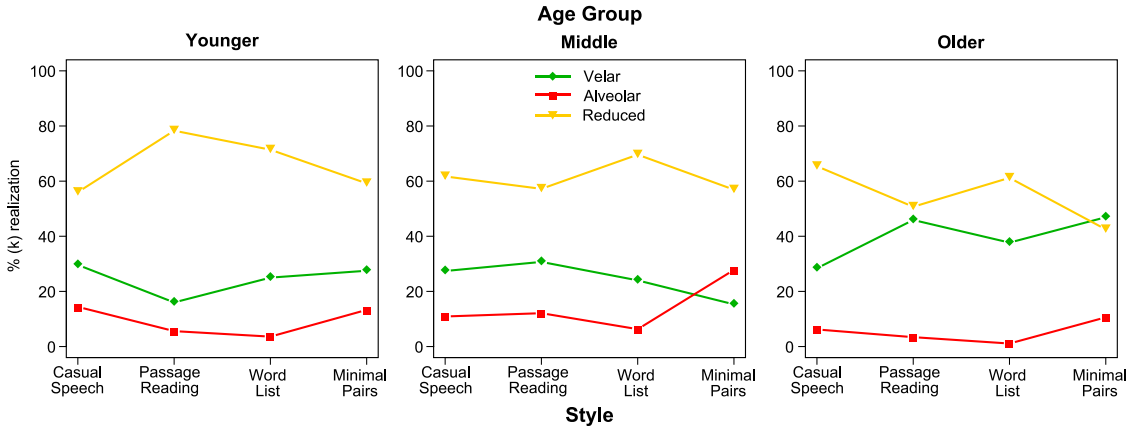


Figure 26. Percentage (k) realization by age group across speech styles

For both (**ŋ**) and (**k**) realization, the middle age group seems anomalous, when compared against younger and older speakers. It can be seen that for (**ŋ**) realization, there was an overall rise in percentage of [velar] realization by younger and older speakers when speech style increased in formality. The same trend can be seen in (**k**) realization, albeit in a less distinctive manner.

What is consistent, however, across all age groups and for both (**ŋ**) and (**k**) realization, was that in the minimal pairs speech style, the proportion of [alveolar] realization went up in conjunction with a decrease in the proportion of [reduced] realizations. This goes against all expectations of what should be obtained in the

minimal pairs reading task. The original prediction was that with minimal pairs reading being the speech setting where speakers typically pay the most attention to their speech production, it would be the least likely setting for speakers to make use of the innovative variant, namely [alveolar] realization. Yet, the results go against this expected trend.

The increase in [velar] realization by the younger and older speakers with an increase in formality of speech styles suggests that for these groups of speakers, [velar] retention might be above the level of conscious awareness, and speakers might be conscious that [velar] realization is more appropriate in formal speech situations. However, even these groups which seem conscious of the role of [velar] retention in more formal speech settings display an increase in [alveolar] realizations in minimal pairs speech styles and an overall decrease in [reduced] realizations. This suggests that the picture might not be as simple as '[velar] equals conservative' and '[alveolar] equals innovative' in this sound change, and that instead perhaps [reduced] realization is the least accepted variant.

5.5 Discussion

Returning to the notion that it is speakers who change their speech behaviour, it has been clear in this section that there is a significant amount of interspeaker variation but that social factors play a significant role in explaining the variation in (**ŋ**) and (**k**) variation. Gender was found to be significant in explaining (**ŋ**) to [reduced] realization, (**k**) to [alveolar] variation, as well as (**k**) to [reduced] variation. In terms of age, roughly the same cohort of speakers as found in Wong's study were making the most progressive use of [alveolar] realization for (**ŋ**)—in this study, it was the speakers in

their 20s and 30s, in Wong's (2005) study it was the 11–20 year olds. As Wong's study did not code for [reduced] variants, it is unclear whether there is any possibility that similar patterns of extensive lenition in could be found in the younger generations in her cohort of informants.

In the case of alveolarization, it seems that in most speakers sampled in this study, there is little conscious social awareness of the actual phonological difference itself. While alveolarization is often described as a feature of *la:nɬjɐm˥* or "lazy speech" (recall discussion in Section 2.2), such discussion is generally limited to those who are educators—or who engaged in social or professional activities, such as those who perform Chinese opera, or broadcast the news, where participants are trained to adopt pronunciation according to the standard description—who have a vested interest in being aware of the "proper" pronunciation. There was also one participant (OF56) in this study who was herself aware of alveolarization through personal observation that the speech of the younger generation was different from her own. But such examples are rather rare, compared to the overwhelming majority of the population that is not consciously aware of the phenomenon.

Yet, even if there is little general social awareness of alveolarization in the speech community and previous studies have found evidence for only those linguistic features that are within speakers' social awareness that change throughout adult life, this is still not a guarantee that the age-related differences in innovative alveolar realization found here are necessarily apparent-time generational changes. However, the evidence so far does suggest that this is likely the case, and it is only through a panel study tracking these same speakers in ten, perhaps twenty years' time that we can be certain that the apparent-time interpretation is in fact the correct one.

It was also found that the middle age group was anomalous in being more progressive than even the younger age group in use of [alveolar] variants. If there was not somehow a sampling bias involved, and if this was also reflected in the overall population, then this anomaly seems to go against the trend in which sound changes have been found to develop in an S-shaped curve, with progressively more and more speakers making use of innovative variants—rather, there seems to be a rise in [alveolar] variants in speakers who are in their 20s and 30s and then a fall in the younger speakers. It is unclear at this stage whether the transition from the adolescent group to adulthood could indeed be an age-grading pattern. Another real possibility for the pattern observed could be if there were a sampling bias. The informants in the youngest age group were chosen by teachers to take part in this study. This could have resulted in teachers selecting a group of elite students who might generally be better academic performers and more likely to want to "speak properly", with the result that the extent of [velar] retention was higher and the rate of [reduced] or [alveolar] realizations were lower than average for their peers. However, this is just a conjecture, and without sampling further from their peers in a less selective way, it would be hard to know whether the sample collected in this study is indeed representative of their generation when it comes to (**ŋ**) and (**k**) realization.

Furthermore, our "naïve" exploration of speech style showed that age groups behaved differently with respect to speech style. Younger speakers and older speakers behaved similarly, in the sense that the amount of [velar] retention went up, as predicted, with increasing formality in speech setting. The middle age group showed anomalous behaviour and did not show the expected increase in [velar] retention with increasing formality. Instead, the amount of [velar] realization went down for this group.

Moreover, what was remarkable was that for all age groups, there was a noticeable increase in alveolarization in the minimal pairs reading task. Even those speakers who normally have a low rate of [alveolar] realization, especially those in the older age group, suddenly seemed to display a spike in [alveolar] variants. If one compares the rate of [alveolar] increase with the rate of decrease in [reduced] realization, it is apparent that these two measures are correlated. It is likely that the minimal pairs reading task prompted the speakers to pay the greatest amount of attention to their own speech, and to fully pronounce the codas where they would not normally do so in daily speech. However, in doing so, it seemed that speakers were increasing their use of [alveolar] variants because they were aware of [reduced] variants being the least desirable.

There is a point to be made here about the lack of orthographic cues when reading Chinese characters. It was almost as if the speakers were faced with a "forced choice" decision task, guessing which realization the codas should take, [velar] or [alveolar], as the orthography shows no clue as to what the coda should be, with Chinese characters lacking any real form of phonetic encoding of sounds. This means perhaps we should reinterpret previous studies of alveolarization (Bauer 1979; Zee 1996; Ding 2010), where subjects were asked to perform a careful reading of single-word lists that contained velar nasal and stop codas. These studies reported that alveolarization was very progressive in the informants that they performed the studies on. However, without also having performed an analysis of these informants' casual speech behaviour, the extent of alveolarization might in fact overall be less extensive than the single-word list reading task suggested. For instance, Ding (2010) has found in his subjects, who were asked to read a list of words that contained final velars, that the rate of coda

alveolarization is fairly high in his subjects, without offering any comparison with casual speech settings. Based on the cross-over finding of unexpectedly high rate of alveolar realization in the minimal pairs task, it is suggested here that future studies on phonetic variation in languages with non-phonetic orthography should not only rely on word reading and minimal pair reading tasks, but also try to incorporate elicitation and analysis of casual speech data. This would help mitigate against the "forced choice decision task" effect found here which might prompt informants to make use of conservative vs. innovative variants in a random chance manner, and might in turn lead to an overestimation of use of innovative variants.

Referring back to our discussion in Chapter 2 on the discourse on "lazy pronunciation", one could invoke the notion of "indexicality" to explain the interacting variation patterns in speech style, age, and gender, as discussed in Section 5.4.3. The notion of indexicality in sociolinguistic variation was introduced by Silverstein (2003) and has been employed in what has been termed "second-wave" sociolinguistic studies (see e.g. Eckert 2008; Leimgruber 2009; Hall-Lew 2014) to analyse how speakers might attach social meaning to linguistic variation patterns, beyond pre-defined social categories such as gender and age. While this notion is typically the foundation of ethnographic studies whereas the current study was not set up to be one, our foregoing discussion on [reduced] realization possibly being the least accepted variant for (**ŋ**) could suggest that speakers were aware of [reduced] realization as being indexical of and associated with social stigmas that go with "sloppy" and "lazy" pronunciations, prompting speakers to opt for using one of the two salient coda realizations [ŋ] and [n] in a fairly random manner in order to *not* produce [reduced] realizations. As for the consistent finding with Wong's 2005 study that the same generation of speakers (those

in their 20s and 30s in this study, and those 11–20 years of age in Wong's) produced the highest rate of [alveolar] variants of (**ŋ**), even higher than the youngest speakers in this study, this could also be indexical of a generational trend.

In our exploration of inter-speaker variation, it was found that there were two distinct outliers, MF35 and MM35, whose speech patterns were very conservative, and much more similar to the older speakers than the middle age group speakers which they were from. In the case of MM35, his mother, OF56, was also included in this study and her speech patterns also showed a high rate of [velar] retention and low rates of innovative [alveolar] (or [reduced]) variants. It could be the case that MM35's anomalous speech behaviour compared to his peer group (his good school friend, MM35, was also included in this study and was found to have the lowest rate of [velar] retention and the highest rate of [alveolar] realization) was in fact due to a close resemblance of his speech characteristics to his mother's. Despite having spent a large amount of time studying and maintaining friendship with peers who have very different speech characteristics from his own, it seems that transmission of (**ŋ**) and (**k**) realization from his mother, as opposed to diffusion of wider innovations from his peer group prevailed (cf. distinction between transmission vs. diffusion of linguistic innovations in Labov 2007). As for the conservative speech behaviour of MF35, it could be explained by the fact that she is a Chinese language and history teacher, and it seems likely that her training of standard Cantonese pronunciation helped her become aware of and maintain the distinction between velar and alveolar codas, even in casual speech. Such outliers remind us that while the adoption of broad social distinctions of "age group" might be useful for obtaining a broad general picture of the extent of variation which might be observed between generations of speakers, researchers should always be aware

that individual speakers might not neatly fit within the broad categories they objectively belong to, and that other forms of self-made identity, such as heritage or profession, might in the end explain a given individual speaker's departure from the wider norm. It also highlights the value of supplementing sociovariationist studies like the current one with ethnographic studies which investigate the indexicality of social meaning (cf. our discussion in Section 5.4.3) in variation in speech patterns.

5.6 Summary

This chapter looked at the language-external factors of (**ŋ**) and (**k**) variation, by examining age, gender, and speech style. In terms of age-based variation, the speakers in the "middle" in their 20s and 30s were the most progressive in alveolarization, while the younger age group made less use of alveolar variants than expected. This cohort of speakers were found to display similarly high rates of innovative alveolar realization for (**ŋ**) as the 11-20 year olds in Wong's (2005) study. The apparent-time interpretation was invoked to explain this variation pattern.

In terms of broad gender-based variation, overall, females were more likely than men to use [reduced] realization for (**ŋ**), and more likely than men to use [alveolar] and [reduced] realizations for (**k**), respectively.

It was also found in our exploration of speech styles that the rate of innovative alveolar realization was unexpectedly high for the minimal pairs task, and it was suggested that the reduced realizations could be above the level of social awareness of speakers, and that the increased conscious awareness encouraged by the minimal pairs reading task led them to decrease the amount of reduced realizations, thereby leaving only [velar] and [alveolar] realizations for speakers to choose from in a "forced choice decision"-like manner.

The next chapter describes the results pertaining to predictions made by the neogrammarian regularity hypothesis that it is the immediate phonological environment which explains sound change.

6 Neogrammarian regularity hypothesis

This section examines the immediate phonological environment of (**ŋ**) and (**k**) with regards to neogrammarian predictions on regularity of alveolarization. It starts by exploring the preceding vowels of (**ŋ**) and (**k**), which are the four long vowels **a**:, **ɛ**:, **œ**:, and **ɔ**:, and three short vowels **ɐ**, **ɪ**, and **ʊ** respectively. These have in previous studies, and in our statistical model in Chapter 4, been found to be strong predictors of alveolarization. Then, the following onset, tautosyllabic onset, and tone on the morphosyllable are also discussed, as we have seen that these phonological environments have an effect on alveolarization or lenition of (**ŋ**) and (**k**).

6.1 Preceding vowels

As mentioned in Section 4.1, it was found that preceding vowels were significant in predicting [alveolar] and [reduced] realizations in (**ŋ**) and (**k**) variation. The results of the effects of preceding vowels from the full mathematical model are summarized in Tables 59 and 60.

Turning our attention first to (**ŋ**) realization as [alveolar], overall, the vowels **a**:, **ɛ**:, **œ**: and **ɔ**:, were significantly different from the vowels **ɪ** and **ʊ**, which highly disfavoured [alveolar] realization, with respect to the reference category **ɐ**. The model confirmed that the preceding vowels **a**:, **œ**: and **ɐ** are the most progressive in favouring [alveolar] realization.

Table 59. Multinomial logistic regression results for preceding vowel effects on (ŋ) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
Internal factors						
(ŋ) > [n] variation						
- Vowel (Ref.: ɐ)						
ɛː	-0.858	0.269	-3.190	1.425×10^{-3}	0.424	-57.6%
ɪ	-4.351	0.354	-12.291	1.013×10^{-34}	0.013	-98.7%
ʊ	-5.253	0.457	-11.495	1.405×10^{-30}	0.005	-99.5%
ɔː	-1.024	0.209	-4.900	9.607×10^{-7}	0.359	-64.1%
œː	-0.153	0.232	-0.659	0.510	0.858	-14.2%
aː	0.356	0.213	1.671	0.095	1.428	42.8%
(ŋ) > [reduced] variation						
- Vowel (Ref.: ɐ)						
aː	1.421	0.167	8.509	1.755×10^{-17}	4.141	314.1%
ɛː	1.134	0.175	6.480	9.172×10^{-11}	3.108	210.8%
ɪ	-1.614	0.194	-8.320	8.827×10^{-17}	0.199	-80.1%
ɔː	1.317	0.151	8.722	2.737×10^{-18}	3.732	273.2%
œː	1.544	0.158	9.772	1.483×10^{-22}	4.683	368.3%
ʊ	-0.767	0.175	-4.383	1.171×10^{-5}	0.464	-53.6%

In terms of (**ŋ**) to [reduced] variation, it was also found that preceding vowels were significant in explaining the variation. There was a statistically significant difference between the preceding vowels **a**ː, **ɛ**ː, **ɪ**, **ɔ**ː, **œ**ː, and **ʊ**, with respect to the reference vowel **ɐ**, in predicting (**ŋ**) to [reduced] realization. The vowels **ɪ** and **ʊ** highly disfavoured [reduced] realization, with respect to the reference category **ɐ**, while the vowels **a**ː, **ɛ**ː, **ɔ**ː and **œ**ː were, respectively, 4.13, 3.11, 3.73, and 4.68 times more likely to favour [reduced] realization over the reference vowel **ɐ**. The results altogether show that the vowels **œ**ː and **a**ː are the most progressive in favouring (**ŋ**) to [reduced] variation, followed by **ɔ**ː and **ɛ**ː.

As for (**k**) to [alveolar] realization (see Table 60), there was a statistically significant difference between the preceding vowels **ʊ**, **ɪ**, **ɔ**ː and **œ**ː ($p = 1.302 \times 10^{-9}$, 6.573×10^{-23}, 1.253×10^{-30}, 1.453×10^{-3}) with respect to the reference vowel **ɐ**, in predicting (**k**) to [alveolar] realization. The vowels **ʊ**, **ɪ**, **ɔ**ː and **œ**ː were far less likely to

favour [alveolar] realization than **ɐ** (99.9%, 99.9%, 72.7%, and 53.3% less likely, respectively).

Table 60. Multinomial logistic regression results for preceding vowel effects on (k) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
Internal factors						
(k) > [t] variation						
- Vowel (Ref.: ɐ)						
ʊ	-7.596	1.252	-6.067	1.302×10^{-9}	5.025×10^{-4}	-99.9%
ɪ	-7.706	0.782	-9.854	6.573×10^{-23}	4.501×10^{-4}	-99.9%
ɔː	-1.300	0.113	-11.504	1.253×10^{-30}	0.273	-72.7%
œː	-0.761	0.239	-3.184	1.453×10^{-3}	0.467	-53.3%
ɛː	-0.398	0.237	-1.679	0.093	0.672	-32.8%
aː	0.013	0.121	0.107	0.914	1.013	1.3%
(k) > [reduced] variation						
- Vowel (Ref.: ɐ)						
ɛː	0.682	0.184	3.707	2.101×10^{-4}	1.978	97.8%
ʊ	-1.254	0.118	-10.627	2.229×10^{-26}	0.285	-71.5%
ɪ	-2.094	0.117	-17.897	1.235×10^{-71}	0.123	-87.7%
aː	0.924	0.106	8.717	2.857×10^{-18}	2.519	151.9%
ɔː	0.554	0.100	5.540	3.025×10^{-8}	1.740	74.0%
œː	0.465	0.192	2.422	1.544×10^{-2}	1.592	59.2%

In terms of (**k**) to [reduced] realization, there was a statistically significant difference between the preceding vowels **ɛː**, **ʊ**, **ɪ**, **aː**, **ɔː** and **œː** with respect to the reference vowel **ɐ**, in predicting (**k**) to [reduced] realization. The vowels **ɪ** and **ʊ** highly disfavoured [reduced] realization, with respect to the reference category **ɐ** (at 2.0, 0.3, 0.1, 2.5, 1.7, and 1.6 times odds-ratio, $p = 2.101 \times 10^{-4}$, 2.229×10^{-26}, 1.235×10^{-71}, 2.857×10^{-18}, 3.025×10^{-8}, and 1.544×10^{-2}, respectively), while the vowels **aː**, **ɛː**, **œː** and **ɔː** were more likely to favour [reduced] realization over the reference vowel **ɐ**.

As mentioned, previous studies have found that the preceding vowel conditioned the realization of (**ŋ**) and (**k**). The above results obtained from the full mathematical model described in Chapter 4 confirm the findings from previous studies reviewed in Sections

2.4 and 2.5. More specifically, the short vowels **ɪ** and **ʊ** were the least likely to favour (**ŋ**) to [alveolar]/[reduced] and (**k**) to [alveolar]/[reduced] realizations. For (**ŋ**) variation, **a**:, **œ**: and **ɐ** were the most progressive in favouring [alveolar] realization, while vowels **a**: and **œ**: were the most progressive in favouring (**ŋ**) to [reduced] variation, followed by **ɔ**: and **ɛ**:. For (**k**) variation, **a**: and **ɐ** were the most progressive in favouring [alveolar] realization followed by **ɛ**:, **œ**:, **ɔ**:, while vowels **a**: **ɛ**:, **ɔ**:, **œ**:, **ɐ** (in decreasing order), are the most progressive in favouring (**k**) to [reduced] variation.

For visual exploration of the data, "naïve" univariate statistics are presented by summarizing the percentage of realization of the variables by preceding vowel. The statistics are summarised in Tables 63 and 64 and shown as graphs in Figures 27 and 28, ranked according to increasing percentage of [velar] realization. These statistics match the results from the full mathematical model and suggest that for both (**ŋ**) and (**k**), the highest rate of [velar] retention occurs when they were preceded by the short vowels **ʊ** and **ɪ**. This is in line with previous studies (Bauer 1979; Shum 1993; Wong 2005). There was almost full velar retention (89.5%) for (**ŋ**) when preceded by **ɪ**, and a slightly lower percentage for **ʊ** (78.4%). The amount of alveolarization for these two phonological environments was almost zero (0.75% for **ʊ** and 1.3% for **ɪ**).

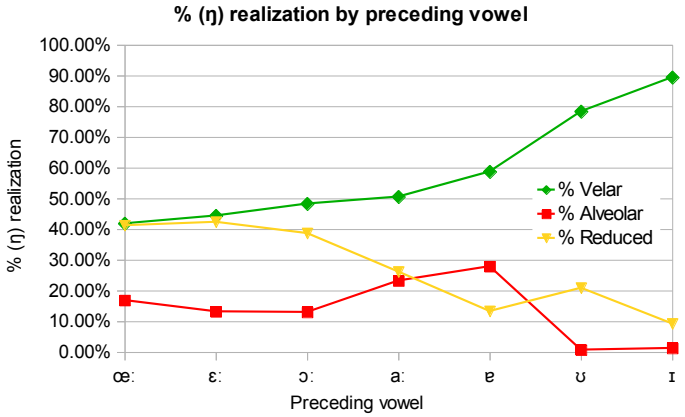


Figure 27. Percentage (ŋ) realization by preceding vowel, all speech styles combined

N=8392

Table 61. Realization of (ŋ) by preceding vowel in all speech styles

	œ:		**ɛ:**		**ɔ:**		**a:**		**ɐ**		**ʊ**		**ɪ**		**Total**	
Velar [ŋ]	548	41.9%	215	44.4%	754	48.3%	472	50.6%	701	58.8%	1256	78.4%	1171	89.5%	**5117**	**61.0%**
Alveolar [n]	221	16.9%	64	13.2%	203	13.0%	217	23.3%	333	27.9%	12	0.75%	17	1.3%	**1067**	**12.7%**
Other	540	41.2%	205	42.4%	604	38.7%	244	26.2%	159	13.3%	335	20.9%	121	9.2%	**2208**	**26.3%**
Total	**1309**		**484**		**1561**		**933**		**1193**		**1603**		**1309**		**8392**	

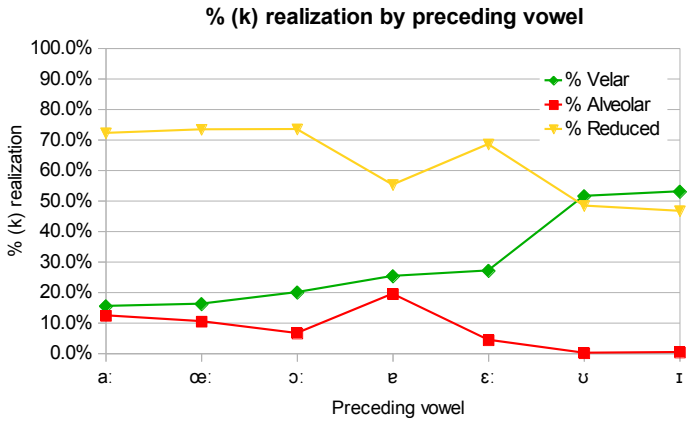


Figure 28. Percentage (k) realization by preceding vowel, all speech styles combined

Table 62. Realization of (k) by preceding vowel in all speech styles

	a:		œ:		ɔ:		ɐ		ɛ:		ʊ		ɪ		Total	
Velar [k]	176	15.5%	34	16.2%	267	20.0%	243	25.3%	74	27.1%	389	51.5%	456	53.0%	**1639**	**29.6%**
Alveolar [t]	141	72.1%	22	73.3%	88	73.4%	187	55.2%	12	68.5%	1	48.3%	3	46.6%	**454**	**8.2%**
Other	821	12.4%	154	10.5%	981	6.6%	530	19.5%	187	4.4%	365	0.1%	401	0.3%	**3439**	**62.2%**
Total	**1138**		**210**		**1336**		**960**		**273**		**755**		**860**		**5532**	

This is similar for (**k**), which showed a numerically lower percentage of velar retention than (**ŋ**), but which was still highest amongst the preceding vowel environments. The remaining vowels **a**:, **ɐ**, **œ**:, **ɔ**:, and **ɛ**: merit discussion in different groupings as the results indicate that (**ŋ**) and (**k**) differed in how they were realized when preceded by these vowels.

Overall, for (**ŋ**), the vowels after which there was the least amount of [velar] realization were, in order, the front vowels **œ**: and **ɛ**: , back vowel **ɔ**:, followed by central vowels long open **a**: and short open-mid **ɐ**. However, as can be seen in Figure 27, across vowels, a lower percentage of [velar] realization did not necessarily equate a higher percentage of [alveolar] realization, as has been assumed in previous studies on alveolarization of (**ŋ**). In fact, an increase in [velar] realization was closely linked to a decrease in [reduced] realization, but only up to and including the short central mid-open **ɐ** vowel, and not including the short vowels **ɪ** and **ʊ**. There does not seem to be a single vowel feature—whether frontness, height or length—which can neatly explain which vowels are associated with [velar] retention. On the basis of this, one cannot easily invoke the principle of least effort to explain the differences in (**ŋ**) realization with respect to preceding vowel. On the other hand, vowel length does, at least in the case of (**ŋ**), predict the likelihood of [velar] retention. When aligning the

vowels according to increasing likelihood of [velar] retention, it can be seen that the vowels are neatly divided into long and short, with longer vowels much more likely leading to non-velar realizations, and short vowels more likely to retain [velar] realization. The phonetic duration of the preceding vowels were not quantified in this study, so it is hard to say at this stage whether there is an inverse linear relationship between quantitative vowel length and [velar] retention, but this question could be investigated in future work.

There seems to, more or less, be an inverse relationship between the amount of [reduced] and [alveolar] realization across the long vowels. Those environments where there were found to be the most [reduced] variants were mid-low (cf. Table 18 on p. 72) vowels **œ**:, **ɛ**: and **ɔ**:, followed by the central open **a**: and central mid-low **ɐ**. These environments also showed increasing amounts of [alveolar] realization.

As for (**k**), the vowels which favoured [velar] retention the least were, in decreasing order, central open-mid **a**:, central close-mid **œ**:, and back close-mid **ɔ**: respectively. As with (**ŋ**), there is no vowel feature such as frontness or height which can explain the rate of [velar] retention.

However, for both (**ŋ**) and (**k**) vowel height seems a fairly good candidate for explaining degree of alveolarization. The more open the vowel, the more likely it seemed that (**ŋ**) and (**k**) were realized as [alveolar], with open vowels **a**: and **ɐ** more likely to be associated with alveolarization than more close vowels such as **œ**:, **ɔ**:, **ɛ**:, **ɪ** and **ʊ**. Again, principle of least effort cannot be invoked to explain why the open **a**: and **ɐ** vowels are more likely to be associated with alveolarization of the following velar consonant, as both **a**: and **ɐ** vowels are central and do not in particular favour

alveolarization any more than they would disfavour velar retention, from the perspective of conservation of articulatory effort.

When the results of the current study are compared with those of Wong (2005), it can be seen that the trends with which the preceding vowels are associated with (**ŋ**) alveolarization are, overall, fairly similar (leaving aside the differences in experimental conditions between the two studies for the moment). The results are shown in Figure 29, which shows percentage of (**ŋ**) realization as [alveolar] in all speech styles.

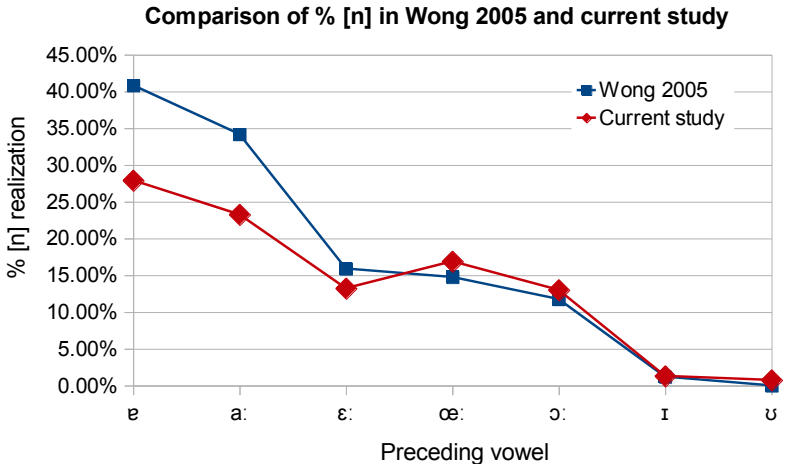


Figure 29. Comparison of percentage realization of (ŋ) as [n] by preceding vowel between Wong (2005) and current study

The only major source of divergence is in the percentages of alveolarization for vowels **a**: and **ɐ**, which are noticeably lower in the current study. While this divergence cannot easily be explained without verifying whether this difference was a real effect or whether it was due to different auditory classification methods used by Wong versus those used in this study, the overall trend shows that the order of vowels associated with alveolarization is strikingly similar, with the only major difference being that the **œ**: vowel was found in this study to be more likely associated with alveolarization than the

ɛ: vowel although the difference is small. The next section examines following onset and how they might influence (**ŋ**) and (**k**) realization.

6.2 Following onsets

The results from Section 4.1 showed that following onsets were significant in predicting [alveolar] and [reduced] realizations in (**ŋ**) and (**k**) variation. The results are summarized in Tables 63 and 64. For a list of how following onset was coded, see coding scheme in Appendix VI.

Table 63. Multinomial logistic regression results for following onset effects on (ŋ) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
Internal factors						
(ŋ) > [n] variation						
- Following Onset (Ref.: Velar)						
Labial	1.852	0.255	7.263	3.793×10^{-13}	6.373	537.3%
Palatal	3.937	0.237	16.612	5.724×10^{-62}	51.265	5026.5%
Alveolar	3.371	0.284	11.870	1.700×10^{-32}	29.108	2810.7%
Other	2.567	0.242	10.607	2.752×10^{-26}	13.027	1202.7%
(ŋ) > [reduced] variation						
- Following Onset (Ref.: Velar)						
Labial	3.080	0.143	21.538	6.793×10^{-103}	21.758	2075.8%
Palatal	2.954	0.177	16.689	1.569×10^{-62}	19.183	1818.3%
Alveolar	1.905	0.142	13.415	4.907×10^{-41}	6.719	571.9%
Other	1.811	0.143	12.664	9.321×10^{-37}	6.117	511.7%

Considering (**ŋ**) to [alveolar] realization, when (**ŋ**) was followed by an onset segment with the feature 'alveolar', the odds with which (**ŋ**) was alveolarized was 29.1 times that when the feature of the following onset is 'velar'. Other following onset features, namely 'labial', 'palatal' and 'other' (which includes pauses, such as phrase boundaries; and zero following onsets, where following syllables do not have an onset consonant—see Section 2.3.2 on the Cantonese syllable structure), were also more likely

(at 6.4, 51.3, and 13.0 times, respectively) to favour [alveolar] realization than a following 'velar' onset. Taken together, and as expected, a following 'alveolar' onset feature was very likely to favour [alveolar] realization of (ŋ), but the 'palatal' onset feature was, unexpectedly, even more highly likely to favour [alveolar] realization. A following 'labial' feature was far less likely to favour [alveolar] realization, followed by a 'velar' feature, which was the least likely to favour [alveolar] realization.

As for (ŋ) to [reduced] variation, when (ŋ) was followed by an onset segment with the feature 'labial' or 'palatal', the odds with which (ŋ) was [reduced] were respectively 21.8 and 19.2 times the odds compared to when the feature of the following onset is 'velar'. Other following onset features, namely 'alveolar' and 'other' (which includes pauses and zero following onsets), were also more likely (at 6.7, and 6.1 times, respectively) to favour [reduced] realization than a following 'velar' onset. Taken together, a following 'velar' feature was the least likely to favour [reduced], followed by 'other' and 'alveolar'. The following features 'labial' and 'palatal' were the most highly likely to favour [reduced] realization of (ŋ).

Table 64. Multinomial logistic regression results for following onset effects on (k) variation

Response Ref. Cat.: [velar]	β (regress. coeff.)	Standard Error	Z-score (β / S.E.)	p-value (2-tailed)	Odds ratios (Exp(β))	% diff. in odds (cf. ref. cat.)
Internal factors						
(k) > [t] variation						
- Following Onset (Ref.: Velar)						
Labial	2.963	0.397	7.463	8.426×10^{-14}	19.356	1835.6%
Palatal	5.915	0.443	13.352	1.151×10^{-40}	370.554	36955.4%
Alveolar	5.301	0.386	13.733	6.427×10^{-43}	200.537	19953.7%
Other	3.756	0.389	9.656	4.458×10^{-22}	42.777	4177.7%
(k) > [reduced] variation						
- Following Onset (Ref.: Velar)						
Labial	3.952	0.185	21.362	3.006×10^{-101}	52.039	5103.9%
Palatal	3.750	0.271	13.838	1.511×10^{-43}	42.521	4152.1%
Alveolar	2.434	0.176	13.829	1.691×10^{-43}	11.404	1040.4%
Other	3.853	0.177	21.768	4.629×10^{-105}	47.134	4613.4%

In terms of (**k**) to [alveolar] variation, when (**k**) was followed by an onset segment with the feature 'alveolar', the odds with which (**k**) was alveolarized was 200.5 times the odds than when the feature of the following onset is 'velar'. Other following onset features, namely 'labial', 'palatal' and 'other' (which includes phrasal level pauses and syllables with zero onsets), were also more likely (at 19.4, 370.6, and 42.8 times, respectively) to favour [alveolar] realization than a following 'velar' onset. Taken together, a following 'velar' feature was the least likely to favour [alveolar] realization, followed by 'labial' and 'other' features. The features 'palatal' and 'alveolar' were both highly likely to favour [alveolar] realization of (**k**).

As for [reduced] variation of (**k**), when (**k**) was followed by an onset segment with the feature 'labial', 'palatal', 'alveolar' and 'other' the odds with which (**k**) was reduced were, respectively, 52.0, 42.5, 11.4 and 47.1 times the odds when the feature of the following onset is 'velar'. Overall, a following 'velar' feature was the least likely to favour [reduced] realization. 'Palatal', 'other' and 'labial' features were then (in increasing order) highly likely to favour [reduced] realization of (**k**).

Summarizing these results, it was found that a following 'velar' feature was consistently disfavouring innovative realizations [alveolar] and [reduced], for both of our variables (**ŋ**) and (**k**). The following features 'palatal' and 'alveolar' were both very likely to favour [alveolar] realization of (**ŋ**) and (**k**). The somewhat unexpected trend is that the following feature 'palatal' favours [alveolar] realization even more strongly than does a following 'alveolar', the latter of which we would expect to display the highest amount of assimilatory effects and to favour [alveolar] realization the most out of all other feature types. One reason could be that there was an imbalance in the data, whereby there was a far higher number tokens containing the short vowels **ɪ** and **ʊ**

(which have been shown to disfavour [alveolar] realization) that are followed by a onset with an 'alveolar' feature. Since there was not sufficient data to allow the interaction between preceding vowels and following onset to be explored in the full model, it was not possible to control for this imbalance in data. This issue will be discussed in more detail in the univariate statistical description of (**ŋ**) realization below.

To provide a sense of the distribution of the raw data with respect to following onsets, the percentages of (**ŋ**) realization by following onset is shown in Figure 30, ranked according to decreasing percentage of [-n] realization. The raw counts which made up these percentages are shown in Table 65, along with percentage [-n] realization. From this "naïve" univariate overview, the following velar onsets, **k**, **ŋ**, **k^{h}**, and **k^{w}** favoured [-n] the least, although this was also to be expected and is validated by results in the full statistical model. The set of coronal onsets in this "naïve" statistical description favoured [-n] the most, the onsets being, in decreasing order, **n**, **t**, **t^{h}**, **ts**, **ts^{h}**, and **l**. On the other end, it seemed that bilabials and palatals highly favour [reduced] realizations of (**ŋ**), which is also validated by the full statistical model.

One unexpected outlier is a following **s**. As a coronal, one would expect it to favour alveolarization. But a closer inspection of the tokens of (**ŋ**) after which an **s** onset followed revealed that these tokens of (**ŋ**) had a high percentage of short preceding vowels. As the results in Section 6.1 on preceding vowels show, the short vowels **ɪ** and **ʊ** were very unlikely to lead to alveolarization, and as a direct consequence of this, those tokens followed by an **s** had a very low rate of alveolarization. This is a confounding factor that needs to be taking into consideration when interpreting the effects of the following onsets, but such interaction between factors is common in natural language.

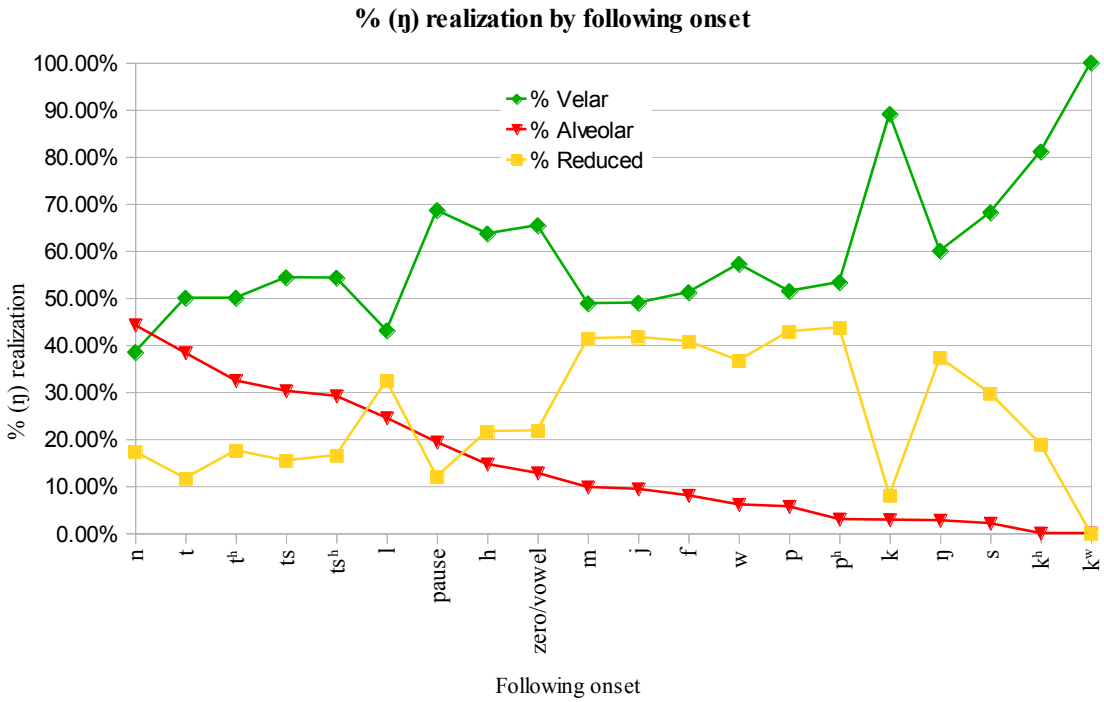


Figure 30. Percentage (ŋ) realization by following onset

N= 8392

Table 65. Raw counts of (ŋ) realization by following onset, including percentage [-n]

N=8392

	n	t	tʰ	ts	tsʰ	l	pause	h	zero/ vowel	m	j	f	w	p	pʰ	k	ŋ	s	kʰ	kʷ
Velar [ŋ]	20	137	37	169	95	109	901	638	87	399	308	236	357	336	72	784	66	227	103	36
Alveolar [n]	23	105	24	94	51	62	254	147	17	80	59	37	38	37	4	25	3	7	0	0
Reduced	9	32	13	48	29	82	158	217	29	338	262	188	229	280	59	71	41	99	24	0
Total	**52**	**274**	**74**	**311**	**175**	**253**	**1313**	**1002**	**133**	**817**	**629**	**461**	**624**	**653**	**135**	**880**	**110**	**333**	**127**	**36**
% velar [ŋ]	38.4%	50.0%	50.0%	54.3%	54.3%	43.1%	68.6%	63.7%	65.4%	48.8%	49.0%	51.2%	57.2%	51.5%	53.3%	89.1%	60.0%	68.2%	81.1%	100.0%
% [n]	44.2%	38.3%	32.4%	30.2%	29.1%	24.5%	19.3%	14.7%	12.8%	9.8%	9.4%	8.0%	6.1%	5.7%	3.0%	2.8%	2.7%	2.1%	0.0%	0.0%
% reduced	17.3%	11.7%	17.6%	15.4%	16.6%	32.4%	12.0%	21.7%	21.8%	41.4%	41.7%	40.8%	36.7%	42.9%	43.7%	8.1%	37.6%	30.0%	18.9%	0.0%

In terms of (**k**) realization, "naïve" univariate statistics suggest that following coronals **tʰ**, **t**, **ts**, and **tsʰ** favoured alveolarization, although this was not the case for coronals **l** and **n**. The following onsets which seemed to favour alveolarization the least were the velars **kʰ** and **kʷ**. As with (**ŋ**), the palatals, **j** and bilabials **w**, **p** were not found

to particularly favour alveolarization, but seemed to be associated with high rates of [reduced] variants. This is validated by the full statistical model. Finally, [velar] retention was associated with following velar onsets, **k**, **k^{h}**, and **k^{w}**, as would be expected, and the results from the full statistical model concur.

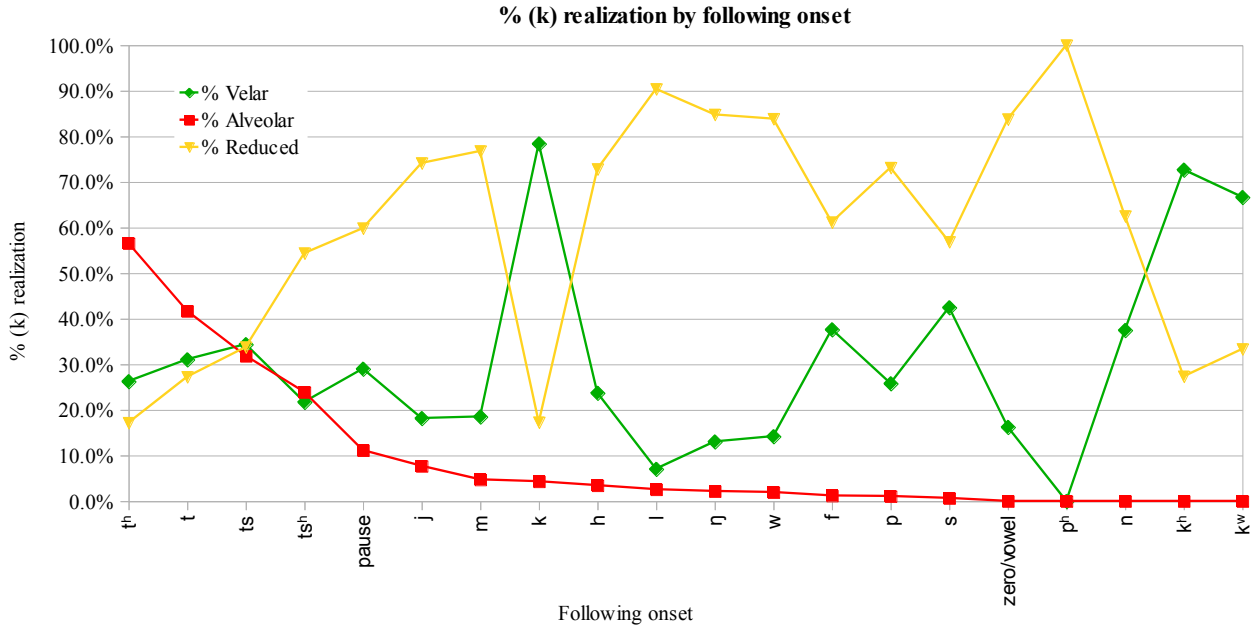


Figure 31. Percentage (k) realization by following onset

Table 66. Raw counts of (k) realization by following onset, including percentage [-t]

	tʰ	**t**	**ts**	**tsʰ**	**pause**	**j**	**m**	**k**	**h**	**l**	**ŋ**	**w**	**f**	**p**	**s**	**zero/ vowel**	**pʰ**	**n**	**kʰ**	**kʷ**
Velar [k]	26	73	54	10	351	26	106	218	206	11	6	51	189	120	65	13	0	3	109	2
Alveolar [t]	56	98	50	11	135	11	27	12	30	4	1	7	6	5	1	0	0	0	0	0
Reduced	17	64	53	25	724	106	439	48	632	141	39	301	307	340	87	67	2	5	41	1
Total	**99**	**235**	**157**	**46**	**1210**	**143**	**572**	**278**	**868**	**156**	**46**	**359**	**502**	**465**	**153**	**80**	2	8	150	3
% velar [k]	26.3%	31.1%	34.4%	21.7%	29.0%	18.2%	18.5%	78.4%	23.7%	7.1%	13.0%	14.2%	37.6%	25.8%	42.5%	16.3%	0.0%	37.5%	72.7%	66.7%
% alveolar [t]	56.6%	41.7%	31.8%	23.9%	11.2%	7.7%	4.7%	4.3%	3.5%	2.6%	2.2%	1.9%	1.2%	1.1%	0.7%	0.0%	0.0%	0.0%	0.0%	0.0%
% reduced	17.2%	27.2%	33.8%	54.3%	59.8%	74.1%	76.7%	17.3%	72.8%	90.4%	84.8%	83.8%	61.2%	73.1%	56.9%	83.8%	100.0%	62.5%	27.3%	33.3%

6.3 Tautosyllabic onset

Recall from Section 1.6 that Labov's reassessment of Skikun velarization found that tautosyllabic onset features 'velar', 'apical', 'labial' and 'nasal' were found to favour velarization, whereas the onset features 'uvular' and 'laryngeal' were found to disfavour velarization. While tautosyllabic onset was not included as an explanatory factor in the full statistical model, "naïve" univariate statistics are explored here to infer the possible broader effects that tautosyllabic onset might have on (**ŋ**) and (**k**) variation.

Figures 32 and 33 show percentage (**ŋ**) and (**k**) realization by tautosyllabic onset. The tautosyllabic onsets are presented according to decreasing percentages of [alveolar] realization.

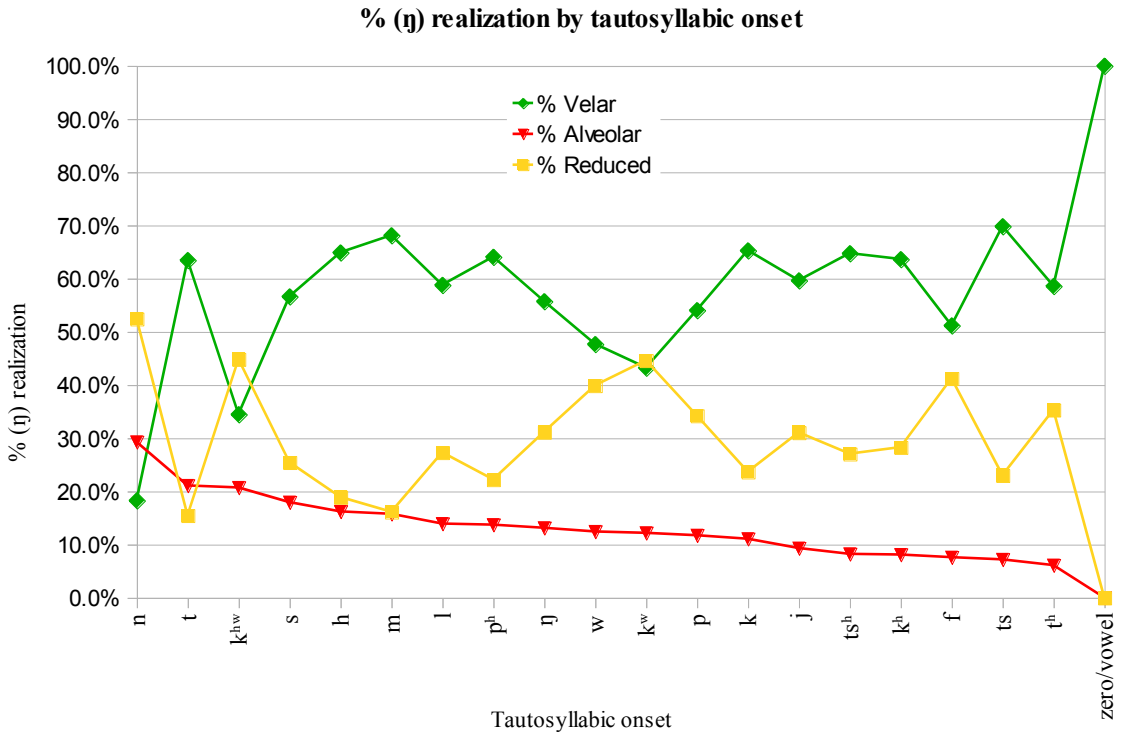


Figure 32. Percentage (ŋ) realization by tautosyllabic onset

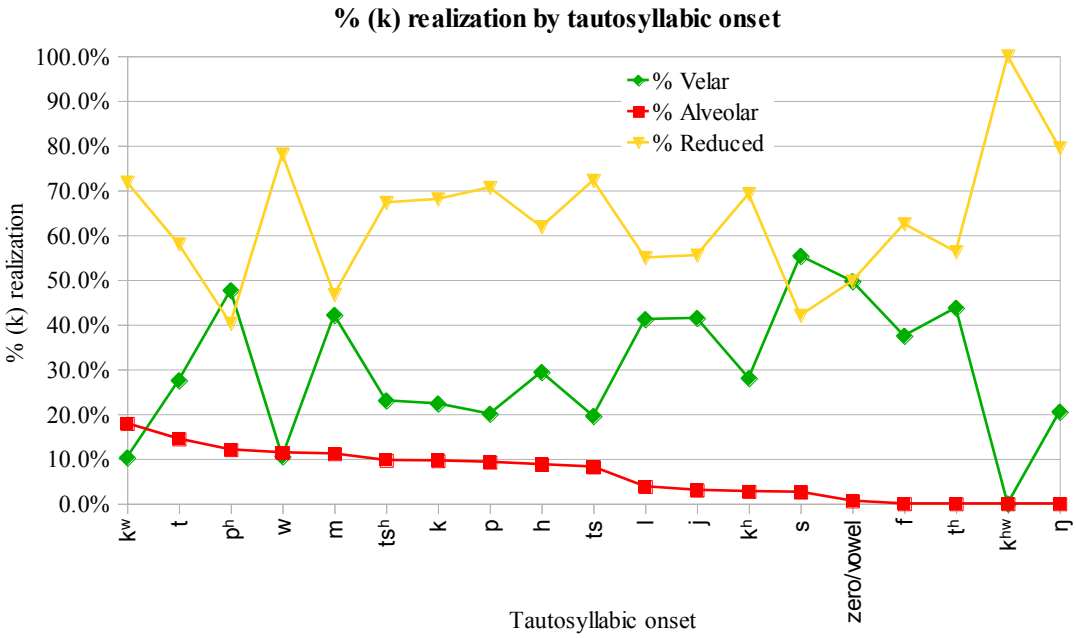


Figure 33. Percentage (k) realization by tautosyllabic onset

It can be seen that the amount of variation of (**ŋ**) and (**k**) realization is high within the groups labials (**m**, **p**, **p^{h}**), apicals (**n**, **t**, **l**, **s**, **ts**, **ts^{h}**), and velars (**k**, **k^{h}**, **ŋ**, **kw**, **k^{hw}**). Unlike following onset, tautosyllabic onset was not a good predictor of (**ŋ**) and (**k**) realization as [velar] or [alveolar]. Thus, dissimilatory effects such as those found by Labov when reassessing the Skikun velarization data, whereby "the place of articulation of the initials is differentiated as much as possible from the place of articulation of the final" (Labov 1994:446) and features such as 'labials', 'apicals', and 'velars' were found to be favouring velarization, and 'uvular' and 'pharyngeal' were found to be disfavouring velarization, could not be found in this study.

6.4 Influence of tone

Neogrammarians have typically dealt with Indo-European languages, not languages which have tonal contrasts. However, as tone is considered part of the phonology of the morphosyllable in Cantonese and within the immediate phonological environment of the velar codas **-ŋ** and **-k**, the effect that tone might have on velar coda realization is also discussed here. This section explores the influence of tone effects on variation of (**ŋ**). Tonal effects were not included in (**k**) variation in the full mathematical model, as inclusion meant that the estimation procedure could not converge. The results of the full model estimate for (**ŋ**) are summarized in Table 67.

Table 67. Multinomial logistic regression results for tone effects on (ŋ) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
Internal factors						
(ŋ) > [n] variation						
- Tone (Ref.: T1)						
T4	-0.274	0.104	-2.635	8.423×10^{-3}	0.760	-24.0%
T6	0.311	0.142	2.190	2.851×10^{-2}	1.36 5	36.5%
T2	0.024	0.121	0.198	0.843	1.024	2.4%
T3	-0.084	0.224	-0.375	0.708	0.919	-8.1%
T5	-0.157	0.197	-0.797	0.425	0.855	-14.5%
(ŋ) > [reduced] variation						
- Tone (Ref.: T1)						
T4	0.313	0.076	4.118	3.815×10^{-5}	1.368	36.8%
T2	0.102	0.101	1.010	0.313	1.107	10.7%
T3	-0.184	0.17	-1.082	0.279	0.832	-16.8%
T5	0.026	0.172	0.151	0.880	1.026	2.6%
T6	0.13	0.108	1.204	0.229	1.139	13.9%

Considering (**ŋ**) to [alveolar] variation, the tones which favour [alveolar] realization the most is T6, followed by T2, T1, T3, T5, with T4 favouring [alveolar] realization the least. As for (**ŋ**) to [reduced] variation, T4 favours [reduced] realization the most, followed by T6, T2, T5, and T1, with T3 favouring [reduced] realization the

least. However, some caution needs to be exercised when interpreting the effect of tone on (**ŋ**) realization, as both tone and preceding vowel are features of the same syllable and are likely to be confounding factors.

Figure 34 is a univariate graph showing percentage of (**ŋ**) realization ranked by decreasing percentage of [alveolar] realization, while Figure 35 shows the same for (**k**).

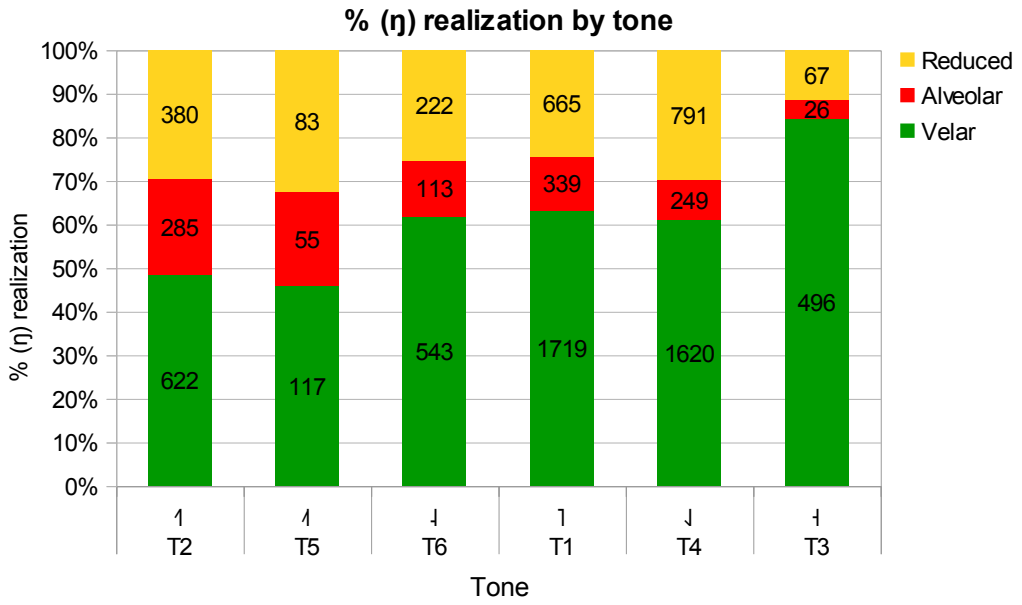


Figure 34. Percentage (ŋ) realization by tone

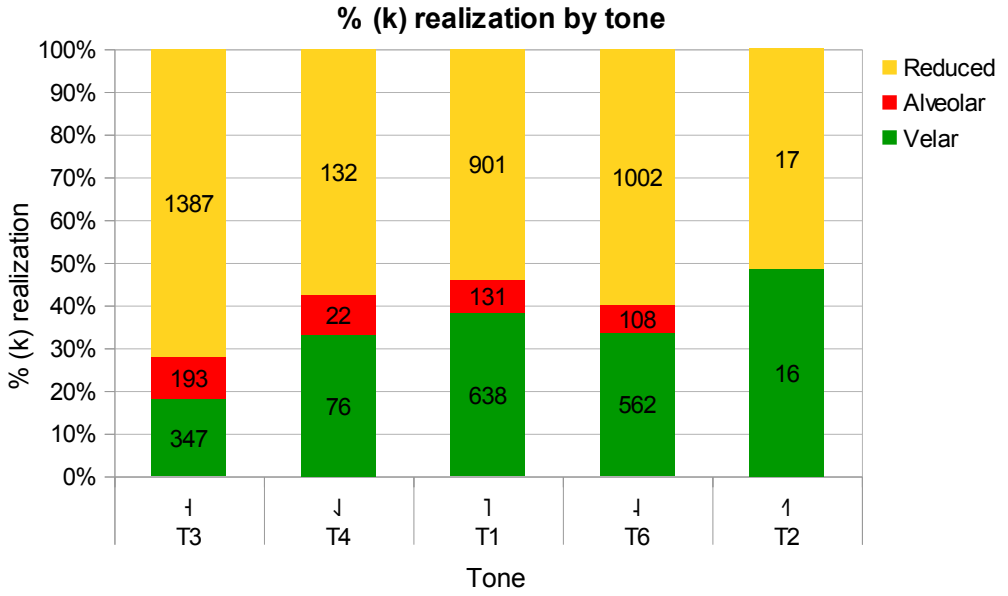


Figure 35. Percentage (k) realization by tone

These figures above are merely provided to give a broad impression of the numbers of overall token distribution across different tones.

6.5 Suspension of phonological contrast

Recall from the discussion at the end of Section 2.3.4 the question which was posed with regards to whether rimes such as **ɪŋ**, **ɪk**, **ɛ**:**ŋ**, **ɛ**:**k**, **ʊŋ**, **ʊk**, **œ**:**ŋ**, and **œ**:**k**, which do not have firmly established alveolar counterparts, undergo alveolarization more readily than rimes such as **ɔ**:**ŋ**, **ɔ**:**k**, **a**:**ŋ**, **a**:**k**, **ɐŋ**, and **ɐk**, for which alveolar counterparts do exist, and for which confusion with existing morphosyllables could result when the coda is alveolarized. From the results presented in Models 5 and 6 (in 0), part of which is duplicated here as Table 68, it can be seen that no support could be found for any difference between the rime set **ɪŋ**, **ɪk**, **ɛ**:**ŋ**, **ɛ**:**k**, **ʊŋ**, **ʊk**, **œ**:**ŋ**, and **œ**:**k** and the rime set **ɔ**:**ŋ**, **ɔ**:**k**, **a**:**ŋ**, **a**:**k**, **ɐŋ**, and **ɐk**, with respect to how likely they are to undergo alveolarization or reduction. In other words, the availability of an alveolar phonemic

counterpart (and the concomitant possibility for suspension of phonological contrast) did not make it more or less likely for velar nasal (**ŋ**) or (**k**) to be realized as [alveolar] or [reduced].

Table 68. Comparison of rate of [n] realization in rime sets with and without alveolar counterparts

(ŋ) → [n] (Ref. category: [ŋ])	Ref. category: ɐ	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score	*p*-value	Odds ratios	% diff. in odds
ɪŋ, ɛːŋ, ʊŋ, œːŋ	no alveolar counterpart exists						
	ɛː	-0.858	0.269	-3.190	1.425×10^{-3}	0.424	-57.60%
	ɪ	-4.351	0.354	-12.291	1.013×10^{-34}	0.013	-98.70%
	ʊ	-5.253	0.457	-11.495	1.405×10^{-30}	0.005	-99.50%
	œː	-0.153	0.232	-0.659	5.096×10^{-1}	0.858	-14.187
ɔːŋ, aːŋ	alveolar counterpart exists						
	ɔː	-1.024	0.209	-4.9	9.607×10^{-7}	0.359	-64.10%
	aː	0.356	0.213	1.671	9.5×10^{-2}	1.428	42.76%

6.6 Discussion

In the above results, it can be seen that both preceding vowel and following onset were good candidates for predicting (**ŋ**) and (**k**) realization. In terms of preceding vowels, the short vowels **ɪ** and **ʊ** were found to be associated with high rates of [velar] retention for both (**ŋ**) and (**k**). What was also consistent with previous studies was that **ɐ** and **a**ː were associated with the highest rates of [alveolar] realization. As for vowel categories **ɛ**ː, **œ**ː and **ɔ**ː, they showed a high amount of variation between the [velar], [alveolar] and [reduced] outcomes. On the basis of these patterns of variation, it seems hard to justify that these patterns arose from assimilation with any single phonetic feature of the

preceding vowel, whether it be tongue height or frontness, or vowel length. The vowels **ɐ** and **a**: are of different length and height but same frontness; the vowels **ɛ**:, **œ**: and **ɔ**: are of the same tongue height and length, with different frontness; and **ɪ** and **ʊ** are of the same tongue height and different frontness. Yet, these vowel sets were found to behave similarly within, and very differently across sets. One could perhaps argue that the best phonetic measures for predicting likelihood of alveolarization could be 'open' and 'long' vowels combined, but it is unclear how alveolarization would supposedly arise from assimilation with a lower tongue height, while the most progressive short **ɐ** vowel would then be excluded from this generalization. Based on these results, it seems difficult to argue that the patterns of variation in (**ŋ**) and (**k**) realization in contemporary Hong Kong Cantonese are due to assimilation to the phonetic feature of tongue height in the preceding vowel, which is the interpretation that Yue (2002) adopted for explaining the tendency of how Middle Chinese ***-k** developed into **-t** across Yue dialects (see Section 2.7). This interpretation of vowel height assimilation is not supported by the results obtained from this study. However, overall, each vowel category still seemed a good predictor of the amount of [velar] retention and [alveolar] realization.

As for following onset, it was also found that (**ŋ**) and (**k**) realization fell fairly neatly into groups according to phonetic features of the following segment, namely with 'apical' associated with the highest rates of [alveolar] realization and 'velar' associated with the highest rates of [velar] retention. The results from our statistical model in Chapter 4 and our exploration of the data in this chapter suggest that assimilation with phonetic features of the following onset is likely a better explanation for patterns of (**ŋ**) and (**k**) realization into [velar], [alveolar] and [reduced] categories.

We found that there was a large amount of variation within each tautosyllabic onset category. Unlike with the following onset segment, (**ŋ**) and (**k**) realization did not correspond neatly to phonetic features in tautosyllabic onsets. So, overall, tautosyllabic onsets seem a bad predictor for [velar], [alveolar] and [reduced] realization of (**ŋ**) and (**k**).

The effect of tone on (**ŋ**) and (**k**) realization was present, but it was discussed that the effects of tone might be confounded with effects of preceding vowel.

Returning to the notion of neogrammarian regularity, which is the overarching objective for exploring the effects of neighbouring phonological environments, it was found that the only phonological environment that seemed a good predictor for (**ŋ**) and (**k**) realization patterns was preceding vowel category. Tautosyllabic onset was found to not be a good predictor, unlike in Labov's reassessment of Skikun velarization. The following onset segment was also found to be a very good predictor of (**ŋ**) and (**k**) realization, although this is not strictly a phonological environment, but a phonetic one, and dependent on the lexemes which collocate with lexemes containing (**ŋ**) and (**k**). Since preceding vowels are good predictors, one can draw the conclusion that neogrammarian regularity is observed in (**ŋ**) and (**k**) variation patterns.

6.7 Summary

In this chapter, the effects of preceding vowel, following onset, tautosyllabic onset and tone on (**ŋ**) and (**k**) realization were explored. It was found that preceding vowel and following onset segment were the best predictors of (**ŋ**) and (**k**) realization patterns as supported by the statistical model presented in Chapter 4. If we are to take the preceding

vowel as the primary phonological environment, then broadly speaking, it can be said that (**ŋ**) and (**k**) display neogrammarian regularity on this account. However, unlike Labov's finding that Skikun velarization was favoured by tautosyllabic onset phonetic features 'apical', 'velar', 'labial' and 'nasal', it was found that phonetic features of the tautosyllabic onset were not good predictors in the case of (**ŋ**) and (**k**) realization.

The next chapter is on lexical diffusion and lexical frequency effects, to test how lexical measures fare against phonological measures in predicting (**ŋ**) and (**k**) realization.

7 Lexical diffusion

This chapter explores the lexical dimension of (**ŋ**) and (**k**) variation. As a short recap of the discussion in Chapter 1 on what would constitute lexical diffusion, it was suggested that if the usage-based prediction that lexical frequency has an effect on sound change can be attested, and frequent (or infrequent) items can be shown to lead a particular sound change, this would constitute strong evidence for the sound change transmitting via lexical diffusion. First, lexical frequency effects are examined, and the findings from Chapter 4 that the higher the lexical frequency, the more likely it was that (**ŋ**) and (**k**) were realized with reduced variants, and the less likely (**k**) was realized with alveolar variants are presented. A qualitative analysis of lexemes which might be leading the change will then be offered. An investigation of homophonous lexemes will also be presented, to offer a comparison with Chen & Hsieh's (1971) study of homonym split from Middle Chinese ***-ai** into modern Peking Mandarin **-ai**/**-a**, discussed in Section 1.5.1.2.

7.1 Lexical frequency effects

In this section, the results from the full model for lexical frequency effects are discussed in detail, summarized again in Tables 69 and 70 for ease of reference.

Table 69. Multinomial logistic regression results for lexical frequency effects on (ŋ) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
Internal factors						
(ŋ) > [n] variation						
- Lexical Frequency (LexFreq)						
log(freq): mean-centered	-0.306	0.264	-1.159	0.246	0.736	-26.4%
(ŋ) > [reduced] variation						
- Lexical Frequency (LexFreq)						
log(freq): mean-centered	0.526	0.174	3.023	2.501×10^{-3}	1.692	69.2%

In terms of describing (**ŋ**) to [n] variation, lexical frequency—expressed as log(frequency)—was found not to be significant. However, log(frequency) had a significant effect on (**ŋ**) to [reduced] variation. The results suggest that for every 1-unit increase in log(frequency) the likelihood of [reduced] realization compared to the reference category [velar] realization increased by 69.2%. In other words, the higher the lexical frequency, the more likely it was that (**ŋ**) was realized with reduced variants, with all else being equal. This finding conforms with the prediction put forward in studies of lexical frequency effects in sound change, discussed in the literature review in Section 1.7.

As for (**k**), the results in Table 70 show that log(frequency) had a significant effect on both realization of (**k**) as [alveolar] and realization of (**k**) as [reduced].

Table 70. Multinomial logistic regression results for lexical frequency effects on (k) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
Internal factors						
(k) > [t] variation						
- Lexical Frequency (LexFreq)						
log(freq): mean-centered	-1.311	0.612	-2.142	0.032	0.270	-73.0%
(k) > [reduced] variation						
- Lexical Frequency (Ref.: Infreq.)						
log(freq): mean-centered	0.492	0.162	3.037	2.389×10^{-3}	1.635	63.6%

For every 1-unit increase in log(frequency) the likelihood of [alveolar] realization decreased by -73.0%, while the likelihood of [reduced] realization increased by 63.6%, with all else being equal. In other words, the more frequent a lexical item, the more likely (**k**) was reduced (as with (**ŋ**) above), but also the less likely (**k**) was realized with [alveolar]. This finding suggests that the reductive sound change of (**k**) is showing lexical frequency effects, with more frequent items leading the change. On the other hand, the sound change involving increased use of alveolar realization is led by infrequent items. These findings will be discussed in further detail in Section 7.5.

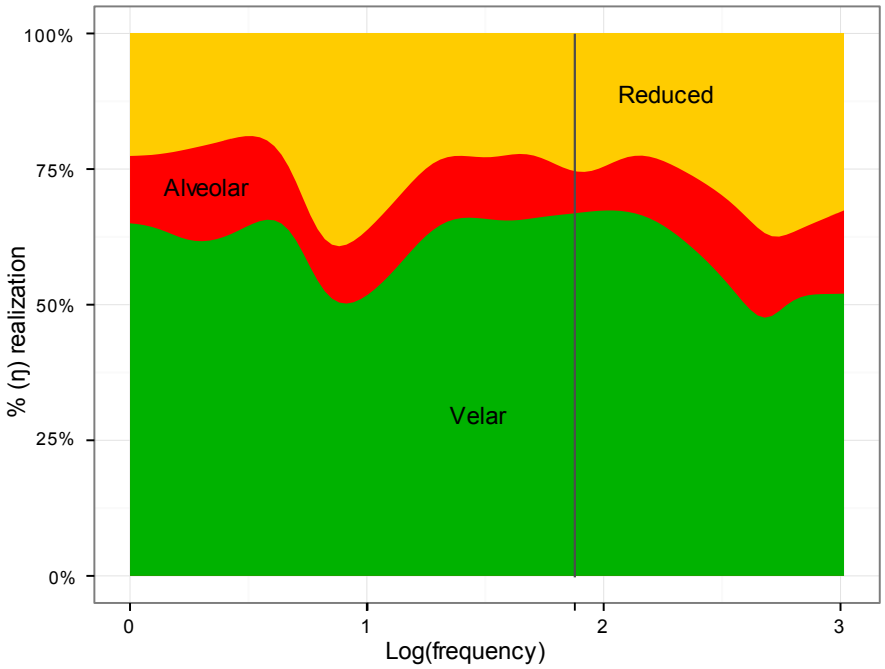
The importance of employing statistical modelling to compare factors becomes apparent when we explore "naïve" univariate statistics and with the use of a conditional probability density plot, which shows the overall probably distribution of and applies smoothing to (**ŋ**) and (**k**) realizations as [velar], [alveolar] and [reduced] over the continuous log(frequency) measure. This is shown in Figures 36 and 37.[34] The proportion of [velar] realization is in green, at the bottom of the graph, [alveolar] realization is in red in the middle, and [reduced] variants in yellow at the top of the graph.

In Figure 36, the plot hardly shows any effect of log(frequency) on (**ŋ**) realization as [alveolar]. Yet, from our full statistical model, with all other factors kept equal, there was a significant effect of log(frequency) on (**ŋ**) realization as [alveolar], with an increase in log(frequency) accounting for a proportional increase in [alveolar].

---

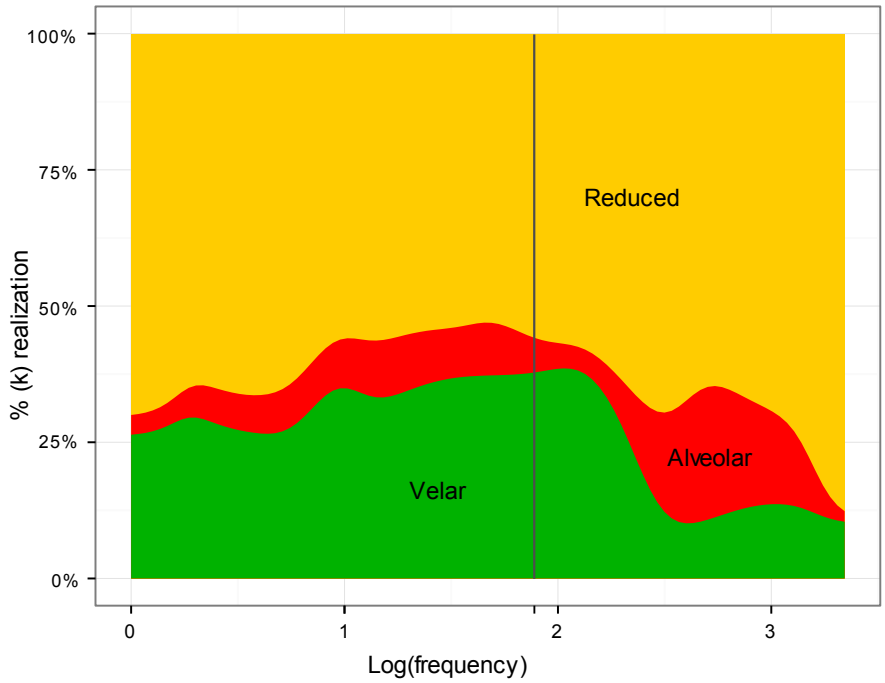
[34] As with conditional probability density graphs for (**ŋ**) and (**k**) presented for age-based variation in Section 5.1, these figures were plotted using the `cdplot` function (written by Achim Zeileis) in the R package `ggplot2` (Wickham & Chang 2013). Function smoothing was obtained by `cdplot` through default kernel density estimation in the `density` function (R Core Team and contributors 2013).

The same is true for Figure 37. While the plot seems to suggest an increase in [alveolar] realization of (**k**) with an increase in log(frequency), when all other factors are controlled for, the overall effect on [alveolar] realization is in fact a decrease. This illustrates the importance of including all explanatory factors in the same model so that all other factors can be kept equal when describing the effect of one explanatory variable on any given dependent variable.



HKCAC frequency median for (**ŋ**) = 75; Log(median) = 1.875, indicated by grey vertical line; N = 8392 tokens of (**ŋ**), including all phonological environments.

Figure 36. Conditional probability density curve of (ŋ) realization by log(frequency)



HKCAC frequency median for (**k**) = 78; Log(median) = 1.89, indicated by grey vertical line; N = 5532 tokens of (**k**).

Figure 37. Conditional probability density curve of (k) realization by log(frequency)

In summary, when generalizing across all the lexemes of (**ŋ**) and (**k**) in this study, there was an overall lexical frequency effect, for (**ŋ**) to [reduced] and (**k**) to [alveolar] as well as [reduced] variation, with all other factors being held equal, when lexical frequency was measured as mean-centered 'log(frequency)'. Overall, the usage-based hypothesis described in Chapter 1, that the higher the lexical frequency, the more likely lenition occurs was borne out.

7.2 Which lexemes lead the change?

The aim in this section is to explore the extent to which certain lexemes lead the change. In order to investigate this, it was considered best to compare like with like as far as possible, which in this case would be lexemes which share the same preceding vowel (i.e. the rimes **-ɐŋ**, **-a**:**ŋ**, **-ɛ**:**ŋ**, **-œ**:**ŋ**, **-ɔ**:**ŋ**, **-ɪŋ**, and **-ʊŋ**, respectively). When lexemes with different preceding vowels are found to respond differently with respect to lexical frequency effects, this could constitute proof that certain lexemes lead the change of (**ŋ**) to [alveolar] or [reduced] variants.

In order to achieve this, let us turn to the results obtained from the full model of (**ŋ**) varation on the interaction between preceding vowels and lexical frequency. The results are summarized in Table 71. As the statistical results show, there is at least a subset of preceding vowels for which there is a significant interaction effect with lexical frequency, expressed as log(frequency), on (**ŋ**) variation. The fact that there is statistically significant difference between (some) preceding vowel categories with respect to lexical frequency *logically implies* that it *must* be the case that lexemes with certain vowel categories show more progressive innovative variants ([alveolar] or [reduced]) with an increase in lexical frequency than other vowel categories.

Table 71. Multinomial logistic regression results for lexical frequency * preceding vowel effects on (ŋ) variation

Response Ref. Cat.: [velar]	*β* (regress. coeff.)	Standard Error (S.E.)	Z-score (*β* / S.E.)	*p*-value (2-tailed)	Odds ratios (Exp(*β*))	% diff. in odds (cf. ref. cat.)
Internal factors						
(ŋ) > [n] variation						
- Lexical Frequency (LexFreq)						
log(freq): mean-centered	-0.306	0.264	-1.159	0.246	0.736	-26.4%
- LexFreq * Vowel(ɛ: vs. ɐ)	-0.651	0.220	-2.960	3.085×10^{-3}	0.522	-47.8%
- LexFreq * Vowel(ɪ vs. ɐ)	-0.884	0.298	-2.966	3.013×10^{-3}	0.413	-58.7%
- LexFreq * Vowel(ɔ: vs. ɐ)	-0.314	0.125	-2.512	1.200×10^{-2}	0.731	-26.9%
- LexFreq * Vowel(œ: vs. ɐ)	-0.284	0.142	-2.000	4.550×10^{-2}	0.753	-24.7%
- LexFreq * Vowel(a: vs. ɐ)	0.2	0.126	1.587	0.112	1.221	22.1%
- LexFreq * Vowel(ʊ vs. ɐ)	-0.061	0.502	-0.122	0.903	0.941	-5.9%
(ŋ) > [reduced] variation						
- Lexical Frequency (LexFreq)						
log(freq): mean-centered	0.526	0.174	3.023	2.501×10^{-3}	1.692	69.2%
- LexFreq * Vowel(ɛ: vs. ɐ)	-0.615	0.166	-3.705	2.115×10^{-4}	0.541	-45.9%
- LexFreq * Vowel(ɔ: vs. ɐ)	-0.591	0.130	-4.546	5.464×10^{-6}	0.554	-44.6%
- LexFreq * Vowel(a: vs. ɐ)	0.012	0.139	0.086	0.931	1.012	1.2%
- LexFreq * Vowel(ɪ vs. ɐ)	-0.163	0.196	-0.832	0.406	0.850	-15.0%
- LexFreq * Vowel(œ: vs. ɐ)	-0.236	0.138	-1.710	0.087	0.788	-21.0%
- LexFreq * Vowel(ʊ vs. ɐ)	0.083	0.155	0.535	0.592	1.087	8.7%

While the interpretation of the exact way in which preceding vowel and lexical frequency interact is rather complex and considered beyond the scope of this dissertation, it might be instructive to explore the univariate plots showing how lexemes with different vowels respond to lexical frequency to obtain an impression of how much variation there can be between different lexemes across the lexical frequency scale. This can be done by listing the lexemes within each rime category by HKCAC lexical frequency to allow us to compare how much each lexeme is produced with [velar], [alveolar], and [reduced] realizations respectively. In the following graphs, only lexemes for which 10 tokens or more could be obtained were included to reduce bias from small token counts. The Chinese character for each lexeme is presented with a phonological transcription. The number in square brackets above the lexeme is the number of tokens from which percentages in the graph were obtained. One has to keep in mind, however, that these are mere overall counts, with other factors not having been

controlled for. Thus, any differences between preceding vowel categories could be due to other confounding effects.

First, lexemes with the **-ɐŋ** rime, are presented in Figure 38 from high frequency (401) to low frequency (1). Overall, there seems to be a gradual increase in [velar] retention when going down from higher frequency lexemes, with the exception of **nɐŋ˩** 能, which has a frequency count of 401, and **pʰɐŋ˩** 朋, with a frequency count of 203, both of which are fairly high. The rates of [velar] retention in these lexemes is particularly low, with concomitant higher rates of [alveolar] and [reduced] realizations. Comparing the difference in percentages of realization of these two lexemes in relationship to other lexemes with the same **-ɐŋ** rime, it seems as if these two lexemes were spear-heading the change from [velar] realization to [alveolar] or [reduced] realizations. However, if we focus on [alveolar] realization alone, we see that the lexemes **hɐŋ˧** 幸 and **pɐŋ˩** 憑, which have relatively low frequencies of 10 and 3, respectively, have fairly high rates of [velar] retention, but also very high rates of [alveolar] realization, and very low rates of [reduced] realization.

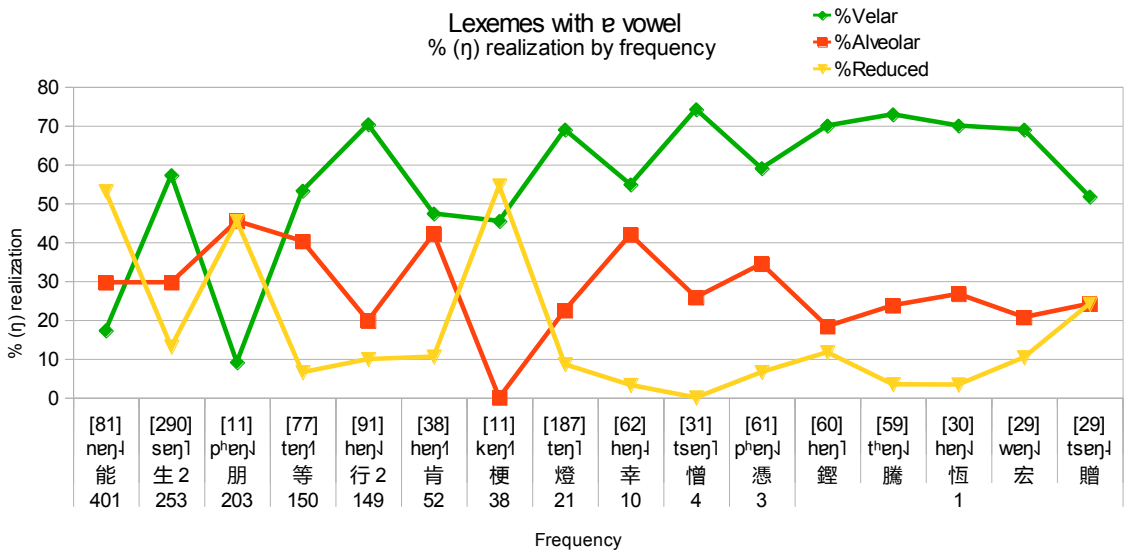


Figure 38. Realization of lexemes with -ɐŋ rime by frequency

For lexemes with the **-a**:**ŋ** rime, shown in Figure 39, there were three that were found to spear-head the change to lenition of [velar]. On the one hand, there is **sa**:**ŋ**˥ 生, which has the highest frequency in this group, of 253 occurrences in the HKCAC, and on the other, there are the lexemes **ka**:**ŋ**˥ 耕 and **k^{w}a**:**ŋ**˦ 逛, which have the lowest possible frequency of 1. All three lexemes seem to behave similarly in that they showed very low rates of [velar] retention, and high rates of [reduced] and [alveolar] variants. On the basis of these lexemes with **-a**:**ŋ** rimes, it is not clear how lexical frequency plays a role in lenition of [velar], as both the most frequent as well as the least frequent lexemes spearhead this change.

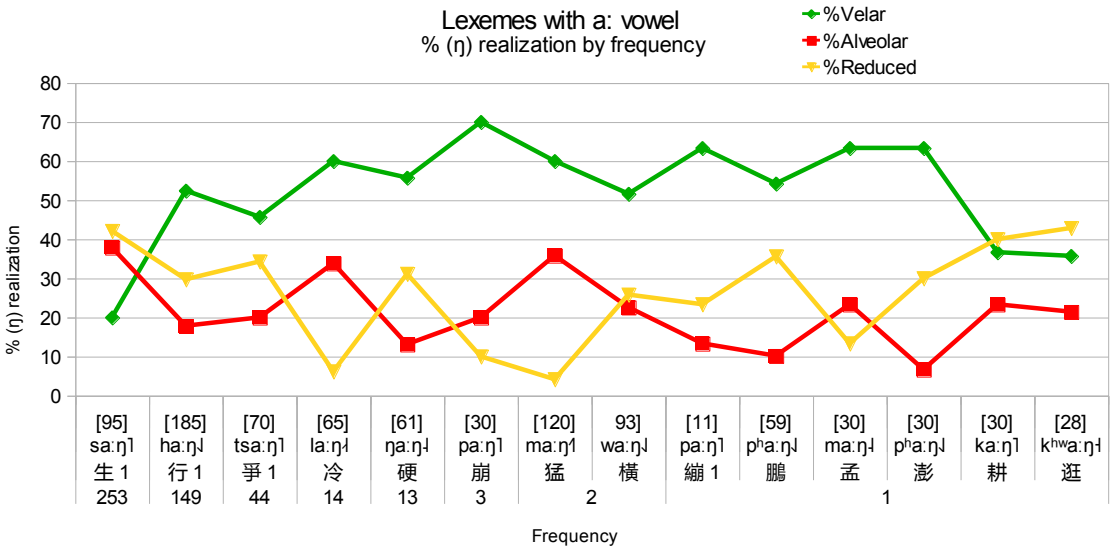


Figure 39. Realization of lexemes with -a:ŋ rime by frequency

Next, we will consider lexemes with the rimes **-ɛ**:**ŋ**, **-œ**:**ŋ**, and **-ɔ**:**ŋ** together. The variation patterns in (**ŋ**) realization for these lexemes are show in Figures 40, 41, and 42. For lexemes with the **-ɛ**:**ŋ** rime, there are lexemes at both ends of the frequency scale which have the lowest amount of [velar] retention. These are the lexemes **tɛ**:**ŋ**˥ 聽 and

tʰɛːŋ˩ 艇, which have frequencies of 248 and 1, respectively. They both show high rates of [reduced] realization, but fairly low rates of [alveolar] realization. As for lexemes with the rimes **-œːŋ** and **-ɔːŋ**, there is a large amount of variation, and it does not seem to be the case that lexemes at either end of the frequency range are leading the change.

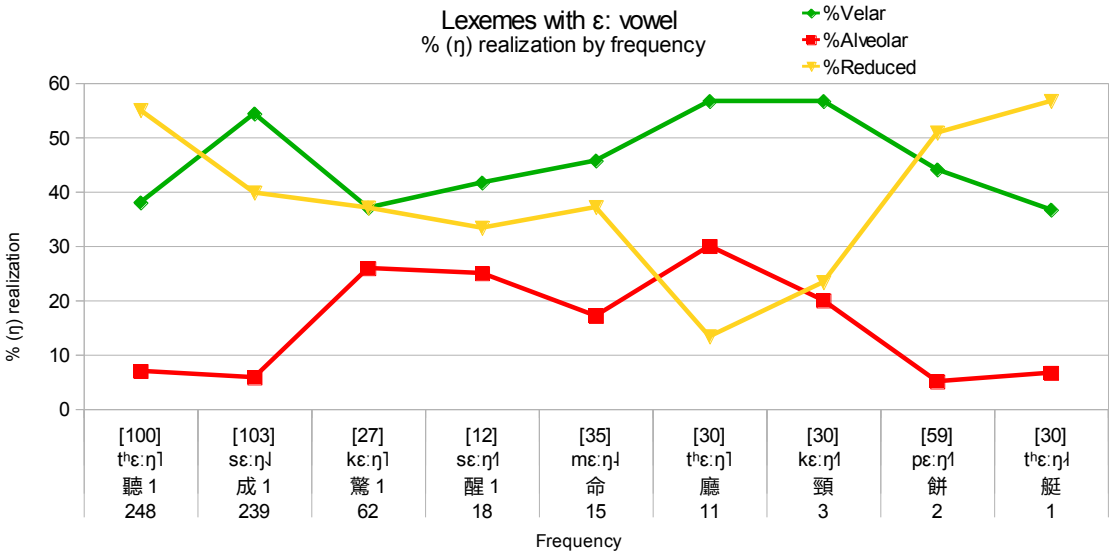


Figure 40. Realization of lexemes with -ɛːŋ rime by frequency

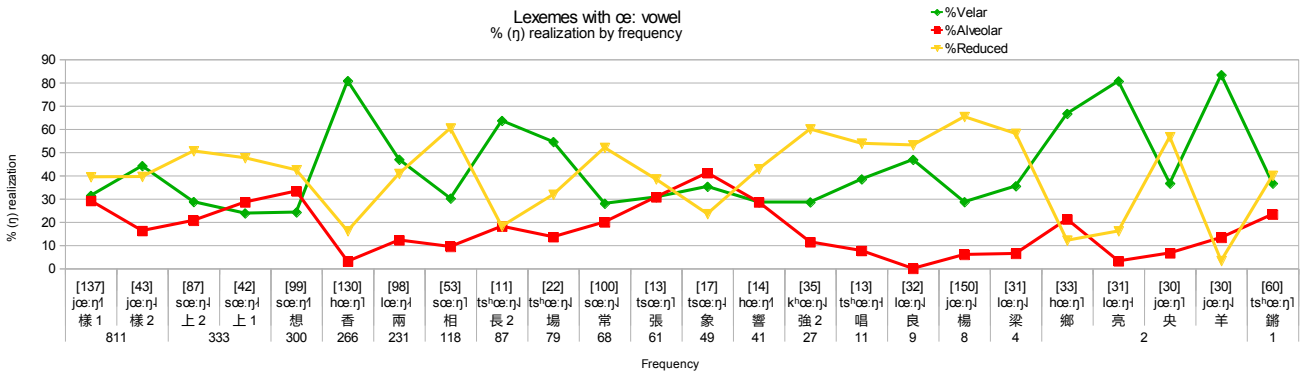


Figure 41. Realization of lexemes with -œːŋ rime by frequency

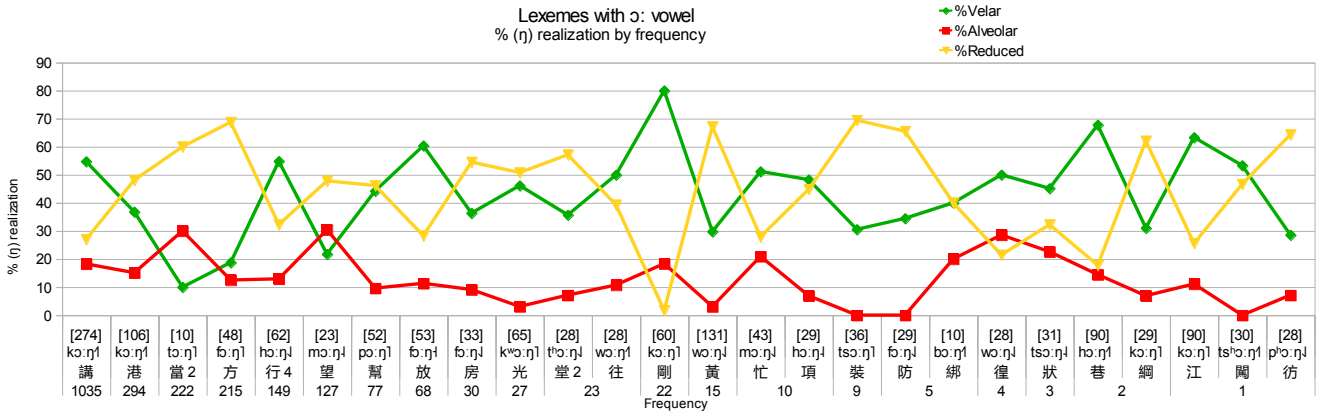


Figure 42. Realization of lexemes with -ɔːŋ rime by frequency

For lexemes with **-ɪŋ** and **-ʊŋ** rimes, as discussed before, the rate of [velar] retention is much higher than for lexemes containing other rimes. Despite this, there are 'spikes' of [reduced] realizations, especially amongst lexemes with higher frequencies, although there is a lexeme with the **-ʊŋ** rime, namely **hʊŋ˥** 恐, which has a frequency of 12 and showed a very high spike in [reduced] realization.

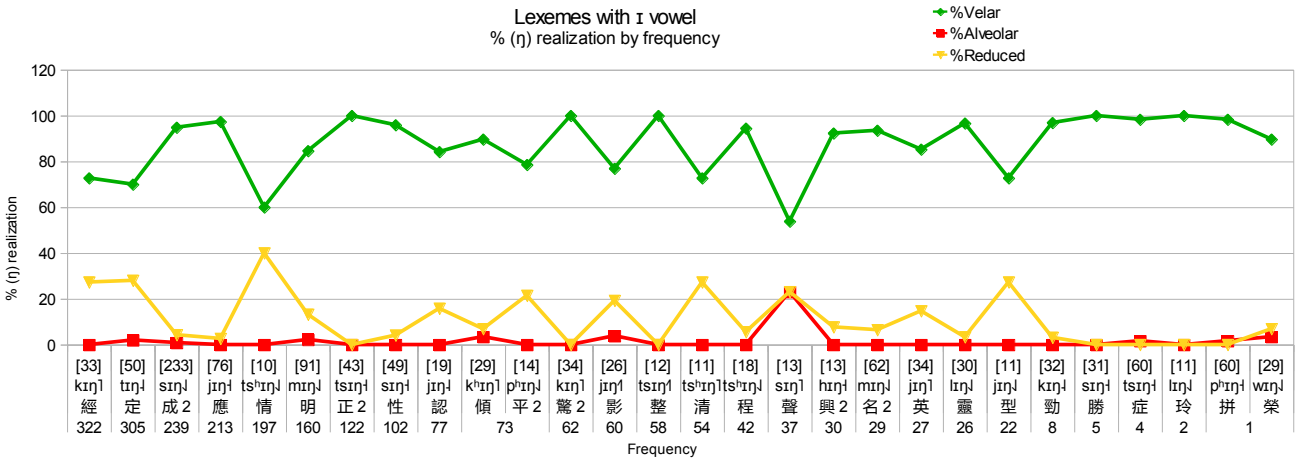


Figure 43. Realization of lexemes with -ɪŋ rime by frequency

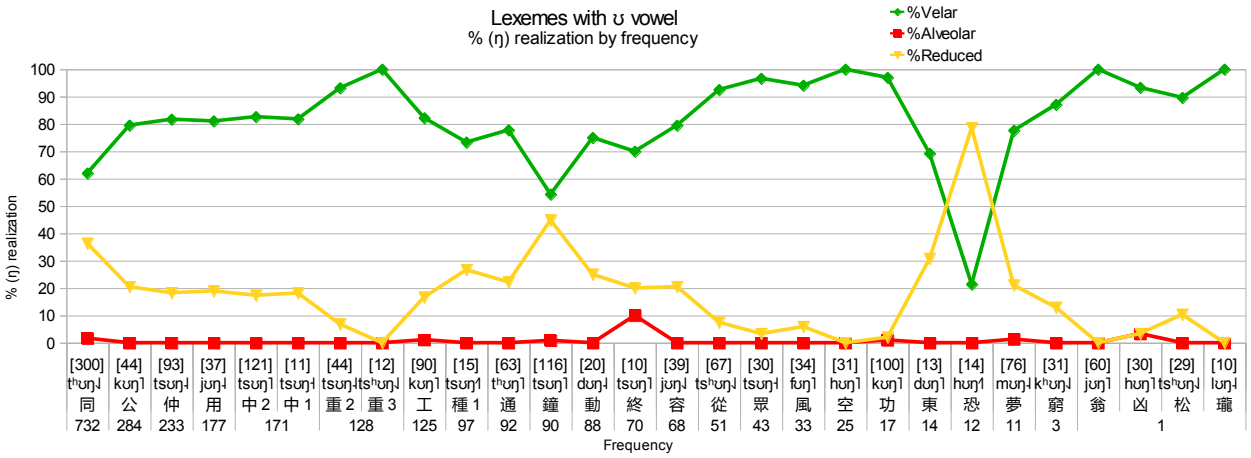


Figure 44. Realization of lexemes with -ʊŋ rime by frequency

One word of caution we must bear in mind, when interpreting results from low frequency lexemes presented here as leading the change to [alveolar] realization, is that many of the very low frequency words were elicited in formal reading tasks. As we saw in Section 5.4, informants seemed to treat such tasks, especially the minimal pairs reading task, as a forced-choice decision task, such that they resorted to the use of either [velar] or [alveolar]. Due to the set-up of this elicitation, and the inescapable fact that low frequency lexemes could only be obtained this way, we must exercise caution when interpreting the realization of infrequent items, as this might not necessarily reflect how infrequent lexemes are realized in casual, natural speech.

Overall, it could be seen that there were indeed some lexemes which were more progressive in [velar] lenition, and which showed concomitant higher rates of [alveolar] and [reduced] realizations. For lexemes with the rimes **-ɐŋ** and **-a**:**ŋ**, the data seemed to point to lexemes on the highest and lowest frequency ends leading the change to [velar] lenition. As for lexemes with the rimes **-ɛ**:**ŋ**, **-œ**:**ŋ**, and **-ɔ**:**ŋ**, the pattern of [velar] lenition and frequency seemed less clear. It should be noted that there was a large amount of variation, and that the patterns described here are only tentative. However, it

certainly seems the case that there were *some* lexemes in each rime group which were leading the change to [velar] lenition, and had a higher rate of [reduced] or [alveolar] realization. One logical follow-up question would be, what might be driving these lexemes to lead the change?

Recall from our discussion of neogrammarian factors in the preceding chapter, that following onset was found to be a good predictor of (**ŋ**) and (**k**) realization. For instance, following onsets with a 'velar' feature successfully predicted a high rate of [velar] retention for (**ŋ**) and (**k**), while an 'alveolar' feature was successful at predicting a higher rate of [alveolar]. Furthermore, 'labial' and 'palatal' features seemed to predict higher rates of [reduced] realizations. Given the influence that the following onset has on (**ŋ**) and (**k**) realization, and given that lexemes in Cantonese do not occur in isolation but appear with other lexemes, one conjecture which might be put forward is that for a given lexeme, frequent collocations with other lexemes containing onsets that promote lenition or alveolarization can make it more likely for this given lexeme to undergo sound change. Due to the limit in scope of this thesis, this hypothesis was not tested in here, but seems a promising direction for future research.

7.3 Homophonous lexemes

Recall the study of the development of homonym pairs from Middle Chinese (Chen & Hsieh 1971) from Section 1.5.1.2, where it was found that originally homophonous pairs developed to have different **-ai**/**-a** rimes in modern (Peking) Mandarin. This was

considered a prime example of lexical diffusion, as there could have been no phonetic conditioning involved to influence which word with the original rime were to split.

In this study, some lexemes were homophonous with one another, i.e. they share the same onset, vowel, rime *and* tone. One question that naturally arises then would be: do such homophonous lexemes show different rates of innovative (**ŋ**) and (**k**) realizations? Here, tokens with (**ŋ**) are compared. Figure 45 presents homophonous pairs (or triplets) side-by-side along with their frequency information. The token numbers used to obtain the percentages is again presented in square brackets above the lexeme. In each homophone pair (or triplet), the lexemes are presented with the highest frequency one first, followed by the lower frequency one(s).

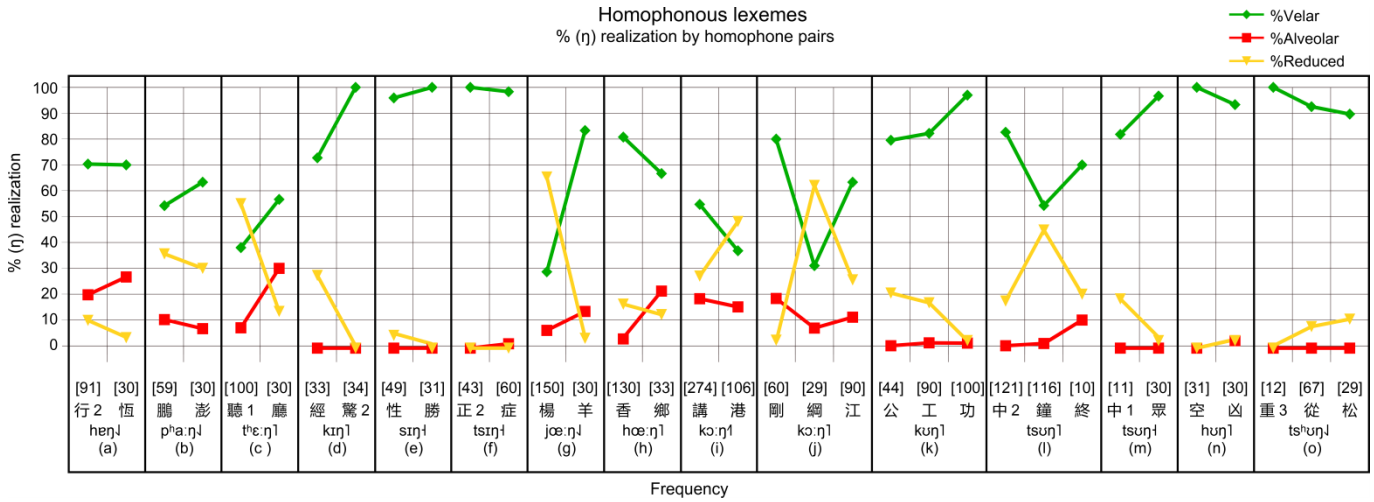


Figure 45. Percentage of (ŋ) realization by homophonous lexemes

As can be seen from the distribution of (**ŋ**) realization, there are two types of homophone pairs/triplets. Those where there is no (or hardly any) difference in (**ŋ**) realization between the lexemes, and those where there is a marked difference between the lexemes. Those in the former group are pairs (a), (b), (e), (f), (m), (n), and (o),

where there is negligible difference between lexemes and there does not seem to be any difference in how a lexeme with (**ŋ**) is realized when the frequency was varied. On the other hand, pairs/triplets where there is a marked difference between lexemes include (c), (d), (g), (h), (i), (j), (k), and (l). Let us focus our attention on this latter group.

First, let us consider [velar] retention. In the pairs:

(c) **tʰɛ**:**ŋ˥**, 聽 1 and 廳,
(d) **kɪŋ˥**,經 and 驚 2, and
(g) **jœ**:**ŋ˩**,楊 and 羊

the higher frequency lexemes seem to display more [velar] lenition than the lower frequency lexemes, although this trend is reversed in pairs:

(h) **hœ**:**ŋ˥**, 香 and 鄉, and
(i) **kɔ**:**ŋ˧˥** 講 and 港.

In triplets:

(j) **kɔ**:**ŋ˥**, 剛, 綱, and 江, and
(l), **tsʊŋ˥**,中 2, 鐘, and 終,

there is a varying range of [velar] retention across frequency.

In terms of [alveolar] realization, in pairs (c), (g), (h), and triplets (j) and (l), there was a higher proportion of [alveolar] realization for infrequent items than for frequent items. Pairs (d) and (i) showed little to no difference in [alveolar] realization.

As for [reduced] realization, pairs (c), (d), (g) and (h) showed higher rates of [reduced] realization in the frequent lexemes than for the infrequent lexemes. In pair (i), this trend was reversed. In the triplets (j) and (l), there was a varying range of [reduced] realization across frequencies.

This qualitative analysis shows that homophonous lexemes do behave differently from one another with respect to how (**ŋ**) is realized as [velar], [alveolar] or [reduced] variants, but it seems hard to attribute the difference to lexical frequency *per se*. It is hard to predict from this analysis, which lexemes might undergo velar lenition or alveolarization decades or centuries from now.

Even though the above picture, of the interaction effects of preceding vowel categories and lexical frequency effects, as well as of how individual lexemes respond to the lexical frequency variable, is far from clear, what the above exploration at least demonstrates is that there is ample evidence for lexical variation, whether this involves major categories of lexemes (e.g. those which share a rime) or whether the variation is between individual lexemes. Such demonstrable lexical variation within homophones is in and of itself problematic for the neogrammarian hypothesis of the regularity of sound change.

Next, instead of collating all lexemes obtained from all speakers, we will consider how much lexical variation exists within a given phonological environment, within individual speakers.

7.4 Inter- and intra-speaker lexical variation

This section investigates how much lexical variation exists within and across the individual speaker, by comparing the extent of (ŋ) realization across lexemes. Recall from Section 1.10, the discussion on what constitutes evidence for lexical diffusion. It was suggested that the strongest possible case for lexical diffusion can be made if lexical diffusion can be attested (a) within a speaker's repertoire, and (b) within a given speech style, and (c) within the same phonetic environment. That is, a given sound change has affected some lexical items but not others that meet all of the above conditions. Since it was not possible to explore the full interactions between the above three factors in a statistical model, this section offers a qualitative (and to a limited extent quantitative) exploration of the amount of lexical variation found across and within speakers.

All the tokens analysed here came from the word list reading task, such that the effect that speech style might have on (ŋ) realization can be controlled for. Figures 46 and 47 show percentage [velar] realization of the most frequent and the least frequent lexemes containing (ŋ) across speakers in the form of heat maps. The darker the green in a given square, the higher the percentage of [velar] retention. It can be seen that, overall, the rate of [velar] retention of least frequent lexemes (Figure 47) is slightly higher than for frequent lexemes (Figure 46), although the differences are not large. What can also be seen is that there is a large amount of interspeaker variation. In the most extreme case of MM26, hardly any velars were retained in the word list reading task, while MM25 and the older speakers consistently show very high rates of [velar] retention. There is also a large amount of variation amongst the younger speakers.

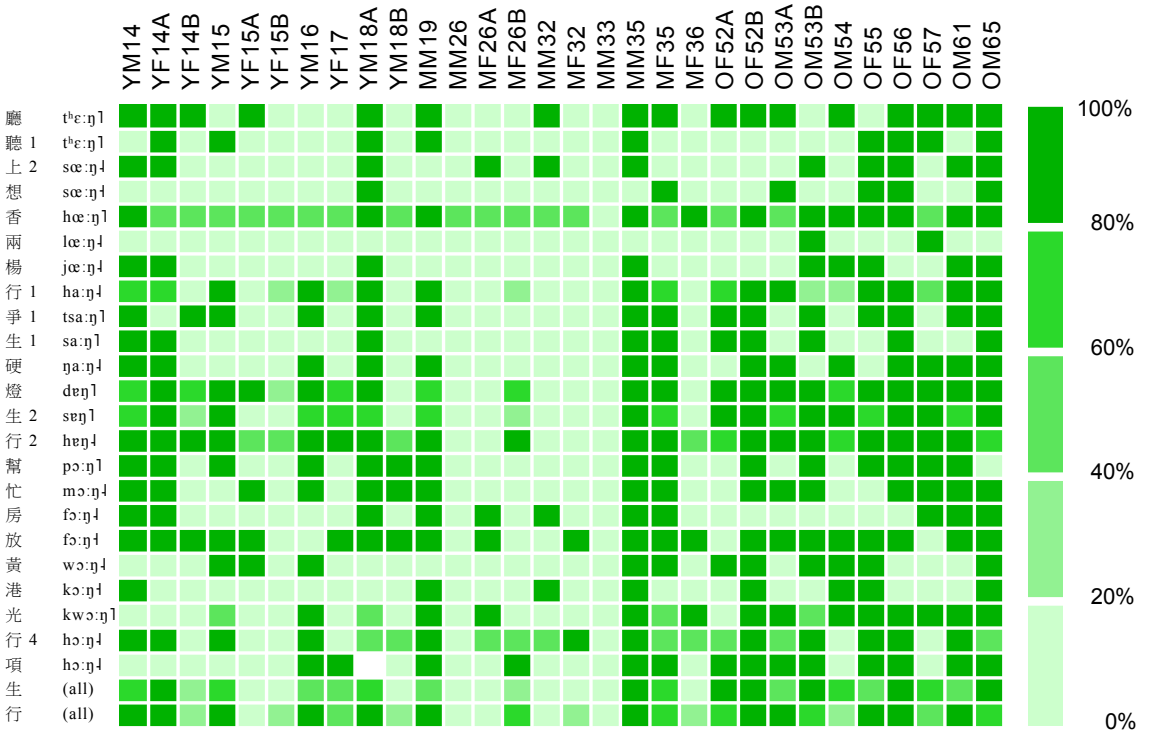


Figure 46. Percentage [ŋ] realization, most frequent characters in word list reading task

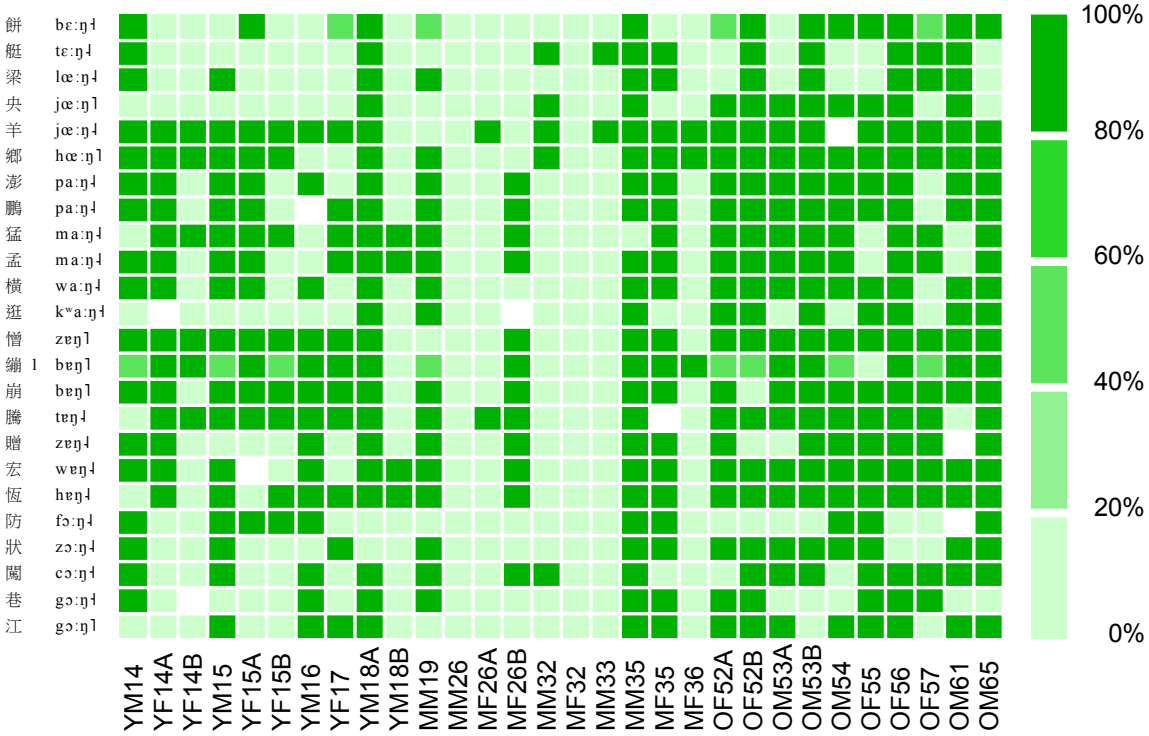


Figure 47. Percentage [ŋ] realization of least frequent characters in word list reading task

Blank = no data available

When comparing between lexemes, it can be seen that for both the most frequent and the least frequent lexemes, there is a great deal of inter-lexeme variation. For instance, the lexeme **lœ**:**ŋ**˧ 兩, which is amongst the most frequent lexemes presented in Figure 46 (6th row from the top), showed high rates of [velar] attrition, while **hœ**:**ŋ**˥ 香 (5th row from the top), which shares the same rime, showed much lower rates of [velar] attrition.

7.5 Discussion

Compared to the neogrammarian factors, namely preceding vowels and following onset, which were found to be good predictors, lexical frequency seemed in comparison to be a weaker predictor of variation in (**ŋ**) and (**k**) realization. However, from our exploration of lexical frequency effects in Section 7.1, our exposition of the fact that there are certain lexemes within preceding vowel categories and within the same homophonous environment for which there is demonstrable lexical frequency effects in Sections 7.2 and 7.3, as well as our demonstration that there are inter- and intra-speaker lexical variation patterns in Section 7.4, there is ample evidence that lexical diffusion could well operate on (**ŋ**) and (**k**) variation.

Moreover, as described in Section 6.2, the following phonetic segment seemed a very good predictor of the variation in (**ŋ**) and (**k**) realization. For instance, when followed by a velar, retention of [velar] realization was found to be much higher in (**ŋ**) and (**k**). Similarly, a following apical onset was associated with higher rates of [alveolar] realization. This leads us on to the logical possibility that any differences in (**ŋ**) and (**k**) realization found between lexemes which share a phonological environment

(preceding vowel), but for which lexical frequency could *not* be found, were in fact due to *which onset segment followed the lexemes*. In other words, certain lexemes could be propelled to higher rates of [velar] lenition, or higher rates of [alveolar] and [reduced] realizations, due to *frequent collocations* which these occurred in. This is a working hypothesis, and studying the effects of frequent collocations on (**ŋ**) and (**k**) realization is beyond the scope of this thesis, but could be a promising direction for future research. Since frequency of collocation is inherently lexically-based, if such effects were indeed found this would be a strong indication that lexically-based variation exists and that frequency of collocation is potentially driving certain lexical items to spearhead the sound change. If frequency of collocation is shown to be a plausible explanation, this would mean that the strongest version of the neogrammarian hypothesis is not in itself sufficient in explaining sound change and the neogrammarian hypothesis of regularity of sound change would need to be relaxed, such that any investigation into sound change should allow for the possibility of units beyond those in the immediate phonological domain (e.g. lexemes, collocations of lexemes) to have a significant effect on variation patterns in a sound change.

To summarize, this chapter has presented findings which demonstrate that (**ŋ**) and (**k**) variation show lexical effects, providing support that lexical diffusion operates on the progression of sound changes involving increased use of [alveolar] or [reduced] variants.

7.6 Summary

An extensive survey of lexical factors of variation in (**ŋ**) and (**k**) realization was presented in this chapter, namely lexical frequency as a continuous variable, the lexemes which are leading the change, variation in homophonous lexemes, and interspeaker variation.

Statistical results of the full model showed that 'log(frequency)' was significantly correlated with (**ŋ**) to [reduced] and (**k**) to [alveolar] / [reduced] variations, thereby confirming the usage-based hypothesis that more frequent items lead sound changes involving lenition.

A qualitative investigation into the lexemes which lead the change to [reduced] and [alveolar] variants in Section 7.2 found that some lexical items seemed more advanced in [velar] attrition and increased [alveolar] and [reduced] variants than other lexical items. The same wide range of inter-lexeme variation was found in a comparison of homophonous lexemes and in Section 7.3 and of interspeaker variation in Section 7.4.

This chapter closes with the suggestion that, given the findings presented here of inter-lexeme differences which seem unexplained by lexical frequency alone, and given that the following onset segments seem better predictors of (**ŋ**) and (**k**) realization as [velar], [alveolar] and [reduced], there is a logical possibility that the inter-lexeme differences can be explained by collocations which the lexeme frequently occur in. This is perhaps a promising direction for future research, and it would suggest that if any such frequency of collocation effects could be found, this would constitute further evidence that lexical diffusion can operate on sound changes.

Taken together, this chapter provided evidence that lexical diffusion operates on the sound change involving (**ŋ**) and (**k**). The next chapter will be devoted to collating and interpreting the results in this dissertation, and returning to the predictions made by the neogrammarian and lexical diffusion models of sound change. In doing so, an attempt is made to answer the central question of which factors are found to be the most significant and robust in explaining how the sound change in velar codas (**ŋ**) and (**k**) in Hong Kong Cantonese is driven.

8 Conclusion

8.1 Brief review

At the outset of this study, a discussion was offered presenting two views of sound change which have been held in opposition to one another: the neogrammarian regularity hypothesis and lexical diffusion model. Throughout much of the history of modern linguistics, the neogrammarian assumption of regularity of sound change has facilitated the reconstruction of earlier sound systems of languages, but the prevalence of this assumption meant that variation within the social and lexical domain were deemed to be of little to no relevance to an understanding of sound change. Throughout the course of structuralist and generativist developments, such a view of regularity was largely maintained. Studies of lexical diffusion were prompted by discussion in recent decades that called for a change in such a view of regularity, and which centred on the idea that it is lexemes which change. More recently, usage-based models have further supported the notion that lexical items are the primary medium through which sound change is transmitted, and that speakers form examplars of phonetic features from exposure to lexical items. An extension of this view was that frequent items were more likely to undergo leniting sound change, as lexemes were more likely to be reduced in frequent repetition. It was proposed that a sociophonetic study of variation along the phonetic and phonological as well as lexical dimensions would allow the predictions made by these models to be tested.

Chapter 2 provided an overview of the linguistic variables investigated in this thesis: the velar codas (**ŋ**) and (**k**) in Hong Kong Cantonese. These codas were found in

previous studies to be undergoing alveolarization. A brief overview of Cantonese phonology was offered to provide an understanding of the phonological environments which (**ŋ**) and (**k**) can occur in. It was also discussed that in most modern Sinitic languages, the Middle Chinese codas **-p**, **-t**, and **-k** have been lost, with Cantonese being one of the few languages still remaining which has preserved these codas. This is therefore a timely investigation into how the coda consonants in Cantonese are changing, and what factors might be driving this change.

Chapter 3 presented the methodology used in this thesis. Sociolinguistic interviews were conducted to sample from a group of speakers stratified by age and gender, and to elicit the tokens of interest within different speech styles. In order to test an extension of the lexical diffusion model, namely that lexical frequency effects might be observed in sound change, lexical frequency information was extracted from the Hong Kong Cantonese Adult Corpus (Leung, Law & Fung 2001). It was also suggested that, instead of treating [velar] and [alveolar] realizations as the only possible outcomes for (**ŋ**) and (**k**), as previous studies have done, it was useful to also investigate the extent to which the codas were undergoing lenition, introduced as a third [reduced] category. This division of possible realizations into a tripartite distinction required the use of multinomial logistic regression models to estimate the effects which different factors might have on (**ŋ**) and (**k**) realization, and a brief introduction to multilevel (or mixed-effects) multinomial logistic regression models was offered.

Chapter 4 presented full statistical models of (**ŋ**) and (**k**) realizations which combined all the factors discussed in this thesis. This ensured, as best as possible, that the effects of each factor was estimated with all other factors controlled for. It was

confirmed through the use of lexical frequency as a continuous variable that frequent items were more likely to be associated with a [reduced] realization, supporting the prediction offered by usage-based models that frequent lexemes are more likely to undergo leniting sound changes. On the other hand, alveolarization did not seem to show any frequency effects. However, any effects that lexicon-wide frequency had on (**ŋ**) and (**k**) realization were smaller, compared to the effects of other factors such as following onset features.

Chapter 5 presented variation of (**ŋ**) and (**k**) realization along the social dimensions of age, gender and speech style. It was found that younger speakers were much more likely to display alveolar and reduced realizations than older speakers, suggesting that sound change in the velar codas was progressing from older to younger generations. The speakers in the "middle" age group in their 20s and 30s were found to show far more alveolarization than even the younger age group. On the other hand, the younger age group displayed a higher rate of [reduced] realization. Invoking the apparent-time interpretation, these patterns seem to suggest that the velar codas are undergoing alveolarization as well as lenition. Speech style was also found to be significant in delineating the variation patterns in (**ŋ**) and (**k**) realizations. An increase in velar retention when going from casual speech style to passage reading and word list reading suggested that more attention to speech prompted speakers to retain higher rates of velar realizations. However, minimal pairs reading presented an anomaly, as it showed speakers consistently produced higher rates of alveolar realizations than expected. It seemed as if speakers were treating minimal pairs reading as a forced choice decision task where one coda or the other had to be produced, making reduced

variants a less acceptable option. On the basis of this anomalous result, a cautious reinterpretation of previous studies (Zee 1996; Ding 2010) was called for, where high rates of alveolarization were produced in single word lists, prompting the authors to suggest that alveolarization was much more prevalent than it might in fact be. This cross-over effect, whereby minimal pairs prompted a higher rate of alveolar realization, suggests that perhaps sound changes operating on languages with an orthography which does not encode phonetic cues progress differently along the social dimensions than do sound changes that operate on languages of which the orthography *does* encode phonetic cues.

Chapter 6 analysed phonological factors, which neogrammarian models deem to be the primary predictors of sound change. In this study, the immediate phonological environment of the velar codas were preceding vowel and tone. It was suggested that vowels and tones could be interacting factors, such that variation in (**ŋ**) and (**k**) realization due to tone might in fact be due to preceding vowels. Tautosyllabic onset and following phonetic onsets were also investigated. Overall, it was found that preceding vowel and following onset were the best out of neogrammarian factors discussed at describing the variation in (**ŋ**) and (**k**) realizations.

Chapter 7 examined lexical factors. The statistical effects of lexical frequency as a continuous variable on (**ŋ**) and (**k**) variation showed that there was a statistically significant effect of 'log(frequency)'. More specifically, 'log(frequency)' was significantly correlated with (**ŋ**) to reduced, (**k**) to alveolar, and (**k**) to reduced variations, thereby confirmed the usage-based hypothesis that more frequent items lead sound changes involving lenition.

Qualitative analyses were also performed to investigate which lexemes were leading the change. While it was certainly found that certain lexemes led the change, a comparison between homophonous lexemes, where phonological environment was strictly controlled for, showed that lexical frequency could not neatly explain why one lexeme as opposed to another was spearheading the shift to [alveolar] or [reduced] realizations. Combining the findings from Chapter 6, where following onset was found to be highly predictive, it was suggested that perhaps what could more prominently drive certain lexemes to adopt the innovative variants was their frequent collocations with lexemes that contained sound shift-promoting onsets. This prediction is borne out by the significant results of lexical frequency on (**ŋ**) to [reduced], and (**k**) to [alveolar] or [reduced] realization, as obtained from the full models in Chapter 4, with all other factors— including following phonological feature —being kept equal.

8.2 Future directions

It has become apparent in our investigation of predictions made by neogrammarian and lexical diffusion models of sound change, that stochastically, immediate phonetic and phonological environments were good predictors of which realization velar codas (**ŋ**) and (**k**) in Cantonese might take, compared to a lexicon-wide frequency measure. However, it was found that lexical frequency (measured as mean-centered 'log(frequency)') was significant in explaining (**ŋ**) to reduced, (**k**) to alveolar, and (**k**) to [reduced] variations, with all else being equal. It was also found in our qualitative investigation, that certain lexemes *do* spearhead the shift to [alveolar] or [reduced] variants. These results suggest that the sound changes involving (**ŋ**) and (**k**) *do* show lexical diffusion, and in addition, lexical frequency effects.

One promising future direction arising from this study would be to investigate the effects of frequency of collocations of lexemes in predicting change in velar coda consonants. As described above, it is hypothesized that this measure might offer a better prediction as to which lexemes lead the change. A study of collocations, rather than single morphosyllables, as has been the case in this study, is therefore put forward as a possible direction for future study.

8.3 Concluding remarks

In the introduction of this thesis, a discussion was presented of the "paradox" or "controversy" based on which this study was designed. Using Labov's words (1994: 541),

> [r]estated most simply, the paradox amounts to this: if Wang and his associates are right about lexical diffusion, and the Neogrammarians were more right than they knew about sound change, how can both be right?

By simultaneously investigating variation patterns along social, phonetic and phonological as well as lexical dimensions, this investigation of velar coda (**ŋ**) and (**k**) alveolarization in Hong Kong Cantonese has shown the complexity of variation patterns involved in sound change. In this evaluation of neogrammarian and lexical diffusion models of sound change, evidence was found against the strongest versions of these two models: lexical variation was found within phonological categories, and phonetic variation was found within a given lexeme. Consequently, a response to the Labov's question would be that, at least in the case of sound change in velar codas of Cantonese, both models were right, when taken *in conjunction*.

The use of three outcome categories, namely 'velar', 'alveolar' and 'reduced', for analysing the variation in (**ŋ**) and (**k**) realization, has allowed us to assess the extent to which these velar codas are undergoing alveolarization and lenition. Using multinomial logistic regression, it was possible to explore an extensive range of factors which seemed to drive the sound change in (**ŋ**) and (**k**), on the one hand, to alveolarization, and on the other, to undergo lenition. In doing so, further support was found for frequency effects in leniting sound change, as proposed in usage-based models, confirming the prediction that frequent lexemes are more susceptible to reductive changes. On the other hand, alveolarization was not found to be sensitive to lexical frequency effects for (**ŋ**) but only for (**k**). The prevalence of lenition in casual speech style in both (**ŋ**) and (**k**) realization when compared to the more formal speech styles also suggests that lenition could create

a precondition for changes in (**ŋ**) and (**k**) realization, and that this change is likely to start with more frequent lexical items. The abrupt substitution of one phonetic realization by another, also dubbed "flip-flops" by Wang (1969), was observed in the minimal pair reading task. The evidence from the minimal pairs reading tasks (with an unexpectedly high rate of alveolarization) suggests that prevalent lenition might create a condition whereby speakers gradually lose the distinctions between velar and alveolar realizations in (**ŋ**) and (**k**). While the distinction between velar and alveolar realizations seems increasingly lost by speakers, what is clear however, is that lenition is above the level of conscious social awareness, and that speakers are mindful to produce *some* coda and that it seemed *least* socially desirable to produce lenited variants of (**ŋ**) and (**k**) realization.

One prediction that arises from these findings, is that the mechanism of consonantal sound changes involving a shift in place of articulation might be different from that of reductive sound changes. While reductive sound changes seem sensitive to lexical frequency effects, sound changes involving a shift in place of articulation, such as from velar to alveolar, might be mediated via more socially-conditioned mechanisms.

A further suggestion put forward here is that the frequency with which a particular lexeme enters into collocations with other lexemes (which contain onsets promoting a shift in place of articulation) might predict whether a given lexeme spearheads a sound change involving a shift in place of articulation. Taken together, this study suggests that not all consonantal sound changes are the same, and underscores the importance of incorporating factors from social, phonetic and phonological, and lexical dimensions in investigating the mechanisms of sound change.

References

Agresti, Alan. 2007. *An Introduction to Categorical Data Analysis*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Agresti, Alan, James G. Booth, James P. Hobert & Brian Caffo. 2000. Random-effects modeling of categorical response data. *Sociological Methodology.* 30(1). 27–80.

Baayen, R.H. 2001. *Word Frequency Distributions*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Baayen, R.H. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Barrie, Mike. 2003. Contrast in Cantonese Vowels. *Toronto Working Papers in Linguistics.* 20(1–19).

Barry, W.J. 1996. Some fundamental problems of looking at connected speech. *Arbeitsberichte*, vol. 31, 113–118. (Arbeitsberichte). Institut für Phonetik und Digitale Sprachverarbeitung, Kiel University.

Bauer, Robert S. 1979. Alveolarization in Cantonese: a case of lexical diffusion. *Journal of Chinese Linguistics.* 7(1). 132–141.

Bauer, Robert S. 1983. Cantonese sociolinguistic patterns: correlating social characteristics of speakers with phonological variables in Hong Kong Cantonese. University of California, Berkeley, PhD thesis. Ann Arbor: University Microfilms International.

Bauer, Robert S & Paul K Benedict. 1997. *Modern Cantonese Phonology*. Berlin: Mouton de Gruyter.

Bauer, Robert S, Kwan-Hin Cheung & Pak-Man Cheung. 2003. Variation and merger of the rising tones in Hong Kong Cantonese. *Language Variation and Change.* 15. 211–225.

Bauer, Robert S & Stephen Matthews. 2003. Cantonese. In Graham Thurgood & Randy J LaPolla (eds.), *The Sino-Tibetan Languages*, 146–155. London: Routledge.

Baxter, Laura. 2010. Lexical Diffusion in the Early Stages of the Merry-Marry Merger. *University of Pennsylvania Working Papers in Linguistics.* 16(2).

Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin / New York: Mouton de Gruyter.

Boersma, Paul & David Weenink. 2013. Praat: Doing Phonetics by Computer (Version 5.3.51). Retrieved June 10, 2013, from http://www.praat.org.

Bourgerie, Dana S. 1990. A Quantitative Study of Sociolinguistic Variation in Cantonese. The Ohio State University, doctoral dissertation.

Browman, C.P. & L. Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49(3–4). 155–180. http://www.haskins.yale.edu/SR/SR111/SR111_03.pdf (16 October, 2011).

Buckingham, B. R. 1936. *A combined word list*. Boston: Ginn and Co.

Bullock, Barbara E., Almeida Jacqueline Toribio & Mark Amengual. 2014. The status of s in Dominican Spanish. *Lingua.* 143. 20–35.

Bybee, Joan. 1999. Usage-based phonology. In Michael Darnell et Al. (ed.), *Functionalism and Formalism in Linguistics. Vol. I: General Papers*, 211–42. Amsterdam: John Benjamins.

Bybee, Joan. 2000a. The phonology of the lexicon: evidence from lexical diffusion. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-Based Models of Language*, 65–86. Stanford: Center for the Study of Language (CSLI).

Bybee, Joan. 2000b. Lexicalization of sound change and alternating environments. In M. Broe & J. Pierrehumbert (eds.), *Laboratory V: Language acquisition and the lexicon*, 250–268. Cambridge: Cambridge University Press.

Bybee, Joan. 2002a. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change.* 14. 261–290.

Bybee, Joan. 2002b. Phonological evidence for exemplar storage of multiword sequences. *Studies in Second Language Acquisition.* 24. 215–221.

Bybee, Joan. 2010. Patterns of lexical diffusion and articulatory motivation for sound change. *Workshop on Sound Change*. Institut d'Estudis Catalans, Barcelona.

Campbell, Lyle. 2004. *Historical Linguistics*. Cambridge, Massachusetts: MIT Press.

Cedergren, Henrietta. 1973. The interplay of social and linguistic factors in Panama. Cornell University, doctoral dissertation.

Census. 2011. Usual residents aged 5 and over by duration of residence in Hong Kong, sex, nationality and usual language. http://www.census2011.gov.hk/en/main-table/F118.html.

Chambers, J. K. & Peter Trudgill. 1998. *Dialectology*. Cambridge: Cambridge University Press.

Chambers, J K. 1992. Dialect acquisition. *Language.* 68(3). 673–705.

Chan, Yuet Hung Cecilia. 2008. A sociolinguistic study of the merging of initial /n-/ with /l-/ in Cantonese and Putonghua by Cantonese speakers in Hong Kong. *Proceedings of the 3rd CLS International Conference on Media in Foreign Language Teaching and Learning*, pp 24-35. Singapore: National University of Singapore, Singapore, 4-7 December 2008.

Chao, Yuen-Ren. 1933. A system of tone letters. *Le Maître Phonétique.* 20. 24–27.

Chao, Yuen Ren. 1947. *Cantonese Primer*. Cambridge, Massachusetts: Harvard University Press.

Chen, Hoi Ying. 1998. Norms of Pronunciation and the Sociolinguistics of Cantonese in Hong Kong. The University of Hong Kong,.

Chen, Matthew. 1972a. Nasals and nasalization in Chinese: an essay in natural phonology. University of California, Berkeley, doctoral dissertation.

Chen, Matthew. 1972b. The Time Dimension: Contribution toward a Theory of Sound Change. *Foundations of Language.* 8(4). Springer. 457–498.

Chen, Matthew & Hsin-I Hsieh. 1971. The time variable in phonological change. *Journal of Linguistics.* 7. 1–13.

Chen, Matthew Y & William S-Y. Wang. 1975. Sound change: actuation and implementation. *Language.* 51(2). 255–281.

Cheng, Chin-Chuan & William S.Y. Wang. 1977. Tone change in Chao-Zhou Chinese: A study in lexical diffusion. In William S.Y. Wang (ed.), *The lexicon in phonological change*, 86–100. The Hague: Mouton.

Cheung, Hung-nin (張洪年). 1972. *香港粵語語法的研究 [Studies in Hong Kong Cantonese Grammar]*. Hong Kong: The Chinese University of Hong Kong Press.

Cheung, Kwan-Hin. 1986. The phonology of present-day Cantonese. University College London, Ph.D. thesis.

Cohen, Paul. 1970. The tensing and raising of short [a] in the metropolitan area of New York City. Columbia University, Master's essay.

Croft, William. 2000. *Explaining language change: An evolutionary approach*. Essex: Pearson Education.

Ding, Picus Sizhi. 2010. Phonological change in Hong Kong Cantonese through language contact with Chinese topolects and English over the past century. In Robert McColl Millar (ed.), *Marginal dialects: Scotland, Ireland and beyond*, 198–218.

Dinkin, Aaron J. 2008. The real effect of word frequency on phonetic variation. *University of Pennsylvania Working Papers in Linguistics.* 14(1). 97–106.

Docherty, Gerard, Jen Hay & Abby Walker. 2006. Sociophonetic patterning of phrase-final /t/ in New Zealand English. In Paul Warren & Catherine I. Watson (eds.), *Proceedings of the 11th Australian International Conference on Speech Science & Technology*, 378–383. University of Auckland, New Zealand.

Dong, Tong-he. 1954. *Zhongguo Yuyin Shi [A Phonological History of Chinese]*. Taipei.

Eckert, Penelope. 1989. *Jocks and Burnouts: social categories and identity in the high school*. New York: Teachers College Press.

Eckert, Penelope. 1997. Age as a sociolinguistic variable. *The Handbook of Sociolinguistics*, 151–67. Oxford: Blackwell.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics.* 12. 453–476.

Ferguson, Charles A. 1975. "Short a" in Philadelphia English. In E. Smith (ed.), *Studies in Linguistics in Honor of George L. Trager*, 259–74. The Hague: Mouton.

Fidelholtz, James L. 1975. Word frequency and vowel reduction in English. *Chicago Linguistic Society.* 11. 200–213.

Fought, John. 1982. The Reinvention of Hugo Schuchardt (Review Article). *Language in Society.* 11(3). 419–436.

Francis, W. Nelson & Henry Kučera. 1982. *Frequency analysis of English usage*. Boston: Houghton Mifflin.

Fung, Roxanna S.Y. & Cathy S.P. Wong. 2011. Acoustic Analysis of the New Rising Tone in Hong Kong Cantonese. *The 17th International Congress of Phonetic Sciences (ICPhS XVII)*.

Gilliéron, Jules. 1902. *Atlas linguistique de la France*. Paris: Champion.

Gordon, Elizabeth, Margaret Maclagan & Jennifer Hay. 2007. The ONZE corpus. *Creating and digitizing language corpora. Volume 2: Diachronic Databases*, 82–104.

Guy, Gregory, Jennifer Hay & Abby Walker. 2008. Phonological, lexical, and frequency factors in coronal stop deletion in early New Zealand English. *Laboratory Phonology.* 11. 53–54.

Hall-Lew, Lauren. 2014. Chinese social practice and San Franciscan Authenticity. In Veronique Lacoste, Jakob Leimgruber & Thiemo Breyer (eds.), *Indexing Authenticity: Sociolinguistic Perspectives*, 55–77. FRIAS Ling. Berlin: de Gruyter.

Hall-Lew, Lauren & Sonya Fix. 2012. Perceptual coding reliability of (L)-vocalization in casual speech data. *Lingua.* 122. 794–809.

Halle, Morris & K. P. Mohanan. 1985. Segmental Phonology of Modern English. *Linguistic Inquiry.* 16. 57–116.

Hansen, Anita Berit. 2001. Lexical diffusion as a factor of phonetic change: the case of Modern French nasal vowels. *Langage Variation and Change.* 13. 209–52.

Harris, John. 1989. Towards a lexical analysis of sound change in progress. *Journal of Linguistics.* 25. 35–56.

Hashimoto, Oi-kan Yue. 1972. *Phonology of Cantonese*. Cambridge: Cambridge University Press.

Hay, J., K. Drager & P. Warren. 2010. Short-term Exposure to One Dialect Affects Processing of Another. *Language and Speech*. 1–25.

Hay, Jennifer B., Jennet B. Pierrehumbert, Abby J. Walker & Patrick LaShell. 2015. Tracking word frequency effects through 130 years of sound change. *Cognition.* 139. 83–91.

Ho, Miu-Tai. 2004. A sociolinguistic investigation of Cantonese in Hong Kong and Guangzhou. Essex: University of Essex, Ph.D. Thesis.

Hockett, Charles F. 1965. Sound change. *Language.* 41(No. 2). 185–204.

Hooper, Joan Bybee. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In W Christie (ed.), *Current progress in historical linguistics*, 96–105. Amsterdam: North Holland.

International Phonetic Association. 1989. Report on the 1989 Kiel Convention. *Journal of the International Phonetic Association.* 19(2). 67–80.

Johnson, David Ezra. 2009. Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass.* 3(1). 359–383.

Johnson, Theodore Clifton. 1983. Phonological free variation, word frequency, and lexical diffusion. University of Washington. University Microfilms International, Ann Arbar.

Jurafsky, D., A. Bell, E. Fosler-Lussier, C Girand & W. Raymond. 1998. Reduction of English function words in Switchboard. *Proceedings of ICSLP-98*. Sydney.

Kao, Diana L. 1971. *Structure of the syllable in Cantonese*. The Hague: Mouton.

Karlgren, Berhard. 1926. *Études sur la phonologie chinoise [Archives d'études orientales, Vol. 15]*. *Archives d'Études Orientales*. Vol. 15. Leiden: E.J. Brill.

Kerswill, Paul & Ann Williams. 2000. Creating a New Town Koine: Children and language change in Milton Keynes. *Language in Society.* 29. 65–115.

Kiparsky, Paul. 1989. Phonological change. In F. J. Newmeyer (ed.), *Linguistics, the Cambridge survey*, 363–415. Cambridge: Cambridge University Press.

Kisler, T., F. Schiel & H. Sloetjes. 2012. Signal processing via web services: the use case WebMAUS. *Proceedings Digital Humanities 2012*, 30–34. Hamburg, Germany.

Krishnamurti, Bh. 1998. Regularity of sound change through lexical diffusion: A study of s > h > Ø in Gondi dialects. *Language Variation and Change.* 10. 193–220.

Labov, William. 1963. The social motivation of a sound change. *Word.* 19(3). 273–309.

Labov, William. 1981. Resolving the Neogrammarian controversy. *Language* 57(2). 267–308.

Labov, William. 1989. The exact description of the speech community: Short a in Philadelphia. In R. Fasold D. Schiffrin (ed.), *Language Change and Variation*, 1–57. Washington D.C.: Georgetown University Press.

Labov, William. 1991. The three dialects of English. In Penelope Eckert (ed.), *New Ways of Analyzing Sound Change*, 1–44. New York: Academic Press.

Labov, William. 1994. *Principles of Linguistic Change. Volume 1: Internal Factors*. Oxford: Basil Blackwell.

Labov, William. 2001. *Principles of linguistic change. Volume 2: Social Factors*. Oxford: Blackwell.

Labov, William. 2007. Transmission and Diffusion. *Language.* 83.

Labov, William, Malcah Yaeger & Richard Steiner. 1972. *A quantitative study of sound change in progress*. Philadelphia: U.S. Regional Survey.

Lavoie, Lisa M. 2002. Subphonemic and allophonic consonant variation: the role of the phoneme inventory. *ZAS Papers in Linguistics.* 28. 39–54.

Law, S.-P., Roxana S.-Y Fung & Robert S Bauer. 2001. Perception and production of Cantonese consonant endings. *Asia Pacific Journal of Speech, Language and Hearing* 6. 179–195.

Lee, Tan, Patgi Kam & Frank K Soong. 2006. Modeling Cantonese Pronunciation Variations for Large-Vocabulary Continuous Speech Recognition. *Computational Linguistics and Chinese Language Processing.* 11(1). 17–36.

Lehmann, Winfred P. 1967. *A reader in nineteenth century historical Indo-European linguistics*. (Ed.) Winfred P. Lehmann. Bloomington: Indiana University Press.

Leimgruber, Jakob R.E. 2009. Modelling Variation in Singapore English. University of Oxford (Pembroke College), DPhil thesis.

Leskien, August. 1876. *Die Declination im Slavisch-litauischen und Germanischen*. Leipzig: S. Hirzel.

Leung, M.-T., S.-P. Law & Roxana S.-Y Fung. 2001. The HKCAC The Hong Kong Cantonese adult language corpus. *International Journal of Corpus Linguistics.* 6. 305–325.

Li, Paul Jen-Kuei. 1982. Linguistic variations of different age groups in the Atayalic dialects. *The Tsing Hua Journal of Chinese Studies (new series).* 14. 167–91.

Li, Xinkui, Jiajiao Huang, Qisheng Shi, Yun Mai & Dingfang Chen. 1995. *廣州方言研究 [Study of the Cantonese Dialect]*. Shaoguan 韶關: Guangdong People's Publishing House 廣東人民出版社.

Lindblom, B. 1990. Explaining phonetic variation: A sketch of the H and H theory. In W. J. Hardcastle & A. Marchand (eds.), *Speech Production and Speech Modelling*, 403–40. Dordrecht: Kluwer Academic Publishers.

Markey, T. (ed. & trans.). 1979. *Hugo Schuchardt: The ethnography of variation. Selected writings on pidgins and creoles (Linguistica extranea, Studio, 3)*. Ann Arbor, Michigan: Karoma.

Matthews, Stephen & Virginia Yip. 1994. *Cantonese: a comprehensive grammar*. New York: Routledge.

Milroy, James. 1992. *Linguistic Variation and Change*. Oxford: Blackwell.

Milroy, Lesley & Matthew J. Gordon. 2003. *Sociolinguistics: method and interpretation*. Oxford: Blackwell.

Mowrey, Richard & William Pagliuca. 1995. The reductive character of articulatory evolution. *Rivista di Linguistica.* 7. 37–124.

Ogura, Mieko. 1995. The development of Middle English ī and ū: a reply to Labov (1992, 1994). *Diachronica.* 12. 31–53.

Oliviera, Marco Antonia De. 2009. The neogrammarian controversy revisited. *International Journal of the sociology of language.* 89(1). 93–106.

Oostdijk, N. 2000. The Spoken Dutch Corpus project. *The ELRA Newsletter.* 5. 4–8.

Orton, Harold & M. F. Wakelin (eds.). *Survey of English Dialects. The Basic Material. Vol. 4 (3 parts): The Southern Counties*. Leeds: E. J. Arnold.

Osthoff, H & K Brugmann. 1878. Einleitung, Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen, I [Preface to Morphological Investigations in the Sphere of the Indo-European Languages, Vol. I]; English translation in W.P. Lehmann (Ed.), (1967). 198–209. Bloomington and London: Indiana University Press.

Palmer, Leonard R. 1972. *Descriptive and Comparative Linguistics*. London: Faber & Faber.

Pan, Peter G. 1982. Hong Kong Cantonese: a sociolinguistic perspective. *Working papers in Linguistics and Language Teaching*. Language Centre, University of Hong Kong.

Payne, A. 1980. Factors controlling the acquisition of the Philadelphia dialect by out of state children. In W Labov (ed.), *Locating Language in Time and Space*, 143–158. New York: Academic Press.

Payne, Arvilla. 1976. The acquisition of the phonological system of a second dialect. University of Pennsylvania, PhD thesis.

Phillips, Betty. 1980. Old English an - on: A new appraisal. *Journal of English Linguistics.* 14. 20–23.

Phillips, Betty. 1984. Word frequency and the actuation of sound change. *Language.* 60. 320–342.

Phillips, Betty. 2006. *Word Frequency and Lexical Diffusion*. New York: Palgrave Macmillan.

Phillips, Betty S. 2001. Lexical diffusion, lexical frequency, and lexical analysis. In Bybee Joan & Paul Hopper (eds.), *Frequency and the emergence of linguistic structure*, 123–136. Amsterdam: John Benjamins.

Phillips, Betty S. 1994. Southern English glide deletion revisited. *American Speech.* 69. 115–27.

Pierrehumbert, Janet. 2001. Exemplar dynamics: word frequency, lenition, and contrast. In Joan L Bybee & Paul Hopper (eds.), *Frequency effects and the emergence of lexical structure*, 137–157. Amsterdam: John Benjamins Publishing Company.

Pluymaekers, Mark, Mirjam Ernestus & R. Harald Baayen. 2005. Lexical frequency and acoustic reduction in spoken Dutch. *Acoustical Society of America.* 118. 2561.

R Core Team. 2013. R: A Language and Environment for Statistical Computing, version 3.0.2. Vienna, Austria: R Foundation for Statistical Computing.

R Core Team and contributors. 2013. The R Stats Package. Version 3.0.2.

Rasbash, J., C. Charlton, W.J. Browne, M. Healy & B. Cameron. 2005. MLwiN Version 2.26. Centre for Multilevel Modelling, University of Bristol.

Rasbash, Jon, Fiona Steele, William J. Browne & Harvey Goldstein. 2012. *A User's Guide to MLwiN*. Bristol: Centre for Multilevel Modelling, University of Bristol.

Rossiter, D. G. & Analía Loza. 2012. *Technical Note: Analyzing land cover change with logistic regression in R*. ITC, Enschede.

Sankoff, Gillian. 2004. Adolescents, Young Adults, and the Critical Period: Two Case Studies from "Seven Up." In Carmen Fought (ed.), *Sociolinguistic Variation: Critical Reflections*, 121–239. Oxford / New York: Oxford University Press.

Sankoff, Gillian & Hélene Blondeau. 2007. Language change across the lifespan: /r/ in Montreal French. *Language.* 83(3). 560–88.

Santa Ana, Otto. 1991. Phonetic simplification processes in English of the barrio: A cross- generational sociolinguistic study of the Chicanos of Los Angeles. Doctoral Dissertation, University of Pennsylvania.

Saussure, Ferdinand de. 1959. *Course in General Linguistics*. (Ed.) Charles Bally, Albert Schehaye & in collaboration with Albert Riedlinger. Translated by Wade Baskin. New York: Philosophical Library.

Scherer, Wilhelm. 1868. *Zur Geschichte der deutschen Sprache*. Berlin.

Schiel, F. 1999. Automatic Phonetic Transcription of Non-Prompted Speech. *Proceedings of the ICPhS*, 607–610.

Schleicher, August. 1848. *Sprachvergleichende (linguistische) Untersuchungen II*. Bonn.

Schmid, Carolin, Sylvia Moosmueller & Christian Kasess. 2015. Sociophonetics of the velarized lateral in the Viennese dialect. In The Scottish Consortium for IcPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: the University of Glasgow. http://www.icphs2015.info/pdfs/Papers/ICPHS0701.pdf.

Schmidt, Johannes. 1872. *Die Werwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar: Böhlau.

Schryver, Johan De, Anneke Neijt, Pol Ghesquière & Mirjam Ernestus. 2008. Analogy, Frequency, and Sound Change. The Case of Dutch Devoicing. *Journal of Germanic Linguistics.* 20(2). 159–195.

Schuchardt, Hugo. 1885. Uber die Lautgesetze: Gegen die Junggrammatiker. [Translation in: Schuchardt, the Neogrammarians, and the transformational theory of phonological change, ed. by Theo Vennemann & Terence Wilbur, 39-72. Frankfurt: Athenaum, 1972.].

Shen, Zhongwei. 1990. Lexical diffusion: a population perspective and a numerical model. *Journal of Chinese Linguistics.* 18. 159–200.

Shum, Josephine. 1993. Sound Change between /-n/ and /-ŋ/ in Word Final Position for Cantonese Spoken in Hong Kong. The University of Hong Kong,.

Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication.* 23. 193–229.

So, Lydia. 1996. Tonal Changes in Hong Kong Cantonese. *Current Issues in Language and Society.* 3(2). 186–189.

Sommerfelt, Alf. 1962. *Diachronic and synchronic aspects of language*. The Hague: Mouton.

Song, Chenqing. 2008. Some Issues in Old Chinese Phonology: Facts and Methodology. University of Winsconsin-Madison.

Stuart. 2007. A sociophonetic investigation of postvocalic /r/ in Glaswegian Adolescents. *ICPhS XVI*, 1449–1452. Saarbruecken.

Tagliamonte, Sali A. 2006. *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.

Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change.* 24. 135–178.

Thorndike, Edward & L. Lorge. 1944. *The teacher's word book of 30,000 words*. New York: Columbia University Press.

Trager, George L. 1930. The pronunciation of "short A" in American Standard English. *American Speech.* 5. 396–400.

Trager, George L. 1934. What conditions limit variants of a phoneme? *American Speech.* 9. 313–15.

Trager, George L. 1940. One phonemic entity becomes two: The case of "short a." *American Speech.* 17. 30–41.

Twisk, Jos W. 2006. *Applied multilevel analysis : a practical guide*. Cambridge, UK; New York: Cambridge University Press.

Vennemann, Theo & Terence H. Wilbur. 1972. *Schuchardt, the neogrammarians and the transformational theory of phonological change: four essays*. (Ed.) Theo Vennemann & Terence Wilbur. Frankfurt: Athenäum.

Walker, James A. 2012. Form, function, and frequency in phonological variation. *Language Variation and Change.* 24. 397–415.

Wang, S.-Y. William. 1969. Competing changes as a cause of residue. *Language.* 45. 9–25.

Wang, William & Chinfa Lien. 1993. Bidirectional diffusion in sound change. In Charles Jones (ed.), *Historical Linguistics: problems and perspectives*, 345–400. London: Longman.

Wang, William S-Y. 1977. *The lexicon in phonological change*. The Hague: Mouton.

Wang, William S-Y. & C.-C. Cheng. 1977. Implementation of phonological change: the Shuang-feng Chinese case. In W Wang (ed.), *The lexicon in phonological change*. The Hague: Mouton.

Wang, William S.Y. 1979. Language change: a lexical perspective. *Annual Review of Anthropology.* 8. 353–71.

Weinreich, Ulrich, William Labov & Marvin Herzog. 1968. Empirical foundations for a theory of language change. In W P Lehmann & Y Makiel (eds.), *Directions in Historical Linguistics: a symposium*. Austin: University of Texas Press.

Wells, J C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Wickham, Hadley & Winston Chang. 2013. An implementation of the Grammar of Graphics. ggplot: Version 0.9.3.1.

Winteler, Jost. 1876. *Die Kerenzer Mundart des Kantons Glarus in ihren Grundzügen dargestellt*. Leipzig & Heidelberg: C. F. Winter'sche Verlagshandlung.

Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann & Han Sloetjes. 2006. ELAN: a Professional Framework for Multimodality Research. *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*. Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics, The Language Archive.

Wong, Shuet-Yin. 2005. Phonological Variation: The Case of the Syllable-final Velar Nasal in Cantonese. The University of Hong Kong, Ph.D. thesis.

Yaeger-Dror, Malcah. 1994. Phonetic evidence for sound change in Quebec French. In P. Keating (ed.), *Phonological Structure and Phonetic Form*, 267–92. Cambridge: Cambridge University Press.

Yan, Margaret Mian. 2006. *Introduction to Chinese dialectology*. Muenchen: LINCOM GmbH.

Yeung, Suk-Wah Helen. 1980. Some Aspects of Phonological Variations in the Cantonese Spoken in Hong Kong. The University of Hong Kong,.

Yue, Anne O. 2002. Development of the Stop Endings in the Yue Dialects. *Dialect Variations in Chinese: Papers from the Third International Conference on Sinology, Linguistics Section*. 217–245.

Zee, Eric. 1996. Phonological Changes in Hong Kong Cantonese. *Current Issues in Language and Society*. 3(2). 192–198.

Zee, Eric. 1999. Chinese (Hong Kong Cantonese). *Handbook of the International Phonetic Association*, 58–60. Cambridge: Cambridge University Press.

Zhan, Bohui & Yat-Shing Cheung. 1987. *珠江三角洲方言字音對照 [A survey of dialects in the Pearl River Delta: a comparative lexicon]. Volume I.* Hong Kong: New Century Publishing House.

Zhan, Bohui & Yat-Shing Cheung. 1990. *珠江三角洲方言綜述 [A Survey of Dialects in the Pearl River Delta, A Synthetic Review]. Volume III.* Hong Kong: New Century Publishing House.

Zhan, Bohui 詹伯慧. 1981. *Xiàndài Hànyǔ fāngyán 現代漢語方言 [The dialects of the modern Han language]*. Wǔhàn 武漢: Húběi Rénmín Chūbǎnshè 湖北人民出版社 .

Zipf, George K. 1949. *Human Behaviour and the principle of least effort*. Cambridge, MA: Addison-Wesley.

Appendices

Appendix I. Questions for casual speech elicitation

Spontaneous Questions

Your name is?

Language

What language do you use at home?
Do you speak any other languages apart from Cantonese?
Do your parents speak other dialects/languages?

School (for school children)

How far is your school from your house?
How do you get to school?
What was your favourite subject? Least favourite subject? Why?
What are teachers like at school?
Are they very strict?
Can you remember any trouble-makers at school? What do/did they do?
How about you, were you a trouble-maker in school?
What kinds of things do/did you do?
How were trouble-makers disciplined?
Did you ever pass notes in school?
What kinds of things did you pass notes about?
Did a teacher ever catch you passing notes?
What happened?
Did you ever have a teacher that was really fair? That you liked?
**Did you ever get blamed for something you didn't do?
Do you have lots of homework?
In what subjects?
How much time do you usually spend doing homework?

Teen life (for school children)

Do you spend a lot of time 'surfing the web'?
Do you have your own website?
Are you better at using the internet/email/MSN than your parents?
Why do you think that is?
Do you like to chat with you friends on MSN?
How much time do you spend on MSN?
Do your parents let you chat as long as you like?
**Did you ever have an argument with someone on MSN? What happened?

Do you like to play video games?
How do your parents feel about it?
What's your favorite game? How do you play it? What was your highest score?
What kind of music do you like? What kind of music are you listening to these days?
Have you ever been to a concert to see one of your favourite groups?
How was it?
What's the best concert you were ever at?
Who's your favorite artist? How come?
What is your favourite movie?
What was it about?
What did you do after school to keep yourself occupied?
Did you play sports?
Were you on any sports teams?
What ball games do you play?
Who do you play with?
Why do you like the game?
Do you own your own equipment?
What do you do during recess time?

Hobbies
Do you have any hobbies? What?
How did you get into that?
Did you ever go into competitions? Win a competition? What happened?
What extra-curricular activities do you join?
Have you ever been on a camp?
Did you ever go on a school trip? Where did you go? Did anything interesting ever happen?
Do you play any musical instruments?
If yes, which ones? For how long?
What made you start? *e.g. school, parents*
If no, is there an instrument you would like to learn to play? Why?

Birthdays
When is your birthday?
Are there any down sides to having your birthday when it is?
e.g. too close to Christmas, school is out for summer
**What is the best birthday party you ever had?
What is the best birthday party you ever went to?
Has anyone ever held a surprise birthday party for you?
Who did it?
Were you really surprised or did you pretend?
Has anyone ever forgotten your birthday?

Peer groups

What do kids do around here?
After school? At night? On weekends?
Is there a bunch of kids you hang around with, that you see almost every day?
Where do you hang out?
Who do you hang out with? [Get names and ages]
Have you ever pulled an 'all-nighter'?
Why did you do it?
Did you make it through the night?
Did you get through the next day?

Fights/Arguments

Have you ever witnessed a fight?
Where was it?
What was it about?
Do you ever have fights around here?
How do they start?
Do girls fight around here?
Did you ever get into a fight with a girl/guy?
Do you remember getting into an argument with someone?
Who was it with?
What was it about?
How did you resolve your differences? How did it all turn out?

Festivals/Traditions

What kinds of traditions can you remember growing up with in your family?
Do you (plan to) keep these traditions alive with your own family?
How did you celebrate Chinese New Year ?
Do you celebrate Christmas?

Neighbourhood

How long have you lived here?
Do you know your neighbours? What are they like?
Is there anyone around here who's a real snoop?
Someone who knows everybody else's business?
Someone who causes trouble?
Someone who plays loud music?
Some people say that nowadays everybody's just too busy to just stop by to chat.
What do you think? Why do you think that has changed?
What do you like best about your neighbourhood? What are the things that make you feel good/bad about your neighbourhood?
Did anything really big ever happen around here that you remember?

Like a big fire? Or a house burned down? Or a murder?
Where? Did you see it?
Did people in the neighbourhood help out? With food, clothes, place to stay?
What about accidents or police investigations?
**Do you remember when SARS broke out?
What did you do when SARS broke out, and were you afraid?

Parents/Family

Where were your parents born and raised?
Your grand parents?
Do you have any idea how long your family has been living here?
Where did they originally come from?
Who brought you up?
What kind of upbringing did you have?
Where your parents really strict?
What sort of person is your father?
What is your mother like?
What kind of kid were you when you were growing up?
How where you punished? By whom?
Did you ever get blamed for something you didn't do?
If you got into trouble from your parents could you talk to them?
Which parent would you choose to talk to? Why?
Did your parents have certain expectations from you?
Do you have siblings? How many?
Were you close to your siblings growing up or did you fight a lot?
How about now?
Did you ever go on vacations as a family?
Where would you go? How would you get there?
Did you get along while you were on vacation?

Dating Practices

Do you remember going out on dates? Do you go out on dates? Have you ever gone on a date with a boy/girl?
When did you (or your friends) start dating?
How did your parents react?
How long did you date?
What kind of guy/girl do you like?
What kind of guy/girl do you want to marry?

Food/remedies

A lot of people remember their mom/dad making special foods? What do you remember about the foods your mom/dad used to make? What did you mother/father, grandmother/grandfather like to eat?

Do you ever take Chinese medicine? Have herbal treatments/foods?

I've heard people say that in the olden days, people relied on traditional medicines and natural remedies, do you remember that?

Did people go to doctors in the olden days? Do you go to the doctor when you're sick?

What do you do to prevent colds/flu? What do you do when you get sick?

Dreams

What is the most vivid dream you've ever had?

Do you remember any nightmares that you had?

Uncommon experiences

Have you ever witnessed a terrible accident or tragic event?

What happened?

Have you ever been injured? How?

Was there ever anything that happened when you were growing up that you couldn't explain?

Language judgments

Some people think that Cantonese has changed/is changing, do you think so?

How about the difference between old and young speakers?

What do you think about the way that youth today sound?

What has changed?

Appendix II. Passage reading

做成功的人

香港，可說有不少百萬富翁。以下是獅子山下的故事，大綱如此：

楊敏先生是一位醫生，在醫院做工。很久以前，他生活艱維，很窮很命苦，常常被別人看不起，以為自己沒福分過好生活，與社會分割開。但有一天他在伯伯黃鵬飛的家中作客，黃伯伯跟他講，很多人之所以失敗是因為不好好裝備自己，結果弱不勝強，被人淘汰，感到彷徨及對世事疑惑不解，做一個沒靈魂的空殼。這番說話就好像明燈，明亮地照著楊敏前路。

所以，楊敏從少便很渴望有成就，很聽話和自覺，如果不跟循社會規則，趕忙拼搏，默默耕耘，為自己的作為負責，分析面前的困難，所有往成功的渠道便會閉塞，機會枯乾，很難離得開貧苦的厄運。結果他很努力地讀書，看名著劇本，慢慢咀嚼，最後不但騰飛猛進，鏗鏗鏘鏘成為一位成功的醫生，賺得很多錢，還成為一個名副其實，有內涵，八面玲瓏的人，為自己爭回一口氣。

相反，楊敏的弟弟很硬頸，質疑黃伯伯番說話。他不但有很多不良癖好，還很狡猾，很懶惰。但正如俗話所說，假如他願意懸崖勒馬的話，重新奮鬥，他有一天也會有收穫，從失敗學習，做到一個成功的人。

Appendix III. Word list

Part 1

1	香港百萬富翁
2	油燈
3	騰飛
4	勒褲帶
5	怪癖
6	伯父
7	福分
8	衛生部
9	作弊
10	硬繃繃
11	通行
12	擱放
13	劇本
14	困惑無助
15	行行出狀元
16	力不從心
17	作客
18	剝下
19	食蕃薯
20	背膂
21	癖好
22	燈不離手
23	拼搏
24	唔爭氣
25	尺幅
26	分析
27	行後巷快啲
28	禿髮症
29	同步
30	毒品
31	中央分局
32	羊角風
33	容貌
34	懸崖勒馬
35	上學期
36	恆香餅模
37	欲哭無淚
38	楊敏先生係醫生
39	項目
40	有兩位觀眾想回應
41	從緩
42	性別
43	逛下街
44	成日
45	點燈候月
46	光復
47	疑惑會解開
48	厄運
49	軟殼蟹
50	闖禍
51	落花流水
52	藥房
53	曲學
54	橫霸
55	白板
56	行為
57	孟買
58	光榮
59	少則得，多則惑
60	逼迫
61	自覺
62	內閣密談
63	逆反
64	凶猛
65	症患
66	客廳
67	贈品
68	俗話說
69	澤被天下
70	宏偉
71	擱筆
72	各謀其事
73	江門市
74	岳飛
75	作為
76	幫忙
77	劇舞劇話
78	塊餅崩埋添!
79	好聽話
80	架遊艇快過車!
81	行快啲
82	客氣
83	北緯
84	梁武帝
85	削髮為尼
86	夢鄉
87	棘馬腳
88	驚喜
89	或許
90	防避

91	提燈會
92	黑飯
93	黃鵬飛
94	雞肋
95	燈謎
96	衛生好重要
97	行勻成校都揾唔到
98	面目可憎
99	松柏後凋
100	澎湃
101	慢慢
102	行乞
103	文質彬彬
104	枯乾
105	分割開
106	薆賣
107	刮目相看
108	閒氣
109	突發
110	身痕
111	散髮

Part 2

101	鏗鏗鏘鏘
102	鼻塞
103	逛勻成條街
104	凱勒・海倫
105	架遊艇會唔會反呢?
106	特別
107	連牆都鑿崩埋!
108	石馬
109	封筆
110	白飯
111	策劃
112	唔好再爭
113	踢下波
114	佢好命苦
115	食麥餅
116	直下
117	搶掠去
118	逛埋下條街
119	剛剛
120	嚇暈咗
121	上樑不正下樑歪
122	滄海
123	講話
124	蒙允
125	祈福會
126	劈開
127	北非
128	名副其實
129	裝備
130	搶掠橫逆
131	等飯食
132	防範
133	成為
134	石壩
135	各奔各向
136	抗衡
137	迎合
138	拼命
139	噩夢
140	佢頸部好乾
141	掌摑下佢，等佢醒返
142	不負國望
143	搭棚
144	腳腕
145	相反
146	長江
147	約莫
148	覺幻
149	默默耕耘
150	彷徨
151	打橫行
152	盈富
153	若無其事
154	爭玩具
155	愛錫花朵
156	雀斑
157	窮苦
158	入黑
159	同人爭武器
160	收穫
161	行後巷會唔會快啲呢?
162	蠱惑反叛
163	拍門
164	呢本書好輕好薄
165	爭番啖氣
166	空殼
167	你好叻!
168	伯伯
169	通風
170	豪宅房地產

171	刻苛
172	礦物質
173	疑惑不解
174	咀嚼
175	內閣會議
176	故事大綱係講…
177	浙江北部
178	硬頸
179	省分
180	歌曲
181	食勻晒
182	弱不勝強
183	肋窩
184	行埋去
185	清明
186	塞血管
187	燈火
188	抱著希望
189	江澤民
190	目無法紀
191	新疆行政區

192	加勒比海
193	救生滅火
194	返工
195	鑿雲石
196	麥米粥
197	曲簿
198	疫患
199	喺呢間廟求籤好靈好
200	凸紋
201	不非
202	遺失
203	黑色嘅雁斑
204	還鄉
205	八萬蚊
206	警察
207	趕往
208	動物餅
209	察看
210	汗漫
211	頸喝

Appendix IV. Minimal pairs

1	上(oe:ng)	下
2	波士	多士
3	那兒	那時
4	新(an+m)聞(an)	生(a:ng+m)聞
5	閣(o:k+h)下	割(o:t+h)下
6	毛	無
7	歌	哥
8	高	低
9	生(ang)	新(an)
10	江(o:ng+m)門(u:n)	乾(o:n+m)門(u:n)
11	麥(ak)	物(at)
12	媽寶	孖寶
13	前(i:n)	後
14	波	多
15	剛(o:ng)	乾(o:n)
16	殼(o:k)	喝(o:t)
17	坦(a:n)	蛋(a:n)
18	薑(an+m)買	等(ang+m)買
19	真(an)	珍(an)
20	毛病(e:ng)	無病(e:ng)
21	物(ath)盒	麥(ak+h)盒
22	巷(o:ng)	刊(o:n)
23	各(o:k)	割(o:t)
24	北(ak)	不(at)
25	左	右
26	歌兒	哥兒
27	飲	忍(an)
28	全(y:n+s)心	全(y:n)新(an)
29	角(o:k+m)膜(o:k)	割(o:t+m)膜(o:kz)
30	七(at)	不(at)
31	策(a:k)	察(a:t)
32	七(at+s)星(ing)	不(at+s)升(ing)
33	真(an+ts)豬	珍(an+ts)珠
34	兒	時
35	憑(ang+m)咩	貧(an+m)咩
36	心	新(an)
37	成(ing+k)功(ung)	失(at+p)敗
38	懶(a:n+h)氣	冷(a:ng+h)氣
39	幸(ang+h)好	恨(an+h)好
40	橫(a:ng)	環(a:n)
41	擦(a:t+m)埋	拆(a:k+m)埋
42	刻(ak+h)好	乞(at+h)好
43	橙(a:ng+s)色	慘色
44	文(an+m)武	神(an+m)武
45	黑(ak)	白(a:k)
46	媽	孖
47	好好	差極(ik)
48	照入	照日(at)
49	特(ak)	凸(at)
50	等(ang)	薑(an)
51	百(a:k)	八(a:t)
52	冷(a:ng)	懶(a:n)
53	中(ung)	忠(ung)
54	豬	書
55	殼(o:k+m)物(at)	喝(o:t+m)物(at)
56	橙(a:ng)	慘
57	特(ak+m)聞(an)	凸(at+m)紋(an)
58	角(o:k)	割(o:t)
59	摑(a:k)	刮(a:t)
60	坦(a:n+p)白(a:k)	蛋(a:n+p)白(a:k)
61	幸(ang)	恨(an)
62	講(o:ng)	趕(o:n)
63	北(ak+m)美	不(at+m)美
64	行(a:ng+m)埋	閒(a:n+m)埋
65	江(o:ng)	乾(o:n)
66	出(ut)	入
67	八(a:t+h)合	百(a:k+h)合
68	大豬	大書
69	摑(a:k+h)下	刮(a:t+h)下
70	巷(o:ng+h)後	刊(o:n+h)後
71	橫(a:ng+p)霸	環(a:n+p)壩
72	嘗(oe:ng+s)試	放(o:ng+h)棄
73	策(a:k+m)名(ing)	察(a:t+m)明(ing)
74	文(an)	神(an)
75	入	日(at)
76	又係	就係
77	講(o:ng+m)乜	趕(o:n+m)乜
78	知	衣
79	猛(a:ng+h)霞	晚(a:n+h)霞
80	閣(o:k)	割(o:t)
81	憑(ang)	貧(an)
82	刻(ak)	乞(at)
83	又	就
84	乾(on+h)好	剛(ong+h)好
85	鐘(ung+j)意	憎(ang+h)恨
86	閒(a:n)	行(a:ng)
87	各(o:k+h)去	割(o:t+h)去
88	猛(a:ng)	晚(a:n)
89	淺(i:n+p)薄(o:k)	深厚
90	努力(ik)	懶(a:n+t)惰
91	拆(a:k)	擦(a:t)
92	中(ung+k)間(a:n)	忠(ung+k)奸(a:n)
93	難(a:n+j)飲	難(a:n+j)忍(an)
94	髮(a:t+s)飾(ik)	法(a:t+s)式(ik)
95	無知	毛衣
96	髮(a:t)	法(a:t)

Appendix V. Text files used for forced alignment

Reading Passage

hoeng gong, ho syut jau bat siu baak maan fu jung. ji haa si si zi saan haa dik gu si, daai gong jyu ci:
joeng man sin saang si jat wai ji sang, zoi ji jyun zou gung. han gau ji cin, taa sang wut gaan wai, han kung han ming fu
soeng soeng bei bit jan hon bat hei, ji wai zi gei mut fuk fan gwo hou sang wut. jyu se wui fan got hoi.
daan jau jat tin taa zoi baak baak wong paang fei dik gaa zung zok haak. wong baak baak gan taa gong
han do jan zi so ji sat baai si jan wai bat hou hou zong bei zi gei
git gwo joek bat sing koeng, bei jan tou taai, gam dou pong wong kap deoi sai si ji waak bat gaai
zou jat go mut ling wan dik hung hok. ne faan syut waa zau hou zoeng ming dang, ming loeng dei ziu zoek joeng man cin lou
so ji, joeng man cung siu bin han hot mong jau sing zau, han ting waa wo zi gok, jyu gwo bat gan ceon se wui kwai zak, gon mong ping bok, mak mak gaang wan
wai zi gei dik zok wai fu zaak, fan sik min cin dik kwan naan, so jau wong sing gung dik keoi dou bin wui bai sak, gei wui fu gon,

han naan lei dak hoi pan fu dik aak wan. git gwo taa han nou lik dei duk syu, hon ming zyu kek bun,
maan maan zeoi zoek, zeoi hau bat daan tang fei maang zeon, hang hang coeng coeng sing wai jat wai sing gung dik ji sang,
zaan dak han do cin, waan sing wai jat go ming fu kei sat, jau noi haam, baat min ling lung dik jan, wai zi gei zang wui jat hau hei
soeng faan, joeng man dik dai dai han ngaang geng, zat ji wong baak baak faan syut waa.
taa bat daan jau han do bat loeng pik hou, waan han gaau waat, han laan do. daan zing jyu zuk waa so syut,
gaa jyu taa jyun ji jyun ngaai lak maa dik waa, cung san fan dau, taa jau jat tin jaa wui jau sau wok, cung sat baai hok zaap, zou dou jat go sing gung dik jan

Word List 1

hoeng gong baak maan fu jung
jau dang
tang fei
lak fu daai
gwaai pik
baak fu
fuk fan
wai sang bou
zok bai
ngaang baang baang
tung hang
gok fong
kek bun
kwan waak mou zo
hong hong ceot zong jyun
lik bat cong sam
zok haak
mok haa
sik faan syu
zek zeoi
pik hou
dang bat lei sau
ping bok
ng zaang hei
cek fuk
fan sik
haang hau hong faai di
tuk faat zing
tung bou
duk ban
zung joeng fan guk
joeng gok fung
jung maau
jyun ngaai lak maa
soeng hok kei
hang hoeng beng mou
juk huk mou leoi
joeng man sin saang hai ji sang
hong muk
jau loeng wai gun zung soeng wui
jing
cung wun
sing bit
kwaang haa gaai
sing jat
dim dang hau jyut
gwong fuk
ji waak wui gaai hoi
aak wan
jyun hok haai
cong wo

lok faa lau seoi
joek fong
kuk hok
waang baa
baak baan
hang wai
maang maai
gwong wing
siu zak dak, do zak waak
bik baak
zi gok
noi gok mat taam
jik faan
hung maang
zing waan
haak teng
zang ban
zuk waa syut
zaak bei tin haa
wang wai
gok bat
gok mau kei si
gong mun si
ngok fei
zok wai
bong mong
kek mou kek waa
faai beng bang maai tim
hou teng waa
gaa jau teng faai gwo ce
haang fai di
haak hei
bak wai
loeng mou dai
soek faat wai nei
mung hoeng
kik maa goek
ging hei
waak heoi
fong bei
tai dang wui
hak faan
wong paang fei
gai lak
dang mai
wai sang hou zung jiu
haang wan seng haau dou wan ng dou
min muk ho zang
cung paak hau diu
paang baai

maan maan
hang hat
man zat ban ban
fu gon
fan got hoi
dan maai
gwaat muk soeng hon
haan hei
dat faat
san han
saan faat
san goeng hang zing keoi
gaa lak bei hoi
gau sang mit fo
faan gung
zok wan sek
mak mai zuk
kuk bou
jik waan
hai ne gaan miu kau cim hou leng
hou zun

Word List 2

hang hang coeng coeng
bei sak
kwaang wan seng tiu gaai
hoi lun hoi lak
gaa jau teng wui ng wui faan ne
dak bit
lin coeng dou zok bang maai
sek maa
fung bat
baak faan
caak waak
ng hou zoi zaang
tek haa bo
keoi hou meng fu
sik mak beng
zik haa
coeng loek heoi
kwaang maai haa tiu gaai
gong gong
haak wan zo
soeng loeng bat zing haa loeng waai
cong hoi
gong waa
mung wan
kei fuk wui
pek hoi
bak fei
ming fu kei sat
zong bei
coeng loek waang jik
dang faan sik
fong faan
sing wai
sek baa
gok ban gok hoeng
kong hang
jing hap
ping ming
ngok mung
kui geng bou hou gon
zoeng kwaat haa keoi, dang keoi seng faan
bat fu gwok mong
daap paang
goek wun
soeng faan
coeng gong
joek mok
gok waan
mak mak gaang wan
pong wong
daa waang haang
jing fu
joek mou kei si
zaang wun geoi
oi sek faa do
zoek baan
kung fu
jap hak
tung jan zaang mou hei
sau wok
haang hau hong wui ng wui faai di ne?
gu waak faan bun
paak mun
ne bun syu hou heng hou bok
zaang faan daam hei
hung hok
nei hou lek
baak baak
tung fung
hou zaak fong dei caan
hak ho
kwong mat zat
ji waak bat gaai
zeoi zoek
noi gok wui ji
gu si daai gong hai gong
zit gong bak bou
ngaang geng
saang fan
go kuk
sik wan saai
joek bat sing koeng
lak wo
haang maai heoi
cing ming

sak hyut gun
dang fo
pou zoek hei mong
gong zaak man
muk mou faat gei
san goeng hang zing keoi
gaa lak bei hoi
gau sang mit fo
faan gung
zok wan sek
mak mai zuk
kuk bou
jik waan

hai ne gaan miu kau cim hou leng
hou zun
dat man
bat fei
wai sat
hak sik ge ngaan baan
waan hoeng
baat maan man
ging caat
gon wong
dung mat beng
caat hon
hon maan
geng hot

Appendix VI. Variables coding scheme

1. Final consonant

Code	Description
g	ŋ
n	n
k	k
t	t
m	m
p	p

2. Final consonant realisation

Code	Description	Remarks
G	ŋ	clear velar closure
N	n	clear alveolar closure
K	k	clear velar closure
T	t	clear alveolar closure
M	m	clear bilabial closure
P	p	clear bilabial closure
@	zero	Final consonant is lost, but no glottalisation
?	glottal	abrupt cessation of pulmonic pulse energy
!	glottalisation	Several glottal pulses before pulmonic pulse ceases

3. Preceding vowel

Code	Description
A	a:
Q	œ:
E	ɛ:
O	ɔ:
a	ɐ
i	ɪ
u	ʊ

4. Nasalisation

Code	Description
~	Full
^	Partial
-	oral/not nasal

5. Following phonetic segment

Code		Description	Recoded	fi_new
#	1	pause	other	2
@	2	zero	other	2
B	3	p	labial	1
P	4	p^h	labial	1
M	5	m	labial	1
F	6	f	labial	1
D	7	t	alveolar	4
T	8	t^h	alveolar	4
N	9	n	alveolar	4
L	10	l	alveolar	4
G	11	g	velar	0
K	12	k	velar	0
&	13	ŋ	velar	0
H	14	h	other	2
$	15	k^w	velar	0
^	16	k^{hw}	velar	0
W	17	w	labial	1
Z	18	ts	alveolar	4
C	19	ts^h	alveolar	4
S	20	s	alveolar	4
J	21	j	palatal	3
Q	22	ʔ	other	2

7. Lexical identity

Code	Description
[Chinese character]	Identity of lexical item

8. Task

Code	Description
i	informal, spontaneous speech
a	arithmetics game
s	shop naming game
r	Reading task 1
r	Reading task 2
w	Wordlist 1
W	Wordlist 2
m	Minimal pairs
p	Perception test

6. Preceding onset

Code	Description
@	zero
B	b
P	p
M	m
F	f
D	t
T	t^h
N	n
L	l
G	k
K	k^h
&	ŋ
H	h
$	k^w
^	k^{hw}
W	w
Z	ts
C	ts^h
S	s
J	j
Q	ʔ

9. Speaker identity

Code	Description
1, 2, …	Number of speaker according to sound recording order

10. Speaker gender

Code	Description
M	male
F	female

11. Speaker age

Code	Description
1	13-18 yrs old
2	19-39 yrs old
3	40-70 yrs old

Appendix VII. Lexical distribution of the HKCAC

A scatterplot of the rank of all the types (or lexemes) in the HKCAC (Leung, Law & Fung 2001) by frequency is as shown in Figure 48:

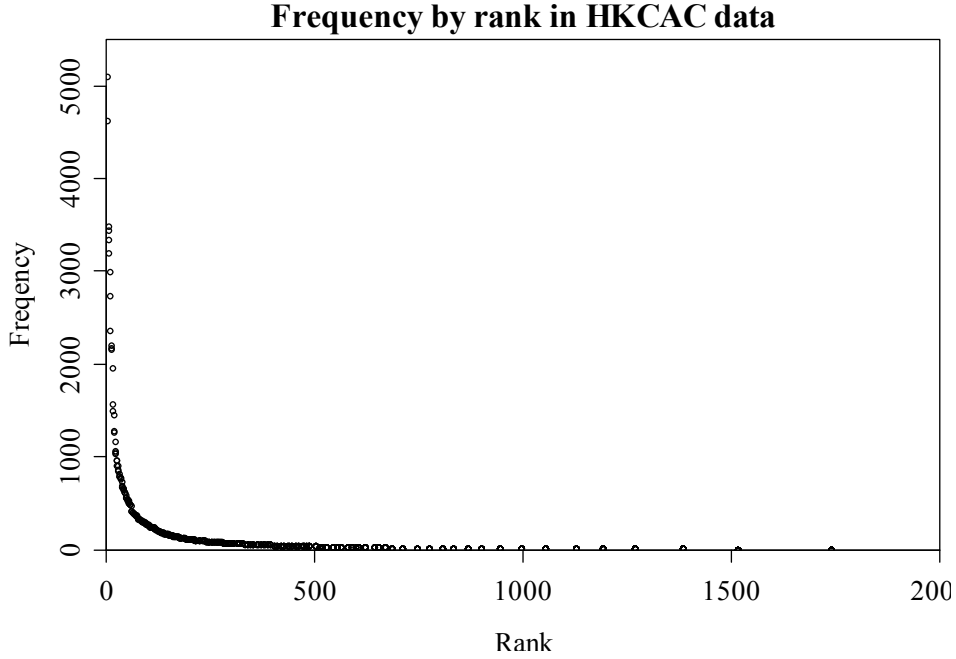
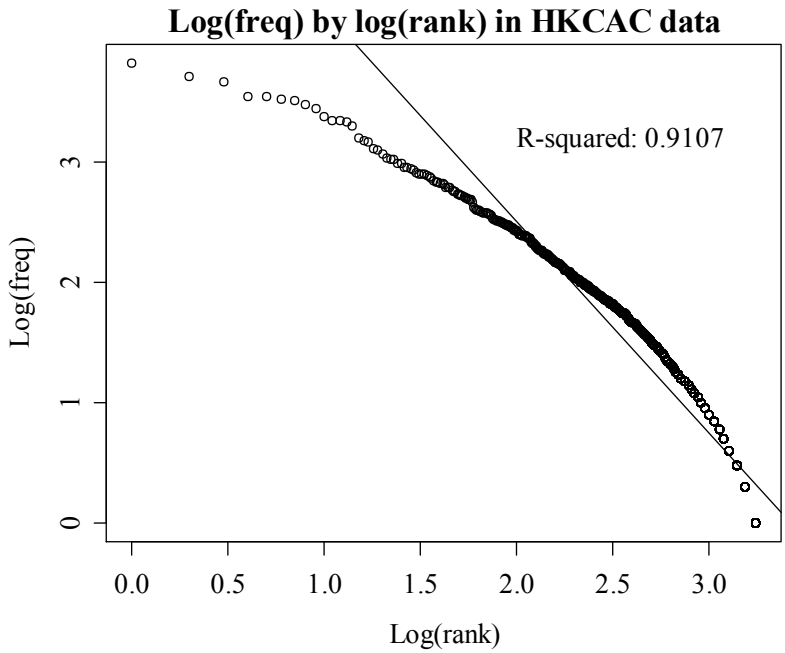


Figure 48. Scatter plot of frequency by rank in HKCAC data

Taking the logarithm of both frequency and rank, we obtain the plot in Figure 49, which shows an inverse linear doubly-logarithmic relationship. The lopsided distribution in this corpus was found to conform quite closely to Zipf's Law, which states that if types in a given corpus are ranked by increasing frequency of occurrence (from high—or numerically low—rank, towards low—or numerically high—word rank), the log-transformed frequency counts will show a negative linear correlation with log-transformed word rank (Zipf 1949).



y = 6.02 -1.76 x;
residual standard error: 0.2308 on 2147 degrees of freedom
R-squared: 0.9107

Figure 49. Scatterplot of log(frequency) by log(rank) of HKCAC data with linear regression line

As an aside, a better inverse linear relationship between frequency and rank is obtained between log-transformed frequency and $1+\sqrt[3]{rank}$, as can be seen in Figure 50. Whichever of the above functions is used to manipulate 'rank' with respect to 'log(frequency)', what is relevant is that in both cases 'log(frequency)' decreases monotonically when rank increases.

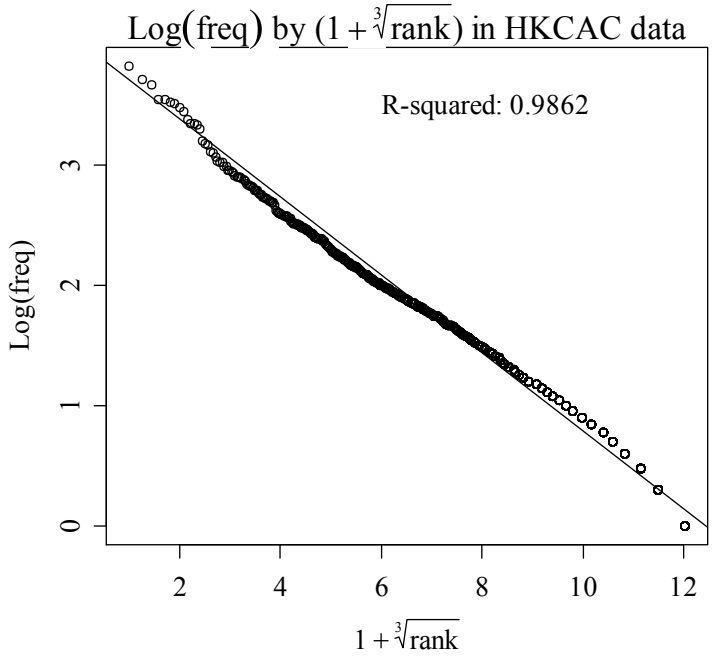


Figure 50. Scatter plot of log(frequency) by $1+\sqrt[3]{rank}$ in HKCAC data with linear regression line

Another way to understand the distribution of types in the HKCAC is by looking at the number of types that have a frequency count of 1 (hapax legomena, which have the lowest rank), a frequency count of 2 (dis legomena, which have the second-lowest rank), a frequency count of 3 (tris legomena, which have the third-lowest rank), and so on (cf. Baayen 2001; Baayen 2008, Ch. 6.5). We would obtain a histogram as shown in Figure 51:

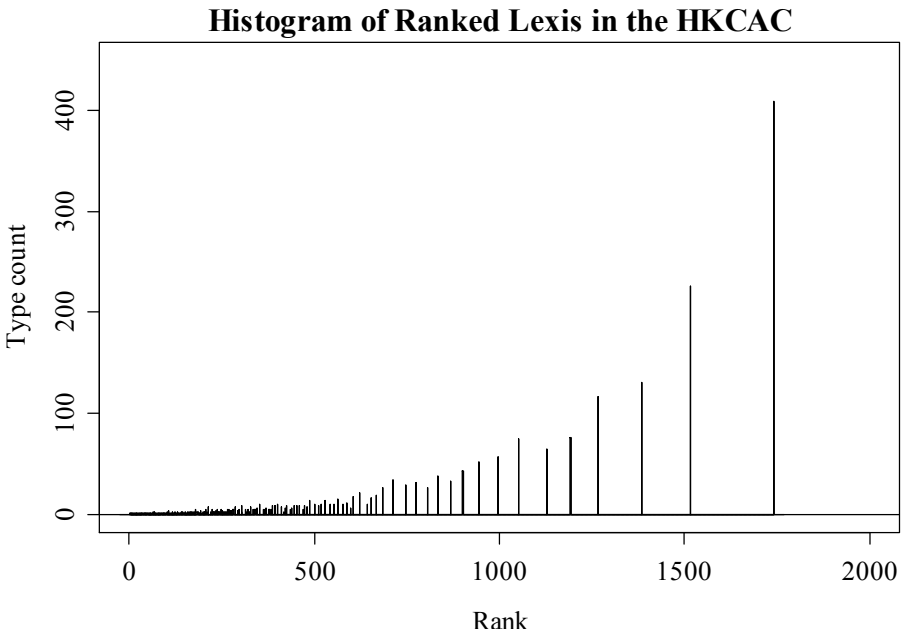


Figure 51. Frequency by rank of words in the HKCAC

Appendix VIII. Raw data

Table 72. Raw counts and percentages of (ŋ) realization by individual speaker

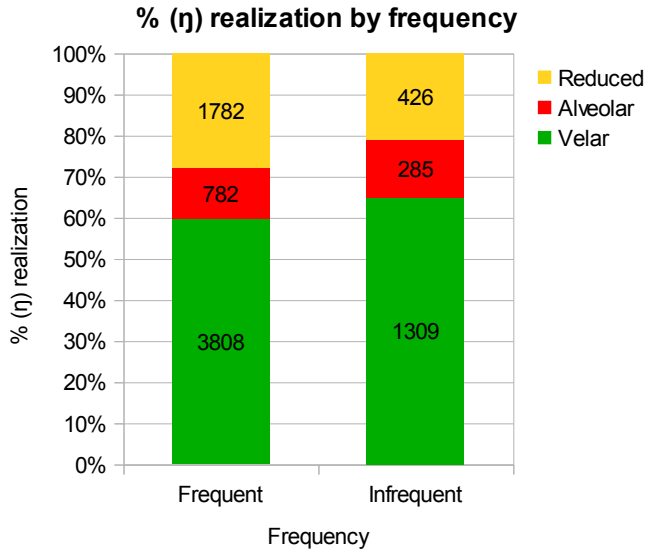
	Velar [ŋ]		Alveolar [n]		Reduced		Total
YM14	203	60.8%	3	11.1	9	28.1	**334**
YF14A	208	70.7%	2	6.8%	6	22.4	**294**
YF14B	128	50.6%	1	5.9%	1	43.5	**253**
YM15	174	63.7%	1	5.1%	8	31.1	**273**
YF15A	120	42.3%	4	16.5	1	41.2	**284**
YF15B	125	43.1%	2	9.0%	1	47.9	**290**
YM16	161	61.9%	1	7.3%	8	30.8	**260**
YF17	172	56.8%	1	6.3%	1	37.0	**303**
YM18	189	69.2%	1	7.0%	6	23.8	**273**
YM18B	135	39.6%	7	22.6	1	37.8	**341**
MM19	211	62.6%	5	15.1	7	22.3	**337**
MM26	122	37.2%	1	32.9	9	29.9	**328**
MF26A	124	41.3%	7	25.3	1	33.3	**300**
MF26B	151	48.7%	8	26.1	7	25.2	**310**
MM32	95	40.8%	6	26.6	7	32.6	**233**
MF32	64	29.0%	9	40.7	6	30.3	**221**
MM33	86	32.5%	1	41.1	7	26.4	**265**
MM35	193	85.8%	3	1.3%	2	12.9	**225**
MF35	180	76.6%	7	3.0%	4	20.4	**235**
MF36	117	40.9%	7	25.9	9	33.2	**286**
OF52A	174	68.8%	1	6.7%	6	24.5	**253**
OF52B	229	81.5%	1	4.6%	3	13.9	**281**
OM53	191	70.5%	1	5.5%	6	24.0	**271**
OM53B	214	77.0%	7	2.5%	5	20.5	**278**
OM54	193	68.7%	1	6.0%	7	25.3	**281**
OF55	235	82.7%	1	5.3%	3	12.0	**284**
OF56	258	92.8%	1	0.4%	1	6.8%	**278**
OF57	206	73.6%	8	2.9%	6	23.6	**280**
OM61	230	83.6%	1	5.5%	3	10.9	**275**
OM65	229	86.1%	5	1.9%	3	12.0	**266**
Total	**5117**		**1067**		**2208**		**8392**

Table 73. Raw counts and percentages of (k) realization by individual speaker

	Velar [k]		Alveolar [t]		Reduced		Total
YM14	76	36.5%	3	47.1%	9	16.3	**208**
YF14A	64	33.5%	1	57.1%	1	9.4%	**191**
YF14B	47	25.1%	1	69.0%	1	5.9%	**187**
YM15	60	33.0%	1	59.3%	1	7.7%	**182**
YF15A	33	16.4%	1	76.6%	1	7.0%	**201**
YF15B	31	14.0%	1	77.8%	1	8.1%	**221**
YM16	33	22.1%	6	73.8%	1	4.0%	**149**
YF17	51	25.9%	8	70.1%	1	4.1%	**197**
YM18	51	29.0%	9	65.9%	1	5.1%	**176**
YM18B	37	19.4%	4	59.7%	1	20.9	**191**
MM19	78	40.2%	1	53.1%	1	6.7%	**194**
MM26	45	25.6%	2	63.1%	1	11.4	**176**
MF26A	33	16.5%	2	72.5%	1	11.0	**200**
MF26B	52	26.5%	3	54.1%	1	19.4	**196**
MM32	35	19.8%	1	73.4%	1	6.8%	**177**
MF32	33	19.8%	2	63.5%	1	16.8	**167**
MM33	30	18.0%	2	68.3%	1	13.8	**167**
MM35	93	65.0%	5	31.5%	4	3.5%	**143**
MF35	52	29.9%	9	64.9%	1	5.2%	**174**
MF36	60	29.3%	2	56.6%	1	14.1	**205**
OF52A	46	25.7%	1	63.7%	1	10.6	**179**
OF52B	85	51.5%	6	44.8%	7	3.6%	**165**
OM53	41	19.8%	7	76.8%	1	3.4%	**207**
OM53B	33	20.4%	2	78.4%	1	1.2%	**162**
OM54	74	42.0%	6	54.5%	9	3.4%	**176**
OF55	75	37.3%	2	52.7%	1	10.0	**201**
OF56	97	51.6%	4	46.3%	8	2.1%	**188**
OF57	80	39.4%	7	57.1%	1	3.4%	**203**
OM61	58	29.1%	6	67.8%	1	3.0%	**199**
OM65	56	37.3%	6	58.7%	8	4.0%	**150**
Total	**1639**		**4**		**3**		**5532**

Table 74. Realization of (ŋ) by frequency as a categorical variable (HKCAC frequency < 7: infrequent; HKCAC frequency ≥ 7 frequent)

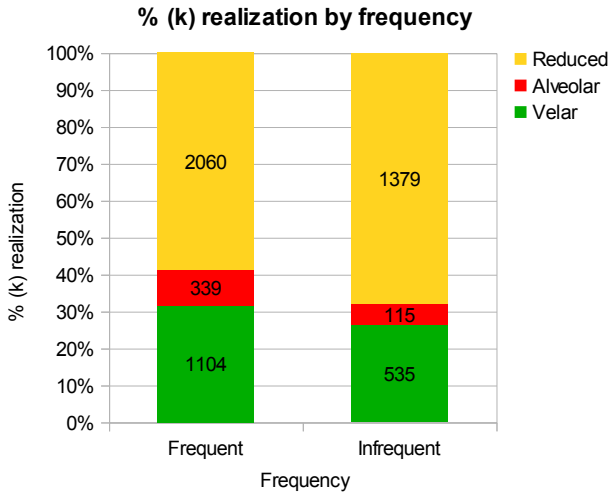
(ŋ) realization	Frequency		
	Frequent	Infrequent	Total
Velar [ŋ]	3808	1309	**5117**
Alveolar [n]	782	285	**1067**
Reduced	1782	426	**2208**
Total	**6372**	**2020**	**8392**



N = 8392

$\chi^2 = 38.0286$, d.f. = 2, p-value = 5.523×10^{-09}

Figure 52. Realization of (ŋ) by frequency as a binary categorical variable (HKCAC frequency < 7: infrequent; HKCAC frequency ≥ 7 frequent)



N = 5532

$\chi^2 = 53.9955$, d.f. = 2, p-value = 1.884×10^{-12}

Figure 53. Realization of (k) by frequency as a categorical variable (HKCAC frequency < 7: infrequent; HKCAC frequency ≥ 7: frequent)

Table 75. Realization of (k) by frequency as a categorical variable (HKCAC frequency < 7: infrequent; HKCAC frequency ≥ 7 frequent)

	Frequency		
(k) Realization	**Frequent**	**Infrequent**	**Total**
Velar	1104	535	**1639**
Alveolar	339	115	**454**
Reduced	2060	1379	**3439**
Total	**3503**	**2029**	**5532**

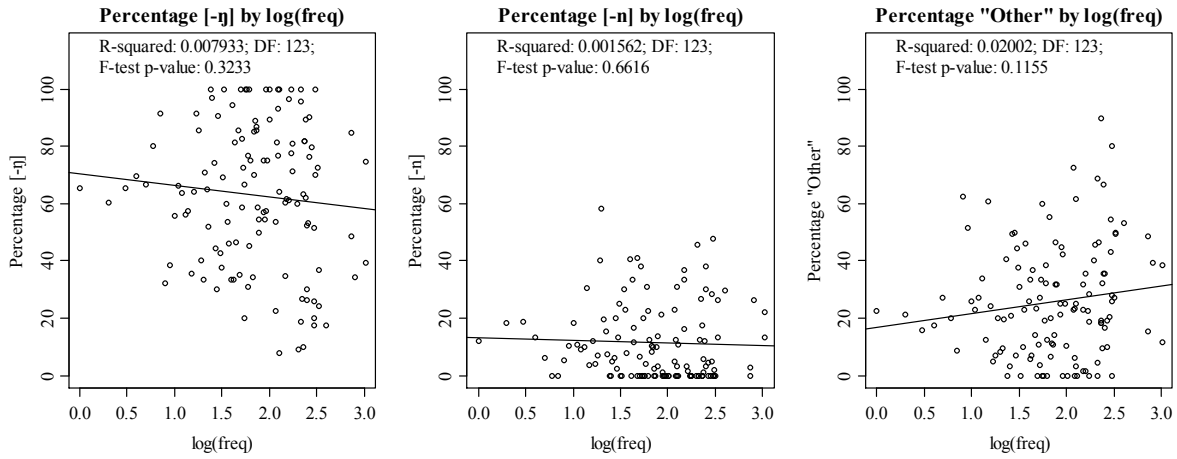


Figure 54. Percentage realization of (ŋ) by log(freq)

Each data point represents one Chinese lexeme out of 125 for (ŋ) used in this study. The percentages of [-ŋ], [-n], and [reduced] realizations were obtained for each lexeme and plotted in separate graphs. The total number of tokens used, N=8392.

This graph does not have "ceiling effects" removed, so all the data points which show no variation are still included in the graph.

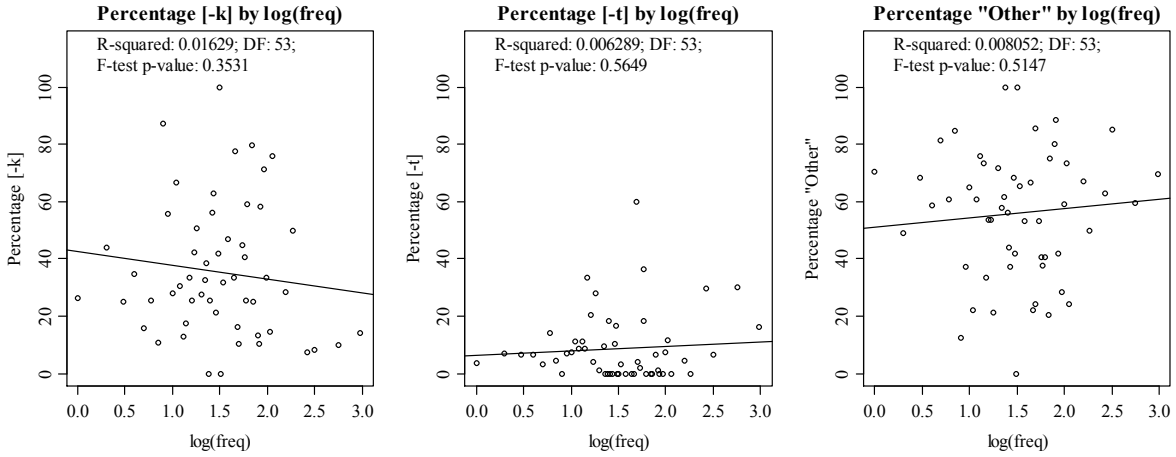


Figure 55. Percentage realization of (k) by log(freq), raw data

Each data point represents one out of 55 lexemes for (k) used in this study. The percentages of [-k], [-t], and [reduced] realizations were obtained for each lexeme and plotted in separate graphs. The total number of tokens used, N = 5532

This graph does not have "ceiling effects" removed, so all the data points which show no variation are still included in the graph.

Appendix IX. Mathematical modelling results

Model 1. "Best" mixed-effects multinomial logistic regression model for response variable "(ŋ) realization" obtained in MLwiN

$$resp_{ijk} \sim \text{Multinomial}(cons_{jk}, \pi_{ijk})$$

$$\log(\pi_{1jk} / \pi_{0jk}) = \beta_{0k}\text{cons.sr_1}_{ijk} + \text{-0.042(0.006)age.sr_1}_{ijk} + \text{0.313(0.354)gender_2.sr_1}_{ijk} +$$
$$\text{-0.748(0.269)style_4.sr_1}_{ijk} + \text{-1.055(0.272)style_6.sr_1}_{ijk} + \text{-0.672(0.324)style_8.sr_1}_{ijk} +$$
$$\text{-0.008(0.008)age.gender_2.sr_1}_{ijk} + \text{-0.120(0.368)gender_2.style_4.sr_1}_{ijk} +$$
$$\text{0.105(0.357)gender_2.style_6.sr_1}_{ijk} + \text{0.201(0.415)gender_2.style_8.sr_1}_{ijk} +$$
$$\text{0.024(0.121)tone_2.sr_1}_{ijk} + \text{-0.084(0.224)tone_3.sr_1}_{ijk} + \text{-0.274(0.104)tone_4.sr_1}_{ijk} +$$
$$\text{-0.157(0.197)tone_5.sr_1}_{ijk} + \text{0.311(0.142)tone_6.sr_1}_{ijk} + \text{1.852(0.255)fi_new_1.sr_1}_{ijk} +$$
$$\text{2.567(0.242)fi_new_2.sr_1}_{ijk} + \text{3.371(0.284)fi_new_3.sr_1}_{ijk} + \text{3.937(0.237)fi_new_4.sr_1}_{ijk} +$$
$$\text{0.356(0.213)aa.sr_1}_{ijk} + \text{-0.858(0.269)ee.sr_1}_{ijk} + \text{-4.351(0.354)i.sr_1}_{ijk} + \text{-1.024(0.209)oo.sr_1}_{ijk} +$$
$$\text{-0.153(0.232)oe.sr_1}_{ijk} + \text{-5.253(0.457)u.sr_1}_{ijk} + \text{-0.306(0.264)(logltfreq-gm).sr_1}_{ijk} +$$
$$\text{0.189(0.262)fi_new_1.(logltfreq-gm).sr_1}_{ijk} + \text{0.399(0.257)fi_new_2.(logltfreq-gm).sr_1}_{ijk} +$$
$$\text{0.454(0.311)fi_new_3.(logltfreq-gm).sr_1}_{ijk} + \text{0.968(0.256)fi_new_4.(logltfreq-gm).sr_1}_{ijk} +$$
$$\text{0.200(0.126)aa.(logltfreq-gm).sr_1}_{ijk} + \text{-0.651(0.220)ee.(logltfreq-gm).sr_1}_{ijk} +$$
$$\text{-0.884(0.298)i.(logltfreq-gm).sr_1}_{ijk} + \text{-0.314(0.125)oo.(logltfreq-gm).sr_1}_{ijk} +$$
$$\text{-0.284(0.142)oe.(logltfreq-gm).sr_1}_{ijk} + \text{-0.061(0.502)u.(logltfreq-gm).sr_1}_{ijk}$$

$$\beta_{0k} = \text{-1.131(0.366)} + v_{0k}$$

$$\log(\pi_{2jk} / \pi_{0jk}) = \beta_{1k}\text{cons.sr_2}_{ijk} + \text{-0.025(0.003)age.sr_2}_{ijk} + \text{0.638(0.209)gender_2.sr_2}_{ijk} +$$
$$\text{-1.676(0.159)style_4.sr_2}_{ijk} + \text{-2.272(0.159)style_6.sr_2}_{ijk} + \text{-3.702(0.295)style_8.sr_2}_{ijk} +$$
$$\text{-0.025(0.005)age.gender_2.sr_2}_{ijk} + \text{0.503(0.212)gender_2.style_4.sr_2}_{ijk} +$$
$$\text{0.808(0.202)gender_2.style_6.sr_2}_{ijk} + \text{0.682(0.375)gender_2.style_8.sr_2}_{ijk} +$$
$$\text{0.102(0.101)tone_2.sr_2}_{ijk} + \text{-0.184(0.170)tone_3.sr_2}_{ijk} + \text{0.313(0.076)tone_4.sr_2}_{ijk} +$$
$$\text{0.026(0.172)tone_5.sr_2}_{ijk} + \text{0.130(0.108)tone_6.sr_2}_{ijk} + \text{3.080(0.143)fi_new_1.sr_2}_{ijk} +$$
$$\text{1.811(0.143)fi_new_2.sr_2}_{ijk} + \text{2.954(0.177)fi_new_3.sr_2}_{ijk} + \text{1.905(0.142)fi_new_4.sr_2}_{ijk} +$$
$$\text{1.421(0.167)aa.sr_2}_{ijk} + \text{1.134(0.175)ee.sr_2}_{ijk} + \text{-1.614(0.194)i.sr_2}_{ijk} + \text{1.317(0.151)oo.sr_2}_{ijk} +$$
$$\text{1.544(0.158)oe.sr_2}_{ijk} + \text{-0.767(0.175)u.sr_2}_{ijk} + \text{0.526(0.174)(logltfreq-gm).sr_2}_{ijk} +$$
$$\text{0.060(0.142)fi_new_1.(logltfreq-gm).sr_2}_{ijk} + \text{-0.354(0.149)fi_new_2.(logltfreq-gm).sr_2}_{ijk} +$$
$$\text{-0.382(0.199)fi_new_3.(logltfreq-gm).sr_2}_{ijk} + \text{-0.093(0.154)fi_new_4.(logltfreq-gm).sr_2}_{ijk} +$$
$$\text{0.012(0.139)aa.(logltfreq-gm).sr_2}_{ijk} + \text{-0.615(0.166)ee.(logltfreq-gm).sr_2}_{ijk} +$$
$$\text{-0.163(0.196)i.(logltfreq-gm).sr_2}_{ijk} + \text{-0.591(0.130)oo.(logltfreq-gm).sr_2}_{ijk} +$$
$$\text{-0.236(0.138)oe.(logltfreq-gm).sr_2}_{ijk} + \text{0.083(0.155)u.(logltfreq-gm).sr_2}_{ijk}$$

$$\beta_{1k} = \text{-1.625(0.223)} + v_{1k}$$

$$\begin{bmatrix} v_{0k} \\ v_{1k} \end{bmatrix} \sim N(0,\ \Omega_v)\ :\ \Omega_v = \begin{bmatrix} \text{1.736(0.159)} & \\ \text{0.133(0.072)} & \text{0.459(0.059)} \end{bmatrix}$$

$$\text{cov}(y_{sjk}, y_{rjk}) = -\pi_{sjk}\pi_{rjk}/cons_{jk} : s \neq r; \quad \pi_{sjk}(1 - \pi_{rjk})/cons_{jk} : s = r;$$

(16784 of 16784 cases in use)

Numerical values and calculations of Model 5

Fixed Part	Beta	S.E.	Beta/S.E.	p-value	Odds-ratio	%Change
cons.sr_1	-1.131	0.366	-3.09016	0.00200046	0.3227104	-67.729
cons.sr_2	-1.625	0.223	-7.287	3.1694E-13	0.1969117	-80.3088
age.sr_1	-0.042	0.006	-7	2.5596E-12	0.9588698	-4.11302
age.sr_2	-0.025	0.003	-8.33333	7.8597E-17	0.9753099	-2.46901
gender_2.sr_1	0.313	0.354	0.884181	0.37659863	1.3675215	36.75215
gender_2.sr_2	0.638	0.209	3.052632	0.00226844	1.8926917	89.26917
style_4.sr_1	-0.748	0.269	-2.78067	0.0054247	0.4733122	-52.6688
style_6.sr_1	-1.055	0.272	-3.87868	0.00010503	0.3481924	-65.1808
style_8.sr_1	-0.672	0.324	-2.07407	0.03807243	0.5106862	-48.9314
style_4.sr_2	-1.676	0.159	-10.5409	5.5972E-26	0.187121	-81.2879
style_6.sr_2	-2.272	0.159	-14.2893	2.5513E-46	0.1031058	-89.6894
style_8.sr_2	-3.702	0.295	-12.5492	4.0179E-36	0.0246741	-97.5326
age.gender_2.sr_1	-0.008	0.008	-1	0.31731051	0.9920319	-0.79681
age.gender_2.sr_2	-0.025	0.005	-5	5.733E-07	0.9753099	-2.46901
gender_2.style_4.sr_1	-0.12	0.368	-0.32609	0.74435856	0.8869204	-11.308
gender_2.style_6.sr_1	0.105	0.357	0.294118	0.76866801	1.1107106	11.07106
gender_2.style_8.sr_1	0.201	0.415	0.484337	0.62814647	1.2226248	22.26248
gender_2.style_4.sr_2	0.503	0.212	2.372642	0.0176614	1.6536749	65.36749
gender_2.style_6.sr_2	0.808	0.202	4	6.3342E-05	2.2434167	124.3417
gender_2.style_8.sr_2	0.682	0.375	1.818667	0.0689623	1.9778294	97.78294
tone_2.sr_1	0.024	0.121	0.198347	0.8427735	1.0242903	2.429032
tone_3.sr_1	-0.084	0.224	-0.375	0.70766047	0.9194313	-8.05687
tone_4.sr_1	-0.274	0.104	-2.63462	0.00842327	0.7603321	-23.9668
tone_5.sr_1	-0.157	0.197	-0.79695	0.42547756	0.8547041	-14.5296
tone_6.sr_1	0.311	0.142	2.190141	0.02851402	1.3647892	36.47892
tone_2.sr_2	0.102	0.101	1.009901	0.31254273	1.1073835	10.73835
tone_3.sr_2	-0.184	0.17	-1.08235	0.27909573	0.8319358	-16.8064
tone_4.sr_2	0.313	0.076	4.118421	3.8148E-05	1.3675215	36.75215
tone_5.sr_2	0.026	0.172	0.151163	0.8798473	1.0263409	2.634095
tone_6.sr_2	0.13	0.108	1.203704	0.22870412	1.1388284	13.88284
fi_new_1.sr_1	1.852	0.255	7.262745	3.7931E-13	6.3725519	537.2552
fi_new_2.sr_1	2.567	0.242	10.60744	2.7519E-26	13.026686	1202.669
fi_new_3.sr_1	3.371	0.284	11.86972	1.7004E-32	29.10762	2810.762
fi_new_4.sr_1	3.937	0.237	16.61181	5.7236E-62	51.264577	5026.458
fi_new_1.sr_2	3.08	0.143	21.53846	6.793E-103	21.758402	2075.84
fi_new_2.sr_2	1.811	0.143	12.66434	9.3206E-37	6.1165609	511.6561
fi_new_3.sr_2	2.954	0.177	16.68927	1.5689E-62	19.182531	1818.253
fi_new_4.sr_2	1.905	0.142	13.41549	4.9066E-41	6.7194076	571.9408
aa.sr_1	0.356	0.213	1.671362	0.0946503	1.4276075	42.76075

ee.sr_1	-0.858	0.269	-3.18959	0.00142474	0.4240093	-57.5991
i.sr_1	-4.351	0.354	-12.291	1.013E-34	0.0128939	-98.7106
oo.sr_1	-1.024	0.209	-4.89952	9.607E-07	0.3591554	-64.0845
oe.sr_1	-0.153	0.232	-0.65948	0.50958581	0.8581297	-14.187
u.sr_1	-5.253	0.457	-11.4945	1.4054E-30	0.0052318	-99.4768
aa.sr_2	1.421	0.167	8.508982	1.7547E-17	4.1412596	314.126
ee.sr_2	1.134	0.175	6.48	9.1723E-11	3.1080639	210.8064
i.sr_2	-1.614	0.194	-8.31959	8.827E-17	0.1990897	-80.091
oo.sr_2	1.317	0.151	8.721854	2.7368E-18	3.7322079	273.2208
oe.sr_2	1.544	0.158	9.772152	1.4827E-22	4.683286	368.3286
u.sr_2	-0.767	0.175	-4.38286	1.1713E-05	0.4644042	-53.5596
(logltfreq-gm).sr_1	-0.306	0.264	-1.15909	0.24641913	0.7363866	-26.3613
(logltfreq-gm).sr_2	0.526	0.174	3.022989	0.00250292	1.6921502	69.21502
fi_new_1.(logltfreq-gm).sr_1	0.189	0.262	0.721374	0.47067941	1.208041	20.8041
fi_new_2.(logltfreq-gm).sr_1	0.399	0.257	1.552529	0.12053566	1.4903336	49.03336
fi_new_3.(logltfreq-gm).sr_1	0.454	0.311	1.459807	0.1443431	1.574598	57.4598
fi_new_4.(logltfreq-gm).sr_1	0.968	0.256	3.78125	0.00015604	2.6326738	163.2674
fi_new_1.(logltfreq-gm).sr_2	0.06	0.142	0.422535	0.67263441	1.0618365	6.183655
fi_new_2.(logltfreq-gm).sr_2	-0.354	0.149	-2.37584	0.0175091	0.701875	-29.8125
fi_new_3.(logltfreq-gm).sr_2	-0.382	0.199	-1.9196	0.0549087	0.6824951	-31.7505
fi_new_4.(logltfreq-gm).sr_2	-0.093	0.154	-0.6039	0.54591272	0.9111935	-8.88065
aa.(logltfreq-gm).sr_1	0.2	0.126	1.587302	0.11244436	1.2214028	22.14028
ee.(logltfreq-gm).sr_1	-0.651	0.22	-2.95909	0.00308548	0.521524	-47.8476
i.(logltfreq-gm).sr_1	-0.884	0.298	-2.96644	0.00301266	0.4131271	-58.6873
oo.(logltfreq-gm).sr_1	-0.314	0.125	-2.512	0.01200491	0.730519	-26.9481
oe.(logltfreq-gm).sr_1	-0.284	0.142	-2	0.04550026	0.7527666	-24.7233
u.(logltfreq-gm).sr_1	-0.061	0.502	-0.12151	0.90328397	0.9408232	-5.91768
aa.(logltfreq-gm).sr_2	0.012	0.139	0.086331	0.93120335	1.0120723	1.207229
ee.(logltfreq-gm).sr_2	-0.615	0.166	-3.70482	0.00021154	0.5406409	-45.9359
i.(logltfreq-gm).sr_2	-0.163	0.196	-0.83163	0.40561633	0.8495912	-15.0409
oo.(logltfreq-gm).sr_2	-0.591	0.13	-4.54615	5.4635E-06	0.5537732	-44.6227
oe.(logltfreq-gm).sr_2	-0.236	0.138	-1.71014	0.08723908	0.7897807	-21.0219
u.(logltfreq-gm).sr_2	0.083	0.155	0.535484	0.59231531	1.0865418	8.654181

Model 2. "Best" mixed-effects multinomial logistic regression model for response variable "(k) realization" obtained in MLwiN

$resp_{ij} \sim Multinomial(cons_j, \pi_{ij})$

$\log(\pi_{1j} / \pi_{0j}) = -1.854(0.411)cons.sr_1_{ij} + -0.047(0.004)age.sr_1_{ij} + 2.622(0.234)gender_2.sr_1_{ij} +$
$-0.871(0.248)style_4.sr_1_{ij} + -1.714(0.256)style_6.sr_1_{ij} + -0.919(0.226)style_8.sr_1_{ij} +$
$-0.045(0.007)age.gender_2.sr_1_{ij} + -0.041(0.426)gender_2.style_4.sr_1_{ij} +$
$0.785(0.409)gender_2.style_6.sr_1_{ij} + -3.255(0.419)gender_2.style_8.sr_1_{ij} +$
$0.004(0.011)age.gender_2.style_4.sr_1_{ij} + -0.011(0.011)age.gender_2.style_6.sr_1_{ij} +$
$0.080(0.010)age.gender_2.style_8.sr_1_{ij} + -0.398(0.237)ee.sr_1_{ij} + -7.596(1.252)u.sr_1_{ij} +$
$-7.706(0.782)i.sr_1_{ij} + 0.013(0.121)aa.sr_1_{ij} + -1.300(0.113)oo.sr_1_{ij} + -0.761(0.239)oe.sr_1_{ij} +$
$2.963(0.397)fi_new_1.sr_1_{ij} + 3.756(0.389)fi_new_2.sr_1_{ij} + 5.915(0.443)fi_new_3.sr_1_{ij} +$
$5.301(0.386)fi_new_4.sr_1_{ij} + -1.311(0.612)(logltfreq\text{-}gm).sr_1_{ij} +$
$2.413(0.627)(logltfreq\text{-}gm).fi_new_1.sr_1_{ij} + 1.363(0.615)(logltfreq\text{-}gm).fi_new_2.sr_1_{ij} +$
$2.232(0.656)(logltfreq\text{-}gm).fi_new_3.sr_1_{ij} + 1.857(0.619)(logltfreq\text{-}gm).fi_new_4.sr_1_{ij}$

$\log(\pi_{2j} / \pi_{0j}) = -1.206(0.207)cons.sr_2_{ij} + 0.001(0.003)age.sr_2_{ij} + 1.770(0.208)gender_2.sr_2_{ij} +$
$-1.597(0.161)style_4.sr_2_{ij} + -1.549(0.143)style_6.sr_2_{ij} + -2.687(0.170)style_8.sr_2_{ij} +$
$-0.034(0.005)age.gender_2.sr_2_{ij} + 0.853(0.338)gender_2.style_4.sr_2_{ij} +$
$0.675(0.282)gender_2.style_6.sr_2_{ij} + -0.478(0.337)gender_2.style_8.sr_2_{ij} +$
$-0.036(0.008)age.gender_2.style_4.sr_2_{ij} + -0.028(0.007)age.gender_2.style_6.sr_2_{ij} +$
$-0.002(0.008)age.gender_2.style_8.sr_2_{ij} + 0.682(0.184)ee.sr_2_{ij} + -1.254(0.118)u.sr_2_{ij} +$
$-2.094(0.117)i.sr_2_{ij} + 0.924(0.106)aa.sr_2_{ij} + 0.554(0.100)oo.sr_2_{ij} + 0.465(0.192)oe.sr_2_{ij} +$
$3.952(0.185)fi_new_1.sr_2_{ij} + 3.853(0.177)fi_new_2.sr_2_{ij} + 3.750(0.271)fi_new_3.sr_2_{ij} +$
$2.434(0.176)fi_new_4.sr_2_{ij} + 0.492(0.162)(logltfreq\text{-}gm).sr_2_{ij} +$
$-0.389(0.181)(logltfreq\text{-}gm).fi_new_1.sr_2_{ij} + -0.360(0.168)(logltfreq\text{-}gm).fi_new_2.sr_2_{ij} +$
$0.870(0.285)(logltfreq\text{-}gm).fi_new_3.sr_2_{ij} + 0.188(0.187)(logltfreq\text{-}gm).fi_new_4.sr_2_{ij}$

$cov(y_{sj}, y_{rj}) = -\pi_{sj}\pi_{rj}/cons_j : s \neq r; \quad \pi_{sj}(1 - \pi_{rj})/cons_j : s = r;$
(11064 of 11064 cases in use)

Numerical values and calculations of Model 6

Fixed Part	Beta	S.E.	Beta/S.E	p-value	odds-ratio	%change
cons.sr_1	-1.854	0.411	-4.51094891	6.4538E-06	0.1566095	-84.33905
cons.sr_2	-1.206	0.207	-5.82608696	5.6742E-09	0.2993925	-70.06075
age.sr_1	-0.047	0.004	-11.75	7.0619E-32	0.9540874	-4.59126
age.sr_2	0.001	0.003	0.333333333	0.73888268	1.0010005	0.10005
gender_2.sr_1	2.622	0.234	11.20512821	3.848E-29	13.763223	1276.322
gender_2.sr_2	1.77	0.208	8.509615385	1.7451E-17	5.8708534	487.0853
style_4.sr_1	-0.871	0.248	-3.51209677	0.00044459	0.4185328	-58.14672
style_6.sr_1	-1.714	0.256	-6.6953125	2.1521E-11	0.1801438	-81.98562
style_8.sr_1	-0.919	0.226	-4.06637168	4.7751E-05	0.3989178	-60.10822
style_4.sr_2	-1.597	0.161	-9.91925466	3.4331E-23	0.2025031	-79.74969

style_6.sr_2	-1.549	0.143	-10.8321678	2.4235E-27	0.2124603	-78.75397
style_8.sr_2	-2.687	0.17	-15.8058824	2.8339E-56	0.0680849	-93.19151
age.gender_2.sr_1	-0.045	0.007	-6.42857143	1.2881E-10	0.9559975	-4.400252
age.gender_2.sr_2	-0.034	0.005	-6.8	1.0462E-11	0.9665715	-3.34285
gender_2.style_4.sr_1	-0.041	0.426	-0.09624413	0.92332668	0.9598291	-4.017087
gender_2.style_6.sr_1	0.785	0.409	1.919315403	0.05494443	2.1924069	119.2407
gender_2.style_8.sr_1	-3.255	0.419	-7.76849642	7.9423E-15	0.0385808	-96.14192
gender_2.style_4.sr_2	0.853	0.338	2.523668639	0.01161373	2.3466763	134.6676
gender_2.style_6.sr_2	0.675	0.282	2.393617021	0.01668316	1.964033	96.4033
gender_2.style_8.sr_2	-0.478	0.337	-1.41839763	0.15607471	0.6200222	-37.99778
age.gender_2.style_4.sr_1	0.004	0.011	0.363636364	0.71612957	1.004008	0.400801
age.gender_2.style_6.sr_1	-0.011	0.011	-1	0.31731051	0.9890603	-1.093972
age.gender_2.style_8.sr_1	0.08	0.01	8	1.2442E-15	1.0832871	8.328707
age.gender_2.style_4.sr_2	-0.036	0.008	-4.5	6.7953E-06	0.9646403	-3.535971
age.gender_2.style_6.sr_2	-0.028	0.007	-4	6.3342E-05	0.9723884	-2.761163
age.gender_2.style_8.sr_2	-0.002	0.008	-0.25	0.80258735	0.998002	-0.1998
ee.sr_1	-0.398	0.237	-1.67932489	0.09308874	0.671662	-32.8338
u.sr_1	-7.596	1.252	-6.06709265	1.3025E-09	0.0005025	-99.94975
i.sr_1	-7.706	0.782	-9.85421995	6.5725E-23	0.0004501	-99.95499
aa.sr_1	0.013	0.121	0.107438017	0.9144415	1.0130849	1.308487
oo.sr_1	-1.3	0.113	-11.5044248	1.2532E-30	0.2725318	-72.74682
oe.sr_1	-0.761	0.239	-3.18410042	0.00145205	0.467199	-53.2801
ee.sr_2	0.682	0.184	3.706521739	0.00021013	1.9778294	97.78294
u.sr_2	-1.254	0.118	-10.6271186	2.2289E-26	0.2853611	-71.46389
i.sr_2	-2.094	0.117	-17.8974359	1.2348E-71	0.1231934	-87.68066
aa.sr_2	0.924	0.106	8.716981132	2.8572E-18	2.5193477	151.9348
oo.sr_2	0.554	0.1	5.54	3.0247E-08	1.7401999	74.01999
oe.sr_2	0.465	0.192	2.421875	0.01544066	1.5920142	59.20142
fi_new_1.sr_1	2.963	0.397	7.463476071	8.4269E-14	19.355953	1835.595
fi_new_2.sr_1	3.756	0.389	9.655526992	4.6576E-22	42.776975	4177.698
fi_new_3.sr_1	5.915	0.443	13.35214447	1.1509E-40	370.5543	36955.43
fi_new_4.sr_1	5.301	0.386	13.73316062	6.4268E-43	200.53725	19953.72
fi_new_1.sr_2	3.952	0.185	21.36216216	3.006E-101	52.039342	5103.934
fi_new_2.sr_2	3.853	0.177	21.76836158	4.629E-105	47.134254	4613.425
fi_new_3.sr_2	3.75	0.271	13.83763838	1.5109E-43	42.521082	4152.108
fi_new_4.sr_2	2.434	0.176	13.82954545	1.6909E-43	11.404409	1040.441
(logltfreq-gm).sr_1	-1.311	0.612	-2.14215686	0.03218086	0.2695504	-73.04496
(logltfreq-gm).sr_2	0.492	0.162	3.037037037	0.00238916	1.6355841	63.55841
(logltfreq-gm).fi_new_1.sr_1	2.413	0.627	3.848484848	0.00011885	11.167413	1016.741
(logltfreq-gm).fi_new_2.sr_1	1.363	0.615	2.216260163	0.02667369	3.9078994	290.7899
(logltfreq-gm).fi_new_3.sr_1	2.232	0.656	3.402439024	0.00066787	9.3184844	831.8484
(logltfreq-gm).fi_new_4.sr_1	1.857	0.619	3	0.0026998	6.4044944	540.4494

(logltfreq-gm).fi_new_1.sr_2	-0.389	0.181	-2.14917127	0.03162083	0.6777343	-32.22657
(logltfreq-gm).fi_new_2.sr_2	-0.36	0.168	-2.14285714	0.03212457	0.6976763	-30.23237
(logltfreq-gm).fi_new_3.sr_2	0.87	0.285	3.052631579	0.00226844	2.3869109	138.6911
(logltfreq-gm).fi_new_4.sr_2	0.188	0.187	1.005347594	0.31472951	1.2068335	20.68335

End of Document