SYLLABLE FUSION IN HONG KONG CANTONESE CONNECTED SPEECH

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Wai Yi Peggy Wong, B.A., M.A.

*****

The Ohio State University
2006

Dissertation Committee:
Professor Mary E Beckman, Adviser
Professor Marjorie K M Chan, Co-Adviser
Professor Chris Brew

_____
Adviser

_____
Co-Adviser
Linguistics Graduate Program

ABSTRACT

This dissertation is about “syllable fusion” in Hong Kong Cantonese. Syllable fusion is a connected-speech phenomenon whereby boundaries between syllables are blurred together in a way that suggests an intermediate level of grouping between the syllable and the larger intonational phrase. Previous studies of this phenomenon have focused on extreme cases — i.e. whole segments (consonants and/or vowels) are deleted at the relevant syllable boundary. By contrast, in this dissertation, “syllable fusion” refers to a variety of changes affecting a sequence of two syllables that range along a continuum from “mild” to “extreme” blending together of the syllables. Less extreme changes include assimilation, consonant lenition and so on, any substantial weakening or effective deletion of the oral gesture(s) of the segment(s) contiguous to the syllable boundary, and the sometimes attendant resyllabifications that create “fused forms”. More extreme fusion can simplify contour tones and “merge” the qualities of vowels that would be separated by an onset or coda consonant at more “normal” degrees of disjuncture between words.

The idea that motivates the experiments described in this dissertation is that the occurrence of syllable fusion marks prosodic grouping at the level of the "foot", a phonological constituent which has been proposed to account for prosodic phenomena such as the process of tone sandhi and neutral tone in other varieties of Chinese.

If syllable fusion marks the grouping together of syllables into feet, there must be factors similar to the factors that influence the occurrence of tone sandhi and neutral tone in other varieties of Chinese that influence the occurrence and the degree of syllable fusion. Discovering these factors would allow us to understand the "foot" structure for Cantonese. Five factors were identified and tested in this dissertation: speech rate, word frequency, word length, morphosyntactic relationship, and prosodic position of words. Three experiments were run to test the five factors.

Results showed main effects for all five factors and interaction effects among them. Unifying these factors, this dissertation proposed that the "foot" be used as an intermediate level of constituent in the Cantonese prosodic hierarchy that captures the phenomenon of syllable fusion in Cantonese.

ACKNOWLEDGMENTS

I always think of a piece of research work as an orchestra — many people offered various kinds of help at various times that shaped the final piece. This dissertation is no exception. I would not have been able to bring this dissertation to a completion all by myself without the help of many, many people, who lovingly offered their prayers, advice, support, encouragement, and actual help. Not only did they offer help, they offered help at the right timing, which could mean inconveniencing their own schedules and sacrificing their own preferences. These acts of love have added much joy and peace to the long process of dissertating. This orchestra will resound in my heart.

My deepest gratitude goes to my adviser, Mary Beckman, who has been a great inspiration in many aspects in this dissertation and a great mentor personally. From the general direction to the finer details of the dissertation, she has given to me a lot of insightful guidance and help. I greatly appreciate her efficiency and patience in reading and correcting my drafts. It was through her corrections that I learned a great deal about writing. Mary always insists on giving me her feedback on time so that I could revise my draft accordingly, even when she was swarmed with work and under time pressure, and

even that meant many hours of sleep needed to be sacrificed. She also goes out of her way to deliberately make an extra trip to come back to Columbus from Georgia for my dissertation defense, even this is her off-quarter. Thanks so much again for all you did, Mary!

I am also very grateful to Margie Chan, my co-adviser, and Chris Brew, for their generous support and patience, especially in the final stage of dissertating. Their "OK's" had become such a comforting word during the times when I was working hard on giving them my dissertation draft, but kept needing to delay. Margie and Chris, thanks again for being so understanding and patient in waiting for my draft, but so efficient in reading it! Margie, thanks also for the various get-togethers at your place, which will become memory of my good old days in Columbus!

I would also like to thank the following people and institutions regarding the use of corpora in this dissertation. Drs. Sam Po Law and Man Tak Leung in the Department of Speech and Hearing Sciences at the University of Hong Kong let me use the Hong Kong Cantonese Adult Language Corpus (HKCAC) for this dissertation. Dr. Shui Duen Chan in the Department of Chinese and Bilingual Studies at the Hong Kong Polytechnic University let me use the Segmentation Corpus and the algorithm and program that did the word segmentation. Mr. Him Man Hui at the Chinese Language Centre at the Hong Kong Polytechnic University gave me instructions about how to use the segmentation program. (The development of the Segmentation Corpus was supported by a grant from

the University Grants Committee of Hong Kong to Prof. Y.-S. Cheung, Department of Chinese and Bilingual Studies, Hong Kong Polytechnic University. The corpus is maintained by Dr. S.-D. Chan.)

My thanks also go to the Graduate School of the Ohio State University. This dissertation was partly supported by an Alumni Grant for Graduate Research and Scholarship (AGGRS) that was granted to me in 2004.

The tasks of reformatting, word segmentation and tagging of the HKCAC, morphosyntactic analysis of the candidate test words, and the actual running of Experiment III all needed to be done one after another (i.e. the sequence of tasks could not be changed), and they needed to be done during the two and a half months that I was in Hong Kong. Thus, the completion of the tasks was under time pressure. Thankfully, help was given one after another at just the right time. I am deeply grateful to all the following people.

After reformatting the HKCAC, I was in need of a person who could quickly write a program to map the segmented orthographic transcription onto the phonetic transcription of the HKCAC. I thank Mr. Dicky Chan for his proficiency and efficiency in getting the work done!

Mr. Samuel P. M. Cheung, Mr. Him Man Hui, and Dr. Alan Yu were kind and brave enough to agree to do the task of morphosyntactic annotation for me, which involved more than 7,000 words! Thank you all so much!

Dr. Cathy Wong let me use the Speech Laboratory in the English Department at the Hong Kong Polytechnic University to run subjects for Experiment III in this dissertation. Ms. Cherry Ng was very helpful in many technical aspects concerning the lab equipment and setting. Many brothers and sisters at the Joshua Fellowship at the Columbus Chinese Christian Church helped me find subjects to participate in my experiment. I would like to thank them all and thank all of my participants who responded so promptly. Just within a few days, about 29 subjects were lined up and I could finish running them all in just about one week! (The next day after finished running the experiment was the day when I had to head back to Columbus.)

I also thank my adviser, Mary Beckman, and my fellow graduate student Kyuchul Yoon for their help on writing of Praat scripts, which speeded up the process of extracting Praat textgrids and the subsequent analysis of data.

In terms of doing statistical analysis, I am thankful to Ms. Zhenhuan Cui, a Ph.D. candidate in the Statistics Department at the Ohio State University and my old neighbor when I was living in the dorm. When I was looking for help, her name just popped up in my head. And it so happened that her research focuses on logistic regression. She had given me valuable advice for the statistical analysis in this dissertation.

I would also like to thank my team-mates in the paidologos (pdlg) project, including, Mary Beckman, Jan Edwards, Junko Davis, Eunjong Kong, Fangfang Li, Laura Slocum

and Asimina Syrika. I thank Mary and Jan for being flexible and considerate in letting me focus on finishing the dissertation first and then make up the work for the project after this dissertation was completed. I also thank the help that my team-mates generously offered to lighten my workload for the project so that I could be more focused on working on this dissertation.

There are many others who were always behind the scene offering invisible but indispensable help. I am thankful for the continuous prayers and the words of encouragement of all those who have been praying for me. They include my fellow grads and alumni in the Linguistics Department: Wes Collins, Hope Dawson, Markus Dickinson, Janice Fon, Anne Hoffman, Crystal Nakatsu, Anton Rytting, Elizabeth Smith, and Lei Xu; all my brothers and sisters at the Joshua Fellowship, the Cantonese Family Fellowship and the Foundation College Fellowship at the Columbus Chinese Christian Church; Pastor J. Chun and his family and many others in the Korean Mission of the Lane Avenue Baptist Church; brothers and sisters at the Boston Chinese Evangelical Church; brothers and sisters at the Joshua Fellowship at the Tai Po Baptist Church; my group-mates and leaders at Bible Study Fellowship (BSF) in Columbus; my colleagues Peggy Mok in the U.K., Zoe Lam at the Chinese University of Hong Kong, K. H. Cheung, Roxana Fung, and Mimi Wu, among others, at the Polytechnic University of Hong Kong; my friend Cherry Cheung and her family. I have to cut short in listing all your names here, but know that your love and care are always in my heart!

My deep gratitude must go to my parents and my sister, Winnie, who have been extremely supportive that no words can describe throughout all these years of my studying at OSU. I would also like to thank my parents-in-law who have been so supportive. I also thank my brother- and my sister-in-law, Peter and Grace Lok, for their hospitality and their generous support. It was enjoyable spending my last half year at OSU living with them, while dissertating. In the house is also their daughter, my niece, baby Chung Yan Lok (樂頌恩), who has always been a source of joy around us!

Last but not least, I am deeply thankful to my beloved husband, John Lok, who has been extremely supportive throughout the process of dissertating. He discussed ideas with me about this dissertation, even though this is not his field. He cooked, even when he, too, was busy preparing for his ministry. Our praying together has been the spiritual support. His persistent reminder of a higher purpose behind the dissertation work has been a source of motivation and of joy and peace even under the pressure of dissertating.

Through all the acts of love given to me during this dissertating process, I see the abundant provision and blessings from God. He has always been the Conductor of this orchestra — the orchestra that composed of different helpers with different gifts playing their roles at just the right time. Through working on this dissertation, I can tell the following saying is true: “And we know that in all things God works for the good of those who love him, who have been called according to his purpose.” (Romans 8:28) I am forever thankful to my Lord.

VITA

December 8, 1971 ……..…. Born – Hong Kong

June, 1993…… B.A.(Hons.), Translation and Chinese, Hong Kong Polytechnic University.

1993 – 1996 …….Research Assistant, Hong Kong Polytechnic University.

June, 1996…… M.A., Chinese Linguistics, Hong Kong Polytechnic University.

August, 1996……...……...…. Certificate, English Phonetics, University College London, U.K.

1996 – 1998 …… Instructor, Department of Chinese and Bilingual Studies, Hong Kong Polytechnic University.

Summer, 1999 …….… Cantonese Language Consultant, Nuance Communications.

Autumn, 2003 …….…… Volunteer Instructor for the literacy mission, Columbus Chinese Christian Church, Ohio.

2005 …….……... Volunteer Simultaneous Interpreter, Columbus Chinese Christian Church, Ohio.

1998 – present …….… Graduate Teaching and Research Associate, Ohio State University.

PUBLICATIONS

1. **Wong, W. Y. P.**, Chan, M. K.-M. and Beckman, M. E. (2005). 'An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese'. In: *Prosodic Typology - The Phonology of Intonation and Phrasing.* (Collection of papers from the 14th International Congress of Phonetic Sciences 1999 Satellite Workshop on "Intonation: Models and ToBI Labelling", San Francisco, CA. 1-7 August 1999.) Edited by Sun-Ah Jun. Oxford University Press.

2. Tserdanelis, G. and **Wong, W. Y. P.** (eds.) (2004). *Language Files 9: Materials for an Introduction to Language and Linguistics*. Ohio State University Press.

3. **Wong, W. Y. P.** (2004). 'Syllable fusion and speech rate in Hong Kong Cantonese'. In: *Proceedings of Speech Prosody 2004*, 255-258. 22-26 March, Nara, Japan.

4. **Wong, W. Y. P.**, Brew, C., Chan S.-D. and Beckman, M. E. (2002). 'Using the Segmentation Corpus to define an inventory of concatenative units for Cantonese speech synthesis'. In: *Proceedings of the First SIGHAN Workshop on Chinese Language Processing in conjunction with COLING 2002*. (Papers presented at the First SIGHAN Workshop on Chinese Language Processing, Academia Sinica, Taipei, Taiwan, 1 September 2002.) Edited by T'sou, B. K., Kwong, O. O. Y. and Lai, T. B. Y.

5. Fung, R. S.Y., Law, S. P., Yu, M. Y. Y. and **Wong, W. Y. P.** (1999). 'Virtual tutorials in phonology: A new approach to raise phonological awareness among Hong Kong students'. In: *Proceedings of First Conference on Quality in Teaching and Learning in Higher Education*, 154-159.

FIELDS OF STUDY

Major Field: Linguistics

Studies in:

Phonetics

Speech prosody

TABLE OF CONTENTS

LIST OF TABLES

Appendix Tables Pages

LIST OF FIGURES

Appendix Figures Pages

LIST OF ABBREVIATIONS

Adj	Adjective
AdjP	Adjectival phrase
Adv	Adverb
AdvP	Adverbial phrase
Aux	Auxiliary verb
CL	Classifier
Conj	Conjunction
Dem	Demonstrative
EXM	Exclamation
F	(Prosodic) foot
Idiom	Idiom / Idiomatic expression
INJ	Interjection
IP	Intonational phrase
LOC	Localizer/Postposition
N	Noun

NEG	Negation
NP	Noun phrase
PN	Proper name
PP	Prepositional phrase
Prep	Preposition
ProN	Pronoun
QF	Quantifier
QM	Question marker
S	Sentence
SP	Sentence particle
Utt	Utterance
V	Verb
VP	Verb phrase
VPT	Verbal particle
σ	Syllable
|	Word boundary

CHAPTER 1

INTRODUCTION

This dissertation is about "syllable fusion" in everyday Hong Kong Cantonese[1], a phenomenon whereby syllables can be blurred together in connected speech. In this dissertation, "syllable fusion" refers to a variety of changes affecting a sequence of two syllables that range along a continuum from "mild" to "extreme" blending together of the syllables. Less extreme changes include assimilation, consonant lenition and so on, any substantial weakening or effective deletion of the oral gesture(s) of the segment(s) contiguous to the syllable boundary and the sometimes attendant resyllabifications that create "fused forms". More extreme fusion can simplify contour tones and "merge" the qualities of vowels that would be separated by an onset or coda consonant at more "normal" degrees of disjuncture between words. Some examples of syllable fusion in Hong Kong Cantonese are shown in Table 1.1 below.

[1] As in Wong, Chan, & Beckman (2005), I use the term "Cantonese" in three senses in this dissertation. The first and oldest sense is in reference to the variety of Chinese spoken in and around the city of Canton (Guangzhou), the provincial capital of Guangdong Province in southern China. Because many Canton City inhabitants have migrated to Hong Kong in the past 150 years or so (Hashimoto 1972:70), the variety of Cantonese spoken in Hong Kong is more similar to that of Canton than are varieties of the language spoken in towns closer to Canton City itself. Hence, socially prestigious "Standard Cantonese" can refer to the Cantonese spoken in Hong Kong as well as in Canton, even though over time a number of differences have emerged in the speech of these two localities (Bauer and Benedict 1997). Finally, the term "Cantonese" has further been extended to refer to the entire group of similar dialects that, together, form one of the major sub-classifications within Chinese, parallel to Mandarin, Wu, Min, and so forth. (One also encounters the term "Yue" for that usage — see, e.g., Hashimoto 1972; Yuan 1983; Grimes 1996.) Although syllable fusion probably occurs to some extent in all varieties of Cantonese in the final sense, the phenomenon seems particularly characteristic of Hong Kong Cantonese — that is, of "Standard Cantonese" or "Cantonese" in the second sense of the term; hence the focus here on syllable fusion in Hong Kong Cantonese.

a.	識唔(識) 'know NEG (know)' $sek^{5}.m^{21}$ → $seʔ^{5}.m^{21}$; $se^{5}.m^{21}$
b.	其實 'in fact' $k^{h}ei^{21}.sɐt^{2}$ → $k^{h}e^{21}.ɐt^{2}$; $k^{h}ɛt^{21+2}$
c.	知道 'know' $tsiː^{55}.tou^{33}$ → $tsiː^{55}.ou^{33}$; $tsiːu^{55+33}$
d.	朝頭(早) 'morning' $tsiːu^{55}.t^{h}ɐu^{21}$ (tsou^{35}) → $tsiː^{55}.ɐu^{21}$ (tsou^{35}) ; $tsiːu^{55+21}$ (tsou^{35})

Table 1.1 Syllable fusion forms resulting from mild or extreme changes.

The phenomenon of syllable fusion was noted at least as early as 1960 (Yuan et al. [1960] 1983), and there is a recent thesis by Lee (2003) giving a phonological account using the stress-based approach and the autosegmental-metrical framework to explain some of the commonly observed cases of more extreme forms of fusion. In her thesis, Lee addresses three questions: why, when and how two syllables are fused together.

In answering the first question of why syllable fusion occurs, Lee assumes that there are underlying stress patterns for the syllables in the Cantonese words and phrases, which are apart of the implicit phonological knowledge of native Cantonese speakers. She proposes that it is the underlying stress patterns that motivate syllable fusion. She also posits that these stress patterns are determined by the syntax internal to a compound word or phrase. Specifically, it is those morphemes that are syntactic non-heads (as defined in Chomsky (1981) and Pollock (1989)) of their compounds or phrases that receive stress. Nonhead stress rule then governs which morphemes can be reduced or merged with another syllable and so give rise to syllable fusion in connected speech.

In order to answer the second question of when does syllable fusion occur, Lee calculated the likelihood of deletion of segments based on the corpus she used. That is, she assumed that syllable fusion is governed also by a phonological process of syllable formation, whereby segments are inserted to fill prosodic positions (C and V slots), with variable rule-like constraints on what segments can go where. In unstressed syllables in connected speech, some of these positions can be deleted, leaving no place for some of underlying segments to be realized.

Finally, to model how syllable fusion occurs, Lee uses an autosegmental framework, proposing that associations would start from both ends of the two syllables to account for the process of how two syllables are fused into one.

Lee's study documented some of the morphosyntactic and segmental factors that are involved in more extreme degrees of fusion in which whole segments adjacent to the syllable boundary can be deleted. However, there is no systematic study of the full range of changes that can occur and of all the factors that might affect the likelihood and the degree of fusion. The purpose of this dissertation is to fill that gap. This dissertation focuses on the *when* and *what* questions — i.e. when do syllable fusion occur and what does syllable fusion inform us about Cantonese speech prosody. It investigates the phenomenon of syllable fusion in Hong Kong Cantonese connected speech at the segmental level. More specifically, it aims to discover the main factors that predict syllable fusion and to propose a convenient phonological representation of any overarching principle that unifies those factors. The idea connecting these two aims is that syllable fusion is a prosodically driven process that might reflect an intermediate level of grouping in the prosodic hierarchy of Cantonese in between the syllable and the intonational phrase. The “foot”, a unit that has been invoked to account for connected speech phenomena observed in other varieties of Chinese, seems a good candidate for the level of the prosodic hierarchy that is involved.

An example of a well-studied foot-level effect is the phenomenon of "third tone sandhi" (hereafter, T3 sandhi) in Beijing Mandarin and in standard Putonghua. The rule of T3 sandhi states that in a sequence of syllables lexically specified as T3, all of the tones except the last are changed into the second tone (T2). Some examples are given in Table 1.2. The T3 sandhi domain may coincide with word boundaries as in Table 1.2a, or it may apply across word and phrasal boundaries as in Table 1.2b. If it can occur within *and* between words, what is the domain for T3 sandhi? Shih (1997) proposes that the minimum domain where T3 sandhi rule application is obligatory in Mandarin is the foot (i.e. the prosodic foot). Thus, for example, /ɕiɑu**3**tçiɛ**3**/ 'young lady' and /tʂan**3**lan**3**/ 'to exhibit' surface as [ɕiɑu**2**tçiɛ**3**] and [tʂan**2**lan**3**] because each constitutes one disyllabic foot.

a. Within word.				
	(i)	/çiɑu**3**tçiɛ**3**/	→ [çiɑu**2**tçiɛ**3**]	'young lady'
	(ii)	/tʂan**3**lan**3**/	→ [tʂan**2**lan**3**]	'to exhibit'
	(iii)	 /lɑu**3**xu**3**tan**3**/ → [lɑu**2**xu**2**tan**3**]		
		tiger gall 'bravery'		
b. Across words.				
	(i)	/kei**3** wo**3** ʂuei**3**/ → [kei**2** wo**2** ʂuei**3**]		
		give me water	'give me water'	
	(ii)	/lɑu**3**li**3** mai**3** xɑu**3** tsiou**3**/		
		Lao Li buys good wine	'Lao Li buys good wine'	
		→ Possible outputs :		
		[lɑu**2**li**3** mai**3** xɑu**2** tsiou**3**]	slow or	low
		[lɑu**2**li**2** mai**3** xɑu**2** tsiou**3**]	speech rate	word freq.
		[lɑu**2**li**3** mai**2** xɑu**2** tsiou**3**]		
		[lɑu**2**li**2** mai**2** xɑu**2** tsiou**3**]	fast	high

Table 1.2 T3 sandhi.

The foot, as Shih suggests, is sensitive to the morphosyntactic branching structure of the string of morpheme-syllables. For example, the compound word 'bravery' given in Table 1.2a(iii) is left branching. T3 sandhi must first apply to the first T3 syllable /lɑu/ as it groups with the following T3 syllable /xu/ to form the word 'tiger' at the lower level of the hierarchy of the compound word. In this compound word /lɑu3xu3tan3/, T3 sandhi

never applies to the second syllable /xu3/ or the third syllable /tan3/ before applying to the first syllable /lɑu3/, as there is the word-internal boundary between /xu3/ 'tiger' and /tan3/ 'gall'.

Moreover, fast speech rate or high frequency of use of expression also affects the foot structure at a higher level. For example, in the extreme case that is the last output in Table 1.2b(ii), in which all T3 syllables preceding the last one became T2, we can think of the entire 5-syllable string as being grouped into one. The two continua to the right represent the fact that such a long prosodic foot is only possible at fast speech rate or when the word is of high frequency. In more typical cases, the foot is just two or three syllables.

In this dissertation, five factors were tested as potentially contributing to the likelihood of syllable fusion in present-day Hong Kong Cantonese connected speech. The first factor was ***speech rate***. Two tempo instructions were used to elicit a variety of rates across the speakers tested. The prediction was that fusion between two syllables would be more likely and the degree of fusion more extreme at fast rate.

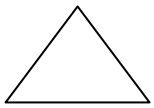
The second factor was ***prosodic position of participating syllables***, with three levels to be tested, namely, intonational-phrase-initial, medial and final positions. The predictions were that fusion between two syllables would be less likely and syllables would be less fused if they are the last two in their intonational phrase.

The third factor was ***word frequency***, with three levels to be tested, namely, high, mid and low word frequency. That is, the target words would be sorted by the log frequency and placed into one of these three frequency categories. The predictions were that some degree of syllable fusion would be more likely between two syllables that belong to the high word frequency category. Also, we predicted that the syllables would be fused into more extreme forms more often in high frequency words.

The fourth factor was ***word-internal morphosyntactic relationship***. Four types of morphosyntactic relationship were tested, namely, “twins”, “sisters”, “aunt-niece”, and “cousins”. Figure 1.1 shows these four types of morphosyntactic relationship.

Within a monomorphemic word, two syllables that are contiguous to each other are twins if the sequence cannot be identified as syllable-morphemes that occur alone as monosyllabic words or as recurring syllable-morphemes in other disyllabic or larger words (see Figure 1.1). That is, the string cannot be further decomposed and there is no word-internal boundary between two syllables that are twins.

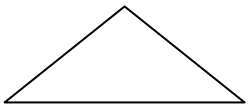
no branching:



先 twins 生

/siːn^{55}saːŋ^{55}/ ‘mister’

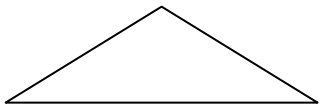
no branching:



鬱 twins 金 twins 香

/wɐt^{5}kɐm^{55}hœːŋ^{55}/ ‘tulip’

no branching:



卡 twins 拉 twins O twins K

/k^{h}aː^{55}laːi^{55}ou^{55}k^{h}ei^{55}/ ‘karaoke’

Figure 1.1 Examples of monomorphemic words. The syllables are “twins”. Relationships between syllables are given at the syllable boundaries.

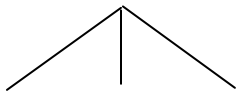
In polymorphemic words, two syllables that are contiguous to each other are sisters if each is a syllable-morpheme and they share the same parent node (see Figure 1.2). That is, there is a word-internal morpheme boundary between two syllables that are sisters to each other.

single-level branching:

靈 sisters 敏

/leŋ^{21}mɐn^{23}/ ‘nimble’

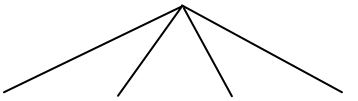
single-level branching:



面 sisters 對 sisters 面

/miːn^{22}tɵy^{33}miːn^{22}/ ‘face-to-face’

single-level branching:



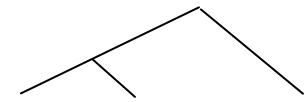
柴 sisters 米 sisters 油 sisters 鹽

/ts^{h}aːi^{21}mɐi^{23}jɐu^{21}jiːm^{21}/ ‘survival essentials’ (literally, ‘firewood, rice, oil, salt’)

Figure 1.2 Examples of polymorphemic words with a completely flat word-internal branching structures. The morpheme-syllables are “sisters”. Relationships between morpheme-syllables are given at the morpheme-syllable boundaries.

Two syllables that are contiguous to each other have an aunt-niece relationship when one of the morphemes is immediately dominated by a node that is higher than the immediate parent of the other one (see Figure 1.3). An aunt-niece relationship implies the existence of a branching structure within the word, e.g. left branching for the words ‘Internet’ and ‘robotic’ and right branching for the words ‘to call’ and ‘four big kings’ in the examples given in Figure 1.3. Thus, this relationship is impossible for two-syllable words.

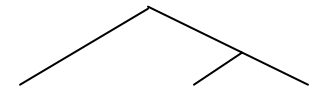
left-branching:



互 sisters 聯 aunt-niece 網

/wuː^{22}lyːn^{21}mɔːŋ^{23}/ ‘Internet’

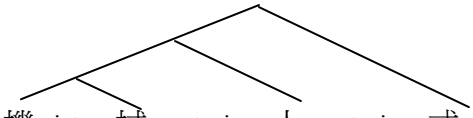
right-branching:



打 aunt-niece 電 sisters 話

/taː^{35}tiːn^{22}waː^{35}/ ‘to call’

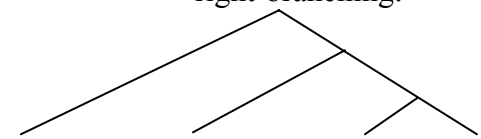
left-branching:



機 sisters 械 aunt-niece 人 aunt-niece 式

/kei^{55}haːi^{22}jɐn^{21}sek^{5}/ ‘robotic’

right-branching:



四 aunt-niece 大 aunt-niece 天 sisters 王

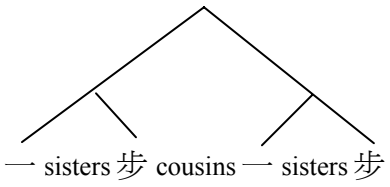
/sei^{33}taːi^{22}thiːn^{55}wɔːŋ^{21}/ ‘four big kings (referring to four top singers in Hong Kong in the 90’s)’

Figure 1.3 Examples of polymorphemic words with different word-internal branching structures. The morpheme-syllables are “sisters” or “aunts-nieces”. Relationships between morpheme-syllables are given at the morpheme-syllable boundaries.

Two syllables that are contiguous to each other are cousins when each of them is a twin or a sister to another syllable-morpheme at a lower level node in the morphosyntactic hierarchy within the word. Two-syllable sequences that were cousins in this dissertation were in words that had a balanced branching structure. The predictions for word-internal morphosyntactic relationship were that fusion would be more likely and

forms would be more fused between two syllables that were twins, as compared to those that were sisters, and sisters would be fused more often and forms would be more fused than those that were in aunt-niece relationship, and that aunts-nieces would be fused more often and forms would be more fused than those that that were cousins.

balanced branching:



/jɐt^{5}pou^{22}jɐt^{5}pou^{22}/ ‘step by step’

Figure 1.4 Examples of polymorphemic words with balanced word-internal branching structure. The morpheme-syllables are “sisters” or “cousins”. Relationships between morpheme-syllables are given at the morpheme-syllable boundaries.

The fifth factor tested was ***word length***. Three levels of word length were tested, namely, two, three, four syllables. The prediction was that two syllables within longer words would be more likely to be fused and would be more fused than two syllables within shorter words. This prediction was based on the idea that word-final position induces some degree of prosodic lengthening, albeit less than the “phrase-final lengthening” associated with the last syllables in an intonation phrase, and a smaller proportion of the syllables in longer words will be in word-final position.

Three experiments were run to test these predictions.

Experiment I asked the question: "Is speech rate a contributing factor for syllable fusion?" Shih (1997) (see above) suggests that fast speech in Mandarin affects T3 sandhi because the foot structure changes when tempo increases. For example, a foot might expand to incorporate syllables that belong together only at a higher level in the morphosyntactic hierarchy of a word. If syllable fusion, similarly, is an indication of the foot structure that a speaker assigns to a string of syllables, we would expect to see that fast speech rate would be conducive to more fused forms and to a higher degree of syllable fusion than normal speech rate. A "pseudo-sentence" repetition task was run to test the prediction. The "pseudo-sentences" were sentences that were made of three short or long words each. The concatenation of the words did not make the sentences meaningful. For each of the target words, fusion forms had been found to occur in everyday Hong Kong Cantonese connected speech. The productions of 32 native Hong Kong Cantonese participants were transcribed. Speech rate was defined as number of syllables per second. A regression analysis supported the prediction that speech rate and the occurrence of syllable fusion were highly correlated. That is, the faster the speech rate, the higher the occurrence of fused forms. Furthermore, the faster the speech rate, the more likely for the vowels of the two participating syllables to be coalesced.

Experiment II asked the question: "Are word frequency, word length and word-internal morphosyntactic relationship contributing factors for syllable fusion?" The transcriptions in a spoken corpus: the Hong Kong Cantonese Adult Language Corpus (HKCAC), were used to probe these factors. I segmented the corpus into words, and target words that covered a good range of values for the three factors were extracted. In the transcriptions by the HKCAC transcribers, five degrees of bisyllabic fusion, from not fused to two syllables completely fused into one, were identified. Multinomial logistic regression analysis was run to test which factor(s) and what combination of factors have an effect on syllable fusion.

For *word frequency*, we would expect that relatively more syllable boundaries would be blurred (i.e. the two syllables that straddle the boundary would be fused) for words with high frequency than for words with mid frequency, and relatively more boundaries blurred for words with mid frequency than words with low frequency.

For *word length*, we would predict that more boundaries would be blurred and blurred to a greater extent within quadrisyllabic words than within trisyllabic words, and more boundaries blurred and blurred to a greater extent within trisyllabic words relative to disyllabic words.

For *word-internal morphosyntactic relationship*, we would predict that word-internal morphosyntactic boundaries would discourage fusion for the two syllables that straddled them and would discourage a higher degree of fusion. Thus, for example, in the quadrisyllabic word /sɐu^{55}jɐm^{55}kei^{55}tsɐi^{35}/ 收音機仔 'small radio' (literally, 'receive sound gadget small'), which is left-branching [[[sɐu^{55}jɐm^{55}]kei^{55}]tsɐi^{35}], the first two syllables (which are sisters to each other) would be more likely to be fused or to be fused to a more extreme form than the second and the third syllables (which are in aunt-niece relationship). Also, the second and the third syllables would be just as likely to be fused and fused to a similar degree as the third and the fourth syllables (which are also in aunt-niece relationship). For syllables that are cousins, we predict that they would be the least likely to be fused. Results showed that word frequency and word-internal morphosyntactic relationship had a significant main effect on syllable fusion, but word length did not. No interaction among the factors was found.

Experiment III followed up on Experiment II. It took a closer look at the influence of word frequency, word length, word-internal morphosyntactic relationship, and also examined a fourth factor — prosodic position of the syllables within the larger prosodic hierarchy of words grouped into intonational phrases. This experiment asked the question: "Are word frequency, word length, morphosyntactic relationship internal to words, and

prosodic position contributing factors for syllable fusion?" A sentence repetition experiment was run using the same set of target words as in Experiment II. The targets were embedded in meaningful sentences so that all four potential factors could be controlled. The experimenter transcribed the prosodic positions of the target words produced by the native speakers using Cantonese ToBI. As in Experiment II, five degrees of bisyllabic fusion, from not fused to two syllables fused into one, were transcribed for all boundaries of all target words. Multinomial logistic regression analysis was run to test which factor(s) and what combination of factors have an effect on syllable fusion. If the factors are found to contribute to syllable fusion, predictions for the first three potential factors would be the same as those for Experiment II above. For *prosodic position*, we predict that with the phenomenon of final-lengthening in Cantonese, intonational phrase-final position might have relatively fewer instances of bisyllabic fusion and a lesser extent of fusion, while the intonational phrase-initial and -medial positions might have relatively more occurrences of fusion and a greater extent of fusion. Results showed a significant main effect for all four factors. Two-term interactions were found for word length with each of the three other factors, and for word frequency with morphosyntactic relationship.

The three experiments above and their results, as well as the corpus that provided the materials for Experiments II and III are described in more detail in chapters 3 to 6 of this dissertation. These four chapters are preceded by a chapter that reviews the prior literature on syllable fusion in Hong Kong Cantonese and discusses phenomena in other Chinese languages that have been described as manifestations of foot level structure. A summary of the results and conclusions are given in chapter 7.

CHAPTER 2

SYLLABLE FUSION IN THE CANTONESE PROSODIC HIERARCHY

If the prosodic system of a language can be thought of as a device for structuring the phonology in ways that function in part to suggest the morphosyntactic structure of an utterance, then the foot in Chinese prosodic systems corresponds to something like the "word" — a morphosyntactic constituent that is intermediate between the smallest "morphemes" (which are typically monosyllabic in Chinese) and the "clause" (which is typically an intonational phrase).

To discover how the prosodic system is structured in a language, we can look for recurrent and predictable landmarks. These landmarks should be reliable in the categorical marking of a certain prosodic entity.

The leading idea in this dissertation is the hypothesis that there is a single process, "syllable fusion", in Cantonese that covers both the nonce minor "coarticulation"-like processes of assimilation or resyllabification and the more drastic lexicalized "contractions" that were noted in earlier descriptions (and in dictionaries). This process serves to mark an intermediate level of prosodic grouping in between the syllable and the intonational phrase — namely the "foot", a constituent which has been posited for other Chinese languages to account for such things as (i) the distribution of unstressed (neutral tone) syllables and the domain of the morphologically regular contextual tone change of "tone 3 sandhi" in Beijing Mandarin and in other dialects of Chinese, and (ii) the domain of tone deletion and tone spreading in Shanghainese.

2.1 Syllable fusion is not "contraction"

Two other possible treatments could be proposed to account for the phenomenon of syllable fusion. First is the "contraction" account used in the past literature that classifies it as a more or less irregular purely morphological process that produces the odd "weak form" lexicalized variant of some very high-frequency phrases. In this section, I will argue that the process of syllable fusion in Cantonese is not simply "contraction" (although contraction can be the diachronic result in some cases).

This contrasts with previous descriptions of Cantonese, where we encounter the terms "縮減" /sok^{5}kaːm^{35}/ (literally 'shrink-reduce') (Yuan et al. [1960] 1983:192), translated as "contracted" in Cheung (1986:242-3), and "contraction" (Hashimoto 1972:114; Li 1986; Wong 1996; Bauer & Benedict 1997; Hsu 2005) to describe the connected speech phenomenon in (Hong Kong) Cantonese as exemplified in Table 1.1. The use of the terms 'shrink-reduce' or "contraction" in Cantonese, however, shows some difference from how "contraction" is typically used to characterize the English alternations between *do not* and *don't*, *I am* and *I'm*, *would not* and *wouldn't* and so on. A quick review of the past literature to see how the term "contraction" was used in Cantonese is in order.

Yuan et al ([1960] 1983:192) already seem to have noted there are intermediate degrees of bisyllabic reduction, such as the form [miː^{55} ɛː^{23}] 'what' that they cite as an intermediate form between bisyllabic [mɐt^{5} jɛː^{23}] and monosyllabic [mɛː^{23}] (lengthened) [1]

[1] The common form that one would encounter in present-day Hong Kong Cantonese has the high-level tone [55] (see Wong, Chan and Beckman 2005), or [55+(2)3] (the high falling-rising tone mark is often used to transcribe the form [mɛː] 'what' in the Hong Kong Cantonese Adult Corpus (Leung & Law 2001)), rather than the low-rising tone [23] as cited in Yuan et al.. This difference could be dialectal difference (Yuan et al. use the Guangzhou standard in their description).

for /mɐt^{5} jɛː^{23}/ ‘what’ (p. 192). Hashimoto (1972:114) also notes cases of bisyllabic reduction in forms such as 細蚊仔 /sɐi^{33} mɐn^{55} (tsɐi^{35})/ → [sɐmɐn^{33+55} (tsɐi^{35})] ‘child’, which is not shown to be further reduced to become monosyllabic.

Li (1986) characterizes “contraction” as follows: “When strings of words are pronounced casually, some of the juxtaposed syllables pair up and merge.” (p.1). Examples are cases such as the less fused forms in Table 1.1. In reduplicated syllables, however, Li points out that “contraction” fuses two syllables into a single syllable with the vowel length prolonged. For example, 清清楚楚 /ts^{h}eŋ^{55} ts^{h}eŋ^{55} ts^{h}ɔː^{35} ts^{h}ɔː^{35}/ is fused to become [ts^{h}eŋ^{55} ts^{h}ɔː^{35}] (both syllables lengthened) ‘clearly’[2] (p.8).

Cheung similarly proposes degrees of “contraction” as follows: “Some contractions involve the deletion of a non-final coda, some involve the deletion of non-marginal coda and onset.” (pp. 236-7). He distinguishes two kinds of “contracted forms”, the less extreme “plain form” and the more extreme “coerced form”. For example, 不如 /pɐt^{5} jyː^{21}/ can be “contracted” into the “plain form” [pɐ^{5} yː^{21}] or the “coerced form” [pyː$^{5+21}$] for ‘it’d be better’.

In Wong (1996), following the previous studies, I also used the term “contraction” in my study. I used the term interchangeably with “bisyllabic fusion” and defined “contraction” (and “bisyllabic fusion”) as follows: “Contraction involves a process where one or more segments in a syllable is deleted when compared with its citation form. Bisyllabic fusion is said to be realised when there is a deletion of non-marginal consonants between two contiguous syllables resulting in a change in the syntagmatic relation between segments.” (pp. 1-2).

[2] I would note that the “contracted” form would be more likely to occur utterance non-finally. Results for Experiment III in this dissertation verified this intuition.

Bauer & Benedict (1997) too define "contraction" as a "phonetic process in which one or more components of syllable is lost and remaining component combines with neighboring syllable" (p. 542; see examples in Table 3.9 on p. 319[3]). Thus, there are intermediate degrees of "contraction" when the term is used to describe the phenomenon as given in Table 1.1 in Cantonese connected speech. In fact, the examples in Table 1.1 are just a few tokens of an extremely fine-grained continuum, with many intermediate degrees of reduction (see Wong 2004).

By contrast, "contraction" as used in English typically does not seem to have much an indication of the intermediate effects. For example, Trask (1993) defines "contraction" as "A single phonological word representing a sequence of two or more separate word forms" (see also Crystal 1992, 1997; Trask 1996).

Note also that some of the intermediate forms produced by syllable fusion do not conform to Cantonese phonological structure constraints, as Cheung (1986) already noted when he observed "coerced" forms that are clearly monosyllabic in terms of the consonant-vowel structure but bisyllabic in terms of having two different tones specified. See an example of his cited in the previous paragraph, where 不如 /pɐt^{5} jyː^{21}/ is coerced to be [pyː$^{5+21}$] 'it'd be better'. That is, the "coerced" form can violate the constraint against the co-occurrence of a labial onset followed by a high front rounded vowel within a syllable. Syllable initial consonant clusters, which do not exist in the Cantonese phonological structure, can occur as a result of deletion of segments contiguous to the syllable boundary, e.g. /k^{h}ek^{5}lek^{5}k^{h}aːk^{5}laːk^{5}/ 'onomatopoeia for clicking sounds' can become [k^{h}lek^{5}k^{h}laːk^{5}], and /hɐm^{22}paːŋ^{22}laːŋ^{22}/ can become [hɐm^{22}plaːŋ^{22}] 'all' (cf. Chan 1984). By contrast, the English forms *don't* and *wouldn't* are all "legal" forms phonologically, with *don't* having the same general structure as the monomorphemic *point* and *wouldn't* having the same general structure as *bottled*.

[3] Some of the tonal realizations given do not seem to be common in Hong Kong Cantonese; dialectal differences could have come into play as examples cited are from, e.g., Bai (1982) and Yuen (1983).

Finally, fusion seems to be an extremely productive process that can affect *any* sequence of morphemes if the sequence becomes frequent enough. For example, when I went home for a visit after being away in the U.S. where I could not keep up with the local morning traffic news, it took me quite a while to recognize the fused form [fyːŋ$^{5+21}$ kaːi^{55}] of the street name 花園街 /faː^{55} jyːn^{21} kaːi^{55}/ 'Garden Street'. Note also that syllable fusion can occur across a word boundary (e.g. 我#覺得 /ŋɔː^{23}#kɔːk^{3}tɐk^{5}/ 'I feel' can be fused to become [ɔː$^{23+3}$tɐk^{5}]). Thus, the fusion process seems to be a productive prosodically driven process, rather than simply a word-formation process.

Therefore, using the term "contraction" could mislead the readers into thinking that "contraction" in Cantonese is analogous to "contraction" in English in many ways (thus, contra Bauer & Benedict who analogize "contraction" in Cantonese as "a process similar to English contraction" (p. 319)), which we have shown that it is not. Calling the above manifestation of fusion "contraction" does not seem to be appropriate given the relative complexity of the phenomenon in Cantonese. This dissertation is going to explore how to characterize the phenomenon and how to predict whether fusion occurs and to what degree.

2.2 Syllable fusion is not "coarticulation"

The second possible account one might use to characterize the phenomenon of syllable fusion is a purely phonetic account that classifies all but the most extreme cases as a low-level process of "gestural overlap" or "coarticulation". While coarticulation processes are probably physiologically motivated would be partly responsible for the phenomenon of syllable fusion in Cantonese, the phenomenon of syllable fusion in Cantonese shows patterns of lenition, deletion, assimilation, etc, that can be quite different from the processes described as "coarticulation" in other languages. For example, vowels (though reduced) tend to be retained in syllable fusion forms that are fused to a lesser degree in Cantonese. Even when the vowel is deleted, the tone (though it

might be truncated) tends to be retained in syllable fusion forms in Cantonese. We can contrast this maintenance of the vowel and especially the tone at the expense of the consonant to English, where only very stressed syllables "bear tone" and where vowels tend to be reduced and even devoiced or deleted in unstressed syllables in connected speech even in low frequency words and at slow speech rates. Thus, we argue that syllable fusion in Cantonese is not coarticulation, in the sense that it is not a merely physiologically motivated process.

2.3 Syllable fusion and the Cantonese prosodic hierarchy

If the phenomenon of syllable fusion in Cantonese is characteristic of some level of prosodic structure between the syllable and the larger intonational phrase, characteristics of syllable fusion as stated in sections 2.1 and 2.2 are just what we might expect. The prosodic group that syllable fusion defines is smaller than the intonational phrase but bigger than the syllable. The current challenge for the investigation of the Cantonese prosodic system lies in the discovery of a structurally well-defined intermediate-level prosodic constituent between the syllable and the intonational phrase.

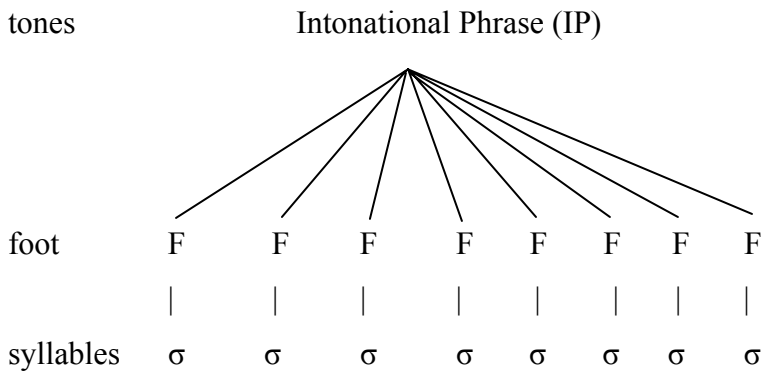
Figure 2.1 shows the current analysis of the intonational structure of Cantonese in the Cantonese Tones and Break Indices (or C_ToBI) system (Wong, Chan & Beckman 2005). Structural alignments are given when syllable fusion is present or absent.

In the C_ToBI analysis, the intonational phrase is defined by the presence of an edge tone. That is, the edge tones are characteristic phonological markers for this level of the hierarchy. These markers are "extra" tones that can be added after the final lexical tone of a prosodic phrase to produce various pragmatic effects. Table 2.1 shows the lexical tone inventory of Hong Kong Cantonese, and Table 2.2 shows the inventory of intonational phrase boundary tones as proposed in Wong, Chan & Beckman (2005).

The Cantonese syllable is defined by the distribution of tones and of particular segment types. That is, each lexical tone marks the head of a syllable and syllable boundaries are clearly identifiable from asymmetries in the distribution of onset and coda segments so that there is nothing like the large-scale ambiguities of segmentation within the stress feet of an English utterance. A hierarchical display of the internal structure of the Cantonese syllable is given in Figure 2.3. Tables 2.3 through 2.5 are the phoneme inventories, and Table 2.6 shows phonemic alternations often found among younger speakers of the language (cf. Hashimoto 1972; Yeung 1980; Cheung 1986; Bauer and Benedict 1997; the more prevalent substitution of final [t] for final /k/ than vice versa see, e.g., Cheung 1999; Zee 1999a; Law, Fung & Bauer 2001; Wong 2002).

The current preliminary account of syllable fusion in C_ToBI is in terms of the "foot" as illustrated in Figure 2.1. In careful speech, where there is no fusion, each syllable is a foot, as in Figure 2.1a. When fusion occurs, then, it marks the "erasure" of the foot boundary between the affected syllables, as shown in Figure 2.1b. Readers who are familiar with morphosyntactic differences across different varieties of Chinese might question the necessity for positing any such intermediate level of prosodic grouping between the intonational phrase and the syllable. However, another approach is to acknowledge that syllable fusion is rare in careful read-speech styles, in which case part of the extremely "monosyllabic" flavor of Cantonese might be ascribed to the preponderance of monosyllabic feet as in citation form utterances Figure 2.1a. The salience of Cantonese syllable edges, however, can be very much obscured in everyday connected speech when syllable are fused as in Figure 2.1b. Fused forms then are exceptions to the regular one-to-one correspondence between foot and syllable. Such exceptions may presage a change toward a system with a more well-defined intermediate level of prosodic grouping between the syllable and intonational phrase, which is a more consistent cue to the "word" constituent in the morphosyntax. Figure 2.2 gives an example utterance to help visualize how the elements at the morphosyntactic level may correspond to the intonational structure.

(a) Alignment of tiers when no syllable fusion occurs across the syllables.



(b) Alignment of tiers when syllable fusion occurs across the syllables.

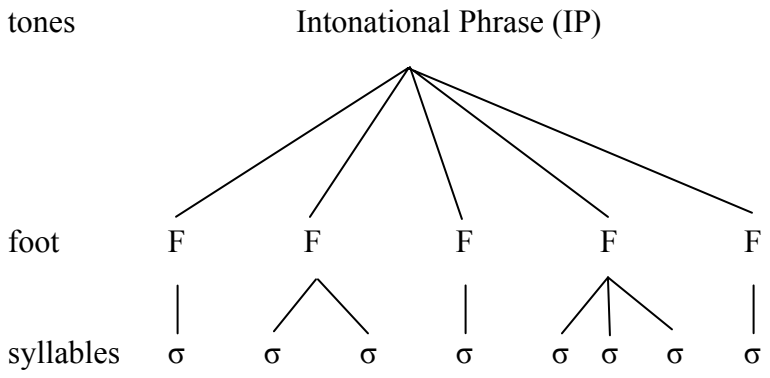


Figure 2.1 The Cantonese intonational structure as in C_ToBI.

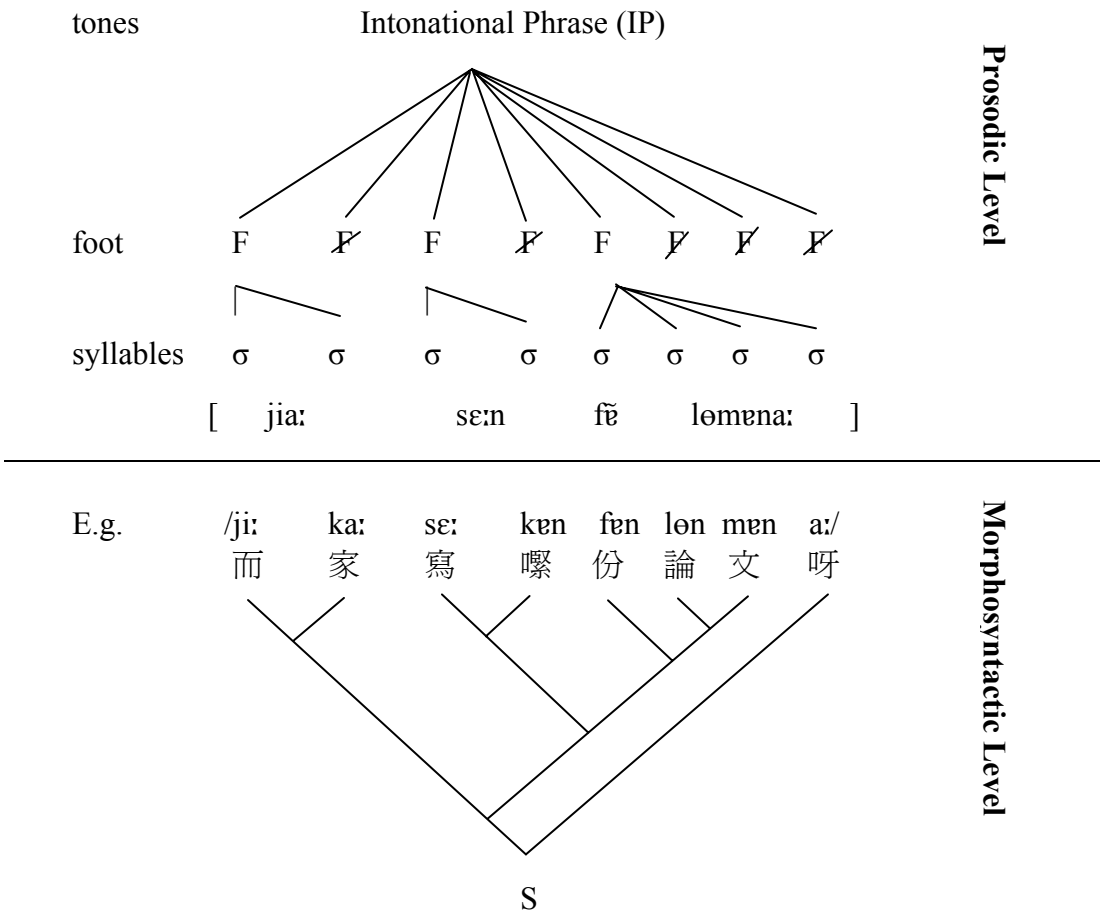


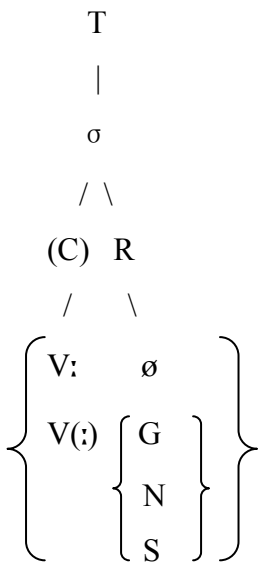
Figure 2.2 The foot level proposed for the Cantonese intonational structure. (A cross on a foot indicates a foot erased.) The utterance ‘now writing the dissertation (literally ‘now writing classifier dissertation sentence-final particle’)’ and its internal morphosyntactic structure is given as an example.

		Syllables *without* final stop consonant	Syllables *with* final stop consonant
Level tones:	High level	55	5
	Mid level	33	3
	Low level	22	2
Rising tones:	High rising	35	35
	Low rising	23	---
Falling tone:	Low falling	21	21

Table 2.1 The tonal inventory of Hong Kong Cantonese represented in Chao-type tone numbers, with '1' to '5' in ascending pitch height (e.g. /55/ is high level, /21/ is low-falling, etc).

Tone types	Descriptions
L%	fall from the final lexical tone
H%	rise from the final lexical tone
H:%	rise from the final lexical tone, with a short plateau at the very end of the rise; incredulity reading accompanied
HL%	final rise and then fall from the final lexical tone
%	phrase-end with no extra tone
-%	truncated rise of the final lexical tone
%fi	frame-initial boundary used to mark the initial particle in phrase-framing particle pairs such as "*mat5 .....me55? (rhetorical question)*"

Table 2.2 The inventory of boundary tones in C_ToBI (details see Wong, Chan & Beckman 2005).



T = lexical tone

C = consonant

R = rime

V: = long vowel

V = short vowel

G = glide

N = nasal

S = oral stop

Syllabic nasals: m̩ ŋ̩

Figure 2.3 The Cantonese syllable.

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Labio-velar	Glottal
Plosive	**p** p^h			**t** t^h			**k** k^h	k^w k^{wh}	
Affricate				ts	ts^h				
Nasal	**m**			**n**			**ŋ**		
Fricative		f		s					h
Approximant						j		w	
Lateral Approximant				l					

Stop coda consonants (marked in bold face) are unreleased.

Table 2.3 Onset / coda consonants: 19 onsets; 6 codas (Zee 1999:59).

front		central	back		
	round			round	
iː	yː			uː	high
e		ɵ		o	mid (short)
ɛː	œː			ɔː	mid (long)
		ɐ	aː		low vowels

$ The adoption of /e/ and /o/ instead of /ɪ/ and /ʊ/ is motivated by the acoustic analysis fo Lee (1999). One would likely to find the alternative analysis of using /ɪ/ and /ʊ/ in the literature (e.g. Hashimoto 1972).

Table 2.4 Eleven nuclear vowels in the Hong Kong Cantonese vowel inventory.

-i	-y	-u
uːi		iːu
ei		ɛːu
ɔːi		ou
ɐi		ɐu
aːi		aːu
	ɵy	

Table 2.5 Diphthongs in Cantonese. (An alternative analysis of these nuclei is vowel followed by a glide.)

Onset	Coda	Syllabic nasals
ŋ ~ ø	n ~ ŋ	ŋ̩ ~ m̩
n ~ l	t ~ k	
k^{w} ~ k / __ ɔː	ʔ ~ t/k	
k^{wh} ~ k^{h} / __ ɔː		

Table 2.6 Phonologically conditioned alternations in present-day Hong Kong Cantonese.

In summary, then, the various predictions that will be tested follow the account of syllable fusion encoded in the current C_ToBI conventions. This account defines syllable fusion as the marker of some intermediate level of grouping. A good candidate for such an intermediate-level of grouping would be the "foot" — the prosodic "foot", which has been invoked to explain several connected speech processes in other varieties of Chinese, such as neutral tone and third tone sandhi in Beijing Mandarin (and standard Putonghua Mandarin), and tone sandhi in Pingyao Mandarin, Shanghainese and Xiamen (e.g. Shih 1997; Yip 1980; Chen 1987, 1990; Selkirk & Shen 1990).

2.4 Is syllable fusion manifestation of "stress"?

Before reviewing the phenomena that have been ascribed to the foot in Chinese, we must distinguish the C_ToBI "pure grouping" hypothesis from an alternative approach that one could entertain, given the literature. This is the hypothesis that the emergence of syllable fusion in Cantonese and tone sandhi in other Chinese dialects such as Beijing Mandarin, Shanghainese and Xiamen are manifestations of an underlying trochaic or predominately iambic stress pattern for words that some researchers posit (e.g. Chan 1985; Lee 2003 for Cantonese; Duanmu 1992, 1993, 1995, 1998, 1999; Wright 1983 for other Chinese varieties such as Beijing Mandarin and Shanghainese). While the idea is interesting for further exploration, we should point out that, while neutral tone

syllables have relatively reliable cues for cueing polysyllabic lexicality in a sequence of syllables, perception studies show that the perception of stress differences among syllables with non-neutral tones in Beijing Mandarin is not conclusive and needs further studies. That is, while some of the segmental and tonal effects of syllable fusion forms in Cantonese are similar to those found in the neutral tone "unstressed" syllables in Beijing Mandarin, this fact does not necessarily warrant an intrinsic trochaic "stress" pattern underlying the words.

Thus, before we can entertain the notion that "stress" is relevant, two points would need clarification. First, at what level in the prosodic structure does the claimed "stress" occur? For example, in Mandarin "stress" has been defined at two levels: (i) the contrast between stressed syllables with full tone and reduced syllables with "neutral" tone, which defines a trochaic foot, and (ii) a purported contrast between two non-neutral toned syllables, which defines an iambic foot. The previous studies (e.g. Lin, Yan & Sun 1984, Yan & Lin 1988, Wang & Wang 1993) on the "stress" on non-neutral tones (or so-called "normal stress") showed "normal stress" syllables do not function as clearly as neutral tone syllables do in cueing polysyllabic lexicality in a sequence of syllables. At best we can say that Wang & Wang's study showed that the Putonghua or Beijing Mandarin listeners *might* use durational cues for word identification in a sequence of non-neutral tone syllables. In their test of trisyllabic words, Yan and Lin (1988) found that pitch range expansion and pitch contour integrity (as the syllables were produced in citation form) among the tokens tended to give rise to the "heaviest" stress (i.e. most stressed) responses, while duration may or may not contribute to the judgment. A further problem is that most studies tested the perception of "stress" in Beijing Mandarin using citation form utterances; it remains to be tested whether the listeners are sensitive to the potential acoustic cues to "normal stress" in connected speech.

Second,since Cantonese has no neutral toned syllables, we would need to develop a clear account of what the actual markers of "stress" are. In particular, how does the pitch pattern contribute to the perception of "stress" on a syllable that retains its lexical

tone specification? This point is quite unclear even for Mandarin Chinese. Suppose it is not duration or pitch alone, but a "package" of cues that might contribute to the perception of "stress" in Beijing Mandarin, then the above studies are ambiguous and it is quite unclear to us how pitch contributes to the perception of the "heaviness" of syllables, and in particular, how pitch might interact with the durational cues? Also, if pitch affects "stress" perception, it is not clear whether it is the pitch contour or some point (e.g. the starting point or the end point) of the pitch contour that is responsible. As for Cantonese, I have not been able to find any perception studies that tested the notion of "stress" as to date.

Having said that, we should now take a closer look at how the prosodic foot that this dissertation plans to invoke is related to "stress" in Beijing Mandarin and/or tone sandhi patterns in various Chinese dialects, and how the foot maps onto morphological and/or syntactic structures.

2.5 The foot in other varieties of Chinese

In this section three phenomena will be described in two dialects of Chinese where the "foot" has been invoked to explain how some domain (e.g., the domain of distribution of neutral tone, and the domain of tone deletion/spread in Shanghainese) in the phonology lines up with the "word" in the morphosyntax.

2.5.1 The foot and the neutral tone in Beijing Mandarin

Beijing Mandarin has four tones and a "neutral tone" in the contrastive tonal inventory. The "neutral tone" differs from the other tones in that a syllable with this "tone" is inherently unspecified for tone: the pitch of the neutral tone varies mostly in accordance with the preceding tone. The neutral tone, therefore, is sometimes described as "atonic" or "toneless". For notation convenience, we will use the label "0" for neutral tone.

The tone shapes of the four tones are summarized in Table 2.7, each illustrated with an example morpheme. Each morpheme in Beijing Mandarin, therefore, has one of these four tones as its "intrinsic / underlying" tone, or it has neutral tone.

Description	Transcriptions		Example		
High-level	55#	T1$	mā	媽	'mother'
High rising	35	T2	má	麻	'hemp'
Falling-rising	21(4)	T3	mǎ	馬	'horse'
High-falling	51	T4	mà	罵	'to scold'
Neutral tone		T0+	ma	嗎	'(particle)'

These are Chao-type tone numbers, with '5' denoting the highest pitch in the pitch range, and '1' the lowest.

$ These are symbols that mimic the pitch contours in the PRC's Pinyin Romanization system: ¯ for tone 1, ´ for tone 2, ˇ for tone 3, and ` for tone 4. Neutral tone is not marked with diacritics in the Pinyin system.

+ T0, merely as a symbol for explicitly marking neutral tone, is used in this dissertation.

Table 2.7 The Beijing Mandarin tonal inventory.

Almost every syllable is a morpheme in Chinese. Beijing Mandarin is no exception. Syllable structure in Beijing Mandarin is rather simple: a syllable has an obligatory nuclear vowel, while onset, medial, and coda are optional. By contrast to the Cantonese syllable, in Mandarin tone is not an obligatory phonetic marker on each syllable, since "neutral tone" means that the phonetic value of the tone is not specified. In most cases a syllable with neutral tone occurs enclitically following a syllable that has one of the four tones, and the actual pitch of the neutral tone is determined by the preceding tone in general (Chao 1968:149-150). In fact, the neutral tone may occur in polysyllabic expressions where every syllable except the first may have the neutral tone.

(Thus, neutral tone syllables never occur in word-initial position.) To the Beijing Mandarin listeners, having a neutral tone syllable is a very clear indication that a recurring syllable sequence is a fully lexicalized polysyllabic word rather than a more decomposable compound or even a phrase, as in the monomorphemic 東西 /toŋ1ɕi0/[4] 'thing' versus the compound 東西 /toŋ1ɕi1/ 'east-west'. On the other side of the coin, Chao observed that some words that did not originally have neutral tone now have optional neutral tone. These are mostly half-new expressions and translation borrowings on the way to being naturalized as everyday expressions (at least during his time). Examples he gave included 希望 [ɕi1uɑŋ0] 'hope', 知道 [tʂɪ1tɑu0] 'know' and 好像 [xɑu3ɕiɑŋ0] 'seems like'. This might lead one to wonder if frequency of use might contribute to the words having "optional" neutral tone — i.e. speakers who pronounce these bisyllabic forms with neutral tone are explicitly marking them phonetically to be more like an indecomposable word. In other words, neutral tone is used as a word-formation process.

While the occurrence of neutral tone is linked to lexical status, it is not as reliable a marker of wordhood as stress is in English, where the vast majority of bisyllabic words contain one foot, consisting of a stressed syllable followed by a reduced syllable. That said, there are still many similarities between Mandarin neutral tone syllables and English reduced syllables.

For example, no content words have only neutral tone syllables, although there are a small number of inherently "unstressed" morphemes, such as suffixes and particles, which are always in the neutral tone (e.g. the nominal suffix 子 /tsə/ in words such as 筷子 /kuai4tsə/ 'chopsticks'). Almost any morpheme in one of the four regular tones can be realized in the neutral tone under certain conditions. For example, there are morphemes

[4] The numbers refer to tone, as in T1, T2, and so on (see Table 2.7).

that have tones when they are used as free-standing words, but lose their tone specification in certain contexts to form bisyllabic words (e.g. 花 /xuɑ1/ ‘flower’ vs. 棉花 /miɛn2xuɑ0/ ‘cotton’; 淨 /tsiɤŋ4/ ‘pure’ vs. 乾淨 /kan1tsiɤŋ0/ ‘clean’) (see Cheng (1973)). According to Cheng (1973), there were perhaps a few hundred words containing a syllable in the neutral tone that must be specified by some feature in the lexicon (see e.g. 兄弟 /ɕioŋ1ti4/ ‘brothers’ vs. 兄弟 /ɕioŋ1ti0/ ‘younger brother’; 大意 /ta4i4/ ‘outline’ vs. 大意 /ta4i0/ ‘careless’). In Yip’s (1980) formulation, that feature would be the foot structure. While each of the regular tone syllables makes a foot, the neutral tone syllable is footed with the preceding syllable to make one branching foot. That is, neutral tone syllables in these words are not treated as the result of a process, but the neutral tone syllables are lexically specified. Figure 2.4 gives the foot structures of the pair of words ‘outline’ vs. ‘careless’. Cheng also proposes that reduplicated nouns (where the second syllable is essentially a reduplicate of the first syllable, except for the tone which becomes neutral tone) must be listed in the lexicon because noun reduplication is not a productive method of generating new items. In other words, when a neutral tone occurs in a context that might be unfamiliar to the Beijing Mandarin listeners, they would still be able to know that there is a word boundary at the right edge of the neutral tone. We might then expect that, by being able to identify the acoustic correlates (see below) of neutral tone, the listeners would still be able to know that this neutral-toned morpheme-syllable must be grouped together with the nearest preceding full-toned syllable into a polysyllabic word.

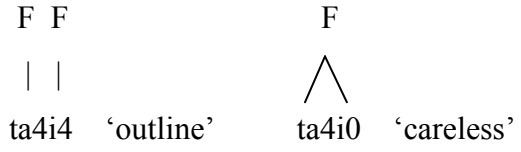


Figure 2.4 Foot alignment for the words 大意 /ta4i4/ ‘outline’ vs. 大意 /ta4i0/ ‘careless’.

The acoustic correlates of Beijing Mandarin neutral tone have been described in terms of fundamental frequency (the term pitch is often used in the literature), duration, intensity, and segmental effects. The actual pitch of neutral tone in Beijing Mandarin is generally described to be dependent on the preceding tone. The influence of the tone preceding the neutral tone may be seen more clearly in a succession of neutral tones. See examples in Table 2.8 below. The examples show that neutral tone follows the drift of the preceding tone, and the drift is upward after T3 (this is probably because T3 is underlyingly a fall-rise tone), but downward elsewhere. Later studies found that neutral tone has pitch contour, rather than just a single pitch height (Dreher and Lee 1966; Wang and Li 1967). Dreher and Lee measured the neutral tone shapes on the five-point scale. The results are shown in Table 2.9.

a.	tsou3 tʂ^{h}u0 lai0	走出來	'come out'
	pitch rising through the two neutral tone syllables		
b	k^{h}an4 tɕ^{h}i0 lai0 .	看起來	'to take a look'
	pitch falling low through the two neutral tone syllables		
c.	tsou3 pu0 k^{h}ai1	走不開	'cannot get away'
	pitch is low for the neutral tone syllable		

Table 2.8 Neutral tones in succession (examples given in Chao 1948)

after first tone:	41
after second tone:	31
after third tone:	23
after fourth tone:	21

Table 2.9 Measurement of the neutral tone shape by Dreher and Lee (1966) on the five-point scale.

We must, however, note in passing all of the above descriptions of neutral tone were based on forms produced in isolation; Shen (1990:40) cited Tseng's (1981) study on spontaneous speech. Tseng has found 29.55% (i.e. less than one-third) of the neutral tone syllables are actually produced in accordance with the phonological predictions. Instead, the great majority of neutral tone syllables are realized with more variations than were expected. Tseng concludes that the phonetic outputs of neutral tone demonstrate that their tonal patterns are not always governed by those of the preceding syllable.)

A characteristic of neutral tone that can cue word identification is its duration. Chao (1968) observed that the duration of neutral tone is relatively short when compared with a syllable that has one of the four tones. Durational measurements by Zadoenko (1958) as cited in Cheng (1973:55) confirm Chao's observation by showing that in a bisyllabic word, the length of a neutral tone syllable is about half that of a non-neutral tone length, and all neutral tone syllables have about the same length, whether they are monophthongal or otherwise. Neutral tone syllables are, hence, also called "unstressed" syllables in Beijing Mandarin (versus "stressed" syllables as having one of the four tones) with durational cues as part of the "stress" cues. Lin (1985) tested factors that might affect the identification of neutral tone among the Beijing Mandarin listeners. The factors were duration, intensity and pitch contours. In his study, Lin synthesized seven pairs of bisyllabic words or phrases, with each pair contrasting in meaning depending on whether the second syllable is perceived as neutral tone ("unstressed") or not. An example pair is the noun phrase 鴨頭 /ia1 t^hou2/ 'duck's head' versus the noun 丫頭 /ia1t^hou0/ 'girl'. Results in Lin's study showed that durational cues are relatively reliable for the identification of neutral tone, although he does not report any statistical tests to tease out whether there were other main effects or interaction effects.

Although there are qualitative descriptions of the tone effects for fused forms in, e.g., Hashimoto 1972; Cheung 1986; Li 1986; Wong 1996; Wong et al. 2002; Lee 2003; Wong 2004; Wong, Chan & Beckman 2005, there are as yet no experimental studies of

the acoustic patterns in terms of F0, duration, intensity for syllable fusion forms in connected Cantonese speech. At the same time, there are segmental effects that Cantonese syllable fusion forms show similarities to segmental effects for neutral tone syllables in Beijing Mandarin. For the neutral tone syllables, the segmental effects can result in "deviation" from the description of the Beijing Mandarin phonology that we give earlier in this section. The "illegal" phonotactics and segments compared to the Beijing Mandarin phonology could probably cue the listeners in identifying whether a sequence of syllables constitute a word.

The segmental effects include deaspiration and voicing of the syllable onset. In Beijing Mandarin, when in a syllable bearing one of the four tones (i.e. "stressed"), unaspirated stops /p, t, k, ts, tɕ, tʂ/ are voiceless, as specified in the phonology of Beijing Mandarin. However, they tend to become voiced when in a syllable in the neutral tone (i.e. "unstressed"). As reported in Cheng (1973:82), in an acoustic study of the neutral tone, Zadoenko (1958:585) observes that aspirated stops in a neutral tone syllable that is preceded by an open syllable also tend to become voiced. The affricate /ts/ is usually not fully voiced in the neutral tone, except that the noun suffix /tsɪ/ is pronounced as [zə] (as in [t^hu4zə] 'rabbit') by many speakers. In other initials voicing does not occur, but aspirated segments tend to become unaspirated, so that the distinction between aspirated and unaspirated consonants is realized instead as a distinction between voiceless and voiced when the syllable is in the neutral tone[5].

Vowel laxing and devoicing have also been found to occur in neutral tone syllables. Vowels in neutral tone tend to become shorter, lax and centralized. For low vowels, they may become schwa. For high vowels, they may become lax but do not

[5] This is a fairly regular pattern in Germanic, too. That is, in English and other Germanic dialects that have a contrast between aspirated and unaspirated plosives initially, the contrast is voiceless unaspirated versus voiced in weak foot-medial position.

become schwa. The vowels are voiced when following tones other than T4 (e.g. 姐夫 /tsiɛ3fu/ → [tsiɛ3fu] 'older sister's husband'), but after a T4, a high vowel in a neutral tone syllable with the initial [f, s, ʂ, ts^h, tʂ^h] tends to be devoiced.

The most extreme case of neutral tone effect on the syllable is probably segmental deletion. For example, a devoiced high vowel can essentially merge with the preceding [f, s, ʂ, ts^h, tʂ^h]. Since these are voiceless consonants that do not become voiced even in the neutral tone, the "standard" onset of the second syllable might then resyllabify with the preceding syllable, resulting in a syllable, with a single tone, having a fricative as the coda, or having a consonant cluster in the coda position. That is, these segmental effects result in "new" syllable types that do not obey the phonotactic constraints of Beijing Mandarin. Also, Chao (1968) and Cheng (1973:34) observed that the rime of a second syllable with a bilabial nasal onset is optionally deleted and the "stranded" nasal onset resyllabifies as the coda of the first syllable (e.g. 我們 /uo3mən0/ → [uom3(+0)] (literally, 'I plural suffix') 'we'). This process gives rise to a syllable structure that is not predicted by the phonology of Beijing Mandarin. In fact, linguists do not see eye to eye about how many syllables this process results in. While Cheng (1973) and Dai (1990) would agree that [uom] would count as one syllable, Chao pointed out that it is not certain that the [m] is not occupying the position of a syllable in the neutral tone, even when there is no vowel (p. 141). This would be a case similar to syllable fusion in Cantonese where the syllable count could become indeterminate when syllables are fused (cf. Wong, Chan and Beckman 2005).

A less controversial phenomenon of segmental deletion can be observed in the process of retroflex suffixation (see Chao 1968; Cheng 1973:24). Retroflex suffixation refers to the deletion of the vowel in the second syllable 兒 /-ɤɹ2/ (the diminutive marker etymologically related to 兒 /ɤɹ2/ 'son') in a compound and resyllabifies [-ɹ], which has become "toneless", as the coda of the first syllable (e.g. 刀兒 /tɑu1-ɤɹ2/ → [tɑuɹ1]

‘knife’; 牌兒 /p^{h}ai2-ɤɹ2/ → [p^{h}əɹ2] ‘signboard’). When [-ɹ] “resyllabifies” with the preceding syllable, various phonetic consequences would occur, e.g., deletion of final nasals, vowel nasalization in the case a final /ŋ/ is deleted, and insertion of schwa between the front vowel in the preceding syllable and [-ɹ]. Thus, segmental deletion and resyllabification signal grouping of syllables, which signal the “words”. This phenomenon can be seen as a “morphologized” end result of the same process of neutral tone reduction, similar to the “true contraction” cases (such as the word ‘what’) in Cantonese.

All of these phenomena have been analyzed by some linguists as foot-level effects. In particular, Yip (1980) proposes that tone is what determines foot construction in Beijing Mandarin. Each syllable that has one of the four regular tones can form a foot by itself, whereas a neutral tone syllable has to be footed together with the nearest preceding regular tone syllable, given that neutral tone does not occur word-initially. Thus, a regular tone syllable can make a non-branching foot by itself, or it makes a branching foot if grouped with the following neutral tone syllable(s). In the latter case, the strong head syllable of the branching foot is the first syllable of the foot. Given the distribution of the neutral tone syllables, it follows that a branching foot always has to be left-branching. Examples are given in Figure 2.5. A word then can consist one or more feet. The relationship between the foot and the word can be seen in the examples given in Figure 2.6. Note that the word level metrical structure in Figure 2.6b suggests that there is no difference between the metrical structure of a phrase and that of a word. There are cases where a word loses its “underlying” regular tone when occurs in running speech (e.g. 看他 /k^{h}an4 t^{h}a1/ → [k^{h}an4 t^{h}a0] (literally, ‘look him/her’) ‘look at him/her’). Yip accounts for these cases of cliticization by the process of “defooting”. That is, 他/她 /t^{h}a1/ ‘him/her’ in the above example is “defooted” and is then footed with the preceding “stressed” syllable to form a trochaic foot; thus, having a configuration no different from Figure 2.5a below.

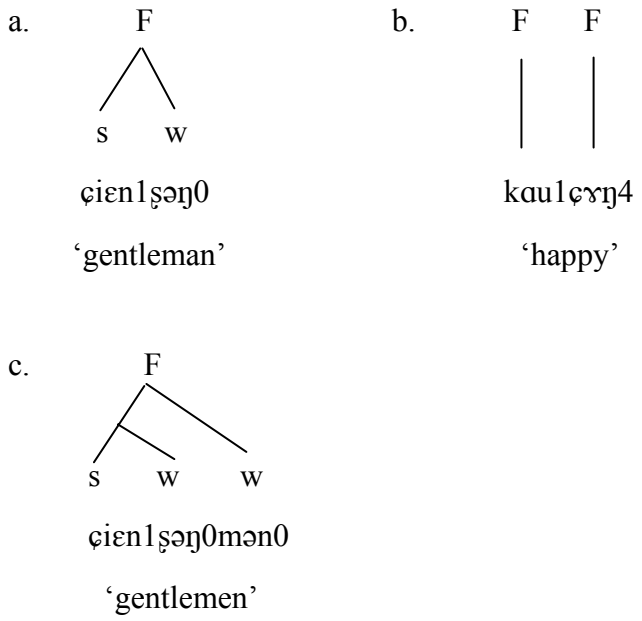


Figure 2.5 Foot structure in Beijing Mandarin (Yip 1980)[6].

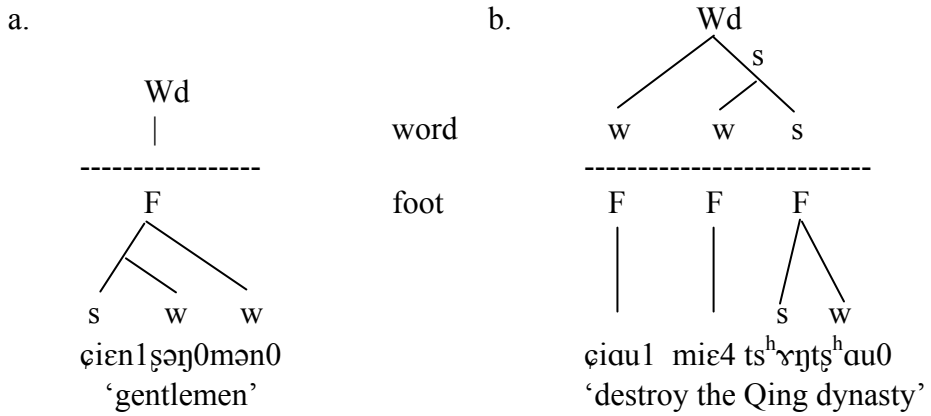


Figure 2.6 Relationship between the foot and the word in Beijing Mandarin (Yip 1980).

[6] Notations such as the symbols for "foot" and for "word" in Yip (1980) are replaced by F and Wd here. "w" means metrically weak syllable and "s" means strong.

Given Yip's formulation, the phonological/metrical characterization of lexical entries would be as follows. First, a word must contain at least one foot (a branching foot or a non-branching foot). Second, a word can have no branching foot, or it can have at least one branching foot. Lastly, the strongest head is always the strong head of the rightmost foot. One implication that arises out of Yip's formulation is that there is no distinction being made between the metrical structure of the word and of larger phrases.

2.5.2 The foot and third tone sandhi in Beijing Mandarin

The "foot" has been invoked also in descriptions of the phenomenon of "third tone sandhi" (hereafter, T3 sandhi) in Beijing Mandarin (although the definition of foot here is somewhat different). As mentioned in chapter 1, the rule of T3 sandhi states that in sequence of third tones (T3s), all third tones, except the last, are changed into the second tone (T2). The question of what is the domain of T3 sandhi arises. T3 sandhi domain may coincide with word boundaries, with or without word-internal branching structures, or it may apply across word and phrasal boundaries. Shih (1997) proposed that the minimum domain for T3 sandhi in Mandarin is the "foot" and that T3 sandhi rule application is obligatory within this prosodic unit. Mandarin foot formation as proposed by Shih is specified in Table 2.10.

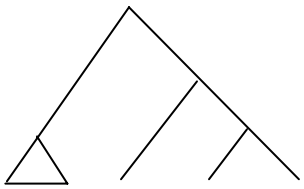
a.	Join immediate constituents into bisyllabic feet.
b.	Scanning from left to right, join monosyllabic constituents into bisyllabic feet.
c.	Join any remaining monosyllable to a neighboring bisyllabic foot to form a 'superfoot' according to the direction of syntactic branching.

Table 2.10 Mandarin foot formation rules (from Shih 1997:98).

It can be seen in the formulation in Table 2.10 that the prosodic foot is sensitive to the morphological/syntactic structure initially (see Table 2.10a), but it ignores syntactic sisterhood at later stages (see Table 2.10b and c). Shih further points out that it is the branchingness that the prosodic foot is sensitive to, but not the node labels (i.e. not the phrasal categories).

T3 sandhi is applied cyclically from the innermost prosodic foot, with a choice of initial cycle. The choice of initial cycle may be higher than the level of a bisyllabic foot depending on speech rate or frequency of use, so that fast speech and frequent usage of the expression tend to lead one to choose a higher level domain as the initial cycle for T3 rule application. Figure 2.7 below shows three choices of initial cycle (conjoined strings at the initial cycles are underlined), with the level of prosodic foot highest in (c), then in (b), and then in (a). Cyclic application of the T3 rule may be seen in (a) and (b). The cyclic application of T3 sandhi and the choice of initial cycle at the word level are given in Figure 2.8.

a.



lau**3**li**3** mai**3** xau**3** tsiou**3** ‘Old Li buys good wine.’

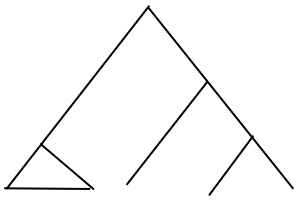
2 **2**

--

2

lau**2**li**2** mai**3** xau**2** tsiou**3**

b.



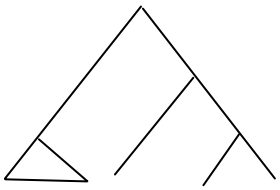
lau**3**li**3** mai**3** xau**3** tsiou**3**

2 **2** **2**

--

lau**2**li**3** mai**2** xau**2** tsiou**3**

c.



lau**3**li**3** mai**3** xau**3** tsiou**3**

2 **2** **2** **2**

lau**2**li**2** mai**2** xau**2** tsiou**3**

Figure 2.7 Three choices of initial cycle for the foot domain (Shih 1997).

a.

tʂɪ**3** <u>lɑu**3**xu**3**</u> paper tiger ‘pretense, coward’

2

--

tʂɪ**3** lɑu**2**xu**3**

b.

<u>tʂɪ**3** lɑu**3**xu**3**</u> paper tiger ‘pretense, coward’

2 2

tʂɪ**2** lɑu**2**xu**3**

c.

<u>lɑu**3**xu**3** </u>tan**3** tiger gall ‘bravery’

2

2

lɑu**2**xu**2** tan**3**

d.

<u>lɑu**3**xu**3** tan**3**</u> tiger gall ‘bravery’

2 2

lɑu**2**xu**2** tan**3**

Figure 2.8 The cyclic application of T3 sandhi and the choice of initial cycle at the word level.

2.5.3 The foot and tone sandhi in Shanghainese

Tone sandhi in Shanghainese is probably the clearest case of indicating what a "word" is among many Chinese dialects. In Shanghainese, when morphemes are strung together, the citation tones of the morphemes are systematically changed to indicate that the string makes a compound word rather than a phrase. Specifically, descriptions in the past literature on the phonology of Shanghainese tone sandhi generally agree that the tone sandhi pattern within a compound is determined by the tone of the leftmost morpheme of the compound (Yip 1980; Wright 1983; Jin 1986; Selkirk & Shen 1990; Duanmu 1998). For all compounds, lexical tones from all but the leftmost syllable-morpheme of a compound are removed and the underlying targets of the leftmost syllable-morpheme are distributed over the whole word in a process of edge-in tone spread.

Zee (2001) (cf. Zee and Maddieson 1979) gives the most phonetic description of the tone sandhi patterns using Chao's five-point tonal scale, with "1" being the lowest and "5" the highest pitch in the local pitch range, based on the production of his native Shanghainese informant. Table 2.11 compares citation tones and polysyllabic word tone sandhi patterns given by Zee.

Lexical tone melodies	Monosyllabic citation	Bisyllabic compounds	Trisyllabic compounds	Quadrisyllabic compounds	Quintesyllabic compounds
A	[51]	[5-1]	[5-3-1]	[5-3$^{\uparrow}$-3$^{\downarrow}$-1][a]	[5-4-3-2-1]
B	[34]	[3-4]	[3-5-1]	[3-5-3-1]	[3-5-3$^{\uparrow}$-3$^{\downarrow}$-1]
C	[13]	[1-3]	[1-5-1]	[1-5-3-1]	[1-5-3$^{\uparrow}$-3$^{\downarrow}$-1]
D	[5]	[4-5]	[4-5-1]	[4-5-3-1]	[4-5-3$^{\uparrow}$-3$^{\downarrow}$-1]
E	[12]	[1-12] ([1-3])[b]	[1-1-12] ([1-5-1])	[1-1-1-12] ([1-5-3-1])	[1-1-1-1-12] ([1-5-3$^{\uparrow}$-3$^{\downarrow}$-1])

[a] The arrows "$^{\uparrow}$" and "$^{\downarrow}$" indicate the relative higher and lower local pitch in the local pitch range.
[b] In parentheses are optional patterns observed in some words.

Table 2.11 Citation tone and tone sandhi patterns in polysyllabic words (Zee 2001). (Lexical tone melodies I through V in his original table correspond to A through E in our table.)

Essentially, there are five lexical tones, and all five are underlyingly contour tones, although tone D surfaces as a single high target in monosyllabic citation form. In compound forms, the targets of the falling tone (tone A) are aligned to the first and last syllables with interpolation in between for compounds longer than two syllables. The targets of the rising tones (types B, C and D), by contrast, align to the first two syllables, with a subsequent fall in citation form. The low tone (tone E) is like third tone in Beijing Mandarin: there is a small rise at the end in citation form.

Shanghainese tone sandhi is sensitive to word versus phrasal boundaries. Table 2.12 below gives some examples.

From Duanmu (1998:172)[7]:				
a.	ts^ho ve → (ts^ho ve) LH LH L H fry rice 'fried rice'		a compound noun	
b.	ts^ho ve → (ts^ho) (ve) LH LH LH LH fry rice 'to fry rice'		a verb phrase	
From Duanmu (1998:177-8):				
c.	(ka) (ɦõ)[#] 'so red'	(lo) (ɦõ) 'very red'	(tsø) (ɦõ) 'most red'	adverbial phrases
d.	(do ɦõ) 'peach red'	(tsz ɦõ) 'purple red'	(du ɦõ) 'big red'	adjectives
e.	k^hø sz →	(k^hø sz)	'reading'	a noun
	→	(k^hø) (sz)	'to read a book'	a verb phrase

Tone was not shown in the paper, but the brackets indicate the occurrence of tone sandhi.

Table 2.12 Sensitivity of Shanghainese tone sandhi to word versus phrasal boundaries.

Note from the above phonetic and distributional characteristics of Shanghainese tone sandhi that monosyllabic words of Shanghainese bear contour tones, except when the monosyllabic words are of Type D tone. That is, the word status of a string of syllables can be identified quite readily. Basically, a level tone would cue the

[7] Note that differing from Jin (e.g. 1997) who posits three tone levels, L, M, and H for Shanghainese, Duanmu ignores the tone registers and segmental factors, positing just two phonological tone levels L and H.

Shanghainese listeners that the morpheme is part of a compound, rather than a monosyllabic word by itself. However, there could be some temporary ambiguity about the word status of a tone when the listener encounters, say, phonetic tone level [5]. In that case, the listeners would probably need to rely on other cues such as segmental cues. For example, in word-initial position underlyingly voiced stops are voiceless (and murmured), while in word-medial position they are voiced (cf. Cao & Maddieson 1992)

A Shanghainese tone sandhi domain can extend to include a function word (e.g. particle and nominal marker) that follows a compound word. This is not surprising if the domain is a prosodic constituent with "initial dominance" since function words are typically enclitics in analogous conditions in other languages. More worthy of note are the cases where the tone sandhi domain extends backwards to include the preceding Wh-marker as shown in Table 2.13. The unusual "backward extension" has led to the proposal that these forms are treated as single words by native speakers (Duanmu 1998:188-9). That is, the backward spread is an indication that the Wh-marker is a productive prefix, so that phrases beginning with the Wh-marker are "words", since in other cases, a string of sandhi tone morphemes contained within one domain is analyzed to be "a lexicalized item" in the mind of the native speakers of Shanghainese.

From Duanmu (1998:178-9):		
a.	(sa ɲĩ)#	
	which person	'who'
b.	(sa me'-z)	
	which thing	'what'
c.	(sa di-fã)	
	which place	'where'
d.	(sa zẽ-k^{w}ã)	
	which time	'when'

Tone was not shown in the paper, but the brackets indicate the occurrence of tone sandhi.

Table 2.13 Phrases proposed to be treated as single words: the lexicality of tone sandhi.

Yip (1980) uses the "foot" as a tonal association domain to account for tone sandhi in Shanghainese, in which there is the deletion of the tones of all but the initial syllable of a word. Yip posits that only the "Designated Terminal Element" is allowed to keep its tone within the foot (p. 103). That is, in the foot that is left-branching, the first syllable is the strong head of the foot, as shown in Figure 2.9.

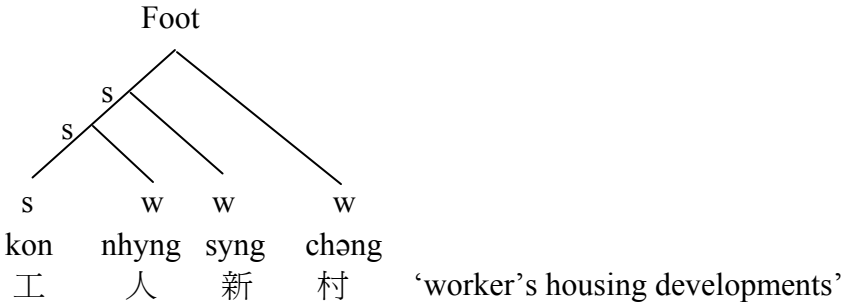


Figure 2.9 Foot structure for a Shanghainese word (Yip 1980:103).

2.6 The foot in Cantonese

The foot in Cantonese will differ from the foot that it has been invoked to describe in these four other Chinese phenomena in the following way. In the other languages, one can talk about culminative prominence alternations (i.e., "stress") or "dominance" (i.e. of the first syllable in Shanghainese), with the lexical tone specification being preserved only for the stronger or the strongest syllable. In Cantonese syllable fusion, by contrast, tone is the last thing to go (cf. e.g. Hashimoto 1972; Bauer & Benedict 1997; examples see Table 2.14), and there is not any sense in which it is possible to say that one of the two fused syllables is "dominant" or winning out over the other. At the same time, syllable fusion bears a strong resemblance to the "edge erasing" potential of segmental effects associated with neutral tone in Beijing Mandarin and probably of similar effects of tone sandhi domain in Shanghainese, where Zee has talked about devoicing of high vowels in non-first syllables, and where Cao & Maddieson (1992) have talked about the voicing of plosives domain medially for the initials of syllables that have low or low-rising tone in citation form.

Unlike other dialects of Chinese, we have not found reliable phonological markers in Cantonese that are comparable to "stress" (syllables with other than neutral tone) in Mandarin Chinese or the "dominant" tone in the tone sandhi domain as in Beijing Mandarin, Xiamen, Shanghainese, etc, that could be interpreted as markers for the grouping of phonological words.

As in other varieties of Chinese such as Beijing Mandarin and Shanghainese, there are segmental effects in the syllable fusion forms in Cantonese. Syllable fusion exhibits grouping that is phonetically and phonotactically identifiable. Phonetically, similar to neutral tone syllables in Beijing Mandarin where segmental lenition or even deletion in the case of [-ɹ] suffixation, there are segmental effects in syllable fusion in Cantonese so that monosyllabic word boundaries are blurred when syllables are fused in

Cantonese. Less extreme fusion shows segmental lenition or deletion, and more extreme fusion can simplify contour tones and "merge" the qualities of vowels that would be separated by an onset or coda consonant at more "normal" degrees of disjuncture between words. However, even at these more extreme junctures, fusion does not usually override the lexical tones of the syllables (Cheung 1986; Li 1986; Wong 1996; Wong et al. 2002; Lee 2003; Wong 2004; Wong, Chan & Beckman 2005). For example, in the case of tentative marker / intensive marker / quantifier 'jɐt'-deletion, the tone of the marker is often retained and attached to the preceding morphemes. This may or may not result in a tone change of the preceding syllable, though it has often been analyzed as cases of "tone change" (e.g. Yip 1980). Table 2.14 lists some examples. It is only in the most extreme cases (e.g. mɐt^{5} jɛː^{23} 'what' → mɛː^{55} 'what') where we see true "tone loss", and the "extremely fused" form is often lexicalized. That is, native speakers of the language seem to treat it as separate entry in their mental lexicon, and often represent it by another Chinese character. The fusion process, then, seems to be a gradient effect. Although it is possible that the most extreme cases of fusion may make the syllable count less determinate, fusion is not necessarily a categorical change from two syllables to one. Phonologically, incomplete merging of vowel qualities or even deletion of segments in a fused form could result in "illegal" segmental sequences outside of the syllable structure in the Cantonese phonology, which is analogous to Beijing Mandarin "final elision" as in /wo.mən/ (I + plural)' → [wom], which shows clearly the word status of the "underlying" bisyllabic form.

a.	Quantifier /jɐt^{5}/ (in reduplicated measure and verbs, and after the existential verb)		
	i.	jɐt5 lɔː21 jɐt5 lɔː21	‘basket by basket’
		jɐt5 lɔː21+5 lɔː21	
	or:	jɐt5 lɔː35 lɔː21	
	ii.	jɐu23 jɐt5 ts^{h}iː33	‘there’s a time…’
		jɐu23+5 ts^{h}iː33	
	or:	jɐu35 ts^{h}iː33	
b.	Tentative marker /jɐt5/		
	i.	maːt3 jɐt5 maːt3	‘give it a little wipe’
		maːt3+5 maːt3	
	or:	maːt35 maːt3	
	ii.	mɐn22 jɐt5 mɐn22	‘try and ask’
		mɐn22+5 mɐn22	
	or:	mɐn35 mɐn22	
c.	Intensive marker (in reduplicated adjectives)		
	i.	hoŋ21 jɐt5 hoŋ21	‘very red’
		hoŋ21+5 hoŋ21	
	or:	hoŋ35 hoŋ21	
	ii.	muːn23 jɐt5 muːn23	‘very full’
		muːn23+5 muːn23	
	or:	muːn35 muːn23	

Table 2.14 Deletion of segments, but tone remains in Cantonese (adapted from Yip 1980:52-3)

In summary, then, syllable fusion in Cantonese could be a revelation of some underlying prosodic motivation in Cantonese that might be identifiable by the Cantonese listeners. This dissertation is going to explore how syllable fusion relates to and thus reveals the foot structure in Cantonese, how the foot structure is mapped onto the word structure, and what are the factors that would condition the mapping.

CHAPTER 3

EXPERIMENT I: TESTING SPEECH RATE

Contrary to the past literature which characterizes syllable fusion in Cantonese as "contraction", suggesting categorical changes of syllable fusion forms, we observed that syllable fusion between two syllables exhibits a very fine-grained continuum of segmental and suprasegmental blending. Less extreme changes include assimilation, consonant lenition, (partial) voicing of voiceless consonants, vowel reduction, any substantial weakening or effective deletion of the oral gesture(s) of the segment(s) contiguous to the syllable boundary, and the sometimes attendant resyllabifications that create "fused forms". More extreme fusion can simplify contour tones and "merge" the qualities of vowels that would be separated by an onset or coda consonant at more "normal" degrees of disjuncture between words. Syllable fusion in Cantonese thus seems to be a prosodically driven phenomenon.

This dissertation defines the term prosodically driven in terms of the things that shape prosodic structure as something linked to, but independent of, for example, morphosyntactic structure. One might then expect a potential factor that affects syllable fusion (and hence the prosodic structure) would affect the phenomenon in a way that varying the parameter of the potential factor would give rise to a varying amount and/or degrees of syllable fusion. Demonstrating a speech rate effect then can be interpreted as a clear indication that the syllable fusion process is prosodically driven (and not simply morphosyntactic).

This chapter describes an experiment that tested the influence of speech rate on syllable fusion using a "pseudo-sentence" repetition task in this experiment. The data were recorded originally in 1996 as part of an earlier study (Wong 1996). in which I noted only whether or not an extreme form of fusion had occurred. For the current re-analysis, I transcribed the repetitions and re-categorized the productions using a two-point scale that distinguishes between fusions that maintain two clear vocalic nuclei and fusions that merge the vowels of the target disyllabic string. I also measured speech rate for each utterance as a whole. As this chapter describes, there was a strong correlation between the percentage of fused forms and the average speech rate in a regression across the 32 speakers who were recorded.

3.1 Method

3.1.1 Materials

The stimuli for the experiment were utterances of "pseudo-sentences" made in 1996 for the study I reported in Wong (1996), in which I studied the influence of speech rate on syllable fusion. The target "pseudo-sentences" were created by recording myself, the experimenter and a native Hong Kong Cantonese, uttering those target "pseudo-sentences" in the simultaneous interpretation laboratory of the Department of Chinese and Bilingual Studies at the Hong Kong Polytechnic University.

The selection criteria for the target words were as follows. First, fusion forms of these words should have been observed in prior studies. The potential fusion forms for this experiment were selected from Cheung (1986:242-9) and his subsequent collection. (The Hong Kong Cantonese Adult Corpus (HKCAC) was not available at the time of the experiment.) Second, they were judged not to be highly lexicalized. For example, the realization [mɛː^{55}] or [mɛː$^{55+3}$] for /mɐt^{5}jɛː^{23}/ 'what' may be said to be lexicalized, as suggested by the fact that the Hong Kong Cantonese speakers would represent it with a

single character 咩. (This phenomenon can also be seen in the orthographic transcription of the HKCAC, Leung & Law 2001.) Also, not every native Hong Kong Cantonese speaker would have the knowledge that [mɛː^{55}] or [mɛː$^{55+3}$] is derived from /mɐt^{5}jɛː^{23}/ 乜嘢 'what'. Third, the target words should be able to build the story for the story-telling task in Wong (1996). That is, the forms elicited to study the influence of speech rate on syllable fusion was half of a larger study for Wong (1996), the other half of which involved a story-telling task using the same target words, which is not reported here.

The selected fusion targets covered a variety of short and long words, including monomorphemic words (e.g. 但係 /taːn^{22}hɐi^{22}/ 'but', 即刻 /tsek^{5}hɐk^{5}/ 'immediate'), compounds comprising two (proper) nouns in a modifier-head relationship (e.g. 超級市場 /ts^{h}iːu^{55}k^{h}ɐp^{5}siː^{23}ts^{h}œːŋ^{21}/ 'supermarket', 收音機 /sɐu^{55}jɐm^{55}kei^{55}/ 'radio', 司徒拔道 /siː^{55}t^{h}ou^{21}pɐt^{2}tou^{22}/ 'Stubbs Road'), compounds comprising a verb and a noun in a verb-complement relationship (e.g. 返屋企 /faːn^{55}ok^{5}k^{h}ei^{35}/ 'go home').

Four "pseudo-sentences" with two different sequences of tempo requirements were written on two note cards. Before the actual listening and recording started, the experimenter said the "pseudo-sentences" to the participants followed by saying the tempo instructions "very fast" or "normal" for demonstration purpose. The participants repeated the utterance of "pseudo-sentence" after hearing the tempo prompt. The note cards are shown in Figure 3.1.

The actual test materials then were 45 short or long words, each consisting of two to four syllables. Among the 45 words, 30 were fusion targets while 15 were fillers. Fillers were short or long words that were not in Cheung (1986:242-9) and his subsequent collection.

These two- to four-syllable words were strung together into meaningless “pseudo-sentences” so that each “sentence” would contain three words. Each of these “sentence” then consisted of seven syllables, except the last two, where there were eight. There were a total of 15 “pseudo-sentences”.

The stimulus utterances were recorded in two blocks: the first block had seven utterances (hence, utterances 1-7) and the second block eight (hence, utterances 8-15). The sequence of tempi requirement for the first block was the reverse of that of the second block. Thus, for the first list of utterances, which comprised two blocks of utterances, the sequence of tempi specified for utterances 1-8 was “very fast” and then “normal”, while for utterances 9-15, “normal” and then “very fast”. Each utterance had the same tempo sequence instruction. Three pseudo-utterances for practice preceded each block. The second list of utterances, which comprised the same two blocks of utterances as in the first list, differed from the first list of utterances only in the sequence of tempi requirement for the subjects. Thus, for the second list of utterances the sequence of tempi specified for utterances 1-8 was “normal” and then “very fast”, while for utterances 9-15, the sequence was “very fast” and then “normal”.

The two lists were recorded on two TDK Type D cassette tapes; thus, test tapes A and B. The order of the utterances on test tape A is shown in Appendix A. (Test tape B contained the second list of pseudo-utterances, with the sequence of tempo instructions being the reverse of that in test tape A for each of the two blocks of pseudo-utterances.) Syllables in the stimulus utterances on the test tapes were produced in non-fused forms, paced by a metronome at the speed of 96 (equivalent to 88 beats per minute, or 1 beat = 0.68 second). (The sound of the metronome was not recorded on the test tapes.) Two syllables made one beat. Each utterance on the tapes was followed by instructions “好快” /hou^{35}faːi^{33}/ (very fast) and “正常” /tseŋ^{33}sœːŋ^{21}/ (normal) to prompt the speech tempo requested of the subjects. Each instruction was followed by a pause, where subjects were to repeat the utterance just heard. A two-beat pause was provided on the tapes for

subjects to repeat the utterance after the instruction “very fast”; for “normal”, a four-beat pause was provided. The duration of the entire set of utterances was 3.29 minutes. After utterance 7 on both tapes, there was a change in the sequence of tempo requirements for the subjects with a new set of three practice trials before test trials 8-15 were presented. A ten-beat pause was inserted after the space for the repetition of utterance 7 so that the experimenter could remind subjects of the change. The entire set of “pseudo-sentences” are given in Appendix A.

Note card 1	English translation of note card 1
好快 - 正常	very fast (f) – normal (n)
非常 鬧鐘 大學生 (好快) - (正常)	very alarm clock university students (f) - (n)
汽水 行路 塗改液 (好快) - (正常)	soda to walk white-out (f) - (n)

Note card 2	English translation of note card 2
正常 - 好快	very fast – normal
手錶 食飯 羽毛球 (正常) - (好快)	watch eat dinner badminton (n) - (f)
波鞋 原來 釘書機 (正常) - (好快)	sports shoes originally stapler (n) - (f)

Figure 3.1 Pre-test utterances for demonstration.

3.1.2 Subjects

32 speakers, 16 males and 16 females, participated in the experiment. All of the speakers, except one, who was a postgraduate, were university undergraduate students at the time of the experiment. The speakers ranged between 18 and 26 in age. All were native speakers of Hong Kong Cantonese born and raised in Hong Kong. The subjects were not paid but were given a beverage upon completion of the experiment as a token of thanks for their participation.

3.1.3 Procedures

The experimenter illustrated the repetition task to each of the subjects the repetition task by reading out the pseudo-sentences and prompts in a simulation of the actual task. That is, after reading one sentence, she said the first tempo prompt. The subject then repeated the pseudo- sentences as heard using the tempo requested. When the subject finished repeating the pseudo-utterance, the experimenter said the second tempo prompt to the subject. The subject then repeated the pseudo-utterance using the tempo requested. If the subject did not make much difference between the fast and the normal repetitions, the experimenter would ask the subject to try to repeat the pseudo-utterance at a faster rate for the fast tempo. If the subject did speed up his/her speech rate, the experimenter reinforced the subject by saying something like: "That's right. You can repeat the pseudo-utterances as fast as how you just did it when you're actually listening to the utterances."

For the actual test, subjects listened to the test materials on the tapes over earphones and were recorded in a quiet, empty office in the company of the experimenter. Half of the subjects listened to tape A and half to tape B. Subjects repeated the utterances at the order of tempi requested after they heard each utterance. Subjects did simple arithmetic (see Appendix B) while repeating the utterances. They could handle the dual-task quite successfully. That is, they very rarely forgot or misremembered words they needed to repeat.

3.1.4 Analyses

3.1.4.1 Measuring speech rate

Speech rate was defined as the average number of syllables per second. Rate was calculated by counting the total number of "underlying" syllables elicited in each of the tempo conditions. The total number of "underlying" syllables elicited in each of the

tempo conditions was then divided by the summed durations of each speaker's repetitions at that tempo, excluding production errors. Thus, two rates were calculated for each subject.

"Underlying" syllables in this dissertation refer to the citation forms that would be produced by the general native Hong Kong Cantonese speakers. These citation forms might differ from forms that are given in Cantonese pronunciation dictionaries such as Wong ([1941], 1997) because there are synchronic alternations among the native Hong Kong Cantonese speakers. Section 3.1.4.2 below describes the synchronic alternations that are relevant to this experiment, and describes what we counted as syllables elicited.

The number of syllables elicited should have been constant across speakers, since the same materials were used for all speakers. However, a small number of production errors were fluent misproductions of the target materials, perhaps due to mishearing or misremembering. These were excluded from both the numerator and the denominator of the rate equation. The majority of production errors were different. They were self-corrected ones.

Production errors, which were very often followed by hesitation pause of various durations, range from vocalization of just one segment to vocalization of one or more syllables with abrupt cut-off in the middle, where speakers might self-correct their productions. The rationale for excluding these error and hesitation pause sequences is that they often prolonged the overall utterance duration, obscuring the fact that the speakers produced the targets or fillers at the requested rate before and after the error and hesitation pause sequences. Therefore, the duration of these error and hesitation pause sequences, marked from the onset of the production errors through the onset of the following target item or filler, was subtracted from the duration of the entire utterance in measuring speech rate. See Figure 3.2 for an illustration.

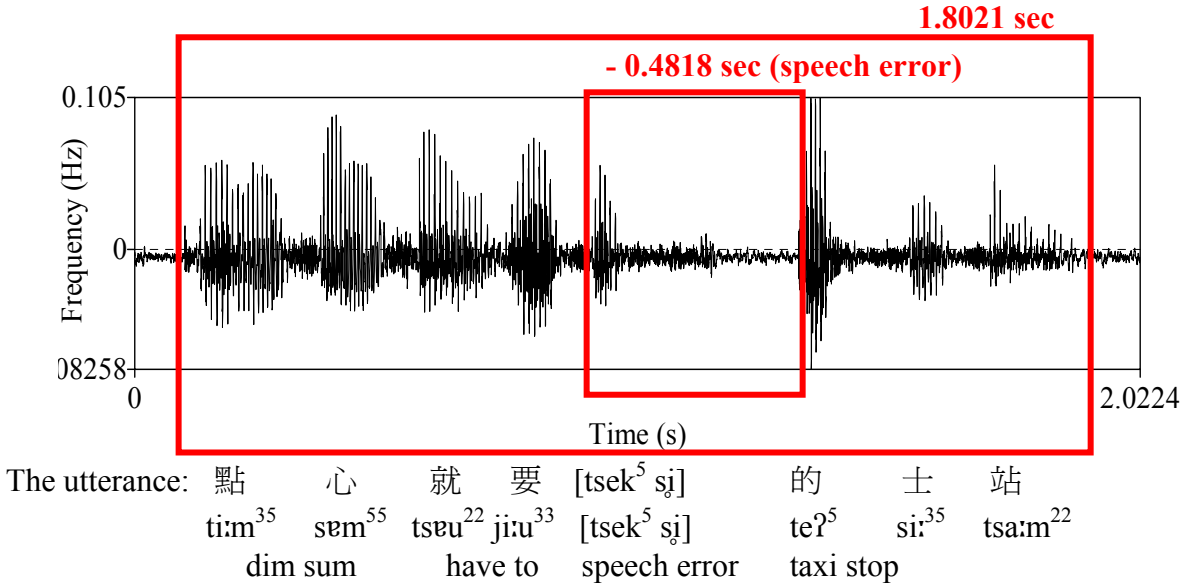


Figure 3.2 The duration of speech error is subtracted from the total duration of an utterance in the calculation of speech rate (syllables per second).

3.1.4.2 Counting elicited fusion targets

Phonologically conditioned and lexically conditioned segmental alternations are found in present-day Hong Kong Cantonese. The subjects might choose to repeat the test utterances with the alternative forms that they preferred. Thus, their productions could be different from the form they were presented aurally, though the different forms actually represented the same word.

A set of criteria was established in this study to make sure that the tokens we counted were the tokens that the subjects repeated as heard. The set of criteria took into consideration phenomena found in present-day Hong Kong Cantonese, with relevance for this study include (a) synchronic alternation between initial [ŋ] and null initial,

between [k^w, k^{wh}] and [k, k^h] initials respectively before [ɔː][1] (cf. Hashimoto 1972; Yeung 1980; Cheung 1986; Bauer and Benedict 1997); the more prevalent substitution of final [t] for final /k/ than vice versa (cf. e.g. Cheung 1999; Zee 1999a; Law, Fung & Bauer 2001; Wong 2002); (b) long/short vowel alternations that are related to literary versus colloquial styles of reading for certain lexical items (e.g. 即刻 [tsek^5 h**ɐ**k^5~ h**aː**k^5] 'immediate'); (c) lexically conditioned segmental alternations (e.g. 佢話 [**k^h**ɵy^{23}~**h**ɵy^{23} waː^{22}] 's/he said'; 朝頭早 [tsiːu^{55} **t^h**ɐu^{21}~**h**ɐu^{21} tsou^{35}] 'morning'); and (d) deletion of place of articulation of final /t/ and final /k/ (though much less so for final /p/), retaining just a certain amount of glottalization or a glottal stop.

The productions of the subjects showed one or more of the above types of phenomena, even though they heard the stimuli on the tapes before they repeated them. Many production tokens would have to be unnecessarily discarded if we had not set criteria that address the above phenomena. The criteria apply to the fusion targets (and fillers). For target forms that were *not fused* to be counted towards the total number of tokens elicited of each subject, the following set of criteria, which took into account (a) through (d) above, must be fulfilled for each syllable in the fusion targets.

(i) The onset consonant was produced in the form given on the test tapes.

(ii) The onset consonant was not produced in the form given on the test tapes, but it could be ascribed to synchronic alternations in that position (cf. (a) above).

(iii) The vowel was produced in the form given on the test tapes.

(iv) The vowel was not produced in the form given on the test tapes, but it could be ascribed to stylistically governed or lexically conditioned segmental alternations (cf. (b) and (c) above).

(v) The coda consonant was one of the final consonants [p, t, k, m, n, ŋ] or [ʔ] (cf. (d) above).

[1] There are also synchronic alternations between [n] ~ [l] initial consonants for initial /n/. Although the [n] ~ [l] alternation is not relevant for the list of items in the repetition task in this experiment, it would be relevant for experiments II and III.

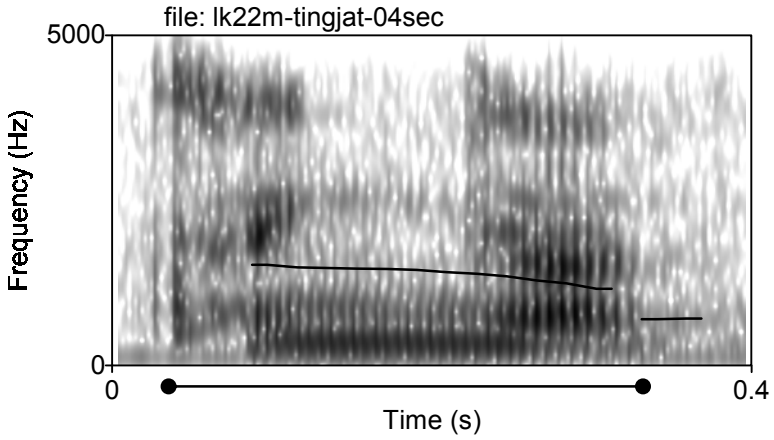
(vi) The lexical tone was produced in the form given on the test tapes.

(vii) The lexical tone was not produced in the form given on the test tapes, but it could be interpreted as having tonal coarticulation, or tonal target undershoot due to, for example, intonation phrase-final effects, etc.

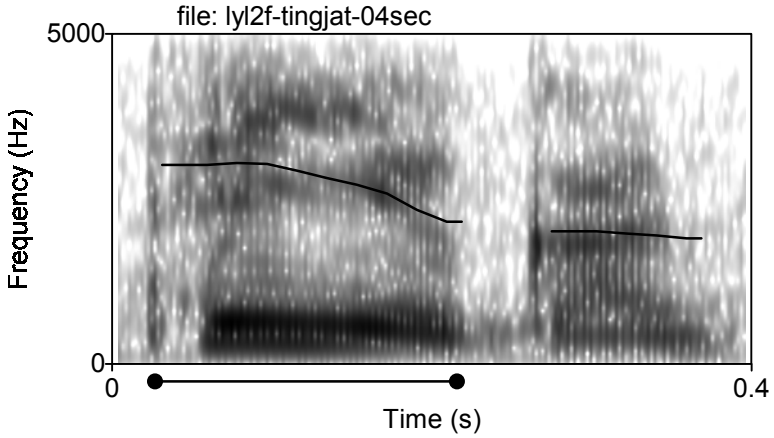
Tokens that did not violate any of the above criteria were non-fused tokens; syllable fusion was evaluated based on the above criteria. Both non-fused forms and fused forms were counted towards the total number of elicited fusion targets for each subject. The total count of fusion targets elicited is subject-dependent, since the number of tokens successfully elicited differed across all subjects.

3.1.4.3 Selecting two degrees of fusion

Two degrees of fusion were noted. First, there is the deletion of at least one segment contiguous to the syllable boundary between two syllables without affecting the vowel count. I term these forms as resulting in a "bisyllabic fusion". Second, there is a "merging" of vowel qualities of adjacent syllables to produce a single intermediate quality or deletion of one of the vowels, with or without tone merger or tone deletion. This "vowel coalescence" result could be interpreted as changing the syllable count. This is the motivation for the otherwise arbitrary choice of these two degrees of fusion in the fusion continuum. Figure 3.3 illustrates the continuum and the two "degrees". Fused forms (a) through (c) in Figure 3.3 were counted into the first category, whereas fused form (d) was counted in the second category.



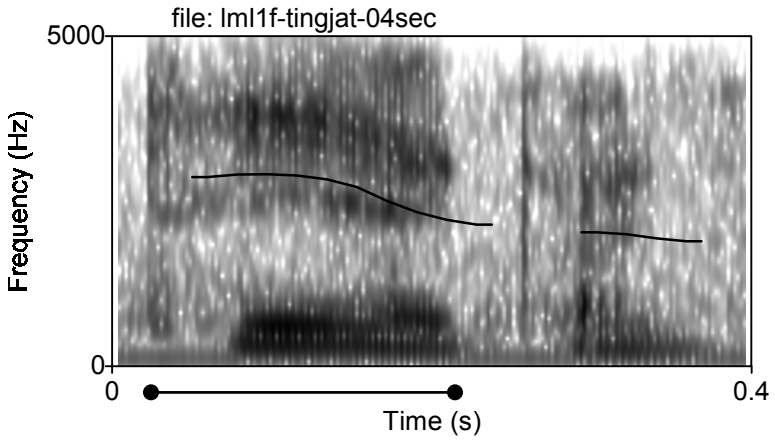
a. /t^{h}eŋ^{55} jɐt^{2}/ ‘tomorrow’ fused to become [t^{h}eŋɐ$^{55+2}$].



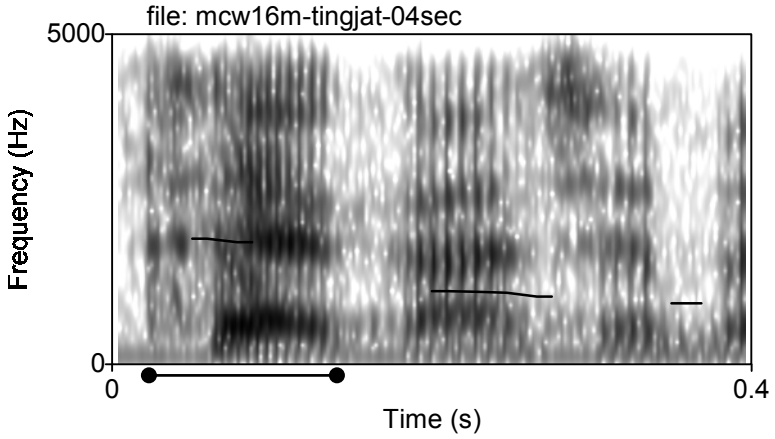
b. /t^{h}eŋ^{55} jɐt^{2}/ ‘tomorrow’ fused to become [$t^{h}\tilde{\varepsilon}^{55}$jɐ^{2}].

Figure 3.3 Four degrees of fusion for the word 聽日 /t^{h}eŋ^{55}jɐt^{2}/ ‘tomorrow’ produced by four speakers. The four tokens are placed on a constant time scale. “●—●” indicates the duration of the word ‘tomorrow’. Notice that the word becomes progressively shorter as more segments are deleted and the vowels are merged. Nevertheless, the tones are not deleted. See the F0 tracks, which all show a falling contour from a high pitch target to a low pitch target.

Figure 3.3 continued



c. /t^{h}eŋ^{55} jɐt^{2}/ ‘tomorrow’ fused to become [$t^{h}\tilde{\varepsilon}$ɐ$^{55+2}$].



d. /t^{h}eŋ55 jɐt^{2}/ tomorrow’ fused to become [$t^{h}\varepsilon^{55+2}$].

Figure 3.3 ended

3.2 Results and discussion

Figure 3.4 shows the speech rate of the subjects. Each data point in the figure represents the number of syllables that was produced per second of a given subject. The figure shows that all subjects, except one, were successful in varying their speech rate. The cloud of the data points lies below the x = y line. That is, the subjects repeated the utterances slower when they were instructed to speak at their normal rate, but faster when they were instructed to speak as fast as they could.

In the fast rate condition, out of a total of 480 test utterances that were presented to all of the 32 subjects, 457 utterances were successfully elicited. 26 out of 457 utterances (or 5.7%) had error and hesitation pause sequences. Duration excluded from these 26 utterances ranged from 0.13 to 1.07 seconds per utterance, with the average being 0.42 second. Thus, the fast rate condition has an average speech rate of 1.00 second per utterance across all subjects after the exclusion of the error and hesitation pause sequences. In the normal rate condition, out of a total of 480 test utterances that were presented to all of the 32 subjects, 458 utterances were successfully elicited. 13 out of the 458 utterances (or 2.8%) had error and hesitation pause sequences. Duration excluded from these 13 utterances ranged from 0.18 to 0.91 second per utterance, with the average being 0.52 second. Thus, the normal rate condition has an average speech rate of 1.43 second per utterance across all subjects after the exclusion of error-hesitation pause sequences.

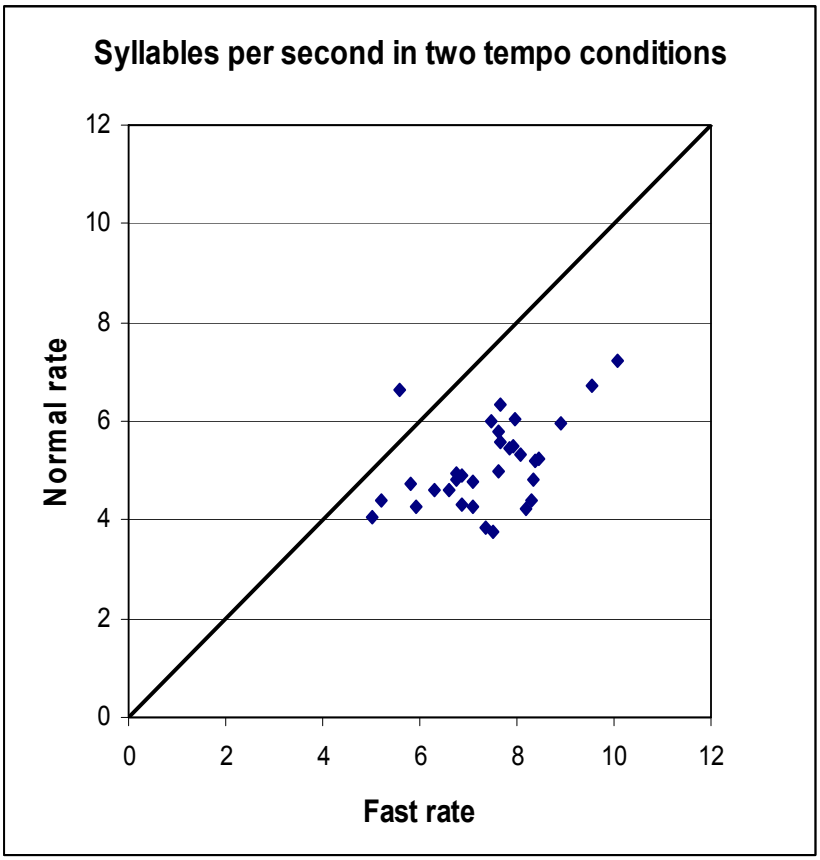


Figure 3.4

Tables 3.1 and 3.2 show the number of fusion targets elicited for each of the 32 subjects at the two tempo conditions. In both the fast and the normal rate conditions, out of the total of 30 fusion targets presented aurally to each of the 32 the subjects, the number of fusion targets elicited ranged from 25 to 30 across the 32 speakers, with the average number of fusion targets elicited being 28 (or 93% of the total number of fusion targets). The fusion targets ranged from 2 syllables long to 4 syllables long.

Tables 3.3 and 3.4 show the number of syllables (both target and foil included) elicited for each of the 32 subjects at the two tempo conditions. In the fast rate condition, out of the total of 107 syllables presented aurally to each of the 32 the subjects, the

number of syllables elicited ranged from 91 to 107 across the 32 speakers, with the average number of fusion targets elicited being 103 (or 96% of the total number of target syllables). In the normal rate condition, out of the total of 107 syllables presented aurally to each of the 32 the subjects, the number of syllables elicited ranged from 92 to 107 across the 32 speakers, with the average number of syllables elicited being 102 (or 95% of the total number of target syllables). Hence, the results showed that the subjects were successful in handling the simultaneous tasks of repeating the test utterances presented to them and the simple arithmetic, and that the majority of the syllables were elicited.

Subject	No. of fusion targets elicited
7	25
17	25
12	26
19	26
14	27
29	27
30	27
6	28
11	28
15	28
16	28
20	28
23	28
24	28
25	28
26	28
27	28
28	28
1	29
9	29
21	29
31	29
2	30
3	30
4	30
5	30
8	30
10	30
13	30
18	30
22	30
32	30

Table 3.1 Number of targets elicited for repetition at fastest possible rate.

Subject	No. of fusion targets elicited
24	25
12	26
19	26
26	26
9	27
15	27
20	27
29	27
30	27
5	28
6	28
11	28
14	28
17	28
23	28
25	28
27	28
28	28
16	29
31	29
32	29
1	30
2	30
3	30
4	30
7	30
8	30
10	30
13	30
18	30
21	30
22	30

Table 3.2 Number of targets elicited for repetition at normal rate.

Subject	No. of syllables elicited
17	91
19	92
7	95
12	99
6	100
11	100
14	100
15	100
21	100
23	100
24	100
29	100
30	100
31	100
9	103
1	107
2	107
3	107
4	107
5	107
8	107
10	107
13	107
16	107
18	107
20	107
22	107
25	107
26	107
27	107
28	107
32	107

Table 3.3 Number of syllables elicited for repetition at fastest possible rate.

Subject	No. of syllables elicited
19	92
26	92
24	93
5	99
9	99
12	99
17	99
6	100
11	100
14	100
15	100
20	100
23	100
25	100
27	100
28	100
29	100
30	100
31	100
1	107
2	107
3	107
4	107
7	107
8	107
10	107
13	107
16	107
18	107
21	107
22	107
32	107

Table 3.4 Number of syllables elicited for repetition at normal rate.

Each data point in Figures 3.5 and 3.6 represents the proportion of fusion forms for each speaker — i.e. the number of fusions (Figure 3.5) or the number of coalesced vowels (Figure 3.6), divided by the number of fusion targets that were not produced with speech errors. Figure 3.5 shows that syllable fusion is highly predictable by speech rate: the faster the speech rate, the more the occurrence of fused forms [r = .854, p =.000, n = 64]. This result may not surprise many, since oral gestures/target shooting is time-linked. Of interest may be how properties of the participating syllables could be merged or changed. Figure 3.6 shows that syllable fusion with participating vowels coalesced is well-predicted by speech rate: the faster the tempo, the more the fusion forms with participating vowels coalesced [r = .759, p = .000, n = 64].

In the fast rate condition in the two figures the average number of syllables per second across all 32 subjects is 7.40. The total number of fusion targets elicited for all 32 subjects is 897 (or 28 targets per subject in average). Among the elicited fusion targets, 546 (or 61%) exhibit bisyllabic fusion; 159 (or 18%) of the elicited fusion targets exhibit vowel coalescence. In the normal rate condition in the two figures the average number of syllables per second across all 32 subjects is 5.11. The total number of fusion targets elicited for all 32 subjects is 902 (or 28 targets per subject in average). Among the elicited fusion targets, 203 (or 23%) exhibit bisyllabic fusion; 31 (or 3%) of the elicited fusion targets exhibit vowel coalescence.

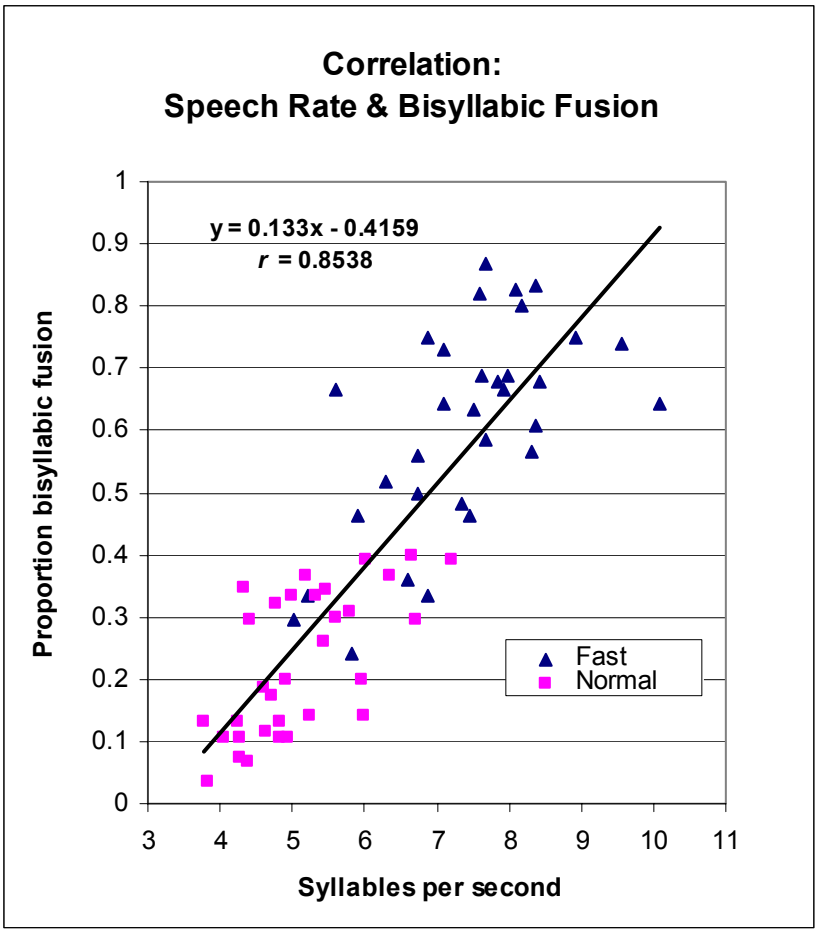

Correlation:
Speech Rate & Bisyllabic Fusion
y = 0.133x - 0.4159
r = 0.8538
Proportion bisyllabic fusion
Syllables per second
Fast
Normal
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
3
4
5
6
7
8
9
10
11

Figure 3.5

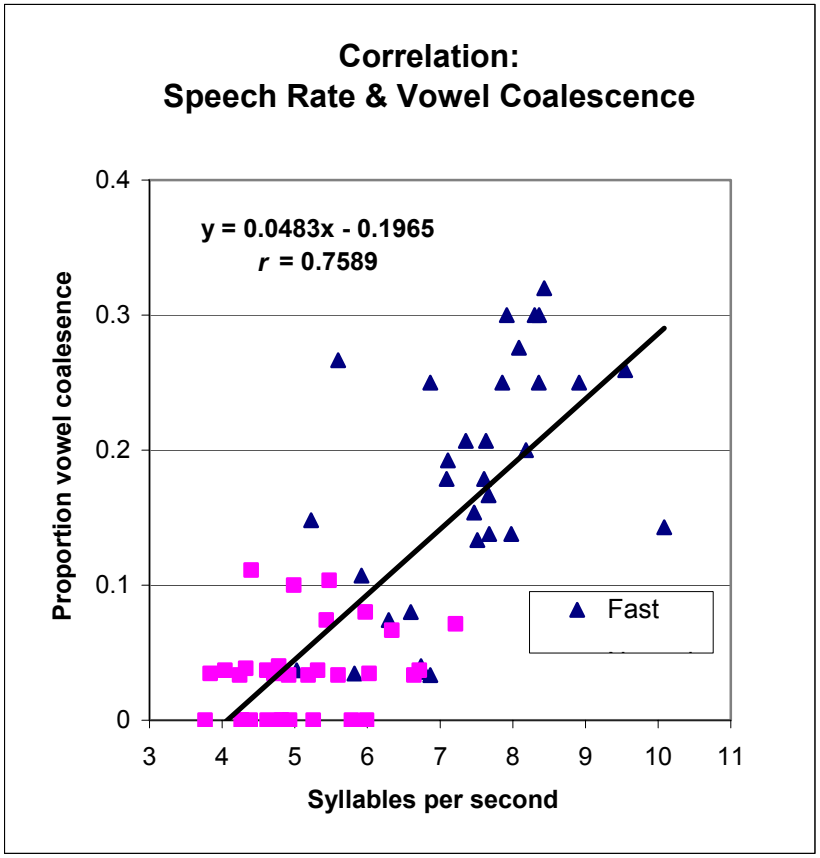
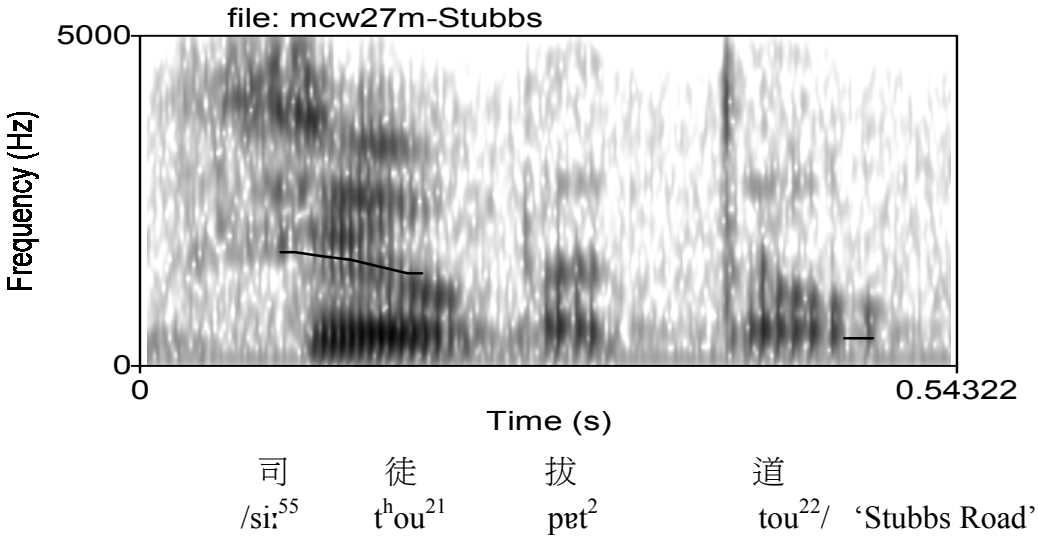
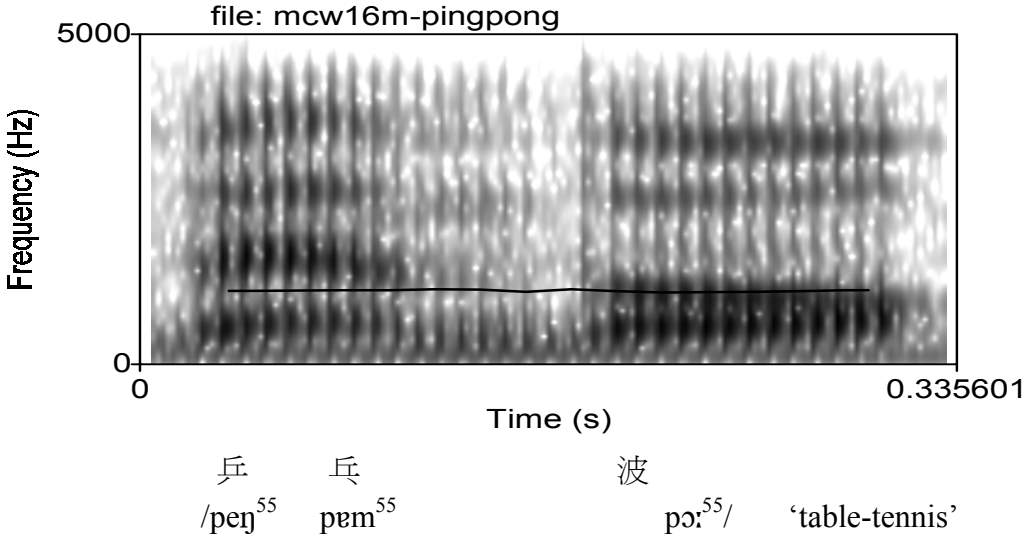


Figure 3.6

An examination of the fused forms shows that fused forms can "violate" the standard Cantonese syllable shapes and phonotactics. Figure 3.7 shows examples of the fused forms of three words 司徒拔道 /siː^{55}t^{h}ou^{21}pɐt^{2}tou^{22}/ 'Stubbs Road, 鍾意 /tsoŋ^{55}jiː^{33}/ 'to like', and 乒乓波 /peŋ^{55}pɐm^{55}pɔː^{55}/ 'table-tennis' produced by three speakers. That is, "triphthong" can occur within a syllable (e.g. [s**iou**] 'Stubbs'), labial onset and coda co-occur within a syllable (e.g. [**p**ɛː**m**] 'table-tennis'), and diphthong that is not in the standard Cantonese diphthongal inventory arises (e.g. [ts**yi**] 'like').



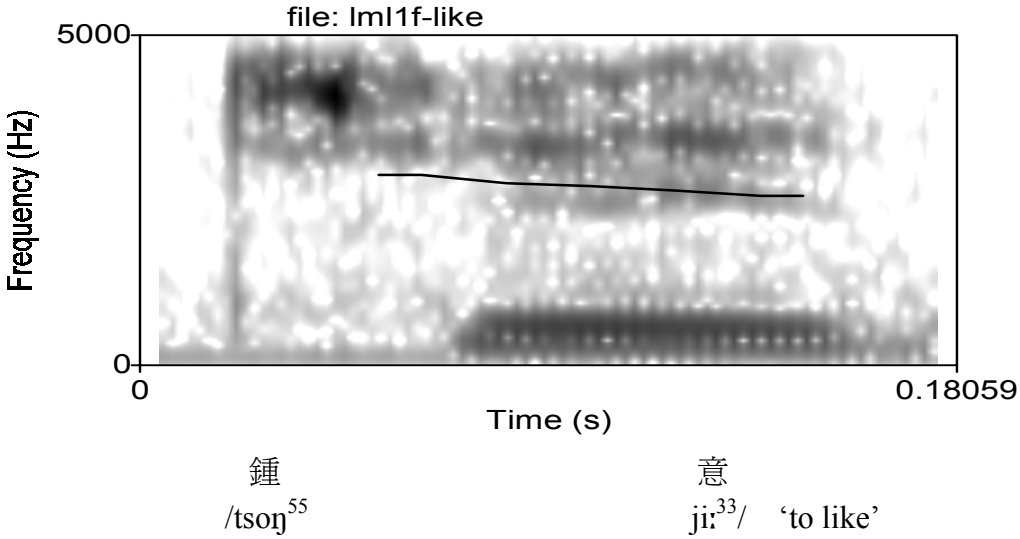
(a) siou55+21pɐt2tou22



(b) pɛm55+55pɔː55

Figure 3.7 Examples of fused forms for the words 'Stubbs Road', 'table-tennis', and 'to like' produced by three speakers. Notice the "violation" of the standard Cantonese syllable shapes and phonotactics in the productions.

Figure 3.7 continued



(c) tsyi55+33

Figure 3.7 ended

While the citation monosyllabic flavor is one major characteristic of Cantonese as described in the past literature (see e.g. Chao 1947), results in this experiment showed varying in speech rate would stretch or shrink the duration of a string of citation (or non-fused) monosyllables. Moreover, accompanying with the varying of speech rate are segmental effects along an extremely fine-grained continuum that can be observed. For example, duration of segments would be changed, consonant(s) and/or vowel(s) contiguous to the syllable boundary could be lenited, or even deleted, and participating vowels and tones could be merged. Results of this experiment, therefore, showed that syllable fusion in Hong Kong Cantonese is clearly prosodically driven. Speech rate is a good predictor for syllable fusion, both in terms of the amount and the degree of fusion. Moreover, syllable fusion not only occurs in fast speech, but also in normal rate of speech.

CHAPTER 4

PREPARATION OF MATERIALS FOR EXPERIMENTS II AND III

Having a list of words that was encoded with word frequency for spoken Cantonese and morphosyntactic relationship was crucial for running Experiments II and III. However, no list of words was readily available for these research purposes. Therefore, it was necessary to create a database of segmented word forms to calculate frequencies and also to extract word types to be tagged for internal structure. To create this database, word segmentation was done on the Hong Kong Cantonese Adult Language Corpus (HKCAC), a spontaneous speech corpus developed by Leung and Law (2001). This process of word segmentation defines what is a word for the current dissertation, provides spoken word token frequency data, and yields a list of word types for testing. The word types selected for use in the two experiments were given to three linguists for morphosyntactic analysis to obtain their word-internal branching structures.

This chapter describes the spontaneous speech corpus from which the test words were obtained. It also explains how the word segmentation was done and it lays out the word segmentation principles that were developed. Descriptions of how words were chosen to represent different word frequency categories and the process of assigning word-internal branching structures to the candidate test words are also given in this chapter.

4.1 Word segmentation for spoken Cantonese

4.1.1 Challenges of word segmentation for Cantonese and for Chinese in general

The question "what is a word" is not easy to answer for any Chinese varieties, in particular, for Cantonese[1]. The reason for this trickiness is partly in the writing system. Orthographically, each syllable is represented by a character (or 漢字 *hanzi*) in Chinese. While languages such as English, French, Greek, and Korean use white spaces to signal word boundaries[2], written forms of the Chinese language, including Cantonese, do not use white spaces to delimit words (see for example the names of radio programs in Table 4.1). In other words, although word segmentation is a real issue for all languages, at least in languages such as English, we can use a rough and ready definition of "word" as "the stuff separated by white space". But this option is unavailable for Chinese.

Morphologically, Cantonese is one of the Chinese dialects that has a strong one-to-one correspondence between the syllable and the morpheme. The vast majority of Cantonese morphemes can stand alone as simple words, while they can also be strung together to form compound words. Moreover, many verb-object compound words can be made discontinuous by morphosyntactic processes. For example, the words 瞓覺 /fɐn^{33}kaːu^{33}/ 'to sleep' (literally, 'to sleep a sleep') and 跳舞 /t^{h}iːu^{33}mou^{23}/ 'to dance' (literally, 'to dance a dance') can be made discontinuous by inserting a quantifier and/or a classifier to become **瞓**兩個鐘頭**覺** /**fɐn^{33}** lœːŋ^{23} kɔː^{33} tsoŋ^{55} t^{h}ɐu^{21} **kaːu^{33}**/ 'to sleep for

[1] The question is not easy to answer for other languages such as Japanese, Korean, Greek, and even English.

[2] I do not intend this statement to mean that the use of delimiters such as white spaces resolves ambiguity of wordhood in the languages mentioned. In English, for example, the decision of whether *choo-choo train* makes one (compound) word, two words, or three words, whether *data base* is one (compound) word (*database, data-base*), or two words (*data base*) may not be as consistent across the native speakers as we might have imagined (see Xia 2000: 4 for more examples). Rather than defining the boundaries of a word, the convention of using white spaces may be simply a means for reflecting probable word boundaries. The use of white spaces nevertheless has made defining what is a word for, say, English, relatively less difficult than defining what is a word for Chinese.

two hours' (literally, 'to sleep two hours sleep') and 跳一隻舞 /**t^hiːu^{33}** jɐt^5 tsɛːk^3 **mou^{23}**/ 'to dance one dance' (literally, 'to dance one classifier dance'). One might have already observed in these above examples that word-internal structures in Cantonese resemble phrasal structures. These characteristics of Cantonese make the determination of a word boundary fuzzy. Native speakers, therefore, often find it hard to judge whether a given morpheme is a bound or a free morpheme.

A partial solution to the above problem is to use the "committee consensus" decisions that are codified in dictionaries such as 廣州話方言詞典 /k^{w}ɔːŋ^{35} tsɐu^{55} waː^{35} fɔːŋ^{55} jiːn^{21} ts^{h}iː^{21} tiːn^{35}/ [Dictionary of the Guangzhou Dialect] by Rao et al. (1991). However, this is only a partial solution because of the problem of out-of-vocabulary (or "unknown") words. For example, 港式廣州話詞典 /kɔːŋ^{35} sek^5 k^{w}ɔːŋ^{35} tsɐu^{55} waː^{35} ts^{h}iː^{21} tiːn^{35}/ [Hong Kong Cantonese Dictionary] by Cheung and Ngai (1999) lists the compound 政府工 /tsiːŋ^{33}fuː^{35}koŋ^{55}/ 'government job' as a word entry, but it does not list very similar trisyllabic strings such as **政府錢 /tsiːŋ^{33}fuː^{35}**ts^{h}iːn^{35}/ 'government money' (as in 用政府錢/joŋ^{22}tsiːŋ^{33} fuː^{35}ts^{h}iːn^{35}/ 'use government money'), or **學校**工 /**hɔːk^{2}haːu^{22}**koŋ^{55}/ 'academic job' (as in 打學校工 /taː^{35}hɔːk^{2}haːu^{22}koŋ^{55}/ 'do an academic job'). When encountering the string 用政府錢 /joŋ^{22}tsiːŋ^{33} fuː^{35}ts^{h}iːn^{35}/ 'use government money'), then, one has to decide whether 'government money' forms one longer compound word or whether it should be analyzed as two or more shorter words. While the analogy to 政府工 /tsiːŋ^{33}fuː^{35}koŋ^{55}/ 'government job' might argue for the first analysis, the fact that 錢 /ts^{h}iːn^{35}/ 'money' and 工 /koŋ^{55}/ 'job' both can stand alone as words just as readily argues for the second. Decisions often times are not easy to make. Factors such as collocation frequency of morphemes and the number of syllables a sequence involves (native speakers tend to judge polymorphemic sequences that have

five or more syllables as having at least two words) are sometimes in conflict, which make it hard for the native speakers to determine the wordhood of a given string of morphemes.

Another example is the string of morphemes 工作日 /koŋ^{55}tsɔːk^{3}jɐt^{2}/, which could be analyzed as one long compound word 工作日 /koŋ^{55}tsɔːk^{3} jɐt^{2}/ 'work day' or two words 工作 /koŋ^{55}tsɔːk^{3}/ 'working' and 日 /jɐt^{2}/ 'day'. Interestingly, the string is highly unlikely to be analyzed as 工 /koŋ^{55}/ and 作日 /tsɔːk^{3} jɐt^{2}/, or as three independent words, even though each of the individual syllable-morphemes within the string can stand alone as a word. To understand why the latter two analyses are strange, one can compare the string with the occurrence of /jɐt^{2}/ in other forms, such as 賽馬日 /**ts^{h}ɔːi^{33}maː^{23}** jɐt^{2}/ 'horse-racing day' and 開放日 /**hɔːi^{55}fɔːŋ^{22}** jɐt^{2}/ 'open day', and the occurrence of /koŋ^{55}tsɔːk^{3}/ in other forms such as 工作室 /**koŋ^{55}tsɔːk^{3}** sɐt^{5}/ 'work room', 工作證 /**koŋ^{55}tsɔːk^{3}** tseŋ^{33}/ 'work permit', etc. These comparisons suggest a fifth analysis in which a loosely conjoined compound word 工作日 /koŋ^{55}tsɔːk^{3}jɐt^{2}/ has embedded within it a more tightly conjoined compound 工作 /koŋ^{55}tsɔːk^{3}/ 'working', yielding the left-branching structure [[工作]$_{V}$ [日]$_{N}$]$_{N}$. Thus, for Cantonese the wordhood of a string of morphemes could be interpreted as a continuum, with one end denoting the loosest relationship between the participating morphemes while the other end the tightest relationship. The judgment of wordhood might then vary along the continuum across different native Cantonese speakers.

Thus, the question "what is a word" is not easy to answer for Cantonese and for Chinese in general. Despite the difficulty, one interesting observation we can make is that it is far easier for Chinese speakers to say where a word boundary *cannot* be and where *a* word boundary *can* be than to say where *the* word boundary *is*. (I had this observation when I was working on the Segmentation Corpus in the Department of Chinese and

Bilingual Studies at the Hong Kong Polytechnic University in the early 90's with researchers from Beijing, Taiwan and Hong Kong. Other researchers such as Sproat et al. 1996, Wu 2003, Gao et al. 2004 also reports this observation.) The difficulty of word segmentation for Chinese and Cantonese, then, may be an artifact of assuming that there exists only one "correct" word segmentation (see Wu 2003). It follows that in devising word segmentation principles, we might have been asking the wrong question; instead of asking where *the* word boundary is, we perhaps should ask where *a* word boundary can be, and then order the potential boundaries in terms of their associated degrees of compositionality, as in the analysis of 工作日 /koŋ^{55}tsɔːk^{3}jɐt^{2}/ 'work day' as [[工作] $_{V}$ [日] $_{N}$] $_{N}$.

4.1.2 Developing word segmentation principles and procedures for spoken Cantonese

In developing the word segmentation principles in this dissertation, consideration was given to the observation that there could be more than one acceptable word boundary parse, involving different levels of grouping, for any given string of morphemes. That is, the set of principles for word segmentation was devised to allow for flexibility in locating the word boundaries, and yet, the word boundaries will not be located at places where the native speakers of Cantonese would say 'this cannot be a word boundary'. Thus, the word segmentation generates long words and short words (interpreted as 'segmentation units' for this dissertation), an idea inspired by Uchimoto et al. (2003) and Wu (2003). A short word can be embedded in a long word. The word segmentation in this dissertation hinges on sentence analysis (cf. Wu 2003). Although the part of speech of each word is not tagged at this stage of research, the word boundaries are tagged for whether they are edges of "short" words with no internal word boundaries or "long" words that have short words embedded within them. Details about the corpus, the word segmentation principles and the procedures are in order.

4.1.3 The Hong Kong Cantonese Adult Language Corpus (HKCAC)

The Hong Kong Cantonese Adult Language Corpus (HKCAC), developed by Leung and Law (2001), is a collection of spontaneous speech segments from seven radio programs (eight hours in total) on five types of topics (see Table 4.1), containing about 170,000 syllables. The corpus comprises orthographic transcriptions, phonetic transcriptions using IPA symbols, and audio files of the recordings. The transcription data files can be downloaded from <http://shs.hku.hk/corpus/index.htm>, although the audio files are not available there. There are 56 original HKCAC data files in Excel format. Each file contains an orthographic transcription of the speech and a phonetic transcription of the speech.

Program name	Theme	Time of broadcast	No. of participants (in the panel discussions)	Recording time
Panel discussions				
政黨論壇 [Political party forum]	Political issues	Morning	11	100 minutes
特區年代財經學人 [Economists of the Special Administrative Region era]	Economic issues	Morning	4	42 minutes
Phone-in programs				
平息你的風波 [To appease your mind]	Current affairs	Morning	15	69 minutes
有冇心情顏聯武 [Are you in the mood – Ngaan Lyun-mou]	Personal	Late night	6	67 minutes
星空奇遇鐵達尼 [Startrek and Titanic]	Personal	Evening	15	73 minutes
海琪的天空 [The sky of Hoi-kei]	Personal	Late night	10	69 minutes
恐怖熱線 [Scary hotline]	Ghost stories	Late night	8	67 minutes

Table 4.1 Details of the content of the HKCAC as adapted from Leung and Law (2001).

4.1.4 Synopsis of the word segmentation exercise

The word segmentation process that yielded word types and word frequencies for this dissertation was based on the transcriptions in the HKCAC. The segmentation process had two stages. First, the HKCAC was segmented into words automatically using the Segmentation Corpus text segmentation algorithm and program (see Leung and Law (2001)). Then, the results of this first-pass segmentation were examined to develop principles for word segmentation more appropriate for spoken Cantonese, and a second-pass segmentation was done using these principles.

Work at stage one involved the following steps. There are 56 original HKCAC data files in Excel format. Each file contains an orthographic transcription of the speech and a corresponding phonetic transcription in alternating rows, and each pair of rows is roughly an utterance as defined by Leung and Law (2001) (see Figure 4.1). To make the data format good to use for subsequent data processing, the orthographic transcription and the phonetic transcription of each of the 56 data files were split to make a set of 56 orthographic transcription files and the corresponding set of 56 phonetic transcription files. All these files were saved in Unicode plain text format (see Figure 4.3). The HKCAC could then be segmented into words by applying the Segmentation Corpus word segmentation algorithm and program to the orthographic transcription files of the HKCAC line by line. The results of this stage could only be a first-pass segmentation. This is because the Segmentation Corpus word segmentation algorithm and program were developed for segmenting sentences of *written texts* into words, and *spoken* Cantonese has many words not seen in written texts.

At stage two, the 56 segmented orthographic transcription files were concatenated together and each line was evaluated so that a set of spoken Cantonese word segmentation principles could be devised to correct the first-pass segmentation. The principles were applied first to get short words. Short words were then grouped into long words according to principles specified. The segmentation yielded words ranging from

one to four syllables in length. The orthographic transcription was then aligned with the phonetic transcription in a concatenated phonetic transcription file, so that the word boundaries and tags could be inserted also into the lines of the phonetic transcription file.

Full details about the data structure of the HKCAC and the segmentation process are given in Appendix C. In the rest of this section of this chapter, we will focus on the things that are directly relevant for this dissertation.

4.1.5 Structure of the HKCAC data

Figure 4.1 shows how the original HKCAC data file is structured with notes (in boxes) about aspects of this structure that are most relevant for understanding the steps of the word segmentation process and the way in which the data were used in this dissertation.

	A	B	C	D	E	F	G	**H**	I	J	K	L	M	N	O	P
1	626	1	O	M	H2	喂	**hello**									
2	626	1	P	M	H2	wɐi˥	hələʊ									
3	626	2	O	F	H1	喂										
4	626	2	P	F	H1	wɐi˥										
5	626	3	O	F	C	喂										
6	626	3	P	F	C	wɐi˥										
7	626	4	O	M	H2	[	係									
8	626	4	P	M	H2	**[**	hɐi˩	**]**								
9	626	5	O	F	H1	[	係	]	你	係						
10	626	5	P	F	H1	[	hɐi˩	]	leiˏ	hɐi˩						
11	626	6	O	F	C	**ɛ˩**	Sand y									
12	626	6	P	F	C	ɛ˩	sænd ɪ									
13	626	7	O	F	H1	Sand y										
14	626	7	P	F	H1	sænd ɪ										
15	626	8	O	F	C	係	呀									
16	626	8	P	F	C	hɐi˩	ja˩									
17	626	9	O	F	H1	係	你	[	好	]						

First row is column code in Excel.

Transcription given in English when speaker code-switched into English. The corresponding phonetic transcription was not given in the Cantonese segment-plus-tone configuration, indicating that the HKCAC transcriber interpreted the speaker to be using an English phonetic form, rather than the Cantonese form for the English loanword sometimes represented as 哈佬 or 蝦佬 /haː^{55}lou^{35}/ 'hello'.

[ ] denotes overlapping speech between two people in the HKCAC transcription

Transcription is given in IPA for the ground-holding morpheme, which does not have an orthographic representation.

First column is row number in Excel.

Figure 4.1(a) Lines 1-17 in the Ghost-26g segment of the HKCAC, with notes in the boxes.

18	626	9	P	F	H1	hɐi˧	lei˨˩˧	[	hou˧˥	]						
19	626	10	O	M	H2	[	O.K.	]								
20	626	10	P	M	H2	[	okei	]								
21	626	11	O	F	C	係	你	好								
22	626	11	P	F	C	hɐi˧	lei˨˩˧	hou˧˥								
23	626	12	O	M	H2	Sandy	有	[	咩	講	]					
24	626	12	P	M	H2	sænd ɪ	jɐu˨˩˧	[	**mɛ**˩˧	kɔŋ˧˥	]					
25	626	13	O	F	C	[	ɛ˧	]	我	就	想	講	我	個	朋	友
26	626	13	P	F	C	[	ɛ˧	]	ɔ˨˩˧	tsɐu˨˩	sœŋ˧˥	kɔŋ˧˥	ŋɔ˨˩˧	kɔ˧	p^hɐn˩˧	jɐu˨˩˧
27	626	14	O	F	C	佢	個	妹	呢	做	嗰	啲	ɛ˧	帶	街	㗎
28	626	14	P	F	C	k^høy˨˩˧	ɔ˧	mui˧˥	lɛ˥	tsou˨˩	kɔ˧˥	ti˥	ɛ˧	tai˧	kai˥	ka˧

A fused form for 乜嘢 /mɐt^{5}jɛː^{23}/ ‘what’.

Figure 4.1(b) Lines 18-28 of Ghost-26g.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	**627**	1	O	M	H 2	喂	hello	係	你	點	稱	呼					
2	627	1	P	M	H 2	wei˥	hələʊ	hɐi ˧	lei˩˨	tim ˥	ts^hi ŋ˥˧	fu˥˧					
3	627	**2**	O	F	H 1	喂	恐	怖	熱	線							
4	627	2	P	F	H 1	wei˥	hoŋ˥	pou ˧	jit˧˨	sin ˧							
5	627	3	O	M	H 2	喂											
6	627	3	P	M	H 2	wei˥											
7	627	4	**O**	F	C	ɛ˧	潘	迢	聰	露	芙	[	係	咪	]	呀	
8	627	4	**P**	F	C	ɛ˧	p^hun˥˧	siu ˩˨	ts^h oŋ˥˧	lou ˧	fu˩˨	[	ɐi˧˨	**mɐ i**˧˨	]	a˧	
9	627	5	O	M	H 2	[	Hello	]									
10	627	5	P	**M**	H 2	[	hələʊ	]									
11	627	6	O	F	H 1	[	係	]									
12	627	6	P	F	**H 1**	[	hɐi˧˨	]									
13	627	7	O	M	H 2	[	係	]									
14	627	7	P	M	H 2	[	hɐi˧˨	]									
15	627	8	O	F	H 1	[	你	好	]								
16	627	8	P	F	H 1	[	lei˩˧	hou ˥	]								
17	627	9	O	F	C	呀	你	好	ɛ˧	我	叫	阿	芝	呀			
18	627	9	P	F	C	a˧	lei˩˧	hou ˥	ɛ˧	ɔ˩˧	kiu ˧	a˧	tsi ˥˧	ja˧			

Code for the radio program and the current file.

Utterance number

A fused form for 唔係 /m̩^{21}hɐi^{22}/ 'no, not'.

A line for ***o***rthographic transcription

A line for ***p***honetic transcription

Gender of speaker: ***m***ale or ***f***emale

Role of speaker: ***h***ost, ***g***uest or ***c***aller

Figure 4.1(c) Lines 1-18 of segment Ghost-27g.

19	627	10	O	F	H1	阿	[	芝	]	係	嘛						
20	627	10	P	F	H1	a˧	[	tsi˥	]	ɐi˧	**ma**˩						
21	627	11	O	M	H2	[	係	]									
22	627	11	P	M	H2	[	hɐi˧	]									
23	627	12	O	F	C	係	呀	係	[	呀	我	]	想	講	嘅	嘢	呢
24	627	12	P	F	C	hɐi˧	ja˧	hɐi˧	[	ja˧	ɔ˩	]	sœŋ˥˩	kɔŋ˥˩	kɛ˧	jɛ˩	lɛ˥
25	627	13	O	F	H1	[	係	你	好	]							
26	627	13	P	F	H1	[	ɐi˧	lei˩	hou˥˩	]							
27	627	14	O	F	C	關	於	係	長	洲	東	堤	㗎				
28	627	14	P	F	C	kwan˥	jy˥	hɐi˧	ts^hœŋ˨˩	tsɐu˥˧	toŋ˥	t^hɐi˨˩	ə˧				

Two possible analyses of this form 嘛 [ma:^{23}]: (i) a fused form whose "underlying" form could be 唔係呀 /m̩^{21}hɐi^{22} a:^{33}/ 'negation-sentence particle', or 咪呀 /mɐi^{22} a:^{33}/ 'negation (fused form)-sentence particle'; (ii) a lexicalized question particle.

Figure 4.1(d) Lines 19-28 of Ghost-27g.

A first thing to note is that, for the most part, the orthographic transcription maps onto the phonetic transcription on a syllable-by-syllable basis. That is, the orthographic transcription of the morphemes is given in Chinese characters with the odd short stretch of roman letters for words of English and other "lettered words" (see Riha 2006), as well as for morphemes that do not have any Chinese character.

The syllable-by-syllable mapping between the orthographic and the phonetic transcriptions holds for fused forms as well. For example, the single Chinese character 咩 was used to transcribe forms such as [mɛː^{55}] and [mɛː$^{5+23}$] (the "underlying" non-fused form is 乜嘢 /mɐt^{5}jɛː^{23}/ 'what') (see an example in lines 23-24 in Figure 4.1b). Similarly, the single Chinese character 咪 was used to transcribe the form [mɐi^{22}] (the "underlying" non-fused form is 唔係 /m̩^{21}hɐi^{22}/ 'no, not') (see an example in line 7-8 in Figure 4.1c). These two cases are examples of fused forms that have been lexicalized to the extent of having a different character for the fused form. For fused forms that are not conventionally represented by a single Chinese character in this way, the HKCAC researchers would represent the fused form by choosing one of the Chinese characters that represents the "underlying" non-fused form. For example, the first Chinese character 就 of the "underlying" form 就係 /tsɐu^{22}hɐi^{22}/ 'is' was chosen to transcribe the fused form [tsɐi^{22}]; the second Chinese character 會 of the "underlying" form 唔會 /m̩^{21}muːi^{23}/ 'will not' was chosen to transcribe the fused form [muːi^{23}].

The syllable-by-syllable mapping did not hold for two types of word-forms in the corpus. Specifically, for loanwords from English with the pronunciation adapted to the Cantonese phonology as well as for code-switching to English by the speakers, orthographic transcriptions were given in English (e.g. *hello* in Figure 4.1a and *case* in Figure 4.2). These two types can be distinguished by phonetic transcriptions. Specifically, transcriptions that *do not* use the Cantonese segment-plus-tone configuration (e.g. *Mummy* [maːmiː] and *Daddy* [dædi]) suggest code-switching. By contrast, transcriptions that used the Cantonese segment-plus-tone configuration (e.g. *Bio* [paːi^{22}ɔː^{55}] 'biology' and *cable* [k^{h}ei^{55}bou^{35}]) suggest phonetic adaptation into the Cantonese phonology. In either case, transcribing an English word (regardless whether the phonetic form was borrowed as well) no longer guarantees that the orthographic transcription maps onto the

phonetic transcription on a syllable-by-syllable basis. That is, the orthographic transcription of the English word could occupy one cell, while one or more syllables could occupy the corresponding cell in the phonetic transcription field. Contrast the word *hello* [hələʊ] in Figure 4.1a and the word *case* [$k^{h}ei^{55}si:^{35}$] in Figure 4.2 for an example.

106	113	O	M	H1	[	咁	你	呀	離	開	咗	]			
106	113	P	M	H1	[	kɐm˥	lei˩	a˧	lei˩	hɔi˥	tsɔ˥	]			
{many lines in-between are omitted here}															
106	256	O	M	C	[	七	]	月	八	號	嗰	個	case	先	啦
106	256	P	M	C	[	$ts^{h}ɐt˥$	]	jyt˧	pat˧	hou˧	ɔ˥	ɔ˧	$k^{h}ei˥si˥$	sin˥	na˥

English translation for utterance 113, literally: So you then left PART. 'So, you left.'

English translation for utterance 256, literally: July 8th that CL. case first PART. 'The case on July 8th first.'

Figure 4.2 A snippet of the original HKCAC Excel data file *ping6-g*, showing an example of a fused form and a borrowing without borrowing the phonetic form.

In summary, then, because of the logosyllabic nature of the writing system, it should be relatively easy to pick out cases of extreme fusion. These will be cases where a disyllabic or longer word has a non-canonical "spelling" in the "O" row where there would be fewer characters than expected for the word. Moreover, because the HKCAC transcribers developed conventions for "spelling" frequent (or lexicalized) fusion forms, it is easy to count these. These "spelling" conventions were developed as part of the general principles for writing morphemes with more than one possible "spelling". Specifically, when more than one written form is possible to represent a certain morpheme, the form that is more commonly used to represent the morpheme would be chosen. (Explanation of what "common" meant is not available in Leung & Law.) An example given in Leung & Law is the passive marker /pei^{35}/, which can be written as 俾 or 畀. 畀was is used to represent the passive marker. For the representation of sentence

final particles, which are pragmatic morphemes attached to the end of an utterance, a single orthographic representation was used to represent a set of pragmatic morphemes that have the same segmental components but not tones. For example, /la:^{55}/, /la:^{33}/ and /la:^{21}/ are represented by 啦, and /a:^{55}/, /a:^{33}/ and /a:^{21}/ by 呀. (These "spelling" conventions also made it easier to add words to the "segmentation dictionary" in the second-pass segmentation, as will be described in section 4.1.8.)

4.1.6 The Segmentation Corpus word segmentation algorithm and program

The word segmentation algorithm and program of the Segmentation Corpus was used to segment the HKCAC into words as a first pass. The Segmentation Corpus is an electronic database of about five million Chinese characters collected from the Chinese newspaper texts on nine different topics. Although the newspaper articles are written in Chinese characters and are thus more or less readable by literate speakers of any variety of Chinese, the texts actually represent three different varieties of Chinese spoken in three regions: Putonghua Mandarin spoken in mainland China, Taiwan Mandarin spoken in Taiwan, and Hong Kong Cantonese.

A set of segmentation principles was developed by researchers at the Chinese Language Centre and Department of Chinese and Bilingual Studies at the Hong Kong Polytechnic University to segment the sequences of Chinese morphemes into "words" in a similar way for all three varieties. That is, the criteria were intended to be applicable to texts in all three Chinese varieties. Therefore, phonological criteria (such as the distribution of "neutral tone" syllables in Putonghua Mandarin) were not invoked. Rather, strings of characters were segmented from the texts based on the syntax, and on the length of the resulting word, its semantic transparency, and its collocational frequency. Therefore, the "words" segmented from the newspaper text database are orthographic strings of characters rather than phonological word forms.

In other words, a "word" as output by the Segmentation Corpus algorithm is a string of Chinese characters that represents a form that satisfies the following tests. First, it is an independent part of speech (e.g. the noun 盒子 /hɐp^{6}tsiː^{35}/ (or /hɐp^{35}tsiː^{35}/ 'box' is one word with the structure stem followed by an suffix. 子 /tsiː^{35}/ is treated as a nominalizing suffix in this context as it cannot stand alone even though 盒 /hɐp^{35}/ (or /hɐp^{6}/) can stand alone to mean 'box'). Second, it has a meaning that is not simply a sum

of its parts (e.g. 火車 is the noun /fɔː^{35}tsʰɛː^{55}/ 'train', not two words 火 /fɔː^{35}/ 'fire' and 車 /tsʰɛː^{55}/ 'vehicle'). Third, it consists of no more than four characters. Fourth, it either is listed in 現代漢語詞典 *Xiandai Hanyu Cidian* [Modern Chinese Dictionary] (1989) or 中國成語大辭典 *Zhongguo Chengyu Dacidian* [Dictionary of Chinese Idioms] (1987), or it meets a predetermined frequency threshold (for strings of text not listed in these two dictionaries).

4.1.7 First-pass segmentation

The Segmentation Corpus word segmentation program was designed to automatically segment Chinese written texts into words based on the 33,000-word lexicon derived from the Segmentation Corpus. However, the program of course does not work for phonetic transcriptions. An English word intermingled with the Chinese characters can be recognized provided the word was in the Segmentation Corpus. This is why I had to split the HKCAC into orthographic and phonetic files. In splitting the Excel files, I extracted only the information that I needed for the current study. Information relevant for the present study are: the orthographic transcription, the phonetic transcription, and the symbols "O" (for orthographic) and "P" (for phonetic) that signaled the beginning of each line, as shown in Figure 4.3 for the first 28 lines of the segment shown in Figure 4.1. The corresponding output of the Segmentation Corpus program is shown in Figure 4.4, with notes about places where the program failed.

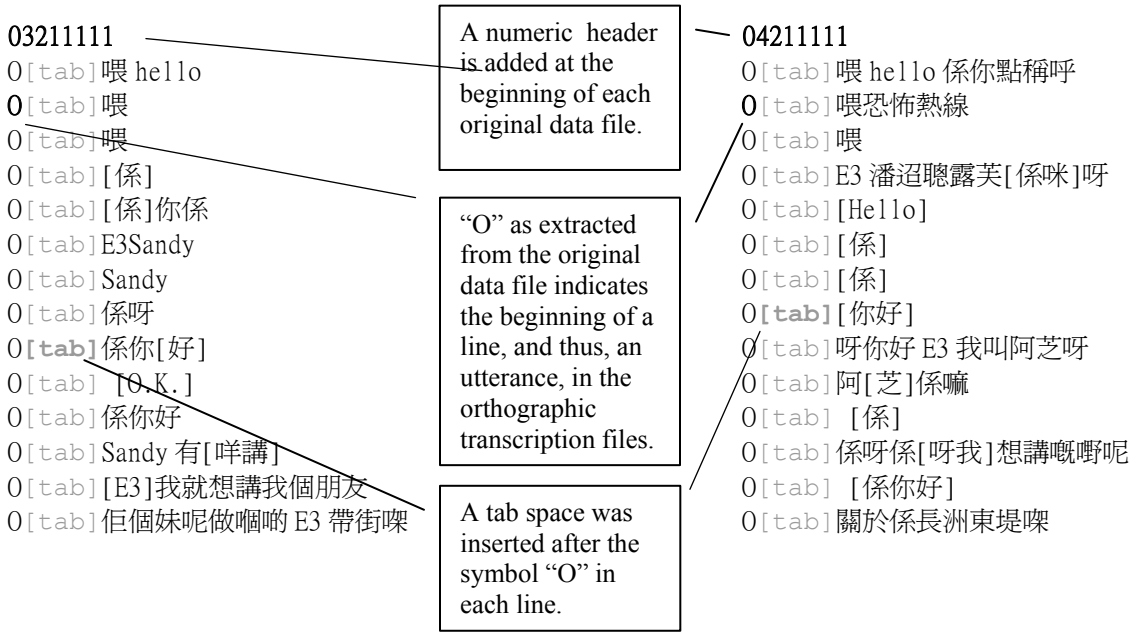


Figure 4.3(a) Orthographic transcription files in plain text format for lines 1-28 of Ghost-26g and Ghost-27g.

Ghost-26g

```
03211111
P     wAi2  hGlGU
P     wAi2
P     wAi2
P     [     hAi6  ]
P     [     hAi6  ]
      lei5  hAi6
P     E3    sYndI
P     sYndI
P     hAi6  ja3
P     hAi6  lei5  [
      hou2  ]
P     [     okei  ]
P     hAi6  lei5  hou2
P     sYndI jAu5  [     mE8
      kON2  ]
P     [     E3    ]     O5
      tsAu6 sJN2  kON2  NO5
      kO3   pHAn4 jAu5
P     kH0y5 O3    mui2  lE1
      tsou6 kO2   ti1   E3
      tai3  kai1  ka3
```

Ghost-27g

```
04211111
P     wei2  hGlGU hAi6
      lei5  tim2
      tsHiN1      fu1
P     wei2  hoN2  pou3
      jit6  sin3
P     wei2
P     E3    pHun1 siu4
      tsHoN1      lou6
      fu4   [     Ai6
      mAi6  ]     a3
P     [     hGlGU ]
P     [     hAi6  ]
P     [     hAi6  ]
P     [     lei5  hou2
      ]
P     a3    lei5  hou2
      E3    O5    kiu3
      a3    tsi1  ja3
P     a3    [     tsi1
      ]     Ai6   ma5
P     [     hAi6  ]
P     hAi6  ja3   hAi6
      [     ja3   O5
      ]     sJN2  kON2
      kE3   jE5   lE1
P     [     Ai6   lei5
      hou2  ]
P     kwan1 jy1   hAi6
      tsHJN4      tsAu1
      toN1  tHAi4 G3
```

A numeric header corresponding to that in the orthographic transcription file was added at the beginning of each original data file.

Tab spaces were inserted corresponding to the boundary of each cell in the original Excel files.

With the conversion from the original Excel format into plain text format, the IPA symbols in the original Excel files were converted into ASCII forms and tone marks became tone numbers 1-6 and 8, where 8 denotes a fall-rise tone as transcribed by the HKCAC transcribers.

"P" as extracted from the original data file indicates the beginning of a line, and thus, an utterance, in the phonetic transcription files.

Figure 4.3(b) Phonetic transcription files in plain text format for the same 1-28 lines.

03211111
0^^喂^^hello^^
0^^喂^^
0^^喂^^
0^^[^^係^^]^^
0^^[^^係^^]^^你^^係^^
0^^E3**^^Sandy^^**
0^^Sandy^^
0^^係^^呀^^
0^^係^^你^^[^^好^^]^^
0^^[^^O.K.^^]^^
0^^係^^你^^好^^
0^^Sandy^^有^^[^^咩^^講^^]^^
0^^[^^E3^^]^^我^^就^^想^^講^^我^^個^^朋友^^
0^^佢^^個^^妹^^呢^^做^^嗰^^啲^^E3**^^帶^^街^^**㗎^^
04211111
0^^喂^^hello^^係^^你^^點^^稱呼^^
0^^喂**^^恐怖^^熱線^^**
0^^喂^^
0^^E3**^^潘^^迢^^聰^^露^^芙^^**[^^係^^咪^^]^^呀^^
0^^[^^Hello^^]^^
0^^[^^係^^]^^
0^^[^^係^^]^^
0^^[^^你^^好^^]^^
0^^呀^^你^^好^^E3^^我^^叫^^阿^^芝^^呀^^
0^^阿^^[^^芝^^]^^係^^嘛^^
0^^[^^係^^]^^
0^^係^^呀^^係^^[^^呀^^我^^]^^想^^講^^嘅^^嘢^^呢^^
0^^[^^係^^你^^好^^]^^
0^^關於^^係**^^長洲^^東^^堤^^**㗎^^

The proper name "Sandy" is not distinguished from other English words in the corpus.

帶街 /ta:i^{33}ka:i^{55}/ is a colloquial Cantonese form for 'tourist guide'. It is a compound word.

恐怖 /hoŋ^{35}pou^{33}/ 'horror' and 熱線 /ji:t^{2}si:n^{33}/ 'hotline' are generic words used as a radio program name here.

The given names 迢聰 /si:u^{21}ts^{h}oŋ^{55}/ 'Siu Chung' and 露芙 /lou^{22}fu:^{21}/ 'Ruth' are not grouped and marked as proper names. The relationship of the former is not obvious with the surname 潘 /p^{h}u:n^{55}/.

The place name 長洲 'Cheung Chau' is not marked as proper name, while 東堤 'Tung Tai' is not grouped and marked as proper name.

Figure 4.4 Output text after applying the Segmentation Corpus word segmentation algorithm on the orthographic transcription files of the 28 lines in Figure 4.3(a). The carats (^^) symbols indicate word boundaries posited by the word segmentation program.

4.1.8 Second-pass segmentation

While the second-pass word segmentation was most immediately necessary for developing the materials for the second and the third experiments in this dissertation, it was also meant to serve as a basis for future development of a word segmentation program for spoken Cantonese. The first step, therefore, was to articulate principles to supplement those developed for the Segmentation Corpus. In developing the principles, I wanted to accomplish two desired outcomes, which were sometimes in conflict.

Firstly, the principles should result in word segmentation at more than one level to generate long words and short words as described above in section 4.1.2. In other words, where the program that produced the first-pass segmentation is required to resolve ambiguities such as 工作日 /koŋ^{55}tsɔːk^{3}jɐt^{2}/ 'work day' into either as [[工作] [日]] or [工作日], I wanted a way to mark both segmentations to indicate [[工作] [日]] resulting in a "short word" embedded in a "long word" as in the segmentation algorithms described in Uchimoto et al. (2003) and Wu (2003).

Secondly, the principles should result in a segmentation that fits the tried-and-true criteria of the Segmentation Corpus algorithm. Essentially, then, the HKCAC word segmentation also hinges on the same kind of sentence analysis done for the original Segmentation Corpus so that the resulting string must be analyzable as a well-defined lexical category, such as a noun or a verb and so on (cf. Wu 2003).

The general guiding principle that I used to resolve conflicts between these two desired outcomes is segmentation consistency. That is, the grouping of the shorter words would be done consistently. I applied such criteria as semantic non-transparency more stringently for shorter words and less stringently when grouping shorter words into longer words, where I developed additional criteria that could be applied to ensure that.

Note also that the first of the desired outcomes led me to override the four-character (or four syllable) constraint of the Segmentation Corpus algorithm when morphosyntactic ambiguity is encountered. That is, since I am not forced to choose a single unambiguous segmentation, I no longer need the length criterion that helped constrain choices in the Segmentation Corpus. Thus, the current word segmentation algorithm generates short words ranging from one to four syllables in length but short words could be grouped into longer words of more than four syllables. These longer forms were generated in the current second-pass segmentation, however, only when there was morphosyntactic ambiguity within the words. Since the maximum word length to be tested in Experiments II and III is four syllables, I did not work on developing criteria for evaluating longer groupings.

The first step in doing the second-pass segmentation of the HKCAC, then, was to evaluate the first-pass segmentation output to correct obvious cases where short words were not recognized so that syllable-morphemes (or stems and affixes) were not grouped appropriately into short words. For example, in the snippet shown in Figure 4.4, the sequence of morphemes 帶 /taːi^{33}/ 'bring' and 街 /kaːi^{55}/ 'street' is a colloquial Cantonese word meaning 'tourist guide', but these characters were not grouped together to make a short word by the text segmentation algorithm. Similarly, the sequence of morphemes 唔 /m̩^{21}/ 'negation' and 該 /kɔːi^{55}/ 'should' did not make a word as parsed by the text segmentation algorithm, despite that the string 唔該 /m̩^{21}kɔːi^{55}/ is a word meaning 'thank you' in Cantonese. Cases like these were corrected using the criterion that they are listed as words in 廣州話口語詞的研究 /k^{w}ɔːŋ^{35} tsɐu^{55} waː^{35} hɐu^{35} jyː^{35} ts^{h}iː^{21} tek^{5} jiːn^{21} kɐu^{33}/ [The Study of Spoken Cantonese] by Qiao (1966), 廣州話方言詞典 /k^{w}ɔːŋ^{35} tsɐu^{55} waː^{35} fɔːŋ^{55} jiːn^{21} ts^{h}iː^{21} tiːn^{35}/ [Dictionary of the Guangzhou Dialect] by Rao et al. (1991), or 港式廣州話詞典 /kɔːŋ^{35} sek^{5} k^{w}ɔːŋ^{35} tsɐu^{55} waː^{35} ts^{h}iː^{21} tiːn^{35}/ [Hong Kong Cantonese Dictionary] by Cheung and Ngai (1999). That is, these three references were added to the list in the first condition of the original fourth criterion described above (see section 4.1.6

above). A word entry in any of the above references makes a short or a long word in the HKCAC. Thus, 帶街 /taːi^{33}kaːi^{55}/ ‘tourist guide’ and 唔該 /m̩^{21}kɔːi^{55}/ ‘thank you’ each makes a short word (i.e. a segmentation unit).

I also added some other out-of-dictionary forms based on transparency and so on. For example, li4li2 臨 [liː^{21}liː^{35}lɐm^{21}] was judged to mean ‘hurry’, a possible variant of 拿拿臨 /laː^{21}laː^{21}lɐm^{21}/ ‘hurry’.

On the other hand, I did not bother at this stage to insert lower-level short-word boundaries into those idioms and idiomatic expressions which were identified as words. That is, these are not tagged differently from other word entries in the Segmentation Corpus and the Cantonese dictionaries, although idioms and idiomatic expressions usually have more complex internal structure than simple and compound words. For example, the string of morphemes 一廂情願 /jɐt^{5}sœːŋ^{55}ts^{h}ɪŋ^{21}jyːn^{22}/ ‘only one of the (two) parties is willing’ (literally, ‘one side willing’) were grouped together as short word because this is an entry in the Segmentation Corpus. The Segmentation Corpus treats a string of morphemes as an idiom or idiomatic expression if the string occurred in the dictionary of idioms 中國成語大辭典 *Zhongguo Chengyu Dacidian* [Dictionary of Chinese Idioms] (1987) or any of the Cantonese / Chinese dictionaries above. Inconsistencies between the Segmentation Corpus word segmentation algorithm and the Cantonese dictionaries were regularized. For example, 第一時間 /tɐi^{22}jɐt^{5}siː^{21}kaːn^{33}/ ‘immediately’ (literally, ‘first/number one time’) is treated as an idiom or idiomatic expression as it occurs in 港式廣州話詞典 /kɔːŋ^{35} sek^{5} k^{w}ɔːŋ^{35} tsɐu^{55} waː^{35} ts^{h}iː^{21} tiːn^{35}/ [Hong Kong Cantonese Dictionary] by Cheung and Ngai (1999). Variant forms of dictionary entries whether as a main entry or as a variant form under another main entry are also treated as idioms or idiomatic expressions. An example is 唔怪得知 /m̩^{21}kwaːi^{33}tɐk^{5}tsiː^{55}/ ‘no wonder’ (literally, ‘not strange able know’) in the HKCAC, for

which Cheung and Ngai (1999) only have the entry 唔怪之 / m̩^{21}kwaːi^{33}tsiː^{55}/ 'no wonder' (literally, 'not strange it'). And they note under the entry that 唔怪之得 /m̩^{21}kwaːi^{33}tsiː^{55}tɐk^{5}/ 'no wonder' (literally, 'not strange it able') is a variant form for the entry.

In addition to identifying morphologically simple short words in this way, I also developed and applied criteria for grouping bound morphemes (clitics) together with their stems to make more short words, as well as for grouping of short words to form more complex long words for testing syllable fusion in this dissertation.

The criteria for grouping were morphosyntactic: shorter words were grouped together as long as the shorter words fell into one of the five lexical categories shown in Table 4.3. The resulting words, long or short, would then fall into one of the 19 lexical categories shown in Table 4.2. These lexical categories are as specified in Cheung (1972) and Matthews & Yip (1994: 57-357). In addition, there was a length criterion, since the maximum length of words that I needed for materials for Experiments II and III was four, I stopped when the maximum length was reached. This maximum word length was chosen because four syllables is a kind of natural limit — as in the original Segmentation Corpus criterion. (At the same time, I did relax the length criterion in cases where there was an ambiguity between two possible groupings — i.e. competing, partially overlapping morphosyntactic analyses that together covered more than four syllables, even though each individual parse did not result in a long word that violated the 4-syllable length constraint — as stated above.)

Before grouping I parsed the syntactic structure of each sentence, although the part of speech of the words and the sentence structure were not explicitly tagged at this stage of the research. Syntactic analysis in this dissertation referenced work on Cantonese

(e.g. Cheung 1972, Matthews & Yip 1994, Law 1990, Cheng & Sybesma 2004), as well as work on Mandarin Chinese (e.g. Li & Thompson 1981, Zhu 1982), since there are syntactic similarities between the two varieties.

	Lexical categories	**Abbrev.**	**Examples**	**Gloss**
1.	Verb	V	追捧 /tsɵy^{55}p^{h}oŋ^{35}/	'to pursue'
			賣 /maːi^{22}/	'to sell'
			蝕 /siːt^{2}/	'to lose'
			去 /hɵy^{33}/	'go'
	(Verbal particles,	(VPT)	賣咗 /maːi^{22}tsɔː^{35}/	'sold'
	including aspect markers and directional verbal particles, are treated as suffixes morphologically, and are thus		賣完 /maːi^{22}jyːn^{21}/	'finished selling'
			賣埋 /maːi^{22}maːi^{21}/	'also sell'
			賣到 /maːi^{22}tou^{35}/	'sold for ($)'
	attached after the verb.)		賣出去 /maːi^{22}ts^{h}ɵt^{5}hɵy^{33}/	'to sell'
2.	Adjective	Adj	簡單 /kaːn^{35}taːn^{55}/	'simple'
			沮喪 /tsɵy^{35}sɔːŋ^{33}/	'frustrated'
			靚 /lɛːŋ^{33}/	'beautiful'
3.	Auxiliary verb	Aux	能夠 /nɐŋ^{21}kɐu^{33}/	'can'
			可以 /hɔː^{35}jiː^{35}/	'can'
			應該 /jeŋ^{55}kɔːi^{55}/	'should'
			寧願 /neŋ^{21}jyːn^{35}/	'would rather'
			要 /jiːu^{33}/	'need to'
			會 /wuːi^{23}/	'will'
			有 /jɐu^{23}/ (as in /jɐu^{23}/+verb)	'has, did'
			冇 /mou^{23}/ (as in /mou^{23}/+verb)	'has not, did not'
			想 /sœːŋ^{35}/	'want to'
			肯 /hɐŋ^{35}/	'willing to'
			敢 /kɐm^{35}/	'dare'
			識 /sek^{5}/	'know'

Table 4.2 Lexical categories (cf. Matthews & Yip 1994).

Table 4.2 continued

4.	Preposition	Prep	經 /keŋ^{55}/ 喺 /hɐi^{35}/ 俾 /pei^{35}/	'via' 'at' 'passive'
5.	Localizer/Postposition	LOC	上面 /sœːŋ^{22}miːn^{22}/ 前 /ts^{h}iːn^{21}/ 中 /tsoŋ^{55}/	'above' 'in the front' 'in the middle'
6.	Classifier	CL	餐 /ts^{h}aːn^{55}/ 隻 /tsɛːk^{3}/ 啲 /tiː^{55}/	(for meals) (for countable objects) (for uncountable objects)
7.	Sentence particle	SP	嘅 /kɛː^{33}/ 喎 /wɔː^{23}/ 㗎 /kaː^{21}/ 呢 /lɛː^{55}/	'declaration, assertion' 'reporting someone's speech' 'rhetorical question' 'question'
8.	Adverb	Adv	啱啱 /ŋaːm^{55}ŋaːm^{55}/ 咁 /kɐm^{33}/ 本來 /puːn^{35}lɔːi^{21}/ 再 /tsɔːi^{33}/	'just' 'so' 'originally' 'again'
9.	Noun	N	公司 /koŋ^{55}siː^{55}/ 生活 /sɐŋ^{55}wuːt^{2}/ 小姐 /siːu^{35}tsɛː^{35}/	'company' 'everyday life' 'miss, lady'
10.	Pronoun	ProN	你 (哋) /nei^{23}(tei^{22})/ 我 /ŋɔː^{23}/ 我哋 /ŋɔː^{23}tei^{22}/ 佢 /k^{h}ɵy^{23}/ 佢哋 /k^{h}ɵy^{23}tei^{22}/	'you (pl. and sing.)' 'I, me' 'we, us' 's/he, him, her, it' 'they, them'
11.	Proper name	PN	中國 /tsoŋ^{55}k^{w}ɔːk^{3}/ 香港 /hœːŋ^{55}kɔːŋ^{35}/ 數碼港 /sou^{33}maː^{35} kɔːŋ^{35}/	'China' 'Hong Kong' 'Cyber Port'

Table 4.2 continued

12.	Demonstrative	Dem	嗰 /kɔː^{35}/ 呢 /liː^{55}/ 今 /kɐm^{55}/	'that' 'this' 'this'
13.	Conjunction	Conj	就算 /tsɐu^{22}syːn^{33}/ 如果 /jyː^{21}k^{w}ɔː^{35}/ 或者 /waːk^{2}tsɛː^{35}/	'even though' 'if' 'or'
14.	Quantifier	QF	每 /muiː^{23}/ 各 /kɔː^{3}/ 所有 /sɔː^{35}jɐu^{35}/ 零 /leŋ^{21}/ 一 /jɐt^{5}/ ... 十 /sɐp^{2}/	'each' 'every' 'all' '0' '1' ... '10'
15.	Question marker	QM	係咪 /hɐi^{22}mɐi^{22}/ 未 /mei^{22}/ 有冇 /jɐu^{23}mou^{23}/ 乜 /mɐt^{5}/ 邊 /biːn^{55}/ 幾 /kei^{35}/ 點 /tiːm^{35}/	'is it…' 'yet' 'did…?' 'what' 'where, who, which' 'how many, how much' 'how, why'
16.	Negation	NEG	無 /mou^{23}/ 未 /mei^{22}/ 唔係 /m̩^{21}hɐi^{22}/ 唔只 /m̩^{21}tsiː^{35}/	'not' 'not yet' 'no, not' 'not only'
17.	Interjection	INJ	嗱 /naː^{21}/	'hey, right'
18.	Exclamation	EXM	嘩 /waː^{33}/ 哎吔 /ɐi^{55}jaː^{23}/	'wow' 'oops'
19.	Idiom / Idiomatic expression	Idiom	於事無補 /jyː^{55}siː^{22}mou^{21}pou^{35}/ 一清二楚 /jɐt^{5}ts^{h}eŋ^{55}jiː^{22}ts^{h}ɔː^{35}/	'not helpful for the matter' 'crystal clear'

Table 4.2 ended

Note that verbs (lexical category 1) could be simple forms such as 賣 /maːi^{22}/ 'to sell' or affixed forms such as 賣咗 /maːi^{22}tsɔː^{35}/ 'sold'. That is, verbal particles (VPT), including aspect markers such as 咗 /tsɔː^{35}/ and 完 /jyːn^{21}/ and directional particles such as 出去 /ts^{h}ɵt^{5}hɵy^{33}/, are treated as suffixes morphologically and necessarily grouped together with a preceding verb stem to make word.

Auxiliary verbs (lexical category 3) are differentiated from full verbs in that (among other things) they cannot take verb particles. That said, it is not always easy to differentiate the two categories, and authors do not always agree on the list of auxiliary forms. For example, I follow Cheung (1972: 122) in treating 有 /jɐu^{23}/ (as in /jɐu^{23}/+verb) and 冇 /mou^{23}/ (as in /mou^{23}/+verb) as auxiliary, but Matthews & Yip (1994) do not include these in their list.

For negation (lexical category 16), there is a set of negation markers that do not stand alone and are treated as prefixes morphologically (cf. Matthews & Yip 1994:33). Thus, they are attached to the beginning of the following word. Some examples are:

唔 /m̩^{21}/ 'un-, im-, in-, non-'

不 /pɐt^{5}/

非 /fei^{55}/

For interjection (lexical category 17), contra Matthews & Yip (1994: 356-7), 咁 /kɐm^{35}/ 'so', 你話 /nei^{23} waː^{22}/ 'you say', and 即係 /tsek^{5}hɐi^{22}/ 'that is' are not treated as interjections in this study, but as words.

For exclamation (lexical category 18), contra Matthews & Yip (1994: 357-8), 死喇 /sei^{35} laː^{33}/ and 弊喇 /pɐi^{22} laː^{33}/, both meaning 'not good', are not treated as exclamations in this study, but as word followed by a particle.

The syntactic analysis according to categories in Table 4.2 yielded more short words because of the following rule:

<u>Rule 1:</u> Bound morphemes are grouped with stems to form short words

Bound morphemes are morphemes that do not stand alone by themselves. At the morphological level, morphological prefixes (e.g. 第 /tɐi^{22}/ 'number…' for counting numbers; 老 /lou^{23}/ (literally 'old') and 阿 /aː^{33}/ that precede a surname to show familiarity), suffixes (e.g. 仔 /tsɐi^{35}/ '-y'; 家 /kaː^{55}/ '-ist, -ian'; 員 /jyːn^{21}/ '-ist'), and infixes (e.g. 鬼 /kwɐi^{35}/ for emphasis as in 麻鬼煩 /maː^{21}kwɐi^{35}faːn^{21}/ 'very troublesome') were grouped with the stem to form short words. Verbal particles (including aspect markers) are suffixes morphologically, and were thus grouped with the preceding verb. Negation markers that do not stand alone (e.g. 唔 /m̩^{21}/, 不 /pɐt^{5}/ and 非 /fei^{55}/) are treated as prefixes morphologically (cf. Matthews & Yip 1994:33). These prefixes were grouped with the following verb or adjective to form short words.

The short words that were identified either by expanding the lexicon (as described at the beginning of this section) or by applying Rule 1 were then grouped into long words. This second stage of grouping used rules like Rule 1. These rules are explained below, but first some general remarks on word formation processes in Chinese are in order.

There is a strong resemblance between morphological structure and syntactic structure in all Chinese varieties (see e.g. Dai 1992, Chung 2004), including Cantonese. Concatenated syntactic constituents can be lexicalized and be parsed morphologically. The degree of lexicalization (i.e. whether a concatenated string of morphemes would be parsed by a native speaker as one compound word or more than one word) can vary from

person to person, since it depends on such properties as the co-occurrence frequency, which is speaker-dependent. Since degree of lexicalization is speaker-dependent, I needed to develop and apply more objective criteria.

That is, no assumption was made about whether a given string of morphemes was lexicalized or not, but operational rules were laid down so that there would be consistency in the grouping. These operational rules specify that short words belonging to any of the five lexical categories listed in Table 4.3 are eligible to be grouped with each other so long as the result is one of the same five lexical categories or reduplicated form. Table 4.4 lists the remaining lexical categories that were not grouped with each other and were not grouped with other categories. When the rules for grouping conflicted with those for non-grouping of words, grouping overrides non-grouping. The process of grouping must begin from the terminal node of a sentence. Furthermore, the grouping had to be done within the same syntactic constituent (i.e. constituents under the same node).

Lexical categories	Remarks
1. Verb (with or without verbal particle)	---
2. Noun (non-pronoun)	---
3. Adjective	---
4. Demonstrative	(one)+classifier were grouped with their preceding demonstrative.
5. Wh-question marker	(i) Wh-question markers is a subset of question markers. The Wh-question markers are: 乜- /mɐt^{5}/ 'what' 邊- /piːn^{55}/ 'which, where, who (depending on the following morpheme)' 幾- /kei^{35}/ 'how many, how much, when (depending on the following morpheme)' 點- /tiːm^{35}/ 'how, why (depending on the following morpheme)' (ii) A Wh-question marker may be grouped with a following classifiers.

Table 4.3 Five lexical categories that were grouped with other categories to form long words for this dissertation.

Lexical categories	Remarks
1. Auxiliary verb	---
2. Preposition	---
3. Localizer/Postposition	---
4. Classifier	(Exception: classifiers were grouped with the demonstrative or a Wh-question marker that preceded it.)
5. Sentence particle	---
6. Adverb	---
7. Pronoun	---
8. Proper name	---
9. Conjunction	---
10. Quantifier	---
11. Non Wh-question marker	---
12. Negation	---
13. Interjection	---
14. Exclamation	---
15. Idiom / Idiomatic expression	---

Table 4.4 Lexical categories that generally were *not* grouped with each other or other categories to form long words for this dissertation.

The five lexical categories specified to be eligible for grouping in this dissertation are nouns, verbs, adjectives, demonstratives, and Wh-question markers (plus classifiers where applicable). Verbs (with or without verbal particles), nouns, and adjectives were chosen because the grouping of them tended to contribute to compounding (Matthews & Yip 1994: 48-54). Note that the grouping included the grouping of more than one word of its own category, e.g., noun+noun, verb+verb, and so on. Demonstratives do not stand alone but must be followed by an appropriate classifier in Cantonese (Matthews & Yip 1994: 89). Thus, a demonstrative was grouped with a following classifier. (Compare Law (1990), who treats the combination of demonstrative+classifier as a single determiner, though the Cantonese speaker in me would say that such a construction has an indefinite

use sometimes.) Under the current grouping specifications, demonstratives were grouped with the four lexical categories above, plus the category of demonstratives themselves, albeit such combination normally does not happen. The morpheme 一 /jɐt^{5}/ 'one' could optionally be inserted between a demonstrative and a classifier, giving rise to the structure demonstrative+(one)+classifier (e.g. 例如係*呢一個*嘅 Yahoo 呀嗰啲呀 /lɐi^{22} jyː^{21} hɐi^{22} **liː^{55} jɐt^{5} kɔː^{33}** kɛː^{33} jaː^{22}huː^{55} aː^{33} kɔː^{35} tiː^{55} aː^{33}/ 'an example is **this** Yahoo thing and such'). The existence or non-existence of 'one' does not change the meaning of the form. Hence, demonstrative were grouped with their following classifier, as well as the morpheme 'one' in between them, if there was one. The main question markers in Cantonese are identified to be 乜- /mɐt^{5}/ 'what', 邊- /piːn^{55}/ 'which, where, who (depending on the following morpheme)', 幾- /kei^{35}/ 'how many, how much, when (depending on the following morpheme)', and 點- /tiːm^{35}/ 'how, why (depending on the following morpheme)' (cf. Matthews & Yip 1994: 323). These markers could be viewed as corresponding to the Wh-question markers in English. I, therefore, labeled this subset of question markers in Cantonese as "Wh-question markers". Like demonstratives, these markers were grouped with classifiers and any of the four lexical categories above, plus the category of Wh-question markers itself (if the latter indeed happens). Note that classifiers were restricted to be grouped with the demonstratives and the Wh-question markers only, but not with other lexical categories.

Where there is equal tension between grouping and non-grouping of words, grouping overrides non-grouping. For example, in the string 第三十九 /tɐi^{22} saːm^{55} sɐp^{2} kɐu^{35}/ '39th', 第 /tɐi^{22}/ 'number…' is a prefix that does not stand alone by itself, whereas 三十九 /saːm^{55}sɐp^{2}kɐu^{35}/ 'thirty-nine' is a stand-alone quantifier that was not to be grouped with other words (cf. Table 4.4). In this case, grouping took precedence over non-grouping, yielding the long word 第三十九 /tɐi^{22} saːm^{55} sɐp^{2} kɐu^{35}/ 'thirty-ninth'. There were only a few cases like this in segmenting the HKCAC.

Reduplicated forms are stand-alone words (see Table 4.5 for examples). Code-switching to English occurs in the HKCAC. Where the English words were produced using the Cantonese phonology (e.g. *Bio* [paːi^{22}ɔː^{55}] 'abbreviation for biology'; *call* [k^{h}ɔː^{55}] 'to call'), the borrowed words were treated as Cantonese and were grouped with their neighboring bound morphemes, if any.

Set (1): "Uninterrupted reduplication". There is an uninterrupted concatenation of an exact copy of a syllable.

AA	講講	/kɔːŋ^{35}kɔːŋ^{35}/	'to say a bit' (lit. 'say say')
AAB	冚冚轉	/t^{h}ɐm^{21}t^{h}ɐm^{21}tsyːn^{33}/	'merry-go-round' (lit. 'onomatopoeic syllables spin'
ABB	試諗諗	/siː^{33}nɐm^{35}nɐm^{35}/	'try to think' (lit. 'try think think')
AABB	嚟嚟去去	/lei^{21}lei^{21}hɵy^{33}hɵy^{33}/	'back and forth' (lit. 'come come go go')

Set (2): "Discontinuous reduplication". There is an interrupted concatenation of an exact copy of a syllable.

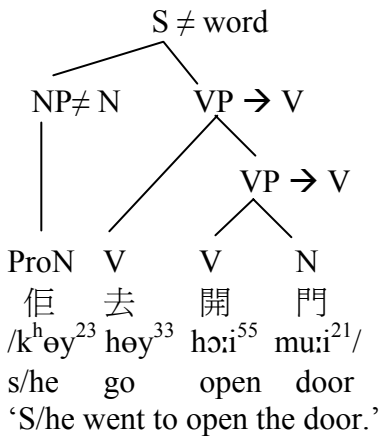
ABAB	逐個逐個	/tsok^{2}kɔː^{33}tsok^{2}kɔː^{33}/	'one by one' (lit. 'each CL each CL')
A 一 A	考一考	/haːu^{35}jɐt^{5}haːu^{35}/	'take a little test' (lit. 'test one test')
A 對 A	面對面	/miːn^{22}tɵy^{33}miːn^{22}/	'face to face' (lit. 'face to face')
有 A 有 B	有根有據	/jɐu^{23}kɐn^{55}jɐu^{23}kɵy^{33}/	'grounded, with reasons' (lit. 'has ground has evidence')
一 A 半 A	一次半次	/jɐt^{5}ts^{h}iː^{33}puːn^{33}ts^{h}iː^{33}/	'once or not even once' (lit. 'one time half time')
A 來 A 去	撳來撳去	/kɐm^{22}lɔːi^{21}kɐm^{22}hɵy^{33}/	'pressing (e.g. a button) many times' (lit. 'press come press go')
A 唔 A B	公唔公道	/koŋ^{55}m̩^{21}koŋ^{55}tou^{22}/	'fair or not' (lit. 'fair not fair')
A 一 A B	唎一唎氣	/t^{h}ɐu^{35}jɐt^{5}t^{h}ɐu^{35}hei^{33}/	'take a breath' (lit. 'breath one breath air')

Table 4.5 Reduplicated forms.

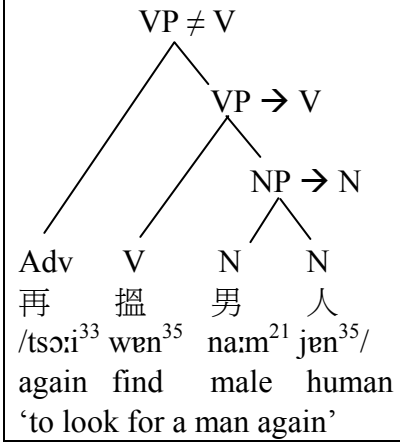
The process of grouping short words began from the terminal node of a sentence. It reiterated until (i) the process encountered a neighboring word that did not belong to any of the five lexical categories specified, or (ii) the process encountered a syntactic boundary that did not share an immediate parent node, or (iii) the process encountered disfluencies, or (iv) the string reached the maximal length of four syllables. For (ii), the relevant syntactic boundary often coincided with the line beginning or end that separates two sentences or stand-alone phrases. (Recall that each line in the HKCAC orthographic transcription file roughly represented a separate utterance.) Criterion (iv) could be temporarily overridden under the condition that I found the internal structure of a long word ambiguous. For example, the long word 結婚週年紀念 /kiːt^{3} fɐn^{55} tsɐu^{55} niːn^{21} kei^{35} niːm^{22}/ '(literally) wedding-annual-commemorate' can be parsed as [結婚 [[週年] [紀念]]] or [[[結婚] [週年]] 紀念], which would yield either 週年紀念 /tsɐu^{55} niːn^{21} kei^{35} niːm^{22}/ '(literally) annual-commemorate' or 結婚週年 /kiːt^{3} fɐn^{55} tsɐu^{55} niːn^{21}/ '(literally) wedding-annual' as a word for testing in this study. So, I grouped this word into a long word of six syllables temporarily, so that I could extract the word and pass it to three other linguists for morphosyntactic structure parsing at a later stage (see section 4.2). I would then modify the grouping according to the linguists' agreed-upon analysis or their majority analysis. (The maximal length that would eventually be extracted for testing would still be four.) Under the same node, either all immediate constituents were grouped, or none of the immediate constituents was grouped. The rules of grouping the Cantonese words also applied to the English words that the speakers produced with the Cantonese phonology. See Table 4.6 for examples of the above.

Although the above principles and operating rules were developed based on the HKCAC and were targeted at the purpose of the current dissertation, they could just as well be used or adapted for word segmentation for other Cantonese spoken corpora. Note that grouping of lexical categories other than those specified in this dissertation would be appropriate for research with a different purpose and/or a different application (for

example, see Wu (2003) for grouping much longer chunks as segmentation units). As noted above, there is no "correct" word segmentation for Chinese (as yet, at least), but word segmentation is purpose-driven.

<table>
<tr><td colspan="2">(i) The grouping process stopped when it encountered a neighboring word that did not belong to any of the five lexical categories specified for grouping.</td></tr>
<tr><td>a.
 

'S/he went to open the door.'</td><td>Grouping of short words into long words (word-internal relationship in parenthesis):

VP → V
去開門 /hɵy^{33}hɔːi^{55}muːn^{21}/ (verb-complement)
VP → V
開門 /hɔːi^{55}muːn^{21}/ (verb-object)</td></tr>
<tr><td>b.
 

'to look for a man again'</td><td>Grouping of short words into long words (word-internal relationship in parenthesis):

VP → V
搵男人 /wɐn^{35}naːm^{21}jɐn^{35}/ (verb-object)
VP → V
男人 /naːm^{21}jɐn^{35}/ (modifier-head)</td></tr>
</table>

Table 4.6 Visualizing the grouping of short words into long words. Note that I make *no* claims here about lexicalization. The groupings simply group things into potential compounds, following the structures in Table 4.7.

Table 4.6 continued

(ii) The grouping process stopped when it encountered a syntactic boundary with an adjacent constituent node that did not share an immediate parent node.
Utt 1: 你 要 boil down to 嘅 就 話 /nei^{23} jiːu^{33} bɔɪl daʊn tuː kɛː^{33} tsɐu^{22} waː^{22}/ you have to boil down to SP is to say 'what you have to boil down to is' Utt 2: 軟件 係 可以 死 嘅 /jyːn^{23} kiːn^{35} hɐi^{22} hɔː^{35} jiː^{23} sei^{35} kɛː^{33}/ software is can dead SP 'software can be dead' Although the first utterance did not seem to be finished, it was considered a separate utterance from the second utterance. It did not share the same syntactic tree with the second utterance. The grouping process, therefore, stopped at 'to say' for the first utterance.
(iii) The grouping process stopped when it encountered disfluencies.
維 ts^{h} 持 到 /wɐi^{21} ts^{h} ts^{h}iː^{21} tou^{35}/ keep ts^{h}- maintain VPT 'can maintain' [ts^{h}-] is a word fragment: the beginning of /ts^{h}iː^{21}tou^{35}/ cut off in a disfluency. 維 ts^{h}持到 were not grouped and thus they remained to be three short words.

Table 4.6 continued

(iv) The grouping process stopped when the string reached the maximal length of four syllables, unless there was an ambiguity of analysis.
a. Grouping of short words into long words (word-internal relationship in parenthesis): VP VP ≠ V (grouping blocked by length of component noun) NP→N 自由競爭 /tsiː^{22}jɐu^{21}keŋ^{33}tsɐŋ^{55}/ (modifier-head) Aux V Adj N 要 用 自由 競爭 /jiː^{33} joŋ^{22} tsiː^{22}jɐu^{21} keŋ^{33}tsɐŋ^{55}/ have to use free competition 'have to use free competition'
b. Grouping of short words into long words: Not applicable because *polyclinic* has four syllables (and it is a code-switched into English). S NP VP VP ≠ V (grouping blocked by length of component noun) ProN V V N 我 去 睇 polyclinic /ŋɔː^{23} hɵy^{33} t^{h}ɐi^{35} pɑlɪklɪnɪk/ I go see polyclinic 'I went to see (a doctor) at the polyclinic.'

Table 4.6 continued

(v) Under the same node, either all immediate constituents were grouped, or none of the immediate constituents was grouped.	
a. PP Prep N V 以 合約 代替 /jiː^{23} hɐp^{2}jœːk^{3} tɔːi^{22}tʰɐi^{33}/ with contract replace 'by means of contract replacement'	Grouping of short words into long words: Not applicable. ← This N V could not be VP→V because PP requires N or NP.
b. PP Prep N V 由 政府 提供 /jɐu^{21} tseŋ^{33}fuː^{35} tʰɐi^{21}koŋ^{55}/ from government provide 'government providing'	Grouping of short words into long words: Not applicable.

Table 4.6 ended

Morphologically, this grouping of words in these lexical categories gives rise to long words having one of the word-internal structures listed in the lower part of Table 4.7. In the table, words are broadly classified into two categories: simple word and complex word. While simple words are monomorphemic and show no internal branching structure, complex words show internal branching structures, which can be at more than one level for words that are of three or four syllables long. In Chinese, including Cantonese, complex words, in particular, 3- or 4-syllable words, idioms and idiomatic expressions, and reduplicated forms could show internal structures that resemble syntactic structures. That is, while the lexical categories in Table 4.2 – 4.5 and the types of word-internal syntactic structures in Table 4.6 – 4.7 were used to identify potential long words, the tags that I used only marked the degree of embedding. They are described in the next section.

Eventually, these categories could be used as tags for the corpus. For the current purpose, however, all I needed was to identify a set of potentially complex 2- to 4-syllable words that could be presented for analysis to three other Cantonese-speaking linguists.

Simple words	Monomorphemic	
	Without internal morphological structure	Examples: 論盡 /lɵn^{22}tsɵn^{22}/ 'clumsy' 海洛英 /hɔːi^{35}lɔːk^{2}jeŋ^{55}/ 'cocaine'
Complex words	Polymorphemic	
	Internal morphological structures:	Examples:
	Subject-predicate 主謂 /tsyː^{35}wɐi^{22}/	胃痛 /wɐi^{22}t^{h}oŋ^{33}/ 'stomach ache' (lit. 'stomach ache') 肚餓 /t^{h}ou^{23}ŋɔː^{22}/ 'hungry' (lit. 'tummy hungry')
	Modifier-head 偏正 /p^{h}iːn^{55}tseŋ^{33}/	法庭 /faːt^{3}t^{h}eŋ^{21}/ 'court' (lit. 'law court') 人權 /jɐn^{21}k^{h}yːn^{21}/ 'human rights' (lit. 'human rights')
	Verb-object 動賓 /toŋ^{22}pɐn^{55}/	讀書 /tok^{2}syː^{55}/ 'to study' (lit. 'study books') 食飯 /sek^{2}faːn^{22}/ 'to eat lunch/dinner' (lit. 'eat rice')
	Verb-complement 動補 /toŋ^{22}pou^{35}/	沾污 /tsiːm^{55}wuː^{55}/ 'to make dirty' (lit. 'stain dirty') 飲醉 /jɐm^{35}tsɵy^{33}/ 'to be drunk' (lit. 'drink drunk')
	Parallel 並列 /peŋ^{22}liːt^{2}/	利益 /lei^{22}jek^{5}/ 'profit' (lit. 'profit advantage') 改變 /kɔːi^{35}piːn^{33}/ 'change' (lit. 'change change')
	Affixation 附加 /fuː^{22}kaː^{55}/	耳仔 /jiː^{23}tsɐi^{35}/ 'ear(s)' (lit. 'ear suffix') 問完 /mɐn^{22}jyːn^{21}/ 'finished asking' (lit. 'ask finished')
	Reduplication 重叠 /tsoŋ^{21}tiːp^{2}/	See Table 4.5

Table 4.7 Simple word and morphological structures for complex words.

4.1.9 Tagging conventions for the HKCAC

Tagging was done in the single large orthographic transcription file whose format was converted from the original HKCAC data files (cf. section 4.1.5; see also Appendix C). The symbols that labeled speech events in the original HKCAC Excel files were translated into XML tags for this study. Table 4.8 – 4.10 list these tags and other tags that were added in the analysis. A program was written to insert the same tags into the corresponding places in the phonetic transcription file.

Speech events	HKCAC symbols	XML tags in this study
pause	…	<pause></pause>
speech overlap between two people	[overlapping speech]	<overlap2p></overlap2p>
speech overlap among three or more people	[[overlapping speech]]	<overlap2+p></overlap2+p>
indecipherable speech regardless of its length	x	<indeci-speech></indeci-speech>
uncertain hearing — uncertain speech that can only be vaguely made out by the transcriber	<uncertain hearing>	<uncertain-speech></uncertain-speech>
non-speech sounds including laughs, coughs, sniffs, burps, yawns, inhalation, exhalation	@	<nspeech></nspeech>

Table 4.8 HKCAC tags and corresponding XML tags.

Speech events	(XML) tags in this study
beginning of a new line in the orthographic transcription file	O
beginning of a new line in the phonetic transcription file	P
section header	<header></header>
word length, where # is the number of syllables ranging from 1 through 8	<L#></L#>
proper name	<pn></pn>
proper name including generic word(s) and proper name(s)	<gpp></gpp>
abbreviation	<abb></abb>
number	<nu></nu>
vocalized hesitation (including slips of the tongue)	<vh></vh>
glottal stop	<glottal-stop></glottal-stop>

Table 4.9 Other tags used in this study.

1.	The word length tags are always the outer-most tags, if a word also has other tags. E.g. a. <L2><pn>香港</pn></L2> <L2><pn>hJN1NON2</pn></L2> 'Hong Kong' b. <L2><pn>Cindy</pn></L2> <L2><pn>sIndi</pn></L2> 'Cindy' c. <L2><abb>環保</abb></L2> <L2><abb>wan4pou2</abb></L2> 'environmental protection'
2.	Shorter words that are embedded in long words have their own word-length tags. E.g. a. <L3><gpp><L1><pn>何</pn></L1><L2>小姐</L2></gpp></L3> <L3><gpp><L1><pn>hO4</pn></L1><L2>siu2tsE2</L2></gpp></L3> 'Miss Ho' b. <L8><nu>二一四二八一零四</nu></L8> <L8><nu>ji6jAt1sei3ji6pat3At1liN4sei3</nu></L8> '21428104 (telephone number)' c. <L4>cov<L3>唔<L2>cover</L2></L3></L4> <L4>kHAp1<L3>mc4<L2>kHAp1fa4</L2></L3></L4> 'cover or not'
3.	However, suffixes and other bound morphemes are not surrounded by word-length tags. E.g. <L3><L2>電視</L2>機</L3> <L3><L2>tin6si6</L2>kei1</L3> 'television set'

Table 4.10 Illustration of conventions for inserting the <L#></L#> word-length tags.

4.2 Word-internal morphosyntactic branching structure annotation for words

The morphosyntactic analysis that I did in the second-pass segmentation can be viewed as the initial phase of developing a fully annotated HKCAC, which can be used for other purposes such as developing a syntactic parser for a text-to-speech system. To be good for such general purposes, these initial analyses should be confirmed in all inter-

annotator reliability experiment that can evaluate the extent to which other speakers of Cantonese share my intuitions about lexical categories and the grouping of particular morphemes into larger words.

To test the influence of the potential factor morphosyntactic relationship on syllable fusion, I did not need such a complete elaborated tag set. However, I did need a set of candidate stimulus words with a reliable analysis of the branching structures. This section explains the process by which this analysis of the internal branching structures of the candidate test words was obtained.

4.2.1 Extracting a list of words for morphosyntactic branching structure annotation

Not all the words in the corpus would be needed for testing syllable fusion. To extract a subset of words from the corpus, a list of all the word entries and their corresponding pronunciation was first generated. Identical entries in the list were collapsed. There were cases where the same word was listed several times because there was pronunciation variation or because different orthographic representations were used for different tokens. These variant forms of the same word were not collapsed into a single entry at this point (although they were collapsed at a later stage, see section 4.4(2)). Words that were potential stimuli for testing syllable fusion were extracted from this word list. The criteria for the selection and exclusion of words are given below. The list of words (potential stimuli) was then given to three linguists who were native speakers of Hong Kong Cantonese to make the word internal morphosyntactic branching structure. The word list that was given to these three annotators contained only the orthographic forms of the words.

4.2.2 Selecting the words for morphosyntactic annotation

The basic principles for selecting the set of words for analysis were determined by the research purpose of this dissertation. Thus, the words to be selected (i) should be at least two syllables long; (ii) should have pronunciation in Cantonese attested in the HKCAC, since this dissertation studies syllable fusion in Cantonese; and (iii) should be able to appear in sentence initial, medial, and final positions. Any words that did not meet all three of these criteria were not selected for analysis. Two other sets of words were also excluded, namely, numbers and proper names. Numbers were not selected for further analysis, since their morphosyntactic properties seem different enough that numbers should be examined in a separate study focusing on how numbers can be fused. Proper names were not selected since their frequencies are more idiosyncratic and time sensitive.

To elaborate, then, all word entries that met the following criteria were included in the set of words given to the three linguists for morphosyntactic analysis.

<u>Selection criteria:</u>

1. The word had at least two syllables.

Monosyllabic words were excluded, since I did not intend to compare word-external to word-internal boundaries. On the other hand, although I intended to use only stimuli that were 2, 3 or 4 syllables in length, I did not exclude longer words. That is, words that were longer than four syllables were also selected since word-internal structure analysis could identify substrings that could be used as stimulus words.

2. The pronunciation of *all* the morphemes in the word entry was in Cantonese.

Cantonese pronunciation of a word here means that the transcribers of the HKCAC transcribed the word using the Cantonese segment-plus-tone configuration (e.g. bai6).

For words written in English orthography, this configuration suggests phonetic adaptation into the Cantonese phonology on the part of the speaker as interpreted by the HKCAC transcribers.

Examples of words selected:

a. <L2>其實</L2> <L2>kHei4sAt6</L2> 'in fact'

b. <L4>cov<L3>唔<L2>cover</L2></L3></L4>
<L4>kHAp1<L3>mc4<L2>kHAp1fa4</L2></L3></L4> 'cover or not'

c. <L2>Bio</L2> <L2>bai6O1</L2> 'biology'

Where production variation of the same word occurred, at least one instance of the word needed to be produced in Cantonese in order for the word to be selected (cf. section 4.2.2 (2)).

Examples of words selected:

a. *okay*

<L2>O.K.</L2> <L2>Okei</L2> [production in English]
<L2>OK</L2> <L2>okei</L2> [production in English]
<L2>Okay</L2> <L2>G1kHei3</L2> [production in Cantonese]
<L2>Okay</L2> <L2>G2kHei2</L2> [production in Cantonese]
<L2>Okay</L2> <L2>GkHei</L2> [production in English]
<L2>Okay</L2> <L2>No3kHei2</L2> [production in Cantonese]
<L2>Okay</L2> <L2>hou2kHei3</L2> [production in Cantonese]
<L2>Okay</L2> <L2>o1kHei1</L2> [production in Cantonese]
<L2>Okay</L2> <L2>o2kHei3</L2> [production in Cantonese]
<L2>Okay</L2> <L2>o3kHei6</L2> [production in Cantonese]
<L2>Okay</L2> <L2>okHei</L2> [production in English]
<L2>Okay</L2> <L2>okei</L2> [production in English]

b. cable

<L2>cable</L2> <L2>kHeIbGl</L2> [production in English]

<L2>cable</L2> <L2>kHei1bou2</L2> [production in Cantonese]

3. Onomatopoeic words were selected. (Note that some Cantonese words have no orthographic representations. These words are transcribed in IPA both in the orthographic and the phonetic transcription files.)

Examples of words selected:

a. <L2>kik6kOk6</L2> <L2> kik6kOk6</L2> 'the sound of walking on wood floor'

b. <L2>pik6pak6</L2> <L2>pik6pak6</L2> 'the sound of something falling down from a height'

4. Abbreviations without an embedded proper name were selected.

Examples of words selected:

a. <L2>三保</L2> <L2>sam1pou2</L2> 'insurance for the third party'

b. <L3><abb>中英文</abb></L3> <L3><abb>tsoN1jiN1mAn4</abb></L3> 'Chinese and English languages'

The followings were not treated as proper names and were thus selected:

(i) Abbreviated names of public service or government offices.

Examples:

a. <L3><abb>醫管局</abb></L3> <L3><abb>ji1kun2kok2</abb></L3> 'Hospital Authority'

b. <L3><L2>審計</L2>處</L3> <L3><L2>sAm2kAi3</L2>tsHy5</L3> 'Audit Department'

(ii) Titles of people.

Examples (the underlined are not treated as proper names):

a. <L3><gpp><L1><pn>蔣</pn></L1><u><L2>教授</L2></u></gpp></L3>

<L3><gpp><L1><pn>tsJN2</pn></L1><u><L2>kau3sAu6</L2></u></gpp></L3>

'Professor Chang'

b. <L3><gpp><L1><pn>曾</pn></L1><u><L2>先生</L2></u></gpp></L3>

<L3><gpp><L1><pn>tsAN1</pn></L1><u><L2>sin1saN1</L2></u></gpp></L3>

'Mr. Tsang'

(iii) Laws, ordinances, and regulations.

Examples:

a. <L3><L2>基本</L2>法</L3> <L3><L2>kei1pun1</L2>fat3</L3>

'Basic Law'

b. <L4><L2>國旗</L2><L2>條例</L2></L4>

<L4><L2>kwOk3kHei4</L2><L2>tHiu4lAi6</L2></L4>

'national flag ordinance'

The following types were excluded from the set of words given to the three linguists for morphosyntactic structure analysis.

<u>Exclusion criteria:</u>

1. One syllable words.

Examples: <L1>我</L1> <L1>O5</L1> 'I/me'

<L1>新</L1><L1>嘅</L1> <L1>sAn1</L1><L1>kE3</L1>

'new nominal particle'

2. Words that were proper names.

Proper names are names of songs, radio programs, people, organizations, declarations and treaties, street names, and countries. The generic part of the name (such as "company", "university", etc) was also treated as part of the proper name, if any.

Examples of proper names:

a. <L4><pn><L2><pn>中華</pn></L2><L2><pn>民國</pn></L2></pn></L4>
<L4><pn><L2><pn>tsoN1wa4</pn></L2><L2><pn>mAn4kOk3</pn></L2>
</pn></L4>

'The Republic of China'

b. <L4><pn><L2><pn>百樂</pn></L2><L2><pn>戲院</pn></L2></pn></L4>
<L4><pn><L2><pn>pak3lOk6</pn></L2><L2><pn>hei3jyn2</pn></L2></pn>
</L4>

'Park Lok Cinema'

c. <L4><pn><L2><pn>人權</pn></L2><L2><pn>公約</pn></L2></pn></L4>
<L4><pn><L2><pn>jAn4kHyN4</pn></L2><L2><pn>koN1jJk3</pn></L2>
</pn></L4>

'Convention on Human Rights'

3. Long words that contained an embedded a proper name (<pn></pn>), regardless of word length.

Examples of long words that were excluded since they were embedded within a proper name (<pn></pn>):

a. <L3><pn><L2><pn>數碼</pn></L2><L1><pn>港</pn></L1></pn></L3>
<L3><pn><L2><pn>sou3ma5</pn></L2><L1>kON2</L1></pn></L3>

'Cyber Port'

b. <L4><L2><pn>中國</pn></L2><L2>概念</L2></L4>
<L4><L2><pn>tsoN1kOk3</pn></L2><L2>kHOi3nim6</L2></L4>
'Chinese concept'

c. <L4><gpp><L1>阿</L1><L3><gpp><L1><pn>蔣</pn></L1><L2>教授
</L2></gpp></L3></gpp></L4>
<L4><gpp><L1>a3</L1><L3><gpp><L1><pn>tsJN2</pn></L1><L2>kau3sAu
6</L2></gpp></L3></gpp></L4>
'Professor Chang'

4. Long words that were embedded with a shorter word that was not produced in Cantonese (cf. section 4.2.2(2)).

Examples of long words that were excluded because they contained an embedded a word that was not produced in Cantonese (cf. section 4.2.2(2)):

a. <L4><L2>自認</L2><L2>dummy</L2></L4>
<L4><L2>tsi6jiN6</L2><L2>dVmi</L2></L4>
'self-professed to be a dummy'

b. <L2>check 咗</L2> <L2>tSHektsO2</L2>'already checked'

c. <L2>U 型</L2> <L2>ju:jiN4</L2> 'U type'

d. <L2>T-shirt</L2> <L2>tHi:-s0t1</L2> 'T-shirt'

5. Word types that did not have at least one token that was produced in Cantonese. That is, every token was represented on the IPA row with a transcription *not* using the Cantonese segment-plus-tone configuration thus suggesting all forms due to code-switching on the part of the speaker.

Examples of words excluded:

a. <L2>CD</L2> <L2>si:di:</L2> ‘CD’

b. <L2>Mummy</L2> <L2>ma:mi:</L2> ‘mother’

c. <L2>Daddy</L2> <L2>dYdi</L2> ‘father’

<L2>Daddy</L2> <L2>dYi</L2>

d. <L2>e-mail</L2> <L2>i-meIl</L2> ‘e-mail’

<L2>email</L2> <L2>jimYl</L2>

6. Sentence particles.

Examples of words excluded:

a. <L1>喺</L1> <L1>wO3</L1>

b. <L1>咋</L1> <L1>tsG3</L1>

c. <L1>嗰</L1> <L1>pO3</L1>

7. Numbers

Examples of words excluded:

a. <L2><nu>十九</nu></L2> <L2><nu>sAp6kAu2</nu></L2> ‘19’

b. <L3><nu>一零四</nu></L3> <L3><nu>jAt1liN4sei3</nu></L3> ‘104’

8. Interjections

Examples of words excluded:

a. <L2>哎吔</L2> <L2>Ai1ja3</L2> ‘ouch’

b. <L2>hm6hm2</L2> <L2>hm6hm2</L2> ‘the sound made to acknowledge the listener has heard the speaker’

9. Laughter

An example of words excluded:
<L2>哈哈</L2> <L2>hG1hG1</L2> 'laughter'

10. Any type that was not speech, including pauses (<pause></pause>), undecipherable speech (<indeci-speech></indeci-speech>), uncertain speech that could only be vaguely made out by the transcriber (<uncertain-speech></uncertain-speech>), non-speech sounds (<nspeech></nspeech>), and vocalized hesitations, including slips of the tongue (<vh></vh>).

Also the one apparent word type that I could not parse was excluded. This one example was: <L2>mE3E3</L2> <L2>mE3E3</L2> '??'.

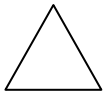
4.2.3 Parsing morphosyntactic branching structure for words

A total of 7414 word types fit the inclusion criteria above and were extracted from the HKCAC. The word length ranged from two to seven syllables. The full list of 7414 words was given to the three linguists for morphosyntactic analysis. Specifically, they were asked to explicitly annotate the morphosyntactic branching structures internal to the words. That is, for each candidate stimulus, the annotators were asked to determine first whether the string of syllables was a monomorphemic or a polymorphemic form, and then for a polymorphemic form longer than 2 syllables, to decide whether there was any word-internal grouping structure below the whole-word node. They would then draw the internal structure on top of the word. For cases where more than one possible analysis was possible, they were asked to give the structure that best agreed with their native intuition. They were not asked to label the nodes, since all that I needed to choose the stimuli for Experiments II and III was the word-internal branching structures. For any word entry that the annotators did not understand, they were asked to put a question mark after it. The annotators were paid for their work.

To help in the analysis, I gave the annotators a set of annotation guidelines along with the word list. The construction of the guidelines made reference to Cheung (1972), Rao et al. (1991:304-327), and Matthews and Yip (1994:31-54). The full set of guidelines is given in Appendix D. To be pointed out here are some of the specific instructions about the basis for any word-internal branching structures.

The distinction between monomorphemic (simple) words and polymorphemic (complex) words needed to be indicated in the annotation because the distinction affected the morphosyntactic relationship between the participating morphemes. In other words, the syllables are twins within a monomorphemic word, but they are sisters within a polymorphemic word. To indicate the distinction, the annotators were asked to indicate *with a horizontal line* within the word for monomorphemic words — i.e. words that *cannot* be further decomposed as the annotators' native intuition judged it (see Figure 4.5 below). For polymorphemic words — i.e. words that *can* be further decomposed into the relationships modifier-head, verb-object, verb-complement, parallel, affixation, or reduplication, as the annotators' native intuition judged it, the annotators were asked to indicate *without a horizontal line* within the word (see Figure 4.6 below). Examples of words that have more than a single level of grouping are given in Figures 4.7 and 4.8. No claims are made about lexicalization, and no distinction is made between phrase and long words. The groupings simply group things into potential compounds, following the structures in Table 4.7.

The annotators were asked to group the elements within the words, from the most closely related elements to the next closely related elements. To illustrate, Figure 4.7 below shows that the string /tsyː^{55}kuː^{55}lek^{5}/ 'chocolate' is a monomorphemic string and the morphemes are twins to each other, and 'chocolate' is then combined with /pɛːŋ^{35}/ 'cookie' to make the long word 'chocolate cookie'. Figure 4.8 shows that 'chocolate cookie' is then combined with /sek^{2}/ 'eat to make an even longer word.



馬　虎

/maː^{55}fuː^{55}/

‘not detail-oriented, careless’

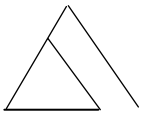
Figure 4.5

炒 麵

/ts^{h}aːu^{35}miːn^{22}/

‘fried noodles’

Figure 4.6

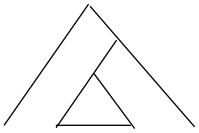


朱古力　餅

/tsyː^{55}kuː^{55}lek^{5}pɛːŋ^{35}/

‘chocolate cookie’

Figure 4.7



食　朱古力　餅

/sek^{2}tsyː^{55}kuː^{55}lek^{5}pɛːŋ^{35}/

‘eat chocolate cookie’

Figure 4.8

4.2.4 Assigning morphosyntactic branching structures to words

The assignment of morphosyntactic branching structures to words observed the majority rule principle — i.e. the words were assigned the structure that at least two out of the three annotators agreed. Where the application of the majority rule would result in inconsistency in the assignment of branching structure to a group of words that did not appear to have different patterns in terms of the concatenation of the morpheme string, lexical category(s) of the morphemes and the meaning of the morphemes, the principle of consistency for the group of words overrode majority ruling for individual words in the group. What appeared to have contributed to the inconsistency was that for the group of words that did not appear to show any difference in their pattern but had more than one possible analysis, at least two annotators agreed on one analysis for some words, while at

least two annotators agreed on a different analysis for other words in the same group. Thus, this inconsistency hinged on the fact that sometimes the individual annotators were not consistent in their analyses of words that do not appear to have different patterns. See Figures 4.9 – 4.12 for two examples of such cases. Figure 4.9 shows the two different majority-rule patterns assigned to verb-object cases where the verb was a monosyllabic stem followed by the aspectual particle 咗 /tsɔː^{35}/. Table 4.11 lists more examples of words that had this syntactic form and which were assigned the two different structures by majority rule. Figure 4.10 shows an analogous case with 得 /tɐk^{5}/, a descriptive complement, Figure 4.11 shows an analogous case with 得 /tɐk^{5}/, a potential complement, and Figure 4.12 shows an analogous case with 唔 /m̩^{21}/.

When the application of the majority rule principle would result in inconsistency in the assignment of branching structure to the words in this way, it was overridden by the principle of consistency, which assigned branching structures that would be consistent across the words that shared the same patterns. See Figures 4.9 – 4.12 for illustrations of the principle of consistency. 1076 entries (or 14.5 % of the total of 7414 word entries) had their structures assigned this way.

The principle of majority rule could not be applied to words that had at least three possible analyses and were analyzed in three different ways (i.e. none of the three annotators agreed on an analysis). In these cases, I chose one that would be consistent with analysis of entries that I interpreted to be variant forms of the entries in question, or assigned one that followed the principle of consistency (45 entries (or 0.6% of the total of 7414 word entries) had their structures assigned this way (see Table 4.12 for the list)), or chose one among the three analyses that agreed with my intuition best.

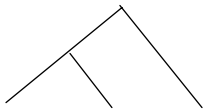
I also rejected the majority parse in favor of a minority analysis in a small handful of words that seemed to be misparsed by two annotators due to insufficient contextual information. There were five entries (or 0.07% of the total of 7414 word entries) that had their structures corrected in this way (see Table 4.13 for the list).

Another type of case where majority rule could not apply was words where one or more annotators failed to parse the form — i.e. they gave "?" as the response. These were very few and were resolved in the same way as cases where the three annotators produced three different analyses. Word entries that all annotators failed to parse were excluded. There were two in total (or 0.03% of the total of 7414 word entries; see Table 4.14 for the list).

The list of circumstances in which the majority rule applied straightforwardly is given in Table 4.15. A total of 6286 entries (or 84.8% of the total of 7414 word entries) had their structures assigned this way. That is, these were the ones where all three annotators provided an analysis and the principle of consistency did not override majority rule in any way.

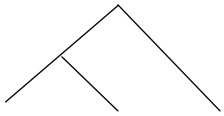
A summary of the final assignment of word-internal branching structures to the words in the word list that was given to the annotators is given in Table 4.17.

(i) Left branching



V	VPT	N
犯	咗	法
/faːn^{22}	tsɔː^{35}	faːt^{3}/
infringe	VPT	law

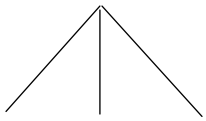
'infringed upon the law'



V	VPT	VPT
入	咗	嚟
/jɐp^{2}	tsɔː^{35}	lei^{21}/
enter	VPT	come

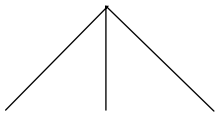
'got in'

(ii) Single-level branching



V	VPT	N
入	咗	稟
/jɐp^{2}	tsɔː^{35}	pɐn^{35}/
file	VPT	report

'file a persecution'



V	VPT	VPT
移	咗	去
/jiː^{21}	tsɔː^{35}	hɵy^{33}/
move	VPT	go

'moved to'

The principle of consistency stipulated that all verbal particles were suffixes that attached to the preceding verb to form short word and the verb (the verb-suffix form) then would group with other morphemes to form a longer word. This stipulation "corrected" the inconsistency in the final assignment of the branching structure to be left branching for the examples above.

Figure 4.9 Verb-complement and verb-object compound words involving the aspectual verbal particle 咗 /tsɔː^{35}/ 'finished'. (i) Left-branching structures vs. (ii) single-level branching structures for trisyllable V-X cases with 咗 /tsɔː^{35}/ 'finished'. The words in question all have the same syntactic pattern of V-咗-X, but those in (ii) would have been assigned different branching structures if majority rule was followed.

(i)

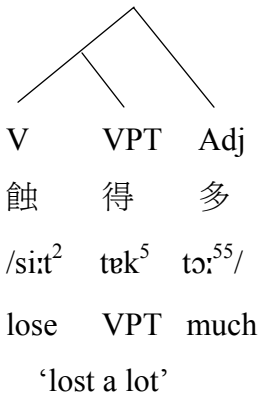
Verb-object compound words:	
上咗車 /sœːŋ^{23}tsɔː^{35}ts^{h}ɛː^{55}/	literally, ‘board VPT car’, ‘got on the car’
做咗嘢 /tsou^{22}tsɔː^{35}jɛː^{23}/	literally, ‘do VPT thing’, ‘did things’
Verb-complement compound words:	
抹咗去 /maːt^{3}tsɔː^{35}hɵy^{33}/	literally, ‘wipe VPT away’, ‘wiped away’
調咗去 /tiːu^{22}tsɔː^{35}hɵy^{33}/	literally, ‘allocate VPT away’, ‘allocated to’

(ii)

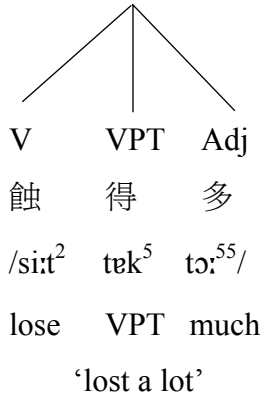
Verb-object compound words:	
撻咗火 /t^{h}aːt^{5}tsɔː^{35}fɔː^{35}/	literally, ‘lit VPT fire’, ‘lit up’
加咗稅 /kaː^{55}tsɔː^{35}sɵy^{33}/	literally, ‘add VPT tax’, ‘increased tax’
Verb-complement compound words:	
彈咗去 /taːn^{22}tsɔː^{35}hɵy^{33}/	literally, ‘bounce VPT go’, ‘bounced away’
攞咗嚟 /t^{h}aːt^{5}tsɔː^{35}fɔː^{35}/	literally, ‘take VPT come’, ‘brought in’

Table 4.11 (i) Examples of other words that were assigned left branching structure by majority rule, and (ii) examples of other words that were assigned with single-level branching structure by majority rule.

(i) Annotators B and C gave a left-branching structure:



Annotator A gave an all-branching structure:



Majority rule assigned the left-branching structure to the word 'lost a lot'.

(ii) All annotators A, B and C gave an all-branching structure (although one would expect annotators B and C to give a left-branching structure as the pattern of this word is not different from that in (i) above):

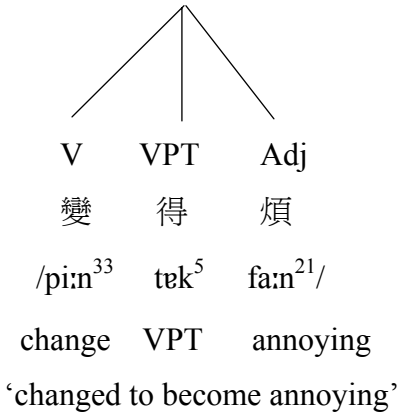


Figure 4.10 Verb-complement compound words involving the verbal particle 得 /tɐk^{5}/, a 'descriptive complement' (see Cheung 1972:125) that is attached to the verb. The adjective that follows the complement describes the verb. In (i) majority rule was left-branching, whereas in (ii) it was single-level branching.

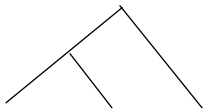
Figure 4.10 continued

Majority rule assigned single-level branching structure to the word ‘changed to become annoying’. However, the pattern of the word for ‘changed to become annoying’ is not different from that of ‘lost a lot’.

The principle of consistency stipulated that all verbal particles were suffixes that attached to the preceding verb to form short word and the verb (the verb-suffix form) then would group with other morphemes to form a longer word. This stipulation “corrected” the inconsistency in the final assignment of the branching structure to be left branching for the examples above.

Figure 4.10 ended

(i) Annotators A and C gave a left-branching structure:



V VPT VPT

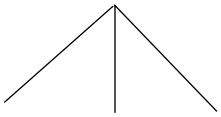
聽 得 到

/t^hɛːŋ^{55} tɐk^5 tou^{33}/

listen VPT reached

'can listen'

Annotator B gave an all-branching structure:



V VPT VPT

聽 得 到

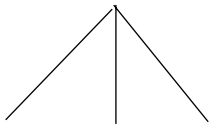
/t^hɛːŋ^{55} tɐk^5 tou^{33}/

listen VPT reached

'can listen'

Majority rule assigned left branching structure to the word 'can listen'.

(ii) All annotators A, B and C gave an all-branching structure (although one would expect annotators A and C to give a left-branching structure as the pattern of this word is not different from that in (i) above):



V VPT VPT

睇 得 到

/t^hɐi^{35} tɐk^5 tou^{33}/

see VPT reached

'can see'

Figure 4.11 Verb-complement compound words involving the verbal particle 得 /tɐk^5/. This 得 /tɐk^5/ is a different verbal particle from the one described in Figure 4.10 above. This 得 /tɐk^5/ is a potential complement (see Cheung 1972:119), which means 'able'. The element (e.g. adjective, directional particle), that follows the potential complement describes the verb.

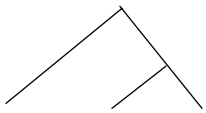
Figure 4.11 continued

Majority rule assigned single-level branching structure to the word ‘can see’. However, the pattern of the word for ‘can listen’ is not different from that of ‘can see’.

The principle of consistency stipulated that all verbal particles were suffixes that attached to the preceding verb to form short word and the verb (the verb-suffix form) then would group with other morphemes to form a longer word. This stipulation “corrected” the inconsistency in the final assignment of the branching structure to be left branching for the examples above.

Figure 4.11 ended

(i) Annotator B gave a right-branching structure:



Adj NEG Adj

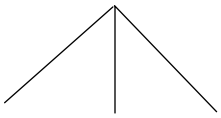
啱 唔 啱

/ŋaːm^{55} m̩^{21} ŋaːm^{55}/

correct NEG correct

'correct or not'

Annotators A and C gave an all-branching structure:



Adj NEG Adj

啱 唔 啱

/ŋaːm^{55} m̩^{21} ŋaːm^{55}/

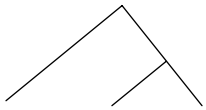
correct NEG correct

'correct or not'

Majority rule assigned right branching structure to the word 'correct or not'.

(ii) Annotator C gave a right branching structure to the word 'okay or not' below (although one would expect annotator C to give an single-level branching structure as the pattern of this word is not different from that in (i) above):

Annotators B and C gave a right-branching structure:



Adj NEG Adj

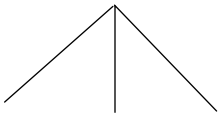
好 唔 好

/hou^{35} m̩^{21} hou^{35}/

okay NEG okay

'okay or not'

Annotator A gave an all-branching structure:



V NEG V

好 唔 好

/hou^{35} m̩^{21} hou^{35}/

okay NEG okay

'okay or not'

Figure 4.12 A-not-A pattern, meaning 'A or not'.

Figure 4.12 continued

Majority rule assigned right branching structure to the word ‘okay or not’. However, the pattern of ‘okay or not’ is not different from that of ‘correct or not’.

The principle of consistency stipulated that all prefixes were attached to the following adjective or verb to form short word first and the short word then would group with other morphemes to form a longer word. This stipulation “corrected” the inconsistency in the final assignment of the branching structure to be right branching for the examples above.

Figure 4.12 ended

Word entries and analysis	Internal structure	Gloss	Remarks
<L3>大不了</L3> /taːi^{22}pɐt^{5}liːu^{23}/	single-level branching	'worst comes to worst'	idiomatic expression; one of the three analyses chosen$
<L3><L2>不得</L2>了</L3> /pɐt^{5}tɐk^{5}liːu^{23}/	right branching	'disastrous, extremely'	idiomatic expression; one of the three analyses chosen
<L3>係咁兒</L3> /hɐi^{22}kɐm^{35}jiː^{35}/	single-level branching	'Do something as a token.'	idiomatic expression; one of the three analyses chosen
<L3>係咁話</L3> /hɐi^{22}kɐm^{35}waː^{22}/	single-level branching	'so to speak'	idiomatic expression; one of the three analyses chosen
<L3>到<L2>陣時</L2></L3> /tou^{33}tsɐn^{22}siː^{21}/	right branching	'when that time comes'	to be consistent with the analysis for <L3>嗰<L2>陣時</L2></L3> /kɔː^{35}tsɐn^{22}siː^{21}/
<L3>潛<L2>意識</L2></L3> /ts^{h}iːm^{21}jiː^{22}sek^{5}/	right branching	'sub-conscious'	one of the three analyses chosen
<L4>公諸於世</L4> /koŋ^{55}tsyː^{55}jyː^{55}sɐi^{22}/	single-level branching	'tell it to the world'	idiom; one of the three analyses chosen
<L4>接<L3>唔<L2>接受</L2></L3></L4> /tsiːp^{3}m̩^{21}tsiːp^{3}sɐu^{22}/	right branching	'accept or not'	assigned a structure that was consistent with the principle of consistency for the analysis of 唔 /m̩^{21}/ as a prefix (see principle of consistency in Table 4.15)

$ This entry might look similar to 'A-not-A' construction at first glance. But unlike 'A-not-A' construction, the last morpheme for this entry might have been a sentence final particle, such that a right branching analysis does not seem appropriate.

Table 4.12 Examples of words whose internal branching structure obtained no agreement among the three annotators.

Misparsed by annotators	Revision	
<L3><L2>女家</L2>鬼</L3> ghost of a female's home	<L3>女<L2>家鬼</L2></L3> a house ghost that is female	/nɵy^{23}kaː^{55}kwɐi^{35}/
<L3><L2>男家</L2>鬼</L3> ghost of a male's home	<L3>男<L2>家鬼</L2></L3> a house ghost that is male	/naːm^{21}kaː^{55}kwɐi^{35}/
<L3><L2>聽</L2>壞機</L3> listen to a broken CD player	<L3><L2>聽壞</L2>機</L3> listen to the CD player and so broke it	/tʰɛːŋ^{55}waːi^{22}kei^{55}/
<L3>做<L2>錯嘢</L2></L3> do things that are already wrong	<L3><L2>做錯</L2>嘢</L3> do things wrong	/tsou^{22}tsʰɔː^{33}jɛː^{23}/
<L3>上<L2>訴庭</L2></L3> (meaning unknown)	<L3><L2>上訴</L2>庭</L3> court of appeal	/sœːŋ^{22}sou^{33}tʰeŋ^{21}/

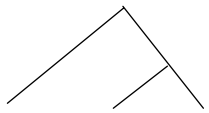
Table 4.13 An exhaustive list of the five words that were misparsed by the annotators because of insufficient contextual information. (I revised the annotation as in the column "Revision".)

公 mui6iN4	/kʊŋ^{55}muːi^{22}ŋ^{21}/
kHAu4kHei4kHei3	/kʰɐu^{21}kʰei^{21}kʰei^{33}/

Table 4.14 The two entries that were excluded because all three annotators failed to parse the form.

1. Regularized analyses for “interrupted” reduplicative forms, which are forms having an interrupted concatenation of an exact copy of a syllable (cf. 4.2).

A 唔 A (A-not-A):



Adj NEG Adj

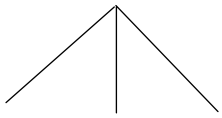
啱 唔 啱

/ŋaːm^{55} m̩^{21} ŋaːm^{55}/

correct NEG correct

‘correct or not’

A 一 A (A-one-A):



V one V

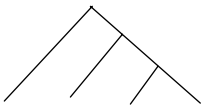
望 一 望

/ŋaːm^{55} m̩^{21} ŋaːm^{55}/

look one look

‘take a look’

A 唔 AB (A-not-AB):



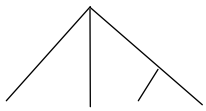
公 唔 公 道

/koŋ^{55}m̩^{21}koŋ^{55}tou^{22}/

fair NEG fair reason

‘fair or not’

A 一 AB (A-one-AB):



插 一 插 咀

/t^{h}aːp^{3}jɐt^{5} t^{h}aːp^{3}tsɵy^{35}/

insert one insert mouth

‘to interrupt a bit’

Tagging:

<L3>A<L2>唔 A</L2></L3>

<L4>A<L3>唔<L2>AB</L2></L3></L4>

<L3>A 一 A</L3>

<L4>A 一<L2>AB</L2></L4>

Table 4.15 Operations that regularized word-internal mophosyntactic branching structure assignment in accord with the principle of consistency.

Table 4.15 continued

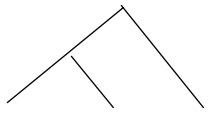
While the pattern ‘A-not-A’ showed inconsistent analysis within the individual annotators (cf. Table 4.15(1)), the single-level branching structure of ‘A-one-A’ obtained majority ruling among the three annotators. The regularizing operation treated ‘A-not-A’ and ‘A-one-A’ as two different structures. For ‘A-not-A’, the negation marker does not stand alone but is prefix to the following verb or adjective. The treatment was to first group the prefix with its following morpheme, and then the prefixed word was grouped with the first morpheme in the word. The structure of ‘A-one-A’ is different from that of ‘A-not-A’. A right branching structure as assigned to ‘A-not-A’ would mislead one into thinking that ‘one’ in ‘A-one-A’ quantifies the following verb or adjective, but ‘V-one-V’ means ‘attempt, to do a bit’, whereas ‘Adj-one-Adj’ means ‘very’. Thus, single-level branching structure which treated ‘one’ in parallel to other elements in the word was assigned to ‘A-one-A’ pattern. The regularizing operation similarly treated ‘A-not-AB’ and ‘A-one-AB’ as two different structures. For ‘A-not-AB’, it can be analyzed as the eclipsed form of ‘AB-not-AB’. The negation marker is a prefix attached to the following verb or adjective. As for ‘A-not-A’ structure, the treatment was to first group the prefix with its following morpheme, and then the prefixed word was grouped with the eclipsed form (i.e. A) in the word. For ‘A-one-AB’, ‘one’ does not quantify the following verb, but it means ‘attempt, to do a bit’. In this pattern, only the first syllable in the disyllabic word can be reduplicated, such that ‘AB-one-AB’ is not acceptable. The treatment was to make ‘one’ in parallel to the other elements in the word.

Table 4.15 continued

2. Verbal particles The principle of consistency stipulated that verbal particles be attached the verbal particles to the verb.		
(i) Directional verbal particles are attached to the preceding verb.		
	V VPT E.g. 擺 落去 /paːi^{35} lɔːk^{2}hɵy^{33}/ put down 'put something down'	Tagging: <L3>擺<L2>落去</L2></L3>
(ii) Aspectual verbal particles are treated as suffixes.		
	V VPT N E.g. 入 咗 稟 /jɐp^{2} tsɔː^{35} pɐn^{35}/ file VPT report 'file a persecution'	Tagging: <L3><L2>入咗</L2>稟</L3>

Table 4.15 continued

(iii) The aspectual particle 咗/tsɔː^{35}/ is attached to the preceding compound word that is in verb-complement relationship.

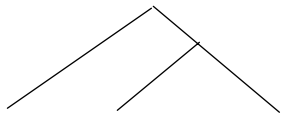


	V	Adj	VPT
E.g.	收	多	咗
	/sɐu^{55}	tɔː^{55}	tsɔː^{35}/
	collect	more	VPT

'more was collected'

Tagging:
<L3><L2>收多</L2>咗</L3>

(iv) The aspectual particle 咗/tsɔː^{35}/ is attached to the preceding compound word that is in subject-predicate relationship.

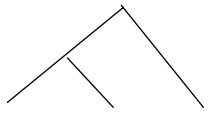


	N	V	VPT
E.g.	質	差	咗
	/tsɐt^{5}	ts^{h}aː^{55}	tsɔː^{35}/
	quality	bad	VPT

'quality got worse'

Tagging:
<L3>質<L2>差咗</L2></L3>

(v) Series of verbal particles are treated as suffixes attached to the verb hierarchically.



	V	VPT	VPT
E.g.	買	過	晒
	/maːi^{23}	k^{w}ɔː^{33}	saːi^{33}/
	buy	VPT	VPT

'buy everything again'

Tagging:
<L3><L2>買過</L2>晒</L3>

Table 4.15 continued

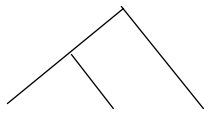
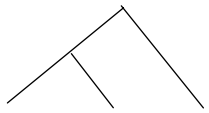
(vi) 得 /tɐk^{5}/, a descriptive complement, is treated as a suffix.  V VPT Adj 蝕 得 多 /siːt^{2} tɐk^{5} tɔː^{55}/ lose VPT much 'lost a lot'	Tagging: <L3><L2>蝕得</L2>多</L3>
(vii) 得 /tɐk^{5}/, a potential complement meaning 'able', is treated as a suffix.  V VPT VPT 聽 得 到 /t^{h}ɛːŋ^{55} tɐk^{5} tou^{33}/ listen VPT reached 'can listen'	Tagging: <L3><L2>聽得</L2>到</L3>

Table 4.15 continued

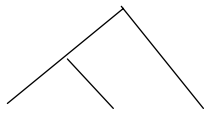
Note: The verbal particle -番 /faːn^{55}/ 'go back, reverse' is exempted from the regularizing operation, but it was the majority ruling among the three annotators applied. In the construction verb+/faːn^{55}/+directional particle (e.g. 擺+番+落去 /paːi^{35}/ + /faːn^{55}/ + /lɔːk^{2}hɵy^{33}/ 'put it back down'), it was difficult to draw a clear-cut line between two possible analyses. 番 /faːn^{55}/ can be the suffix to the preceding verb to mean reversal (of state, etc) or it can be grouped with the following directional verbal particle (cf. Cheung 1972). The morpheme could cross these two categories often times. Assignment of word-internal structure for this construction, therefore, followed the majority ruling among the annotators.
3. Negation 唔 /m̩^{21}/ (i) The negation was treated as a prefix of the following verb.  E.g. NEG V N 唔 入 袋 /m̩^{21} jɐp^{2} tɔːi^{35}/ NEG put in bag 'not put into a bag' Tagging: <L3><L2>唔入</L2>袋</L3>

Table 4.15 continued

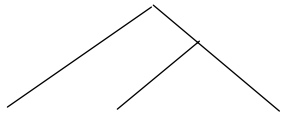
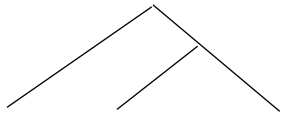
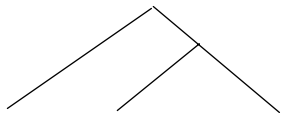
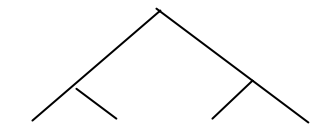
(ii) The negation was treated as part of the verbal particle.  V NEG VPT E.g. 瞓 唔 到 /fɐn^{33} m̩^{21} tou^{35}/ sleep NEG VPT 'cannot sleep'	Tagging: <L3>瞓<L2>唔到</L2></L3>
(iii) For an adjectival compound word that is made up of an adjective followed by a verb, the principle of consistency subjected the compound under the scope of negation.  NEG Adj V E.g. 唔 難 明 /m̩^{21} naːn^{21} meŋ^{21}/ NEG difficult understand 'not difficult to understand'	Tagging: <L3>唔<L2>難明</L2></L3>
4. Ellipsis  N N N E.g. 中 英 文 /tsoŋ^{55} jeŋ^{55} mɐn^{21}/ Chinese English language 'the Chinese and the English languages'	Tagging: <L3><abb>中英文</abb></L3> The 'underlying' form for the eclipsed form was analyzed to be 中文 Conj 英文 language ('Chinese language Conj English language') with ellipsis.

Table 4.15 continued

5. Idioms, which usually involving classical Chinese components, were parsed as if they were their literal components.

E.g.

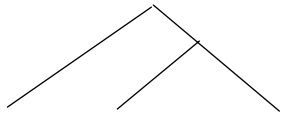


QF	N	V	ProN
一	笑	置	之
/jɐt^{5}	siːu^{33}	tsiː^{33}	tsiː^{55}/
one	smile	take care	it

'dismissed with a smile'

Tagging:
<L4>一笑置之</L4>

6. Treatment of the verb 去 /hɵy^{33}/#

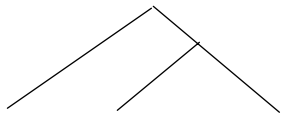


E.g.

V	V	N
去	畀	錢
/hɵy^{33}	pei^{35}	ts^{h}iːn^{35}/
go	give	money

'go give money (to someone)'

Tagging:
<L3>去<L2>畀錢</L2></L3>



E.g.

V	V	N
去	幫	人
/hɵy^{33}	pɔːŋ^{55}	jɐn^{21}/
go	help	people

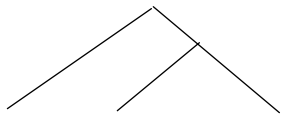
'go help people'

Tagging:
<L3>去<L2>幫人</L2></L3>

Table 4.15 continued

7. Treatment of the verb 有 /jɐu^{23}/ ‘have’ and 冇 /mou^{23}/ ‘does not have’

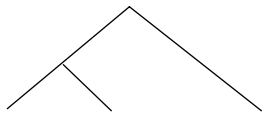
有 /jɐu^{23}/ ‘have’ and 冇 /mou^{23}/ ‘not have’ are verbs when not followed by a verb, but are treated as auxiliary verbs when followed by a verb.



	V	Wh-	V
E.g.	有	咩	事
	/jɐu^{23}	mɛː^{55}	siː^{22}/
	have	what	matter

‘what is the matter’

Tagging:
<L3>有<L2>咩事</L2></L3>



	V	Wh-	V
# E.g.	有	咩	講
	/jɐu^{23}	mɛː^{55}	kɔːŋ^{35}/
	have	what	say

‘what does someone want to say’

Tagging:
<L3><L2>有咩</L2>講</L3>

The construction is called 連謂結構 /liːn^{21}wɐi^{22}giːt^{3}kɐu^{33}/ ‘serial verb construction’ (cf. e.g. Cheung 1972: 106; Li & Thompson 1981: 607-622; Zhu 1982: 160-173). The hierarchical analysis given was my working hypothesis, since I did not found a model that I could follow, apart from descriptions of the linear components.

Table 4.15 ended

Majority rule was used under the following circumstances:
1. The assignment of whether a word was monomorphemic or polymorphemic.
2. The assignment of structures to structurally ambiguous word entries, where the structural ambiguity was not due to a lack of contextual information for the annotators. For example, majority rule decided that the long word 結婚週年紀念 /kiːt^{3} fɐn^{55} tsɐu^{55} niːn^{21} kei^{35} niːm^{22}/ '(literally) wedding-annual-commemorate' should be parsed as [[[結婚] [週年]] 紀念] and not as [結婚 [[週年] [紀念]]].
3. Any other circumstances not stated in Table 4.15.

Table 4.16 The use of majority rule.

Categories	No. of entries	%
Majority rule applied	6230	84.0%
Regularizing operations in accord with principle of consistency applied	1132	15.3%
None of the annotators agreed#	45	0.6%
Word entries misparsed by annotators	5	0.07%
Word entries that all annotators failed to parse	2	0.03%
Total	7414	100%

There were 186 word entries that none of the annotators agreed, among which 141 were regularized by regularizing operations that were in accord with the principle of consistency. These 141 entries were counted into the category of "Regularizing operations in accord with principle of consistency applied", and thus, 45 were in the current category.

Table 4.17 Summary of final assignment of word-internal branching structures to the words in the word list that was given to the annotators to analyze.

4.3 Word list generated after re-tagging the relevant words

7414 out of the total of 9458 words (or 78.4% of the words) were re-tagged according to the principles for morphosyntactic analysis as illustrated above. The 2044 word entries (or 21.6% of the words) that did not need to be re-tagged were proper names, English words and numbers. These word entries were not given to the annotators to analyze. XML tags were used according to the tagging conventions laid out in section 4.1.9. The program that was written for automatically mapping the orthographic transcription in the orthographic transcription file onto their corresponding phonetic transcription in the concatenated phonetic transcription file (which was the concatenation of all the 56 original transcription files) was run again. (Recall that the automatic mapping was run the first time after the word segmentation principles for spoken Cantonese were devised but before the above principles for morphosyntactic analysis were laid down, cf. section 4.1.4.) To see an example of what the two concatenated orthographic and phonetic transcription files looked like after being segmented and tagged, see Appendix C. After the automatic mapping between orthography and phonetic transcriptions was finished, a word list containing all the words from 2 to 4 syllables long was generated. The frequency of occurrence was given for each unique combination of orthography and phonetic transcription. See an example in section 4.4(1) below.

4.4 Extracting target words for testing

After extracting the target words for testing syllable fusion from the word list that contained all the words from 2 to 4 syllables long (see section 4.3), I took the following further steps to "proofread" the list.

1. I looked through the word list to make sure that homophones with the same orthographic representation were not actually different words, which would make separate word entries with the meaning of each word added after the orthographic representation. (For example, /siːn^{55}saːŋ^{55}/ 先生 means 'mister', 'teacher' or 'husband', and thus, /siːn^{55}saːŋ^{55}/ made three separate entries.)

2. I collapsed the counts of variant forms of words. Variant forms were those that involved different orthographic representations, orthographic typos or variant forms of pronunciation of the same word in running speech. For example, the counts of the following entries

orthography	pronunciation	count
<L2>已經</L2>	<L2>ji5kiN1</L2>	83
<L2>已經</L2>	<L2>ji5iN1</L2>	80
<L2>以經</L2>	<L2>ji5kiN1</L2>	1

were collapsed because of the orthographic typo 以 for 已 /jiː^{23}/ and that the different pronunciations were interpreted to be variant forms of pronunciation in running speech. The word <L2>已經</L2> /jiː^{23}keŋ^{55}/ then had the count equaled to 164.

3. I eliminated the ones listed in section 4.2.2 'exclusion criteria'.

4. I sorted the word list by the count and computed the log frequency ($\log_{10}$) of each word entry. The resulting log frequency ranged from 0-2.9. The words then could be neatly divided into three categories:

   low frequency range = 0-0.9 (7947 word entries)
   mid frequency range = 1.0-1.9 (595 word entries)
   high frequency range = 2.0-2.9 (54 word entries)

Table 4.18 displays the distribution of the number of word entries available for selection and the number of target words to be selected. For 4-syllable words, there were cases where the embedded short words crossed frequency category. For example, there

were six 4-syllable words in the mid frequency category, in which the embedded short words were of a different frequency category. These six words were excluded from being selected. Similarly, for 3-syllable words in the mid frequency category, forms which included an embedded shorter word that was in the high or the low frequency category were excluded from selection. These words were excluded from selection because research shows that frequency differences between morphemes within a word would affect the perception of wordhood (see Hay 2002). A total of 173 words would be selected as targets. It can be seen in the table that the frequency of the words of different length and of different internal branching structures were not evenly distributed across the three frequency categories. So, the distribution of the stimuli could not be balanced across all cells. Also, for the sake of the time that would be needed for testing one subject and for the sake of having enough types within a cell for subsequent statistical analysis, not all the structures attested in the HKCAC were tested. This is particularly true for 4-syllable stimuli.

Frequency category / Word length	High	Mid	Low	Total
2 syllables no. available (no. selected)	54 (20)	542 (20)	4496 (20)	5092 (60)
3 syllables no. available (no. selected)	0 (0)	47 (20)	2065 (20)	2112 (40)
4 syllables no. available (no. selected)	0 (0)	6* (0)	1386 (73)	1392 (73)
Total	54 (20)	595 (40)	7947 (113)	8596 (173)

* The embedded short words for these six word entries were of non-mid frequency categories; thus, these six words were excluded from being selected.

Table 4.18 Distribution of candidate stimuli and the distribution of stimuli in parenthesis.

The following considerations led to the numbers of stimuli specified in each cell of Table 4.19. First, the in high frequency category, some cells are empty because only 2-syllable words occurred at rates above the cut-off for this frequency category. Among them, only 6 were monomorphemic and all 6 monomorphemic words were selected. The 14 polymorphemic 2-syllable words selected were the highest frequency words in the high frequency category.

In the mid frequency category, the average log frequency was (0+2.9)/2 = 1.45. The monomorphemic 2-syllable words were first extracted to make a list sorted by frequency. Among these monomorphemic 2-syllable words, the 5 words at and above the average log frequency, and 5 words just below the average log frequency, were selected as stimuli[3]. The polymorphemic 2-syllable words (with the monomorphemic 2-syllable words excluded) were then sorted by frequency. There were 23 candidate stimuli, each having the log frequency of 1.45. These 23 candidate words were randomized. The first 20 were selected as stimuli for the test.

In the low frequency category, only word entries with log frequency = 0 were considered as candidate stimuli. The monomorphemic 2-syllable words were first extracted to make a list. There were 62 of them. After randomization, the bottom 20 were selected as stimuli. The polymorphemic 2-syllable words (with the monomorphemic 2-syllable words excluded) made another list. There were 2352 candidate words, each having the log frequency of 0. These 2352 candidate stimuli were randomized. The bottom 20 were selected as stimuli for the test.

Analogous considerations applied in the selection of the 3-syllable stimuli. First, in the high frequency category, there were no 3-syllable words. In the mid frequency

[3] Later, one stimulus, /siːn^{55}saːŋ^{55}/ 先生 'mister', among the 5 words chosen below the average log frequency was found to have been overlooked as homophones of two other words meaning 'teacher' and 'husband'. The counts of these three words were, therefore, adjusted accordingly. As a result, the log frequency of this stimulus 'mister' was 1.08 in the mid frequency category. The range of log frequency of the other 9 stimuli was between 1.76 and 1.34.

category, all forms which included an embedded shorter word that was in the high or the low frequency category were excluded from selection. There was only one 3-syllable word that was all-branching in its internal structure. The first 10 stimuli were chosen for the right-branching structure from a randomized list of the 3-syllable candidate stimuli, and the first 9 were chosen for the left-branching structure. In the low frequency category, only word entries with log frequency = 0 were made candidate stimuli. Again, all forms that contained embedded short words that were of the high or the mid frequency category were excluded from selection. Only 3 words were monomorphemic and they were all selected as stimuli. For the other three branching structures, 3-syllable candidate stimuli of each structure made a list, and thus, three separate lists. Stimuli for each structure were selected from the top of the randomized lists.

Similarly for the 4-syllable stimuli, the count is 0 in the high frequency category, because there were no 4-syllable words. In the mid frequency category, once all forms containing embedded short words that were of the high or the low frequency category were excluded from selection, there were no 4-syllable words. In the low frequency category, word entries with log frequency = 0 were made candidate stimuli. However, since there were internal structures that had very limited candidate stimuli, word entries with log frequency = 0.3 were selected as well. There were 3 (out of 73) stimuli that belonged to this case. 4-syllable candidate stimuli of each branching structure made a list, and thus, 7 separate lists (the balanced branching category had 3 lists). Stimuli for each structure were selected from the top of the randomized lists.

The distribution of stimuli by word-internal branching structure is given in Table 4.20, and distribution of stimuli by word-internal morphosyntactic relationships in Table 4.20. The list of stimuli for Experiments II and III is given in Appendix E.

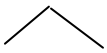
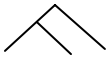
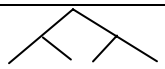
Frequency category / Word length	High	Mid	Low	Word-internal branching structure (cf. Figs.1.1 – 1.4)
2 syllables	14	10	10	
	6	10	10	
+)	(20)	(20)	(20)	
3 syllables	0	1	4	
	0	0	3	
	0	9	6	
	0 0	1 9	0 7	 
+)	(0)	(20)	(20)	
4 syllables	0	0	10	
	0	0	2	
	0 0	0 0	1 14	 
	0	0	16	
	0	0	30 break downs: [13] [4] [13]	 A A B B A B A B Others
+)	(0)	(0)	(73)	

Table 4.19 Distribution of morphosyntactic structures internal to the stimuli.

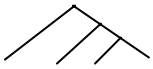
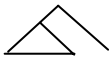
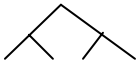
Structure	Visualization			Row total
	2 syllables	3 syllables	4 syllables	
No branching				31 words
Single-level branching				29 words
Right branching	---			30 words
	---	---		
Left branching	---			31 words
	---		---	
Balanced branching	---	---		30 words

Table 4.20 Distribution of stimuli by word-internal branching structure as summarized from Table 4.19.

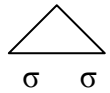
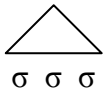
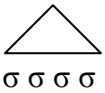
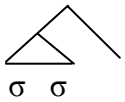
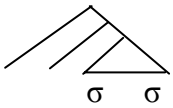
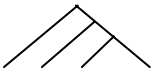
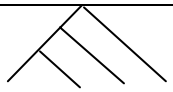
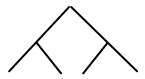
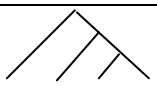
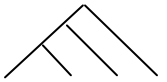
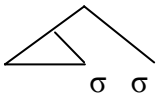
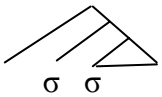
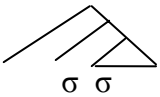
Relationships	Relevant branching structures				Pairs of syllables
Twins					
					40 (11.1%)
Sisters					
					
					195 (54.3%)
Aunt-niece					
					
					94 (26.2%)
Cousins					30 (8.4%)
Total					359 (100%)

Table 4.21 Distribution of stimuli by word-internal morphosyntactic relationships as summarized from Table 4.19.

CHAPTER 5

EXPERIMENT II: TABULATING SYLLABLE FUSIONS IN THE HKCAC

The 173 target words that were selected in the process described in chapter 4 were then used in a quantitative study of the HKCAC. This quantitative study is Experiment II. That is, Experiment II tested the influence of word frequency, word length, and word-internal morphosyntactic relationship on syllable fusion by examining the transcriptions given for all tokens of these words in the spontaneous speech corpus — the Hong Kong Cantonese Adult Language Corpus (HKCAC).

5.1 Procedures

For each of the 173 target words, variant pronunciations were identified in the transcriptions, and tokens of each variant were counted. Each variant was then examined to classify the degree of fusion at each word-internal syllable boundary. To facilitate the examination, the HKCAC transcription for each word entry (see "trans.HKCAC" in Table 5.1) was copied into another column in an Excel file. (IPA equivalents are given in the columns in the table. See also Appendix F for the list of ASCII phonetic symbols.) The XML tags were removed from the transcription while the tones given in numbers were replaced by a syllable boundary marker "|" (see the column "copy.trans.HKCAC" in the table). The columns "1st b", "2nd b", and "3rd b" (where "b" stands for "boundary") transcribed the degree of segmental fusion of each pair of two syllables from left to right that straddled across the syllable boundary. Five degrees from 0 (vowel merge) to 4 (no fusion) were identified. The definitions of each degree will be given later in Table 5.3.)

orthography	trans.HKCAC	copy.trans.HKCAC	1st b	2nd b	3rd b
<L2>呢個</L2> *'this one'*	<L2>ji1G3</L2> IPA: ji1ə3$ *ji1-kO3*	ji|G IPA: ji|ə	2	n/a	n/a
<L2>我哋</L2> *'we'*	<L2>NO3ei6</L2> IPA: ŋɔ3ei6 *NO5-tei6*	NO|ei IPA: ŋɔ|ei	2	n/a	n/a
<L2>我哋</L2> *'we'*	<L2>O3ei6</L2> IPA: ɔ3ei6 *O5-tei6*	O|ei IPA: ɔ|ei	2	n/a	n/a
<L3>嗰<L2>方面</L2></L3> *'that aspect'*	<L3>kO2<L2>fON1in6</L2></L3> IPA: kɔ2fɔŋ1in6 *kO2-fON1-min6*	kO|fON|in IPA: kɔ|fɔŋ|in	4	2	n/a
<L3>返<L2>屋企</L2></L3> *'go home'*	<L3>fa1<L2>ok1kHei2</L2></L3> IPA: fa1ok1kʰei2 *fan1-ok1-kHei2*	fa|ok|kHei IPA: fa|ok|kʰei	1	4	n/a
<L4><L2>拗下</L2><L2>拗下</L2></L4> *'scratching'*	<L4><L2>Au1ha5</L2><L2>au1ha5</L2></L4> IPA: ɐu1ha5au1ha5 */au1-ha5-/au1-ha5*	Au|ha|au|ha IPA: ɐu|ha|au|ha	4	4	4
<L4><L2>拗下</L2><L2>拗下</L2></L4> *'scratching'*	<L4><L2>NAu1ha5</L2><L2>Nau1wa5</L2></L4> IPA: ŋɐu1ha5ŋau1wa5 *Nau1-ha5-Nau1-ha5*	NAu|ha|Nau|wa IPA: ŋɐu|ha|ŋau|wa	4	4	3
<L4><L3><L2>搵嚟</L2>工</L3>做</L4> *'finding a job'*	<L4><L3><L2>wAn2NAn2</L2>koN1</L3>tsou6</L4> IPA: wɐn2ŋɐn2koŋ1tsou6 *wAn2-kAn2-koN1-tsou6*	wAn|NAn|koN|tsou IPA: wɐn|ŋɐn|koŋ|tsou	3	4	4

Table 5.1 Examples of the 173 target words showing the transcription of the degree of syllable fusion for pairs of syllables that crossed the word-internal syllable boundary in the HKCAC. The forms in italics are information supplied for the readers here. They are the English gloss, under orthography, and the 'underlying' non-fused forms of the words, under trans.HKCAC. Note that there are two different underlying forms for 'we' and 'scratching' because of a synchronic sociolinguistic alternation between pronunciations with initial /ŋ/ and pronunciations with no onset consonant for all historically /ŋ/-initial words. (Also note that IPA transcriptions are shown here to explain the phonetic transcriptions in Columns 2 and 3, but the IPA does not appear in the data file.)

5.2 Analysis

Tokens that shared the same variant pronunciations of the 173 target words (see chapter 4 on the selection of target words and Appendix E for the list of target words) were collapsed for analysis, but variant pronunciations of the same word were *not* collapsed for the current analysis. That is, for example, the following entries for /jɐn^{55}wɐi^{22}/ 因爲 'because' and their counts were not collapsed.

orthography	trans.HKCAC	(in IPA)	count
<L2>因爲</L2>	<L2>jAn1Ai6</L2>	jɐn1ɐi6	19
<L2>因爲</L2>	<L2>jG1Ai6</L2>	jə1ɐi6	18
<L2>因爲</L2>	<L2>jG1At6</L2>	jə1ɐt6	1
<L2>因爲</L2>	<L2>jG1G1</L2>	jə1ə1	8
<L2>因爲</L2>	<L2>jG1wai6</L2>	jə1wai6	1

Regarding the criteria for making sure that the variant forms were tokens of any given target word type, I relied on the interpretations of the HKCAC transcribers. It is possible to deduce these interpretations because of the nature of the Cantonese writing system, which is a logosyllabary, and because fused forms as well as non-fused forms in the HKCAC were transcribed both phonetically and orthographically by the HKCAC transcribers, who were all native speakers of Cantonese. They interpreted what word a syllable or a string of syllables represented and gave an orthographic transcription of each of the syllables according to their interpretation. To transcribe the orthographic form that was interpreted to be a fused form, the transcribers used the conversational contexts of the speakers and their judgment as native speakers to recover the 'underlying' non-fused form of the word and gave orthographic transcription using the transcription principles the HKCAC researchers set forth (cf. section 4.1.5). Thus, the same word transcribed with the same orthographic representation might have variant pronunciations that might be, for example, ascribed to synchronic variation (cf. section

3.1.4.2 for examples) or to different types or different degrees of syllable fusion, or tonal coarticulation, etc. For the current analysis, tokens of the same target word were matched by the same orthographic representation. I went back to the original context to check whenever the following appeared:

(i) For a given word entry, there was the possibility of homophones using the same orthographic representation. For example, tokens of the sequence /siːn^{55}saːŋ^{55}/ 先生 had to be allocated among three words with different frequencies of occurrence since this word-form corresponds to three word meanings in the corpus: 'mister' (12 times), 'teacher' (4 times) or 'husband' (14 times).

(ii) For a given word entry, there were suspected typos. For example, the four entries given in Table 5.2 were discarded because I could not be sure whether the orthographic or the phonetic transcription was a typo. Even though the original context suggested which was more likely to be a typo, I had no access to the audio files to be very sure.

orthography	trans.HKCAC	count	word indicated by the orthography	word indicated by the IPA
<L2>嗰個</L2>	<L2>kO2ti1</L2> IPA: kɔ2ti1	1	/kɔː^{35}kɔː^{33}/ 'that one'	嗰啲 /kɔː^{35}tiː^{55}/ 'that CL.'
<L2>但係</L2>	<L2>tai6hOk6</L2> IPA: tai6hɔk6	1	/taːn^{22}hɐi^{22}/ 'but'	大學 /taːi^{22}hɔːk^{2}/ 'university'
<L2>咁樣</L2>	<L2>kAm2kON2</L2> IPA: kɐm3kɔŋ2	1	/kɐm^{35}jœːŋ^{35}/ 'this way'	咁講/kɐm^{35}kɔːŋ^{35}/ 'say something this way'
<L2>做嘢</L2>	<L2>kON2jE5</L2> IPA: kɔŋ2jɛ5	2	/tsou^{22}jɛː^{23}/ 'do work'	講嘢 /kɔːŋ^{35}jɛː^{23}/ 'to speak'

Table 5.2 An exhaustive list of entries discarded in the HKCAC because of suspected errors.

After removing these suspected types, there were a total of 626 entries representing a transcribed pronunciation of at least one token of one of the 173 target

word types. Word-internal boundaries between all sequences of two syllables in the target form for each of these entries were examined regarding syllable fusion. There were 866 boundaries.

For each of these 866 boundaries, I categorized the pronunciation for degree of fusion transcribed for the two-syllable sequence indicated by that transcription. Five degrees of fusion were noted along the continuum from not fused at all to extremely fused. The definitions for the five degrees of syllable fusion from 0 (highly fused) to 4 (not fused), as well as examples, are given below.

	Degree of fusion	Definition
highly fused	0	***Merging of the vowels*** in the two syllables at a syllable boundary to become a single vowel, whether or not the original target vowels are the same, and whether or not the resulting merged vowel is of an intermediate quality (i.e. vowel coalescence). For example, [kɔː$^{35+33}$] for /kɔː^{35}kɔː^{33}/ 嗰個 'that one', [t^{h}ɛː$^{55+2}$] for /t^{h}eŋ^{55}jɐt^{2}/ 聽日 'tomorrow', and [tsɐi^{22}] for /t^{h}ɐu^{22}hɐi^{22}/ 就係 'is' are transcribed with the fusion degree "0" for the between the two target syllables. (Examples from the HKCAC were not available since there was no instance of syllable fusion to the degree '0'.)
	1	For ***each*** of the two participating syllables at a syllable boundary, there is deletion of ***at least one*** segment that is immediately contiguous to the syllable boundary. However, the whole syllable is not deleted for either of the two syllables. (Nil onset syllables that are not produced with an initial glottal stop are *not* counted as segmental deletion.) Examples of deletion from the HKCAC are [tə^{22}ɐi^{22}] for /taːn^{22}hɐi^{22}/ 但係 'but', [jɐ^{55}ɐi^{22}] for /jɐn^{55}wɐi^{22}/ 因爲 'because', [jɛ^{21}ə^{22}] for /jiːn^{21}hɐu^{22}/ 然後 'then', and [jyː^{2}ei^{21}jyːt^{2}] for /jyːt^{2}lei^{21}jyːt^{2}/ 越嚟越 'more and more'.

Table 5.3 Definitions for the five degrees of syllable fusion from 0 (highly fused) to 4 (not fused).

Table 5.3 continued

	2	For ***either*** of the two participating syllables at a syllable boundary, there is deletion of ***at least one*** segment immediately contiguous to the syllable boundary. However, the whole syllable is not deleted for either of the two syllables. Examples of deletion from the HKCAC are [mɔː^{23}lɔːk^{2}] for /mɔːŋ^{23}lɔːk^{2}/ 網絡 'network' and [jyː^{21}ɔː^{5}] for /jyː^{21}kwɔː^{35}/ 如果 'if'.
	3	There are ***phonological processes*** weakening the segment(s) contiguous to the boundary between the two participating syllables. However, *neither* of the two participating syllables had segments deleted at the syllable boundary. Phonological processes include place/manner/voicing assimilation (which includes partial voicing of voiceless phonemes, where partial voicing is defined as at least half of the segment is voiced, but not the whole segment is voiced), liason, vowel reduction, consonant lenition, incomplete deletion of consonant (e.g. deletion of place for coda stops but retaining the glottal stop that accompanies the coda stops: [ʔ] for /ʔt/), consonantal target undershoot, etc. The third degree of syllable fusion was also used for monophthongization of diphthongs or for reduction in the number of vowels contiguous to the bisyllabic boundary. Examples of assimilation from the HKCAC are [jɐm^{55}mɐi^{22}] or [jɐm^{55}wɐi^{22}] for /jɐn^{55}wɐi^{22}/ 因爲 'because', [kɐm^{35}mœːn^{35}] for /kɐm^{35}jœːŋ^{35}/ 咁樣 'this way', and [tsɐu^{22}wɐi^{22}] for /tʰɐu^{22}hɐi^{22}/ 就係 'is'. Examples of monophthongization of diphthongs are [mu̲^{23}jɛː^{23}tsou^{22}] for /mou̲^{23}jɛː^{23}tsou^{22}/ 冇嘢做 'nothing to do', and [m̩^{21}kɪ^{33}tɐ^{5}] for /m̩^{21}kei^{33}tɐk^{5}/ 唔記得 'not remember'.
not fused	4	No segments are deleted and no phonological processes occurred to the segments of the two participating syllables that cross the syllable boundary.

Table 5.3 ended

5.3 Results

Results showed that there were no instances of fusion for degree zero (i.e. the most extreme fusion with merging of the vowels). Figures 5.1 through 5.3 show the trends for the other four degree of fusion categories. Multinomial logistic regression was employed to find out which potential factor influenced the degree of syllable fusion.

The principles of multinomial logistic regression are familiar to linguists who have used logistic regression analysis of sociolinguistic variables (i.e. VARBRUL, see Sankoff (1978) or Reitveld & van Hout (1993)). Essentially, a maximum likelihood model is built with no factors as the control case. The multinomial case differs from this in involving more than two possible outcomes. Models are difficult to fit analytically, and so hill-climbing techniques are used. A significant effect, then, is one where the model fitting converges on a single most likely solution and that most likely solution is a significantly better predictor of the category distributions than the base control model.

In the present study, factors were added one by one to the model to look for the best fit of the model, and a Chi-square test ($p<.05$) was performed to test the significance of each factor. The input factors were: word length (2, 3, 4 syllables), word frequency (high, mid, low), and morphosyntactic relationships (twins, sisters, aunts-nieces, cousins). Four different sets of analysis were done to test the four different possible outcomes, since no degree zero was found in the target words in the HKCAC as transcribed by the corpus transcribers. The outcomes were the four degrees of fusion (1, 2, 3 and 4). Results showed significant main effects ($p<.05$) for two input factors: word frequency (wfc) and morphosyntactic relationship (mr), but no main effect of word length (wl) was found. An interaction term effect was tested for word frequency and morphosyntactic relationship (wfc * mr), but not for either of these factors in combination with word length (wl), since word length did not have a main effect. No significant interaction effect was found. The tests for main and for interaction effects of the three potential factors on the influence of syllable fusion are given in Table 5.4 at the end of this chapter.

Figure 5.1 shows the distribution of the degrees of fusion by word frequency. The distribution was as predicted. That is, more extreme fusion values occurred relatively more often for words with high frequency than for words with mid frequency, and similarly for words with mid frequency relative to words with low frequency.

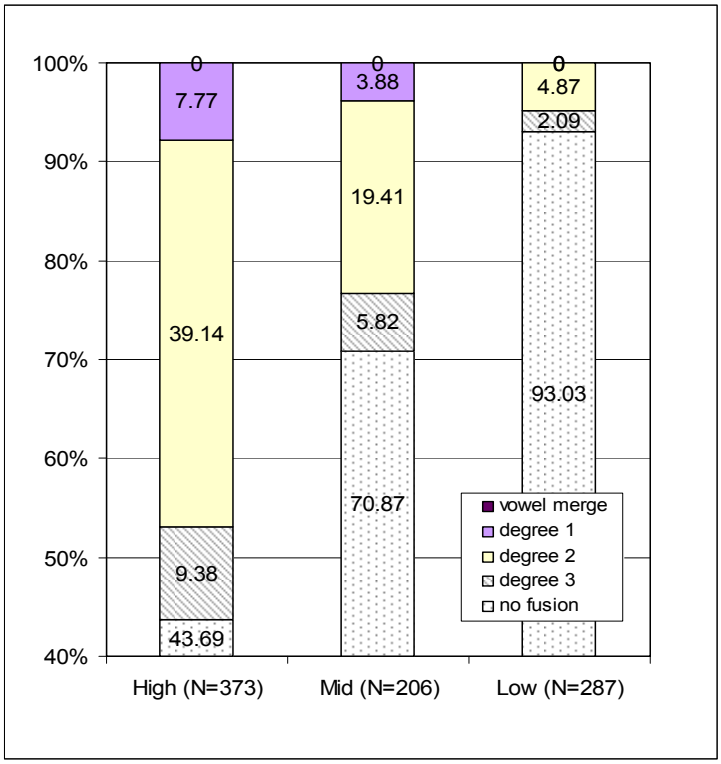


Figure 5.1 Word frequency and syllable fusion in the HKCAC. There was a significant main effect ($p \approx 0$).

Figure 5.2 shows the distribution of the degree of fusion by morphosyntactic relationships. Results showed that word-internal morphosyntactic boundaries discouraged more extreme forms of fusion, such that adjacent syllables in monomorphemic words or monomorphemic subparts of longer words were more likely to show more extreme forms of fusion than pairs of syllables in polymorphemic words separated by morpheme

boundaries. Among syllable sequences in polymorphemic words, sisters were more likely to be fused than pairs of syllables in an aunt-niece relationship, while cousins were least likely to be fused.

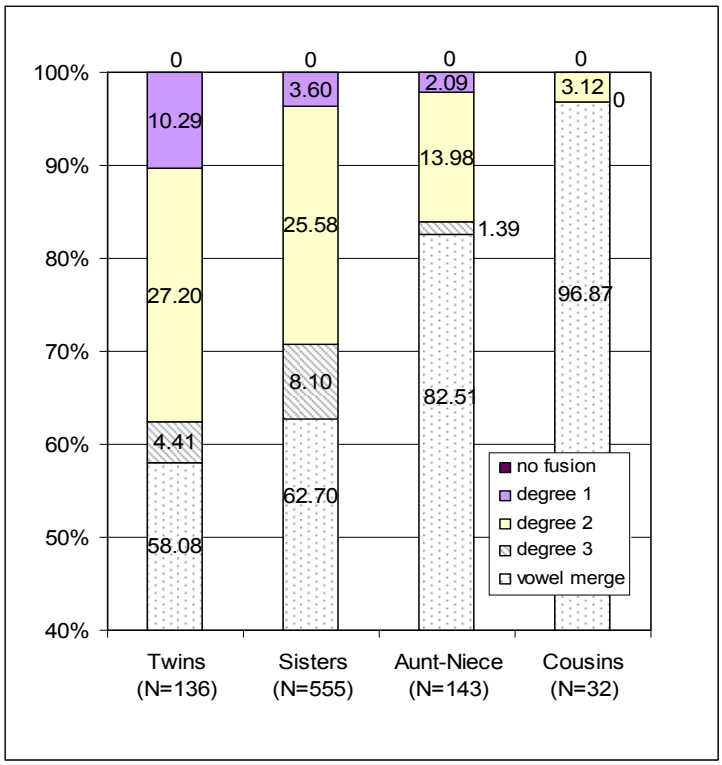


Figure 5.2 Morphosyntactic relationship and syllable fusion in the HKCAC. There was a significant main effect (p = 0.027).

A closer examination of the effect of morphosyntactic relationship is shown in Figure 5.3, which expands each of the bars in Figure 5.2 to show the distribution of syllable fusion in different word-internal branching relationship.

Since fusion is more likely in syllables that are grouped together, the distribution of syllable fusion in the graph suggests some rhythmic pressure in Cantonese that tends to reinforce the morphosyntactic grouping of the words. Rhythmic pressure in Cantonese will be discussed in more details in the next chapter. Basically, my intuition of the rhythm for the Cantonese words is that for 3-syllable words, the first two syllables are more likely to be grouped together than the last two syllables, whereas for 4-syllable words, the first two syllables tend to be grouped together, as well as the last two syllables. For 4-syllable words that are left-branching, however, the first three syllables tend to be grouped together, while the last syllable would make a group by itself.

Figure 5.3b shows syllable fusion of the 3-syllable words. The distribution shows that fusion was more likely for twins and sisters in the left branch of a word than for twins and sisters with single-level morphosyntactic structures. If the rhythmic pressure indeed has some influence on fusion, it might influence fusion to the extent of overriding the word-internal branching structures. We can see this effect in the figure, where aunts-nieces in the right-branching structure were fused more often and fused into more extreme forms than aunts-nieces in the left-branching structure. That right-branching sisters were fused more often and more fused than left-branching sisters appears to be due to the presence of more "fusion felicitous" segments (more in details in the next chapter). For 4-syllable words shown in Figure 5.3c, left-branching sisters were fused more often, while right-branching sisters were not found to be fused.

Because I did not have access to the audio files to note the intonational phrasing in the utterances of the HKAC, I could not directly examine the effect of prosodic position. However, a closer examination of the effect of morphosyntactic relationship suggests that prosodic position did affect degree of fusion. Figure 5.3c shows how position in the word interacted with the morphosyntactic relationship in the longer words. There seems to be an effect of "word-final lengthening" that comes out when the branching structure reinforces final lengthening over a more rhythmic grouping. That is,

where fusion is expected because of the grouping by the branching structures, fusion did not occur. For example, no fusion was found for the twins that are in the right branch, and no fusion was found for the right-branching sisters.

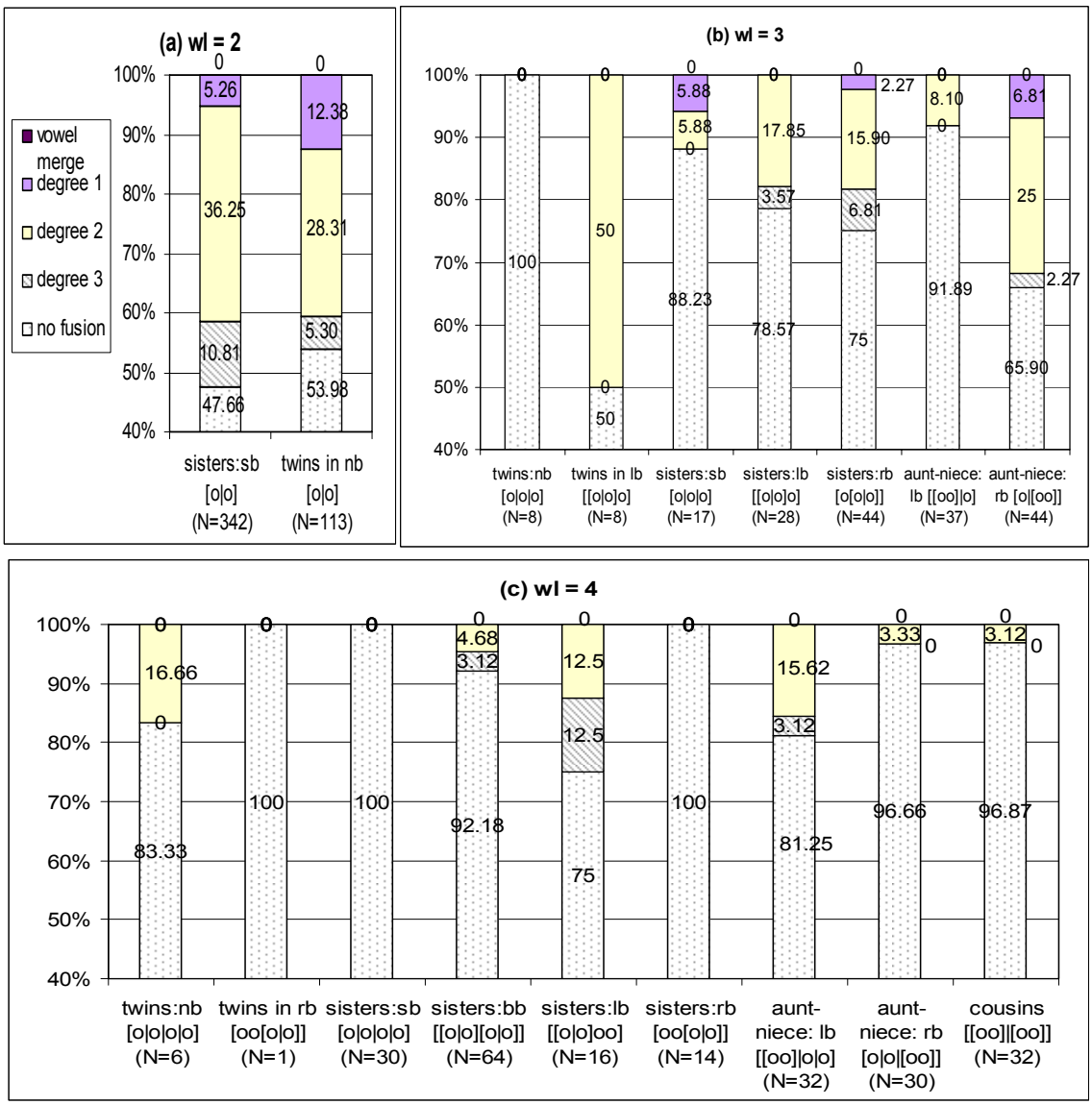


Figure 5.3 Degree of syllable fusion by word-internal branching structures in the HKCAC. (On the x-axis, an "o" represents a syllable; vertical line "|" indicates the target boundary, whereas square brackets the beginning and the end of the word.)

Figure 5.4 shows the distribution of the degree of fusion by word length. Word length was not found to have a significant effect. Nevertheless, we can still observe from Figure 5.4 a general pattern that there was relatively more extreme fusion at syllable boundaries in bisyllabic words than in trisyllabic words, and in trisyllabic words than in quadrisyllabic words. Given the positional effect suggested in Figure 5.3, this tendency for shorter words to show more fusion is the opposite of what we would expect. That is, whereas any boundary in a 2-syllable word involves the word-final syllable, only one of the two boundaries in a 3-syllable word is in that position, so we would expect less fusion overall in 2-syllable words relative to 3-syllable words. Similarly, we would expect less fusion overall in 3-syllable words relative to 4-syllable words, where only a third of the boundaries is subject to the word-final lengthening effect. However, this tendency for shorter words to show more fusion is almost surely an artifact of the word frequency effect shown in Figure 5.1 in combination with the impossibility of balancing frequency levels across word lengths. That is, the only words that occur with high frequency in the HKCAC are 2-syllable ones and all 4-syllable words are low frequency. At the same time, since the frequencies are defined by the count in the HKCAC, there is necessarily a small sample size for the low frequency 2-syllable condition. In the next chapter, the effects of these three factors, plus the potential factor prosodic position, were tested in a second more controlled experiment using the same target words in a repetition task so that cells could be more balanced.

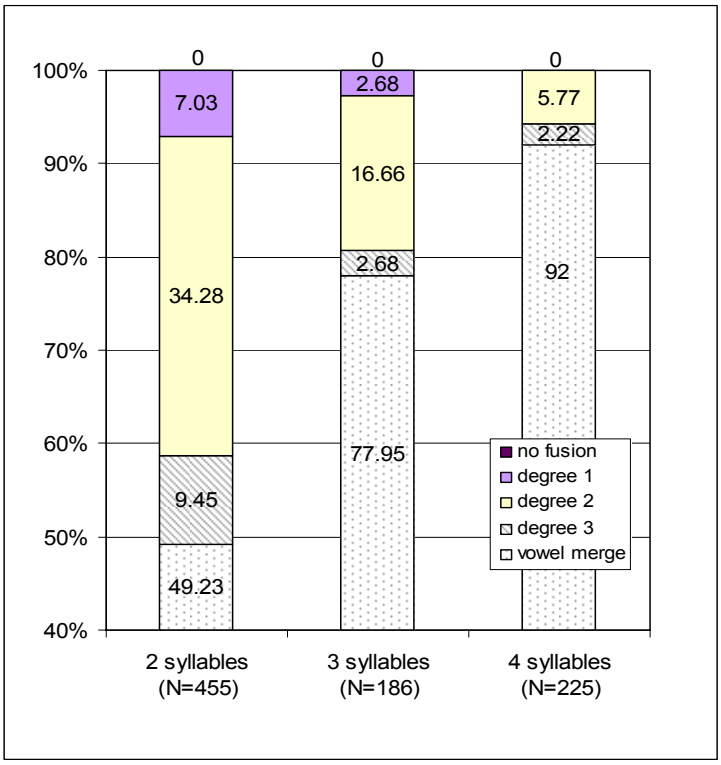


Figure 5.4 Word length and syllable fusion in the HKCAC. There was no significant main effect.

(i) Testing main effects

```
factors:
wl = word length
wfc = word frequency category
mr = morphosyntactic relationship
outcome:
tbd = degree of fusion (1, 2, 3, 4) at each word-medial syllable
boundary

frequency of each unique combination of factor and outcome:
freq = the number of times a unique combination of factors and
outcome emerged
```

A control case built:

```
> mult.0=multinom(tbd~rep(1,length(freq.data$mr)),
weight=freq.data$freq, data=freq.data)
# weights:  12 (6 variable)
initial  value 1200.530917
final  value 792.713411
converged
```

Set mult.1.1 to be the input factor wfc:

```
> mult.1.1=multinom(tbd~factor(wfc), weight=freq.data$freq,
data=freq.data)
# weights:  16 (9 variable)
initial  value 1200.530917
iter  10 value 715.988691
iter  20 value 689.486917
iter  20 value 689.486914
final  value 689.486914
converged
```

Table 5.4 Statistical results using multinomial logistic regression for testing the main effect and interaction effect of the three potential factors on syllable fusion. (Edited output of R.)

Table 5.4 continued

Significance of the input factor wfc tested by Chi-square test. Result was p ≈ 0.

```
> anova(mult.0, mult.1.1)
Likelihood ratio tests of Multinomial Models

Response: tbd
                          Model Resid. df Resid. Dev   Test    Df
LR stat. Pr(Chi)
1 rep(1, length(freq.data$mr))       429   1585.427
2                  factor(wfc)       423   1378.974 1 vs 2     6
206.453       0
```

Set mult.1.2 to be the input factor mr:

```
> mult.1.2=multinom(tbd~factor(wfc)+factor(mr),
weight=freq.data$freq, data=freq.data)
# weights:  28 (18 variable)
initial  value 1200.530917
iter  10 value 697.143740
iter  20 value 680.276640
iter  30 value 680.134759
final  value 680.133765
converged
```

Significance of the input factor mr tested by Chi-square test. Result was p = 0.027.

```
> anova(mult.1.1, mult.1.2)
Likelihood ratio tests of Multinomial Models

Response: tbd
                     Model Resid. df Resid. Dev   Test    Df LR
stat.    Pr(Chi)
1              factor(wfc)       423   1378.974
2 factor(wfc) + factor(mr)       414   1360.268 1 vs 2     9
18.70630 0.02780928
```

Table 5.4 continued

Set mult.1.3 to be the input factor wl.

```
> mult.1.3=multinom(tbd~factor(wfc)+factor(mr)+factor(wl),
weight=freq.data$freq, data=freq.data)
# weights:  36 (24 variable)
initial  value 1200.530917
iter  10 value 706.079095
iter  20 value 677.316179
iter  30 value 675.977583
iter  40 value 675.948294
final  value 675.947867
converged
```

Significance of the input factor wl tested by Chi-square test. Result was p = 0.212.

```
> anova(mult.1.2, mult.1.3)
Likelihood ratio tests of Multinomial Models

Response: tbd
                                  Model Resid. df Resid. Dev
Test    Df LR stat.
1              factor(wfc) + factor(mr)       414   1360.268
2 factor(wfc) + factor(mr) + factor(wl)       408   1351.896 1 vs
2     6 8.371796
    Pr(Chi)
1
2 0.2121100
```

(ii) Testing interaction effect for wfc * mr

Set mult.2.1 to be the interaction term wfc * mr:

```
> mult.2.1=multinom(tbd~factor(wfc)*factor(mr)+factor(wfc)+
factor(mr)+factor(wl), weight=freq.data$freq, data=freq.data)
# weights:  60 (42 variable)
initial  value 1200.530917
iter  10 value 702.443720
iter  20 value 673.219609
iter  30 value 671.546295
iter  40 value 671.488469
final  value 671.487964
converged
```

Table 5.4 continued

Significance of the interaction term wfc * mr tested by Chi-square test. Result was p = 0.835.

```
> anova(mult.1.2, mult.2.1)
Likelihood ratio tests of Multinomial Models

Response: tbd

Model Resid. df Resid. Dev
1                                          factor(wfc) +
factor(mr)       414   1360.268
2 factor(wfc) * factor(mr) + factor(wfc) + factor(mr) +
factor(wl)       390   1342.976
    Test    Df LR stat.   Pr(Chi)
1
2 1 vs 2    24 17.29160 0.8359615
```

Coefficients for the final modal of the data in Experiment II (i.e. main effects of wfc and mr)

```
> summary(mult.1.2)$coefficients
  (Intercept) factor(wfc)low factor(wfc)mid factor(mr)cousins factor(mr)sisters
2    1.445248      13.95522      0.1887652         2.991616        0.4441804
3   -1.627057      14.93997      0.8831903        -6.701214        2.1906118
4    1.283610      16.83005      1.4933455         3.712065        0.7512605
  factor(mr)twins
2      -0.5558462
3       0.4312379
4      -0.4583886
```

Table 5.4 ended

CHAPTER 6

EXPERIMENT III: ELICITING FUSIONS WITH A SENTENCE-REPETITION TASK

This experiment followed up on Experiment II. Since the frequencies of the target words were derived from the number of occurrences in the Hong Kong Cantonese Adult Language Corpus (HKCAC), it was not easy to know whether the frequency effect in Experiment II was due to the “inherent” frequency or to a repetition effect. By “inherent" frequency”, I mean the general probability that a word will occur in any discourse, as measured by token counts in very large, balanced corpora such as the Brown Corpus (Kucera & Francis 1967). Work such as Bybee (2000) shows that phonological processes such as assimilation and segment deletion tend to affect more frequent words first, and this is how we interpreted the very large effect of word frequency on syllable fusion in Chapter 5. At the same time, another body of research has found that at the second mentions of word types in discourses, vowels and word duration tend to be reduced, and speakers tend to produce the words with less articulatory clarity than the first mention of the word types (cf. e.g. Fowler & Housum (1987)). Because the token counts that I used to test the effect of word frequency on syllable fusion in the HKCAC were derived from the HKCAC itself, it is possible that the frequency effect shown in Figure 5.1 is due to this kind of repetition effect for second and later tokens of words in particular segments of the corpus rather than to the more general frequency of the different word types. Another related problem is that 3- and (especially) 4-syllable words are inherently less frequent and the only high frequency words were two syllables, which made it very difficult to test the effect of word length. That is, recall that no significant effect of word length was found in Experiment II, but that could have been simply because of the

necessary imbalance among the cells, such that all of the boundaries in the 4-syllable words were in low frequency words, whereas the boundaries in the 2-syllable words were a mix of types taken from a few low frequency words and many high frequency words.

Thus, one thing that Experiment III was designed to do was to test "inherent" frequency without any confound from a possible corpus-specific "repetition" frequency. This experiment took a more controlled look at the influence on syllable fusion of word frequency by eliciting a balanced number of tokens of words in all three frequency categories, while systematically varying word length, word-internal morphosyntactic relationship, as well as the new factor of prosodic position. The same target words that were tested in Experiment II were used in a sentence repetition task. Embedded in the sentences were the target words.

6.1 Method

6.1.1 Materials

This experiment used the same 173 target words that were tested in Experiment II (see Appendix E for the list of words). Each of the 173 words was tested in three prosodic positions: initial, medial, and final within a larger prosodic constituent such as the intonational phrase. In order to elicit the words in these positions, I designed a set of sentences that were short enough that each was likely to be pronounced as a single intonational phrase. In order to shorten the experimental time for each subject by minimizing the number of sentences to be repeated, I combined target words into the same sentence. That is, each sentence was designed to contain three target words, with the target words coming from different frequency categories insofar as that was possible given the imbalanced distribution.

The three target words in the sentences were separated optionally by intervening words that helped make the sentence meaningful. Meaningful sentences were used because they are what everyday natural speech involves, and this would help elicit natural connected speech. The number of intervening words was not controlled except to insure that none of the sentences was too long for the subjects to remember and to repeat. Monosyllabic sentence-final particles were placed at the end of a sentence when needed, since it would be unnatural if sentence final particles were altogether omitted in spoken Cantonese. Each sentence had between 15 and 20 syllables. Three example sentences are given in Table 6.1. The full set of test sentences is given in Appendix G. There were two sentences (numbers 173 and 174) that had two target words only, since I could not think of a meaningful way to include a third word. There were a total of 174 test sentences.

Before testing subjects, all the test sentences were proof-read by a linguistically trained speech therapist. Modifications of the sentences were made where necessary.

1.	I<L2>嗰個</L2> 傻婆 M<L3><L2>突然</L2>間</L3> 喺個 頭度 F<L4><L2>拗下</L2><L2>拗下</L2></L4> I<L2>kɔː^{35}kɔː^{33}</L2> sɔː^{21}p^{h}ɔː^{35} M<L3><L2> tɐt^{2}jːn^{21}</L2>kaːn^{55}</L3> hɐi^{35}kɔː^{33} t^{h}ɐu^{21}tou^{22} F<L4><L2>ŋaːu^{55}haː^{23}</L2><L2>ŋaːu^{55}haː^{23}</L2></L4> That funny lady suddenly scratched her head.
2.	I<L2>我哋</L2> 呢啲初哥 M<L2>投資</L2> 起上嚟真係 F<L2>論盡</L2> 呀 I<L2>ŋɔː^{23}tei^{22}</L2> lei^{55}tiː^{55}ts^{h}ɔː^{55}kɔː^{55} M <L2>t^{h}ɐu^{21}tsiː^{55}</L2> hei^{55}sœːŋ^{35}lei^{21}tsɐn^{55}hɐi^{22} F<L2>lɵn^{22}tsɵn^{22}</L2>aː^{33} We the first-time investors are really clumsy in doing investment.
3.	I<L2>呢個</L2> nil M<L2>市民</L2> 鬧房屋署做嘢 F<L4><L2>拖下 </L2><L2>拖下</L2></L4>咁 I<L2> lei^{55}kɔː^{33}</L2> nil M<L2>siː^{23}mɐn^{21}</L2> laːu^{22}fɔːŋ^{21}ok^{5}ts^{h}yː^{23}tsou^{22}jɛː^{23} F<L4><L2>t^{h}ɔː^{55}haː^{23}</L2><L2>t^{h}ɔː^{55}haː^{23}</L2></L4>kɐm^{35} This citizen scolded the Housing Authority as it procrastinated in doing things.

Table 6.1 A few examples of the test sentences, in which target words were embedded. (Underlined are target words. I/M/F indicate the position, initial, medial or final, of the target words within the sentence.)

6.1.1.1 Recording and editing the test sentences

An aural stimulus corresponding to each test sentence was made. I recorded myself, a native Hong Kong Cantonese female speaker, producing the test sentences in the sound-proof booth in the Speech Laboratory in the English Department at the Hong Kong Polytechnic University. My production was recorded directly on compact discs (CDs) using Marantz CDR300 CD recorder through an AKG C5900 (hypercardioid; frequency range: 20 - 22,000Hz) microphone. The microphone was clipped on a microphone stand at a fixed position throughout the entire process of recording. I read the test sentences syllable by syllable. That is, in reading each sentence, I paused after producing each syllable to make sure there was no fusion with the following syllable. Recording of all the test sentences was done in one sitting. Note that this syllable-by-syllable pronunciation is not as unnatural as it would be for English, since in many sentences each syllable is a morpheme and children learn to read by learning the character for each morpheme separately. Thus, a syllable-by-syllable reading is equivalent to the word-by-word reading that an English-acquiring child might hear when first learning to read words in sentences.

Some of the characters have more than one pronunciation. For alternations that are lexically conditioned, I adopted the following principles for my recording. If a target word has variant pronunciations, but occurred just once in the HKCAC, I said the word as it was produced in the HKCAC. For example, 麻雀 'sparrow' has variant pronunciations [maː^{21}tsœːk^{3}] and [maː^{21}tsœːk^{35}]. I used the former form in recording the stimuli because that was how the speaker said it in the HKCAC. Other examples are 到 'able to' as in 學習得到 'can learn'. It was recorded as [tou^{35}], despite its variant pronunciations of [tou^{35}] or [tou^{33}]. 房 'room' as in 匿咗入房 'hid in the (bed)room' was recorded as [fɔːŋ^{35}], despite its literary style of pronunciation could be [fɔːŋ^{21}]. If a target word occurred in the HKCAC more than once with varying pronunciations, I chose the pronunciation that would be more natural for me. For example, 呢 'this' as in 呢一家

‘this one CL.’, 呢一個 ‘this one CL.’, and 呢個 ‘this CL.’ were all produced as [lei^{55}], despite that it can be produced as [nei^{55}], [niː^{55}], [lei^{55}], [liː^{55}], [jiː^{55}] or [tiː^{55}]. 而家 ‘now’ was produced as [jiː^{21}kaː^{55}] in my recording, despite that it has variant pronunciations, namely, [jiː^{21}kaː^{55}] or [jiː^{55}kaː^{55}].

For alternations that are phonologically conditioned, the principle that I followed was to choose the pronunciation that would be more natural for me, and hence, for the recording. Thus, initial /n/’s were replaced by [l]’s in my recording. For example, 諗 ‘think’ as in 去諗辦法 ‘to think out a way’ and in 諗 咩 ‘thinking about what’ were recorded as [lɐm^{35}], instead of [nɐm^{35}]. (To some, the latter would sound more literary/formal in terms of style.) Another example was 電腦 ‘computer’, which I produced as [tiːn^{22}lou^{23}] for the stimuli. In addition, in my recording I used the dental or alveo-palatal affricates, which are allophonic alternations before rounded vowels. (Some Hong Kong Cantonese speakers would have the allophonic alternation for /s/ in the same vowel context.) In my recording, distinctions were also made between /k$^{w(h)}$/ and /k$^{(h)}$/ before the vowel /ɔː/, and between final dentals and velars.

After recording the stimuli, I edited them to remove the beginning silence and the pauses between syllables. Each syllable was cut at the beginning of a glottal pulse and at the final glottal pulse. Syllables beginning with a stop onset were cut right before the stop burst. For syllables ending with a stop coda, 11 milliseconds of silence was left after the final glottal pulse for subjects to perceive the checked syllable. All cuts were placed at zero-crossing.

The effect of cutting out the pauses between syllables was that although the test sentences would be comprehensible to the subjects, they would be rather unnatural. I explained the unnaturalness of the sentences by telling the subjects that the sentences were produced by a primitive speech synthesis system. The synthesis system took recordings of the individual syllables and simply put them together to make the sentences.

6.2 Procedures

Before the experiment started, each subject read the debriefing material (see Appendix H) about the task, signed a participation consent form and an audio release consent form (see Appendices I and J), and asked any questions that they had. They were instructed to turn off their mobile phones throughout the experiment. After they finished the experiment, they filled out a questionnaire that inquired about their language background (see Appendix K for the questionnaire).

In the experiment, all the aural stimuli were played from a Compaq 1715AP notebook computer. The subjects listened to the stimuli through the Sennheiser HMD25-1 headphones. Their productions were recorded using the AKG C5900 microphone in the sound-proof booth in the Speech Laboratory in the English Department at the Hong Kong Polytechnic University. The microphone was clipped on a microphone stand at a fixed position throughout the entire process of recording. The subjects' productions were recorded directly onto compact discs (CDs) using a Marantz CDR300 CD recorder.

Each subject listened to the test sentences and repeated the sentences once. The subjects were told that the utterances were produced by a poorly designed speech synthesizer, and that I needed their productions of the same utterances as a better model for the synthesizer. They were encouraged to repeat the utterances as fast as possible because the interval for them to repeat was pretty short. They were also encouraged to repeat the utterances in a way that they would produce them in everyday natural speech without adding or deleting words. They were also told that some utterances would show

inversion or topicalization, which was not uncommon in everyday spoken Cantonese. Also, some of the syllables might sound fast and others sound slow, and that was part of the reason why the production of the utterances sounded unnatural. Furthermore, the utterances were all meaningful, although they might sound a bit strange by being out of context.

There were four randomized lists of test sentences. Sentences in each list made three blocks. Each block had 58 sentences. The subjects were encouraged to take breaks between blocks.

The stimuli were played by running a Praat script. The script played each test sentence twice, with a 0.5 second pause between the first and the second playing. After the sentence was played the second time, there was a 2.5 second pause during which the subject was to repeat the sentence. The blank interval was kept short so that the subject could not speak too slowly, since slow tempo would discourage the occurrence of fusion forms.

Unlike in Experiment I, the subjects in this experiment did not have to do simple arithmetic while repeating the sentences. That is, I did not try to encourage fusion forms by dividing attention using a second concurrent task, since the stimuli in this experiment were two or three times longer than those in Experiment I and the task of remembering the sentence was difficult enough.

6.3 Subjects

There were 13 male and 14 female Hong Kong Cantonese speakers between 18-23 years of age who participated in the experiment. They were mainly recruited through friends, although flyers were also sent out (see Appendix L for the recruitment flyer). They were all undergraduate students majoring in areas other than linguistics, Chinese

language, or translation. These majors were excluded because I found in a pilot study for my other research that these students tend to produce clearly-articulated forms, rather than the fusion forms I was trying to elicit.

The participants were all born and raised in Hong Kong. Their parents had also been living in Hong Kong for eighteen or more years. None of the participants was monolingual, and it would have been impossible to find monolingual speakers of the targeted dialect. However, all subjects stated that Cantonese was their most proficient spoken language. All participants reported no record of speech or hearing problems. The subjects were paid on completion of the experiment, and were provided with soft drinks for their participation in the experiment.

6.4 Transcribing elicited productions

The productions of 5 male and 5 female speakers were selected for transcription[1]. The criterion for choosing these 10 speakers was that they were the 5 men and 5 women who repeated the most number of sentences, and I wanted there to be a balance between the genders. Even so, speakers sometimes failed to repeat sentences because they could not remember the full sentence. Also, a number of repetitions had to be discarded because of the noise, such as the ringing of the experimenter's mobile phone that needed to be left on to receive messages from other potential participants. Elicited productions were transcribed by the experimenter using Praat. The transcription was done in the textgrids using Praat. Each sentence was linked to a textgrid, in which there were seven tiers. An example showing the test sentence number 110 as repeated by a subject and transcribed by the experimenter is shown in Figure 6.1. Information about the target words of each sentence was imported from a text file into the textgrids. The intervals in this tier were

[1] My original intent was to hire two labelers to do the labeling and to recheck, but I did not get the funds that I had applied for to do that. Thus, I did the transcription all by myself, and I could only transcribe 10 subjects, rather than all of the 27 subjects who participated in the experiment.

not made to be time-linked to the waveform because the native transcriber knew whether and where the target words were produced in the speech signal. The locations of the intervals were predetermined by the Praat script.

In the textgrids, there were two tiers that the transcriber did not have to make changes to. First, the "targets" tier (an interval tier), in which the 'underlying' segmental forms of the target words (details see section 6.1.1.1) were given using WorldBet symbols. The vertical mark (" | ") here indicates the 'underlying' syllable boundary. Stop codas in Cantonese are accompanied by a glottal stop. Although transcribing the accompanying glottal stop is not the common practice for Cantonese, the WorldBet symbol '?' for glottal stop was added in the transcription before the stop coda [p], [t], or [k] for the target words. Second, the "tones" tier (a point tier), in which the tones of the target words using Chao's five tone numbers were given. The transcriber did not have to make changes to this tier because this dissertation restricts the investigation of syllable fusion to the segmental level.

The transcriber's task was to make changes as appropriate in the other four tiers, starting with the third tier — i.e. the trans.PW tier. The trans.PW tier was a copy of the targets tier by default. That is, it contained a copy of the 'underlying' segmental forms of the target words (details see section 6.1.1.1). The transcriber modified the segmental transcription in the trans.PW tier where appropriate according to the productions of the speakers.

The first tier was a point tier, which transcribed the degree of fusion of the word-internal boundaries of the target words and the intonational phrasing. Each vertical mark corresponded to an index of the degree of syllable fusion in the first tier. Five degrees of syllable fusion from 0 to 4 were chosen and transcribed, using the same definitions in Experiment II. The definitions for the five degrees of syllable fusion are reproduced here

as Table 6.2. Each point was set as ‘4’ (i.e. no fusion) as the default value, and the transcriber’s task was to change the degree as appropriate based on the transcription she gave in the trans.PW tier.

	Degree of fusion	Definition
highly fused	0	***Merging of the vowels*** for the two participating syllables at a syllable boundary to become a single vowel, whether or not the participating vowels are the same, and whether or not the merged vowel is of an intermediate quality (i.e. vowel coalescence).
	1	For ***each*** of the two participating syllables at a syllable boundary, there is deletion of ***at least one*** segment that is immediately contiguous to the syllable boundary. However, the whole syllable is not deleted for either of the two syllables. (Nil onset syllables that are not produced with an initial glottal stop are *not* counted as segmental deletion.)
	2	For ***either*** of the two participating syllables at a syllable boundary, there is deletion of ***at least one*** segment immediately contiguous to the syllable boundary. However, the whole syllable is not deleted for either of the two syllables.
	3	There are ***phonological processes*** occurring for the segment(s) contiguous to the boundary between the two participating syllables. However, *neither* of the two participating syllables had segments deleted at the syllable boundary. Phonological processes include place/manner/voicing assimilation (which includes partial voicing of voiceless phonemes, where partial voicing is defined as at least half of the segment is voiced, but not the whole segment is voiced), liason, vowel reduction, consonant lenition, incomplete deletion of consonant (e.g. deletion of place for coda stops but retaining the glottal stop that accompany the coda stops: [ʔ] for /ʔt/), consonantal target undershoot, etc. The third degree of syllable fusion was also used for monophthongization of diphthongs or for reduction in the number of vowels contiguous to the bisyllabic boundary.
not fused	4	No segments are deleted and no phonological processes occurred to the segments of the two participating syllables that cross the syllable boundary.

Table 6.2 Definitions for the five degrees of syllable fusion from 0 (highly fused) to 4 (not fused).

The transcriber followed the guidelines and labeling conventions below to transcribe the target words segmentally in the third tier and to label the degree of fusion in the first tier.

(i) *Alternations*. The transcriber took into account the lexically conditioned alternations and the phonologically conditioned synchronic alternations that are present in Hong Kong Cantonese (cf. sections 3.1.4.2 and 6.1.1.1). When there was uncertainty between assimilation / segmental deletion and alternations of the sorts above, the transcriber would consistently treat those cases as alternations. An example is 電腦 'computer', which some subjects produced as [tiːn^{22}nou^{23}] (even though I produced it as [tiːn^{22}lou^{23}] in all of the stimuli) (see Figure 6.3). The initial [n-] for the second morpheme could be the result of the subject assimilating the nasal feature of the preceding segment across the syllable boundary, or it could be the 'underlying' /n/ that the subject had for the morpheme to be conservative. Cases like this were given the fusion degree '4', meaning no fusion occurred for the segments that straddled across the bisyllabic boundary.

(ii) *Deletion*. Deletion was defined as deletion of the time slot for a segment that was contiguous to the syllable boundary in the underlying form. That is, even if certain features of the segment remained (e.g. labiality or nasality), if the realization of the features did not define a separate time slot so that the segment was completely coarticulated with a neighboring segment, then the token was transcribed as deletion of the segment. The fusion degree would be '1' or '2' depending on whether or not deletion occurred in both syllables. That is, '1' meant that in a pair of syllables both the final segment and the initial segment were deleted. Figure 6.4 shows an example where one segment (/k/) was deleted for the first target word tsi:|kei (for /tsiː^{22}kei^{35}/ 'self'). Figure 6.5 shows deletion of a segment, with the nasal feature of the segment taken up by the following consonant across the syllable boundary. Deletion could occur partially for a segment, such as the final stops in Cantonese, which are accompanied by a glottal stop. Speakers could delete the place of articulation of the stops while retaining the

accompanying glottal stop. These cases of incomplete deletion of segments were labeled with the fusion degree ‘3’. See examples in Figures 6.1 and 6.3.

(iii) *Assimilation*. A segment was not deleted, but was changed to become a form different from the ‘underlying’ form. The change could be attributed to being influenced by a neighboring segment. An example of assimilation is shown in Figure 6.4. The third target word 辛苦 ‘torturing’ /sɐn^{55}fuː^{35}/ was produced as [sɐn^{55}vuː^{35}] in the sentence. The voiceless fricative at the bisyllabic boundary was voiced but not deleted, presumably taking on the voice feature from the preceding code.

(iii) *Ambiguity between deletion and assimilation*. When there was ambiguity between deletion and assimilation, the case was treated as deletion. Examples can be found in Figures 6.1 and 6.3. There was ambiguity between deletion and assimilation when a coda and its following consonant were homorganic. ‘Underlyingly’, the coda for the first syllable and the onset for the second syllable for the word 出出入入 /ts^{h}ɵt^{5}ts^{h}ɵt^{5}jɐp^{2}jɐp^{2}/ ‘going in and out’ are homorganic. It was hard to draw a line of which portion of the closure belonged to which of the segments. For cases like this, the place of articulation of the coda of the first syllable was treated to be deleted. The degree of fusion was ‘3’, if the accompanying glottal stop was not deleted. Another example is between the second and the third syllables for the word 返屋企 /faːn^{55}ok^{5}k^{h}ei^{35}/ ‘go back home’ in Figure 6.3., where the place of articulation for the final velar was consistently treated as deleted, rather than assimilated to the following velar onset.

(iv) *Liason*. Cases of liason were transcribed without the syllable boundary mark “ | ”. Such cases typically occurred for syllables with null onset following a nasal. For example, in Figure 6.3 返屋企 /faːn^{55}ok^{5}k^{h}ei^{35}/ was produced as [faːn^{55}noʔ^{5}k^{h}ei^{35}], with the final [n] in [faːn^{55}] being ambisyllabic. The transcriber consistently treated cases as such as liason, and transcribed without the syllable boundary marker, e.g. [faːnoʔ]. (One could assume the ‘underlying’ form for [oʔ^{5}] was /ŋok^{5}/ for the speaker, and treated it as a case of place assimilation. Since synchronic alternations took precedence over assimilation, the ‘underlying’ form this case was assumed to be /ok^{5}/.)

(v) *Precedence*. (i) through (iv) above can be summarized in precedence terms of analysis. In cases of more than one possible analysis, treating a case as an instance of lexically conditioned alternation or phonologically conditioned synchronic alternation took precedence over assimilation or deletion (cf. (i)). In case where deletion or assimilation were plausible, deletion took precedence over assimilation (cf. (iii)). In cases where assimilation or liason were plausible, liason took precedence over assimilation (cf. (iv). A lower fusion value (for more fused) took precedence over a higher fusion value (for less fused) in cases where two different degrees of fusion applied to the fusion phenomena across the bisyllabic boundary. For example, Figure 6.6 shows that there was a reduction in the number of vowels contiguous to the syllable boundary for the second syllable, which would be given the fusion degree '3'. However, the deletion of the onset for the third syllable would render the case to be given the fusion degree '2'. A lower fusion value (for more fused) took precedence over a higher fusion value (for less fused) in cases like this rendered the degree of fusion for the case to be '2'.

(vi) *Labeling conventions for segments*. See Appendix F.

(vii) *Missing tokens*. When the subject stumbled or made some other error such as not repeating a target exactly, the whole word was labeled as a missing token by substituting the 'NA' ('not available') label for all default labels in the degree of fusion, transcription, and the prosodic position tiers. Figure 6.2 shows an example in which the subject did not repeat the target word 打過去問 /ta:^{35}k^{w}ɔ:^{33}hɵy^{33}mɐn^{22}/ 'has called to ask' but substituted 打電話去問/ta:^{35}ti:n^{22}wa:^{35}hɵy^{33}mɐn^{22}/ 'called by phone to ask'. A word token was labeled as missing if the subject deleted the target word in repeating the sentence, or if the subject hesitated in the middle of producing the target word, or if the subject substituted some other morpheme for one or more of the morphemes of the target word, or the subject made some production errors such as a phoneme substitution that is not a legal allomorph. A set of criteria was used to make sure that the tokens transcribed and would be analyzed were the same tokens that the subjects repeated as they had heard them. The criteria were the same as those used in Experiment I (see section 3.1.4.2).

The first tier also transcribed the beginning and the end of an intonational phrase indicated by the labels '<SIL>' and '%' respectively. The default was set such that each sentence made one intonational phrase. The transcriber changed the default by adding or removing points in the tier to indicate the existence of more intonational phrases for the sentence where appropriate. The addition or removal of points was done with respect to the location of the fusion degree labels, which were located within the target word intervals. The example in Figure 6.1 shows that the speaker made two intonational phrases as analyzed by the transcriber PW. The placement of the pair of '<SIL>' and '%' labels indicated that the first target word belonged to the first intonational phrase, while the other two words belonged to the second. Definitions of an intonational phrase followed the preliminary analysis of Wong, Chan & Beckman 2005, although the transcriber did not have to mark the tonal event (H, L, HL, tonal prolongation and truncation). Intonational phrase-end would be indicated by any of the following phenomena. There would be a fall from the final lexical tone; rise from the final lexical tone; a rise from the final lexical tone, with a short plateau at the very end of the rise, sometimes accompanied by incredulity reading; a final rise and then fall from the final lexical tone; no extra tone added after the lexical tone of the final syllable; or a truncated rise of the final lexical tone. The transcriber added the label 's' between the intonational phrase-end marker '%' and the preceding fusion value for cases where the speaker added sentence-final particle(s) between the last syllable of the target word and the phrase-end. Examples can be found in Figures 6.1 and 6.4.

The fourth tier (a point tier) transcribed the prosodic position ('I' for initial, 'M' for medial, and 'F' for final) of the target words in their respective sentences. There were cases where the target words occupied the phrase-initial and -final prosodic positions. The label 'I&F' was used to mark those cases. The points were made midpoints of each of the target word intervals. A target word would be labeled as initial when it occurred at the very beginning of an intonational phrase. A target word would be labeled as final when it occurred at the very end of an intonational phrase, or just right before the sentence-final particle. A target word would be labeled as medial when it did not occur at

the very beginning of an intonational phrase, and did not occur at the very end of an intonational phrase, or before the sentence final particle. The transcriber changed the prosodic position labels for the target words as appropriate.

The posWord.PW tier was added to the textgrids after all the transcriptions were done. Different from the position.PW tier, where the labels “I” (initial), “M” (medial), “F” (final) and “I&F” (initial&final) indicated the prosodic position of the word with respect to the intonational phrase, the same labels in the posWord.PW tier indictated the prosodic position of the boundary within the word with respect to the intonational phrase. The posWord.PW tier was obtained automatically by running a script that took into account the labels in the position.PW tier and the “s” label (for the existence of sentence-final particle(s)) in the degFusion.PW tier. Thus, for example, in Figure 6.1 the last target word 袋落咗去 /tɔːi^{22}lɔːk^{2}tsɔː^{35}hɵy^{33}/ ‘put into’ was labeled as “F” in the position.PW tier; in the posWord.PW tier, the first boundary of this target word was then labeled as “M”. The last boundary of this word was labled as “M” because boundary was not the last boundary in the utterance, but it was the second last one, as indicated by the “s” label (for sentence-final particle) in the degFusion.PW tier. The analysis of prosodic positions in this experiment took into account labels in this posWord.PW tier, which pertained to the prosodic position of boundaries within words. Lastly, the notes tier at the bottom was for the transcriber to make notes.

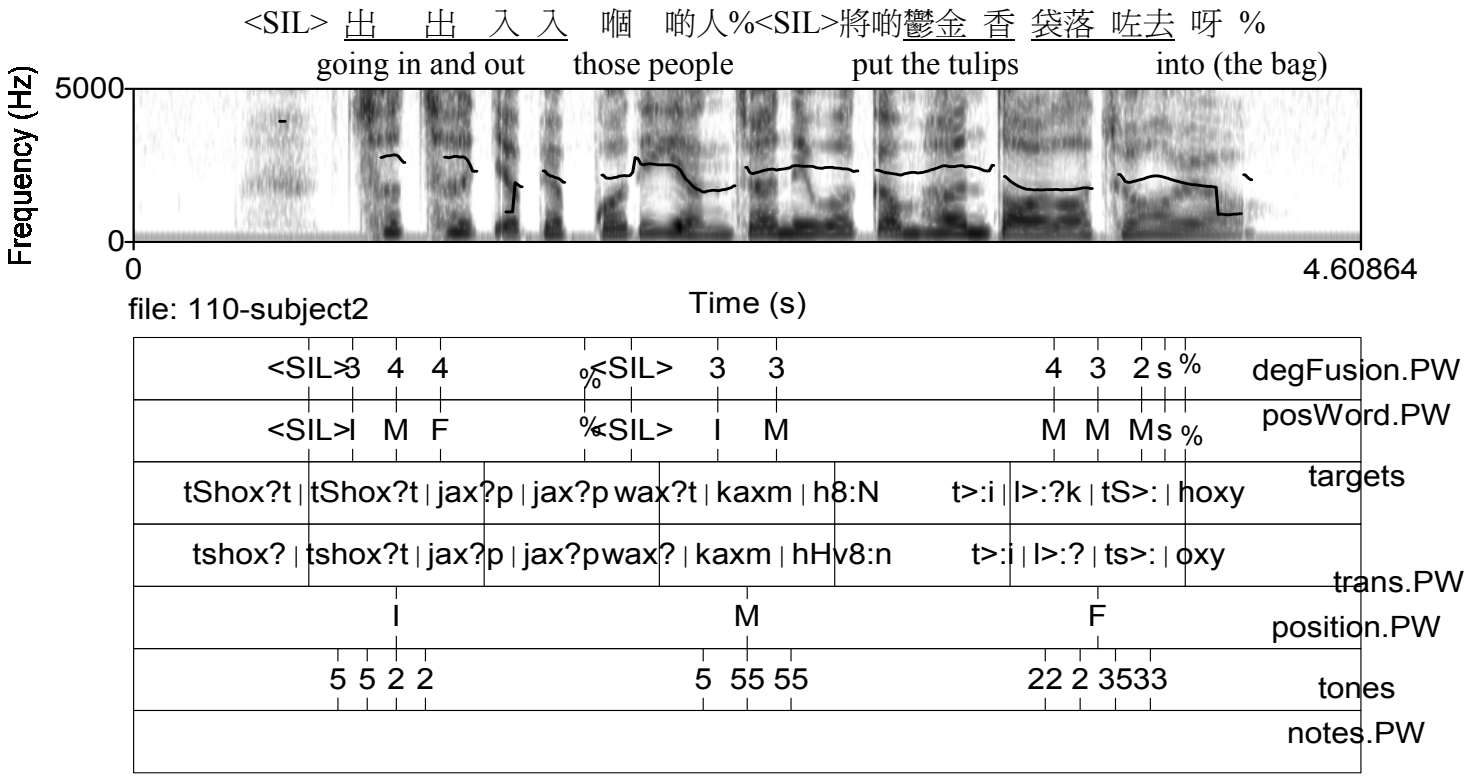


Figure 6.1 The example shown is the test sentence no.110 as repeated by subject no. 2. The target words were 出出入入 /ts^{h}ɵt^{5}ts^{h}ɵt^{5}jɐp^{2}jɐp^{2}/ ‘going in and out’, 鬱金香 /wɐt^{5}kɐm^{55}hœːŋ^{55}/ ‘tulip’, and 袋落咗去 /tɔːi^{22}lɔːk^{2}tsɔː^{35}hɵy^{33}/ ‘put into’. They are underlined in the caption above the spectrogram.

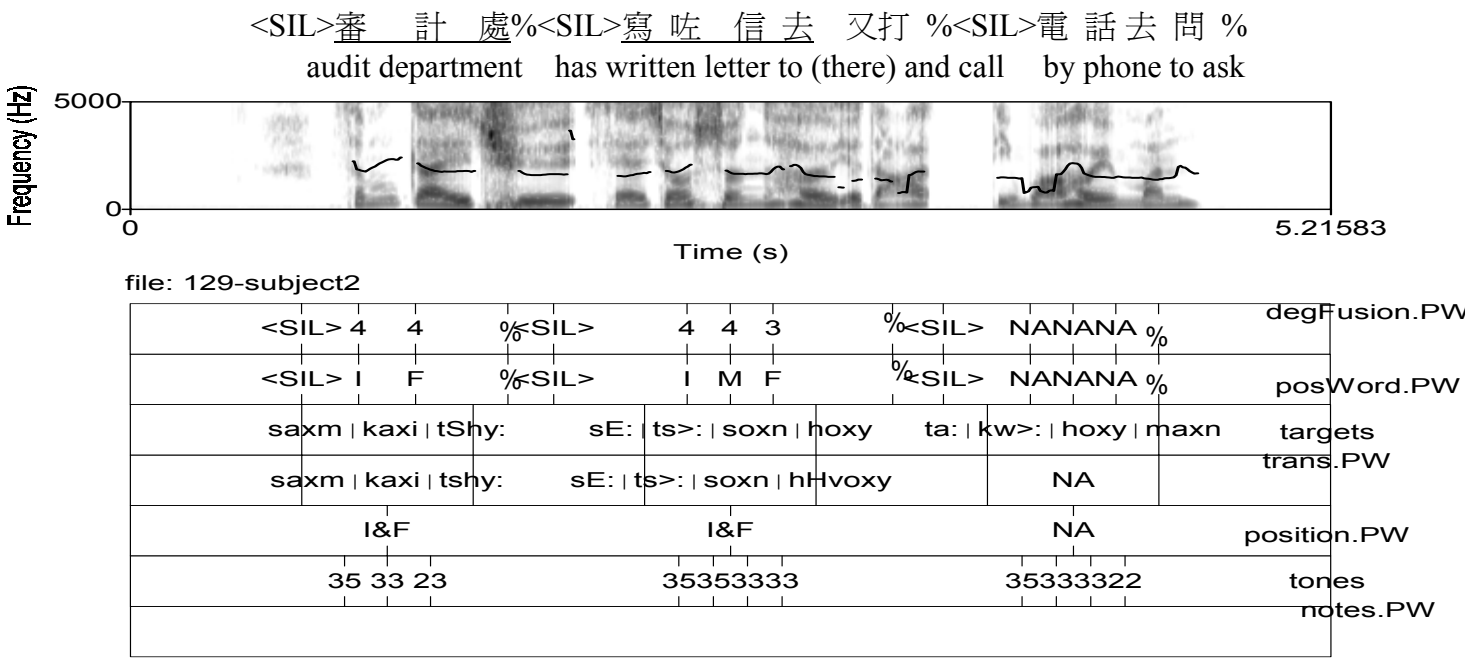


Figure 6.2 The example shown is the test sentence no.129 as repeated by a subject. The target words were 審計處 /sɐm^{35}kɐi^{33}ts^{h}yː^{23}/ 'audit department', 寫咗信去 /sɛː^{35}tsɔː^{55}sɵn^{33}hɵy^{33}/ 'has written letter to', and 打過去問 /taː^{35}k^{w}ɔː^{33}hɵy^{33}mɐn^{22}/ 'has called to ask'. Each of the first two target words made an intonational phrase and simultaneously occupying the initial and final position (and thus the label 'I&F' on the position.PW tier) of the phrase, while for the last target word, the subject did not produce the target form and instead substituted another word 打電話去問 /taː^{35}tiːn^{22}waː^{35}hɵy^{33}mɐn^{22}/ 'called by phone to ask' that was not in the stimulus (and thus the label 'NA'). The target words produced are underlined in the caption above the spectrogram.

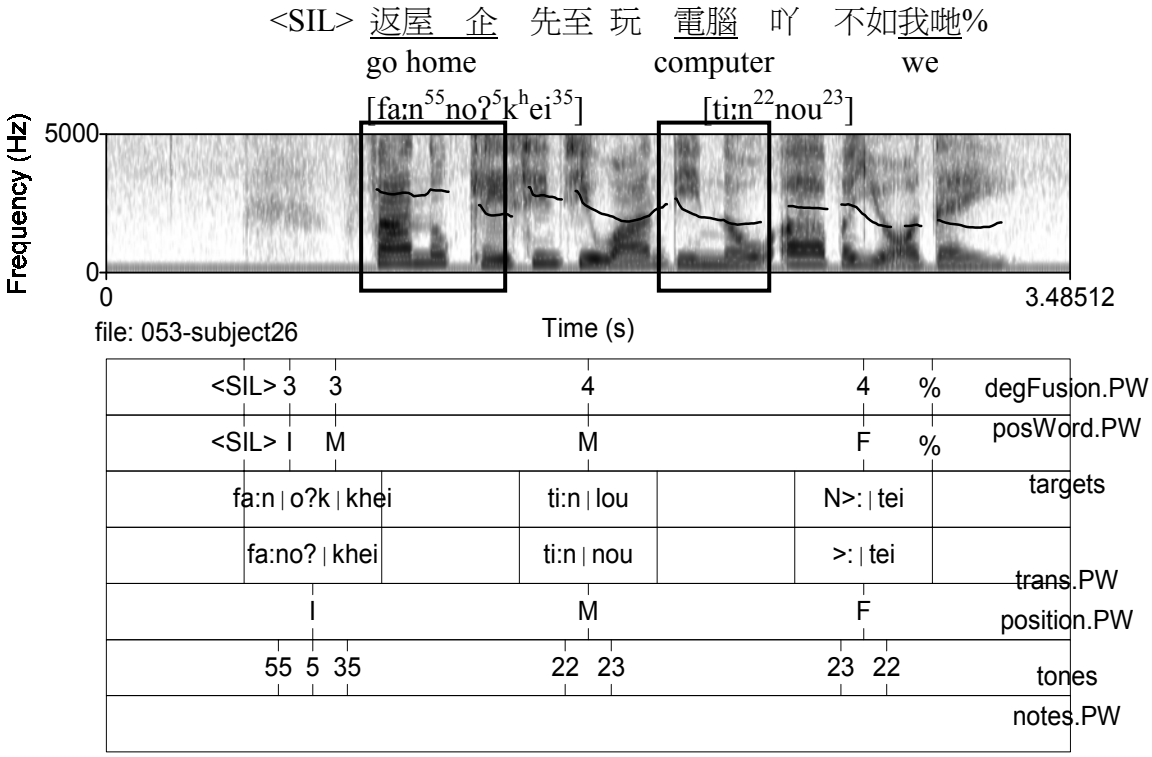


Figure 6.3 The production of [n-] for the second syllable of the second target word 電腦 'computer' [ti:n^{22}nou^{23}] could be ascribed to assimilating the nasal feature from the preceding coda or the synchronic alternations between [n-] and [l-] for /n-/. The transcriber consistently treated such case as alternation, and thus, not a case of syllable fusion. The degree of fusion was '4'. The first target word shows that liason between the first two syllables was transcribed without the vertical boundary marker. The degree of fusion was '3'. For the second syllable, the velar place of articulation was deleted for the coda, although the glottal stop was retained. The degree of fusion for cases as such was '3'.

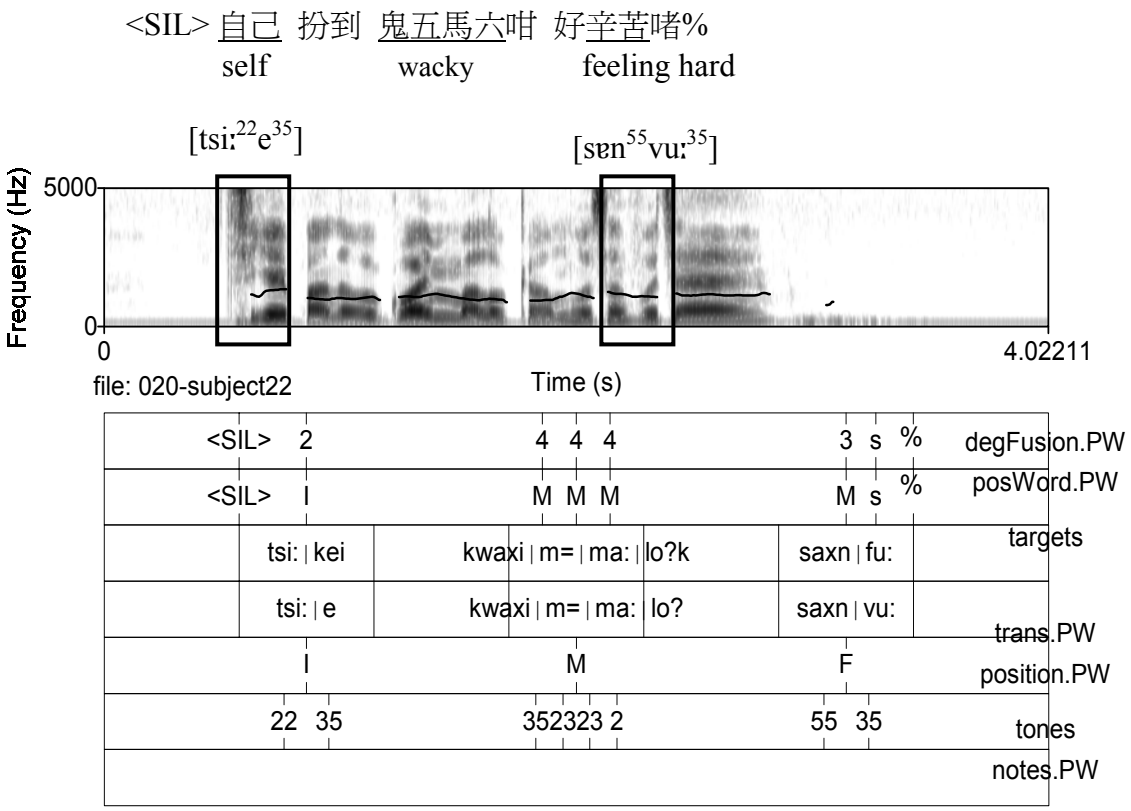


Figure 6.4 An example of assimilation as shown by the third target word 辛苦 'feeling hard' [sɐn^{55}vuː^{35}] /sɐn^{55}fuː^{35}/ in the sentence. The voiceless fricative at the bisyllabic boundary was voiced but not deleted. Deletion occurred for the first target word 自己 /tsiː^{22}kei^{35}/ 'self', in which the onset for the second syllable was deleted. (The fact that [i] for /kei^{35}/ was deleted was not relevant for coding syllable fusion in this dissertation, but it would be relevant for its relationship with the following syllable, if the following syllable were a target.)

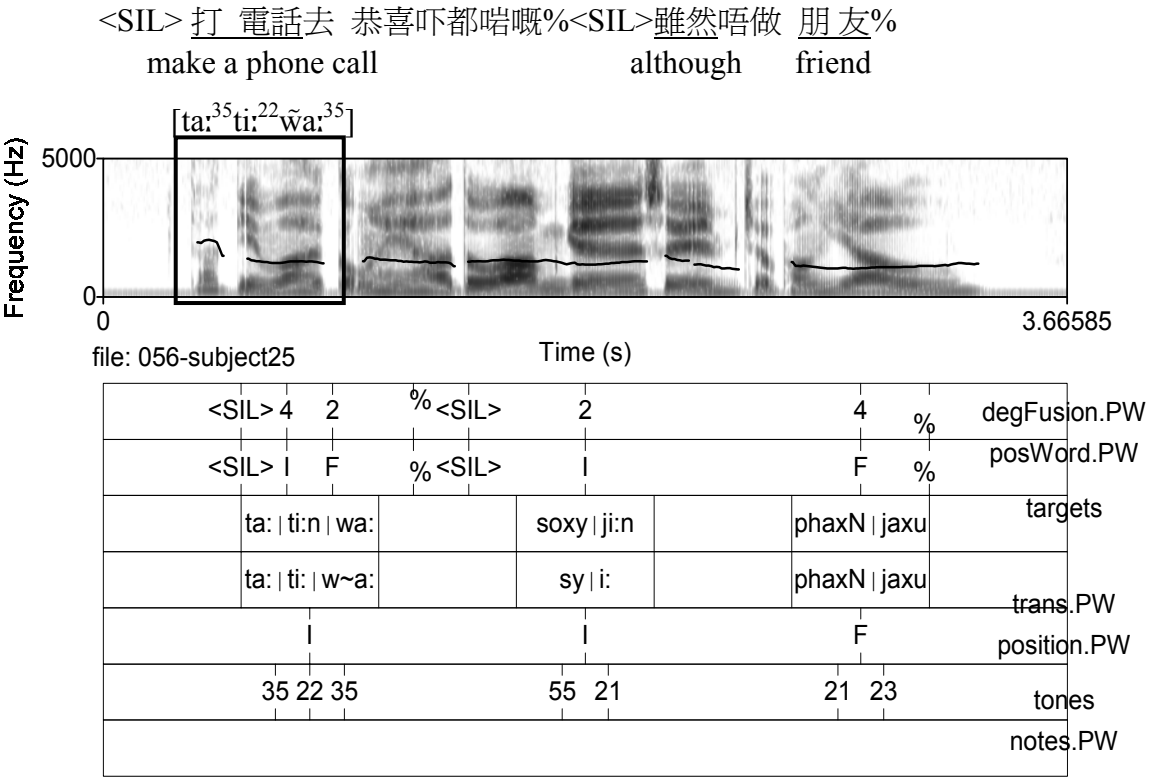


Figure 6.5 An example of segmental deletion and assimilation. In the first target word /ta:^{35}ti:n^{22}wa:^{35}/ 打電話 'to make a phone call' in the sentence, the final nasal for the second syllable was deleted. The nasal feature was coarticulated with the following bilabial approximant. This was counted as a case of deletion. The degree of fusion was '2'.

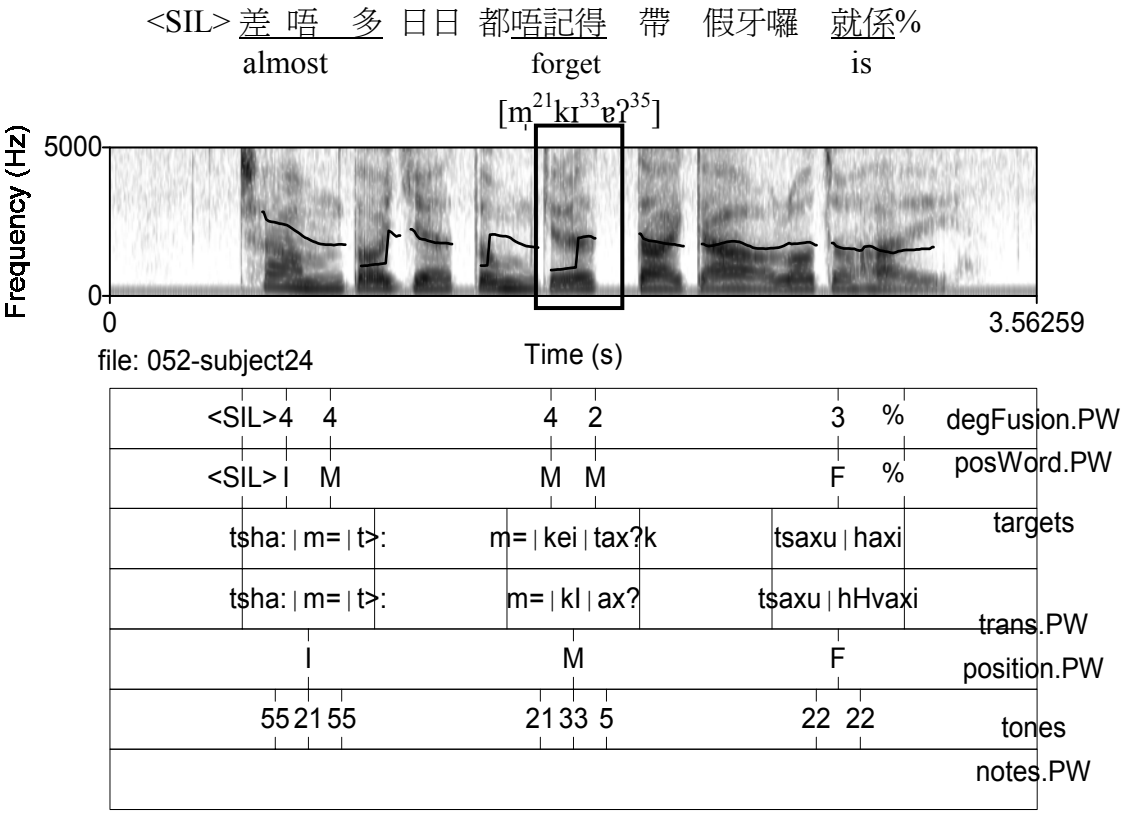


Figure 6.6 For the second target word 唔記得 /m̩21kei33tɐk5/ 'forget', the reduction in the number of vowels contiguous to the syllable boundary for the second syllable would be given the fusion degree '3', whereas the deletion of the onset for the third syllable would render the case to be given the fusion degree '2'. A lower degree of fusion took precedence over a higher degree of fusion in cases like this rendered the degree of fusion for the case to be '2'.

6.5 Analysis

A total of 1733 usable sentence productions out of the theoretically possible total of 1740 sentences were elicited and analyzed. Labels from the Praat textgrids were extracted into a text file, and those for word tokens that were labeled as 'NA' (i.e. missing tokens) were deleted. Recall that the analysis of prosodic positions in this experiment took the labels in the posWord.PW tier, which pertained to the prosodic position of boundaries within words, rather than the labels in the position.PW tier, which described the position of the target word as a whole. Altogether, there were 4695 word tokens containing 9628 word-internal syllable boundaries.

6.6 Results and discussion

As in Experiment II, stepwise multinomial logistic regression was used to see which potential factors and which combinations of potential factors significantly influenced syllable fusion. The tests for main and interaction effects and the coefficients for the final model are given at the end of this chapter in Table 6.7.

Results showed significant main effects of all four input factors and significant two-term interactions were tested for all combinations of the four factors. Interaction effects were found for word length with each of the three other factors, and for word frequency by morphosyntactic relationship. Because there were these significant two-way interactions, the three-term interaction among morphosyntactic relationship, word length, and word frequency was also tested, but it was not found to be significant. Figure 6.7 shows the amount of syllable fusion and the distribution of the degree of fusion for each of the four factors, thus giving an overview of each main effect. Figures 6.8 through 6.12 plot the significant interaction effects. In each of these figures, the bars are divided into regions for each degree of fusion, from no fusion (degree 4) at the bottom to extreme fusion (degree 0) at the top.

As Figure 6.7a shows, high frequency was highly conducive to the occurrence of syllable fusion. More than half (66.9%) of the boundaries in high frequency words had at least some degree of fusion. By contrast, the majority (54%) of boundaries within mid and low frequency words showed no fusion.

Figure 6.7b shows the main effect of morphosyntactic relationship. Just as in Experiment II, pairs of syllables that were in a cousins relationship were least likely to be fused (62% not fused). However, results of other comparisons were counter to the results of Experiment II. Specifically, pairs of syllables that were in an aunt-niece relationship were more likely to have some degree of fusion than pairs which were sisters, and sisters were more likely to have some degree of fusion than twins. As Figures 6.8 through 6.12 will suggest, this difference between the two experiments may be due to the interaction effects that were uncovered when more tokens of mid and low frequency forms could be examined.

Figure 6.7c shows the main effect of word length. This significant effect is in the opposite direction of the unexpected tendency in Figure 5.3 in Experiment II, supporting our earlier interpretation of that tendency as an artifact of the frequency effect in combination with the inability to balance the cells for different frequencies across different word lengths. That is, in Experiment III, where we can tabulate instances of syllable fusion in many more mid frequency and low frequency 2-syllable tokens than we could examine in Experiment II, we find fusion to be least likely in the shortest words and most likely in the longest words. Since we find this effect in spite of having only low frequency tokens for 4-syllable words, this result lends further support to our explanation for the unexpected tendency in Figure 5.4. The result will be explored further in interaction plots, particularly in Figure 6.9.

Figure 6.7d shows the distribution of the degrees of syllable fusion as influenced by prosodic position — the factor that could not be tested in Experiment II. As expected, "Initial&Final" boundaries (i.e. syllable sequences that were complete intonational phrases by themselves) were least likely to show fusion. Other comparisons, however, were not as predicted. Again, explanations will be sought by looking at interaction plots.

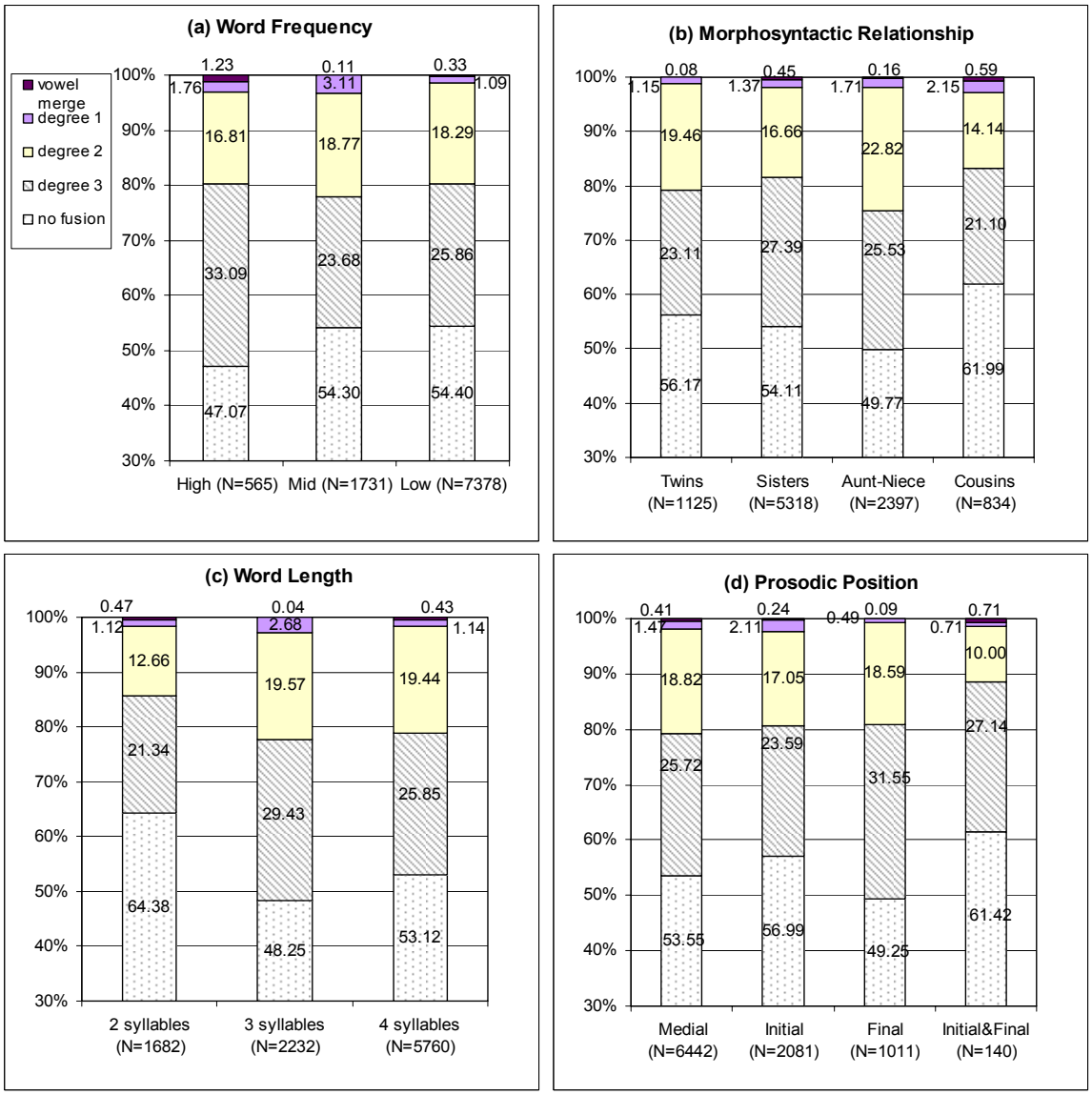


Figure 6.7 Main effects in Experiment III. Each plot shows the proportion of tokens that had each degree of fusion, from no fusion (at the bottom) to most fused (at the top).

Figures 6.8, 6.9, and 6.13 through 6.15 give several interaction plots for the factor word length, which interacted with all three other factors. From the interaction plots, the behavior of word-internal boundaries in different positions in the word can be seen.

Figure 6.8 plots interactions between morphosyntactic relationship, word frequency, and word length. Interactions can be seen between different types of morphosyntactic relationship and word frequency, and between different types of morphosyntactic relationship with word length.

Interactions between word frequency and the different types of morphosyntactic relationship can be seen in twins and sisters in the various panels of Figure 6.8. As predicted, twins generally were more likely to have some degree of fusion and were fused to a greater extent more often than sisters, but this was not true for all cases. In the high and the mid frequency categories for the 2-syllable words (see Figure 6.8 (a-b)), twins were less likely to have some degree of fusion and were fused to a lesser extent than sisters. High frequency might lead the speakers to treat sisters (comparatively more decomposable forms) as twins (comparatively less decomposable forms), and thus giving rise to more fused forms for sisters.

Another interaction effect between word frequency and the different types of morphosyntactic relationship can be seen in sisters and aunts-nieces for the 3-syllable words (see Figure 6.8 (d-e)). That is, while, as expected, aunts-nieces were less likely to have some degree of fusion than sisters in the mid frequency category aunts-nieces were more likely to have some degree of fusion than sisters in the low frequency category. A possible explanation for this interaction will be suggested when we examine the word-internal branching structures of the words in Figure 6.9, and when we examine the segmental sequences that were contiguous to the boundary between two syllables.

In Figure 6.8, interactions between word length and the different types of morphosyntactic relationship can also be seen. A comparison between twins and sisters for the 2- and the 3-syllable words (see Figure 6.8 (b) and (d)) regarding the amount of fusion shows that for 2-syllable words, sisters were (just a bit) more likely to have some degree of fusion than twins, and sisters had more extreme fused forms than twins. This was in reverse to the pattern as seen in the 3-syllable words (Figure 6.8d), where twins were fused more often and fused into more extreme forms than sisters. As suggested above, higher frequency might lead the speakers to treat sisters (comparatively more decomposable forms) as twins (comparatively less decomposable forms), and thus giving rise to more fused forms for sisters.

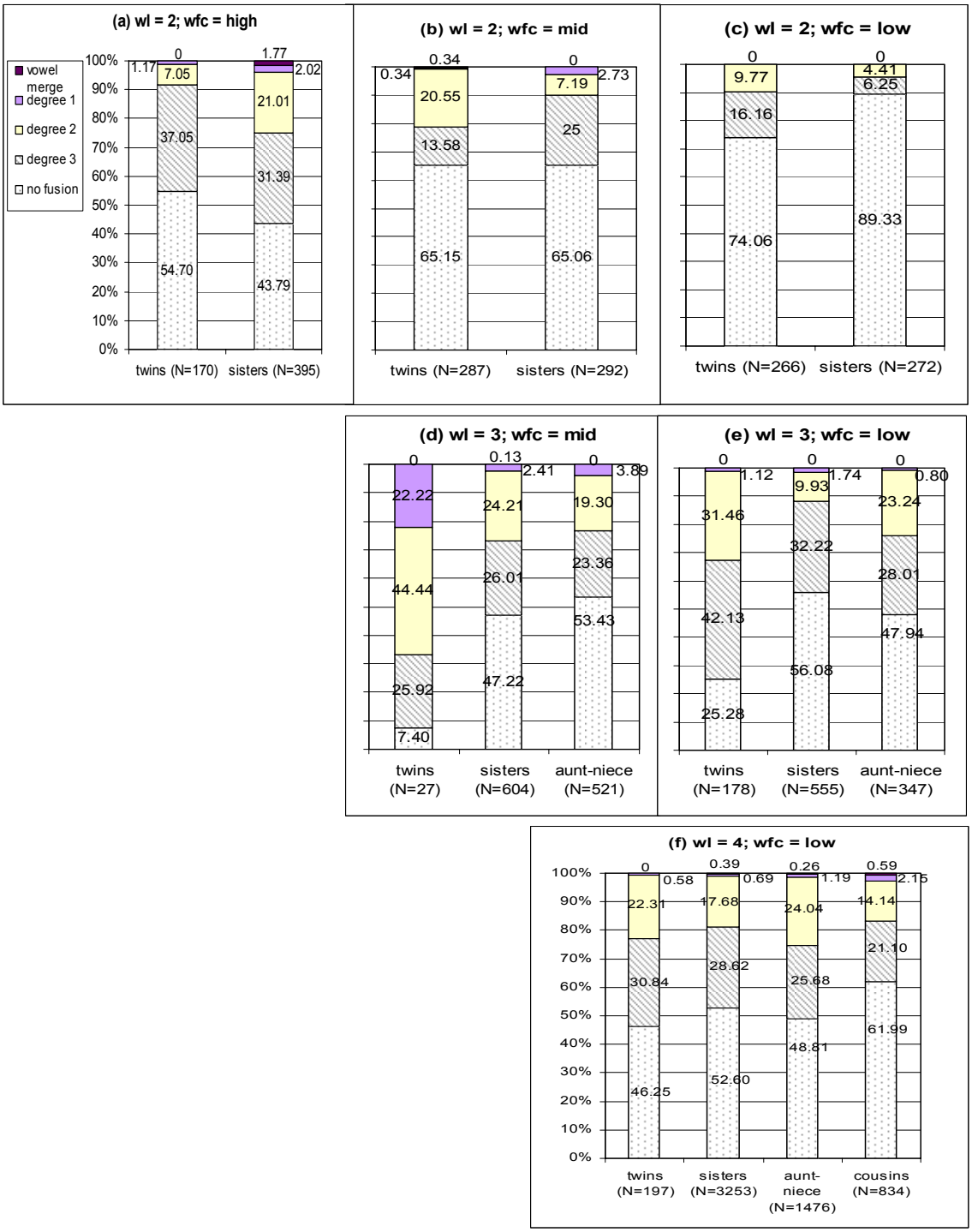


Figure 6.8 Plots to show interactions between morphosyntactic relationship and word frequency, and between morphosyntactic relationship and word length.

Figure 6.9 gives expanded plots for Figure 6.8(d-f) — i.e. the 3- and the 4-syllable words — to take a closer look at the internal branching structures of the four different morphosyntactic relationships, with a view to finding explanations for the unexpected patterns observed in Figure 6.8.

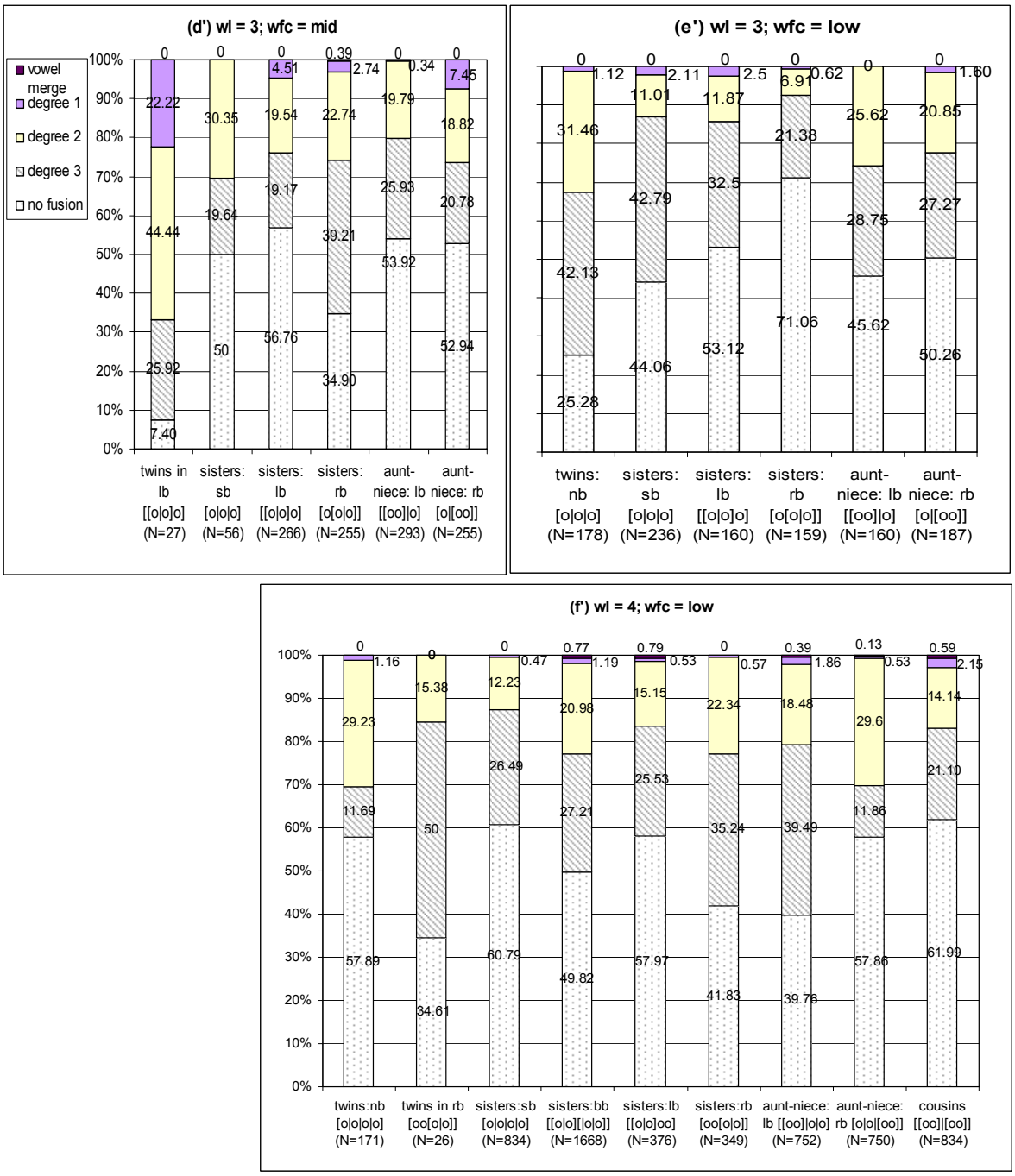


Figure 6.9 Expanded plots of Figure 6.8(d, e) and (f) to show the internal branching structures of the four different morphosyntactic relationships. (On the x-axis, "Twins in lb" means twins at the lowest level node in the left branch; an "o" represents a syllable; and a vertical line "|" indicates the target boundary, whereas square brackets the beginning and the end of the word.)

Figure 6.9(d'), which corresponds to Figure 6.8d, shows that aunts-nieces that were in the left-branching node were fused less often and fused to less extreme forms than aunts-nieces that were in the right-branching node, suggesting that "word-final lengthening" is at work. That is, the lengthening of the last syllable discouraged fusion between the last two syllables.

The interaction of word-final lengthening with the branching structure might also be understood as a rhythmic effect, if the added length expresses the second beat of a "defectively non-branching foot". Rhythm for 4-syllable words is transparently obvious, where rhythmic balance would dictate two equal-sized bisyllabic "foot" groups — i.e. [σ σ][σ σ] (cf. Figure 6.12). Figure 6.10 gives a representation to analogize the rhythm for Cantonese 3-syllable words with musical rhythm.

musical rhythm [𝅗𝅥 𝅗𝅥][𝅗𝅥 𝅝]

syllables [σ σ][σ]

Figure 6.10 Rhythm for Cantonese 3-syllable words in analogy with musical rhythm.

My intuition is that the first two syllables in a 3-syllable word in (Hong Kong) Cantonese tend to be grouped together, with each syllable having a half note, while the third syllable makes a full note (i.e. it is lengthened). The intuition mainly comes from when the words themselves are completely an intonational phrase by themselves. To represent the intuition in terms of musical rhythm, Figure 6.10 shows that the first two syllables would make a full note, while the lengthened third syllable alone makes a full note. In other words, in a 3-syllable word, the first two syllables would be more likely to be fused than the second and the third syllables as facilitated by the rhythmic pressure. This rhythmic pressure could override word-internal branching structure.

As far as the branching structure for the aunt-niece relationship is concerned, we would then predict that aunts-nieces that are in the left-branching structure would be fused less often than aunts-nieces that are in the right-branching structure. (Figure 6.11 gives a schematic representation of the two branching structures mapping onto the rhythm for Cantonese 3-syllable words.) That aunts-nieces in the left-branching were fused more often and had more extreme fusion more often than aunts-nieces in the right-branching structures in Figure 6.9(d') was thus predicted. As for sisters, the mapping between the rhythmic and the branching structures leads us to predict that sisters in the left-branching structure would be fused more often and fused more extremely than sisters in the right-branching structure. The predictions were borne out by the sisters in the left-branching and the right-branching structures in Figure 6.9(e').

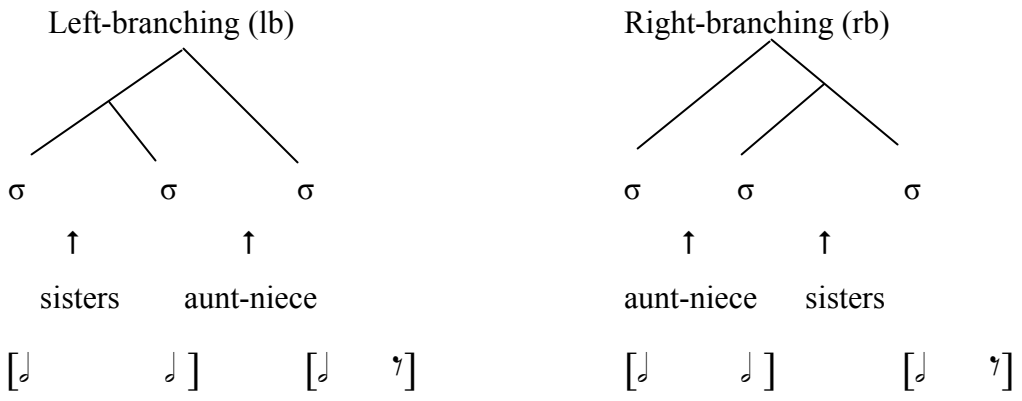


Figure 6.11 Sisters and aunt-niece relationships in two branching structures mapping onto the rhythm for Cantonese 3-syllable words.

These predictions, however, were not borne out by the aunts-nieces in the left- and right-branching structures for the low frequency words in Figures 6.9(e') and (f'), nor did the predictions borne out by the sisters in the left- and the right-branching structures for the mid and low frequency words in Figure 6.9(d') and (f'). The segments that were

contiguous to the syllable boundaries for the sisters and the twins could have given rise to these unexpected results, and hence, the interactions. We will first identify the segments that are likely to induce fusion. Then, we will examine the target words in Experiment III.

Given how degrees of fusion are defined in this dissertation, to incur some degree of fusion, one of the segments that is contiguous to the boundary between two syllables needs to be prone to be assimilated or deleted. This can cause segmental artifacts in fine-grained interaction plots such as the graphs in Figures 6.8 and 6.9. In order to determine whether some of the interaction effects are artifacts of the materials used, therefore, we review the segmental effects that were identified in Lee (2003) and Wong (1996), the two previous quantitative studies of syllable fusion.

Glides and fricatives were the segments that were most likely to be deleted in these two studies (see e.g. Wong 1996; Lee 2003:47). A closer examination of Lee's data and my earlier data in 1996 shows that not all the glides and fricatives encourage fusion equally. The segments that were most likely to be assimilated or deleted, and thus, creating some degree of fusion were [h] and [j]. Beside these two sounds, some degree of fusion was likely to incur for syllables ending with the coda stops [t] and [k] before the syllable boundary, because speakers tend to delete the place of articulation even in citation form, although the glottal stop that accompanies the final place of articulation tends to retain (see Wong 2002).

Another way to incur some degree of fusion is that one of the segments that is contiguous to the boundary between two syllables needs to be likely to spread its feature(s) to the segment across the syllable boundary, *and* the feature-spreading segment itself needs to be resistant to be assimilated or deleted. Lee (2003:47) found that bilabials [p], [p^h] and [m], as well as syllabic [m̩] were least likely to be deleted in her corpus. (For each of these two groups of sounds, not more than 1.5% were deleted.) In transcribing the subjects' productions in Experiment III, I observed that [p, p^h, m] (when occurred as

onsets) and [m] (when occurred as codas) were likely to spread features to the consonant across the syllable boundary such that the consonant contiguous to the syllable boundary tends to be assimilated as a labial or the consonant contiguous to the syllable boundary is deleted. (Although onsets [k(wh)] and [f] were also found to be least likely to be deleted in Lee's corpus, I rarely observed spreading of features to the consonant across the syllable boundary such that some degree of fusion was incurred.)

Thus, the presence of any of the codas [m, t, k], onsets [j, h, m, p, p^h], or the syllabic nasal [m̩] at the boundary between two syllables was likely to induce some degree of fusion. We now can take a look at the target words in Experiment III to see if the presence of the above segments at the syllable boundary helps in explaining the distribution of fused forms between the left- and the right-branching structures for the sisters and the aunts-nieces in Figure 6.9.

Table 6.3 lists the mid frequency 3-syllable target words, containing disyllabic sequences that were targets for the relationships sisters and aunt-niece. Among the 9 target words in the left-branching structure, there were 44% of the sisters that had one or more of segments that induced fusion across the syllable boundary. For the sisters in the right-branching structure, however, there were more (55%) of those fusion-inducing segments. In light of this segmental distribution, the outcome that left-branching sisters were fused less often and be less fused than the right-branching sisters in Figure 6.9 (d') makes sense.

The distribution of fusion forms between the mid frequency left-branching sisters and the right-branching aunts-nieces also makes sense if viewed in the light of the segments that were contiguous to the boundaries across the two syllables. The rhythmic pressure that we identified as a possible explanation for many of the positional effects in Figure 6.9 predicts that there would be more fused forms for the left-branching sisters than for the right-branching aunts-nieces. An examination of the segments that were

contiguous to the syllable boundaries as listed in Table 6.3 reveals that the unexpected outcome could be due to the fact that the left-branching sisters had fewer fusion-inducing segments (44%) than the right-branching aunts-nieces (55% of the targets had fusion-inducing segments).

words	'underlying' form	gloss	presence of fusion-inducing segments: codas [m, t, k], onsets [j, h, m, p, p^h], or [m̩] across syllable boundary of…	
mid frequency 3 syllables				
left branching			**sisters**	**aunt-niece**
<L3><L2>影響</L2>到</L3>	jeŋ35hœːŋ35tou35	affect	yes	
<L3><L2>預算</L2>案</L3>	jyː22syːn33ŋɔːn33	budget		
<L3><L2>基本</L2>法</L3>	kei55puːn35faːt3	Basic Law	yes	
<L3><L2>唔該</L2>晒</L3>	m̩21kɔːi55saːi33	thank you	yes	
<L3><L2>冇嘢</L2>做</L3>	mou23jɛː23tsou22	nothing to do		
<L3><L2>細路</L2>仔</L3>	sɐi33lou22tsɐi35	children		
<L3><L2>私營</L2>化</L3>	siː55jeŋ21faː33	privatization	yes	
<L3><L2>財政</L2>司</L3>	tshɔːi21tseŋ33siː55	financial secretary		
<L3><L2>互聯</L2>網</L3>	wuː22lyːn21mɔːŋ23	Internet		yes
		% yes:	**44%**	**11%**

Table 6.3 Presence of target segmental sequences across syllable boundary that might have encouraged fusion for mid frequency 3-syllable words.

Table 6.3 continued

words	'underlying' form	gloss	presence of fusion-inducing segments: codas [m, t, k], onsets [j, h, m, p, p^h], or [m̩] across syllable boundary of…	
mid frequency 3 syllables				
right branching			**aunt-niece**	**sisters**
<L3>一<L2>部份</L2></L3>	jɐt^{5}pou^{22}fɐn^{22}	a portion	yes	
<L3>返<L2>屋企</L2></L3>	faːn^{55}ok^{5}k^{h}ei^{35}	go back home		yes
<L3>呢<L2>一個</L2></L3>	nei/ni/lei/li/ji^{55}jɐt^{5}kɔː^{33}	this one	yes	yes
<L3>唔<L2>記得</L2></L3>	m̩^{21}kei^{33}tɐk^{5}	forget	yes	
<L3>嗰<L2>方面</L2></L3>	kɔː^{35}fɔːŋ^{55}miːn^{22}	that aspect		yes
<L3>大<L2>姑奶</L2></L3>	taːi^{22}kuː^{55}naːi^{55}	sister-in-law		
<L3>打<L2>電話</L2></L3>	taː^{35}tiːn^{22}waː^{35}	to call		
<L3>差<L2>唔多</L2></L3>	tshaː^{55}m̩^{21}tɔː^{55}	almost	yes	yes
<L3>搵<L2>唔到</L2></L3>	wɐn^{35}m̩^{21}tou^{35}	can't find	yes	yes
		% yes:	**55%**	**55%**

Table 6.3 ended

The unexpected outcome for the aunts-nieces for the low frequency 3-syllable words in Figure 6.8(e') can also find explanation from the presence (or absence) of fusion-inducing segments. Table 6.4 lists the low frequency 3-syllable target words for the relationships sisters and aunt-niece. Among the 7 target words in the left-branching structure, 57% of the aunt-niece sequences had one or more of segments that induced fusion across the syllable boundary. For the aunts-nieces in the right-branching structure,

however, there were fewer (50%) of the fusion-inducing segments. Thus, the outcome that left-branching aunts-nieces were fused more often and were fused to a greater extent than the right-branching aunts-nieces makes sense.

It also makes sense for the comparison between left-branching aunts-nieces and the right-branching sisters in Figure 6.8(e'). Sisters were expected to be fused more often than aunts-nieces, but the result did not come out as expected. Table 6.4 shows that among the 6 right-branching sisters, only 33% of them had the fusion-inducing segments at the boundary that crosses two syllables, while 57% of the left-branching aunts-nieces had the fusion-inducing segments.

A comparison between left-branching sisters and right-branching aunts-nieces showed another unexpected outcome. That is, left-branching sisters were expected to have more fused forms than right-branching aunts-nieces, but the prediction did not borne out. Table 6.4 shows that among the 7 left-branching sisters, 71% of them had the fusion-inducing segments at the boundary that crosses two syllables, while only 50% of the right-branching aunts-nieces had the fusion-inducing segments. The outcome, then, makes sense in light of the presence (or absence) of fusion-inducing segments at the boundary that crosses two syllables.

words	'underlying' form	gloss	presence of fusion-inducing segments: codas [m, t, k], onsets [j, h, m, p, p^h], or [m̩] across syllable boundary of…	
low frequency 3 syllables				
left branching			**sisters**	**aunt-niece**
<L3><L2>音樂</L2>會</L3>	jɐm^{55}ŋɔːk^{2}wuːi^{35}	concert	yes	yes
<L3><L2>諗嚟</L2>咩</L3>	nɐm^{35}kɐn^{35}mɛː^{55}	thinking what	yes	yes
<L3><L2>勞工</L2>界</L3>	lou^{21}koŋ^{55}kaːi^{33}	labor sector		
<L3><L2>審計</L2>處</L3>	sɐm^{35}kɐi^{33}ts^{h}yː^{23}	audit department	yes	
<L3><L2>準備</L2>咗</L3>	tsɵn^{35}pei^{33}tsɔː^{35}	prepared	yes	
<L3><L2>超越</L2>到</L3>	ts^{h}iːu^{55}jyːt^{2}tou^{35}	surpass		yes
<L3><L2>不得</L2>了</L3>	pɐt^{5}tɐk^{5}liːu^{23}	intensified	yes	yes
		% yes:	**71%**	**57%**

Table 6.4 Presence of target segmental sequences across syllable boundary that might have encouraged fusion for low frequency 3-syllable words.

Table 6.4 continued

words	'underlying' form	gloss	presence of fusion-inducing segments: codas [m, t, k], onsets [j, h, m, p, p^h], or [m̩] across syllable boundary of…	
low frequency 3 syllables				
right branching			**aunt-niece**	**sisters**
<L3>一<L2>大輪</L2></L3>	jɐt^{5}taːi^{22}lɔn^{21}	a round of	yes	
<L3>傾<L2>唔妥</L2></L3>	k^{h}eŋ^{55}m̩^{21}t^{h}ɔː^{23}	can't come to agreement	yes	yes
<L3>女<L2>家鬼</L2></L3>	nɵy^{23}kaː^{55}kwɐi^{35}	female house ghost		
<L3>唔<L2>合法</L2></L3>	m̩^{21}hɐp^{2}faːt^{3}	illegal	yes	
<L3>呢<L2>一家</L2></L3>	nei/ni/ji^{55}jɐt^{5}kaː^{33}	this one		yes
<L3>上<L2>天堂</L2></L3>	sœːŋ^{23}t^{h}iːn^{55}t^{h}ɔːŋ^{21}	go to heaven		
		% yes:	**50%**	**33%**

Table 6.4 ended

For the 4-syllable words, my intuition for the rhythm of the left-branching words in (Hong Kong) Cantonese is that the first three syllable would be grouped together and distributed over two half notes (although I do not have a strong sense of how the three syllables would be distributed among the two half notes), while the last syllable would make a full note all by itself. It follows that fusion would be more likely for any or all of the pairs of two syllables among the first three syllables, but less likely between the third and the last syllables. For 4-syllable words that have a right-branching structure, my intuition is that they share the same rhythmic pattern with balanced branching, single-level branching and no branching structures, in all of which the first two syllables make a full note and so do the last two syllables. It follows that the first two syllables would be more or less as likely to be fused as the last two syllables within the word. Figure 6.12 gives representations for the mapping between the musical rhythm and five different branching structures for the 4-syllable words in (Hong Kong) Cantonese.

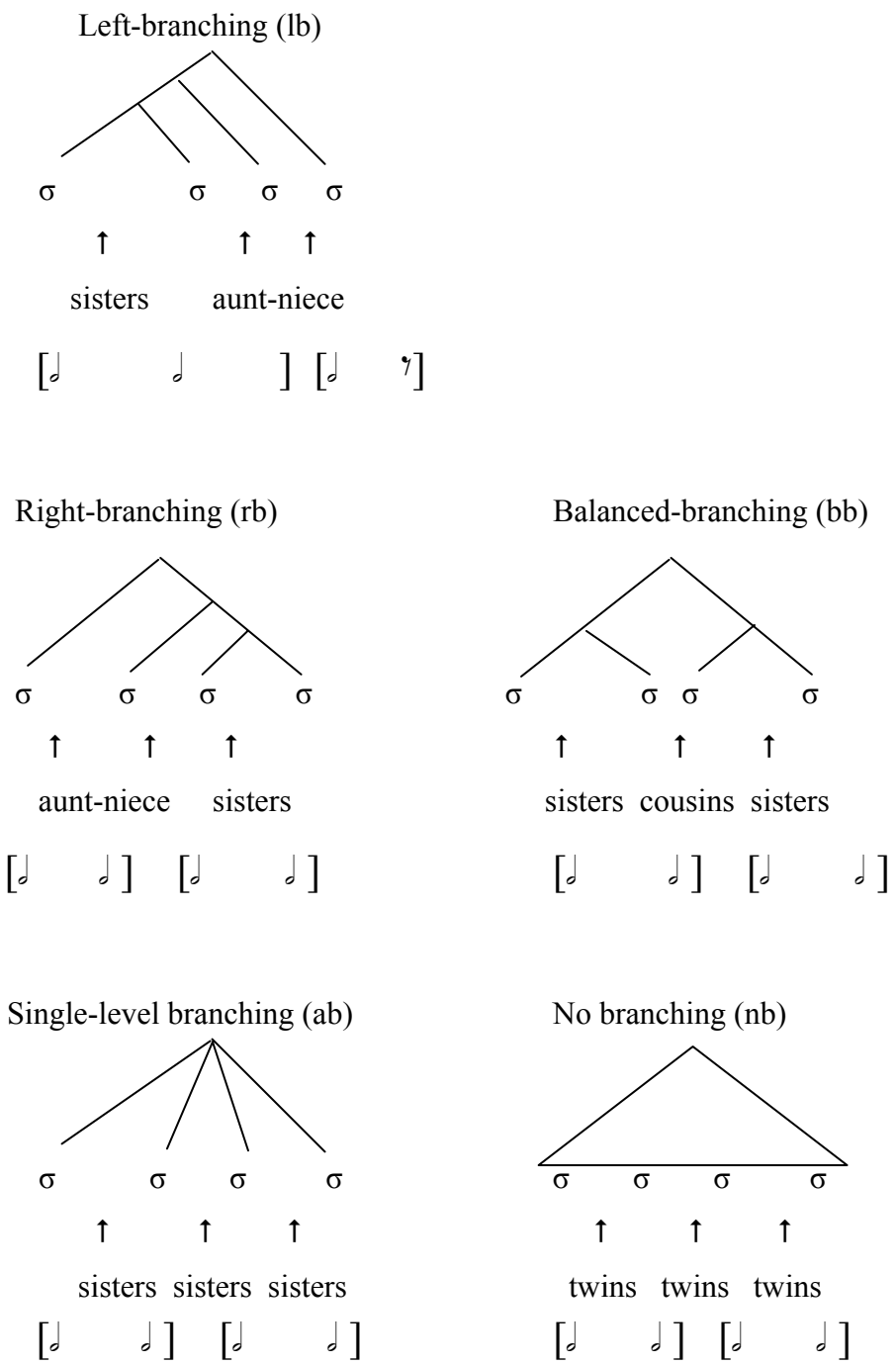


Figure 6.12 Syllables in five different branching structures for the 4-syllable words in (Hong Kong) Cantonese.

If my intuition about the rhythmic patterns for the 4-syllable words holds some truth, then some outcomes for the low frequency 4-syllable words in Figure 6.9(f') were unpredicted. First, left-branching sisters had fewer fused forms and were less fused than right-branching, which was unexpected. Table 6.5 lists the low frequency 4-syllable target words, within each were the relationships sisters and aunt-niece. Among the 16 target words in the left-branching structure, there were 38% of the sisters that had one or more of segments that induced fusion across the syllable boundary. For the right-branching sisters, however, there were more (40%) of the fusion-inducing segments. Thus, when viewed in the light of the presence (or absence) of fusion-inducing segments that were contiguous to the syllable boundaries, the outcome that left-branching sisters were fused more often and were more fused than the right-branching sisters makes sense.

Second, for the aunts-nieces in Figure 6.9(f'), the outcomes were unexpected. The rhythm for the 4-syllable left- and right-branching words would predict similar amount of fused forms for aunts-nieces in the two branching structures. However, the left-branching aunts-nieces came out to fewer fused forms than the right-branching aunts-nieces. An examination of the presence (or absence) of fusion-inducing segments that were contiguous to the syllable boundaries would also predict more fused forms for the right-branching aunts-nieces than for the left-branching aunts-nieces. A closer look at Figure 6.9(f') shows that although the relative amount of non-fused forms between the right-branching aunts-nieces than for the left-branching aunts-nieces were not expected, the right-branching aunts-nieces actually had more extreme form of fusion (to the extent of deleting at least one segment contiguous to the syllable boundary) more often than the left-branching aunts-nieces. Table 6.5 lists the low frequency 4-syllable target words, within each were the relationships sisters and aunt-niece. Among the 16 target words in the left-branching structure, an average of (44%+50%)/2 = 47% of the aunts-nieces had one or more of segments that induced fusion across the syllable boundary. For

the right-branching aunts-nieces, however, there were more (an average of (87%+47%)/2 = 67%) of the fusion-inducing segments. Thus, the outcome that left-branching aunts-nieces were fused less often in terms of more extreme forms of fusion than the right-branching aunts-nieces makes sense.

words	'underlying' form	gloss	presence of fusion-inducing segments: codas [m, t, k], onsets [j, h, m, p, p^h], or [m̩] across syllable boundary of…		
low frequency 4 syllables					
left branching			**sis.**	**aunt-n.**	**aunt-n.**
<L4><L3><L2>起得</L2>番</L3>身</L4>	hei^{35}tɐk^{5}faːn^{55}sɐn^{55}	can stand back up		yes	
<L4><L3><L2>學習</L2>得</L3>到</L4>	hɔːk^{3}tsaːp^{2}tɐk^{5}tou^{35}	can learn	yes		yes
<L4><L3><L2>醫番</L2>好</L3>咗</L4>	jiː^{55}faːn^{55}hou^{35}tsɔː^{35}	be healed		yes	
<L4><L3><L2>預見</L2>得</L3>到</L4>	jyː^{22}kiːn^{33}tɐk^{5}tou^{35}	can foresee			yes
<L4><L3><L2>機械</L2>人</L3>式</L4>	kei^{55}haːi^{22}jɐn^{21}sek^{5}	robotic	yes	yes	
<L4><L3><L2>匿咗</L2>入</L3>房</L4>	lei^{55}tsɔː^{35}jɐp^{2}fɔːŋ^{35}	hide in the room		yes	
<L4><L3><L2>唔見</L2>咗</L3>人</L4>	m̩^{21}kiːn^{33}tsɔː^{35}jɐn^{21}	the person disappeared	yes		yes
<L4><L3><L2>順住</L2>嚟</L3>擺</L4>	sɵn^{22}tsyː^{22}lei^{21}paːi^{35}	to put in order			yes
<L4><L3><L2>寫咗</L2>信</L3>去</L4>	sɛː^{35}tsɔː^{35}sɵn^{33}hɵy^{33}	wrote a letter to somewhere			yes
<L4><L3><L2>釋放</L2>番</L3>晒</L4>	sek^{5}fɔːŋ^{33}faːn^{55}saːi^{33}	released	yes		
<L4><L3><L2>打過</L2>去</L3>問</L4>	taː^{35}kwɔː^{33}hɵy^{33}mɐn^{22}	called to ask		yes	yes

Table 6.5 Presence of target segmental sequences across syllable boundary that might have encouraged fusion for low frequency 4-syllable words.

Table 6.5 continued

words	'underlying' form	gloss	presence of fusion-inducing segments: codas [m, t, k], onsets [j, h, m, p, p^h], or [m̩] across syllable boundary of…		
low frequency 4 syllables					
left branching			**sis.**	**aunt-n.**	**aunt-n.**
<L4><L3><L2>代落</L2>咗</L3>去</L4>	tɔːi^{22}lɔːk^{2}tsɔː^{35}hɵy^{33}	putt in (some stuff in somewhere)		yes	yes
<L4><L3><L2>站得</L2>住</L3>腳</L4>	tsaːm^{22}tɐk^{5}tsyː^{22}kœːk^{3}	can stand up (under some conditions)	yes	yes	
<L4><L3><L2>走番</L2>出</L3>廳</L4>	tsɐu^{35}faːn^{55}ts^{h}ɵt^{5}t^{h}ɛːŋ^{55}	went back out to the living room			yes
<L4><L3><L2>沖完</L2>涼</L3>走</L4>	ts^{h}oŋ^{55}jyːn^{21}lœːŋ^{21}tsɐu^{35}	left after shower	yes		
<L4><L3><L2>搵嚟</L2>工</L3>做</L4>	wɐn^{35}kɐn^{35}koŋ^{55}tsou^{22}	looking for a job			
		% yes:	**38%**	**44%**	**50%**

Table 6.5 ended

Table 6.5 continued

words	'underlying' form	gloss	presence of fusion-inducing segments: codas [m, t, k], onsets [j, h, m, p, p^h], or [m̩] across syllable boundary of…		
low frequency 4 syllables					
right branching			**aunt-n.**	**aunt-n.**	**sis.**
<L4>去<L3>諗<L2>辦法</L2></L3></L4>	hɵy^{33}nɐm^{35}paːn^{22}faːt^{3}	go think out a way		yes	
<L4>去<L3>買<L2>外賣</L2></L3></L4>	hɵy^{33}maːi^{23}ŋɔːi^{22}maːi^{22}	go buy carry-out	yes		yes
<L4>去<L3>唔<L2>去街</L2></L3></L4>	hɵy^{33}m̩^{21}hɵy^{33}kaːi^{55}	go out or not	yes	yes	
<L4>有<L3>咩<L2>聯想</L2></L3></L4>	jɐu^{23}mɛː^{55}lyːn^{21}sœːŋ^{35}	has any association	yes		
<L4>容<L3>有<L2>不當</L2></L3></L4>	joŋ^{21}jɐu^{23}pɐt^{5}tɔːŋ^{33}	bear with wrong doings	yes		yes
<L4>記<L3>唔<L2>記到</L2></L3></L4>	kei^{33}m̩21kei^{33}tou^{35}	can remember or not	yes	yes	
<L4>亂<L3>拋<L2>垃圾</L2></L3></L4>	lyːn^{22}p^{h}aːu^{55}laːp^{2}saːp^{3}	littering	yes		
<L4>心<L3>有<L2>靈犀</L2></L3></L4>	sɐm^{55}jɐu^{23}leŋ^{21}sɐi^{55}	read each other's mind	yes		
<L4>嗰<L3>一<L2>類型</L2></L3></L4>	kɔː^{35}jɐt^{5}lɵy^{22}jeŋ^{21}	that type	yes	yes	yes
<L4>心<L3>有<L2>不妙</L2></L3></L4>	sɐm^{55}jɐu^{23}pɐt^{5}miːu^{22}	make one feel at peace in the heart	yes		yes
<L4>四<L3>大<L2>天王</L2></L3></L4>	sei^{33}taːi^{33}tiːn^{55}wɔːŋ^{21}	four top singers			
<L4>當<L3>唔<L2>準確</L2></L3></L4>	tɔːŋ^{33}m^{21}tsɵn^{35}k^{h}ɔːk^{3}	be regarded as inaccurate	yes	yes	
<L4>當<L3>冇<L2>件事</L2></L3></L4>	tɔːŋ^{33}mou^{23}kiːn^{22}siː^{22}	pretend that nothing has happened	yes		
<L4>足<L3>唔<L2>足夠</L2></L3></L4>	tsok^{5}m̩^{21}tsok^{5}kɐu^{33}	enough or not	yes	yes	yes
<L4>鐘<L3>唔<L2>鐘意</L2></L3></L4>	tsoŋ^{55}m̩^{21}tsoŋ^{55}jiː^{33}	like it or not	yes	yes	yes
		% yes:	**87%**	**47%**	**40%**

Table 6.5 ended

In short, the reasons for the interactions between sisters and aunts-nieces that we saw between the mid and the low frequency 3-syllable words, and between the 3- and the 4-syllable words become clearer when the internal-branching structures of the words were examined. The rhythmic pressure I proposed for (Hong Kong) Cantonese and the presence (or absence) of fusion-inducing segments that were contiguous to the syllable boundaries seemed to interplay with the internal-branching structures that created the interaction effects.

Figure 6.13 replots data in Figure 6.8(a-e) so as to emphasize interaction between word frequency (bars) and each type of morphosyntactic relationship (columns). The frequency effect was bigger for the syllables that were sisters (polymorphemic words) than for twins (monomorphemic words) for the 2-syllable words, which might mean that sisters are increasingly more decomposable into smaller units than twins as word frequency decreases.

The frequency effect was not as big for the syllables that were sisters than for twins in the 3-syllable words. Also, an interaction can be observed for pairs of syllables that were in aunt-niece relationship in the mid and the low frequency. As we saw in Figure 6.9, the internal branching structures and the presence (or absence) of fusion-inducing segments that were contiguous to the syllable boundaries might have came into play and thus giving rise to these results.

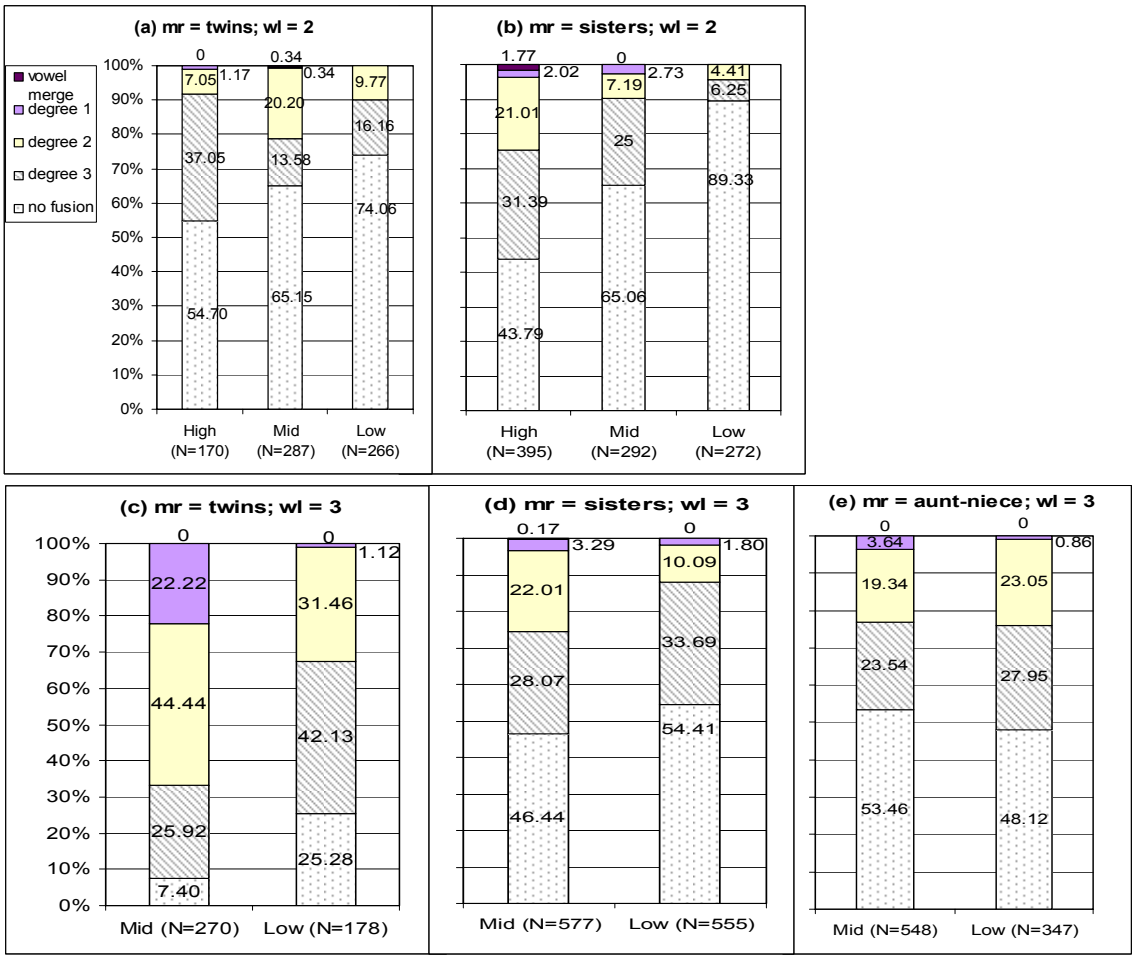


Figure 6.13 Replotting data in Figure 6.8(a-e) so as to emphasize interaction between word frequency (bars) and morphosyntactic relationship (columns).

Figure 6.14 plots the interactions between prosodic position and word length on syllable fusion. Results showed that across all prosodic positions (initial, medial and final) in an intonational phrase, 2-syllable words were least likely to have some degree of fusion and were least likely to have more extreme fusion forms to the extent of deleting at least one segment that is contiguous to the syllable boundary. The outcomes were expected, given that 2-syllable words have one cross-syllable boundary and that boundary is subject to word-final lengthening, whereas 4-syllable words have two-third of the cross-syllable boundaries that are not subject to word-final lengthening.

The interaction effect in Figure 6.14 can be seen among the three different word lengths. While for 3- and 4-syllable words, some degree of syllable fusion in the intonational phrase medial position was more likely when compared with syllable fusion in the phrase-initial position, some degree of syllable fusion in the intonational phrase medial position was less likely when compared with syllable fusion in the phrase-initial position for 2-syllable words. Again, this could be related to the rhythmic pressure that the 3- and the 4-syllable words have, but the 2-syllable words do not share, such that more fused forms were encouraged for the 3- and the 4-syllable words. Other possible motivations for this interaction would need further exploration.

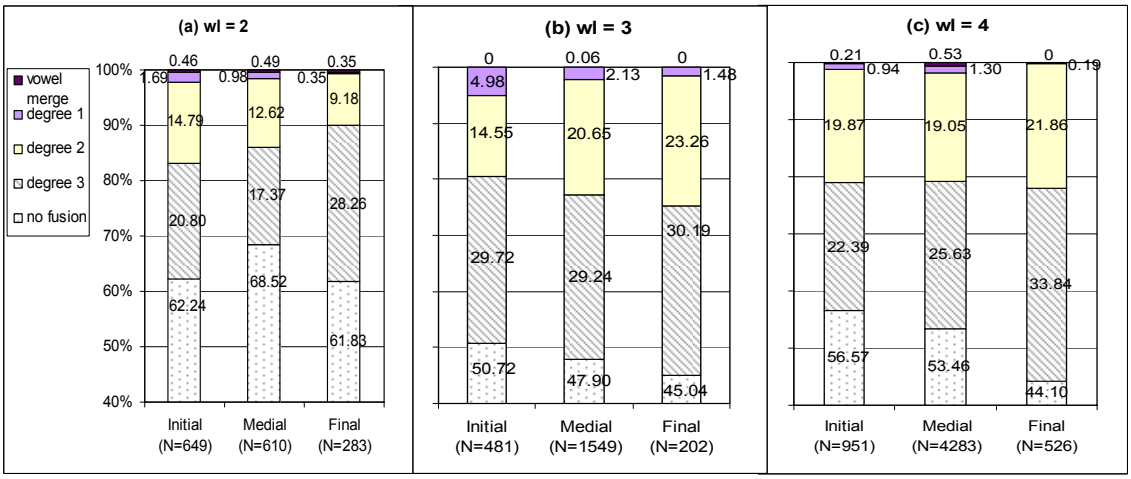


Figure 6.14 Plotting the interactions between prosodic position and word length on syllable fusion.

Figure 6.15 isolates the word length effect by plotting the interaction of word length with prosodic position for *just the low frequency words* in the two morphosyntactic relationships that are common to all three lengths. The figure shows that the general pattern for both of the two types of morphosyntactic relationship is that the 3-syllable words were most likely to have some degree of fusion, while the 4-syllable words were next likely to have some degree of fusion, except in the prosodic phrase-final position for the 4-syllable sisters. The 2-syllable words then were least likely to have some degree of fusion. Thus, an interaction effect can be seen in the low frequency 4-syllable words in the intonational phrase-final position (see Figure 6.15f). The interaction could be motivated by the rhythmic pressure on the 3- and the 4-syllable words. The rhythmic pressure would make the last two syllables less likely to be fused when compared with fusion between the last two syllables in the 4-syllable words. (See Figures 6.11 and 6.12. Note that the intuition about the rhythm of the (Hong Kong) Cantonese words comes from when the word makes an intonational phrase all by itself. That is, the intuition would be more directly applicable to intonational phrase-initial and phrase-final phenomena). The fact that 3- and 4-syllable twins (Figure 6.15c) did not show similar

pattern as their sisters counterparts (Figure 6.15f) could be due to the presence of more fusion-inducing segments (100%) at the boundary across the last two syllables in the 3-syllable than in the 4-syllable words (0%). See Table 6.6, in which the low frequency 3- and 4-syllable twins were listed regarding whether the fusion-inducing segments were present across the syllable boundary.

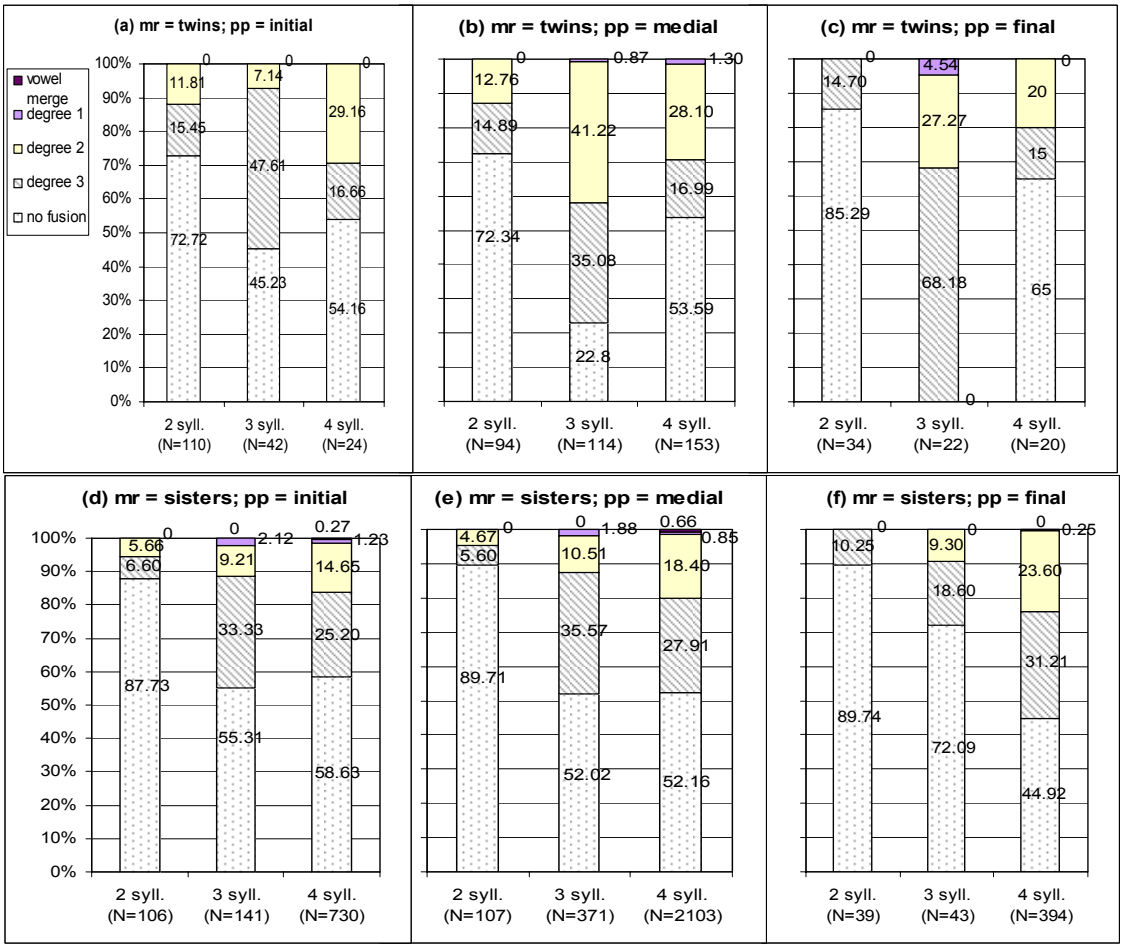


Figure 6.15 Isolating the word length effect by plotting the interaction of word length with prosodic position for *just the low frequency* words in the two morphosyntactic relationships that are common to all three lengths.

words	'underlying' form	gloss	presence of fusion-inducing segments: codas [m, t, k], onsets [j, h, m, p, p^h], or [m̩] across syllable boundary of…
low frequency 3-syllable twins			**the last two syllables (sisters)**
<L3>乜鬼嘢</L3>	mɐt^{5}kwɐi^{35}jɛː^{23}	whatever	yes
<L3>面對面</L3>	miːn^{22}tɵy^{33}miːn^{22}	face to face	yes
<L3>避一避</L3>	pei^{22}jɐt^{5}pei^{22}	avoid	yes
<L3>追趕跑</L3>	tsɵy^{55}kɔːn^{35}p^{h}aːu^{35}	chase	yes
		% yes:	**100%**
low frequency 4-syllable twins			
<L4>卡拉 OK</L4>	k^{h}aː^{55}laːi^{55}ou^{55}k^{h}ei^{55}	karaoke	
<L4>歇斯底里</L4>	kiːt^{3}siː^{55}tɐi^{35}lei^{23}	hysterical	
		% yes:	**0%**

Table 6.6 Presence of target segmental sequences across syllable boundary that might have encouraged fusion for low frequency 3- and 4-yllable twins.

(i) Testing main effect

```
factors:
wl = word length
wfc = word frequency category
mr = morphosyntactic relationship
pp = prosodic position

outcome:
tbd = degree of fusion (0, 1, 2, 3, 4) at each word-medial
syllable boundary

frequency of each unique combination of factor and outcome:
freq = the number of times a unique combination of factors and
outcome emerged
```

A control case built:

```
> mult.0=multinom(tbd~rep(1,length(freq.data$mr)),
weight=freq.data$freq, data=freq.data)
# weights:  15 (8 variable)
initial  value 15569.702365
iter  10 value 10412.758139
final  value 10412.562723
converged
```

Table 6.7 Statistical results using multinomial logistic regression for testing the main effect and interaction effect of the three potential factors on syllable fusion. (Edited output of R.)

Table 6.7 continued

Set mult.1.1 to be the input factor wl:

```
> mult.1.1=multinom(tbd~factor(wl), weight=freq.data$freq,
data=freq.data)
# weights:  20 (12 variable)
initial  value 15569.702365
iter  10 value 11072.942908
iter  20 value 10340.952495
iter  30 value 10340.036065
iter  40 value 10339.952371
iter  40 value 10339.952369
iter  40 value 10339.952369
final  value 10339.952369
converged
```

Significance of the input factor wl tested by Chi-square test. Result was p ≈ 0.

```
Likelihood ratio tests of Multinomial Models

Response: tbd
                          Model Resid. df Resid. Dev   Test    Df
LR stat. Pr(Chi)
1 rep(1, length(freq.data$mr))      2876   20825.13
2                   factor(wl)      2868   20679.90 1 vs 2     8
145.2207       0
```

Table 6.7 continued

Set mult.1.2 to be the input factor wfc:

```
> mult.1.2=multinom(tbd~factor(wl)+factor(wfc),
weight=freq.data$freq, data=freq.data)
# weights:  30 (20 variable)
initial  value 15569.702365
iter  10 value 10742.945988
iter  20 value 10282.603907
iter  30 value 10261.599054
iter  40 value 10261.364422
final  value 10261.363729
converged
```

Significance of the input factor wfc tested by Chi-square test. Result was p ≈ 0.

```
> anova(mult.1.1, mult.1.2)
Likelihood ratio tests of Multinomial Models

Response: tbd
                     Model Resid. df Resid. Dev   Test    Df LR
stat. Pr(Chi)
1               factor(wl)      2868   20679.90
2 factor(wl) + factor(wfc)      2860   20522.73 1 vs 2     8
157.1773       0
```

Set mult.1.3 to be the input factor mr.

```
> mult.1.3=multinom(tbd~factor(wl)+factor(wfc)+factor(mr),
weight=freq.data$freq, data=freq.data)
# weights:  45 (32 variable)
initial  value 15569.702365
iter  10 value 10692.533238
iter  20 value 10284.973004
iter  30 value 10210.922009
iter  40 value 10207.671442
iter  50 value 10207.537488
final  value 10207.533722
converged
```

Table 6.7 continued

Significance of the input factor wl tested by Chi-square test. Result was p ≈ 0.

```
Likelihood ratio tests of Multinomial Models

Response: tbd
                                   Model Resid. df Resid. Dev
Test    Df LR stat. Pr(Chi)
1              factor(wl) + factor(wfc)      2860   20522.73
2 factor(wl) + factor(wfc) + factor(mr)      2848   20415.07 1
vs 2    12 107.6600       0
```

Set mult.1.4 to be the input factor pp.

```
> mult.1.4=multinom(tbd~factor(wl)+factor(wfc)+factor(mr)+
factor(pp), weight=freq.data$freq, data=freq.data)
# weights:  60 (44 variable)
initial  value 15569.702365
iter  10 value 10911.737785
iter  20 value 10360.574538
iter  30 value 10217.926421
iter  40 value 10186.862224
iter  50 value 10185.393030
iter  60 value 10185.294698
iter  70 value 10185.244022
final  value 10185.243811
converged
```

Significance of the input factor pp tested by Chi-square test. Result was p = 0.000.

```
> anova(mult.1.3, mult.1.4)
Likelihood ratio tests of Multinomial Models

Response: tbd
                                                Model Resid. df
Resid. Dev   Test    Df
1              factor(wl) + factor(wfc) + factor(mr)      2848
20415.07
2 factor(wl) + factor(wfc) + factor(mr) + factor(pp)      2836
20370.49 1 vs 2    12
  LR stat.      Pr(Chi)
1
2 44.57982 1.216042e-05
```

Table 6.7 continued

(ii) Testing 2-term interaction effect

Set mult.2.1 to be the interaction term wl * wfc:

```
> mult.2.1=multinom(tbd~factor(wl)*factor(wfc)+factor(mr)+
factor(pp), weight=freq.data$freq, data=freq.data)
# weights:  80 (60 variable)
initial  value 15569.702365
iter  10 value 12027.161286
iter  20 value 11058.869508
iter  30 value 10488.230389
iter  40 value 10242.277239
iter  50 value 10172.260558
iter  60 value 10167.783162
iter  70 value 10167.457282
iter  80 value 10167.435201
iter  90 value 10167.431305
iter  90 value 10167.431243
iter  90 value 10167.431243
final  value 10167.431243
converged
```

Significance of the interaction term wl * wfc tested by Chi-square test. Result was p = 0.003.

```
> anova(mult.1.4, mult.2.1)
Likelihood ratio tests of Multinomial Models

Response: tbd
                                                  Model Resid. df
Resid. Dev   Test    Df
1 factor(wl) + factor(wfc) + factor(mr) + factor(pp)      2836
20370.49
2 factor(wl) * factor(wfc) + factor(mr) + factor(pp)      2820
20334.86 1 vs 2    16
  LR stat.     Pr(Chi)
1
2 35.62514 0.003260809
```

Table 6.7 continued

Set mult.2.2 to be the interaction term wl * mr:

```
> mult.2.2=multinom(tbd~factor(wl)*factor(wfc)+factor(wl)*
factor(mr)+factor(mr)+factor(pp), weight=freq.data$freq,
data=freq.data)
# weights:  110 (84 variable)
initial  value 15569.702365
iter  10 value 11850.106492
iter  20 value 11276.195812
iter  30 value 10609.296177
iter  40 value 10262.052624
iter  50 value 10161.635097
iter  60 value 10135.909711
iter  70 value 10132.300973
iter  80 value 10132.004613
iter  90 value 10131.982309
final  value 10131.980339
converged
```

Significance of the interaction term wl * mr tested by Chi-square test. Result was p = 0.000.

```
> anova(mult.2.1, mult.2.2)
Likelihood ratio tests of Multinomial Models

Response: tbd

Model Resid. df
1                           factor(wl) * factor(wfc) +
factor(mr) + factor(pp)      2820
2 factor(wl) * factor(wfc) + factor(wl) * factor(mr) +
factor(mr) + factor(pp)      2796
  Resid. Dev   Test    Df LR stat.      Pr(Chi)
1   20334.86
2   20263.96 1 vs 2    24  70.9018 1.595158e-06
```

Table 6.7 continued

Set mult.2.3 to be the interaction term wl * pp:

```
> mult.2.3=multinom(tbd~factor(wl)*factor(wfc)+factor(wl)*
factor(mr)+factor(wl)*factor(pp)+factor(mr)+factor(pp),
weight=freq.data$freq, data=freq.data, maxit=10000)
# weights:  140 (108 variable)
initial  value 15569.702365
iter  10 value 11665.625629
iter  20 value 11056.690503
iter  30 value 10455.713532
iter  40 value 10270.102103
iter  50 value 10152.655234
iter  60 value 10123.548834
iter  70 value 10113.296086
iter  80 value 10111.298154
iter  90 value 10111.113721
iter 100 value 10111.092314
final  value 10111.091250
converged
```

Significance of the interaction term wl * pp tested by Chi-square test. Result was p = 0.013.

```
> anova(mult.2.2, mult.2.3)
Likelihood ratio tests of Multinomial Models

Response: tbd

Model
1                           factor(wl) * factor(wfc) +
factor(wl) * factor(mr) + factor(mr) + factor(pp)
2 factor(wl) * factor(wfc) + factor(wl) * factor(mr) +
factor(wl) * factor(pp) + factor(mr) + factor(pp)
  Resid. df Resid. Dev   Test    Df LR stat.    Pr(Chi)
1      2796   20263.96
2      2772   20222.18 1 vs 2    24 41.77818 0.01366254
```

Table 6.7 continued

Set mult.2.4 to be the interaction term wfc * mr:

```
> mult.2.4=multinom(tbd~factor(wl)*factor(wfc)+factor(wl)*
factor(mr)+factor(wl)*factor(pp)+factor(wfc)*factor(mr)+
factor(mr)+factor(pp), weight=freq.data$freq, data=freq.data,
maxit=10000)
# weights:  170 (132 variable)
initial  value 15569.702365
iter  10 value 11529.769972
iter  20 value 11153.567807
iter  30 value 10468.990479
iter  40 value 10281.405630
iter  50 value 10126.533099
iter  60 value 10085.963770
iter  70 value 10069.266770
iter  80 value 10063.817872
iter  90 value 10062.380242
iter 100 value 10062.231628
iter 110 value 10062.205262
final  value 10062.203531
converged
```

Significance of the interaction term wfc * mr tested by Chi-square test. Result was p = 0.000.

```
> anova(mult.2.3, mult.2.4)
Likelihood ratio tests of Multinomial Models

Response: tbd

Model
1                            factor(wl) * factor(wfc) +
factor(wl) * factor(mr) + factor(wl) * factor(pp) + factor(mr) +
factor(pp)
2 factor(wl) * factor(wfc) + factor(wl) * factor(mr) +
factor(wl) * factor(pp) + factor(wfc) * factor(mr) + factor(mr)
+ factor(pp)
  Resid. df Resid. Dev   Test    Df LR stat.      Pr(Chi)
1      2772   20222.18
2      2748   20124.41 1 vs 2    24 97.77544 7.177348e-11
```

Table 6.7 continued

Set mult.2.5 to be the interaction term wfc * pp:

```
> mult.2.5=multinom(tbd~factor(wl)*factor(wfc)+factor(wl)*
factor(mr)+factor(wl)*factor(pp)+factor(wfc)*factor(mr)+
factor(wfc)*factor(pp)+factor(mr)+factor(pp),
weight=freq.data$freq, data=freq.data, maxit=10000)
# weights:  200 (156 variable)
initial  value 15569.702365
iter  10 value 11452.643522
iter  20 value 11184.593491
iter  30 value 10445.357381
iter  40 value 10237.985238
iter  50 value 10122.096752
iter  60 value 10077.124968
iter  70 value 10059.817873
iter  80 value 10053.929047
iter  90 value 10051.143299
iter 100 value 10050.331467
iter 110 value 10050.272217
final  value 10050.267236
converged
```

Significance of the interaction term wfc * pp tested by Chi-square test. Result was p = 0.468.

```
> anova(mult.2.4, mult.2.5)
Likelihood ratio tests of Multinomial Models

Response: tbd

Model
1                            factor(wl) * factor(wfc) +
factor(wl) * factor(mr) + factor(wl) * factor(pp) + factor(wfc)
* factor(mr) + factor(mr) + factor(pp)
2 factor(wl) * factor(wfc) + factor(wl) * factor(mr) +
factor(wl) * factor(pp) + factor(wfc) * factor(mr) + factor(wfc)
* factor(pp) + factor(mr) + factor(pp)
  Resid. df Resid. Dev   Test    Df LR stat.   Pr(Chi)
1      2748   20124.41
2      2724   20100.53 1 vs 2    24 23.87259 0.4689021
```

Table 6.7 continued

Set mult.2.6 to be the interaction term mr * pp:

```
> mult.2.6=multinom(tbd~factor(wl)*factor(wfc)+factor(wl)*
factor(mr)+factor(wl)*factor(pp)+factor(wfc)*factor(mr)+
factor(mr)*factor(pp)+factor(mr)+factor(pp),
weight=freq.data$freq, data=freq.data, maxit=10000)
# weights:  215 (168 variable)
initial  value 15569.702365
iter  10 value 11511.943803
iter  20 value 11063.070873
iter  30 value 10417.830051
iter  40 value 10237.799397
iter  50 value 10117.206609
iter  60 value 10072.335731
iter  70 value 10053.690233
iter  80 value 10048.473955
iter  90 value 10045.267620
iter 100 value 10043.963005
iter 110 value 10043.780715
iter 120 value 10043.772762
iter 120 value 10043.772684
iter 120 value 10043.772683
final  value 10043.772683
converged
```

Significance of the interaction term mr * pp tested by Chi-square test. Result was p = 0.428.

```
> anova(mult.2.4, mult.2.6)
Likelihood ratio tests of Multinomial Models

Response: tbd

Model
1                           factor(wl) * factor(wfc) +
factor(wl) * factor(mr) + factor(wl) * factor(pp) + factor(wfc)
* factor(mr) + factor(mr) + factor(pp)
2 factor(wl) * factor(wfc) + factor(wl) * factor(mr) +
factor(wl) * factor(pp) + factor(wfc) * factor(mr) + factor(mr)
* factor(pp) + factor(mr) + factor(pp)
  Resid. df Resid. Dev   Test    Df LR stat.   Pr(Chi)
1      2748   20124.41
2      2712   20087.55 1 vs 2    36  36.8617 0.4288636
```

Table 6.7 continued

(iii) Testing 3-term interaction effect

Set mult.3.1 to be the interaction term mr * wl * wfc:

```
> mult.3.1=multinom(tbd~factor(wl)*factor(wfc)+factor(wl)*
factor(mr)+factor(wl)*factor(pp)+factor(wfc)*factor(mr)+
factor(mr)*factor(wl)*factor(wfc)+factor(mr)+factor(pp),
weight=freq.data$freq, data=freq.data, maxit=10000)
# weights:  230 (180 variable)
initial  value 15569.702365
iter  10 value 11525.439263
iter  20 value 11142.127543
iter  30 value 10575.161132
iter  40 value 10291.741990
iter  50 value 10141.554349
iter  60 value 10087.502764
iter  70 value 10066.655717
iter  80 value 10059.371989
iter  90 value 10057.353177
iter 100 value 10057.197430
iter 110 value 10057.181384
final  value 10057.179364
converged
```

Significance of the interaction term mr * wl * wfc tested by Chi-square test. Result was p = 1.

```
> anova(mult.2.4, mult.3.1)
Likelihood ratio tests of Multinomial Models

Response: tbd

Model
1                                          factor(wl) *
factor(wfc) + factor(wl) * factor(mr) + factor(wl) * factor(pp)
+ factor(wfc) * factor(mr) + factor(mr) + factor(pp)
2 factor(wl) * factor(wfc) + factor(wl) * factor(mr) +
factor(wl) * factor(pp) + factor(wfc) * factor(mr) + factor(mr)
* factor(wl) * factor(wfc) + factor(mr) + factor(pp)
  Resid. df Resid. Dev   Test    Df LR stat. Pr(Chi)
1      2748   20124.41
2      2700   20114.36 1 vs 2    48 10.04834       1
```

Table 6.7 continued

Coefficients for the final modal of the data in Experiment III (i.e. main effects and 2-term interactions)

```
> summary(mult.2.4)$coefficients
  (Intercept) factor(wl)3 factor(wl)4 factor(wfc)low factor(wfc)mid factor(mr)cousins
1    17.27676    20.91672   12.716755      -20.10470       16.95334      -0.04302124
2    19.94306    11.17009   -4.637388       16.68381       10.22842      -0.38534033
3    21.22956    10.47431   -4.760836       15.83217       10.69575      -0.28060499
4    21.46954    10.14500   -5.468207       17.25716       11.02317      -0.13333464
  factor(mr)sisters factor(mr)twins factor(pp)I factor(pp)I&F factor(pp)M
1        -17.93922        9.014940  1.24552328    -0.1421248   0.5873678
2        -17.38549        9.005976  0.08266047    -0.9212225  -0.1196072
3        -17.73771        9.821788 -0.62887031    -0.9676403  -0.8273585
4        -17.83768        9.844286 -0.46736837    -1.1162724  -0.4069767
  factor(wl)3:factor(wfc)low factor(wl)4:factor(wfc)low factor(wl)3:factor(wfc)mid
1               30.829847               12.716755               -9.913128
2                5.012358               -4.637388                6.157735
3                5.321804               -4.760836                5.152507
4                4.469987               -5.468207                5.675014
  factor(wl)4:factor(wfc)mid factor(wl)3:factor(mr)cousins factor(wl)4:factor(mr)cousins
1                          0                           0                -0.04302124
2                          0                           0                -0.38534033
3                          0                           0                -0.28060499
4                          0                           0                -0.13333464
  factor(wl)3:factor(mr)sisters factor(wl)4:factor(mr)sisters factor(wl)3:factor(mr)twins
1                     -22.17131                      -6.08177                    8.324395
2                     -27.90961                     -10.37692                   -1.359982
3                     -27.77179                     -10.82024                   -1.663562
4                     -28.36798                     -11.38352                   -2.629653
  factor(wl)4:factor(mr)twins factor(wl)3:factor(pp)I factor(wl)4:factor(pp)I
1                    27.36111                3.116727                -21.52555
2                    17.85666                2.741143                -22.05853
3                    16.82358                3.890693                -21.67867
4                    17.59501                3.852201                -21.17204
  factor(wl)3:factor(pp)I&F factor(wl)4:factor(pp)I&F factor(wl)3:factor(pp)M
1                         0                         0                -8.672380
2                         0                         0                -8.301207
3                         0                         0                -7.497284
4                         0                         0                -7.868229
```

Table 6.7 continued

```
  factor(wl)4:factor(pp)M factor(wfc)low:factor(mr)cousins
1             -21.78257                  -0.04302124
2             -22.91433                  -0.38534033
3             -22.30893                  -0.28060499
4             -22.34381                  -0.13333464
  factor(wfc)mid:factor(mr)cousins factor(wfc)low:factor(mr)sisters
1                        0                   22.99230
2                        0                   26.71288
3                        0                   27.87638
4                        0                   28.53578
  factor(wfc)mid:factor(mr)sisters factor(wfc)low:factor(mr)twins
1                         5.332935                  -13.660201
2                        10.752384                   -3.876723
3                        11.047587                   -4.225170
4                        11.390061                   -4.459658
  factor(wfc)mid:factor(mr)twins
1                      -44.01743
2                      -34.99888
3                      -37.31636
4                      -36.64880
>
```

Table 6.7 ended

CHAPTER 7

CONCLUSION

In this dissertation, we have shown that syllable fusion in Hong Kong Cantonese is clearly a prosodically driven process. For example, in Experiment I, we showed that the occurrence of syllable fusion is highly dependent on speech rate. The faster the speech rate, the more likely the occurrence of fused forms and the more likely that the target sequences would fuse to a greater extent resulting in the merging of two vowels. Similarly, in Experiment III, we showed that the degree of syllable fusion depends on prosodic position and on word length. Two phrase-medial syllables are more likely to show more extreme degrees of fusion than two phrase-final syllables. Also, syllable sequences in three- and four-syllable words (where they are less likely to be word-final) are more likely to show more extreme degrees of syllable fusion than syllable sequences in disyllabic words. These results help to make clear that syllable fusion should not be confused with the morphological phenomenon of contraction, as syllable fusion exhibits a very fine-grained continuum, while contraction suggests categorical changes. Although we selected two or (in Experiments II and III) five degrees of syllable fusion for the analysis for our experiments, the number of degrees of fusion apparently far exceeds the number selected.

That syllable fusion is a prosodically driven process suggests that syllable fusion marks some constituent in the prosodic hierarchy of Cantonese. The current description of the Cantonese prosodic system that is assumed in the Cantonese ToBI annotation system employs the prosodic construct, the foot, to capture the phenomenon of syllable fusion. That is, the Cantonese ToBI system posits this structurally well-defined

intermediate level of phrasing between the syllable and the intonational phrase, thus identifying syllable fusion to phenomena such as tone sandhi in Mandarin and Shanghainese, varieties that have a much less "monosyllabic" feel than Cantonese does. By this account, then the “monosyllabic” flavor of Cantonese (is ascribed to the preponderance of monosyllabic feet in careful read-speech styles). Fused forms then can be analyzed as erasure of foot boundaries. In other words, syllable fusion forms are exceptions to the regular one-to-one correspondence between foot and syllable.

What motivates the erasure of a foot boundary? Are there factors that influence the fusion of syllables, and thus the formation of a foot domain? Results of the three experiments reported in this dissertation showed that there are. As noted already, Experiment I, in which subjects repeated sequences of target words and fillers in a divided attention task, showed that the incidence of syllable fusion is gradiently related to the specch rate. Experiment II, in which spontaneous speech data in the Hong Kong Cantonese Adult Language Corpus (HKCAC) was analyzed and three input factors were tested, showed that the factors word frequency and morphosyntactic relationship significantly affected the occurrence of syllable fusion and the degree of fusion, while word length was not found to be a significant factor. No interactions among the factors were found. Experiment III was a more controlled experiment. It was a sentence-repetition task that used the same target words in Experiment II, testing the same three factors as in Experiment II — word frequency, morphosyntactic relationship, and word length — as well as the new factor prosodic position. All of these factors were controlled for in the target words. Results of Experiment III showed that all four factors significantly affected the occurrence of syllable fusion and the degree of fusion, and that two-term interaction effects were found among the factors. Word length interacted with each of the other three factors, while word frequency also interacted with morphosyntactic relationship.

The main effects and interactions in the two experiments that used the same set of materials were not the same, which might be due to the fact that the design of the two experiments was not the same. Word frequency tested in Experiment II might have been corpus-specific, rather than the "inherent" frequency of the words for general Hong Kong Cantonese speakers, whereas in Experiment III, the "inherent" frequency of words was being tested. The experimental design would affect the contexts in which the target words occurred, which obviously were not the same between the two experiments. The target words that were collected from the HKCAC occurred in a larger discourse context as speech in the HKCAC was from people who called in to the radio programs discussing certain topics or issues, while the contexts for the target words in Experiment III were the sentences that the subjects repeated. Hence, information load of the tokens of the target words in the two experiments could be different. Whether a token of the target word carried new or given information would be subject to the analysis of the speakers in a given context. A growing body of research has found that speakers tend to have phonological processes such as reduction, shortening, deletion, etc, when words are predictable, i.e. less informative, or whose (local) frequency of use is or becomes high (e.g. Jurafsky, Bell, Gregory & Raymond 2001, Gregory 2001 and references cited there in, Raymond, Dautricourt & Hume (submitted), etc). While speech in the HKCAC gives completely spontaneous speech data for analysis, control for the prosodic position of the target words was hardly possible. Prosodic position of the target words in Experiment II thus might be confounded with the other two factors tested in the experiment. In Experiment III, the prosodic position of the target words was able to be controlled for. Although speech elicited in a laboratory setting might not be completely spontaneous as speech that occurs in everyday life, the laboratory setting allows all the factors identified to be tested in this dissertation to be controlled for in the target words.

Taken together results from Experiments II and III, word frequency, word length, morphosyntactic relationship, and prosodic position all significantly affect the occurrence of syllable fusion and the degrees of fusion. While the test statistics cannot be interpreted in a way that we can rank the relative importance of each main factor, patterns of how

each factor influence the occurrence of syllable fusion can be observed in the distribution of fusion forms in the test results. Overall, results for word frequency in Experiments II and III (with interactions teased apart from the main effect) showed that, high frequency words were fused more often and to a greater extent than mid frequency words, and mid frequency words were fused more often and to a greater extent than the low frequency ones.

For morphosyntactic relationship, the overall test results from the two experiments showed that pairs of syllables that were cousins were least likely to be fused. Pairs of syllables that were twins would be fused more often and to a greater extent than those that were sisters, except that in the high frequency category, sisters were fused more often and to a greater extent than twins. The subjects might have treated sisters with high collocational frequency to be even more indecomposable than twins. Word-internal branching structures of the words played a role in syllable fusion. For pairs of syllables that were in a sisters or in an aunt-niece relationship, word boundaries within a word were more likely to be fused and fused to a greater extent when the boundaries occurred in a left-branching structure than in right-branching structure. This might be partly due to the effects of word-final lengthening in Cantonese. Rhythmic pressure in Cantonese also plays a role, which tends to group two syllables from the left to the right. In addition, fusion is partly facilitated by the presence of fusion-inducing segments at the boundaries across two syllables.

For word length, although it was not found to be a significant factor in Experiment II, it was found to have significant effect on syllable fusion in Experiment III. Results of Experiment III showed that, as expected, 2-syllable words were least likely to be fused when compared with 3- and 4-syllable words, contra the results in Experiment II, where 2-syllable words were found to be fused most often. That 2-syllable words were fused most often in Experiment II appears to be an artifact of the word frequency effect in

combination with the inability to balance the cells for different frequencies across different word lengths. The results could also be partly due to the fact that prosodic position was not controlled for in Experiment II.

For prosodic position, results in Experiment III showed that in general syllables (exclusively in 2-syllable words in this experiment) that made a complete intonational phrase were least likely to be fused. Between intonational phrase-initial, -medial and –final positons, syllable fusion was least likely at intonational phrase-initial position, but most likely at intonational phrase-final position. This ranking could be different for speech that is completely spontaneous.

The implications of the above findings for the prosodic hierarchy for Cantonese is that there are factors that contributes to defining a prosodic domain in an intermediate level of constituent below the larger intonational phrase and above the syllables in the prosodic hierarchy for Cantonese as proposed in Wong, Chan & Beckman (2005). Within the prosodic domain are fused forms whose syllable edges are blurred. We propose that the foot be used to capture the prosodic domain. With the assumption that the feet are monosyllabic feet, syllable fused forms then exhibit a weaker prosodic disjuncture between the syllables, and can be interpreted as erasure of foot boundaries. The edges of a prosodic foot exhibit "normal" disjuncture between syllables (i.e. no syllable fusion occurred). A schematic representation of the prosodic structure is given in Figure 7.1. It also gives an example utterance with the internal morphosyntactic structure of the utterance.

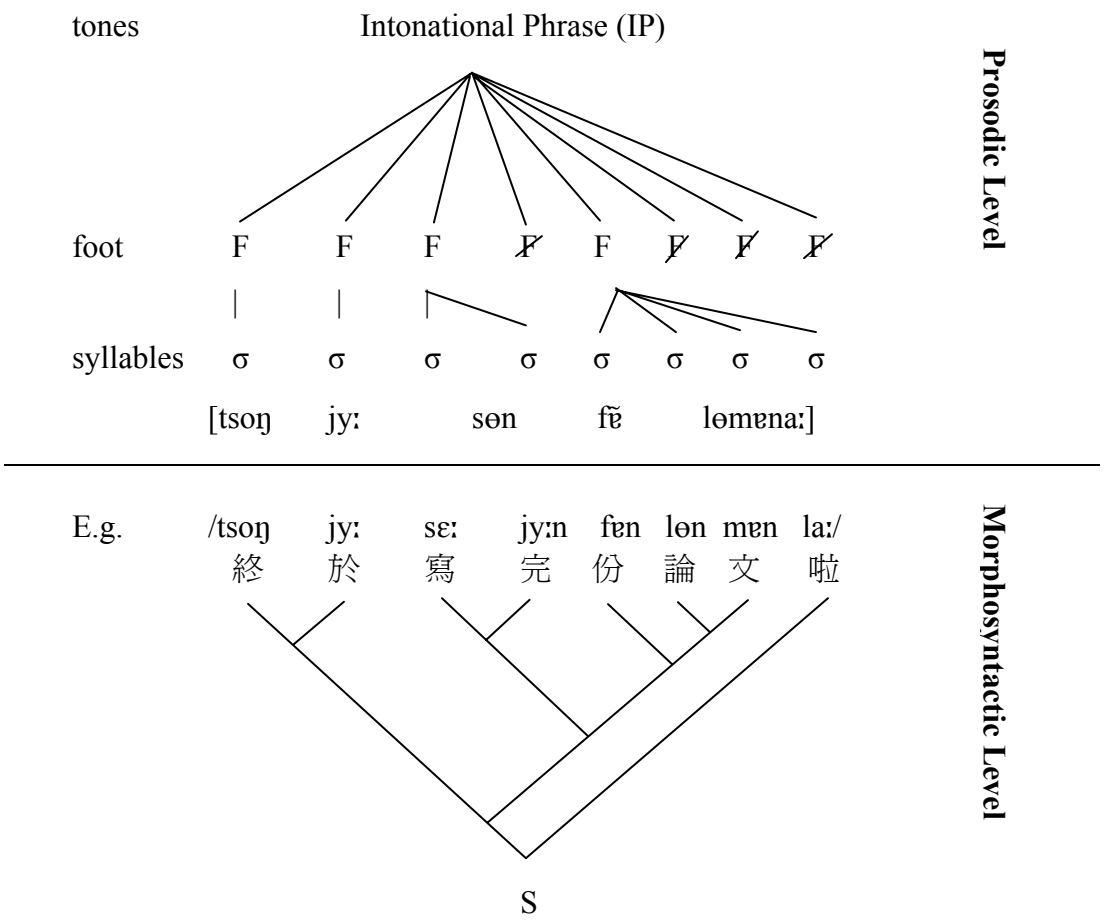


Figure 7.1 The foot level proposed for the Cantonese intonational structure. (A cross on a foot indicates a foot erased.) The utterance 'finally finished writing the dissertation (literally 'finally write finished classifier dissertation sentence-final particle')' and its internal morphosyntactic structure is given as an example.

Part of the Cantonese foot formation process would involve parsing the utterance from left to right with each syllable being a foot, joining immediate constituents into short words, and joining short words into longer words. The construction of a foot domain is sensitive to speech rate, word frequency, morphosyntactic relationship and word-internal branching structure, word length, and prosodic position of the word. For example, very fast speech would give rise to a "longer foot" (more syllables fused together within the foot), but fewer feet within an intonational phrase, when compared

with slow speech. In the most extreme case, an intonational phrase can be mapped to only one prosodic foot. The factors would interact to influence the likelihood of syllable fusion at the syllable boundary between any two syllables.

This dissertation only examined syllable fusion for target words with length up to four syllables. However, syllable fusion did occur between the target words and the adjacent non-target words in the subjects' productions. Also, there were cases showing that syllable fusion can occur across more than two syllables. (See the example utterance in Figure 7.1.) Thus, to have a more complete picture of the phenomena of syllable fusion, further investigation is needed to find out how syllable fusion works beyond the scope of our present study. There must be factors other than the five identified in this study that influence the occurrence of syllable fusion, such as segmental combinations. Some findings can be found in Lee (2003). An examination of the segments that are contiguous to the syllable boundaries in analyzing the results in Experiment III also suggests that the presence or absence of certain segments might influence the occurrence and probably even the degree of fusion. Further investigation would be desired, for example, to see whether segmental combination would be a significant contributing factor for syllable fusion, and if so, whether it would interact with any factors tested in this dissertation. Syllable fusion could also be defined at the tonal level. Future work should be an instrumental and perceptual probe into this area of tonal manifestation of syllable fusion in Cantonese. Other potential factors to look at for a more comprehensive understanding of syllable fusion in Cantonese would be information load of words, focus placement, dialectal background of the speakers. It is not clear how fused forms in Cantonese would affect comprehension. Psycholinguistic tests are desirable to find out how the Cantonese listeners resolve ambiguities, if any, in sentence processing. At the application level, it remains to be seen what would be a good method for modeling syllable fusion in a Cantonese text-to-speech synthesis (TTS) system which aims for naturalness and good comprehensibility in modeling casual Cantonese connected speech. If the concatenative approach is used to develop the Cantonese TTS system, what are the strategies for designing an inventory of a limited number of basic concatenative units that can capture

and generate fused forms wherever appropriate. While all of these questions are beyond the scope of the current study, the results reported in this dissertation contribute to a more developed foundation for formulating future experiments and studies that can address these questions.

BIBLIOGRAPHY

Bauer, R. S. and Benedict, P. K. (1997). *Modern Cantonese Phonology*. *Trends in Linguistics: Studies and Monographs 102*.Berlin and New York: Mouton de Gruyter.

Bai, W. (1982). '廣州方言連讀音變舉例' [Some examples of changed sounds through connected pronunciation in Cantonese dialect]. *Fanyan* 2:81-92.

Bybee, J. L. (2000). The phonology of the lexicon: evidence from lexical diffusion. In: M. Barlow & S. Kemmer (Eds.), *Usage Based Models of Language*. Stanford, CA: CSLI Publications

Cao, J. & Maddieson, I. (1992). 'An exploration of phonation types in Wu Dialects of Chinese'. *Journal of Phonetics* 20.1, 77-92.

Chan, M. K.-M. (1984). 'Initial consonant clusters in Old Chinese: evidence from sesqui-syllabic words in the Yue dialects'. *Fangyan* 4, 300-313.

Chan, M. K.-M. (1985). 'A non-linear analysis of tone and stress'. University Microfilms International, Ann Arbor, Michigan.

Chan, M. K.-M. & Ren, H. (1989). 'Wuxi tone sandhi: from last to first syllable dominance'. *Acta Linguistica Hafniensia* 21.2, 35-64.

Chan, M. K.-M. (1991). 'Contour-tone spreading and tone sandhi in Danyang'. *Phonology* 8.2, 237-259.

Chao, Y.-R. (1947). *Cantonese Primer.* Cambridge, Mass., Published for the Harvard-Yenching Institute by Harvard University Press.

Chao, Y.-R. (1968). *A Grammar of Spoken Chinese*, Berkeley: University of California Press.

Chen, M. Y. (1987), 'The syntax of Xiamen tone sandhi.' *Phonology Yearbook* 4, 109-149.

Chen, M. Y. (1990). 'What must phonology know about syntax.' In: *The Phonology-Syntax Connection*. Inkelas, S. and Zec, D (ed.), Chicago: University of Chicago Press.

Cheng, C. C. (1973). *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton.

Cheng, L.-S. & Sybesma, R. (2004). 'Postverbal 'can' in Cantonese (and Hakka) and Agree.' *Lingua* 114 (2004), 419-445.

Cheung, K.-H. (1986). 'The Phonology of Present-day Cantonese.' Ph.D. dissertation, University College London.

Cheung, L. Y. and Ngai, L. W. (1999). 港式廣州話詞典 [Dictionary of Hong Kong Style Cantonese]. Wan Li Book Company.

Cheung, P. M. (1999). 'Merging of consonantal endings in Hong Kong Cantonese: A microhistoric study'. M.Phil. thesis. The Hong Kong Polytechnic University.

Cheung, S. H.-N. (1972). 香港粵語語法的研究 [Cantonese as spoken in Hong Kong.] Hong Kong Chinese University.

Chung, K. S. (2004). 'Mandarin compound verbs.' Ph.D dissertation. Department of Comparative Linguistics, Leiden University.

Chung, K. S. (2004). 'Mandarin Compound Verbs.' Ph.D. dissertation. Leiden University.

Crystal, D. (1992). *An Encyclopedic Dictionary of Language and Languages*. Blackwell.

Crystal, D. (1997). *Dictionary of Linguistics and Phonetics*. Blackwell.

Dai, J. X. (1992). 'Chinese Morphology and its Interface with Syntax.' Ph.D dissertation. Department of Linguistics, Ohio State University.

Dreher, J. and Lee, P. (1966). 'Instrumental investigation of single and paired Mandarin tonemes.' Douglas Advanced Research Laboratory.

Duanmu, S. (1992). 'End-based theory, cyclic stress, and tonal domains.' In: *Paper presented at CLS 28 Parasession on the Cycle*, April 23 - 25, Chicago.

Duanmu, S. (1993). 'Rime length, stress, and association domains'. *Journal of East Asian* Linguistics 2, 1-44.

Duanmu, S. (1995). 'Metrical and tonal phonology of compounds in two Chinese dialects.' *Language* 71.2, 225-259.

Duanmu, S. (1998). 'Wordhood in Chinese.' In: *New Approaches to Chinese Word Formation : Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*, Packard, J. L. (ed.). Berlin; New York: Mouton de Gruyter.

Duanmu, S. (1999). 'Metrical structure and tone: evidence from Mandarin and Shanghai.' *Journal of East* Asian *Linguistics* 8, 1-38.

Fowler, C. & Housum, J. (1987). 'Talkers' signaling of 'new' and 'old' words in speech and listeners' perception and use of the distinction." *Journal of Memory and Language* 26: 489-504.

Gao, J.; Li, M.; Wu, A.; Huang, C-N. (2004). 'Chinese word segmentation: a pragmatic approach.' Microsoft Research Technical Report, MSR-TR-2004-123.

GB/T 13715-92 (1993). Contemporary Chinese language word-segmentation for information processing. Technical report, Beijing.

Gregory, M. (2001). 'Linguistic informativeness and speech production: an investigation of contextual and discourse-pragmatic effects on phonological variation'. Ph.D. dissertation. University of Colarado.

Hashimoto, O.-K. Y. (1972). *Phonology of Cantonese*. Cambridge, England: Cambridge University Press.

Hay, J. (2002). 'From Speech Perception to Morphology: Affix Ordering Revisited.' *Language*. Vol. 78, No. 3, September 2002, pp. 527-555.

Hong Kong Supplementary Character Set (HKSCS). <http://www.info.gov.hk/digital21/eng/hkscs/introduction.html>

Hsu, H. (2005). 'An Optimality-Theoretic Analysis of Syllable Contraction in Cantonese.' *Journal of* Chinese *Linguistics*. Vol.33, No.1 (67).

Jin, S. (1986). *Shanghai Morphotonemics: A Preliminary Study of Tone Sandhi Behavior Across Word Boundaries*. M. A. thesis, University of Pittsburgh. Bloomington, Indiana: Indiana University Linguistics Club. Reproduced by Indiana University Linguistics Club, Bloomington, Indiana.

Jurafsky, D.; Bell, A.; Gregory M.; Raymond, W. (2001). 'Probabilistic relations between words: Evidence from reduction in lexical production.' In: *Frequency and the Emergence of Llinguistic Structure*. Edited by Joan Bybee and Paul Hopper. Amsterdam: John Benjamins.

Kucera, H. & Francis, W. N. (1967). *Computational Analysis of Present Day American English*. Providence RI: Brown University Press.

Law, S.-P. (1990). 'The syntax and phonology of Cantonese sentence-final particles.' Ph.D dissertation. Boston University.

Law, S.-P., Fung, S.-Y. and Bauer, R. S. (2001). 'Perception and production of Cantonese consonant endings', *Asia Pacific Journal of Speech, Language and Hearing*, 6:179-195.

Lee, K. M. G. (2003). 'Syllable fusion in Cantonese connected speech.' M.Phil. thesis, University of Hong Kong.

Lee, W.-S. (1999). 'A Phonetic Study of the Speech of the Cantonese-Speaking Children in Hong Kong.' Ph.D. dissertation, City University of Hong Kong.

Leung, M.-T. and Law, S.-P. (2001). 'HKCAC: The Hong Kong Cantonese adult language corpus.' *International Journal of Corpus Linguistics* 6, 305-325. (Corpus data available at < http://shs.hku.hk/corpus/index.htm>.)

Li, C. N. & Thompson, S. A. (1981). *Mandarin Chinese: A Functional Reference Grammar.* University of California Press.

Li, P. Y.-C. (1986). 'Contraction in Cantonese: A first probe.' Manuscript. University of California, San Diego.

Lin, M., Yan, J. and Sun, G. (1984). 北京話兩字組正常重音的初步實驗 [A preliminary experimental study on the normal stress in the Beijing dialect]. *Fangyan* 1, 57-73.

Lin, T. (1985). 探討北京話輕聲性質的初步實驗 [Preliminary experiments in the exploration of the nature of Mandarin neutral tone]. 北京語音實驗錄 [Peking University Working Papers on Phonetics], 1-26. Beijing University Press, Beijing.

Matthews, S. and Yip, V. (1994). *Cantonese: A Comprehensive Grammar*. London: Routledge.

Qiao, Y. (1966). 廣州話口語詞的研究 [A Study of Spoken Cantonese]. 香港華僑語文出版社 Huaqiao Yuwen Chubanshe.

Rao, B.-C., Ouyang, J.-Y. and Zhou, W.-J. (eds.) ([1981] 1991). 廣州話方言詞典 [Dictionary of the Guangzhou Dialect]. Hong Kong: The Commercial Press.

Raymond, W. D.; Dautricourt, R.; Hume, E. (submitted). 'Word-medial /t, d/ deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors.' *Language Variation and Change*.

Reitveld, T. and van Hout, R. (1993). *Statistical Techniques for the Study of Language and Language Behavior*. Mouton de Gruyter, Berlin.

Riha, H. (2006). 'The Morphology and Semantics of Roman Letter Words in Mandarin Chinese'. M.A. thesis, Department of Linguistics, Ohio State University.

ROCLING Segmentation Principle for Chinese Language Processing (1997). <http://godel.iis.sinica.edu.tw/ROCLING/juhuashu1.htm>

Sankoff, D. (ed.) (1978). *Linguistic Variation: Models & Methods*. Academic Press, Incorporated, New York.

Selkirk, E. O. and Shen, T. (1990). 'Prosodic domains in Shanghai Chinese.' In: *The Phonology-Syntax Connection*, Inkelas, S. and Zec, D. (eds.). Chicago: University of Chicago Press.

Shen, X. S. (1990). *The Prosody of Mandarin Chinese*. University of California Press.

Shih, C. (1997). 'Mandarin third tone sandhi and prosodic structure.' In: *Studies in Chinese Phonology*. Wang, J. and Smith, N. (eds.). Mouton de Gruyter.

Sproat, R; Shih, C; Gale, W.; Chang, N. (1996). 'A stochastic finite-state word-segmentation algorithm for Chinese.' *Computational Linguistics.* 22(2):377-404.

Trask, R. L. (1993). *A Dictionary of Grammatical Terms in Linguistics*. Routledge.

Trask, R. L. (1996). *A Dictionary of Phonetics and Phonology*. Routledge.

Tseng, C. (1981). 'An Acoustic Phonetic Study on Tones in Mandarin Chinese.' Ph.D. dissertation, Brown University.

Uchimoto, K.; Nobata, C.; Yamada, A.; Sekine, S.; Isahara, H. (2003) 'Morphological analysis of the corpus of spontaneous Japanese.' Workshop on Spontaneous Speech Processing and Recognition.

Wang, J. and Wang, L. (1993). 普通話多音節詞音節時長分布模式 [The types of relative lengths of syllables in polysyllabic words in Putonghua]. 中國語文 *Zhongguo Yuwen*, 2, 112-116.

Wang, W. S.-Y. and Li, K.-P. (1967). 'Tone 3 in Pekinese'. *Journal of Speech and Hearing Research* 10, 629-636.

Wong, S.-L. ([1941], 1997). *A Chinese Syllabary Pronounced According to the Dialect of Canton*. (Hong Kong: Chung Hwa Book Co. Ltd.).

Wong, W. Y. P. (1996). 'Tempo, processing rate and clarity drive in Hong Kong Cantonese connected speech.' M.A. thesis, Hong Kong Polytechnic University.

Wong, W. Y. P. (2002). 'Cues and constraints: Production and perception of coda place in modern Hong Kong Cantonese.' First Pre-Candidacy paper. Department of Linguistics, The Ohio State University.

Wong, W. Y. P. (2004). 'Syllable fusion and speech rate in Hong Kong Cantonese'. In: Proceedings *of Speech Prosody 2004*, 255-258. 22-26 March, Nara, Japan.

Wong, W. Y. P., Brew, C., Chan, S.-D. and Beckman, M. E. (2002). 'Using the Segmentation Corpus to define an inventory of concatenative units for Cantonese speech synthesis.' In: *Proceedings of the First SIGHAN Workshop on Chinese Language Processing in conjunction with COLING 2002*. Pp. 119-123. (Papers presented at the First SIGHAN Workshop on Chinese Language Processing, Academia Sinica, Taipei, Taiwan, 1 September 2002.) T'sou, B. K., Kwong, O. O. Y. and Lai, T. B. Y. (eds.).

Wong, W. Y. P., Chan, M. K.-M. and Beckman, M. E. (2005). 'An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese'. In: *Prosodic Typology - The Phonology of Intonation and Phrasing*. (Collection of papers from the 14th International Congress of Phonetic Sciences 1999 Satellite Workshop on "Intonation: Models and ToBI Labelling", San Francisco, CA. 1-7 August 1999.) Edited by Sun-Ah Jun. Oxford University Press.

Wright, M. S. (1983). *A metrical approach to tone sandhi in Chinese dialects*. Ph.D. Dissertation. Univ. of Mass. Ann Arbor, Michigan: University Microfilms International.

Wu, A. (2003). 'Customizable segmentation of morphologically derived words in Chinese.' *Computational Linguistics and Chinese Language Processing*, vol. 8, no. 1, February 2003, pp. 1-28.

Xia, F. (2000). The Segmentation Guidelines for the Penn Treebank (3.0). Technical report. University of Pennsylvania. <http://www.cis.upenn.edu/~chinese/>

*Xiandai Hanyu Cidian*現代漢語詞典 [Modern Chinese Dictionary] (1989). Compiled by the Chinese Academy of Social Sciences Institute of Linguistics. The Commercial Press.

Yan, J. and Lin, M. (1988). 北京話三字組重音的聲學表現 [Acoustic characteristics of the stress in Beijing trisyllables]. Fangyan 3, 227-237.

Yeung, S.-W. H. (1980). 'Some aspects of phonological variations in the Cantonese spoken in Hong Kong'. M.A. thesis., University of Hong Kong,

Yip, M. (1980). *The Tonal Phonology of Chines*e. Ph.D dissertation, MIT.

Yu, S. (1999). Guidelines for the annotation of contemporary Chinese texts: word segmentation and POS-tagging. Institute of Computational Linguistics, Beijing University, Beijing.

Yuan, J.-H. et al. ([1960] 1983). 漢語方言概要 [Hanyu Fangyan Gaiyao]. Wenzi Gaige Chubanshe.

Zee, E. (1995), 'Temporal organisation of syllable production in Cantonese Chinese', *Proceedings of ICPhS 95, Stockholm*.

Zee, E. (1999a), 'An acoustical analysis of the diphthongs in Cantonese', *Proceedings of ICPhS 99, San Francisco*.

Zee, E. (1999b), 'Resonance Frequency and Vowel Transcription in Cantonese', *Proceedings of the 10th North American Conference of Chinese Linguistics and the 7th Annual Meeting of the International Association of Chinese Linguistics*.

Zee, E. (2001). 'The sound pattern of Shanghai.' In: *The Sino-Tibetan Languages*, Thrugood, G. and LaPolla, R. (eds.).

Zee, E. and Maddieson, I. (1979).'Tones and tone sandhi in Shanghai: phonetic evidence and phonological analysis.' *UCLA Working Papers in Phonetics* 45, 93-129.

Zhongguo Chengyu Dacidian 中國成語大辭典 [Dictionary of Chinese Idioms] (1987). Edited by T. Wang et al. 上海辭書出版社. Shanghai Cishu Shubanshe.

Zhu, D. (1982). *Yufa Jiangyi* 語法講義. The Commercial Press (Beijing).

APPENDIX A

PSEUDO-SENTENCES AND PRACTICE TRIALS FOR EXPERIMENT I

Tape A

Sequence of tempi requested after each utterance was heard:
好快 'very fast' – 2-beat pause – *正常 'normal'* – 4-beat pause

Practice trials

梗係	直情	等一陣
kɐŋ^{35}hɐi^{22}	tsek^{6}ts^{h}eŋ^{21}	tɐŋ^{35}jɐt^{5}tsɐn^{22}
of course	must be	wait a second
如果	但係	大學生
jyː^{21}k^{w}ɔː^{35}	taːn^{22}hɐi^{22}	taːi^{22}hɔːk^{2}saːŋ^{55}
if	but	university students
成日	譬如	細路仔
sɛːŋ^{21}jɐt^{2}	p^{h}ei^{33}jyː^{21}	sɐi^{33}lou^{22}tsɐi^{35}
always	for example	children

Test trials

1.	聽日	間尺*	乒乓波
	t^{h}eŋ^{55}jɐt^{2}	kaːn^{33}ts^{h}ɛːk^{3}	peŋ^{55}pɐm^{55}pɔː^{55}
	tomorrow	ruler	table-tennis
2.	嗰個	但係	大帽山*
	kɔː^{35}kɔː^{33}	taːn^{22}hɐi^{22}	taːi^{22}mou^{22}saːn^{55}
	that CLASS.	but	Tai Mo Shan
3.	不過	書包*	朝頭早
	pɐt^{5}k^{w}ɔː^{33}	syː^{55}paːu^{55}	tsiːu^{55}t^{h}ɐu^{21}tsou^{35}
	but	school bag	morning
4.	零食*	啱啱	收音機
	leŋ^{21}sek^{2}	aːm^{55}aːm^{55}	sɐu^{55}jɐm^{55}kei^{55}
	snacks	just	radio
5.	住喺	睇吓	返屋企
	tsyː^{22}hɐi^{35}	t^{h}ɐi^{35}haː^{23}	faːn^{55}ok^{5}k^{h}ei^{35}

	live in	take a look	go home
6.	記得	辣椒*	大笨象*
	kei^{33}tɐk^{5}	laːt^{2}tsiːu^{55}	taːi^{22}pɐn^{22}tsœːŋ^{22}
	remember	chilli	elephant
7.	鋼琴*	淨係	星期日
	kɔːŋ^{33}k^{h}ɐm^{21}	tseŋ^{22}hɐi^{22}	seŋ^{55}k^{h}ei^{21}jɐt^{2}
	piano	only	Sunday

(A 10-beat pause inserted here)

Sequence of tempi requested after each utterance was heard was reversed:
正常 'normal' – 4-beat pause – *好快 'very fast'* – 2-beat pause

Practice trials

貼紙	打掃	鉛筆刨
t^{h}iːp^{3}tsiː^{35}	taː^{35}sou^{33}	jyːn^{21}pɐt^{5}p^{h}aːu^{35}
stickers	to clean up	pencil sharpener
電腦	飲茶	洗衣機
tiːn^{22}lou^{23}	jɐm^{35}ts^{h}aː^{21}	sɐi^{35}jiː^{55}kei^{55}
computer	drink tea	washing machine
以前	雪糕	牛頭角
jiː^{23}ts^{h}iːn^{21}	syːt^{3}kou^{55}	ŋɐu^{21}t^{h}ɐu^{21}kɔːk^{3}
in the past	ice-cream	Ngau Tau Kok

Test trials

8.	鍾意	即刻	奧地利*
	tsoŋ^{55}jiː^{33}	tsek^{5}hɐk^{5}	ou^{33}tei^{22}lei^{22}
	to like	immediate	Austria
9.	粉筆*	佢話	嗰陣時
	fɐn^{35}pɐt^{5}	k^{h}ɵy^{23}waː^{22}	kɔː^{35}tsɐn^{22}siː^{21}
	chalk	s/he said	at that time
10.	邊個	真係	洛杉磯*
	piːn^{55}kɔː^{33}	tsɐn^{55}hɐi^{22}	lɔːk^{2}ts^{h}aːm^{33}kei^{55}
	who	really	Los Angeles

11.	客人	出去	四呀[22]歲
	haːk^{3}jɐn^{21}	tsʰɵt^{5}hɵy^{33}	sei^{33}aː^{22}sɵy^{33}
	guest	go out	forty years old
12.	而家	奶茶*	一陣間
	jiː^{21}kaː^{55}	laːi^{23}tsʰaː^{21}	jɐt^{5}tsɐn^{22}kaːn^{55}
	now	milk tea	in a moment
13.	點心*	就要	的士站*
	tiːm^{35}sɐm^{55}	tsɐu^{22}jiːu^{33}	tek^{5}siː^{35}tsaːm^{22}
	dim sum	have to	taxi stop
14.	知道	膠袋*	超級市場
	tsiː^{55}tou^{33}	kaːu^{55}tɔːi^{35}	tsʰiːu^{55}kʰɐp^{5}siː^{23}tsʰœːŋ^{21}
	know	plastic bag	supermarket
15.	就嚟	風扇*	司徒拔道
	tsɐu^{22}lei^{21}	foŋ^{55}siːn^{33}	siː^{55}tʰou^{21}pɐt^{2}tou^{22}
	about to	fan	Stubbs Road

Notes:
1. Asterisks are filler short or long words that were not in Cheung (1986:242-9) and his subsequent collection.
2. Transcribed are the non-fused forms of the syllables.
3. Underlined are the syllables where bisyllabic fusion was expected to occur.
4. English translations are provided here, but not on the test tapes.
5. Words in italics are for illustration here, but were not on the test tapes.
6. 正常 'normal' and 好快 'very fast' in parentheses after each utterance were aural prompts recorded on the test tapes.
7. Numbering each utterance is for the sake of easy identification here. The recorded utterances were not aurally numbered.

APPENDIX B

SIMPLE ARITHMETIC

Subjects were to perform the calculation task below while repeating the utterances.

FOR CALCULATION

4 + 5 =	10 + 9 =	6 + 3 =	25 + 3 =
6 + 5 =	61 + 7 =	36 + 3 =	63 + 2 =
8 + 1=	6 + 4 =	6 + 21 =	2 + 54 =
12 + 7 =	20 + 8 =	32 + 2 =	1 + 8 =
2 + 16=	11 + 3 =	18 + 1 =	13 + 6 =
5 + 4 =	22 + 4 =	5 + 5 =	50 + 2 =
2+ 8 =	71 + 3 =	12 + 4 =	17 + 2 =
32 + 7 =	23 + 4 =	50 + 2 =	70 + 9 =
61+ 8 =	31 + 5 =	9 +10 =	31 + 5 =
21 + 6 =	52 + 9 =	32 + 3 =	71 + 2 =
8 + 4 =	6 + 3 =	41 + 7 =	31 + 7 =
20+ 8 =	2 + 7 =	33 + 4 =	11 + 8 =
50+ 9 =	30 + 5 =	25 + 2 =	0 + 6 =
1+ 9 =	12 + 7 =	26 + 1 =	6 + 41 =
40+ 7 =	4 + 3 =	51 + 4 =	26 + 3 =
24+ 4 =	12 + 5 =	8 + 11 =	11 + 6 =
73 + 6 =	21 + 4 =	12 + 7 =	22 + 6 =
82 + 6 =	3 + 5 =	12 + 3 =	12 + 3 =
12+ 5 =	9 + 3 =	10 + 9 =	82 + 5 =
22+ 7 =	82 + 2 =	16 + 2 =	97 + 1 =
3 + 7 =	30 + 6 =	2 + 8 =	11 + 7 =
6 + 11 =	2 + 3 =	31 + 7 =	13 + 5 =
5 + 3 =	41 + 5 =	9 + 1 =	35 + 3 =
8 + 2 =	17 + 2 =	6 + 13 =	60 + 8 =
40 + 9 =	5 + 5 =	4 + 13 =	7 + 12 =
62 + 6 =	8 + 1 =	41 + 5 =	9 + 10 =
13 + 5 =	13 + 5 =	13 + 1 =	6 + 3 =
31 + 4 =	3 + 3 =	40 + 6 =	8 + 2 =
6 + 12 =	15 + 2 =	15 + 3 =	7 + 12 =
12 + 7 =	70 + 3 =	22 + 4 =	51 + 4 =
6 + 3 =	3 + 6 =	6 + 23 =	26 + 4 =
3 + 5 =	3 + 5 =	12 + 7 =	78 + 1 =
4 + 10 =	6 + 10 =	11 + 3 =	33 + 6 =
12 + 7 =	8 + 2 =	6 + 4 =	14 + 5 =
11 + 8 =	3 + 31 =	31 + 2 =	47 + 2 =
6 + 30 =	9 + 0 =	35 + 4 =	2 + 5 =
51 + 2 =	8 + 1 =	61 + 3 =	21 + 5 =
6 + 11 =	13 + 6 =	70 + 9 =	44 + 3 =
90 + 7 =	8 + 1 =	71 + 3 =	14 + 3 =
1 + 8 =	60+ 9 =	8 + 41 =	90 + 2 =
0 + 6 =	7 + 31 =	2 + 7 =	3 + 11 =
13 + 5 =	13 + 6 =	8 + 13 =	6 + 10 =
71 + 6 =	53 + 4 =	11 + 7 =	2 + 8 =
10 + 6 =	21 + 4 =	6 + 13 =	6 + 71 =

82 + 6 =	3 + 5 =	12 + 3 =	12 + 3 =
12+ 5 =	9 + 3 =	10 + 9 =	82 + 5 =
8 + 2 =	17 + 2 =	6 + 13 =	60 + 8 =
40 + 9 =	5 + 5 =	4 + 13 =	7 + 12 =
62 + 6 =	8 + 1 =	41 + 5 =	9 + 10 =
13 + 5 =	13 + 5 =	13 + 1 =	6 + 3 =
31 + 4 =	3 + 3 =	40 + 6 =	8 + 2 =
6 + 12 =	15 + 2 =	15 + 3 =	7 + 12 =
12 + 7 =	70 + 3 =	22 + 4 =	51 + 4 =
6 + 3 =	3 + 6 =	6 + 23 =	26 + 4 =
3 + 5 =	3 + 5 =	12 + 7 =	78 + 1 =
4 + 10 =	6 + 10 =	11 + 3 =	33 + 6 =
12 + 7 =	8 + 2 =	6 + 4 =	14 + 5 =
11 + 8 =	3 + 31 =	31 + 2 =	47 + 2 =
6 + 30 =	9 + 0 =	35 + 4 =	2 + 5 =
51 + 2 =	8 + 1 =	61 + 3 =	21 + 5 =
6 + 11 =	13 + 6 =	70 + 9 =	44 + 3 =
90 + 7 =	8 + 1 =	71 + 3 =	14 + 3 =
1 + 8 =	60+ 9 =	8 + 41 =	90 + 2 =
0 + 6 =	7 + 31 =	2 + 7 =	3 + 11 =
13 + 5 =	13 + 6 =	8 + 13 =	6 + 10 =
71 + 6 =	53 + 4 =	11 + 7 =	2 + 8 =
10 + 6 =	21 + 4 =	6 + 13 =	6 + 71 =
4 + 5 =	10 + 9 =	6 + 3 =	25 + 3 =
6 + 5 =	61 + 7 =	36 + 3 =	63 + 2 =
8 + 1=	6 + 4 =	6 + 21 =	2 + 54 =
12 + 7 =	20 + 8 =	32 + 2 =	1 + 8 =
2 + 16=	11 + 3 =	18 + 1 =	13 + 6 =
5 + 4 =	22 + 4 =	5 + 5 =	50 + 2 =
2+ 8 =	71 + 3 =	12 + 4 =	17 + 2 =
32 + 7 =	23 + 4 =	50 + 2 =	70 + 9 =
61+ 8 =	31 + 5 =	9 +10 =	31 + 5 =
21 + 6 =	52 + 9 =	32 + 3 =	71 + 2 =
8 + 4 =	6 + 3 =	41 + 7 =	31 + 7 =
20+ 8 =	2 + 7 =	33 + 4 =	11 + 8 =
50+ 9 =	30 + 5 =	25 + 2 =	0 + 6 =
1+ 9 =	12 + 7 =	26 + 1 =	6 + 41 =
40+ 7 =	4 + 3 =	51 + 4 =	26 + 3 =
24+ 4 =	12 + 5 =	8 + 11 =	11 + 6 =
73 + 6 =	21 + 4 =	12 + 7 =	22 + 6 =
3 + 7 =	30 + 6 =	2 + 8 =	11 + 7 =
6 + 11 =	2 + 3 =	31 + 7 =	13 + 5 =
5 + 3 =	41 + 5 =	9 + 1 =	35 + 3 =

APPENDIX C

DETAILS OF DATA STRUCTURE IN THE ORIGINAL HKCAC DATA FILES AND PROCESSING OF THE FILES FOR WORD SEGMENTATION

There were 56 original data files in the HKCAC. The files were in Excel format. Each of the files contains pairs of lines: a line of orthographic transcription followed by a line of phonetic transcription of the recorded radio programs. Each pair of line in the data files corresponds roughly to an utterance, and each line begins with codes that represent information about the speakers. Figure C.1(a) shows the first 14 utterances (i.e. 28 lines) of two of the original Excel data files.

	A	B	C	D	E	F	G	**H**	I	J	K	L	M	N	O	P
1	626	1	O	M	H2	喂	**hello**									
2	626	1	P	M	H2	wɐi˥	hələʊ									
3	626	2	O	F	H1	喂										
4	626	2	P	F	H1	wɐi˥										
5	626	3	O	F	C	喂										
6	626	3	P	F	C	wɐi˥										
7	626	4	O	M	H2	[	係									
8	626	4	P	M	H2	[	hɐi˧	]								
9	626	5	O	F	H1	[	係	]	你	係						
10	626	5	P	F	H1	[	hɐi˧	]	lei˩	hɐi˧						
11	626	6	O	F	C	**ɛ˧**	Sandy									
12	626	6	P	F	C	ɛ˧	sænd ɪ									
13	626	7	O	F	H1	Sandy										
14	626	7	P	F	H1	sænd ɪ										
15	626	8	O	F	C	係	呀									
16	626	8	P	F	C	hɐi˧	ja˧									
17	626	9	O	F	H1	係	你	[	好	]						
18	626	9	P	F	H1	hɐi˧	lei˩	[	hou˥	]						
19	626	10	O	M	H2	[	O.K.	]								
20	626	10	P	M	H2	[	okei	]								
21	626	11	O	F	C	係	你	好								

First row is column code in Excel.

Transcription given in English when speaker code-switched into English. The corresponding phonetic transcription was not given in the Cantonese segment-plus-tone configuration, indicating that the HKCAC transcriber interpreted the speaker to be using an English phonetic form, rather than the Cantonese form for the English loanword sometimes represented as 哈佬 or 蝦佬 /ha:^{55}lou^{35}/ 'hello'.

[ ] denotes overlapping speech between two people in the HKCAC transcription

Transcription is given in IPA for the ground-holding morpheme, which does not have an orthographic representation.

First column is row number in Excel.

Figure C.1 Examples of the original HKCAC Excel data, the processing of the original data files, and the subsequent steps in the word segmentation exercise in this dissertation.

Figure C.1 continued

(i) *Ghost-26g*

22	626	11	P	F	C	hɐi˧	lei˨˩	hou˥								
23	626	12	O	M	H2	Sandy	有	[	咩	講	]					
24	626	12	P	M	H2	sænd ɪ	jɐu˨˩	[	**mɛ** ˧˩	kɔŋ˥	]					
25	626	13	O	F	C	[	ɛ˧	]	我	就	想	講	我	個	朋	友
26	626	13	P	F	C	[	ɛ˧	]	ɔ˨˩	tsɐu ˧	sœŋ ˥	kɔŋ ˥	ŋɔ ˨˩	kɔ ˧	p^hɐn ˩	jɐu ˨˩
27	626	14	O	F	C	佢	個	妹	呢	做	嗰	啲	ɛ˧	帶	街	㗎
28	626	14	P	F	C	k^høy˨˩	ɔ˧	mui˥	lɛ˥˧	tsou ˧	kɔ˥	ti˥˧	ɛ˧	tai ˧	kai˥˧	ka˧

A fused form for 乜嘢 /mɐt^{5}jɛː^{23}/ ‘what’.

Figure C.1 continued

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	**627**	1	O	M	H 2	喂	hello	係	你	點	稱	呼					
2	627	1	P	M	H 2	wei ˥	hələʊ	hɐi˧	lei˨˩˦	tim ˥	ts^{h}iŋ ˥˧	fu ˥˧					
3	627	**2**	O	F	H 1	喂	恐	怖	熱	線							
4	627	2	P	F	H 1	wei ˥	hoŋ˥˩	pou ˧	jit˧	sin ˧							
5	627	3	O	M	H 2	喂											
6	627	3	P	M	H 2	wei ˥											
7	627	4	**O**	F	C	ɛ˧	潘	迢	聰	露	芙	[	係	咪	]	呀	
8	627	4	**P**	F	C	ɛ˧	p^{h}un ˥˧	siu˩˧	ts^{h}oŋ ˥˧	lou ˧	fu˩˧	[	ɐi˧	**mɐi** ˧	]	a˧	
9	627	5	O	M	H 2	[	Hello	]									
10	627	5	P	**M**	H 2	[	hələʊ	]									
11	627	6	O	F	H 1	[	係	]									
12	627	6	P	F	**H 1**	[	hɐi˧	]									
13	627	7	O	M	H 2	[	係	]									
14	627	7	P	M	H 2	[	hɐi˧	]									
15	627	8	O	F	H 1	[	你	好	]								
16	627	8	P	F	H 1	[	lei˨˩˦	hou ˥	]								
17	627	9	O	F	C	呀	你	好	ɛ˧	我	叫	阿	芝	呀			
18	627	9	P	F	C	a˧	lei˨˩˦	hou ˥	ɛ˧	ɔ˨˩˦	kiu˧	a˧	tsi ˥˧	ja˧			

Code for the radio program and the current file.

Utterance number

A fused form for 唔係 /$m̩^{21}$hɐi^{22}/ ‘no, not’.

A line for ***o***rthographic transcription

A line for ***p***honetic transcription

Gender of speaker: ***m***ale or ***f***emale

Role of speaker: ***h***ost, ***g***uest or ***c***aller

Figure C.1 continued

(ii) *Ghost-27g*

19	627	10	O	F	H1	阿	[	芝	]	係	嘛						
20	627	10	P	F	H1	a˧	[	tsi˥	]	ɐi˧	**ma**˩˧						
21	627	11	O	M	H2	[	係	]									
22	627	11	P	M	H2	[	hɐi˨	]									
23	627	12	O	F	C	係	呀	係	[	呀	我	]	想	講	嘅	嘢	呢
24	627	12	P	F	C	hɐi˧	ja˧	hɐi˨	[	ja˧	ɔ˩˧	]	sœŋ˥	kɔŋ˥	kɛ˧	jɛ˩˧	lɛ˥
25	627	13	O	F	H1	[	係	你	好	]							
26	627	13	P	F	H1	[	ɐi˧	lei˩˧	hou˥	]							
27	627	14	O	F	C	關	於	係	長	洲	東	堤	㗎				
28	627	14	P	F	C	kwan˥	jy˥	hɐi˨	tsʰœŋ˩˨	tsɐu˥	toŋ˥	tʰɐi˩˨	ə˧				

Two possible analyses of this form 嘛 [maː^{23}]: (i) a fused form whose "underlying" form could be 唔係呀 /m̩^{21}hɐi^{22} aː^{33}/ 'negation-sentence particle', or 咪呀 /mɐi^{22} aː^{33}/ 'negation (fused form)-sentence particle'; (ii) a lexicalized question particle.

In each original Excel data file, the first column contains the code for the radio program and the file. The second column is the utterance number. The third column indicates whether the line gives an orthographic or a phonetic transcription of the utterance. The fourth column shows whether the speaker is male (M) or female (F). The fifth column states whether the speaker is a host (H), a guest (G), or a caller (C) in the program. Orthographic and phonetic transcriptions start from the sixth column. The orthographic transcription of the HKCAC uses Chinese characters that can be found in the Big-5 traditional Chinese character set and in the Hong Kong Supplementary Character Set (HKSCS). The HKSCS has two versions: the ISO 10646 version and the Big-5 version. Either version is good for viewing the HKCAC data files. Both versions are freely downloadable at <http://www.info.gov.hk/digital21/eng/hkscs/download.html>. Phonetic transcription of the corpus is available in IPA. The IPA fonts are downloadable at <http://shs.hku.hk/corpus/index.htm>.

The orthographic and the phonetic transcriptions in each of the original HKCAC data files were split and saved in separate files, yielding 56 orthographic transcription files and the corresponding 56 phonetic transcription files (see Figure C.2).

(i) Orthographic transcription files in plain text format

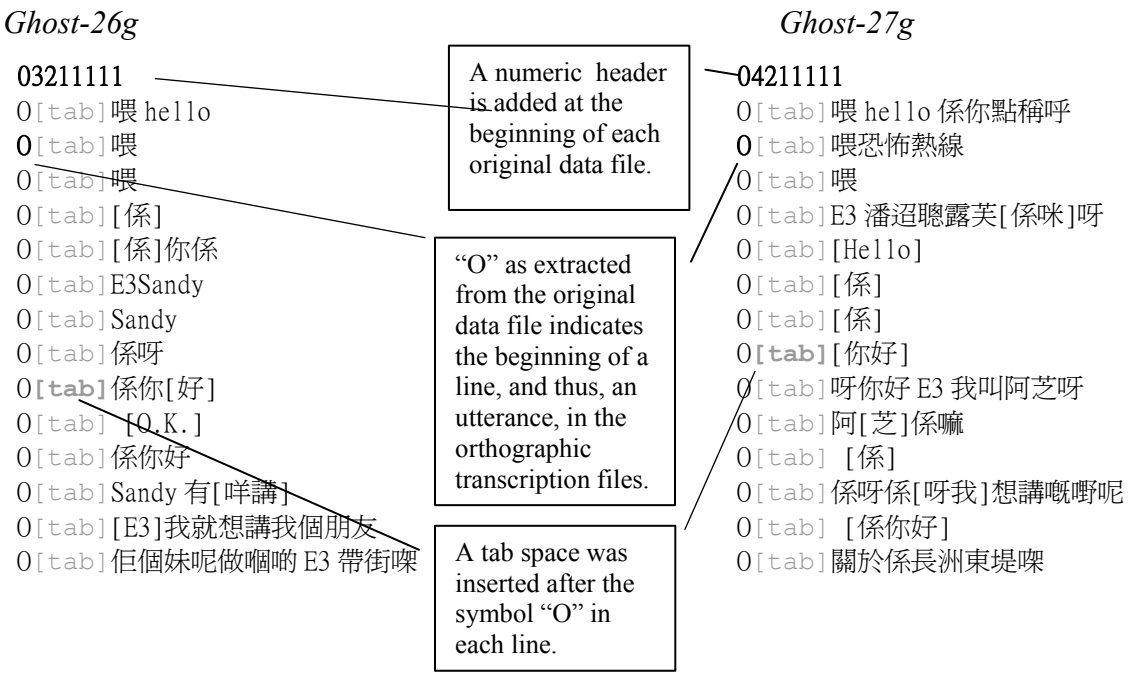


Figure C.2 In the processing of the original HKCAC, each original Excel data file was split to make two files: an orthographic transcription file and the corresponding phonetic transcription file. Each of these files retains only information that is relevant for this project. The files were then saved in plain text format.

Figure C.2 continued

(ii) Phonetic transcription files in plain text format

Ghost-26g

```
03211111
P     wAi2  hGlGU
P     wAi2
P     wAi2
P     [     hAi6  ]
P     [     hAi6  ]
      lei5  hAi6
P     E3    sYndI
P     sYndI
P     hAi6  ja3
P     hAi6  lei5  [
      hou2  ]
P     [     okei  ]
P     hAi6  lei5  hou2
P     sYndI jAu5  [     mE8
      kON2  ]
P     [     E3    ]     O5
      tsAu6 sJN2  kON2  NO5
      kO3   pHAn4 jAu5
P     kH0y5 O3    mui2  lE1
      tsou6 kO2   ti1   E3
      tai3  kai1  ka3
```

A numeric header corresponding to that in the orthographic transcription file was added at the beginning of each original data file.

Tab spaces were inserted corresponding to the boundary of each cell in the original Excel files.

With the conversion from the original Excel format into plain text format, the IPA symbols in the original Excel files were converted into ASCII forms and tone marks became tone numbers 1-6 and 8, where 8 denotes a fall-rise tone as transcribed by the HKCAC transcribers.

"P" as extracted from the original data file indicates the beginning of a line, and thus, an utterance, in the phonetic transcription files.

Ghost-27g

```
04211111
P     wei2  hGlGU hAi6
      lei5  tim2
      tsHiN1      fu1
P     wei2  hoN2  pou3
      jit6  sin3
P     wei2
P     E3    pHun1 siu4
      tsHoN1      lou6
      fu4   [     Ai6
      mAi6  ]     a3
P     [     hGlGU ]
P     [     hAi6  ]
P     [     hAi6  ]
P     [     lei5  hou2
      ]
P     a3    lei5  hou2
      E3    O5    kiu3
      a3    tsi1  ja3
P     a3    [     tsi1
      ]     Ai6   ma5
P     [     hAi6  ]
P     hAi6  ja3   hAi6
      [     ja3   O5
      ]     sJN2  kON2
      kE3   jE5   lE1
P     [     Ai6   lei5
      hou2  ]
P     kwan1 jy1   hAi6
      tsHJN4      tsAu1
      toN1  tHAi4 G3
```

Figure C.2 ended

In the orthographic transcription files, Chinese characters and English words that sometime intermingle with the Chinese text need not be space delimited, although English words that occur in a row need to be space delimited. In the phonetic transcription files, cells in the original data files are reflected by tab spaces. Lines of transcription in the orthographic and the phonetic transcription files begin with the symbol “O” (for orthography) or “P” (for phonetic) followed by a tab space. The inclusion of the two symbols was to help visualize the beginning (and the end) of a line (i.e. an utterance according to Law & Leung 2001). (Being able to visualize the beginning of lines was useful when I re-segmented the words after the Segmentation Corpus word segmentation program was run.) File names of the texts need to consist of eight digits. All the files were renamed with 8-digit codes (i.e. [01-56]211111), which were also added as headers at the beginning of the files. The correspondence between the original file names and the 8-digit codes is given in Table C.1.

HKCAC original file names	New file names in this study	HKCAC original file names	New file names in this study
Finance-21g	012111111	ping-55g	292111111
Ghost-25g	022111111	ping-56g	302111111
Ghost-26g	032111111	sky-42g	312111111
Ghost-27g	042111111	sky-43g	322111111
Ghost-28g	052111111	sky-44g	332111111
Ghost-45g	062111111	sky-63g	342111111
Ghost-46g	072111111	sky-64g	352111111
Ghost-47g	082111111	sky-65g	362111111
Ghost-48g	092111111	sky-66g	372111111
ngaan-8g	102111111	sky-67g	382111111
ngaan-9g	112111111	sky-69g	392111111
ngaan-10g	122111111	Titanic-11g	402111111
ngaan-29g	132111111	Titanic-12g	412111111
ngaan-30g	142111111	Titanic-13g	422111111
ngaan-31g	152111111	Titanic-14g	432111111
ping-1g	162111111	Titanic-15g	442111111
ping-2g	172111111	Titanic-32g	452111111
ping-3g	182111111	Titanic-33g	462111111
ping-4g	192111111	Titanic-34g	472111111
ping-5g	202111111	Titanic-35g	482111111
ping-6g	212111111	Titanic-36g	492111111
ping-7g	222111111	Titanic-37g	502111111
ping-49g	232111111	Titanic-38g	512111111
ping-50g	242111111	Titanic-39g	522111111
ping-51g	252111111	Titanic-40g	532111111
ping-52g	262111111	Titanic-41g	542111111
ping-53g	272111111	Zing-16g	552111111
ping-54g	282111111	zing-57g	562111111

Table C.1 The HKCAC file names were renamed for the convenience of running the Segmentation Corpus word segmentation algorithm over the HKCAC texts. There are 56 HKCAC original data files in total.

The files were stored in two different folders, namely, the orthographic transcription folder and the phonetic transcription folder. The Segmentation Corpus word segmentation program was run on the files in the orthographic transcription folder only, but not on the files in the phonetic transcription folder. After running the word segmentation program on the 56 orthographic transcription files of the HKCAC, the individual files were concatenated to make one big orthographic transcription file. Likewise, the 56 phonetic transcription files were concatenated to form one big phonetic transcription file. The numeric headers are section dividers within each of the big files. The big orthographic transcription file had two formats — MS Word and plain text, both having Traditional Chinese Big-5 encoding. The big phonetic transcription file had only the plain text format in ASCII. The IPA symbols and the tone marks in the HKCAC original Excel data files were automatically translated into ASCII. See Appendix 5.1 for the chart of correspondence between the IPA and other speech transcription systems encountered in this project. The processing of the HKCAC in the subsequent steps refers to the processing of these two big files in plain text format.

After applying the Segmentation Corpus word segmentation algorithm on the 56 orthographic transcription files of the HKCAC, the individual files were concatenated to make one big orthographic transcription file in plain text format. See Figure C.3. ("^^" tags word boundary in the Segmentation Corpus.) The subsequent application of the spoken Cantonese word segmentation algorithm that was developed in this dissertation was applied to this big file.

03211111
0^^喂^^hello^^
0^^喂^^
0^^喂^^
0^^[^^係^^]^^
0^^[^^係^^]^^你^^係^^
0^^E3**^^Sandy^^**
0^^Sandy^^
0^^係^^呀^^
0^^係^^你^^[^^好^^]^^
0^^[^^O.K.^^]^^
0^^係^^你^^好^^
0^^Sandy^^有^^[^^咩^^講^^]^^
0^^[^^E3^^]^^我^^就^^想^^講^^我^^個^^朋友^^
0^^佢^^個^^妹^^呢^^做^^嗰^^啲^^E3**^^帶^^街^^**㗎^^
04211111
0^^喂^^hello^^係^^你^^點^^稱呼^^
0^^喂**^^恐怖^^熱線^^**
0^^喂^^
0^^E3**^^潘^^迢^^聰^^露^^芙^^**[^^係^^咪^^]^^呀^^
0^^[^^Hello^^]^^
0^^[^^係^^]^^
0^^[^^係^^]^^
0^^[^^你^^好^^]^^
0^^呀^^你^^好^^E3^^我^^叫^^阿^^芝^^呀^^
0^^阿^^[^^芝^^]^^係^^嘛^^
0^^[^^係^^]^^
0^^係^^呀^^係^^[^^呀^^我^^]^^想^^講^^嘅^^嘢^^呢^^
0^^[^^係^^你^^好^^]^^
0^^關於^^係**^^長洲^^東^^堤^^**㗎^^

The proper name "Sandy" is not distinguished from other non-proper names in the corpus.

帶街 /ta:i^{33}ka:i^{55}/ is a colloquial Cantonese form for 'tourist guide'. It is a compound word.

恐怖 /hoŋ^{35}pou^{33}/ 'horror' and 熱線 /ji:t^{2}si:n^{33}/ 'hotline' are generic words used as a radio program name here.

The given names 迢聰 /si:u^{21}ts^{h}oŋ^{55}/ 'Siu Chung' and 露芙 /lou^{22}fu:^{21}/ 'Ruth' are not grouped and marked as proper names. The relationship of the former is not obvious with the surname 潘 /p^{h}u:n^{55}/.

The place name 長洲 'Cheung Chau' is not marked as proper name, while 東堤 'Tung Tai' is not grouped and marked as proper name.

Figure C.3 A snippet of the texts from the HKCAC after the Segmentation Corpus word segmentation algorithm was applied on the 56 orthographic transcription files.

After applying the spoken Cantonese word segmentation algorithm to segment the big orthographic transcription file of the HKCAC into words, short words were yielded. See Figure C.4. ("^^" tags word boundary and "**" proper names in the Segmentation Corpus.)

03211111
0^^喂^^hello^^
0^^喂^^
0^^喂^^
0^^[^^係^^]^^
0^^[^^係^^]^^你^^係^^
0^^E3**Sandy^^
0**Sandy^^
0^^係^^呀^^
0^^係^^你^^[^^好^^]^^
0^^[^^O.K.^^]^^
0^^係^^你^^好^^
0**Sandy^^有^^[^^咩^^講^^]^^
0^^[^^E3^^]^^我^^就^^想^^講^^我^^個^^朋友^^
0^^佢^^個^^妹^^呢^^做^^嗰啲^^E3**^^帶街^^**嘞^^
04211111
0^^喂^^hello^^係^^你^^點^^稱呼^^
0^^喂**恐怖**熱線^^
0^^喂^^
0^^E3**^^潘**迢聰**露芙^^**[^^係咪^^]^^呀^^
0^^[^^Hello^^]^^
0^^[^^係^^]^^
0^^[^^係^^]^^
0^^[^^你^^好^^]^^
0^^呀^^你^^好^^E3^^我^^叫**阿芝^^呀^^
0**阿[芝^^]^^係^^嘛^^
0^^[^^係^^]^^
0^^係^^呀^^係^^[^^呀^^我^^]^^想^^講^^嘅^^嘢^^呢^^
0^^[^^係^^你^^好^^]^^
0^^關於^^係**長洲**東堤^^**嘞^^

The "Sandy" is marked as proper name.

帶街 /ta:i^{33}ka:i^{55}/, a colloquial Cantonese form for 'tourist guide', is grouped as a word.

恐怖 /hoŋ^{35}pou^{33}/ 'scary' and 熱線 /ji:t^{2}si:n^{33}/ 'hotline' are generic words used as a radio program name here. They are marked as proper names.

The proper names 迢聰 'Siu Chung' and 露芙 'Ruth' are grouped and marked as proper names. The surname 潘 /p^{h}u:n^{55}/ is not tagged as proper name in the Segmentation Corpus.

Notations that are irrelevant for the current project (such as "[ ]" for overlapping speech) would be deleted.

The place name 長洲 'Cheung Chau' is marked as proper name, and 東堤 'Tung Tai' is grouped and marked as proper name.

Figure C.4 Short words were yielded after applying the spoken Cantonese word segmentation algorithm to segment the big orthographic transcription file of the HKCAC into words.

In this big orthography file, the short words were grouped into long words according to a set of grouping principles devised in this dissertation. The words were then tagged according to the set of tagging conventions developed for this dissertation (see Figure C.5).

Tagging the header with an XML tag <header></header>.

The proper name "Sandy" is marked as having two syllables (<L2></L2>) and a proper name (<pn></pn>).

Short words 有 /jɐu^{35}/ 'have' and 咩 /mɛː$^{55+23}$/ 'what' are grouped to become a long word, which is then grouped with 講 /kɔːŋ^{35}/ 'say' to become another long word, according the principles in this project.

恐怖熱線 /hoŋ^{35}pou^{33}jiːt^{2}siːn^{33}/ 'scary hotline' are generic words used as a radio program name here. They are marked as proper name in this project.

The surname 潘 /p^{h}uːn^{55}/ is tagged as proper name in this project. The surname is grouped with the given name 迢聰 /siːu^{21}ts^{h}oŋ^{55}/ 'Siu Chung'. 迢聰 and 露芙 'Ruth' are grouped and marked as proper names.

```
<header>03211111</header>
0<L1>喂</L1><L2>hello</L2>
0<L1>喂</L1>
0<L1>喂</L1>
0<L1>係</L1>
0<L1>係</L1><L1>你</L1><L1>係</L1>
0<L1>E3</L1><L2><pn>Sandy</pn></L2>
0<L2><pn>Sandy</pn></L2>
0<L1>係</L1><L1>呀</L1>
0<L1>係</L1><L1>你</L1><L1>好</L1>
0<L2>O.K.</L2>
0<L1>係</L1><L1>你</L1><L1>好</L1>
0<L2><pn>Sandy</pn></L2><L3><L2>有咩</L2>講</L3>
0<L1>E3</L1><L1>我</L1><L1>就</L1><L1>想</L1><L1>講</L1><L1>我</L1><L1>個
</L1><L2>朋友</L2>
0<L1>佢</L1><L1>個</L1><L1>妹</L1><L1>呢</L1><L1>做</L1><L2>嗰啲
</L2><L1>E3</L1><L2>帶街</L2><L1>㗎</L1>
<header>04211111</header>
0<L1>喂</L1><L2>hello</L2><L1>係</L1><L1>你</L1><L3>點<L2>稱呼</L2></L3>
0<L1>喂</L1><L4><pn><L2><pn>恐怖</pn></L2><L2><pn>熱線</pn></L2></pn></L4>
0<L1>喂</L1>
0<L1>E3</L1><L3><L1><pn>潘</pn></L1><L2><pn>迢聰</pn></L2></L3><L2><pn>露芙
</pn></L2><L2>係咪</L2><L1>呀</L1>
0<L2>hello</L2>
0<L1>係</L1>
0<L1>係</L1>
```

Figure C.5 In the big orthography file, the short words were grouped into long words and the words were tagged using XML tags.

Figure C.5 continued

0<L1>你</L1><L1>好</L1>
0<L1>呀</L1><L1>你</L1><L1>好</L1><L1>E3</L1><L1>我</L1><L1>叫
</L1><L2><gpp><L1>阿</L1><L1><pn>芝</pn></L1></gpp></L2><L1>呀</L1>
0<L2><gpp><L1>阿</L1><L1><pn>芝</pn></L1></gpp></L2><L1>係</L1><L1>嘛</L1>
0<L1>係</L1>
0<L1>係</L1><L1>呀</L1><L1>係</L1><L1>呀</L1><L1>我</L1><L1>想</L1><L1>講
</L1><L1>嘅</L1><L1>嘢</L1><L1>呢</L1>
0<L1>係</L1><L1>你</L1><L1>好</L1>
0<L2>關於</L2><L1>係</L1><L2><pn>長洲</pn></L2><L2><pn>東堤</pn></L2><L1>㗎
</L1>

Notations that are irrelevant for the current project (such as "[ ]" for overlapping speech) are deleted.

The place names 長洲 'Cheung Chau' and 東堤 'Tung Tai' are marked as proper names.

Figure C.5 ended

The orthographic transcription in the orthographic transcription file was then mapped onto their corresponding phonetic transcription in the phonetic transcription file. The phonetic transcription file was then segmented and tagged according to how the morphemes in the orthographic transcription file was segmented and tagged. See Figure C.6.

```
<header>03211111</header>
P<L1>wAi2</L1><L2>hGlGU</L2>
P<L1>wAi2</L1>
P<L1>wAi2</L1>
P<L1>hAi6</L1>
P<L1>hAi6</L1><L1>lei5</L1><L1>hAi6</L1>
P<L1>E3</L1><L2><pn>sYndI</pn></L2>
P<L2><pn>sYndI</pn></L2>
P<L1>hAi6</L1>
P<L1>hAi6</L1><L1>lei5</L1><L1>hou2</L1>
P<L2>okei</L2>
P<L1>hAi6</L1><L1>lei5</L1><L1>hou2</L1>
P<L2><pn>sYndI</pn></L2><L3><L1>jAu5</L1><L1>mE8</L1><L1>kON2</L1
></L3>
P<L1>E3</L1><L1>O5</L1><L1>tsAu6</L1><L1>sJN2</L1><L1>kON2</L1><L
1>NO5</L1><L1>kO3</L1><L2>pHAn4jAu5</L2>
P<L1>kH0y5</L1><L1>O3</L1><L1>mui2</L1><L1>lE1</L1><L1>tsou6</L1>
<L2>kO2ti1</L2><L1>E3</L1><L2>tai3kai1</L2>><L1>ka3</L1>
<header>04211111</header>
P<L1>wei2</L1><L2>hGlGU</L2><L1>hAi6</L1><L1>lei5</L1><L3>tim2<L2
>tsHiN1fu1</L2></L3>
P<L1>wei2</L1><L4><pn><L2><pn>hoN2pou3</pn></L2><L2><pn>jit6sin3<
/pn></L2></pn></L4>
P<L1>wei2</L1>
P<L1>E3</L1><L3><L1><pn>pHun1</pn></L1><L2><pn><pn>siu4tsHoN1</pn
></L2></L3><L2><pn>lou6fu4</pn></L2><L2>Ai6mAi6</L2><L1>a3</L1>
P<L2>hGlGU</L2>
P<L1>hAi6</L1>
P<L1>hAi6</L1>
P<L1>lei5</L1><L1>hou2</L1>
P<L1>a3</L1><L1>lei5</L1><L1>hou2</L1><L1>E3</L1><L1>O5</L1><L1>k
iu3</L1><L2><gpp><L1>a3</L1><pn>tsi1</pn></L1></gpp></L2><L1>ja3<
/L1>
P<L2><gpp><L1>a3</L1><L1><pn>tsi1</pn></L1></gpp></L2><L1>Ai6</L1
><L1>ma5</L1>
P<L1>hAi6</L1>
P<L1>hAi6</L1>ja3<L1>hAi6</L1>ja3<L1>O5</L1><L1>sJN2</L1><L1>kON2
</L1><L1>kE3</L1><L1>jE5</L1><L1>lE1</L1>
P<L1>Ai6</L1><L1>lei5</L1><L1>hou2</L1>
P<L2>kwan1jy1</L2><L1>hAi6</L1><L2><pn>tsHJN4tsAu1</pn></L2><L2><
pn>toN1tHAi4</pn></L2><L1>G3</L1>
```


Figure C.6 A snippet taken from the big phonetic transcription file (i.e. 56 original files concatenated), in which the phonetic transcription had been segmented and tagged according to how their corresponding orthographic transcription in the big orthographic transcription file were segmented and tagged.

APPENDIX D

GUIDELINES FOR MORPHOSYNTACTIC ANALYSIS
(GLOSS AND IPA TRANSCRIPTION ARE SUPPLIED HERE, BUT NOT IN THE ORIGINAL VERSION.)

A. What do you find in this list?

- 2+ syllable words with pronunciations in Cantonese.
- Each entry you see in this list is called a word in this project.
- These words were segmented out of the spoken Cantonese corpus HKCAC.
- Pronunciation in "/ /" or meaning in "( )" would be given next to the word in case of possible multiple readings.
- Pronunciations in "/ /" are given in Jyutping (Jyutping Romanization Scheme attached).
- There are English words in the list, but they are pronounced in Cantonese way, e.g. *music* [miu1 sik4] (in Jyutping).
- Sentence final particles are not included in the list.

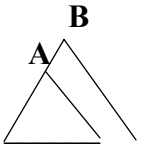
B. What are you going to do with the words?

- Draw a tree to indicate the internal morphosyntactic structure (see section C below) for each word entry.
- If you find there could be more than one possible analysis of the internal morphosyntactic structure for a given word, draw the structure that you think would agree with your intuition best.
- You DO NOT have to label the nodes of the trees.
- Please draw trees with a pen. This will save you lots of time!

C. Guidelines for drawing internal morphosyntactic structures for the words

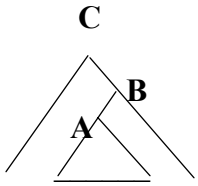
1. **Grouping**. Group the elements that are most closely related first (let's call it A). Then, group A with element(s) that is/are closely related to it (let's call it B), so on and so forth.

E.g. 1



朱古力　餅　　(You don't have to write the letters A, B, C, etc.)
/tsyː^{55}kuː^{55}lek^{5} pɛːŋ^{35}/
'chocolate cookie'

E.g. 2



食　朱古力　餅　　(You don't have to write the letters A, B, C, etc.)
/sek^{2} tsyː^{55}kuː^{55}lek^{5} pɛːŋ^{35}/
'eat chocolate cookie'

2. **Distinguish monomorphemic (simple) words from polymorphemic (complex) words**.

Monomorphemic (simple) words are words that CANNOT be further decomposed, as your native intuition judges it. E.g. 徘徊 /p^{h}uːi^{21}wuːi^{21}/ 'walking back and forth'; 葡萄 /p^{h}ou^{21}t^{h}ou^{21}/ 'grapes'.

Polymorphemic (complex) words are words that CAN be further decomposed, as your native intuition judges it.

Polymorphemic words include words that you judge to have any of the following structures (a-g):

a. 主謂 (subject-predicate): e.g. 腳軟 /kœːk^{3}jyːn^{23}/ ‘sore feet’, 頭痛 /t^{h}ɐu^{21}t^{h}oŋ^{33}/ ‘headache’

b. 偏正 (modifier-head): e.g. 靚女 /lɛːŋ^{33}nɵy^{35}/ ‘pretty girl’, 窮人 /k^{h}oŋ^{21}jɐn^{21}/

c. 動賓 (verb-object): e.g. 食粥 /sek^{2}tsok^{5}/ ‘eat congee’, 踢波 /t^{h}ɛːk^{3}pɔː^{55}/ ‘playing soccer’

d. 動補 (verb-complement): e.g. 打死 /taː^{35}sei^{35}/ ‘beaten to death’, 洗乾淨 /sɐi^{35}kɔːn^{55}tsɛːŋ^{22}/ ‘washed clean’, 行落去 /haːŋ^{21}lɔːk^{2}hɵy^{33}/ ‘walk down’ (see note (i) after (g))

e. 聯合/並列 (parallel): e.g. 男女 /naːm^{21}nɵy^{23}/ ‘male and female’, 阿爸阿媽 /aː^{33}paː^{21}aː^{33}maː^{55}/ ‘dad and mom’

f. 附加 (加上詞綴的詞 ‘words formed by affixation’), e.g.:

前綴 (prefix, attached before a root): 阿妹 /aː^{33}muːi^{35}/ ‘younger sister’, 第二 /tɐi^{22}jiː^{22}/ ‘the second’, 唔食 /m̩^{21}sek^{2}/ ‘do not eat’

中綴 (infix, embedded inside a root): 麻鬼煩 /maː^{21}kwɐi^{35}faːn^{21}/ ‘very troublesome’ (see note (iii) after (g))

後綴 (suffix, attached after a root): 做完 /tsou^{22}jyːn^{21}/ ‘finished’, 食晒 /sek^{2}saːi33/ ‘eat all’, 參加過 /ts^{h}aːm^{55}kaː^{55}kwɔː^{33}/ ‘participated’, 經濟學 /keŋ^{55}tsɐi^{33}hɔːk^{2}/ ‘economics’, 藝術家 /ŋɐi^{22}sɵt^{2}kaː^{55}/ ‘artists’ (see note (i) after (g))

g. 重疊 (reduplication: affixing an element which copies some feature from the root)[1], e.g.:

AA	AA 哋/tei^{35}/	AABB	AAB	ABB
A 一/jɐt^{5}/A	A 兩/lœːŋ^{23}/A	AA 聲/sɛːŋ^{55}/		
A 鬼/kwɐi^{35}/A	A 馬/maː^{23}/	(see note (ii) below)		

Notes

i. 動補 (verb-complement) and 附加 (affixation) 結構 structure: (cf. Cheung 1972) 補語 (complement) can be in the form of verbal particles. These verbal particles are attached after the verb to indicate the result, state, direction, degree, etc, of the verb they attached to. These verbal particles, therefore, can be treated as suffixes. Some examples of these verbal particles are:

完 /jyːn^{21}/ 'finished' 晒 /saːi^{33}/ 'all'

嚟 /kɐn^{35}/ 'continuous' 吓 /haː^{23}/ 'to do a bit'

到 /tou^{35}/ 'able to'：做到 /tsou^{22}tou^{35}/ 'can do'

過 /kwɔː^{33}/ 'experienced'：做過 /tsou^{22}kwɔː^{33}/ 'have done'

定 /teŋ^{22}/ 'in advance'：做定 /tsou^{22}teŋ^{22}/ 'do in advance'

過 /kwɔː^{33}/ 'over'：跳過 /tsou^{22}kwɔː^{33}/ 'jump over'

上嚟 /sœːŋ^{23}lei^{21}/ 'up' 上去 /sœːŋ^{23}hɵy^{33}/ 'up'

落嚟 /lɔːk^{2}lei^{21}/ 'down' 落去 /lɔːk^{2}hɵy^{33}/ 'down'

開嚟 /hɔːi^{55}lei^{21}/ 'come here' 開去 /hɔːi^{55}hɵy^{33}/ 'go there'

埋嚟 /maːi^{21}lei^{21}/ 'here' 埋去 /maːi^{21}hɵy^{33}/ 'there'

出嚟 /ts^{h}ɵt^{5}lei^{21}/ 'out' 出去 /ts^{h}ɵt^{5}hɵy^{33}/ 'out'

[1] http://www.ling.lu.se/education/homepages/ALS052/handouts/TL5.pdf

出 /ts^{h}ɵt^{5}/ ‘emerge’：諗出 /nɐm^{35}ts^{h}ɵt^{5}/ ‘think up’ ; 教出 /kaːu^{33}ts^{h}ɵt^{5}/ ‘teach’

低 /tɐi^{55}/：放低 /fɔːŋ^{33}tɐi^{55}/ ‘put down’; 寫低 /sɛː^{35}tɐi^{55}/ ‘write down’

番/返 /faːn^{55}/ ：飛返 /fei^{55}faːn^{55}/ ‘fly back’

起 /hei^{35}/ ‘up’：掛起 /kwaː^{33}hei^{35}/ ‘hang up’

開 /hɔːi^{55}/ ‘away’：行開 /haːŋ^{21}hɔːi^{55}/ ‘go away’

落去 /lɔːk^{2}hɵy^{33}/ ‘continue’：喊落去 /haːm^{33}lɔːk^{3}hɵy^{33}/ ‘keep crying’

埋 /maːi^{21}/ ‘together’：喺埋 /hɐi^{35}maːi^{21}/ ‘gather together’

Verbal particles may be separated from the verb by the modal 得 or the negation marker 唔, e.g:

食得晒 /sek^{2}tɐk^{5}saːi^{33}/ ‘can eat all up’

食唔晒 /sek^{2}m̩^{21}saːi^{33}/ ‘cannot eat all up’

食得落 /sek^{2}tɐk^{5}lɔːk^{2}/ ‘can eat, have the appetite to eat’

食唔落 /sek^{2}m̩^{21}lɔːk^{2}/ ‘cannot eat, do not have the appetite to eat’

The modal 得 and the negation marker 唔 are treated as affixes in this project. It’s up to you whether you want to treat them as infixes or suffixes in a given context. Treating them as infixes would give the tree structure as shown in e.g. 3; treating them as infixes would give the tree structure as shown in e.g. 4.

E.g. 3 得 /tɐk^{5}/ ‘can’ as infix

食 **得** 飽
/sek^{2}tɐk^{5}paːu^{35}/
‘eat and feel full’

E.g. 4 得 /tɐk^{5}/ ‘can’ as suffix

食 **得** 晒
/sek^{2}tɐk^{5}saːi^{33}/
‘can eat all up’

ii. 重疊 (reduplication):

The assumption for reduplication is that for the two identical forms of morphemes (e.g. 乾乾 /kɔːn^{55}kɔːn^{55}/), the second element is the copy of the first element and it affixes after the first element. That is, it is not the case that the first element is the copy of the second element.

E.g. 5-8 below demonstrate how the tree structures of the followings could be like.

AABB: e.g.5 乾乾淨淨 /kɔːn^{55}kɔːn^{55}tsɛːŋ^{22}tsɛːŋ^{22}/ 'clean'

ABAB: e.g.6 逐個逐個 /tsok^{2}kɔː^{33}tsok^{2}kɔː^{33}/ 'one by one'

AAB: e.g.7 騰騰震 /t^{h}ɐn^{21}t^{h}ɐn^{35}tsɐn^{33}/ 'shivering'

ABB: e.g.8 肥騰騰 /fei^{21}t^{h}ɐn^{21}tsɐn^{21}/ 'fat'

E.g. 5

乾 乾 淨 淨

/kɔːn^{55}kɔːn^{55}tsɛːŋ^{22}tsɛːŋ^{22}/

E.g. 7

騰 騰 震

/t^{h}ɐn^{21}t^{h}ɐn^{35}tsɐn^{33}/

E.g. 6

逐 個 逐 個

/tsok^{2}kɔː^{33}tsok^{2}kɔː^{33}/

E.g. 8

肥 騰 騰

/fei^{21}t^{h}ɐn^{21}tsɐn^{21}/

iii. The tree structure of a word with an infix:

E.g. 9 鬼 /kwɐi^{35}/ ‘ghost’ as an infix

麻 鬼 煩
/maː^{21}kwɐi^{35}faːn^{21}/
‘very troublesome’

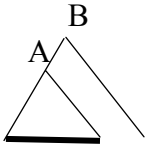
3. **Indicating monomorphemic (simple) words vs. polymorphemic (complex) words**.

Indicate monomorphemic (simple) words <u>WITH a horizontal line within the word</u>, as in e.g. 10 and 11.

E.g. 10

馬 虎 (i.e. you indicate that 馬虎 /maː^{55}fuː^{55}/, meaning 求其 /k^{h}ɐu^{21}k^{h}ei^{21}/
/maː^{55}fuː^{55}/ ‘carelessly, not detail oriented’ is a monomorphemic word)
‘careless’

E.g. 11



Bible 堂 (i.e. you indicate that *Bible* is a monomorphemic word)
/paːi^{55}pou^{21} t^{h}ɔːŋ^{21}/
‘Bible class’

Indicate polymorphemic (complex) words <u>WITHOUT a horizontal line within the word</u>, as in e.g. 12 and 13.

E.g. 12

A

炒 麵

/ts^{h}aːu^{35}miːn^{22}/

'fried noodle'

(i.e. you indicate that 炒麵 /ts^{h}aːu^{35}miːn^{22}/ 'fried noodle' is a polymorphemic word)

E.g. 13

A

做 完

/tsou^{22}jyːn^{21}/

'done'

(i.e. you indicate that 做完 /tsou^{22}jyːn^{21}/ 'done' is a polymorphemic word)

4. For word entries that you cannot make sense out of it, put a "**?**" after it.

E.g. 芝辜 /tsiː^{55}kuː^{55}/ **?**

If you have questions, please feel free to contact me anytime:

mobile: 98662071

e-mail: pwong@ling.ohio-state.edu

THANK YOU SO VERY MUCH FOR YOUR GENEROUSLY HELP !!

========= End of guidelines =========

香港語言學學會粵語拼音方案 (1993)

The Jyutping Cantonese Romanization Scheme developed by the Linguistic Society of Hong Kong (LSHK)

聲母 (19個) 19 onsets

b (巴)[$]	p (怕)	m (媽)	f (花)	
d (打)	t (他)	n (那)		l (啦)
g (家)	k (卡)	ng (牙)*	h (蝦)	
gw (瓜)	kw (誇)			w (蛙)
z (渣)	c (叉)		s (沙)	j (也)

$ In parenthesis are example words.

*零聲母不用字母作標記, 如 "呀"只拼作 aa.

*Null onset is not coded, e.g., the Jyutping symbol "aa" is used.

韻腹 (9個) 9 nuclear vowels

aa (沙)	i (詩/星/識)	u (夫/風/福)	e (些/四)	o (疏/蘇)
yu (書)	oe (鋸)			
a (新)	eo (詢)			

韻尾 (8個) 8 codas

p (濕)	t (失)	k (塞)
m (心)	n (新)	ng (笙)
	i (西 / 需)	u (收)

- 鼻音單獨成韻 syllabic nasals: m (唔) ng (吳)

字調 Lexical tone

調號 tone numbers：	1 (夫/福)	2 (虎)	3 (副/霍)
	4 (扶)	5 (婦)	6 (父/服)

標調位置放在音節後，如：

Tone numbers are placed after the syllable, e.g.:

fu1 (夫)	fu2 (虎)	fu3 (副)
fu4 (扶)	fu5 (婦)	fu6 (父)

APPENDIX E

A LIST OF 173 WORDS FOR TESTING IN EXPERIMENTS II AND III

Below is the list of the 173 stimuli for testing the influence of the identified potential factors on syllable fusion. The list is first arranged by word frequency (high, mid, low), then by word length, and lastly by word-internal branching structure.

(i) High frequency words ('m' means monomorphemic within the innermost pair XML tag <L#></L#>)

words	**'underlying' form**	**gloss**	**count**	**m**	**$\log_{10}$ freq**
2 syllables					
no branching					
<L2>自己</L2>	tsiː^{22}kei^{35}	self	394	m	2.595496
<L2>或者</L2>	waːk^{2}tsɛː^{35}	or	251	m	2.399674
<L2>朋友</L2>	pɐŋ^{21}jɐu^{23}	friend	215	m	2.332438
<L2>所以</L2>	sɔː^{35}jiː^{23}	therefore	177	m	2.247973
<L2>知道</L2>	tsiː^{55}tou^{33}	know	118	m	2.071882
<L2>不過</L2>	pɐt^{5}kwɔː^{33}	but	104	m	2.017033
single-level branching					
<L2>呢個</L2>	nei/ni/ji^{55}kɔː^{33}	this one	799		2.902547
<L2>我哋</L2>	ŋɔː^{23}tei^{22}	we	642		2.807535
<L2>嗰個</L2>	kɔː^{35}kɔː^{33}	that one	552		2.741939
<L2>但係</L2>	taːn^{22}hɐi^{22}	but	527		2.705864
<L2>其實</L2>	k^{h}ei^{21}sɐt^{2}	in fact	493		2.692847
<L2>唔係</L2>	m̩^{21}hɐi^{22}	NEG	482		2.683047
<L2>就係</L2>	tsɐu^{22}hɐi^{22}	is	478		2.679428
<L2>可以</L2>	hɔː^{35}jiː^{23}	can	464		2.666518
<L2>覺得</L2>	kɔːk^{3}tɐk^{5}	feel	458		2.660865
<L2>因爲</L2>	jɐn^{55}wɐi^{22}	because	448		2.651278
<L2>咁樣</L2>	kɐm^{35}jœːŋ^{35}	this way	444		2.647383
<L2>如果</L2>	jyː^{21}kwɔː^{35}	if	393		2.594393
<L2>而家</L2>	jiː^{21}kaː^{55}	now	343		2.535294
<L2>可能</L2>	hɔː^{35}nɐŋ^{21}	perhaps	327		2.514548

(ii) Mid frequency words ('m' means monomorphemic within the innermost pair XML tag <L#></L#>)

words	'underlying' form	gloss	count		log_{10} freq
2 syllables					
no branching					
<L2>以前</L2>	jiː^{23}tsʰiːn^{21}	in the past	58	m	1.763428
<L2>經濟</L2>	keŋ^{55}tsɐi^{33}	economics	36	m	1.556303
<L2>然後</L2>	jiːn^{21}hɐu^{22}	then	32	m	1.50515
<L2>考慮</L2>	haːu^{35}lɵy^{22}	consider	28	m	1.447158
<L2>辛苦</L2>	sɐn^{55}fuː^{35}	feeling hard	26	m	1.414973
<L2>網絡</L2>	mɔːŋ^{23}lɔːk^{2}	network	26	m	1.414973
<L2>以爲</L2>	jiː^{23}wɐi^{21}	misinterpret	24	m	1.380211
<L2>雖然</L2>	sɵy^{55}jiːn^{21}	although	24	m	1.380211
<L2>等如</L2>	tɐŋ^{35}jyː^{21}	equals	22	m	1.342423
<L2>先生</L2> = mister	siːn^{55}saːŋ^{55}	mister	12	m	1.079181
single-level branching					
<L2>放棄</L2>	fɔːŋ^{33}hei^{33}	give up	28		1.447158
<L2>行爲</L2>	hɐŋ^{21}wɐi^{21}	behavior	28		1.447158
<L2>做嘢</L2>	tsou^{22}jɛː^{23}	do work	26		1.414973
<L2>唔敢</L2>	m̩^{21}kɐm^{35}	dare not	28		1.447158
<L2>發覺</L2>	faːt^{3}kɔːk^{3}	find out	28		1.447158
<L2>報紙</L2>	pou^{33}tsiː^{35}	newspaper	28		1.447158
<L2>市民</L2>	siː^{23}mɐn^{21}	citizen	28		1.447158
<L2>投資</L2>	tʰɐu^{21}tsiː^{55}	invest	28		1.447158
<L2>電腦</L2>	tiːn^{22}nou^{23}	computer	28		1.447158
<L2>重要</L2>	tsoŋ^{22}jiːu^{33}	important	28		1.447158
no branching					
---					
3 syllables					
single-level branching					
<L3>越嚟越</L3>	jyːt^{21}ei/lɐi^{21}jyːt^{2}	more and more	14		1.146128
right branching					
<L3>一<L2>部份 </L2></L3>	jɐt^{5}pou^{22}fɐn^{22}	a portion	15		1.176091

words	‘underlying’ form	gloss	count		log_{10} freq
<L3>返<L2>屋企</L2></L3>	fa:n^{55}ok^{5}k^{h}ei^{35}	go back home	10		1
<L3>呢<L2>一個</L2></L3>	nei/ni/lei/li/ji^{55}jɐt^{5}kɔ:^{33}	this one	92		1.963788
<L3>唔<L2>記得</L2></L3>	m̩^{21}kei^{33}tɐk^{5}	forget	14		1.146128
<L3>嗰<L2>方面</L2></L3>	kɔ:^{35}fɔ:ŋ^{55}mi:n^{22}	that aspect	12		1.079181
<L3>大<L2>姑奶</L2></L3>	ta:i^{22}ku:^{55}na:i^{55}	sister-in-law	10		1
<L3>打<L2>電話</L2></L3>	ta:^{35}ti:n^{22}wa:^{35}	to call	27		1.431364
<L3>差<L2>唔多</L2></L3>	tsha:^{55}m̩^{21}tɔ:^{55}	almost	18		1.255273
<L3>搵<L2>唔到</L2></L3>	wɐn^{35}m̩^{21}tou^{35}	can’t find	11		1.041393
left branching					
<L3><L2>影響</L2>到</L3>	jeŋ^{35}hœ:ŋ^{35}tou^{35}	affect	10		1
<L3><L2>預算</L2>案</L3>	jy:^{22}sy:n^{33}ŋɔ:n^{33}	budget	32		1.50515
<L3><L2>基本</L2>法</L3>	kei^{55}pu:n^{35}fa:t^{3}	Basic Law	17		1.230449
<L3><L2>唔該</L2>晒</L3>	m̩^{21}kɔ:i^{55}sa:i^{33}	thank you	22		1.342423
<L3><L2>冇嘢</L2>做</L3>	mou^{23}jɛ:^{23}tsou^{22}	nothing to do	11		1.041393
<L3><L2>細路</L2>仔</L3>	sɐi^{33}lou^{22}tsɐi^{35}	children	12		1.079181
<L3><L2>私營</L2>化</L3>	si:^{55}jeŋ^{21}fa:^{33}	privatization	35		1.544068
<L3><L2>突然</L2>間</L3>	tɐt^{2}j:n^{21}ka:n^{55}	suddenly	25	m	1.39794
<L3><L2>財政</L2>司</L3>	ts^{h}ɔ:i^{21}tseŋ^{33}si:^{55}	financial secretary	12		1.079181
<L3><L2>互聯</L2>網</L3>	wu:^{22}ly:n^{21}mɔ:ŋ^{23}	Internet	22		1.342423

(iii) Low frequency words (‘m’ means monomorphemic within the innermost pair XML tag <L#></L#>)

words	‘underlying’ form	gloss	count	m	$\log_{10}$ freq
2 syllables					
no branching					
<L2>虛僞</L2>	hɵy^{55}ŋɐi^{22}	hypocritical	1	m	0
<L2>論盡</L2>	lɵn^{22}tsɵn^{22}	clumsy	1	m	0
<L2>零舍</L2>	leŋ^{21}sɛː^{33}	especially	1	m	0
<L2>麻雀</L2>	maː^{21}tsœːk^{3}	sparrow	1	m	0
<L2>朦糊</L2>	moŋ^{21}wuː^{21}	blurred	1	m	0
<L2>惡劣</L2>	ŋɔːk^{3}lyːt^{2}	awkward	1	m	0
<L2>蕭條</L2>	siːu^{55}t^{h}iːu^{21}	slow down	1	m	0
<L2>打算</L2>	taː^{35}syːn^{33}	plan	1	m	0
<L2>糟蹋</L2>	tsou^{55}tsɐt^{5}	torture	1	m	0
<L2>污衊</L2>	wuː^{55}miːt^{2}	smear	1	m	0
single-level branching					
<L2>去傾</L2>	hɵy^{33}k^{h}eŋ^{55}	discuss	1		0
<L2>開住</L2>	hɔːi55tsyː^{22}	opened / turned on	1		0
<L2>應承</L2>	jeŋ^{55}seŋ^{21}	promise	1		0
<L2>搞嚟</L2>	kaːu^{35}kɐn^{35}	doing	1		0
<L2>掛名</L2>	kwaː^{33}mɛːŋ^{35}	to register a name	1		0
<L2>靈敏</L2>	leŋ^{21}mɐn^{23}	nimble	1		0
<L2>外資</L2>	ŋɔːi^{22}tsiː^{55}	foreign capital	1		0
<L2>把持</L2>	paː^{35}ts^{h}iː^{21}	hold on to	1		0
<L2>新嘢</L2>	sɐn^{55}jɛː^{23}	new things	1		0
<L2>沮喪</L2>	tsɵy^{35}sɔːŋ^{33}	discouraged	1		0
3 syllables					
single-level branching					
<L3>乜鬼嘢</L3>	mɐt^{5}kwɐi^{35}jɛː^{23}	whatever	1		0
<L3>面對面</L3>	miːn^{22}tɵy^{33}miːn^{22}	face to face	1		0
<L3>避一避</L3>	pei^{22}jɐt^{5}pei^{22}	avoid	1		0
<L3>追趕跑</L3>	tsɵy^{55}kɔːn^{35}p^{h}aːu^{35}	chase	1		0

words	‘underlying’ form	gloss	count	m	log_{10} freq
no branching					
<L3>冚唪爛</L3>	hɐm^{22}paːŋ^{22}laːn^{22}	all	2		0.30103
<L3>海洛英</L3>	hɔːi^{35}lɔːk^{2}jeŋ^{55}	cocaine	1		0
<L3>鬱金香</L3>	wɐt^{5}kɐm^{55}hœːŋ^{55}	tulip	1		0
right branching					
<L3>一<L2>大輪</L2></L3>	jɐt^{5}taːi^{22}lɵn^{21}	a round of	1		0
<L3>傾<L2>唔妥</L2></L3>	k^{h}eŋ^{55}m̩^{21}t^{h}ɔː^{23}	can’t come to agreement	1		0
<L3>女<L2>家鬼</L2></L3>	nɵy^{23}kaː^{55}kwɐi^{35}	female house ghost	1		0
<L3>唔<L2>合法</L2></L3>	m̩^{21}hɐp^{2}faːt^{3}	illegal	1		0
<L3>呢<L2>一家</L2></L3>	nei/ni/ji^{55}jɐt^{5}kaː^{33}	this one	1		0
<L3>上<L2>天堂</L2></L3>	sœːŋ^{23}t^{h}iːn^{55}t^{h}ɔːŋ^{21}	go to heaven	1		0
left branching					
<L3><L2>音樂</L2>會</L3>	jɐm^{55}ŋɔːk^{2}wuːi^{35}	concert	1		0
<L3><L2>諗嚟</L2>咩</L3>	nɐm^{35}kɐn^{35}mɛː^{55}	thinking what	1		0
<L3><L2>勞工</L2>界</L3>	lou^{21}koŋ^{55}kaːi^{33}	labor sector	1		0
<L3><L2>審計</L2>處</L3>	sɐm^{35}kɐi^{33}ts^{h}yː^{23}	audit department	1		0
<L3><L2>準備</L2>咗</L3>	tsɵn^{35}pei^{33}tsɔː^{35}	prepared	1		0
<L3><L2>超越</L2>到</L3>	ts^{h}iːu^{55}jyːt^{2}tou^{35}	surpass	1		0
<L3><L2>不得</L2>了</L3>	pɐt^{5}tɐk^{5}liːu^{23}	intensified	1		0
4 syllables					
no branching					
<L4>卡拉 OK</L4>	k^{h}aː^{55}laːi^{55}ou^{55}k^{h}ei^{55}	karaoke	1		0
<L4>歇斯底里</L4>	kiːt^{3}siː^{55}tɐi^{35}lei^{23}	hysterical	2		0.30103

words	'underlying' form	gloss	count	m	$\log_{10}$ freq
single-level branching					
<L4>可免則免</L4>	hɔː^{35}miːn^{23}tsɐk^{5}miːn^{23}	avoid if one can	1		0
<L4>可大可小</L4>	hɔː^{35}taːi^{22}hɔː^{35}siːu^{35}	big or small	1		0
<L4>人之常情</L4>	jɐn^{21}tsiː^{55}sœːŋ^{21}ts^{h}eŋ^{21}	human nature	1		0
<L4>一沉百踩</L4>	jɐt^{5}ts^{h}ɐm^{21}paːk^{3}ts^{h}aːi^{35}	to be walked on by others	1		0
<L4>由始至終</L4>	jɐu^{21}ts^{h}iː^{35}tsiː^{33}tsoŋ^{55}	from beginning to end	1		0
<L4>另作別論</L4>	leŋ^{22}tsɔːk^{3}piːt^{2l}ɵn^{22}	another story	1		0
<L4>五顏六色</L4>	ŋ^{23}ŋaːn^{21}lok^{2}sek^{5}	colorful	1		0
<L4>五臟俱全</L4>	m̩^{23}tsɔːŋ^{22}k^{h}ɵy^{55}ts^{h}yːn^{21}	has everything	1		0
<L4>柴米油鹽</L4>	ts^{h}aːi^{21}mɐi^{23}jɐu^{21}jiːm^{21}	firewood, rice, oil, and salt	1		0
<L4>情有可原</L4>	ts^{h}eŋ^{21}jɐu^{23}hɔː^{35}jyːn^{21}	forgivable	1		0
balanced branching: ABAB					
<L4><L2>拗下 </L2><L2>拗下 </L2></L4>	ŋaːu^{55}haː^{23}ŋaːu^{55}haː^{23}	to scratch	2		0.30103
<L4><L2>一步 </L2><L2>一步 </L2></L4>	jɐt^{5}pou^{22}jɐt^{5}pou^{22}	step by step	1		0
<L4><L2>賣下 </L2><L2>賣下 </L2></L4>	maːi^{22}haː^{23}maːi^{22}haː^{23}	to sell	1		0
<L4><L2>拖下 </L2><L2>拖下 </L2></L4>	t^{h}ɔː^{55}haː^{23}t^{h}ɔː^{55}haː^{23}	to procrastinate	2		0.30103
balanced branching: AABB					
<L4>肥肥大大</L4>	fei^{21}fei^{21}taːi^{22}taːi^{22}	fat and big	1		0
<L4>輕輕鬆鬆</L4>	heŋ^{55}heŋ^{55}soŋ^{55}soŋ^{55}	relaxing	1		0
<L4>原原本本</L4>	jyːn^{21}jyːn^{21}puːn^{35}puːn^{35}	in original form	1		0
<L4>驚驚青青</L4>	kɛːŋ^{55}kɛːŋ^{55}ts^{h}ɛːŋ^{55}ts^{h}ɛːŋ^{55}	hesitant	1		0
<L4>嚟嚟去去</L4>	lei^{21}lei^{21}hɵy^{33}hɵy^{33}	back and forth	1		0

words	‘underlying’ form	gloss	count	m	log_{10} freq
<L4>迷迷糊糊</L4>	mɐi^{21}mɐi^{21}wuː^{21}wuː^{21}	vague	1		0
<L4>平平穩穩</L4>	p^{h}eŋ^{21}p^{h}eŋ^{21}wɐn^{35}wɐn^{35}	stable	1		0
<L4>實實際際</L4>	sɐt^{2}sɐt^{2}tsɐi^{33}tsɐi^{33}	practically	1		0
<L4>偷偷摸摸</L4>	t^{h}ɐu^{55}t^{h}ɐu^{55}mɔː^{55}mɔː^{55}	sneaky	1		0
<L4>拖拖拉拉</L4>	t^{h}ɔː^{55}t^{h}ɔː^{55}laːi^{55}laːi^{55}	to procrastinate	1		0
<L4>追追逐逐</L4>	tsɵy^{55}tsɵy^{55}tsok^{2}tsok^{2}	chasing after	1		0
<L4>出出入入</L4>	ts^{h}ɵt^{5}ts^{h}ɵt^{5}jɐp^{2}jɐp^{2}	going in and out	1		0
<L4>正正當當</L4>	tseŋ^{33}tseŋ^{33}tɔːŋ^{33}tɔːŋ^{33}	proper	1		0
balanced branching: others					
<L4><L2>法治</L2><L2>不張</L2></L4>	faːt^{3}tsiː^{22}pɐt^{5}tsœːŋ^{55}	not ruled by law	1		0
<L4><L2>汽車</L2><L2>隧道</L2></L4>	hei^{33}ts^{h}ɛː^{55}sɵy^{22}tou^{33}	car tunnel	1		0
<L4><L2>人人</L2><L2>平等</L2></L4>	jɐn^{21}jɐn^{21}p^{h}eŋ^{21}tɐŋ^{35}	all man are equal	1		0
<L4><L2>有碟</L2><L2>話碟</L2></L4>	jɐu23tiːp2waː22tiːp2	to tell the truth	1		0
<L4><L2>有朝</L2><L2>一日</L2></L4>	jɐu^{23}tsiːu^{55}jɐt^{5}jɐt^{2}	one day	1		0
<L4><L2>個別</L2><L2>人士</L2></L4>	kɔː^{33}piːt^{2}jɐn^{21}siː^{22}	individual beings	1		0
<L4><L2>鬼五</L2><L2>馬六</L2></L4>	kwɐi^{35}m^{23}maː^{23}lok^{2}	wacky	1		0
<L4><L2>保留</L2><L2>條文</L2></L4>	pou^{35}lɐu^{21}t^{h}iːu^{21}mɐn^{21}	to retain an ordinance	1		0
<L4><L2>收費</L2><L2>水平</L2></L4>	sɐu^{55}fɐi^{33}sɵy^{35}p^{h}eŋ^{21}	level of charge	1		0
<L4><L2>動議</L2><L2>辯論	toŋ^{22}jiː^{23}piːn^{22}lɵn^{22}	to move a motion	1		0

words	‘underlying’ form	gloss	count	m	log_{10} freq
</L2></L4>					
<L4><L2>產後 </L2><L2>護理 </L2></L4>	$ts^{h}a:n^{35}hɐu^{22}wu:^{22}lei^{23}$	after-birth care	1		0
<L4><L2>中醫 </L2><L2>中藥 </L2></L4>	$tsoŋ^{55}ji:^{55}tsoŋ^{55}jœ:k^{2}$	Chinese medical treatment and medicine	1		0
<L4><L2>黃金 </L2><L2>時段 </L2></L4>	$wɔ:ŋ^{21}kɐm^{55}si:^{21}ty:n^{22}$	golden time slot	1		0
right branching					
<L4>去<L3>諗<L2>辦法 </L2></L3></L4>	$hɵy^{33}nɐm^{35}pa:n^{22}fa:t^{3}$	go think out a way	1		0
<L4>去<L3>買<L2>外賣 </L2></L3></L4>	$hɵy^{33}ma:i^{23}ŋɔ:i^{22}ma:i^{22}$	go buy carry-out	1		0
<L4>去<L3>唔<L2>去街 </L2></L3></L4>	$hɵy^{33}m̩^{21}hɵy^{33}ka:i^{55}$	go out or not	1		0
<L4>有<L3>咩<L2>聯想 </L2></L3></L4>	$jɐu^{23}mɛ:^{55}ly:n^{21}sœ:ŋ^{35}$	has any association	1		0
<L4>容<L3>有<L2>不當 </L2></L3></L4>	$joŋ^{21}jɐu^{23}pɐt^{5}tɔ:ŋ^{33}$	bear with wrong doings	1		0
<L4>記<L3>唔<L2>記到 </L2></L3></L4>	$kei^{33}m̩^{21}kei^{33}tou^{35}$	can remember or not	1		0
<L4>亂<L3>拋<L2>垃圾 </L2></L3></L4>	$ly:n^{22}p^{h}a:u^{55}la:p^{2}sa:p^{3}$	littering	1	m	0
<L4>心<L3>有<L2>靈犀 </L2></L3></L4>	$sɐm^{55}jɐu^{23}leŋ^{21}sɐi^{55}$	read each other’s mind	1		0
<L4>嗰<L3>一<L2>類型 </L2></L3></L4>	$kɔ:^{35}jɐt^{5}lɵy^{22}jeŋ^{21}$	that type	1		0
<L4>心<L3>有<L2>不妙 </L2></L3></L4>	$sɐm^{55}jɐu^{23}pɐt^{5}mi:u^{22}$	make one feel at peace in the heart	1		0
<L4>四<L3>大<L2>天王 </L2></L3></L4>	$sei^{33}ta:i^{33}ti:n^{55}wɔ:ŋ^{21}$	four top singers	1		0
<L4>當<L3>唔<L2>準確 </L2></L3></L4>	$tɔ:ŋ^{33}m^{21}tsɵn^{35}k^{h}ɔ:k^{3}$	be regarded as inaccurate	1		0
<L4>當<L3>冇<L2>件事 </L2></L3></L4>	$tɔ:ŋ^{33}mou^{23}ki:n^{22}si:^{22}$	pretend that nothing has happened	1		0
<L4>足<L3>唔<L2>足夠 </L2></L3></L4>	$tsok^{5}m̩^{21}tsok^{5}kɐu^{33}$	enough or not	1		0

words	‘underlying’ form	gloss	count	m	log_{10} freq
<L4>鐘<L3>唔<L2>鐘意</L2></L3></L4>	tsoŋ^{55}m̩^{21}tsoŋ^{55}jiː^{33}	like it or not	1		0
left branching					
<L4><L3><L2>起得</L2>番</L3>身</L4>	hei^{35}tɐk^{5}faːn^{55}sɐn^{55}	can stand back up	1		0
<L4><L3><L2>學習</L2>得</L3>到</L4>	hɔːk^{3}tsaːp^{2}tɐk^{5}tou^{35}	can learn	1		0
<L4><L3><L2>醫番</L2>好</L3>咗</L4>	jiː^{55}faːn^{55}hou^{35}tsɔː^{35}	be healed	1		0
<L4><L3><L2>預見</L2>得</L3>到</L4>	jyː^{22}kiːn^{33}tɐk^{5}tou^{35}	can foresee	1		0
<L4><L3><L2>機械</L2>人</L3>式</L4>	kei^{55}haːi^{22}jɐn^{21}sek^{5}	robotic	1		0
<L4><L3><L2>匿咗</L2>入</L3>房</L4>	lei^{55}tsɔː^{35}jɐp^{2}fɔːŋ^{35}	hide in the room	1		0
<L4><L3><L2>唔見</L2>咗</L3>人</L4>	m̩^{21}kiːn^{33}tsɔː^{35}jɐn^{21}	the person disappeared	1		0
<L4><L3><L2>順住</L2>嚟</L3>擺</L4>	sɵn^{22}tsyː^{22}lei^{21}paːi^{35}	to put in order	1		0
<L4><L3><L2>寫咗</L2>信</L3>去</L4>	sɛː^{35}tsɔː^{35}sɵn^{33}hɵy^{33}	wrote a letter to somewhere	1		0
<L4><L3><L2>釋放</L2>番</L3>晒</L4>	sek^{5}fɔːŋ^{33}faːn^{55}saːi^{33}	released	1		0
<L4><L3><L2>打過</L2>去</L3>問</L4>	taː^{35}kwɔː^{33}hɵy^{33}mɐn^{22}	called to ask	1		0
<L4><L3><L2>代落</L2>咗</L3>去</L4>	tɔːi^{22}lɔːk^{2}tsɔː^{35}hɵy^{33}	putt in (some stuff in somewhere)	1		0
<L4><L3><L2>站得</L2>住</L3>腳</L4>	tsaːm^{22}tɐk^{5}tsyː^{22}kœːk^{3}	can stand up (under some conditions)	1		0
<L4><L3><L2>走番</L2>出</L3>廳</L4>	tsɐu^{35}faːn^{55}ts^{h}ɵt^{5}t^{h}ɛːŋ^{55}	went back out to the living room	1		0
<L4><L3><L2>沖完</L2>涼</L3>走</L4>	ts^{h}oŋ^{55}jyːn^{21}lœːŋ^{21}tsɐu^{35}	left after shower	1		0
<L4><L3><L2>搵嚟</L2>工</L3>做</L4>	wɐn^{35}kɐn^{35}koŋ^{55}tsou^{22}	looking for a job	1		0

APPENDIX F

PHONETIC SYMBOL CHART

A correspondence table showing the Jyutping Romanization Scheme (1993) for Cantonese <http://cpct92.cityu.edu.hk/lshk/>, the WorldBet symbols and other diacritics (the "phones" symbols) that were use in transcribing fusion forms in this dissertation, the ASCII phonetic symbols used in the HKCAC data files, and the corresponding IPA symbols (length distinction of nuclear vowels are indicated (cf. Lee 1999; Zee 1995, 1999a, 1999b)). Correspondence regarding the lexical tones of Cantonese is also given in the table.

Cantonese consonants							
IPA	**WorldBet**	**Jyutping**	**HKCAC (ASCII)**	**IPA**	**WorldBet**	**Jyutping**	**HKCAC (ASCII)**
p^h	ph	p	p^h (pH)	h	h	h	h (h)
p	p	b	p (p)	w	w	w	w (w)
t^h	th	t	t^h (tH)	j	j	j	j (j)
t	t	d	t (t)	l	l	l	l (l)
k^h	kh	k	k^h (kH)	ts^h	tsh	c	ts^h (tsH)
k	k	g	k (k)	ts	th	z	ts (ts)
m	m	m	m (m)	k^{wh}	kwh	kw	k^{wh} (kwH)
n	n	n	n (n)	k^w	kw	gw	k^w (kw)
ŋ	N	ng	ŋ (N)				
f	f	f	f (f)				
s	s	s	s (s)				

Table F.1 Phonetic chart.

Table F.1 continued

Cantonese vowels				Cantonese diphthongs			
IPA	**WorldBet**	**Jyutping**	**HKCAC (ASCII)**	**IPA**	**WorldBet**	**Jyutping**	**HKCAC (ASCII)**
iː	i:	i	(i)	uːi	u:i	ui	ui (ui)
yː	y:	yu	(y)	ei	ei	ei	ei (ei)
uː	u:	u	(u)	ɔːi	>:i	oi	oi (oi)
e	e	i	(e)	ɐi	axi	ai	ɐi (Ai)
o	o	u	(o)	aːi	a:i	aai	ai (ai)
ɛː	E:	e	(E)	iːu	i:u	iu	iu (iu)
œː	8:	oe	(J)	ɛːu	E:u	eu	ɛːu (Eu)
ɵ	ox	eo	(0)	ou	ou	ou	ou (ou)
ɔː	>:	o	(O)	ɐu	axu	au	ɐu (Au)
ɐ	ax	a	(A)	aːu	a:u	aau	aːu (au)
aː	a:	aa	(a)	ɵy	oxy	eoi	ɵy (0y)

Table F.1 continued

Cantonese syllabic nasals				Others			
IPA	**WorldBet**	**Jyutping**	**HKCAC (ASCII)**	**IPA**	**WorldBet**	**Jyutping**	**HKCAC (ASCII)**
m̩	m=	m	m̩ (mc)	ə	&	---	ə (G)
ŋ̩	N=	ng	ŋ̩ (N)	ʃ	S	s	ʃ (S)
				æ	@	---	---
				ɯ	4	---	---
				ts^{j}	tsj	z	---
				$ts^{h\,j}$	tshj	c	---
				ʊ	V[	---	---
				tʃ	tS	---	
				$tʃ^{h}$	tSh	---	
				ʔ	?	---	---
				ɾ	d(	---	---
				ˬ	Hv (breathy voiced of preceding segment)	---	---

Table F.1 continued

Diacritics used in transcription in Experiment 3				
IPA	**Exp.3 Diacritics**	**Jyutping**	**HKCAC (ASCII)**	**Remarks**
ʔ	(modified WorldBet) qq	---	---	Glottalization. Placed after the segment in question (typically vowels). Glottalization in this project is defined by the presence of at least two glottal pulses at the end of the segment in question.
~	~	---	---	Nasalization. Placed after a vowel or a consonant to indicate nasalization of the vowel or the consonant (typically approximant).
ʲ	~	---	---	Approximating palatal place. Placed after the velar or the dental nasal consonants (ŋ/n) to indicate the consonants were produced approximating the palatal place, but not the velar or dental place.
---	(pv)	---	---	Partially voiced. Placed after a voiceless consonant to indicate that the consonant was partially voiced, which was defined as voicing of at least half of the consonant.
---	[ ]	---	---	Target undershot. Placed around the consonant in question to indicate undershoot of the segmental target.
---	0	---	---	Devoiced. Zero was put after a segment indicating devoicing of the segment.

Table F.1 continued

Cantonese lexical tones			
Description	**Chao's tone number (non-checked syllables)**	**(checked syllables)**	**Jyutping**
Level	55	5	1
	33	3	3
	22	2	6
Rising	35	35	2
	23	--	5
Falling	21	21	4

Table F.1 ended

APPENDIX G

TEST SENTENCES FOR EXPERIMENT III

The following is the entire list of test sentences for Experiment III. These were the sentences that were recorded by the experimenter and played to the participants. (Only the Cantonese syllables were recorded and played. Symbols such as I/M/F/<L#> were for the experimenter's information only.) For each sentence in the list, the first line is the Cantonese version, while the second line is the English translation of the target words. Cantonese words without native orthography are written in romanization for the experimenter to know what syllable to be recorded. The participants heard and repeated only the Cantonese version; the English translation is provided for the readers here.

Keys:
I = sentence-initial target word
M = sentence-medial target word
F = sentence-final target word
<L#></L#> = number of syllables involved
The number, if any, that follows the particle in the particle column denotes the lexical tone to be recorded.

sent no.

1 I<L2>嗰個</L2>傻婆 M<L3><L2>突然</L2>間</L3>喺個頭度 F<L4><L2>拗下</L2><L2>拗下</L2></L4>
That funny lady suddenly scratched her head.

2 I<L2>我哋</L2>呢啲初哥 M<L2>投資</L2>起上嚟真係 F<L2>論盡</L2>呀
We the first-time investors are really clumsy in doing investment.

3 I<L2>呢個</L2>nilM<L2>市民</L2>鬧房屋署做嘢 F<L4><L2>拖下</L2><L2>拖下</L2></L4>咁
This citizen scolded the Housing Authority as it procrastinated in doing things.

4 I<L2>但係</L2>一睇就知呢啲 M<L2>行爲</L2>係對抗啦喺度 F<L2>搞嚟</L2>
But just by taking a look at it, one knows such behavior means to protest.

5 I<L2>其實</L2>我真係 M<L3><L2>唔該</L2>晒</L3>你俾我 F<L2>掛名</L2>呀
In fact, I really thank you for letting me register my name.

6 I<L2>唔係</L2>阿明 M<L3>打<L2>電話</L2></L3>俾你就要 F<L2>應承</L2>㗎
It's not that you must say yes whenever Ah Ming calls you up.

7 I<L2>就係</L2>覺得 M<L4>心<L3>有<L2>不妙</L2></L3></L4>先至搵醫生 F<L2>去傾</L2>
It's exactly the lack of peace in the heart that one sought advice from the doctor.

8 I<L2>可以</L2>喺發展科技 M<L3>嗰<L2>方面</L2></L3>去吸引 F<L2>外資</L2>
Can attract foreign capital by developing technology.

9 I<L2>覺得</L2>啲顧客 M<L4>由始至終</L4>都唔敢試吓啲 F<L2>新嘢</L2>嘅 2
Feels like the customers never want to try new things.

10 I<L2>不過</L2>我知道 M<L3>呢<L2>一個</L2></L3>手機爆起上嚟係 F<L3><L2>突然</L2>間</L3>㗎
But I know the blast of this cell phone happened all of a sudden.

11 I<L2>可能</L2>賓賓 M<L4>去<L3>買<L2>外賣</L2></L3></L4>嗰陣見到 F<L3>大<L2>姑奶</L2></L3>呢
Maybe Ben Ben met his/her sister-in-law on his/her way to buy carry-out food.

12 I<L2>咁樣</L2>拎住棵菜喺度 M<L4><L2>賣下</L2><L2>賣下</L2></L4>唔算係 F<L2>投資</L2>架
It is no investment by selling vegetables like this.

13 I<L2>因爲</L2>激嬲咗 M<L4>肥肥大大</L4>嘅老板咪搞到 F<L3><L2>冇嘢</L2>做</L3>囉
As s/he offended the big fat boss, s/he lost the job.

14 I<L2>如果</L2>所有 M<L3><L2>勞工</L2>界</L3>嘅人團結起嚟就 F<L3>差<L2>唔多
</L2></L3>
It would work if all the workers join together.

15 I<L2>或者</L2>香港嘅 M<L4>人之常情</L4>就係睇錢 F<L2>重要</L2>呢
Maybe the common attitude among Hong Kong people is to see that money is important.

16 I<L2>所以</L2>唔係要等到 M<L4>一沉百踩</L4>嘅時候先至 F<L2>發覺</L2>嘅
So it doesn't have to wait to realize it until everybody walks on you when you're down.

17 I<L2>朋友</L2>能夠 M<L4><L3><L2>起得</L2>番</L3>身</L4>係靠好嘅 F<L2>經濟</L2>嘅
That the friend can stand up again is because of the good economy.

18 I<L2>知道</L2>社會變得 M<L4><L2>法治</L2><L2>不張</L2></L4>就唔應該 F<L2>放棄
</L2>
Know that the society is becoming less ruled by law, and one shouldn't give up.

19 I<L2>而家</L2>我 M<L3>上<L2>天堂</L2></L3>狂搵都 F<L3>搵<L2>唔到</L2></L3>啦
Even if now I go find it in heaven crazily, I would not find it.

20 I<L2>自己</L2>扮到 M<L4><L2>鬼五</L2><L2>馬六</L2></L4>咁好 F<L2>辛苦</L2>者 1
It's torturing dressing oneself up like a clown.

21 I<L2>投資</L2>嗰個阿嬸 M<L2>其實</L2>淨係用咗身家嘅 F<L3>一<L2>部份</L2></L3>咋
The lady who invested actually used only a portion of all she had.

22 I<L2>放棄</L2>份工 M<L2>但係</L2>又唔係返屋企湊 F<L3><L2>細路</L2>仔</L3>喎
Though quitted the job, s/he didn't stay home to take care of the children.

23 I<L3><L2>細路</L2>仔</L3>nilM<L2>可能</L2>最鍾意玩 F<L3>打<L2>電話</L2></L3>
Maybe kids like dialing phone most.

24 I<L3><L2>私營</L2>化</L3>係 M<L2>可以</L2>透過 F<L3><L2>互聯</L2>網</L3>嘅
Privatization can be done through the Internet.

25 I<L3><L2>互聯</L2>網</L3>上面可以睇到 M<L2>嗰個</L2>香港 F<L2>先生</L2>喎
Can see that Mr. Hong Kong in the Internet.

26 I<L2>重要</L2>嘅就係 M<L2>呢個</L2>香港嘅 F<L3><L2>基本</L2>法</L3>勒
Important is this Basic Law of Hong Kong.

27 I<L2>唔敢</L2>再 M<L2>咁樣</L2>nilF<L2>以爲</L2>喇
Dare not think like this again.

28 I<L2>經濟</L2>要發展 M<L2>唔係</L2>日日念住 F<L2>以前</L2>嘅
It's no use keep thinking of the past, if the economy is to develop.

29 I<L2>報紙</L2>鬧啲官員 M<L2>不過</L2>又點呢 F<L2>然後</L2>
The newspapers scolded the government officials, but then what's the use?

30 I<L3>一<L2>部份</L2></L3>人唔高興係 M<L2>因爲</L2>佢平日嘅 F<L2>行爲</L2>
Some people got upset because of his/her everyday behavior.

31 I<L2>電腦</L2>今次 M<L2>如果</L2>整得番好就真係 F<L3><L2>唔該</L2>晒</L3>勒
If the computer can be fixed this time, I really have to thank you.

32 I<L2>發覺</L2>瞓覺嗰陣 M<L2>就係</L2>俾呢啲噪音 F<L3><L2>影響</L2>到</L3>
Found that when sleeping, it was this noise that was affecting me.

33 I<L3><L2>基本</L2>法</L3>翻譯完晒 M<L2>所以</L2>咪轉工囉 F<L2>考慮</L2>
The translation of the Basic Law is finished, and so s/he thinks of switching job.

34 I<L2>做嘢</L2>nilM<L2>或者</L2>最吃香嗰行仍然係 F<L2>電腦</L2>
Perhaps the field that makes most money is still computer science.

35 I<L2>考慮</L2>參選 M<L2>我哋</L2>就考慮咗好耐喇 F<L2>雖然</L2>
Although we have been considering participating in the election for a long time,…

36 I<L2>雖然</L2>係 M<L2>朋友</L2>都唔可以立亂洗人 F<L2>做嘢</L2>㗎2
Although we are friends, we cannot make the other one work at our will.

37 I<L2>以前</L2>係人都 M<L2>知道</L2>學英文就要睇 F<L2>報紙</L2>啦
In the past, all knew that one had to read the newspaper to learn English.

38 I<L2>市民</L2>nilM<L2>而家</L2>連食海鮮都 F<L2>唔敢</L2>喇
The citizens don't even dare eat seafood now.

39 I<L3><L2>突然</L2>間</L3>發覺 M<L2>自己</L2>老到乜嘢都 F<L3>唔<L2>記得
</L2></L3>lu3
Suddenly I find that I am too old to remember anything.

40 I<L3><L2>影響</L2>到</L3>政府去將 M<L2>以前</L2>嘅政策改成而家 F<L2>咁樣</L2>
It effected the government to change the old policies to become the present ones.

41 I<L3><L2>預算</L2>案</L3>宣讀嗰陣我 M<L2>發覺</L2>佢瞓著覺喎 F<L2>不過</L2>
But I found that s/he fell asleep when the budget proposal was being read.

42 I<L2>先生</L2>小姐 M<L2>辛苦</L2>減肥嚟冇用呢 F<L2>但係</L2>
But it's no use for the gents and the ladies to lose weight in a hard way.

43 I<L3>呢<L2>一個</L2></L3>nilM<L3><L2>細路</L2>仔</L3>好聰明㗎F<L2>其實</L2>
In fact, this child is very smart.

44 I<L3><L2>冇嘢</L2>做</L3>就 M<L3>差<L2>唔多</L2></L3＞放工喇 F<L2>可以</L2>
Can almost get off from work if there is nothing else to do.

45 I<L3>唔<L2>記得</L2></L3>女人係有 M<L3>越嚟越</L3>靚嘅 F<L2>可能</L2>添
Forget that it's possible for women to become more and more beautiful.

46 I<L3>大<L2>姑奶</L2></L3>有 M<L3>一<L2>部份</L2></L3>麻雀腳坐緊船 F<L3>返<L2>屋企</L2></L3>
Some of the mahjong players of sister-in-law are going back home by ferry.

47 I<L3>嗰<L2>方面</L2></L3>嘅買家話搞 M<L3><L2>私營</L2>化</L3>要靠 F<L2>網絡</L2>㗎
Buyers from that side said that privatization had to be done through network.

48 I<L3><L2>唔該</L2>晒</L3>你喺 M<L2>重要</L2>關頭搵到份 F<L3><L2>預算</L2>案</L3>喎
Thanks for digging out the budget proposal at the critical moment.

49 I<L2>行爲</L2>方面犯錯 M<L2>然後</L2>坐監嘅明星就係 F<L2>嗰個</L2>勒
That is the movie star who committed behavioral misconduct and was then sentenced to jail.

50 I<L2>以爲</L2>請食飯就可以 M<L3><L2>影響</L2>到</L3>大局囉 F<L2>因爲</L2>
Because they thought they could influence the results by buying people meals.

51 I<L2>等如</L2>有份 M<L2>報紙</L2>話呢個世界係冇 F<L2>如果</L2>嘅
It equals to what a newspaper said, 'there is no if in this world'.

52 I<L3>差<L2>唔多</L2></L3>日日都 M<L3>唔<L2>記得</L2></L3>帶假牙囉 F<L2>就係</L2>
Forgot to bring the false teeth almost every day.

53 I<L3>返<L2>屋企</L2></L3>先至玩 M<L2>電腦</L2>吖不如 F<L2>我哋</L2>
How about playing with the computer when we get home?

54 I<L3>搵<L2>唔到</L2></L3>阿旺就 M<L3>返<L2>屋企</L2></L3>咁吓先啦 F<L2>或者</L2>
Perhaps let's go home and take a rest if we can't find Ah Wong.

55 I<L3>越嚟越</L3>識得 M<L2>考慮</L2>美容護膚 F<L3>嗰<L2>方面</L2></L3>喇
Know to pay more and more attention to the aspects of beauty and skin care.

56 I<L3>打<L2>電話</L2></L3>恭喜吓都啱嘅 M<L2>雖然</L2>唔做 F<L2>朋友</L2>
It's right to call to congratulate him/her even if you can't be friends anymore.

57 I<L2>然後</L2>走上五十六樓就 M<L2>等如</L2>鍛煉吓 F<L2>自己</L2>
Then ran up to the 56th floor as if s/he was to train up oneself.

58 I<L2>辛苦</L2>湊大個仔嚟都唔知做乜囉 M<L2>覺得</L2>nilF<L2>而家</L2>
Now s/he doesn't understand why went through the toil to bring up his/her son.

59 I<L3><L2>財政</L2>司</L3>個 M<L3><L2>預算</L2>案</L3>寫得好好吖我 F<L2>覺得</L2>
The budget proposal of the financial secretary is well-written, I think.

60 I<L2>網絡</L2>客戶嘅錢搵極都 M<L3>搵<L2>唔到</L2></L3>囉 F<L2>所以</L2>
So, can't reap the fortune from the Internet customers no matter how.

61 I<L4>實實際際</L4>咁參予咗 M<L3>追趕跑</L3>嘅項目吖我哋 F<L2>知道</L2>
We know they have in fact participated in the items chasing, hurrying, and running.

62 I<L4><L2>鬼五</L2><L2>馬六</L2></L4>嘅 M<L3>大<L2>姑奶</L2></L3>喺度扮 F<L3><L2>財政</L2>司</L3>喎
The wacky sister-in-law is imitating the financial secretary.

63 I<L4><L3><L2>打過</L2>去</L3>問</L4>nilM<L3><L2>財政</L2>司</L3>點解經濟咁 F<L2>蕭條</L2>啦
S/he called to ask the financial secretary why the economy had so much slowed down.

64 I<L3>呢<L2>一家</L2></L3>nilM<L3><L2>互聯</L2>網</L3>搞嚟 F<L3><L2>私營</L2>化</L3>
This Internet provider is working on privatization.

65 I<L4><L2>有朝</L2><L2>一日</L2></L4>nilM<L2>經濟</L2>復甦就幫到 F<L2>市民</L2>勒
One day when the economy revives, it would help the citizens.

66 I<L4><L2>人人</L2><L2>平等</L2></L4>都 M<L2>唔敢</L2>奢望喇 F<L3>越嚟越</L3>
Increasingly, s/he dares not to hope for equality for all people.

67 I<L4>迷迷糊糊</L4>咁 M<L2>做嘢</L2>嘅就係 F<L3>呢<L2>一個</L2></L3>勒
This is the one who does things unconsciously.

68 I<L4><L3><L2>預見</L2>得</L3>到</L4>nilM<L3><L2>冇嘢</L2>做</L3>就神仙啦 F<L2>等如</L2>

If one could predict that there would be nothing to do, one would be a fairy being.

69 I<L4><L2>黃金</L2><L2>時段</L2></L4>講 M<L3><L2>基本</L2>法</L3>嘅人就係 F<L2>呢個</L2>囉

This is the one who talks about the Basic Law during the golden time zone.

70 I<L3><L2>不得</L2>了</L3>啦如果你而家 M<L2>放棄</L2>nilF<L4>去<L3>諗<L2>辦法</L2></L3></L4>

It would be a big problem if you give up thinking up a solution now.

71 I<L4><L2>汽車</L2><L2>隧道</L2></L4>嘅 M<L2>網絡</L2>發展益晒啲 F<L4><L3><L2>搵嚟</L2>工</L3>做</L4>嘅

The development of the tunnel network for vehicles benefits all who are looking for jobs.

72 I<L3>冚唪爛</L3>人都 M<L2>以爲</L2>佢俾阿媽打完就 F<L4><L3><L2>匿咗</L2>入</L3>房</L4>

All people thought that s/he hid himself/herself in the bedroom after being spanked by his/her mother.

73 I<L4>有<L3>咩<L2>聯想</L2></L3></L4>嘅都 M<L4>可免則免</L4>擺喺 F<L4><L2>黃金</L2><L2>時段</L2></L4>啦

Better to avoid placing programs that have any suggestive association in the golden time zone.

74 I<L4><L3><L2>順住</L2>嚟</L3>擺</L4>紅到 M<L4>五顏六色</L4>嘅 F<L4>四<L3>大<L2>天王</L2></L3></L4>

Placing in order the four king singers who are very popular.

75 I<L2>零舍</L2>nilM<L2>論盡</L2>嘅大蝦跌落坑渠仲 F<L4><L3><L2>起得</L2>番</L3>身</L4>喎 3

Big Shrimp who is particularly clumsy could get back up after falling into the drainage.

76 I<L2>糟蹋</L2>完旺財之後就 M<L4><L2>一步</L2><L2>一步</L2></L4>咁 F<L4><L3><L2>走番</L2>出</L3>廳</L4>

After torturing Wong Choi, s/he went back to the living room step by step.

77 I<L4>可免則免</L4>就 M<L4>當<L3>冇<L2>件事</L2></L3></L4>後果 F<L4>可大可小</L4>架

Avoid it the best one can and think it is okay, and the consequences could be big or small.

78 I<L4><L2>拖下</L2><L2>拖下</L2></L4>nilM<L4>嗰<L3>一<L2>類型</L2></L3></L4>嘅人做嘢唔會 F<L4>實實際際</L4>嘅

People belonging to the procrastinating type will not do things practically.

79 I<L4><L3><L2>寫咗</L2>信</L3>去</L4>問 M<L4>四<L3>大<L2>天王</L2></L3></L4>攞簽名又話 F<L2>唔係</L2>

Denied having sent a letter to the four king singers for their signatures.

80 I<L4><L3><L2>代落</L2>咗</L3>去</L4>之後就 M<L4>驚驚青青</L4>咁囉 F<L4>由始至終</L4>都

From the beginning to the end, they were scared after putting in (the stuff in a bag).

81 I<L3><L2>超越</L2>到</L3>nilM<L4>平平穩穩</L4>嘅生活就睇你 F<L4>鐘<L3>唔<L2>鐘意</L2></L3></L4>勒

It depends on you whether you like to break through the stable life.

82 I<L3>鬱金香</L3>nilM<L3>冚唪爛</L3>都係 F<L4><L3><L2>順住</L2>嚟</L3>擺</L4>㗎

All the tulips are arranged in order.

83 I<L4><L2>一步</L2><L2>一步</L2></L4>咁 M<L2>去傾</L2>點樣做到 F<L4><L2>人人</L2><L2>平等</L2></L4>

To talk about how to achieve human equality step by step.

84 I<L4>追追逐逐</L4＞咁喺 M<L4><L2>汽車</L2><L2>隧道</L2></L4>度玩就要 F<L3>避一避</L3>啦

It's necessary to avoid those who are chasing one another in the car tunnels.

85 I<L4>由始至終</L4>都 M<L4><L3><L2>唔見</L2>咗</L3>人</L4>嘅餅舖就係 F<L3>呢<L2>一家</L2></L3>勒

This is the cake shop whose keeper didn't show up from the beginning to the end.

86 I<L2>惡劣</L2>情形裡面 M<L4><L3><L2>站得</L2>住</L3>腳</L4>之後就會 F<L4>輕輕鬆鬆</L4>勒

Being able to stand up under the turmoil and one would feel at ease afterwards.

87 I<L4><L2>賣下</L2><L2>賣下</L2></L4>nilM<L2>麻雀</L2>就轉咗賣 F<L4>柴米油鹽</L4>咯

Sold sparrows at first and then switched to selling firewood, rice, oil, and salt.

88 I<L4>輕輕鬆鬆</L4>咁帶埋啲 M<L4>柴米油鹽</L4>去參加 F<L4><L2>產後</L2><L2>護理</L2></L4>

Brought along at ease the firewood, rice, oil, and salt to attend the after-birth care seminar.

89 I<L2>虛僞</L2>咁 M<L2>應承</L2>我以後唔再 F<L4>亂<L3>拋<L2>垃圾</L2></L3></L4>呀 4

Promised me hypocratically that sh/e won't litter again.

90 I<L4>平平穩穩</L4>咁 M<L2>搞嚟</L2>嗰間茶餐廳係 F<L4>五臟俱全</L4>嘅

The café that is being run stably has everything in it.

91 I<L4><L3><L2>學習</L2>得</L3>到</L4>nilM<L4>去<L3>唔<L2>去街</L2></L3></L4>都答得 F<L4>迷迷糊糊</L4>

Learned to answer vaguely whether or not s/he wanted to go out.

92 I<L3>傾<L2>唔妥</L2></L3>點樣 M<L4><L2>保留</L2><L2>條文</L2></L4>咪 F<L2>沮喪</L2>囉

Failed to negotiate how to retain the stipulations, and so s/he was frustrated.

93 I<L4>一沉百踩</L4>nilM<L3>呢<L2>一家</L2></L3>嘢諗倒都覺得 F<L2>惡劣</L2>呀

It feels nasty even to think of it when everybody hits you when you are down.

94 I<L4><L2>收費</L2><L2>水平</L2></L4>nilM<L3>傾<L2>唔妥</L2></L3>就搞到佢咁 F<L4>歇斯底里</L4>勒

Failed to reach an agreement on the fee to be charged and so s/he became hysterical.

95 I<L4>正正當當</L4>咁 M<L2>掛名</L2>就唔怕唔 F<L3>唔<L2>合法</L2></L3>啦

Get registered appropriately and you don't have to fear for being illegal.

96 I<L4><L3><L2>走番</L2>出</L3>廳</L4>就 M<L2>朦糊</L2>咁見到隻 F<L3>女<L2>家鬼</L2></L3>呀

Saw a female ghost when went back out to the living room.

97 I<L4>偷偷摸摸</L4>咁 M<L4><L3><L2>代落</L2>咗</L3>去</L4>之後就 F<L4>去<L3>買<L2>外賣</L2></L3></L4>

Putting the stuff in sneakily and went out to buy carry-out food.

98 I<L2>麻雀</L2>nilM<L4>鐘<L3>唔<L2>鐘意</L2></L3></L4>唱 F<L4>卡拉 OK</L4>呢

Do sparrows like to sing karaoke?

99 I<L4>去<L3>買<L2>外賣</L2></L3></L4>嗰陣 M<L3>面對面</L3>撞到都 F<L4>情有可原</L4>

It's forgivable if met face to face on the way to buy carry-out food.

100 I<L4>容<L3>有<L2>不當</L2></L3></L4>而又 M<L4><L2>拖下</L2><L2>拖下</L2></L4>咁就 F<L3><L2>不得</L2>了</L3>勒

Let the wrong doings continue and procrastinate in fixing them would make it worse.

101 I<L4>原原本本</L4>咁講 M<L4><L2>產後</L2><L2>護理</L2></L4>嘅常識 F<L3>一<L2>大輪</L2></L3>

Has been talking about after-birth care in its original sense for a long time.

102 I<L2>去傾</L2>點樣 M<L4><L3><L2>預見</L2>得</L3>到</L4>將來發展係 F<L4>平平穩穩</L4>嘅

To talk about how to foresee that the future development is stable.

103 I<L4>去<L3>唔<L2>去街</L2></L3></L4>都答得 M<L2>零舍</L2>nilF<L2>虛僞</L2>嘅

Answered in a particular hypocritical way about whether or not she wanted to go out.

104 I<L3>面對面</L3>咁去買 M<L3>海洛英</L3>就 F<L4>可免則免</L4>勒

Should avoid buying cocaine face to face.

105 I<L2>新嘢</L2>喺 M<L4><L2>黃金</L2><L2>時段</L2></L4>播出係不能 F<L4>容<L3>有<L2>不當</L2></L3></L4>嘅

Should not allow new things that are broadcast during the golden time zone to contain any inappropriateness.

106 I<L3>女<L2>家鬼</L2></L3>nilM<L4>偷偷摸摸</L4>咁食完就 F<L4><L3><L2>唔見</L2>咗</L3>人</L4>

The female ghost disappeared after eating sneakily.

107 I<L4><L3><L2>唔見</L2>咗</L3>人</L4>nilM<L3>一<L2>大輪</L2></L3>都係佢囉 F<L4>嚟嚟去去</L4>

S/he is the one who disappeared for a long while.

108 I<L2>把持</L2>權力即使 M<L3>唔<L2>合法</L2></L3>都好過 F<L4>拖拖拉拉</L4>吖

It's better to hold the authority than to procrastinate, despite that it is not legal.

109 I<L4>嚟嚟去去</L4>都仲係 M<L3><L2>諗嚟</L2>咩</L3>nilF<L3>傾<L2>唔妥</L2></L3>

Back and forth, s/he was still thinking about what made it fail to come into an agreement.

110 I<L4>出出入入</L4>啲人將啲 M<L3>鬱金香</L3>nilF<L4><L3><L2>代落</L2>咗</L3>去</L4>

The people who go in and out put the tulips into it.

111 I<L3><L2>諗嚟</L2>咩</L3>就 M<L4><L2>有碟</L2><L2>話碟</L2></L4>唔洗 F<L4>驚驚青青</L4>

Tell frankly what you think and don't be scared.

112 I<L3><L2>勞工</L2>界</L3>nilM<L3><L2>準備</L2>咗</L3>討論 F<L4><L2>收費</L2><L2>水平</L2></L4>勒

The labor sector is ready to discuss the level of fee to be charged.

113 I<L4>亂<L3>拋<L2>垃圾</L2></L3></L4>之後咪 M<L4><L3><L2>匿咗</L2>入</L3>房</L4>餵 F<L2>麻雀</L2>囉

He hid himself in the bedroom and fed the sparrow after littering.

114 I<L4>五顏六色</L4>咁 M<L4><L3><L2>順住</L2>嚟</L3>擺</L4>啲 F<L3>鬱金香</L3>

Arrange the tulips in order and in a colorful way.

115 I<L4>人之常情</L4>就係 M<L2>把持</L2>唔住都 F<L4>當<L3>冇<L2>件事</L2></L3></L4>

The human nature is to see it nothing even if failed to stand firm.

116 I<L4>驚驚青青</L4>咁 M<L4><L2>動議</L2><L2>辯論</L2></L4>咪俾人丙到 F<L4>五顏六色</L4>囉

Scared in moving the motion and so s/he was scolded badly.

117 I<L2>應承</L2>nilM<L4><L2>有朝</L2><L2>一日</L2></L4>會獎勵 F<L4><L2>個別</L2><L2>人士</L2></L4>嘅

Promise that individuals will be rewarded one day.

118 I<L2>朦糊</L2>咁用 M<L4><L2>中醫</L2><L2>中藥</L2></L4>就包你快啲 F<L3>上<L2>天堂</L2></L3>呀

Use Chinese medical treatment and Chinese medicine without clarifying what they are and I can assure you that you'll get to the heaven soon.

119 I<L4>記<L3>唔<L2>記到</L2></L3></L4>nilM<L2>外資</L2>點解要 F<L4><L2>保留</L2><L2>條文</L2></L4>呀

Can you remember why the foreign investors asked for retaining the stipulations?

120 I<L4>卡拉 OK</L4>雖細而 M<L4>五臟俱全</L4>嘅就係 F<L4>嗰<L3>一<L2>類型</L2></L3></L4>勒

That's the type of karaoke, which, though small, has everything in it.

121 I<L4><L3><L2>沖完</L2>涼</L3>走</L4>嗰位 M<L2>先生</L2>做乜咁 F<L4>偷偷摸摸</L4>嘅

Why did that mister who just took a bath look so sneaky?

122 I<L3>一<L2>大輪</L2></L3>nilM<L2>惡劣</L2>嘅 F<L4>追追逐逐</L4>

A long and nasty chasing.

123 I<L4>情有可原</L4>嘅就係感到 M<L2>沮喪</L2>而冇 F<L2>打算</L2>

It's forgivable when one feels frustrated and has no planning.

124 I<L3>上<L2>天堂</L2></L3>nilM<L4><L3><L2>沖完</L2>涼</L3>走</L4>係唔同嘅 F<L2>零舍</L2>

It feels particularly different by going to the heaven and taking a bath.

125 I<L4><L3><L2>醫番</L2>好</L3>咗</L4>個病就試 M<L2>新嘢</L2>喇 F<L4><L2>有朝</L2><L2>一日</L2></L4>

One day when the illness has been healed, s/he will try new things.

126 I<L2>論盡</L2>嘅肥妹同 M<L2>靈敏</L2>嘅廋仔直情係 F<L4>心<L3>有<L2>靈犀</L2></L3></L4>呀

The clumsy fat gal and the nimble slim guy can really read each other's mind.

127 I<L4>五臟俱全</L4>嘅 M<L3><L2>審計</L2>處</L3>仲喺度諗地方 F<L4>足<L3>唔<L2>足夠</L2></L3></L4>呀 4

The audit department, which already has everything in it, is thinking about whether the department has enough space.

128 I<L4><L2>中醫</L2><L2>中藥</L2></L4>nilM<L4><L3><L2>醫番</L2>好</L3>咗</L4>啲 F<L3>乜鬼嘢</L3>話

What success did the Chinese medical treatment and the Chinese medicine have in treatments?

129 I<L3><L2>審計</L2>處</L3>nilM<L4><L3><L2>寫咗</L2>信</L3>去</L4>之餘又 F<L4><L3><L2>打過</L2>去</L3>問</L4>

The audit department has written and called to ask.

130 I<L3>乜鬼嘢</L3>都話 M<L4>心<L3>有<L2>靈犀</L2></L3></L4>梗係 F<L4>當<L3>唔<L2>準確</L2></L3></L4>啦

Of course it was regarded as inaccurate since s/he attributed everything to the ability of reading each other's mind.

131 I<L3><L2>準備</L2>咗</L3>nilM<L4>出出入入</L4>都要用 F<L4><L2>中醫</L2><L2>中藥</L2></L4>喇

S/he has prepared to use Chinese medical treatment and Chinese medicine when going in and out.

132 I<L3><L2>音樂</L2>會</L3>嗰陣 M<L4>迷迷糊糊</L4>咁聽到你問佢 F<L4>去<L3>唔<L2>去街</L2></L3></L4>嘅

In the concert, I vaguely heard you asking him/her whether s/he wanted to go out.

133 I<L4>四<L3>大<L2>天王</L2></L3></L4>都 M<L4><L3><L2>搵嘢</L2>工</L3>做</L4>令你 F<L4>有<L3>咩<L2>聯想</L2></L3></L4>呀

What would you think when even the four king singers can't find a job?

134 I<L4><L2>動議</L2><L2>辯論</L2></L4>嘅人 M<L4>記<L3>唔<L2>記到</L2></L3></L4>邊個屬於 F<L3><L2>勞工</L2>界</L3>架 3

Can those who move the motive remember who belong to the labor sector?

135 I<L4><L2>產後</L2><L2>護理</L2></L4>個導師 M<L4><L3><L2>打過</L2>去</L3>問</L4>點樣恢復 F<L2>靈敏</L2>

The instructor for after-birth care has called to ask how to resume agileness.

136 I<L4>鐘<L3>唔<L2>鐘意</L2></L3></L4>nilM<L2>虛僞</L2>嘅人喺政壇度 F<L4><L3><L2>站得</L2>住</L3>腳</L4>呀

Do you like to see hypocrites standing firm in the political arena?

137 I<L4><L3><L2>釋放</L2>番</L3>晒</L4>nilM<L3>乜鬼嘢</L3>種類嘅 F<L3>海洛英</L3>話

What kinds of cocaine have been released?

138 I<L4><L2>法治</L2><L2>不張</L2></L4>而 M<L4>容<L3>有<L2>不當</L2></L3></L4>就 F<L4>另作別論</L4>

It's another story when it's no longer governed by law and wrong doings are allowed.

139 I<L4><L2>個別</L2><L2>人士</L2></L4>nilM<L4>有<L3>咩<L2>聯想</L2></L3></L4>都係好 F<L2>朦糊</L2>嘅

Whatever association individuals makes, it is vague.

140 I<L3>追趕跑</L3>嘅項目直情係 M<L2>糟蹋</L2>我哋啲生得 F<L4>肥肥大大</L4>嘅

The items of chasing, hurrying, and running are really torturing to us who are fat and big.

141 I<L4>可大可小</L4>嘅爭拗都係 M<L3>避一避</L3>好過 F<L3>面對面</L3>勒

Arguments, big or small, should be avoided rather than facing them face to face.

142 I<L4>嗰<L3>一<L2>類型</L2></L3></L4>嘅 M<L3>女<L2>家鬼</L2></L3>實會 F<L4><L3><L2>沖完</L2>涼</L3>走</L4>啦 1

That type of female ghost would certainly leave after taking a bath.

143 I<L2>打算</L2>討論 M<L4><L2>收費</L2><L2>水平</L2></L4>嗰陣時要 F<L4><L2>有碟</L2><L2>話碟</L2></L4>

Plan to speak frankly when discuss about the level of fee to be charged.

144 I<L4>歇斯底里</L4>嘅 M<L4><L2>個別</L2><L2>人士</L2></L4>咪繼續 F<L3>追趕跑</L3>囉

The hysterical individuals would continue to chase, hurry, and run.

145 I<L4>當<L3>唔<L2>準確</L2></L3></L4>嘅文件都 M<L4>原原本本</L4>咁交番俾 F<L3><L2>審計</L2>處</L3>勒

Documents that were regarded as inaccurate were sent back to the audit department as they were.

146 I<L4>當<L3>冇<L2>件事</L2></L3></L4>反而仲 M<L4>輕輕鬆鬆</L4>唔係 F<L4>人之常情</L4>

It's not human nature to think nothing has happened and even to relax.

147 I<L4>另作別論</L4>嘅就係當鄭秀文 M<L3><L2>超越</L2>到</L3>王菲嘅 F<L3><L2>音樂</L2>會</L3>

That's another story if Sammi Cheng's concert surpasses Faye Wong's.

148 I<L4>拖拖拉拉</L4>諗住 M<L4>情有可原</L4>我都 F<L4><L3><L2>預見</L2>得</L3>到</L4>

I could foresee that he would procrastinate and wish to be forgiven.

149 I<L4><L3><L2>機械</L2>人</L3>式</L4>到 M<L3><L2>不得</L2>了</L3>棧俾人 F<L2>糟蹋</L2>

Treating things in such a robotic way would deem to be tortured by others.

150 I<L4>心<L3>有<L2>靈犀</L2></L3></L4>nilM<L4>足<L3>唔<L2>足夠</L2></L3></L4>令你哋著到 F<L4><L2>鬼五</L2><L2>馬六</L2></L4>呀

Is being able to read each other's mind sufficient to cause you to dress up in such a wacky way?

151 I<L2>外資</L2>繼續 M<L4>正正當當</L4>咁做生意而唔理人哋嘅 F<L2>污衊</L2>

The foreign investors continued to do their business righteously without paying attention to others' smearing.

152 I<L4>去<L3>諗<L2>辦法</L2></L3></L4>nilM<L2>污衊</L2>香港話呢度 F<L4><L2>法治</L2><L2>不張</L2></L4>

Try to think up some ways to smear Hong Kong by saying that it is not governed by law here.

153 I<L2>污衊</L2>呢個 M<L3><L2>音樂</L2>會</L3>嘅人都唔係 F<L4>正正當當</L4>嘅

Those who try to smear this concert are not righteous.

154 I<L4><L3><L2>起得</L2>番</L3>身</L4>同埋可以 M<L2>開住</L2>架車過 F<L4><L2>汽車</L2><L2>隧道</L2></L4>

Can get up again and can drive the car through the car tunnel.

155 I<L4><L3><L2>匿咗</L2>入</L3>房</L4>再 M<L4><L3><L2>走番</L2>出</L3>廳</L4>就將啲臭氣 F<L4><L3><L2>釋放</L2>番</L3>晒</L4>

Hide oneself in the room and then come back out to the living room, and one can release the odor.

156 I<L4>肥肥大大</L4>嘅肥 BM<L4><L2>拗下</L2><L2>拗下</L2></L4>個頭咁喺度 F<L4>出出入入</L4>

The fat and big Fat B goes in and comes out while scratching his/her head.

157 I<L4><L3><L2>搵嚟</L2>工</L3>做</L4>嘅時候就 M<L4><L3><L2>學習</L2>得</L3>到</L4>搵工要 F<L4><L2>一步</L2><L2>一步</L2></L4>

When looking for a job, one can learn to look for a job step by step.

158 I<L4><L2>拗下</L2><L2>拗下</L2></L4>隻腳板就 M<L4><L3><L2>釋放</L2>番</L3>晒</L4>一直嘅 F<L2>把持</L2>呀 4

Can one release one's stronghold by scratching one's sole?

159 I<L4><L2>保留</L2><L2>條文</L2></L4>去罰 M<L4>亂<L3>拋<L2>垃圾</L2></L3></L4>嘅論點你 F<L4>記<L3>唔<L2>記到</L2></L3></L4>呀

Can you remember the arguments for retaining the stipulation to penalize those who litter?

160 I<L2>靈敏</L2>嘅旺財喺度 M<L4>追追逐逐</L4>令我覺得 F<L4>心<L3>有<L2>不妙</L2></L3></L4>

The agile Wong Choi is chasing and running, which doesn't make me feel at peace in the heart.

161 I<L4>柴米油鹽</L4>要點配都係 M<L4><L3><L2>機械</L2>人</L3>式</L4>咁 F<L4><L3><L2>學習</L2>得</L3>到</L4>咋

Only learned how to use firewood, rice, oil, and salt in a robotic way.

162 I<L3>海洛英</L3>啲買家喺後巷度 M<L4>嚟嚟去去</L4>都唔知 F<L3><L2>諗嚟</L2>咩</L3>

The cocaine buyers were walking back and forth in the alley without knowing what they were thinking about.

163 I<L2>開住</L2>音樂 M<L4>去<L3>諗<L2>辦法</L2></L3></L4>提出 F<L4><L2>動議</L2><L2>辯論</L2></L4>

Turned on the music and tried to think of how to move the motion.

164 I<L3>避一避</L3>嗰班 M<L4>歇斯底里</L4>嘅人唔洗咁 F<L4><L3><L2>機械</L2>人</L3>式</L4>嘅

It doesn't have to act in such a robotic way to avoid that group of hysterical people.

165 I<L4>心<L3>有<L2>不妙</L2></L3></L4>咪 M<L4>實實際際</L4>咁 F<L4><L3><L2>寫咗</L2>信</L3>去</L4>囉

Not feeling at peace in the heart and so practically wrote a letter to ask.

166 I<L2>蕭條</L2>嘅時候啲 M<L4>卡拉 OK</L4>有優惠架 F<L3>冚唪爛</L3>都

All karaokes give discounts when the economy is slow.

167 I<L4><L3><L2>站得</L2>住</L3>腳</L4>就唔怕經濟 M<L2>蕭條</L2>時俾人 F<L4>一沉百踩</L4>啦

Stand up firm and one wouldn't fear being hit by others when the economy goes slow.

168 I<L2>掛名</L2>話要 M<L4><L2>人人</L2><L2>平等</L2></L4>嘅真相露出嚟喇 F<L4>原原本本</L4>咁

The truth of the hypocritical saying that all should be equal has come out in its original form.

169 I<L4>足<L3>唔<L2>足夠</L2></L3></L4>所有人用我都 M<L2>打算</L2>nilF<L3><L2>準備</L2>咗</L3>先

Whether it's going to be enough for all, I plan to get prepared first.

170 I<L2>沮喪</L2>可以 M<L4>另作別論</L4>都唔係將個碳爐 F<L2>開住</L2>掛

Even it can be another story for feeling frustrated, one shouldn't leave the coal oven on.

171 I<L2>搞嚟</L2>嘅計畫 M<L4>可大可小</L4>到連李氏都 F<L3><L2>超越</L2>到</L3>呀
The project in progress that is either big or small has such potential as to surpass the Lee's.

172 I<L4><L2>有碟</L2><L2>話碟</L2></L4>呢個測驗 M<L4>當<L3>唔<L2>準確
</L2></L3></L4>nilF<nil>而唔信都得格
It's okay to take the test "be frank" as inaccurate and do not trust the test.

173 I<nil>你都仲 M<L4>拖拖拉拉</L4>又點會係 F<L4><L3><L2>醫番</L2>好</L3>咗</L4>呢
You are still procrastinating and how could you get healed?

174 I<L3>唔<L2>合法</L2></L3>咁拎啲手巾仔出街 M<nil>nilF<L4><L2>賣下</L2><L2>賣下
</L2></L4>咁
Sell the handkerchiefs illegally on the street.

APPENDIX H

DEBRIEFING MATERIALS


香港廣東話“自然而言”
Hong Kong Cantonese spoken in a natural, casual style.

簡介

☺ 歡迎參與這項研究 ☺

這項研究的目的是要了解香港廣東話的口語特徵，研究結果將有助廣東話語音合成系統的發展。研究共分四個實驗單元，完成後回答一份簡單的語言背景問卷，你的任務便大功告成。整個研究預計需時二至三個小時，而最多不會超過三小時。報酬以小時爲單位，不足一小時亦作一小時計，最多不超過三小時。每小時美金十元，即港幣七十八元。報酬將以港幣發放。

<u>更多有關這項研究的資料</u>

1. 可預見的風險或不安：無。
2. 利益：除了認識他們所說的語言的口語特徵外，實驗參與人並不會獲得任何直接的利益。這項研究的貢獻是可以幫助我們了解廣東話的口語特徵，而研究結果將有助廣東話語音合成系統的發展。
3. 其他研究程序：沒有其他研究程序可爲這個研究項目提供更佳的效益。
4. 保密：我們確保你的姓名將不會以任何形式在研究結果上出現。研究收集得來的數據資料或會讓其他研究員使用，但所有數據資料的保存和分發都會以完全匿名

的方式進行，而不會用上能識別身份的標記。

5. 自願參與：你參與這項研究是自願的；拒絕參與這項研究不會爲你帶來懲罰性的後果，你應得的利益也不會因爲你拒絕參與這項研究而受損。在實驗進行期間，你可隨時收回協議，終止參與實驗，此舉並不會令你權益受損。
6. 聯絡資料：Prof. Mary Beckman, 222 Oxley Hall, 1712 Neil Avenue, Columbus, OH43210, USA. 電話：1-(614) 292-9752；電子郵件：mbeckman@ling.osu.edu。聯絡本地聯絡人，請致電：98662071；電子郵件：pwong@ling.osu.edu。

如有疑問，請隨時發問。

你的熱心幫助，讓我先表感謝!

第一單元　簡介

在這單元，你會透過耳機聽到一些意思完整的廣東話句子。可是，這些句子聽起來並不像我們日常說廣東話時那麼自然和地道，原因是製造這些句子時，電腦只把句子中的字粗糙的拼合起來。我們希望透過你們的參與，爲我們提供自然和地道的廣東話樣本，幫助我們改善這個情況。

你的任務是甚麼？

在這單元中，你的任務簡單而有趣。你會透過耳機聽句子，聽到句子後，便在錄音帶上所提供的頗短的空檔，用廣東話自然地把句子重複一次。怎樣才算自然？你平時怎樣用廣東話和朋友／家人說話，你現在就怎樣用廣東話重複句子，這樣就是自然了。重複句子時，請不要加添或減少句子的字。

舉例：

［你透過耳機聽到・・・］

一班六合彩迷坐冷氣巴去中環置地廣場買朱古力｛空檔｝

［在這頗短的空檔內，你重複・・・］

一班六合彩迷坐冷氣巴去中環置地廣場買朱古力

［你聽到下一句句子・・・］

肥仔今朝出門口嗰陣時唔記得咗帶假牙｛空檔｝

［在這頗短的空檔內，你重複・・・］

肥仔今朝出門口嗰陣時唔記得咗帶假牙

如此類推・・・

第二單元　簡介

在這單元，你會透過耳機再次聽到第一單元那些意思完整的廣東話句子。可是，你今次不用把句子重複一次。

你的任務是甚麼？

在這單元中，你的任務是聽到句子後，便按着句子的內容用廣東話自然地在一個頗短的空檔內回答問題（隨着句子而來的問題可多可少）。就如第一單元一樣，你平時怎樣用廣東話和朋友／家人說話，你現在就怎樣用廣東話回答問題，這就是自然了。

舉例：

［你透過耳機聽到下面的句子：］

一班六合彩迷坐冷氣巴去中環置地廣場買朱古力

［你透過耳機聽到第一條問題：］

嗰班六合彩迷去邊度呀？｛空檔｝

［在這頗短的空檔內，你答：］

嗰班六合彩迷去中環置地廣場

［你透過耳機聽到第二條問題：］

班六合彩迷點去㗎？｛空檔｝

［在這頗短的空檔內，你答：］

佢哋坐冷氣巴去嘅

［你透過耳機聽到另一句句子：］

肥仔今朝出門口嗰陣時唔記得咗帶假牙

［你透過耳機聽到第一條問題：］

咦，肥仔今朝做乜出咗門口又走番轉頭嘅？｛空檔｝

［在這頗短的空檔內，你答：］

佢唔記得咗帶假牙呀

如此類推・・・

第三單元　簡介

你的任務是甚麼？

在這單元中，你的任務十分簡單，就是讀單字。研究員會給你一個字表，你用廣東話把單字讀出來便可以了。[1]

第四單元　簡介

你的任務是甚麼？

在這單元中，你會用你對香港廣東話的語感，把廣東話的詞語結構畫出來。研究員將會給你一套指引，幫助你執行這個任務。[1]

問卷

以上四個單元完成後，回答一份簡單的語言背景問卷，你的任務便大功告成。

[1] This part was not analyzed in this dissertation because I lacked the resources.

Debriefing English translation

☺ Welcome to participating in this research ☺

This research study aims to contribute to our understanding of connected speech processes in Hong Kong Cantonese, and to inform the development of a speech synthesis system for Cantonese. The experiment for this research study has four units, followed by a language background questionnaire. We expect that it will take between 2 and 3 hours for you to finish the study. Payment is based on a per hour basis, i.e. less than an hour is counted as an hour, for up to 3 hours. The payment is US$10 per hour, which equals to HK$78 per hour. Payment will be made in HK dollars.

Additional information about this study:

1. Foreseeable risks or discomforts: None.
2. Benefits: Besides learning about speech production in their native language, participants will not receive any direct benefits. The study will contribute to our understanding of connected speech processes in Hong Kong Cantonese and inform the development of a speech synthesis system for Cantonese.

3. Alternative procedures: No alternative procedures offer greater advantage to the subject.

4. Confidentiality: We assure you that your name will not be associated in any way with the research findings. Data collected from this experiment may be made available to other researchers, but all data will be preserved or distributed in a completely anonymous format – with no personal identification marks.

5. Participation is voluntary: Your participation is voluntary. Refusal to participate will involve no penalty or loss of benefits to which you would otherwise be entitled. You are free to withdraw consent at any time and to discontinue participation in the study without prejudice.

6. Contact information: Prof. Mary Beckman, 222 Oxley Hall, 1712 Neil Avenue, Columbus, OH43210. Telephone: 1-(614)292-9752; e-mail address: mbeckman@ling.osu.edu. For local contact, please call telephone number 98662071, or send an e-mail to pwong@ling.osu.edu.

If you have any questions, feel free to ask anytime.

Thank you in advance for your great help!

Unit One Introduction

In this unit, you will listen to meaningful sentences spoken in Cantonese over the headphones. However, the sentences you will hear are not spoken as natural and native as the way we speak when we speak in Cantonese in our everyday life. The reason is that when making these sentences, the computer only strings words together in a crude way. Through your participation, we hope to obtain natural and native spoken Cantonese samples to help us improve the situation.

What is your task?

In this unit, your task is both simple and interesting. You will be listening to the meaningful sentences spoken in Cantonese over the headphones. After each sentence, repeat the sentence in Cantonese once **in a natural way** in the short blank interval provided on the tape. What is natural? Say the sentence in the way that you would talk to your friends/family in your everyday life and that is natural. When you repeat the sentences, please **do not** add or delete words of the sentences.

Examples:

[You hear over the headphones…]

A group of lottery fans rode an air-conditioned bus to Landmark in the Central District to buy chocolate {blank interval}

[In the short blank interval, you repeat…]

A group of lottery fans rode an air-conditioned bus to Landmark in the Central District to buy chocolate

[You hear over the headphones the next sentence…]

FeiZai forgot to bring his false teeth when he went out the door this morning {blank interval}

[In the short blank interval you repeat…]

FeiZai forgot to bring his false teeth when he went out the door this morning

And so on …

Unit Two Introduction

In this unit, you will hear the same set of meaningful sentences spoken in Cantonese over the headphones that you heard in Unit One again. However, you do not have to repeat the sentences once.

What is your task?

In this unit, your task is to answer questions in Cantonese according to the contents of the sentences. You will answer questions **in a natural way** within the short blank interval provided after each question (the number of questions that follow the sentence may vary). Just as in Unit One, say the sentence in the way that you would talk to your friends/family in your everyday life and that is natural.

Examples:

[You hear the following sentence over the headphones:]

A group of lottery fans rode an air-conditioned bus to Landmark in the Central District to buy chocolate

[You hear the first question over the headphones following the sentence:]

Where did the lottery fans go? {blank interval}

[In the short blank interval, you answer:]

They went to Landmark in the Central District

[You hear the second question over the headphones following the sentence:]

How did they get there? {blank interval}

[In the short blank interval, you answer:]

They took an air-conditioned bus

[You hear another sentence over the headphones:]

FeiZai forgot to bring his false teeth when he went out the door this morning

[You hear the first question over the headphones following the sentence:]

Hey, why did FeiZai come back in after he went out the door this morning? {blank interval}

[In the short blank interval, you answer:]

He forgot to bring his false teeth

And so on …

Unit Three Introduction

What is your task?

In this unit, your task is very simple. It is to read words. The experimenter will give you a word list, and all you have to do is to read aloud the words in Cantonese. [1]

Unit Four Introduction

What is your task?

In this unit, you will use your intuition as native Hong Kong Cantonese speaker to draw the structure of Cantonese words. The experimenter will give you a set of guidelines to help you do the task.[1]

Questionnaire

After finished doing the above units, you will fill out a simple language background questionnaire, and you are all set.

APPENDIX I

AUDIO RECORDING RELEASE FORM

美國俄亥俄州立大學　　草案編號：2003B0102

聲音錄製及使用授權書

本人同意讓研究員在這實驗中錄下本人的聲音。

本人明白在某程度上，說話人的身份可憑其聲音給辨認出來。

本人也明白本人在這實驗中的錄音可能會在試演、演講時播放，又或作實驗刺激物之用，可是，本人的錄音將不會被指名道姓地使用，而使用時亦只會用上很短（三秒或更短）的錄音片段。

最後，本人確認本人已細閱並完全明白這份聲音錄製及使用授權書的內容。本人在不受約束和自願的情況下，在下面簽名作實，並存副本乙份。

日期：_____

姓名：_____

簽名：_____

(實驗參與人)

簽名：_____

(主要研究員或其授權代理人)

English translation Protocol no. ___2003B0102___

Audio recording release form

I consent to have an audio recording of my voice made during this experiment.

I understand that to a certain extent people can be identified by the way their voice sounds.

I also understand that audio recordings made in this research may be played in demos, presentations, or as stimuli in experiments, but that my recorded voice will not be identified by name and that only short segments (3 seconds or shorter) will be so used.

Finally, I acknowledge that I have read and fully understand this release form. I sign it freely and voluntarily. A copy has been given to me.

Date: _____

Name: _____

Signature: _____

(Participant)

Signature: _____

(Principle investigator or his/her

authorized representative)

APPENDIX J

CONSENT FORM

美國俄亥俄州立大學　　草案編號：2003B0102

社會及行爲研究參與人同意書

本人同意參與下列研究：

Syllable Fusion in Hong Kong Cantonese Connected Speech

主要研究員 Mary E. Beckman 教授或其授權代理人已向本人解釋上述研究的目的、程序，及預計所需的時間。主任研究員或其授權代理人並爲本人解釋上述研究可能帶來的效益。如有其他適用於上述研究的研究程序可供選擇，主任研究員或其授權代理人亦已向本人講述。

本人明白上述研究收集得來的數據資料或會讓其他研究員使用，但所有數據資料的保存和分發都會以完全匿名的方式進行，而不會用上能識別身份的標記。

本人同意讓研究員使用本人爲上述研究所錄製的錄音，本人也明白研究員將會無限期保存本人的錄音，並可能在論文演講或／及會議上使用本人的錄音。

本人證實主任研究員或其授權代理人曾予以機會，讓本人就上述研究提出問題，以取得更多與本研究相關的資料。所作的提問，本人已得到完全滿意的答覆。此外，本人明白在實驗進行期間，本人可隨時收回協議，終止參與上述實驗，而此舉並不會爲本人帶來任何懲罰性的後果。

如有疑問，本人可透過電郵 mbeckman@ling.osu.edu 聯絡上述研究的主要研究員，或／及透過郵電 pwong@ling.osu.edu 聯絡上述研究的授權代理人。有關本人作爲研究參與人權益的問題，本人可向下列機構／人士查詢：(i) 美國俄亥俄州立大學研究風險保護處 (Office of Research Risks Protection) (美國電話號碼：1-(614) 688-4792；電郵地址：kelsey.18@osu.edu)；(ii)上述研究的授權代理人(香港電話號碼：9866 2071；電郵地址：pwong@ling.osu.edu)。

最後，本人確認本人已細閱並完全明白這份同意書的內容。本人在自願和不受約束的情況下，在下面簽名作實，並存副本乙份。

日期：_____

姓名：_____

簽名：_____

(研究參與人)

簽名：_____

(主要研究員或其授權代理人)

Protocol no. __2003B0102___ English translation

Consent for participation in social and behavioral research

I consent to my participation in the research entitled:

_____Syllable Fusion in Hong Kong Cantonese Connected Speech_____

The principal investigator, Professor Mary E. Beckman, or her authorized representative has explained to me the purpose of this study, the procedures to be followed, and the anticipated duration of my participation. Possible benefits of the study have been described, as have alternate procedures, if such procedures are applicable and available.

I understand that data collected from this experiment may be made available to other researchers, but that all data will be preserved or distributed in a completely anonymous format — with no personal identification marks.

I consent to the use of audio recordings of my voice. I understand that the investigator will keep the audio recordings of my voice indefinitely, and that the audio recordings of my voice may be used by the investigator in presentations and/or conferences.

I acknowledge that I have had the opportunity to obtain additional information regarding the study and that any questions that I have raised have been answered to my full

satisfaction. Furthermore, I understand that I am free to withdraw consent at any time and to discontinue participation in the study without penalty.

If I have questions, I can contact the principal investigator of this study at mbeckman@ling.osu.edu, and/or her authorized representative at pwong@ling.osu.edu. If I have questions about my rights as a research participant, I can contact the following organization and people: (i) the Office of Research Risks Protection of the Ohio State University (United States) (telephone number in the US: 1-(614) 688-4792; e-mail address: kelsey.18@osu.edu); (ii) the authorized representative of this study (telephone number in Hong Kong: 9866 2071; e-mail address: pwong@ling.osu.edu).

Finally, I acknowledge that I have read and fully understood this form. I sign it freely and voluntarily. A copy has been given to me.

Date: _____

Name: _____

Signature: _____

(Participant)

Signature: _____

(Principal investigator or his/her authorized representative)

APPENDIX K

LANGUAGE BACKGROUND QUESTIONNAIRE

*（*請刪去不適用者。）*

I.　關於實驗參與人

中文姓名：_____　　日期：_____年____月____日

年齡：____　　性別：* 男 ／ 女

出生地點：_____　　居港年期：_____

職業：_____　　教育程度：_____

聯絡電話：_____ (家) _____ (工作) _____ (手提)

請列出你會聽會說的語言和方言	流利程度 (見下面問題 1)

1. 上列各語言／方言中，你個人認爲，在聽和說兩方面，你使用得較自如和較不自如的是哪種（或哪些）語言／方言？請在 "流利程度" 一欄填寫："1" 表示在聽和說兩方面都最流利，"2" 次之，如此類推。如適用，你可給予不同語言／方言相同的流利程度。
2. 你有沒有言語及/或聽覺疾障的紀錄?　* 有 ／ 沒有

II.　關於實驗參與人的父母

1a.　父親

年齡：___

出生地點：_____　　居港年期：_____

你從小時到現在，他是不是一直都用廣東話和你談話？　* 是 ／ 不是

1b.　母親

年齡：___

出生地點：_____　　居港年期：_____

你從小時到現在，他是不是一直都用廣東話和你談話？　* 是 ／ 不是

～ 問卷完畢，謝謝您的幫忙！～

English translation **Linguistic Background Questionnaire**

*（*Delete where inappropriate。）*

I. About the participant

Chinese name：_____ Date：_____Year____Month____Day

Age：_____ Sex：* Male ／ Female

Birth place：_____ Years of residency in Hong Kong：_____

Occupation：_____ Education level：_____

Telephone：_____ (home) _____ (work) _____ (mobile)

Please list the language(s) and dialect(s) you can comprehend and speak	**Level of proficiency** **(see question 1 below)**

1. In the above language(s) / dialect(s), which one(s) do you think you master well and not as well in terms of comprehension and speaking? In the column "Level of proficiency", please write "1" to indicate that you master the language/dialect best in terms of comprehension and speaking, use "2" to indicate that you master the language/dialect second best in terms of comprehension and speaking, and so on. If appropriate, you can give the same rating for different languages / dialects.
2. Do you have record of speech and/or hearing disorder? * Yes / No

II. About the parents of the participant

1a. Father

Age：____

Birth place：_____ Years of residency in Hong Kong：_____

From childhood to now, has he been speaking to you in Cantonese? * Yes / No

1b. Mother

Age：____

Birth place：_____ Years of residency in Hong Kong：_____

From childhood to now, has she been speaking to you in Cantonese? * Yes / No

～ End of questionnaire. Thank you! ～

APPENDIX L

PARTICIPANT RECRUITMENT FLYER

香港廣東話“自然而言”

Hong Kong Cantonese spoken in a natural, casual style.

誠招　研究參與人

本人現正進行博士論文研究，以了解香港廣東話的口語特徵。你的參與將有助廣東話語音合成系統的發展。這研究共有四個實驗單元（主要是聽和說廣東話），以及一份簡單的語言背景問卷。本研究歡迎有興趣而又符合下列條件的同學參加：

- 你在香港出生和長大，父母已經居港十八年或以上，而廣東話是你和你父母說得最流利的語言；
- 你的主修科目不是語言學、中國語文、翻譯；
- 你的年齡介乎 **18** 至 **23** 歲之間；及
- 你現在是香港其中一所大學的本科生

(undergraduate student)（剛在這暑假畢業的本科生也歡迎參加）。

實驗地點：紅磡 香港理工大學 英文系 語音實驗室 **(AG423)**

預計需時：每人 **2** 至 **3** 個小時

現金報酬：報酬以小時爲單位，不足一小時亦作一小時計，最多不超過 **3** 小時。每小時美金 **10** 元，即港幣 **78** 元。參加者將獲發港幣報酬。

查詢詳情或預約時間，請與黃小姐聯絡。

電話： **98662071 (2004** 年 **8** 月 **10** 至 **8** 月 **30** 日**)**

E-mail：**pwong@ling.ohio-state.edu (**任何時間**)**

English translation

香港廣東話“自然而言”

Hong Kong Cantonese spoken in a natural, casual style.

Research Participants Wanted

I am a PhD student working on my dissertation. This research study aims to contribute to our understanding of connected speech processes in Hong Kong Cantonese. Your participation will inform the development of a speech synthesis system for Cantonese. The experiment for this research study has four units (mainly speaking and listening tasks in Cantonese), followed by a simple language background questionnaire. This research welcomes interested parties who meet the following requirements to participate:

- You were born and grew up in Hong Kong, and your parents have been living in Hong Kong for 18+ years. Cantonese is the most proficient language for you and your parents;
- Your major is <u>not</u> linguistics, Chinese language, or translation;
- You are between 18 and 23 in age; and
- You are undergraduate student in one of the universities in Hong Kong (those who just graduated this summer are welcome, too).

Place of experiment: AG423, Speech Laboratory, Department of English, Hong Kong Polytechnic University, Hung Hom, Kowloon.

Expected duration: Between 2 and 3 hours for each participant.

Pay rate (cash): Payment is on a per hour basis, i.e. less than an hour is counted as an hour, for up to 3 hours. The payment is US$10 per hour, which equals to HK$78 per hour. Payment will be made in HK dollars.

For details or to make appointments, please contact Ms Wong.

Phone: 98662071 (August 10 – 31, 2004)

E-mail: pwong@ling.ohio-state.edu (anytime)