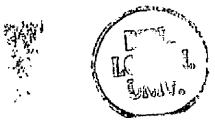
INTONATION IN CANTONESE

Choi-Yeung-Chang FLYNN

Thesis submitted to the
Department of Linguistics
School of Oriental and African Studies
University of London
in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy

2001



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ACKNOWLEDGMENTS

Words cannot express my gratitude to David C. Bennett, my supervisor, for providing the stimulus for this thesis, for his constant encouragement and for so much work in helping me to improve my drafts, without which the work would not have been possible.

I wish to thank Katrina Hayward for her inspiring lectures on phonetics and experimental phonetics which opened the door for my research in this area; and the Phonetics Laboratory at SOAS for allowing me to use the laryngograph recording facilities, the computer programmes such as SPG and the Speech Workstation.

Finally, I would like to thank my husband, Paul, for his unfailing support in every way, especially for cooking excellent Irish dinners while I was kept out late doing research and keeping me laughing when I was mentally exhausted.

ABSTRACT

This thesis develops a system for describing intonation in Cantonese, a language having six phonological tones employing both pitch and slope. It analyses the utterance intonation contour into major intonation units, intonation units and feet. It defines what criteria those units meet and how they relate to each other.

The intonation contours, constructed with a string of lexical tones, are described in terms of prosodic units which separate themselves in terms of pitch height and pitch span. The demarcation of the units is an innovation of the thesis. The different F_0 values of identical phonological tones in an utterance are found to be in gradual descent if they are within an intonation group, and an intonation group is depicted more clearly when the two fitted lines which cover the top and the bottom are parallel and declined. A major intonation group is the largest prosodic unit in utterances. It is decided by a larger size of resetting of pitch span. An intonation group and a major intonation group each represent a unit of information which is semantically and syntactically coherent. The most prominent syllable in an intonation group is the tonic.

An acoustic analysis of all possible combinations of the lexical tones of disyllabic and trisyllabic tonal sequences shows that tonal coarticulation is an important factor in modifying the F_0 contours. The modification can affect both the pitch height and the slope of the F_0 contours, and is also realised in both anticipatory and carryover effects. Prominence is examined, both at the level of words and of utterances, and a description of its prosodic parameters is developed with supporting evidence from the discussion of tonics.

ABBREVIATIONS

T	-	tone
t	-	one cycle of vocal fold vibration
F_0	-	fundamental frequency
CQ	-	closed quotient
Qx	-	closed quotient contour
Lx	-	laryngograph waveform
Sp	-	speech waveform
H	-	high tone
L	-	low tone
h	-	high intensity
l	-	low intensity
Hz	-	Hertz
ms	-	minisecond
F	-	foot
s		stressed
u	-	unstressed
S	-	sentence
NP	-	noun phrase
VP	-	verb phrase
AdjP	-	adjective phrase
AdvP	-	adverbial phrase
ConjP	-	conjunction phrase
IntP	-	interjection phrase
LocP	-	localiser phrase
PrepP	-	preposition phrase
NC	-	noun compound
VC	-	verb compound
AdjC	-	adjective compound
AdvC	-	adverbial compound
ConjC	-	conjunction compound
LocC	-	localiser compound
PrepC	-	preposition compound
PronC	-	pronoun compound
PC	-	particle clusters
n	-	noun
v	-	verb/auxiliary verb
det	-	determinative
adj	-	adjective
adv	-	adverb
int	-	interjection
m	-	measure (including classifier)
loc	-	localiser
p	-	particle
pl	-	plural marker
pron	-	pronoun
prep	-	preposition
q	-	quantifier

CONTENTS

List of Figures

List of Tables

CHAPTER ONE

INTRODUCTION

1.1. AIM AND SCOPE

The thesis describes intonation in Cantonese. Cantonese refers to Guangzhouhua, a branch of the Yue dialect group in China. It is spoken by the populations of Guangdong province along the Pearl River, Hong Kong, Macao, and varieties in Guangxi province and overseas Chinese communities. As the economy in Hong Kong is regarded as the most advanced, the Hong Kong variety of Cantonese spoken by six and a half million people enjoys higher social esteem. The present research deals with Hong Kong Cantonese. Unlike Mandarin, which is widely spoken in China, Cantonese has no standard writing system and no standard spelling system. In terms of the phonemic transcription using IPA symbols, there is no common agreement among linguists who work on Cantonese as to a preferred number of phonemes or tonemes. Section 2 of Chapter 1 is concerned with what phonemes and tonemes the thesis will adopt. Very little work has been done on Cantonese intonation in the past. Nevertheless, those contributions which have been produced are valuable and are reviewed in Section 3 of Chapter 1.

Intonation here is construed as speech prosody expressing linguistic and paralinguistic meaning above word level. Prosody is seen in terms of its phonetic nature, that is, as correlates that involve the acoustic parameters of pitch, duration, and intensity. The phonetic representation of intonation is ultimately realised in the F_0 contour, as we know that the fundamental frequency is produced by the vibration of the vocal folds. The choice of using the F_0 contour as the phonetic representation of intonation in my discussion is motivated by practical considerations. As Pierrehumbert (1980: 12) put it: " F_0 contours can be obtained in quantity with the aid of a computer programme for pitch tracking; articulatory data of

intonation must be obtained by much more difficult and painstaking techniques, such as electromyographic studies of the laryngeal muscles, and tracheal punctures."

The realisation of F_0 contours in utterances results from other factors in addition to intonation, the most important one being lexical tones. Lexical tones in Cantonese have their own inherent pitch and slope, which can be tracked in their F_0 contours. Thus, the first difficulty we encounter is how lexical tones affect each other when they are adjacent. A close look at the characteristics of F_0 contours of lexical tones appears, therefore, to be necessary. Chapter 2 is devoted to investigating the phonetic variation of lexical tones when they occur in different sequential combinations. Another difficulty we encounter depends on how stress affects the F_0 contour in utterances on the assumption that speakers are free to highlight any part of the utterance they like and that such vocal activity will have an effect on the F_0 contours. Before dealing with this question, we have to ask whether there is stress at word level in Cantonese. If yes, what primary acoustic parameter(s) does it use? Chapter 3 presents a detailed account of stress and prominence at word level. After discussing the impacts of lexical tones and lexical prominence on the F_0 contour, the last major aspect in our description of contour is how to delimit larger units than individual syllables in the F_0 contour of utterances. For this, we will use the framework established in Halliday (1967 and 1970). Our first concern is to determine the smallest prosodic unit - rhythmic foot. The second concern is with the demarcation of the larger prosodic units – minor and major intonation units, which contain the tonic prominence. Chapter 4 provides a detailed discussion of the characteristics of F_0 contours when tone and intonation interact in utterances. Chapter 5 ends with a brief conclusion summarising the thesis' findings.

1.2. SOUND SYSTEM OF CANTONESE

1.2.1. Syllable

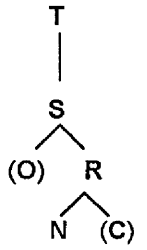
Cantonese is usually described in terms of syllable types. A syllable is divided into an onset, a rhyme and a tone:

a. The onset (O) is optional within a syllable (S);

b. A rhyme (R) consists of an obligatory nuclear (N) and an optional coda (C).

c. Tone (T) is suprasegmental (in this study tone will be referred to by tone number on the right of the rhyme with superscript).

Here is the composition of a syllable:



Almost all morphemes in Cantonese are monosyllabic. One morpheme can be one word and can be represented by one character (where there is one - some colloquial expressions in Cantonese have no written form).

1.2.2. Onset

The onset segment of a Cantonese syllable can be either a consonant (a sonorant or an obstruent) or a glide. The Cantonese onsets are tabulated as follows:

Labials	p	p^h	m	f		
Dental-alveolar	t	t^h	n			l
Alveolo-palatal	tɕ	$tɕ^h$		ɕ	j	
Velars	k	k^h	ŋ			
Labialised-velars	k^w	k^{hw}			w	
Glottal				h		

1.2.2.1. Notes on Cantonese Onsets

1.2.2.1.1. Plosive

I observe that the plosives are sometimes voiced in connected speech, as are the affricates. (The latter is also suggested by Vance 1976: 383).

1.2.2.1.2. VOT

The VOT in aspirated plosives is generally considerably longer than in their unaspirated counterparts. Table 1.1 illustrates the results from Clumect *et at* (1981: 214) and Lisker and Abramson (1964: 384).

Table1.1.The measurements of VOT in ms taken from Clumeck *et al* and Lisker and Abramson.

	/p	p^h	t	t^h	k	k^h/
Lisker & Abramson	9	77	14	75	34	87
Clumeck *et al*	10	74	10.5	83	26	90.8

As shown in Table 1.1, the VOT in velar plosives is approximately three times longer than in bilabials and twice as long as in alveolars. The difference is much less in the aspirated series. The difference in VOT between the unaspirates and aspirates is higher in the bilabials than the alveolars and lowest in velars.

1.2.2.1.3. Glide

The glides /w-/ and /j-/ are sometimes produced with friction (reported by Yuan *et al* (1960: 183) and Gao (1980: 2)). The friction can be spread to the following vowel.

1.2.2.1.4. Sibilant

The sibilants are usually pronounced as alveolo-palatal in my idiolect, similar to IPA [ɕ] and [tɕ] [$tɕ^h$], though alternative symbols are used by other scholars. For example, Jones & Woo choose /ts ts' s/ (1912: xi), Wang chooses /tʃ tʃ ' ʃ/ (1936-37: 641) and Chao /tɕ tɕ' ɕ/ (1947: 28). Regarding the place of articulation, apart from the influence from the following vowel, a given speaker varies the sound from time to time and different speakers vary the sound from each other - the variation is not phonetically conditioned.

1.2.2.1.5. /n-/ and /l-/

/n-/ and /l-/ are not homorganic in a narrow sense. Commonly the nasalised [l̃-] is a realisation for the /n-/ and for the /l-/. Note that not all Cantonese speakers have lost their /n-/, even though the contrast of /n-/ and /l-/ is no longer generally maintained. We are not concerned here with the historical development of /n-/ and /l-/ or with pan-dialectal phonology, but since the distinction of the pair is still used in dictionaries, grammar books and phonological works (including Cheung (1986) who does not recognise /n-/), for convenience I shall continue to use the theoretical /n-/ and /l-/.

1.2.2.1.6. Zero Onset

We have said that the onset is optional. Syllables with initial vowels may be pronounced with initial [ʔ]. Note that this is not obligatory and certainly not before the particle and the prefix /a:/. It may be therefore desirable to consider zero onset as an underlying onset type, potentially realised as glottal stop [ʔ], occasionally realised as a voiced pharyngeal fricative [ʕ] in low tone syllables as, in [ʕa:i^{4}] (to endure). (Setting up a category of zero-initial is a way of adopting traditional Chinese phonology as in Wang (1935-36: 640) and Yuan *et al* (1960: 182).)

The zero onset and the velar nasal onset do not contrast any more. Some words retain the velar nasal (e.g., /ŋɐn^{4}/ (silver) and /ŋɔ:ŋ^{6}/ (stupid)) and some drop it (e.g., /ɐw^{4}/ (cow) and /ɔ:^{5}/ (I, me)). Some words even add a historically unjustified velar nasal before an initial vowel (e.g., /ŋa:n^{3}/ (noon) and /ŋɔ:n^{1}/ (safe)). Again, for convenience I shall continue to use the theoretical /ŋ-/ and /Ø-/.

1.2.2.1.7. Labial-velar

The loss of lip-rounding in the labial-velars /k^{w}/ and /k^{hw}/ before /-ɔ:/ is an on-going sound change and is a process of fusion of two lip-rounded sounds comparable to what happened in the history of the English word 'sword', now pronounced /sɔ:d/ (Robins 1964: 317). This new loss is in evidence in the young generation but some speakers are uneasy about it (e.g., /k^{w}ɔ:k^{3}/ > /kɔ:k^{3}/ (country)). Regarding the sequences /k^{w}u:/ and /k^{hw}u:/, the degree of lip-rounding varies from individual to individual and from word to word, but in general I do not consider that the labial-velar in /k^{w}/ and /k^{hw}/ before the high back rounded vowel has been lost.

1.2.2.2. Examples of Cantonese Onsets

/p/	/pa:^{1}/	father

/p^{h}/	/p^{h}a:^{1}/	procumbent
/m/	/ma:^{1}/	mother
/f/	/fa:^{1}/	flower
/t/	/ta:^{1}/	a dozen
/t^{h}/	/t^{h}a:^{1}/	he
/n/	/na:^{4}/	to fetch
/l/	/la:^{1}/	a particle
/tɕ/	/tɕa:^{1}/	to hold
/tɕ^{h}/	/tɕ^{h}a:^{1}/	fork
/ɕ/	/ɕa:^{1}/	sand
/j/	/ja:^{5}/	also
/k/	/ka:^{1}/	family
/k^{h}/	/k^{h}a:^{1}/	to block
/ŋ/	/ŋa:^{4}/	tooth
/k^{w}/	/k^{w}a:^{1}/	melon
/k^{hw}/	/k^{hw}a:^{1}/	to exaggerate
/w/	/wa:^{1}/	frog
/h/	/ha:^{1}/	prawn
/Ø/	/a:^{1}/	a particle

1.2.3. Rhyme

1.2.3.1. Nucleus

A nucleus can be a long or a short vowel. The distinction between long and short vowels is a prominent characteristic of the Cantonese vowel system. The length distinction is also accompanied by quality differences. The Cantonese vowels are tabulated as follows.

Long vowels	i:	ɛ:	a:	y:	œ:	u:	ɔ:
Short vowels	e	ɐ	ɵ	o			

The corresponding vowel chart is given below: [(1)-(8) indicate the place of Cardinal vowels.]

(1) i: y: u: (8)

(2) e o (7)

ɵ

(3) ɛ: œ: ɔ: (6)

ɐ

(4) a: (5)

1.2.3.2. Coda

The coda segment of a Cantonese rhyme can be either an obstruent or a glide. The coda is exhibited as below:

nasal	m	n	ŋ		
plosive	p	t	k		
glide	w			j	y

1.2.3.3. Rhyme

The following table 1.2 gives an exhaustive list of all the combinations of nuclei and codas in Cantonese. All short vowels have to be followed by a coda.

Table 1.2. Combinations of nuclei and codas in Cantonese.

	Vowel	Glide coda		Nasal coda			Plosive coda		
L	a:	a:j	a:w	a:m	a:n	a:ŋ	a:p	a:t	a:k
S		ɐ j	ɐ w	ɐ m	ɐ n	ɐ ŋ	ɐ p	ɐ t	ɐ k
L	ɛ:		ɛ:w	ɛ:m	ɛ:n	ɛ:ŋ	ɛ:p	ɛ:t	ɛ:k
S		e j				e ŋ			e k
L	œ:					œ:ŋ			œ:k
L	ɔ:	ɔ:j			ɔ:n	ɔ:ŋ		ɔ:t	ɔ:k
S			o w	om		o ŋ	o p		o k
S		ɵ y			ɵ n			ɵt	
L	i:		i:w	i:m	i:n		i:p	i:t	
L	u:	u:j			u:n			u:t	
L	y:				y:n			y:t	
Syllabic				m		ŋ			

1.2.3.3.1. Notes on Cantonese Rhymes

1.2.3.3.1.1. Vowel Values

The vowel values are based on Lee (1983: 108) which accords with my own impressionistic account. Tongue height in vowels is fairly variable, especially for /y:/, /u:/, /a:/ and /ɐ/. The variability of tongue height for /a:/ and /ɐ/ results in a large area of overlap of their F1 - F2 space. This can be seen in Lee (1983) where the F1 and F2 of eleven vowels are measured. (Alternative phonemes were chosen by other scholars (e.g., Jones and Woo 1912, Wang 1935-36, Wong 1941, Zee 1993 and Bauer and Benedict 1997, etc.): /e/, /i/, /ɛ/, /ɪ/ and /e^{j}/ were used for describing the short vowel which is [- low, - back, - round]; /u/, /o/, /ʊ/ and /o^{w}/

were used for describing the short vowel which is [- low, - front, + round]; /œ/, /ø/ and /ɵ/ were used for describing the short vowel which is [-high, - low, - back, + round]; /a/, /ɐ/ and /ɑ/ were used for describing the short vowel which is [- high, - mid, - round].)

1.2.3.3.1.2. Length

1.2.3.3.1.2.1. Vowel

The length of long vowels is approximately twice as long as that of short vowels. This is taken from the measurements of Kao (1971: 49) and Lee (1985: 31), who both suggest that duration difference is the primary cue to distinguish Cantonese vowels. Despite their different phonetic environments among the test tokens, the measurements are very similar. Table 1.3 illustrates vowel duration in different types of codas in ms according to the results taken from two works (from Lee I take his Hong Kong informants only).

Table 1.3. The measurements of vowel length taken from Kao (1971: 49) and Lee (1985: 31).

	Long vowels	(Kao/Lee)	Short vowels	(Kao/Lee)
open vowel:	V:	(308/283)		
with nasal:	V:N	(203/193)	V N	(100/93)
with plosive:	V:P	(169/146)	V P	(89/75)

Their results coincide with Chao's impressionistic account (1947: 23) - a vowel before a plosive coda is naturally shorter than before a nasal, the relative difference in vowel length being still maintained (this is further confirmed by an experiment conducted by Fok (1974: 21)). However, I consider that it would be misleading to conclude that the long-short distinction is the primary distinction in Cantonese vowels, if for no other reason than the length of vowels depends on vowel quality. More significantly, the duration of a vowel can be stretched or shortened but the vowel quality is maintained.

1.2.3.3.1.2.2. Nasal Coda

The nasals are in a pattern of long and short according to whether they follow a short or long vowel. These make up a complementary distribution: long nasal follows a short vowel, and vice versa. Long and short are described by Chao (1947) as strong and weak, respectively.

Thus there is a long-short duration difference and also a concomitant strong-weak energy difference.

1.2.3.3.1.2.3. Rhyme

Long vowels can be twice as long as short vowels within the same type of rhymes, and the codas have a long-short pattern depending on the short or long vowels they follow (following Chao 1947 and Zee 1993:98). I consider that this kind of complementary distribution of long-short or short-long between vowels and coda is a characteristic of Cantonese rhymes.

1.2.3.3.1.3. Plosive-coda

The plosives are unreleased. The plosive articulation is held tightly for a reasonably long time. At the same time the glottis also closes tightly. As a result of this, they are usually realised as [ʔp], [ʔt], [ʔk]. They are labelled as glottalized stop in Fok (1974: 8). Cheung also observes the characteristic of the glottal closure of these codas (1986: 193).

1.2.3.3.2. Examples of Cantonese Rhymes

/a:/	/a:^{1}/	a particle
/a:j/	/a:j^{1}/	to lean on
/a:w/	/a:w^{1}/	to scratch
/a:m/	/ka:m^{1}/	jail
/a:n/	/ka:n^{1}/	middle
/a:ŋ/	/ka:ŋ^{1}/	to plough
/a:p/	/a:p^{3}/	duck
/a:t/	/a:t^{3}/	to press
/a:k/	/ka:k^{3}/	to separate
/ɐ j/	/ɐj^{2}/	short
/ɐ w/	/ɐw^{1}/	Europe
/ɐ m/	/ɐm^{1}/	nunnery
/ɐ n/	/kɐn^{1}/	root
/ɐ ŋ/	/ɐŋ^{1}/	oriole
/ɐ p/	/kɐp^{1}/	hurry
/ɐ t/	/pɐt^{1}/	pen
/ɐ k/	/pɐk^{1}/	north
/ɛ:/	/ɕɛ:^{1}/	some
/ɛ:w	/pɛ:w^{6}/	to jostle with the hip
/ɛ:m/	/lɛ:m^{2}/	to lick

/ɛ:n/	/wɛ:n^{1}/	mini-bus
/ɛ:ŋ/	/t^{h}ɛ:ŋ^{1}/	lounge
/ɛ:p/	/kɛ:p^{6}/	to grip
/ɛ:t/	/p^{h}ɛ:t^{1}/	buttocks
/ɛ:k/	/tɕɛ:k^{3}/	a measure
/e j/	/pej^{1}/	humble
/e ŋ/	/jeŋ^{1}/	hero
/e k/	/jek^{1}/	benefit
/œ:/	/hœ:^{1}/	boot
/œ:ŋ/	/hœ:ŋ^{1}/	fragrant
/œ:k/	/kœ:k^{3}/	foot
/ɔ:/	/ɔ:^{1}/	to defecate
/ɔ:j/	/ɔ:j^{1}/	sad
/ɔ:n/	/ɔ:n^{1}/	safe
/ɔ ŋ/	/kɔ:ŋ^{1}/	river
/ɔ:t/	/hɔ:t^{3}/	to drink
/ɔ:k/	/ɔ:k^{3}/	ferocious
/o w/	/kow^{1}/	tall
/o m/	/mom^{1}/	meal (baby talk)
/o ŋ/	/koŋ^{1}/	work
/o p/	/pop^{4}/	heart-beat kind of sound
/o k/	/ok^{1}/	house
/ɵ y/	/kɵy^{1}/	to reside
/ ɵ n/	/tɕ^{h}ɵn^{1}/	spring
/ɵ t/	/tɕ^{h}ɵt^{1}/	to go out
/i:/	/ji:^{1}/	clothes
/i:w/	/ji:w^{1}/	waist
/i:m/	/ji:m^{1}/	flood
/i:n/	/ji:n^{1}/	smoke
/i:p/	/ji:p^{6}/	leaf
/i:t/	/ji:t^{6}/	hot
/u:/	/wu:^{1}/	black
/u:j/	/wu:j^{4}/	return
/u:n/	/wu:n^{2}/	bowl
/u:t/	/wu:t^{6}/	alive
/y:/	/jy:^{1}/	in
/y:n/	/jy:n^{1}/	injustice
/y:t/	/jy:t^{6}/	moon
/m/	/m^{4}/	not
/ŋ/	/ŋ^{5}/	five

1.2.4. Tone

Cantonese is known for having a rich tonal system in terms of quantity: the tones can be classified from six to eleven depending upon the criteria used. We dismiss the over-differentiated system and adopt the six tones system here:

Table 1.4. Tonal code, tone shape and tone letter of the six tones.

Tonal code	T1	T2	T3	T4	T5	T6
Tone shape	high-level	high-rise	mid-level	low-fall	low-rise	low-level
Tone letter	˥^{55}or˥˧^{53} or˥^{5}	˧˥^{35}	˧^{33}or ˧^{3}	˨˩^{21}or ˩^{11}	˩˧^{23}	˨^{22} or ˩^{2}

1.2.4.1. Notes on Cantonese Tones

1.2.4.1.1. Tone Inventory

Other than our six-tone system, there are other views on the tone inventory. Let us start with traditional Chinese phonology where nine tones are recognised.

Table 1.5. Tonal inventory in traditional Chinese phonology.

	/p^{h}eŋ^{4}/ (Even)	/ɕœ:ŋ^{5}/ (Rising)	/hɐy^{3}/ (Going)	/jɐp^{6}/ (Entering)
/jɐm^{1}/	˥ T1	˨˧ T2	˧ T3	˥ T7 /ɕœ:ŋ^{6}/ (high) ˧ T8 /tɕoŋ^{1}/ (mid)
/jœ:ŋ^{4}/	˩˧ T4	˩ T5	˩ T6	˨ T9 /ha:^{6}/ (low)

The traditional tone labels in Chinese phonology are divided into two registers: /jɐm^{1}/ and /jœ:ŋ^{4}/ - indicating high and low respectively. Each of them possesses four basic tones: /p^{h}eŋ^{4}/, /ɕœ:ŋ^{5}/, /hɐy^{3}/ and /jɐp^{6}/, /jɐm^{1} jɐp^{6}/ is further divided into /ɕœ:ŋ^{6} jɐp^{6}/ and /tɕoŋ^{1} jɐp^{6}/. /jɐp^{6}/ tones only occur in plosive coda syllables. In other words, the /jɐp^{6}/ group contrasts with others in terms of the duration of the syllable, not in the pitch height. The tone-letters indicating the tone values for the traditional tone terms are taken from Wang (1936-37: 651). Ultimately the scheme of representing tones as a sequence of digits

indicating the starting pitch and final pitch is from Chao (1930: 24) and is now adopted in IPA. I have set out the conventional tone codes above for the purpose of referencing the traditional terms in our discussion below.

We do not recognise T7, T8 and T9 in our tone inventory. They are equated with T1, T3 and T6 respectively in pitch. The plosive-codas are actually homorganic counterparts of nasal-codas, i.e., /m/ and /p/, /n/ and /t/ and /ŋ/ and /k/. They differ from their counterparts segmentally rather than suprasegmentally. They are shorter than their non-plosive-coda-tone counterparts as discussed in the section on vowels. Therefore, in this study the tones with the plosive-coda will not be treated as a separate category just because of their particularly short length, although there is a logic to the traditional analysis when one considers that the 'tonal' distinctions of languages such as Thai and Vietnamese do not depend solely on pitch. They are treated as variants of T1, T3 and T6 respectively.

1.2.4.1.2. Tone Value

The tone values are adopted from Fok's experiments (1974) and Vance's computation (1977) (the latter's data is taken from Hashimoto 1972) which accord with the acoustical recordings I have made, except for T1. I consider T1 is realised as level more than as a fall. (Note that Fok and Hashimoto's measurements were made in the early sixties and early seventies.) This can be confirmed by Chan (1987), who observes that in Hong Kong today, a number of young people simply have a high level for T1 in all contexts and that T1 behaves like a level tone in the interaction of tone and melody (Chan 1987: 2). T1 is level in general and has a fall variant in addition, while T4 is falling in general and has a level variant in addition. The interchange of level and fall is not phonetically conditioned and there is no sandhi rule governing it. It varies from person to person or is an intrapersonal variation.

Regarding the actual tonal pitch of the six tones: Hashimoto (1972) has a different tonal pitch description from the other works. Her tonal pitches indicated by tone letters are crowded on the upper part of the scale while Chao's (1947) and others are crowded at the bottom. Vance (1977: 96) averages F_0 values of peaks and dips taken from the isolation-

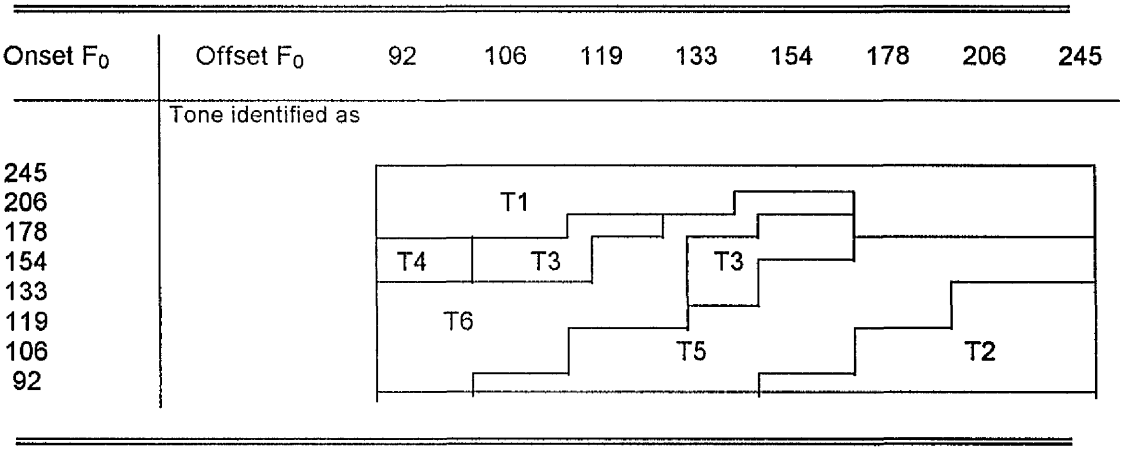
form tokens measured by Hashimoto (1972: 122) and sums up the average values converted to the five-point scale. He then concludes that Chao's description reflects F_0 more accurately. We now summarise his conversion in the following table 1.6:

Table 1.6. A summary of Vance's (1977: 96) F_0 values of peaks and dips of the six tones which are converted from Hashimoto's (1972: 122) data.

T1	T2	T3	T4	T5	T6
5.0-3.2	2.4-4.4	3.3-2.8	1.9-1.0	2.2-3.1	2.9-2.4

Vance further conducts a perceptual experiment using 64 pitch patterns – all combinations of eight F_0 values as the starting pitch with the same eight F_0 values as the end pitch – on a synthetic version of the syllable /jɐw/. Here are his identification results. The value in each case is the one indicated by the largest percentage of subjects (over fifty percent in each case except in the case of T4 as indicated in the caption of the following figure 1.1.)

Fig. 1.1 A summary of the response data from Vance's (1977: 96) identification results of a synthetic version of the syllable /jɐw/. Stimuli on which there was over fifty percent agreement (except T4 which was identified only nine times out of twenty) (cf. Cheung 1986).



Vance concludes that the important cue to distinguish T2 and T5 is the gradiant of the contour; within level tones, the highest pitch is perceived as T1, the mid as T3 and lowest as

T6; T1 does have two forms - level and falling; the end-point of tones is more important than the starting point for tone-identification.

Fok also conducts a perceptual experiment which shows that F_0 is the primary cue for tone-identification by comparison with duration and amplitude (she excludes T7, T8 and T9 in this experiment). Once the F_0 information is removed, tone cannot be named at all though duration and amplitude information remained (1974: 87).

However, scholars differ as to their ideas regarding tone value. A summary of disagreements regarding numbers and value and shape of tones is listed in table 1.7.

Table 1.7. Disagreements as to numbers and value of tones among modern works.

	T1	T2	T3	T4	T5	T6	T7	T8	T9
Jones *et al* (1912:xiiv)	upper falling	upper rising	upper level	lower falling	lower rising	lower level	-	-	-
Chao (1947:24)	53/55	35	33	21	23	22	5	33	2/22
Wong (1941:69)	52/55	25	33	11/1	23	22	5	3	2
Kao (1971:84)	55/53	35	33	21	23	22	-	-	-
Hashimoto (1972:92)	53/55	35	44	21/22	24	33	5	4	3/33
Fok (1974:12)	53/55	35	33	21	23	22	5	33	2
Vance (1977:95)	55	35	33	11	13	22	-	-	-
Rao *et al* (1981:275)	53/55	35	33	11	13	22	55	33	22
Ching (1981:44)	55	35	33	21	23	22	5	3	2
Kwok (1984: 24)	55	45	44	21	34	33	5	4	3
Cheung (1986:182)	44/42	24	33	22/11	23	22	-	-	-
Kong (1987:385)	HL	HR	ML	LL	LR	M-LL	-	-	-
Bauer&Benedict(1997:119)	55/52	25	33	21	23	22	5	33	22

[Note: Jones *et al* use musical notes on top of the labels given above; Wong interprets Jones *et al*'s musical notes with straight-lines into a stave (the description above are interpreted by me); Cheung uses his own four-point scale; the others use Chao's five-point scale; Kong declares he uses Chao's scale but does not indicate any details with numbers.]

1.2.4.1.3. Tone Shape

All the level tones, namely, T1, T3 and T6 (with their short variants) are observed to trail off in a slight falling contour. The two rising tones, namely, T2 and T5 are observed to have a dip after the onset before the rise. Phonetically, they show a large rise following a short fall. The frequency range of the fall variant of T1 and the rising T2 is twice as great as T4 and T5 respectively, according to Fok (1974: 26). T4 is observed to have a varied degree of falling shape and to be accompanied by creaky voice for some speakers. There is a confusion between T3 and T5 in certain words, particularly /ɑi:/, according to Fok (1974) and Vance (1976). However, since most of the words maintain their contrast, it is appropriate to maintain the distinction.

Regarding T1 and T4, there is disagreement as to whether these two tones illustrate one or two values. Jones and Chao were excellent phoneticians and claim to perceive two values (Jones *et al* 1912 and Chao 1947). Fok (1974: 27) and Gandour (1981: 26-27) claim to prove two values in their experiments. Bauer and Benedict (1997: 124) consider that T1 is falling in nature but becomes level for Hong Kong speakers while with Guangzhou speakers it remains falling for most words. Other scholars, including Rao *et al* (1981: 277) have found that in T1, most words are falling, only a few words are level. Chao (1947: 27) considers that there is a sandhi rule governing the change between the high falling and high level, which can be summarised as: 53>55/_53/55/5. Hashimoto (1972: 187) formulated a sandhi rule to describe the change between falling and level applying to both high and low, in which the rule regarding the low can be summarised as: 21>11/_21.

1.2.4.1.4. Intensity

"No regular overall intensity pattern can be assigned to any one tone in particular and there is no evidence for a certain tone to be lower or higher in intensity than the rest of the tones. However, T1 is the loudest and T4 the softest." This is adopted from Fok (1974: 24).

1.2.4.1.5. Length

Studies do not assign standard lengths to tones. Vance suggests that vowel length apparently does not have a role in differentiating the Cantonese tones (T1 to T6) (1976: 383). He reached this conclusion after an experiment in which he placed a test word /ɕi:/ (/si:/ in his notation) in different positions in utterances – medial, final and contrastive context. Kong (1987) also looks at the influence of tones upon vowel duration in Cantonese. He employs three informants to pronounce /ɕi:/ (/si:/ in his notation) with six tones in a frame sentence and concludes that the vowel duration is different under the influence of tones: T2 (272ms) is the longest, then T3 (264ms), T5 (260ms), T6 (257ms), T1 (236ms) and T4 (229ms) in descending order. In other words, rising tones (T2 and T5) are longer than others; with the other tones, the one at the mid point of the frequency range (T3) is the longest and the two at both ends of the frequency range (T1 and T4) are the shortest. He considers the results are significant. His claim does not agree with Fok's, which shows that T4 is the shortest but T1 or T5 is the longest, taking the voiced segments in /ɕeŋ/ (/si:ŋ/ in her notation) pronounced in isolation by one male and one female speaker (1974: 24-25). Among the plosive-coda tones, T8 is the longest (200ms) and T7 the shortest (150ms). The difference between T1 and T4 was 230ms for the male speaker and that between T5 and T4 200ms for the female speaker. However, Fok considers that the duration of each of the normal tones (T1-T6) or each of the plosive-coda tones (T7, T8 and T9) seems to be arbitrary: the length of the tones fluctuates but in no regular pattern (1974: 34).

1.2.4.1.6. /pi:n^3 jɐm^1/ (Changed-tone)

In addition, there are two so-called /pi:n^3 jɐm^1/ (changed-tone or modified-tone) which are listed as a separate tonal category in some works (e.g., Chao 1947, Wong 1941, Bauer and Benedict 1997, etc). The high-rising /pi:n^3 jɐm^1/ has a high-rising contour as in T2 and is considered to have a higher pitch at the end (e.g., Yuan *et al* 1960). Chao perceives it as having a lower start and uses the pitch pattern of 25 for it. It is usually marked as T2*. The high-level /pi:n^3 jɐm^1/ is considered to have a higher pitch than T1 and is usually marked as T1*. /pi:n^3/ means `to derive', /jɐm$^1$/ means `the tone'. It means a changed-tone of a word

possessing a new meaning is derived from its historical cognate which is the underlying form (and can be of any tone). Usually the underlying form and the changed-tone share the same character. For example:

/pi:n^{3} jɐm^{1}/	Derived from
/ji:n^{1*}/ (smoke of tobacco)	/ji:n^{1}/ (smoke in general)
/t^{h}ɔ:ŋ^{2*}/ (candy)	/t^{h}ɔ:ŋ^{4}/ (sugar)

Hashimoto (1972) and Rao *et al* (1981) provide a detailed discussion regarding the morphological and syntactical environment under which the /pi:n^{3} jɐm^{1}/ occurs. Whitaker (1955-56) and Wong (1981 & 1982) also wrote at length about the /pi:n^{3} jɐm^{1}/ phenomenon. Bauer and Benedict (1997) follow Chao's view and set up a separate tonal category for /pi:n^{3} jɐm^{1}/ besides the nine tones. They claim that the contour of the high-level /pi:n^{3} jɐm^{1}/ is considerably higher than that of any other tone contour (p.142). The contours of the high-rising /pi:n^{3} jɐm^{1}/ and the regular T2 are identical (p.143).

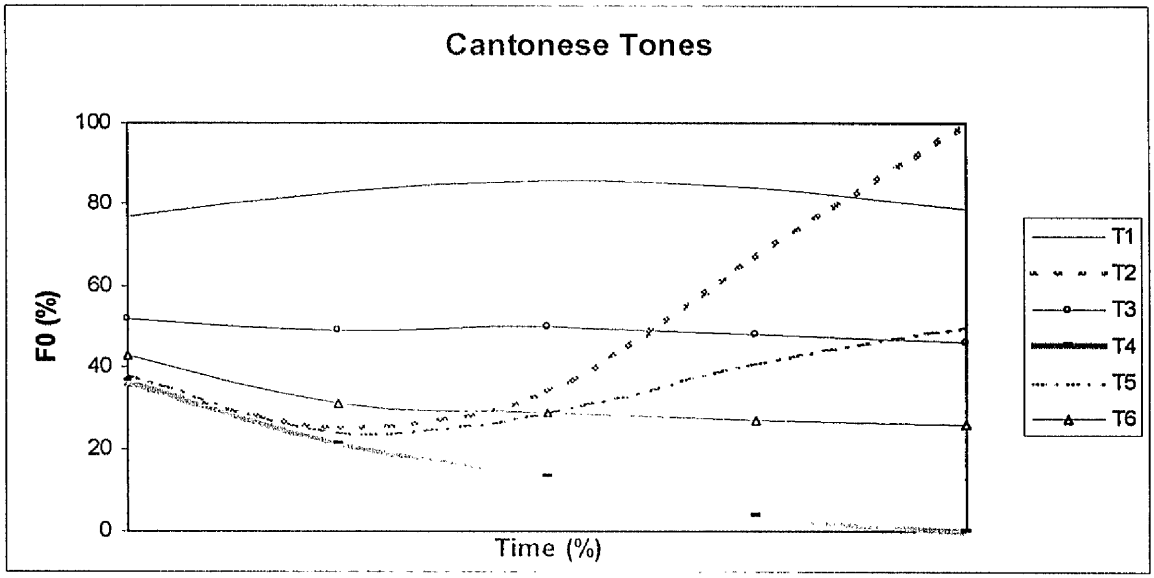
A study of /pi:n^{3} jɐm^{1}/ (changed-tone) would take us into considerations of morphology, etymology and derivation from historical cognates. We are not directly concerned with these matters here. They are psychologically perceived with a higher pitch: the high-level changed-tone is higher than T1 and the high-rise changed-tone is higher than T2. However, we are only concerned with their material pitch movement. First, are they higher in pitch in all environments? If we read through all the examples listed in Wong (1941) or Rao *et al* (1981), etc., we will find that the changed tone occurs in the last syllable of a disyllabic word, the last syllable being given more time to complete the level or the rise contour than the first syllable. Second, are they actually higher in pitch than their regular T1/T2 counterparts as minimal pairs? I find no instrumental work done on this, except Bauer and Benedict (1997: 250-265) which shows no evidence of a higher peak or a lower onset for the high-rising /pi:n^{3} jɐm^{1}/ than the regular T2 in the contours obtained from /ji:^{2}/ (chair) (p.264) and /ji:^{2*}/ (the number two servant) (p.265), nor a higher pitch for high-level /pi:n^{3} jɐm^{1}/ than the regular T1

from /ji:^{1}/ (doctor) (p.268, 270-1) and /ji:^{1*}/ (auntie) (p.266). As the changed-tones do not contrast with the regular T1 and T2, they receive no special treatment, i.e. do not have a separate tonal category in this study.

1.2.4.2. Examples of Cantonese Tones

T1	/ji:^{1}/	aunt
T2	/ji:^{2}/	chair
T3	/ji:^{3}/	meaning
T4	/ji:^{4}/	son
T5	/ji:^{5}/	ear
T6	/ji:^{6}/	two

Fig. 1.2. The six Cantonese tones. Mean F_0 contours of the 6 tones pattern in monosyllables produced by 3 informants (details see next chapter), on the set of words listed above. Each syllable is normalised for duration on a percentage scale and F_0 contours are also normalised for each of the speakers.



1.3. SURVEY OF LITERATURE ON CANTONESE INTONATION

Research on intonation in Cantonese is not a virgin field but did not start until recent years. The bibliographical works of Yang (1974) and (1981), Yuyan Yanjiusuo (1978) and (1983), Lucas (1985) and Cheng (1993) deserve special mention as they have provided me with

some very helpful source references. Major contributions are found in Barnett (1949-1950), Cheung (1972), Cheung (1986), Kwok (1984), Hashimoto (1972), Fok (1974), Vance (1976), Vance and Walker (1976), Kwok and Luke (1986), Johnson (1987), Bauer and Benedict (1997). The first four works mentioned above can be grouped under the heading of the phonemic approach and in addition Cheung (1972) and Kwok (1984) represent the syntactic description and Barnett (1949-1950) and Cheung (1986) represent the phonological description of the language, while the others can be categorised as following a prosodic approach.

1.3.1. Cheung (1972) and Kwok (1984)

Cheung (1972), a Chao follower, quotes the six forms of /hɐu^{2} hej^{3}/ (mouth-air) 'tone' (note the term here refers to manner of expression in speech, e.g., a serious tone) proposed by Chao (1926) in the study of northern dialects and extends the discussion of one of the forms - particles. (The rest of the six forms are: substantial words, adverbs or conjunctions, inflections, interjections and intonation.) He mentions that Chao used to treat two forms of intonational endings as part of Chinese sentence intonation, but later found it better to treat them as particles (Chao 1968: 182). The two forms of intonational endings – rising ending and falling ending, have no segmental phoneme in their own right but have to reside parasitically on the last morpheme by prolonging it for the length of a neutral-tone syllable. In the discussion of particles, Cheung considers that particles are function words but they cannot be independent as they can only appear at the ends of a sentence or phrase (p.169). They are start-bound and end-free. He goes on to propose a question intonation: /35:/ in Cantonese. Two examples are given:

	Statement	Statement + Question intonation /35:/
a.	/ɕek^{6} fa:n^{6}/ (eat-rice.)	/ɕek^{6} fa:n^{6}/ + /35:/ (eat-rice?)
b.	/tɕɐn^{1} kɛ:^{3}/ (really-particle - Really.)	/tɕɐn^{1}kɛ:^{3}/ +/35:/ (really-particle - Really?)

He explains that the tone value of the syllable preceding the added question intonation (i.e., the syllable /fa:n^{6}/ or / kɛ:^{3}/) is changed once the /35:/ intonation is superimposed on it. The

sentence particle /kɛ:^{3}/ in the second example has two forms in the two sentences, and it is not necessary to regard them as two different particles (p.170).

However, he decides not to go any further in the discussion of the intonation of modal particles simply because there is no previous work done in the field of Cantonese intonation. Even so, in the analysis of the function of particles, he cannot help touching on intonation on two occasions:

(a) Sometimes the variation of tempo or loudness (i.e. quick or slow, strong or weak) of a particle would give a different meaning to it, say /pɔ:^{3}/ (p.175). Unfortunately, he does not include this feature in his examples.

(b) A rising or falling intonation is superimposed on the sentence particle, whereby the mood is changed (p.194). Examples are:

/ɕek^{6} fa:n^{6} kə lə a:/	(eat-rice-particles)
/ɕek^{6} fa:n^{6} kɛ:^{3} la:^{1}/	(I think it is time to have a meal.)
/ɕek^{6} fa:n^{6} kɛ:^{3} la:^{3}/	(Have a meal, please!)
/ɕek^{6} fa:n^{6} kɛ:^{3} la:^{4}/	(Is it time to eat now?)

He notes that the sentence particle here has no fixed tone and does not belong to anywhere in the toneme inventory; its change upward or downward pitch movement would depend on the influence of intonation. Particles are usually characterised as a light tone where the vowel has a tendency to reduce to a schwa (p.170).

Cheung (1972) is the first I know of to have touched on the field of Cantonese intonation. Nevertheless, he exhibits a fair understanding of the correlations between intonation, tone, sentence final particle and the final syllable modification in his brief discussion, which has presumably had an influence on Cheung (1986) even though the latter makes no reference to his work.

Whereas Cheung (1972) mixes the particles and intonation in one pot, Kwok (1984) makes clear that it does matter to distinguish between the intonation system and particles - the former is non-segmental and the latter is segmental in nature (p.28). She claims that "the same intonation pattern superimposed on different syntactical structures may carry different meanings" (p.30) and then refers to three examples of intonational modification in the following. She specifies that all the sentences so modified are zero-particled.

(a) A declarative sentence [+ rise] changes a statement to a question (i.e. intonational question). Varying degrees of surprise or even doubt may be indicated by differences in 'key' and by the extent of the rise.

e.g. /nej^{5} wu:j^{5} hɐy^{3}/ (you-will-go.) -> (You will go?)

(b) A declarative sentence [+ rise-fall] indicates disagreement. The intensity of disagreement may be varied according to the extent of the rise-fall, the length of the last syllable.

e.g. /ma:j^{5} ɕi:^{1} mi:n^{4}/ (buy-silk-cotton.) -> (Buy silk cotton but not ordinary cotton.)

/tɐk^{1} ɕɐp^{6} kej^{2} pu:n^{2}/ (only-ten-more-classifier -You have ten odd books.) ->
(You have ten odd books. Implies: You don't have a large number.)

(c) An imperative sentence [+ fall] denotes a lack of civility or mild rudeness to down-right hostility, depending in part on the degree of loudness given to the syllables, especially the last syllable, and the extent of the fall.

e.g. /ha:ŋ^{4} k^{w}ɔ:^{3} lɐj^{4}/ (walk-pass-come - Come over here.)

The (a) idea is an idea from Cheung (1972). The examples in (b) and (c) do not seem well chosen for her observations. Her first sentence in (b) has a tonal combination of low-high-low (i.e., T5-T1-T4) and thus exhibits a contour of rise-fall. Her claim that it has a [+ rise-fall] intonation appears redundant in the light of this. The latter sentence in (b) with the word /tɐk^{1}/ literally means `only ten odd books' and has an implication of `you don't have a large number'. Similarly, the sentence in (c) has a [+ fall] ending corresponding with the tonal

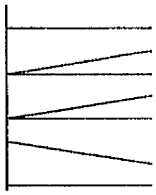
combination of the last two syllables: T3 and T4, making it unnecessary to suppose the intonation feature of [+ fall].

Kwok declared her belief that "certain points made about the functions of English intonation are also valid for Cantonese" (p.28). She quotes from Crystal who mentioned the "predominantly structural stimulated approach which sees intonation playing a primarily grammatical role" (1969: 287). She also adopts Halliday's view that "all contrasts in meaning can be stated either in grammar or in lexis" and if the contrasts made by intonation are not lexical, they are grammatical (Halliday 1967: 10). She believes that "Halliday does not deny that intonation has an attitudinal function, he nevertheless uses the term `grammatical' to include the attitudinal aspects of intonation" (p.29). Therefore, she insists that it is unnecessary to maintain a rigid distinction between what is 'grammatical', and what is 'attitudinal', as 'grammatical' refers to meanings associated with the sentence as a whole.

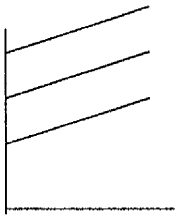
She states that "the tones may be modified to indicate contrasts in meaning which are non-lexical in nature. Such modification, which bring about a change in non-lexical meaning, may be termed intonation" (p.24). She explains that intonation manifests itself in a reduction of exaggeration of the contour of the lexical tones, or even temporarily obliterating their contrasts. She gives the following figures 1.3(a-c) to show how tonal contours are transformed when the intonation feature [+ rise] or [+ fall] is superimposed on them.

She explains that in the case of the [+ rise] modification, "the level and falling tones are made to rise within the syllable" (p.25), while rising tones rise higher and more abruptly. The clipped tones are made much longer. In the case of [+ fall], the fall occurs rather rapidly soon after the beginning of the level, falling and clipped tones; while it occurs after the rise, causing the modified tone to become a rise-fall in the rising tones.

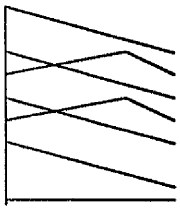
Fig. 1.3. (a) The 6 lexical tones. (b) The three level tonal contours transformed when the intonation feature [+ rise] is superimposed. (c) The 6 tonal contours transformed when the intonation feature [+ fall] is superimposed. (Figures are taken from Kwok 1984: 26).



1.3.a. Lexical tones



1.3.b. [+ rise]



1.3.c. [+ fall]

In her brief account of the field of intonation, Kwok exhibits her understanding of intonation in terms of meaning, forms and function under the influence of modern western linguistic literature, such as Crystal (1969) and Halliday (1967).

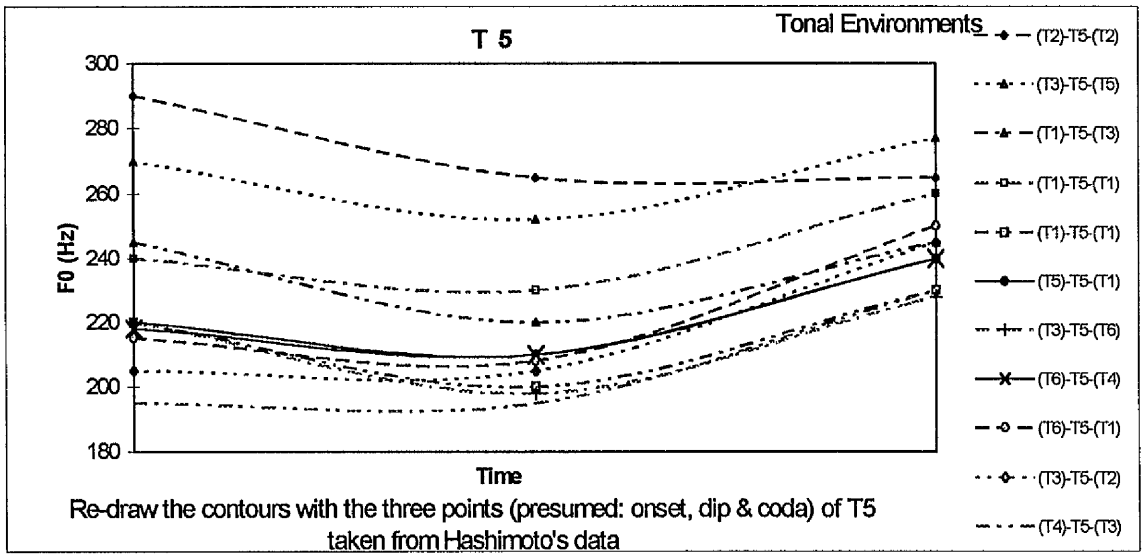
1.3.2. Hashimoto (1972) and Barnett (1949-50)

Hashimoto (1972) has no direct discussion on intonation, but she was the first to make impressionistic observations on tonal behaviour in utterances. She is the first one also to observe that, in normal speech "the onset and coda" of the Cantonese tones "are modified, each according to its immediately preceding and following tones" (p.93). As we can see from her raw data, which are the figures of three points of measurement of the frequency of the tones, no tone displays a level or a straight-line contour, whether the two rising tones (T2 and T5), the level tone (T6) or the falling tone (T4), - all show a fall-rise shape. She considers this modification as the direct variation with the neighbouring tones, "for example, the rising tones in context tend to have a greater fall at the onset if preceded by higher frequency tones and a more prominent rise if followed by the same" (p.93). In other words, the higher the frequency of the preceding tone, the greater the fall is; the higher the frequency of the following tone, the more prominent the rise is. It is not clear whether she meant this is anticipatory modification or carryover modification or both, nor whether the tones preserve their original shape under those modifications of onset and coda. She measures the three points of the frequency of four tones (T4, T6, T5 and T2) with four syllables (/ma:/, /wɐj/, /ma:/ and /hɐw/, respectively) in different tonal environments spoken by a native speaker and gives the frequency of the fall and the rise of the interval but does not mention at what points

and at what times the figures were chosen to measure (p.125). According to the figures, I can only assume those points are taken from the onset, the dip and the offset of the tone with no time specified. I re-draw the contour with these three points linked in single lines in order to give a clearer picture of the fall-rise movement of the tones in normal speech, and then I summarise them by picking out the widest and the narrowest frequency range of each tone and the longest and shortest fall and rise of each tone.

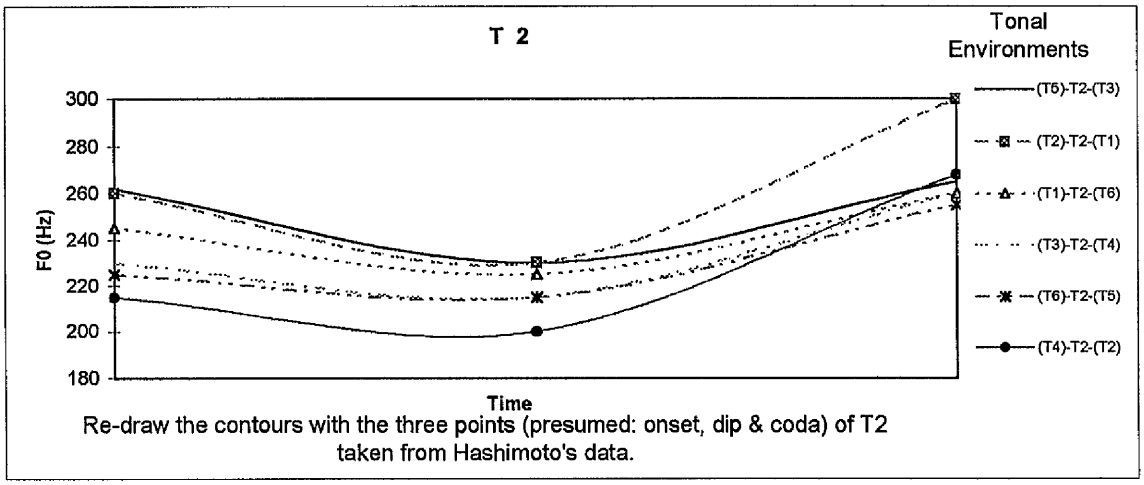
Fig.1.4. (a-d) Re-drawing the contours with the three point measurements (presumed: onset, dip and coda) of the four tones (T5, T2, T4 and T6) in different tonal environments taken from Hashimoto's raw data (1972: 125).

1.4.a.



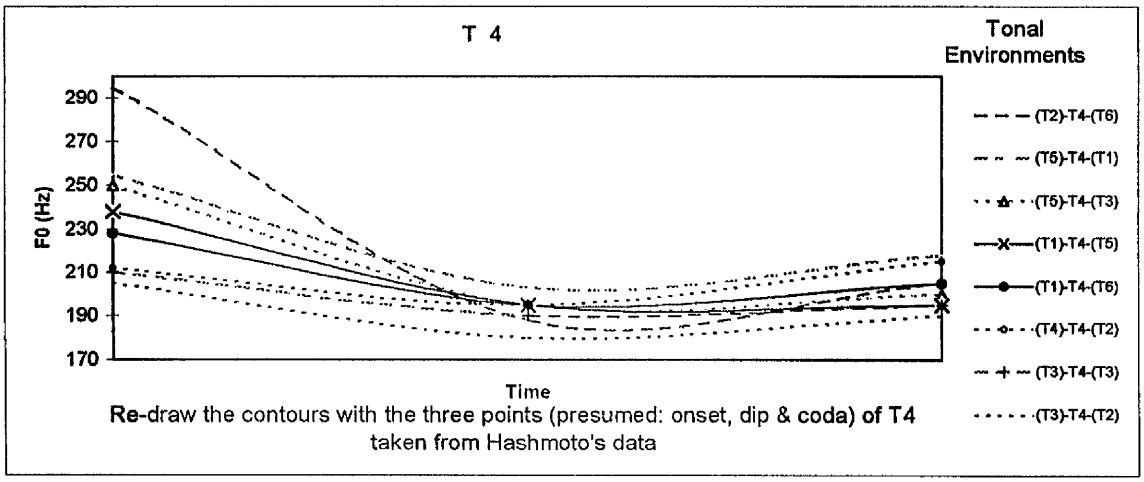
Re-draw the contours with the three points (presumed: onset, dip & coda) of T5 taken from Hashimoto's data

1.4.b.



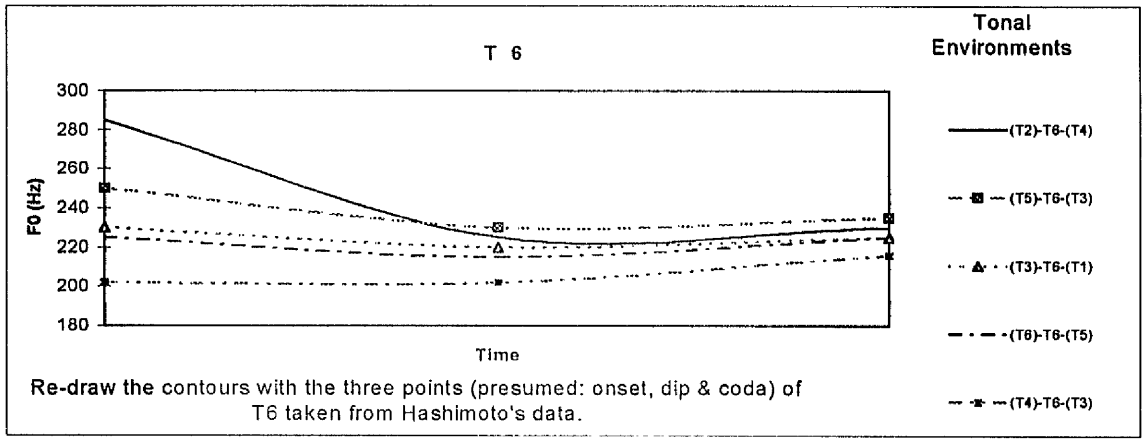
Re-draw the contours with the three points (presumed: onset, dip & coda) of T2 taken from Hashimoto's data.

1.4.c.



Re-draw the contours with the three points (presumed: onset, dip & coda) of T4 taken from Hashmoto's data

1.4.d.



Re-draw the contours with the three points (presumed: onset, dip & coda) of T6 taken from Hashimoto's data.

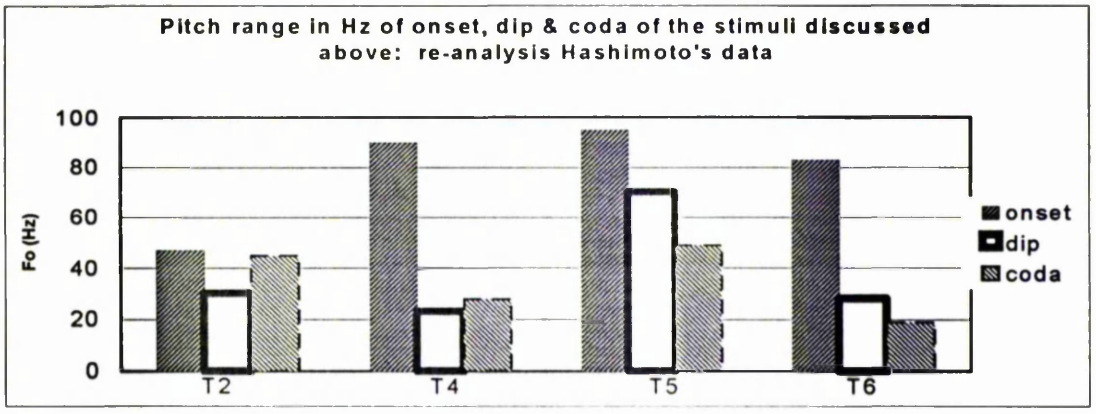
A summary of the figures 1.4(a-d) is given in table 1.8 and in the discussion following. I illustrate the widest and the narrowest frequency range and the longest and shortest fall and rise of each tone (T2, T4, T5 and T6).

Table 1.8. The widest and the narrowest frequency range and the longest and the shortest fall and rise of each of T2, T4, T5 and T6 in different tonal environments. The tone numbers given in the table indicate the preceding and/or following tones while the target tones are measured in Hz. Data are taken from Hashimoto (1972: 125).

Tone	Segments	Pitch Range		Interval			
				Fall		Rise	
		Widest (Hz)	Narrowest (Hz)	Longest	Shortest (Hz)	Longest	Shortest (Hz)
T4	/ma:/	T2-90-T6	T4-3-T2	T2-107	T4-17	20-T3	0-T5
T6	/wɐj/	T2-55-T4	T6-0-T5	T2-60	T4-0	14-T3	5-T1/T4
T5	/ma:/	T3-40-T2	T1-0-T3	T2-25	T4/T3-0	42-T1	0-T2
T2	/hɐw/	T4-53-T2	T5-3-T3	T5-32	T6-10	70-T1	35-T6/T3

The figures 1.4.a-d and the table 1.8 both show that the degree of the falls and rises of the tones complies in principle with what she concludes but reveals a more complicated picture than her discussion suggests: the heights of the falls and the rises of the tones indeed are not necessarily in direct ratio with the frequency of their neighbouring tones. Say, in low rising tone (T5), two of the stimuli are overlapped from the dip to the offset but one is followed by a high tone (T1) and one is followed by a low falling tone (T4) (see the solid lines in figure 1.4a above). One would ask why the one followed by a high tone does not rise more? Another example is the low falling tone (T4) where two stimuli are both preceded by a high tone (T1) but one falls 43Hz and one falls 33Hz down to the dip (see the solid lines in figure 1.4c above): why do they not overlap each other? Moreover, some of the interesting figures slip her attention. Let us look at the table 1.9 below which shows the pitch range in Hz for the three points of the four tones we have discussed above. We find that the range of the pitch occupied by the dips of the stimuli in the low falling tone (T4) is just 23Hz compared with 90Hz for the onset and 28Hz for the offset. Why is the elasticity on the onset three times more than it is on the dip or offset? Is it telling us that for a falling tone, the dip and offset tend to preserve their shape after serious modification of the onset? The low level tone (T6) has a similar phenomenon in the pitch range of the onset, the dip and the offset of the stimuli, which are 83, 28 and 19Hz respectively, showing that the end point is more consistently a level tone. However, for the high rising tone (T2) and the low rising tone (T5), the frequency range of these three points are: 47, 30, 45 and 95, 70, 49Hz respectively, showing that the dip is more stable for the former and the offset for the latter. All these urge us to look at a more complicated picture and to raise questions such as, what else would affect the tonal contour in connected speech apart from the frequency of the neighbouring tones and how do tones preserve their shape in utterances.

Table 1.9. Frequency range of onset, dip and offset of the four tones (T2, T4, T5 and T6) reconstructed from Hashimoto's data in 1972: 125.



Another contribution in Hashimoto (1972) is the discussion of stress. Before her, I can only find Barnett (1949-50) but no one else apart from a brief description in Chao (1947: 36): "stress is not a constituting element of words in Cantonese". By following Chao, Hashimoto (1972) considers that stress "does not play a distinctive role in the phonological system" (p.101), but phonetically there are three perceptible gradations: the neutral stress, the normal stress, and the emphatic stress. She sees that three dimensions are involved in the modification of stress: pitch, duration and loudness. While you give emphatic stress, you are exaggerating the pitch contour and the loudness, but either lengthening or shortening the duration of the syllable (shortening the falling tones) in order to give more prominence; while you are not emphatically stressing the particles, both the length and loudness of such syllables are the weakest or may become atonic and so take the neutral stress. Unfortunately she does not discuss normal stress. She observes that the high level or high falling tone and the short high level tone, the high rising tone and the changed tone, and the low falling tone are often heard as more strongly stressed than others. (cf, Fok's finding that T1 is the loudest and T4 the softest in her perceptual experiments (1974: 34).) Given that she also observes that the preferred stress pattern for nouns is to have the last syllable more heavily stressed, she does not mention what would be the stress pattern if a high level or high falling tone and the short high level tone, the high rising tone and the changed tone, and the low falling tone is in a non-final syllable position of a noun. Then she further confirms this by discussing the transliteration of monosyllabic or disyllabic words in English. She finds a

correlation between the stressed syllable in the English word and the high pitch syllables in Cantonese.

Barnett (1949-1950) has a comparatively detailed description of stress, in which he starts by disagreeing with Chao's claim by comparing the different meaning of /tɕi:^{3}/ when the stress is present and absent: "the most" for the former and "only then" for the latter. His example is given as:

/t^{h}eŋ^{1} jɐt^{6} tɕi:^{3} how^{2} hɐy^{3} mɐn^{6} ha:w^{6} tɕœ:ŋ^{2}/
tomorrow-day-[the most/ only then]-good-go-ask-school-master –
(You'd best ask the headmaster tomorrow.) - When /tɕi:^{3}/ is stressed.
(Don't on any account ask the headmaster before tomorrow.) - When /tɕi:^{3}/ is unstressed.

He points out that Chao also gives this distinction in his index at the end of the "Cantonese Primer" (1947). I cannot ascertain how people distinguished this word in the same context at that time, but I do remember it was quite confusing trying to distinguish the two meanings when I was small. The listener might have to ask for classification and the speaker might have to rephrase the sentence by using a synonym. Nowadays you hardly ever hear the example he quotes above. Using /tɕi:^{3}/ (short for /tɕi:^{3} how^{2}/) for the two meanings has become old fashioned. /tɕɐy^{3} how^{2}/ (most-good) for `the most' and /ɕi:n$^{1}$ how$^{2}$/ (first-good) or /ɕi:n$^{1}$ tɕi:$^{3}$ how$^{2}$/ (first-then-good) for `only then' are common.

He then describes the stress placement in monosyllables, disyllables, trisyllables and tetrasyllables accordingly.

First, he classifies the stress placement in monosyllables into three cases with the presence and absence of stress: (a) has no inherent stress; (b) has an inherent stress; and (c) has an inherent negation of stress (enclitic or proclitic). The overwhelming majority of words come within (a) and can be stressed according to context and emphasis, and words coming within (b) and (c) distinguish themselves in meaning from other words. Examples in (a) are:

/k^{h}ɵy^{5} hɵy^{3} kɐw^{2} loŋ^{4}/ (he-go-nine-dragon – he went to Kowloon) /hɵy^{3}/ is unstressed.
/k^{h}ɵy^{5} kej^{2} ɕi:^{4} hɵy^{3} nɛ:^{1}/ (he-how-time-go-particle -- when did he go?) /hɵy^{3}/ is stressed.

In examples (b) and (c) he refers to his previous page. I am not sure which examples he refers to. It seems the word /pɐt^{1}/ in /k^{h}ɐw^{4} tɕi:^{1} pɐt^{1} tɐk^{1}/ (seek-it-not-obtain - what a stroke of luck) is the one for (c) as it is the only example of negation. Hashimoto comments that these three types of cases can either be explained by syntactic or by stylistic factors (1972: 128).

Then he goes on to discuss the stress placement in polysyllables. He warns that "while there are many words whose stress seems to be dictated purely by euphony or the requirements of emphasis, some generalisation can safely be made" (p.742). The stress placement in disyllables is classified into four cases (s for stressed and u for unstressed):

Examples

(d) trochee (i.e., s-u) /mɐt^{6} t^{h}ɔ:ŋ^{4}/ (honey-sugar - honey)
(e) iambus (i.e., u-s) /mɐt^{6} foŋ^{1}/ (honey-bee - bee)
(f) spondee (i.e., s-s) /ma:^{5} mej^{5}/ (horse-tail – horse tail)
(g) pyrrhic (i.e., u-u) /kɛ:^{3} lɔ:^{3}/ (particles: final affirmative)

The most numerous is (d) and the least numerous (g). A number of words in (e) may also belong to (f). Regarding trisyllables, he classifies these as:

(h) amphimacer (i.e., s-u-s) /mɐt^{6} foŋ^{1} tɐw^{3}/ (honey-bee-nest - beehive)

(i) the negative /m^{4}/ divides into
unstressed (u-s-u): /m^{4} ɕy:^{1} fok^{6}/ (not-comfortable)
(u-s-s): /m^{4} how^{2} ɕek^{6}/ (not-good-eat - not delicious)
and secondary stress (s-u-s): /m^{4} tɕi:^{6} ji:n^{4}/ (not-natural)

(j) marked - anapaest (i.e., u-u-s) /ti:^{1} ti:^{1} ta:j$^{6/1}$ / (little-little-big - tiny)
- dactyl (i.e., s-u-u) /ka:w^{2} m^{4} ti:m^{6}/ (work-not-succeed - not successful)

The case (h) refers to the predominant tendency of stressing on the first and third syllables. The case (i) refers to the disyllabic adjective-group preceded by the negative /m^{4}/: where

/m^{4}/ is unstressed when it is followed by a trochee or a spondee and takes a secondary stress when it is followed by an iambus. The case (j) refers to other stresses which are rare. Finally, he classifies tetrasyllables into two:

(k) choriamb (i.e., s-u-u-s): /ɕej^{3} ɕy:^{1} ŋ^{5} keŋ^{1}/ (four-book-five-classics - The Classics)

(l) double trochee (i.e., s-u-s-u) /p^{h}ow^{3} t^{h}ɐw^{2} ta:m^{1} pow^{2}/
(shop-head-bear-guarantee - shop guarantee)

(k) is commoner. It involves stress on the first and the fourth syllables. It changes the spondees /ɕej^{3} ɕy:^{1}/ and /ŋ^{5} keŋ^{1}/ to choriambs where the tetrasyllables are found. The tendency of latter two syllables to become an iambus is defeated by an original trochee provided they are also the principal part of the combination as in (l). However, Hashimoto thinks "the stress patterns he assigned to polysyllables do not seem to have general validity"; for example, it is difficult to maintain the stress pattern of a phrase with choriamb and not double trochee (Hashimoto 1972: 128).

I think it is difficult to decide what is trochee and what is iambus or spondee in Cantonese in the first place, not to mention going further to decide what is amphimacer and choriamb or double trochee. I do not even see any definition of stress in Barnett but I take it that it has to be inferred from his references to the prominence revealed by impressionistic observation of the effect of stress on loudness, or intensity, because he says he does not propose to say anything here "about the effect of stress on duration and tonality" (p.745). If this assumption is right, we can easily turn to another picture by looking at the intensity of the examples cited by him. In the following I give the results of one of my acoustic investigations taking his words as the material. Here I use "h" for higher intensity and "l" for lower intensity. My results contrast with his classification and the difference is revealed.

Examples are taken from Barnett		Syllables	Barnett's Classification	Intensity (my findings)
/ɕej^{3} ɕy:^{1} ŋ^{5} keŋ^{1}/	(four-book-five-classic - The Classics)	4	choriamb	l-h-h-l
/mɐt^{6} foŋ^{1} tɐw^{3}/	(honey-bee-nest - beehive)	3	amphimacer	l-h-l
/p^{h}ow^{3} t^{h}ɐw^{2}/	(shop-head - a shop)	2	iambus	h-l

1.3.3. Fok (1974) and Vance (1976)

Fok (1974) contributes much work on tones related to intonation based on experiments. She was the first to treat the prosodic elements of tone in different expressions scientifically. She tests the nine tones in isolation as well as in frame sentences with four versions: neutral expression, angry expression, emphatic expression and interrogative expression. She finds that with tones in isolation:

(a) Intensity does not offer a reliable clue to the expression, except that the intensity is always lowest in neutral expressions;

(b) Duration of tone is longest in interrogative expressions, decreasing for emphatic, neutral and angry expressions in that descending order;

(c) The F_0 range expands from neutral expressions to angry expressions, almost doubles in emphatic expressions, with interrogative expressions having the largest. Angry expressions raise the range in general;

(d) The same pattern of relationship among the tones is maintained, but once the structure is disturbed as in the tones said with an interrogative expression - all tones end with a rising tail and end at about the same frequency region - they become less marked. However, they still maintain their initial frequency distinction (p.29).

However, she does not discuss the cause of the differences of the absolute pitch values of the sequence of tones - say, the F_0 of the test word in T5 is higher than in T2 (p.31) showing that she takes less interest in tones within a sentence frame. Her findings here are that the F_0 range covered by the tones said in isolation is wider than that of the ones extracted from a carrier sentence and the rising tail in interrogative expressions is shifted to the last word of the sentence. She notices that the duration of the two sets of tones (i.e., produced in isolation and in the carrier sentence) are about the same but she does not think this finding is valid as she feels that her informants pay too much attention to the test words.

In another of her experiments dealing with perception with isolated tones, she finds that the identification of tones is as accurate with angry expressions as with neutral expressions, (about eighty percent). Less accurate are emphatic expressions and the worst is with interrogative expressions due to the rising tail, about fifty-fifty.

Vance (1976) is a similar version of Vance and Walker (1976). I will only discuss the former.

Vance (1976) is one of the earlier experimental investigations on tone and intonation in Cantonese, which consists of two experiments: one is to test possible intonational effects of putting test words in medial and final positions and in contrastive contexts in three declarative sentences, respectively; one is to test how much confusion there would be if relative F_0 is the sole acoustic cue used to differentiate the tones, by splicing out and interchanging the test words in each pair which are considered the most confusing, e.g. T3 and T5 or T3 and T6 from the data recorded in the previous experiment; and he uses the same subjects to make the identification. His conclusions may be summarised thus:

(a) Tonal distinction does exist in Cantonese - relative pitch relationships among the tones are maintained in all three contexts (p. 368 and 382);

(b) No tone sandhi exists (i.e., 55 and 53 do not contrast with each other.) (p.374 and 382);

(c) A gradual fall in F_0 occurs almost uniformly on level tones (p.382);

(d) Vowel length is not involved in differentiating the tones (p.383);

(e) Contrastive stress does not significantly affect F_0 (p. 368 and 380), but according to impressionistic judgement, it does affect the length of both consonants and vowels and intensity (p.382); the increase of F_0 on the test words, if any, in the contrastive context may simply reflect an increase in overall pitch level (p. 381);

(f) Sentence-final tone lowering does occur in Cantonese (p.376).

He also compares the same words appearing at different positions in two different sentences and concludes that there is a sentence-final tone lowering. His approach is questionable: on the one hand he makes comparisons by using two absolute F_0 values from two different sentences, completely ignoring the fact that a speaker may start with a different register in his voice range on different occasions - for example, the comparison of the sentence-initial word /k^hɵy^5/ (he) in the medial and final contexts. On the other hand, the measurements of tonal contour are not made consistently - some are at the beginning; some are at the end of the vocalic portion; some are both, and some are at the peak of the slope of the contour (p.376-380), which makes it difficult to believe his calculation is right. For example, /how^2/ (very) and /ɕɛ:^{2}/ (write) are both T2 but the former is measured at the peak and the latter at the onset; the test word /ɕi:^{2}/ (history) is measured at the beginning and the end. The fact that the peak of /how^2/ is higher than the figures of its preceding word /ɕi:^{2}/ has escaped his attention. Cheung (1986: 253) points out something that Vance ignores: "...the sentence-initial word /k^hɵy^5/ (he) is constantly of higher pitch than the sentence-medial word /ɕi:^{5}/ (city) (which is one of the six test words used in the experiment). But it is only natural that he was not aware of this interesting fact, for he actually thought that /k^hɵy^5/ was in T2." Cheung is right in this last point, but overlooks Vance's data of /k^hɵy^5/ and /ɕi:^{5}/. On the other hand, if Cheung is correct in the first point, Vance's conclusion about the final lowering is convincing. The fact is that out of five informants, three produce a lower peak in /k^hɵy^5/ than at the beginning and end points in /ɕi:^{5}/. Cheung also wrongly accuses Vance of ignoring his own data which show that purely even tones, i.e. T3 and T6, have a strong tendency to be realised as a fall, not only in a final position but also in a medial position in the sentence. As we can see from above, Vance understands this fall as a 'uniform'. However, as Cheung points out, Vance has mistaken the citation tone of some individual test words which really weakens his claim. He is aware that his informants complained that /ɕi:^{3}/ (try) and /ɕi:^{5}/ (city) were the most difficult of all the test words to tell apart. Further, he mistook a T3 for a

T6 in two words: /k^{w}ɔ:^{3}/ (pass) and /kɔ:^{3}/ (measure word). This leads him to conclude that it is almost certain that T5 was identified as T3, just as T6 was identified as T3 in his analysis.

1.3.4. Kwok and Luke (1986) and Johnson (1987)

Kwok and Luke (1986) give the first general description of Cantonese intonation and describe three forms of intonation based on their observations in experiments with pitch contour. They use the example /k^{h}ɐy^{5} m^{4} hɐy^{3}/ (he-not-go - he will not go) to demonstrate the three meanings expressed by the three forms of intonation which thereby serve as three grammatical functions:

(a) Pitch contour falls slowly - so as to be an answer to "will he go?" - a statement;

(b) Pitch contour rises slowly - equal to "why won't he go?" or "won't he go?" - a question;

(c) Pitch contour rises first then falls - equal to "of course he won't go." or " who says he will go" - to emphasise or to contrast, for arguing.

They see the function of intonation in Cantonese as to distinguish grammatical structure and to express the speaker's attitude, emotion and social background. The forms of intonation are usually demonstrated at the last word of the utterance where the value of the canonical form of the lexical tone is changed. They do not explain how this finding can be reconciled with (a), (b) and (c) above. They declare that each intonation has its own form and meaning, each form of intonation expresses a different meaning under different grammatical structure. This is called the content of intonation (p.34). This is probably the first work to make such a clear declaration of the relationships between the form and content of intonation and the grammatical structure of Cantonese. They further give examples of six forms of intonation, three of which are repeated in Kwok (1984) and were discussed in (1.3.1) above. For convenience, I repeat them here and list them before the three new forms:

(d) The declarative sentence [+ rise] changes a statement to a question:
e.g. /nej^{5} wu:j^{5} hɐy^{3}/ (you-will-go.) -> (You will go?)

(e) The declarative sentence [+ rise-fall] indicates disagreement:

e.g. /ma:j^{5} ɕi:^{1} mi:n^{4}/ (buy-silk-cotton.) -> (Buy silk cotton but not ordinary cotton.)

/tɐk^{1} ɕɐp^{6} kej^{2} pu:n^{2}/ (only-ten-more-measure - You have ten odd books) ->
(You have ten odd books. Implies: You don't have a large number.)

(f) The imperative sentence [+ fall] is to denote a lack of civility:

e.g. /ha:ŋ^{4} k^{w}ɔ:^{3} lɐj^{4}/ (walk-pass-come - Come over here.)

(g) A question in a rising form becomes an echo question:

e.g. /k^{h}ɵy^{5} tej^{6} kej^{2} ɕi:^{4} fa:n^{1}/ (he-plural-what-time-return - When will they be back?) - with a rising form the meaning is changed to: "Did you ask me ...?"

(h) An imperative in a rising form means a change to a softened voice, to consult or to suggest:

e.g. /ɕi:^{3} ha:^{5} joŋ^{6} la:m^{4} ɕek^{1}/ (try-a bit-use-blue-colour - Try blue.) - a rising form and a low pitch at both the beginning and the end of the utterance, changes the meaning to: "Let me try ...."

(i) A statement in a gradual (as opposed to sharp) rise-fall form with the last syllable prolonged is ironic:

e.g. /ŋɔ:^{5} m^{4} fa:n^{4}/ (I-not-bore - I am not bored.) - with a gradual rise-fall form the meaning is: "You will be surprised if I am not bored."

Six informants served in the experiment. Only one figure of the pitch contour of each example sentence together with the figure of their counterpart original sentences (except (h) which does not have a figure for the original sentence) are shown. No segment or duration in any figures are indicated, nor are measurements taken. Under each figure of the original sentences, the same Chinese scripts as in their counterpart sentences (i.e., the examples above) are given to indicate the context of the sentence, but no information about the intonation forms (e.g., [+rise], [+fall]) is given. I assume that the original counterpart sentences have a neutral content. They do not specify whether the eleven figures of pitch contour are all taken from only one informant or from a number of persons or from all six of them. It is obvious that all their conclusions are based on casual observation without any measurements. Looking at the figures, there are a couple of things that seem to have escaped the authors' attention:

First, the durations of the complete sentences (e) and (f) are twice as long as their original counterpart sentences. Sentence (i) is three times as long as its original sentence. This may

be due to the emphasis of the new meanings assigned when compared with their original sentences which have a neutral content. Note that the lengthening takes place not only on the last syllable but on the whole sentence. On the other hand, sentence (g) is slightly shorter than its original sentence. This may be due to the effect of an echo question which repeats the old information: "Did you ask me...?" The same explanation can be given to sentence (d) which has the same length as its original sentence. It is not a genuine question but asks for confirmation, and therefore repeating the old information is necessary.

Second, the F_0 range in the examples above is wider and higher than their counterparts. This seems to be a characteristic of all those intonation forms.

Cheung (1986) criticizes the examples (f) and (h), saying they are "badly chosen" if the respective intonation contrasts are valid (p.255). He refers to the examples and discussion in Kwok and Luke (1983) and (1985) which I have no access to and which, in any event, duplicate those in Kwok and Luke (1986) quoted above. He raises the question - in the absence of other examples, how do we know that the [+rise] in (h) is not due to the final T1, the highest tone and that [+fall] for (f) is not due to the final T4, the lowest tone? He considers that the reason given for the lack of control sentences in those examples is that the wording restricts the connotational possibilities. Thus he casts doubt on Kwok and Luke's claim about the correspondence extracted between form and content. He further criticizes the authors' use of intonational contrasts, normally only used for intonation languages (p.254). For Cantonese, he points out that unlike in English the realisations of such intonational contrasts are all utterance-final phenomena. The last point he makes is unfair to the authors who are not in conflict with him - we have quoted above that the form of intonation is usually demonstrated at the last word of the utterance. Otherwise his comments are fair and accurate. It is true that looseness of specification of the phonetic, semantic and syntactic environment in the examples chosen can lead the experimenter to draw an untenable conclusion. Moreover, he criticizes the authors for not using sentence(-final) particles in all examples given when they discuss Cantonese intonation. He suggests that the connotations can easily override the denotation of the two sentences under discussion

(i.e. (f) and (h)) simply by adding sentence(-final) particles to them. The result of this addition is: the same sentence can have both consultation/polite request and rude instruction/command connotations. The former is realised as a high-rise /˝ə/ and the latter a high-fall /a:/ or a mid-fall /la:/, which suggests the primacy of sentence(-final) particle over final-tone modification. I take it, following this, that a final rise intonation seems to imply the consultation connotation and a final fall intonation the rude command connotation in the circumstances of both the presence and absence of the sentence(-final) particles, in the discussion from the two works. However, I consider that the words /ɕi:^{3} ha:^{5}/ (to try) can be either a polite invitation or a threat (i.e. `dare you!') semantically. Given this, the imperative sentence of /ɕi:$^{3}$ ha:$^{5}$ ha:ŋ$^{4}$ k$^{w}$ɔ:$^{3}$ lɐj$^{4}$/ can mean `try to come over here' or `how dare you come over here'. It also can sound a threat to me if we apply a [+rise] intonation form defined by Kwok and Luke (1986) or a /˝ə/ sentence(-final) particle defined by Cheung (1986). It totally contrasts with their descriptions of an imperative with a polite invitation or a rude command connotation. If this is accepted, the analysis of the correlations between forms and the content of intonation in the above discussion is really brought into question.

Johnson's (1987) main contribution is to define that a declination component is involved in the production of intonation in Cantonese. He uses six paragraphs each of which consists of three sentences. Each sentence has two test words which are divided by long and short intervals. The result of his experiment shows that under all circumstances the F_0 of the first test word was on average higher than the F_0 of the second test word in a sentence, and the short interval condition has a smaller amount of decline than does the long interval condition. Greater change in the length of the interval corresponds to greater change in the ratio of the pitch of the test words. He also indicates that the numbers of sentence initial syllables which precede the first test word have an effect on the mean F_0 ratio of the test words. The correlation between overall declination and the position of the first test word close to the beginning of the sentence is evidence for initial raising which is one type of boundary effect. However, he dismisses some of the other contributing factors to the downtrend, like final lowering and tonal interaction (p.11). He considers these as sources of F_0 patterns to exhibit

downtrend (p.2). He distinguishes the three terms: (a) downtrend, (b) declination and (c) final lowering. Downtrend refers to "a tendency for F_0 to decline over the course of an utterance" (p.1). Declination is defined as "a gradually changing backdrop to local F_0 events" (he quotes from Ladd 1983: 54). "Unlike boundary effects and tonal interaction, declination is an effect which operates over the phonological unit" (p.2). Final lowering refers to "the tendency for F_0 to fall at the ends of prosodic units" (p.1). Note that in his reading corpus the tone of the test words is the same but their neighbouring environment is different. Following this, I am surprised that he ignores the fact that the F_0 difference of the test words could be induced by their neighbouring F_0. [Hashimoto (1972) has a comprehensive description of this effect as we discussed above.] This leads him to draw a conclusion that there is "no evidence in favour of positing any tonal interaction effect" (p.11) even though his measurement is only at the end point of the syllables.

In his pilot experiment he finds that "a method of accounting for tone preservation which included a declining tone space was better suited to the task than one which assumed a level tone space"(p.1). The difference of the two models which he uses to test with seven single sentences is the F_0 envelopes, i.e., he names them `tone space'. The top F_0 envelope is defined by an early high tone in the sentence and one at the end of the sentence, and the bottom F_0 envelope is defined by a low tone at the beginning of the sentence and one late in the sentence in the declining tone space model; while in the level tone space model the two assumed level F_0 envelopes are defined by an early high tone and an early low tone. The result of these two models allowed him to conclude that both the models succeed in keeping the lexical tones separate from each other when averages are considered. However, he is not happy with the greater variation in the position of a tone within the tone space in the level tone space. In other words, the level tone space model suffered from the limitation that the value of a tone within the tone space is dependent upon its location in the sentence and "this runs counter to the native speaker's intuition that tones are preserved in sentence contexts" (p.6). He gives no discussion about how the native speaker's intuition preserves the tones, so we are not clear what he means by "runs counter to". He gives further evidence by computing an r value for a correlation between "position in a sentence" (as indicated by

counting words from the beginning) and "position within the tone space" in both models. The results of the r values show that in all cases the correlation is lower for the declining tone space than for the level tone space. Only "the rising tones are the exceptions" and "it is not clear why". It is very clear to a native speaker that the explanation for the "exceptions" is his mistaking of the high rising T2 for the low rising T5 in his example of /hɐj/ (at) shown in figure 3 and 4.

1.3.5. Cheung (1986) and Bauer and Benedict (1997)

Having accepted the four points of view of intonation in a tone language proposed by Cruttenden, Cheung (1986) casts a fresh light on the view that the "final tone of the utterance may be modified in various ways" (Cruttenden 1986: 10-11). He considers that "the most salient feature of Cantonese intonation is ironically that the language has little intonation" (p.250). The reason given is that the rich tonal system and the rich system of sentence(-final) particles together leave very little room for, respectively, the intonational manipulation of pitch and the sentential connotation expression in intonation. He highlights the importance of the utterance-final contrasts which are considered as the whole realisation of the Cantonese intonation. Therefore, the sentence(-final) particle inevitably becomes the centre of his discussion. He recognises that the sentence(-final) particle shares the content (i.e., sentential connotations) of intonation while the tone shares the form (i.e., the pitch shape) (p.255). By tackling the relationship between tone, intonation and sentence(-final) particle, he highlights the importance of the role of the latter which he describes as fulfilling more or less the same function as intonation. Here is the summary of his treatment of the sentence(-final) particle:

(a) The form of sentence(-final) particles is: segmental or segmentless;

(b) The intrinsic pitch shape of sentence(-final) particles is: intonational, not tone;

  (i) The tone identified for a sentence(-final) particle is the result of coercion; ('coercion' is defined as "a process whereby uncanonical phonetic forms, which arise as the output of casual speech processes, are replaced by canonical forms." (p.2)
  (ii) The sentence(-final) particle and the sentence-final syllable together are subject to contraction;

(c) The function of the sentence(-final) particle is: to convey the sentential connotation together with final tone modification.

As such in (a), (b.ii) and (c), he proposes two segmentless sentence(-final) particles and uses the idea of autosegmental morae in representing them (p. 257 and 259):

(i) Echo-question sentence(-final) particle: H / M (by attaching a high-pitched tail to any utterance.)

(ii) Irony sentence(-final) particle: RF / M (by attaching a rise-fall tail to an utterance.)

(H: high M: mora RF: rise-fall)

The example (g) in Kwok and Luke (1986) can serve as the echo-question sentence(-final) particle and the example (i) can serve as the irony sentence(-final) particle.

Considering (b), he gives the following examples to identify the two sentence(-final) particles and to demonstrate that they cannot be lumped together and assigned T3 (p.256):

"I am back." /ŋɔ:^{5} fa:n^{1} lɐj^{4} la:^{3} (information furnishing)

/ŋɔ:^{5} fa:n^{1} lɐj^{4} la: ↘ / (Look!)

The second sentence(-final) particle serves as a good example of (b) and the first of (b.i) (i.e., the sentence(-final) particles do not have intrinsic pitch shape but the pitch is the result of coercion.). Apart from treating intonation forms as segmentless sentence(-final) particles, he takes the view that intonation is in the form of final tone modification.

This contribution adds little beyond Cheung (1972) to the exploration of Cantonese intonation. There is an inevitable limitation as he discusses intonation without referring to any of the acoustical studies. For example, he claims that downdrift does not apply when monosyllables are cited in isolation, and only considers that the realisation of Cantonese intonational contrasts are all localised - are all utterance-final phenomena which is the foundation of his treatment of intonation. The former conflicts with the experimental result from Fok (1974) and the latter is open to question as we discussed above.

Bauer and Benedict (1997) record eight declarative sentences with neutral expression and then display their F_0 contours. There are about four to six syllables in each sentence where the lexical tones are in random order. They describe in detail the modification of the tone shape and conclude that "Cantonese declarative utterances exhibit the phenomenon of downdrift in which tone contours tend to become progressively lower as the speaker arrives at the end of the utterance" (p.154). In their description of the modification of tone shape they repeatedly assert that the modified tone shapes are due to the F_0 value of the following lexical tone.

Unfortunately they neglect many interesting facts displayed in the F_0 contours. For example, there are four pairs of T3-T2 combinations which have different contour shape: one T3 displays a level shape (p.259), one a fall (p.263), one a fall-rise (p.483) and one a rise-fall (p.485); the peak of one T2 is the same (p.264) as its preceding T3, one is much higher (p.483) and one is much lower (p.266) than their preceding T3. This phenomenon can be found in other pairs, for example, the T5-T5 combinations (P.264, 266 and 343) and the T3-T3 combinations (p.265 and 266). I am afraid that their claims of anticipation modification and downdrift cannot satisfactorily explain this simple phenomenon.

In their discussion of syllable-timed rhythm in Cantonese, they suggest that a few syllables may receive relatively lighter stress yet still generally retain their tones (p.317). They include the four words: (a) /m^4/ (no, not); (b) /a:^{3}/ (a prefix); (c) /kɛ:^{3}/ (a suffix); (c) /kɔ:^{3}/ (a measure). They note that the last three words all have T3 which lies in the middle of the speaker's pitch range and may sound like a neutral tone that is neither high nor low (p.317). This does not explain anything to me as there are many T3 words which also lie in the middle of the speaker's pitch range, but do not sound like a neutral tone nor receive a lighter stress as is the case with these three particular words? Secondly, is there any correlation between the "lighter stress" and the "neutral tone"? Thirdly, what is a neutral tone by definition? (Recall that Barnett has a comprehensive discussion of the negative /m^4/ which never takes

a primary stress. Yip (1992: 28) also points out that the prefix /a:^{3}/ maps into the iambic template in disyllables.)

[Note:

(1) Cheung (1972) and Kwok and Luke (1986) are all written in Chinese. Quotations are in English, my own translation. Summaries are also by me.

(2) Different annotations are employed in different works. They are converted into the system I set out in section 1.2.]

CHAPTER TWO

TONAL COARTICULATION

2.1. TONAL BEHAVIOUR IN TWO-TONE SEQUENCES (EXPERIMENT I)

2.1.1. Introduction

2.1.1.1. Intonation and Tone

Intonation is speech melody whose primary acoustic correlate is fundamental frequency. Tone is known to make use of pitch to convey lexical meaning while intonation is known to make use of pitch to convey linguistic and paralinguistic meaning at utterance level. Different syntactical meanings in utterances may require different types of pitch pattern to realise them, and different lexical meanings require different tones to realise them. When a language like Cantonese uses pitch for tone and intonation in utterances simultaneously, the resulting pitch pattern is necessarily very complicated. When we look at a pitch contour extracted from an utterance, we wonder how it comes about. We ask why two identical syllable tones display different lengths or different heights or shapes in the utterance pitch contour. Or, how the utterance contour would be if one of the syllable tones were omitted, etc. If we fail to answer these basic questions regarding tonal behaviour in intonation, we will be tangled in a puzzle. There is no way to open up the veil of intonation in a tone language if the different behaviours of tones are not discerned. Chao (1968: 39) uses small ripples riding on large waves as a metaphor for comparing tone and sentence intonation. Here my first concern – looking at the small ripples within the large waves - is with the phenomenon of tonal coarticulation: do tones interact with their neighbours in connected speech? And if so, how?

2.1.1.2. Defining Tonal Coarticulation

Coarticulation is defined by Ladefoged as the overlapping of adjacent articulations involved in utterances (1982: 82). In speech production, the vocal organs never move in separate

steps. When a string of sounds is pronounced, the vocal tract is continuously in motion rather than exhibiting a sequence of steady states. The articulation is continually changing and sounds inevitably influence each other. As a result, sounds assimilate to each other. However, the concept of coarticulation is generally restricted to segments. For instance, Keating writes (1990: 452): "coarticulation refers to articulatory overlap between neighbouring segments, which results in segments generally appearing assimilated to their contexts." It is easier to grasp the process of coarticulation between segments as we understand that it takes time for one articulator, for example, the tip of the tongue, to move from one position to another. There is never a precise locus in gesture when producing more than one sound. Sometimes we can even feel the presence of such a movement physically. As regards pitch, we understand that the varying of the pitch of a sound depends on the number of vibrations of the vocal folds in a particular time, but we cannot feel the number of vibrations themselves. However, there is no reason to believe that the phenomenon of coarticulation happens only to segments and not to tones. Tone needs segments to carry it. During the coarticulation of segments, gestures may overlap in time (Browman and Goldstein 1986: 219). Tone, as a suprasegmental phenomenon associated with segments, may also participate in overlapping gestures. Moreover, the vocal folds may also need time to adjust their movement from producing one segment or one tone to another. These adjustments may only take some milliseconds or a few pulses of vibration to accomplish, but this is the crucial thing we need to look at. We presume that there must be a transition time for the vocal folds to adjust their movement from one tone to another. Such a simultaneous or overlapping process is called tonal coarticulation. As a result of tonal coarticulation, the phonetic value of tones is assumed to be varied. One objective of the present study is to examine the phonetic variation of tones when undergoing tonal coarticulation.

2.1.1.3. Directions of Tonal Coarticulation

Crystal (1985: 52) elaborates two types of segmental coarticulation. One is anticipatory coarticulation, which happens when "an articulator not involved in a particular sound begins to move in the direction of articulation needed for a later sound in the utterance". The other is the coarticulation "when a sound retains a characteristic deriving from an earlier

articulation". (We call the latter carryover coarticulation, or carryover assimilation.) It is reported that while English has a tendency to show larger anticipatory effects, French and Italian show larger carryover effects (Ladefoged 1982: 53). In Cantonese, Rao *et al* (1982: 294) reports that segmental assimilation effects are in both directions: anticipatory assimilation as in /ka:p^3 ŋa:ŋ^2/ > /ka:m^3 ŋa:ŋ^2/ (to force) and carryover assimilation as in /kɐm^1 jɐt^6/ > /kɐm^1 mɐt^6/ (today). He further points out that in the case of the elision of the verb suffix /tɕɔ:^{2}/ (past tense suffix), the verb takes over the tone from the suffix even though its segments are elided, as in /ɕek^6 tɕɔ:^{2} fa:n^6 mej^6/ > /ɕek^2 fa:n^6 mej^6/ (eat-ed-rice-not yet - have you eaten yet?). Recent acoustic research on tonal coarticulation confirms the existence of bi-directional assimilation and its effect on both tone contour and tone height. For Vietnamese, Han and Kim (1974) report that a carryover effect is greater than an anticipatory effect. Experiments on Thai by Gandour *et al* (1992 and 1994) and Potisuk *et al* (1997) show that different tones react differently to different tonal contexts in two-tone or three-tone sequences. Interaction between tones is restricted to adjacent tones. An assimilatory effect is greater than a dissimilatory effect, and a carryover effect is greater than an anticipatory effect. The extent of the carryover effect can last for 60% or 70% of the duration of the affected syllable from the onset. On Mandarin, Shen (1990) and Xu (1994) both claim a significant effect on height but disagree regarding the effect on contour. An anticipatory effect and a carryover effect are found to be equally common by Shen while Xu finds that the target tone is influenced more by the preceding tone than by the following tone. Shen also reports that the entire syllable from onset to offset is affected in pitch as a result of assimilation. I am aware of no acoustic study on Cantonese tonal coarticulation. Hashimoto (1972) is probably the only one to observe that, in normal speech the onset and coda of the Cantonese tones "are modified, each according to its immediately preceding and following tones... for example, the rising tones in context tend to have a greater fall at the onset if preceded by higher frequency tones and a more prominent rise if followed by the same" (p.93). This indicates that anticipatory and carryover effects are both present and contours are affected positively in Cantonese. The present study attempts to provide a systematic exploration of the phonetic behaviour of individual tones in sequences.

2.1.2. Method

2.1.2.1. Material

After considering various factors which affect F_0 and examining several alternatives, it was decided to use real words rather than nonsense words, and words voiced throughout rather than words with voiceless onset or coda, for the investigation of tonal coarticulation in Cantonese. There are thirty-six combinations, each with two sets of words, one with /ji:/ as the first syllable and one with /ji:/ as the last. All possible combinations of disyllabic words from the six tones which consist of /ji:/ in two-tone sequences are used: for example, /**ji:**^{1} ji:^{1}/, /**ji:**^{1} jy:n^2/, /**ji:**^{1} wɐj^3/, /**ji:**^{1} li:w^4/, /**ji:**^{1} leŋ^5/, /**ji:**^{1} mi:n^6/ and /ɐw^1 **ji:**^{1}/, /jɐw^2 **ji:**^{1}/, /a:^{3} **ji:**^{1}/, /meŋ^4 **ji:**^{1}/, /jy:^{5} **ji:**^{1}/, /ji:^{6} **ji:**^{1}/, etc, (the target syllable is in bold. For the word list, see appendix I.) The target syllable /ji:/ is one of the few words which can carry all the six tones. It also has an advantage in that there is a similarity between the consonant and the vowel, as the approximant /j/ is known as a nonsyllabic version of the high vowel /i:/; and is described as usually consisting of a rapid glide from a high vowel to the following vowel when it occurs at the beginning of a syllable (Ladefoged 1975: 209). It is also known that the average F_0 of vowels shows a systematic correlation with vowel height. The higher the vowel, the higher the F_0 (Lehiste and Peterson 1961: 419, and Ohala 1978: 29). The difference in F_0 between high and low vowels may be as much as 25 Hz. The choice of /j/ and /i:/ sequences is to minimise any effect on F_0 due to vowel height.

/ji:/ is placed to follow or precede a syllable beginning and ending with sonorants in order to see what will happen in the course of voicing through the two syllables. Shen (1990) and Gandour *et al* (1992) both use voiceless obstruents preceding vowels. In that case, it is easy to work out the segmentation. However, the burst of the obstruents after VOT would affect the beginning of the pitch of the following vowels. This effect is known to last for 100ms after vowel onset (Hombert 1980: 81). This F_0 perturbation gives investigators a problem when it comes to deciding where to measure the onset of the tone. Shen decided that the "tonal

onset was defined after voiceless stops as the first F_0 value" and "after voiced stops as the first F_0 value after the initial abrupt F_0 increase" (1990: 283), while Gandour *et al* decided that "[t]o minimise the effect of consonantal perturbations at the beginning of the F_0 contours, changes in F_0 that were 10Hz or greater across a 10ms time frame were excluded" (1992: 114). In the former case, the perturbation after the burst at the beginning of the F_0 contour may raise the value of the onset of the tone. Shen measures the onsets, the turning points and the offsets of tones, and almost all tonal contours turn out to have a V shape. Gandour *et al* do not take the real onset of the tone but start some way in from the beginning, and the gap between the two vowels in the F_0 contour seems to block the transition between the two tones. Han and Kim (1974: 224) use some voiced initials but decide not to use their F_0 values as the tonal onset because they believe that they do not contribute to the shaping of the tone contour. Xu (1994: 2242) decided to measure from the start of the vocalic segments only, regardless of the manners of the initial consonants. What occurs at the transition of the two tones appears to me to be the crucial point of the coarticulation of the two tones. If we want to see the actual transition of the two tones through the F_0 contour, we have to let the vocal folds vibrate throughout two syllables. Sequences of /ji:/ with a syllable beginning and ending with a sonorant appear to be a desirable choice to maintain the vocal cord tension during the continuous voicing.

All of the disyllabic words are represented with characters and are natural sequences, i.e. the words make sense to native speakers. Thus, the requirements of both voicing throughout and meaning place limits on the choice of possible combinations. Under these requirements, choosing absolutely identical phonetic elements and syntactic structure seems to be impossible. Shen (1990: 283) uses nonsense sequences for the test words but has to instruct the informants not to use the normal stress pattern and this results inevitably in an artefact. Potisuk *et al* (1997: 25) claim that all their test words are in a fixed prosodic environment, but alternative sonorants are inevitably used. As the aim of this study is to investigate tonal coarticulation, only the target syllable is restricted to having the same

phonemes. The disyllabic words used in this investigation are isolated words, i.e., not in carrier sentences, and are placed in a reading list in order to avoid tonal assimilation between the test words and other neighbouring words.

2.1.2.2. Speakers

The informants, JCCF, JHDG and LTHJ, are all native speakers of Cantonese born in Hong Kong and are currently students at SOAS. One female (JCCF) and two males (JHDG and LTHJ); two are aged 20+ (JHDG and LTHJ) and one 40+ (JCCF); one undergraduate (JHDG) and two postgraduates (JCCF and LTHJ).

2.1.2.3. Recording

All words are written with characters in a randomised list and were read by the three informants six times in two sessions in a natural and non-emotional voice, at a consistent normal speaking rate and avoiding a list-reading voice. A sufficient pause between items is provided (about 4 seconds interval) to enable the informants to feel comfortable and to avoid the list-reading effect. In order to avoid unwanted start and end effects, a few extra words are added at the beginning and the end of the reading. The microphone-to-mouth distance was maintained at 30cm. A phone-hood covered the microphone to improve the quality of recording. The recording was made in a sound-proofed booth at SOAS using a laryngograph and a Bruel and Kjœr Electret condenser microphone with a measuring amplifier type 2609 into a Sony PCM. The PCM converts the analogue signal into a digital signal which is further adapted to a video signal and recorded onto a Sony Betamax Video Recorder. The recordings are then played into a laryngograph processor (this allows levels to be adjusted) and then into the PCLx SPG and the Speech Workstation (SW) software packages which are installed in a 486 Dx 66 16 M Ram PC computer.

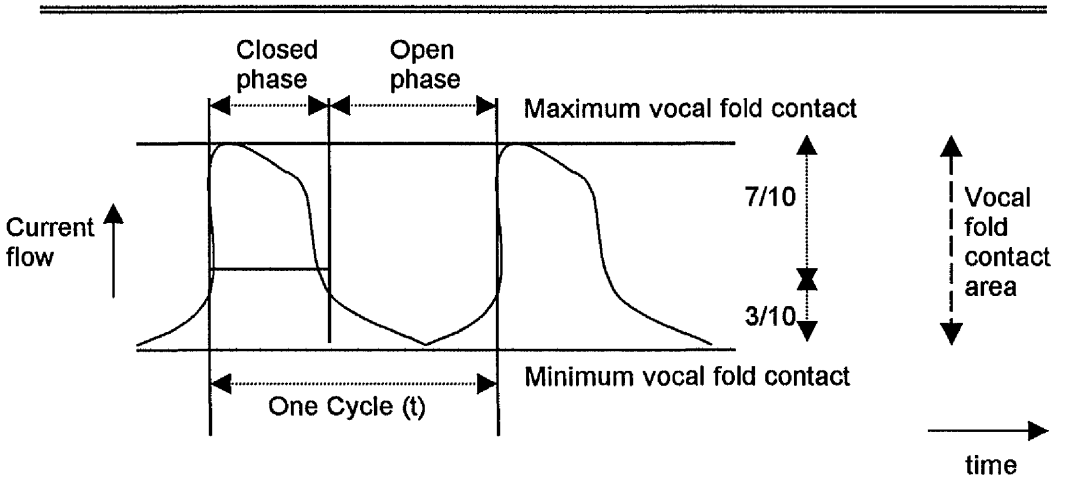
2.1.2.3.1. Laryngograph

The laryngograph is a device enabling the vocal fold contact area to be investigated noninvasively (Fourcin *et al* 1971: 172). Two electrodes, with the help of a bondage wrapping on the neck, are placed on both sides externally at the level of the larynx. When

the vocal folds vibrate, the current flow passing between the electrodes changes. With no vocal fold contact, the current flow between the electrodes is at the minimum; with full vocal fold closure, the current flow is at the maximum. The output of such changing current flow which passes between the electrodes is called the laryngograph waveform (Lx). In one cycle of vocal fold vibration (t) in a laryngograph waveform, the frequency (f) of the vibration is calculated as f = (1/t).

Each waveform cycle (t) can be analysed into two parts: one corresponds to the closed glottis and is called closed phase and the remaining part is the open phase. The closed phase part of the cycle expressed as a percentage is known as the closed quotient (CQ). The closure is from the first vocal fold contact to the maximum vocal fold contact in the waveform. This is based on the assumption that the vocal folds snap together more rapidly than they part. The ratio between the closed phase and the open phase in this case is 7:3. This ratio is essentially arbitrary, but has been found convenient and reliable in previous studies (Davies *et al* 1986: 539). The area of the maximum vocal fold contact is used by the laryngograph analyser software to calculate the periodicity of the waveform. The closed quotient is calculated as CQ = [(closed phase / fundamental cycle of Lx) x 100]%) (Howard 1995: 165). Increases in the closed quotient are in response to the vocal folds abduction resulting in glottal closure.

Fig. 2.1. Illustration of laryngograph output waveform and the closed quotient measurement.



2.1.2.4. Measurement

The laryngograph recordings can be displayed as waveforms and spectrograms in both packages; F_0 contour, Lx waveform and closed quotient contour (Qx) can be derived from the laryngograph trace and displayed simultaneously in the SPG. The packages provide the optimum display for clarity of measurement. As each single vocal pulse can be clearly demonstrated in the Lx waveform, the duration of the disyllabic words is taken from the first vocal pulse to the last. All tokens are digitised at a 10 kHz sampling rate.

2.1.2.4.1. Segmentation

Time points of syllable boundaries are determined by inspection of the speech waveform, the wideband spectrogram, and the Qx and by listening to the signals. The segmentation for intervocalic nasals is relatively straightforward, since the velum opening of the nasals shows an abrupt change between the offglide movement from the preceding vowel to the steady formant pattern, then to rapid onglide movement of the following vowel in the wideband spectrogram. Also the speech waveform can clearly reveal the boundaries between the nasal and its neighbouring vowels - vowels have higher amplitude than the nasals. The starting point of a lateral approximant /l-/ as the initial of a syllable is defined by observing the abrupt stop of the off-glide from a vowel of the preceding syllable to the steady formants and by the noticeably sudden change of the waveform pattern. The F2 movement provides a good cue for defining the boundary of an intervocalic glide /w/. The region in which the F2 slips down and meets with F1 is considered as the beginning of an initial /w-/. The region in which the F2 shows the start of a rapid rise is considered as the end of an offglide /-w/ and the beginning of a high vowel. The segmentation for the semi-vowel /j/ seems to be problematic. However, /j/ is stable in formants, the frequency of its F3 is much higher than that for other vowels. The energy of its F3 is much greater than that of its other formants, especially at the beginning or soon after the beginning of the F3. Thus, a steady state of high frequency and high energy of F3 serves as the starting point of /j/. When an onset /j-/ follows a coda /-j/ or a vowel /-i/, i.e., in sequences of /-i/ # /j-/ and /-j/ # /j-/, its formants still enjoy their steady state, but the energy is weaker compared with its predecessor(s). /-i:/ is a

full vowel which has steady formants and usually has high energy in all formants during voicing. The F2 of /i:/ is rather concrete in comparison with /j/ whose F2 is rather faded. The formants of a coda /-j/ are always affected by its preceding main vowel. F2 especially performs a rapid rising from a preceding open or back vowel to the /-j/. /j/ is fully voiced but has less friction in an onset position than in a coda position. Thus, the starting point of steady formants of /j-/ where it shows a relatively weak intensity - is considered as the syllable boundary between the coda /-j/ and the onset /j-/. Apart from the conventional segmentation method, the Qx is useful because it demonstrates the increase in the closed quotient of the glottal cycle associated with vocal fold abduction. Generally speaking, the higher the Qx, the lower the frequency - i.e., the Qx contour and the F_0 contour are in inverse direction. Any turning point in the F_0 contour more or less accompanies (but not necessarily) a turning sign in the Qx contour at the same time. Changing in segments is also signalled in the Qx contour by increasing or decreasing the closed quotient. The first F_0 value is taken as the tonal onset of the first syllable and the last F_0 value is taken as the offset of the last syllable. It is not abnormal to see the last few flaps of the vocal folds drop the F_0 contour. This signals that the vibration of the vocal folds is weakened when approaching the end of the vibrations. In this case, the last F_0 value is not necessarily reliable, so I use the F_0 value just before these drops (usually it sits approximately 20ms before the end). When the same phenomenon occurs at the beginning of the F_0 contour, i.e., the vocal folds have one or two weak vibrations at the beginning, and this would lead the F_0 contour to show a weak start, the first F_0 value after the weak start is used. It is observed that the F_0 contour is not necessarily a smooth contour. It has a few turning points. The turning points are observed to occur especially more or less a quarter of the duration of the syllable after the tonal onset or before the tonal offset. Therefore, it was decided to take the F_0 value at three points apart from the onset and offset: a quarter, half and three quarters of the total duration of a syllable. The values of these three points plus the values of the onset and offset are taken into computation to determine the F_0 contours. It is not uncommon to

produce a creaky voice in low frequency words in which the F_0 contours become aperiodical. A handful of tokens which result from this and have a long period of aperiodical F_0 contour were excluded from the investigation together with some tokens judged to be unrepresentative. Subsequently many tokens had to be discarded in order to obtain an equal number of tokens for each set of the two-tone sequences. 648 tokens in total were involved: 6 tones x 12 preceding/following contexts x 3 repetitions x 3 informants. Thus a total of 3240 F_0 values (5 measurement locations x 648 tokens) were taken into computation.

2.1.2.4.2. Normalisation

For the reason that not all tokens of each tone are of equal duration, each syllable is normalised for duration on a percentage scale. F_0 contours are also normalised for each of the speakers by using the formula: $F' = 100^* (F_0 - B_{Hz}) / R$. ($F'$: the normalised value; R: F_0 Range of each speaker = $T_{Hz} - B_{Hz}$; T_{Hz} and B_{Hz}: top and bottom Hz values of each speaker's F_0 range estimated from a reasonable sample of their speech). By such normalisation, F_0 contours may be compared across speakers. F_0 contours were smoothed by curve-fitting for display purposes only.

2.1.3. Results and Discussion

2.1.3.1. Monosyllables

2.1.3.1.1. Shape and Height in the Six Tones

First of all, I recorded eighteen monosyllables of /ji:/ with six tones, three different lexical items for each of the six tones and took their F_0 mean in order to give a brief description of the characteristics of the tones when they are pronounced in isolation.

The following figures 2.2-2.3 display the six tones in monosyllabic utterances. Figure 2.2 illustrates the 6 tonal patterns from the individual informants and figure 2.3 illustrates the 6 tones separately across the three informants.

Fig. 2.2 (a-c). Mean F_0 contours of the 6 tones pattern in monosyllables produced by JCCF, JHDG and LTHJ, respectively, on the following set of words: /ji:[1]/ (medical/ cloth/ aunt), /ji:[2]/ (chair/ rely on/ lean on), /ji:[3]/ (meaning/ Italy/ thought), /ji:[4]/ (son/ and/ move), /ji:[5]/ (ear/ already/ discuss), /ji:[6]/ (two/ righteousness/ easy).

2.2.a 2.2.b 2.2.c

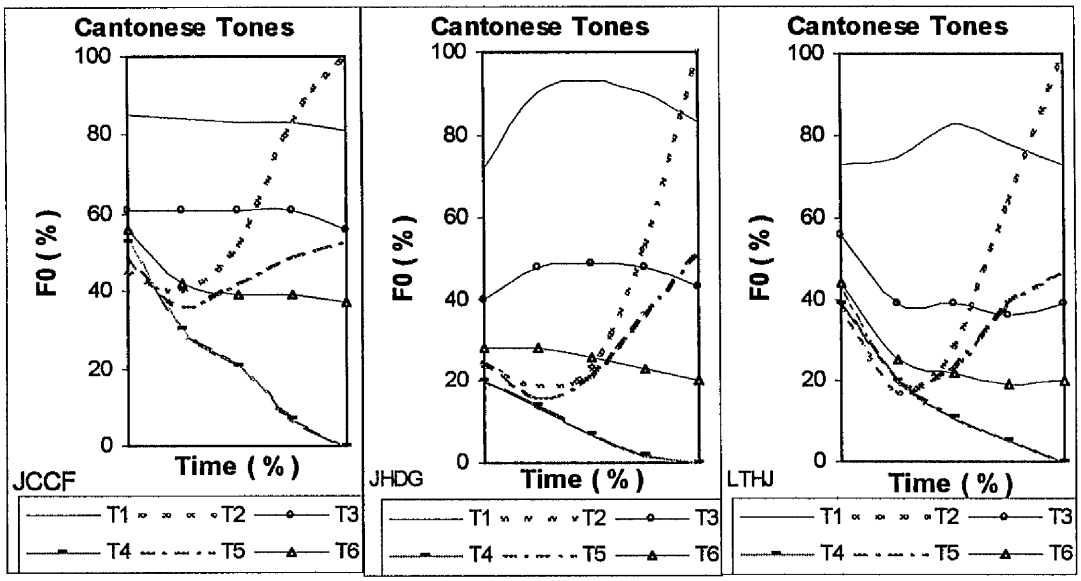
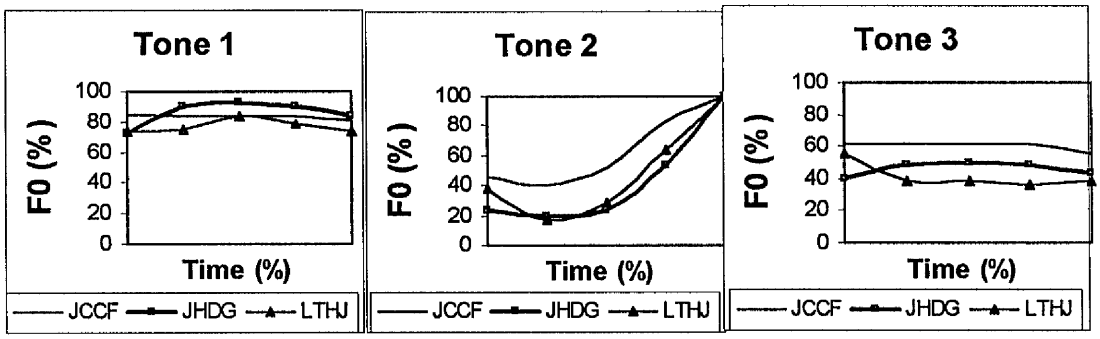
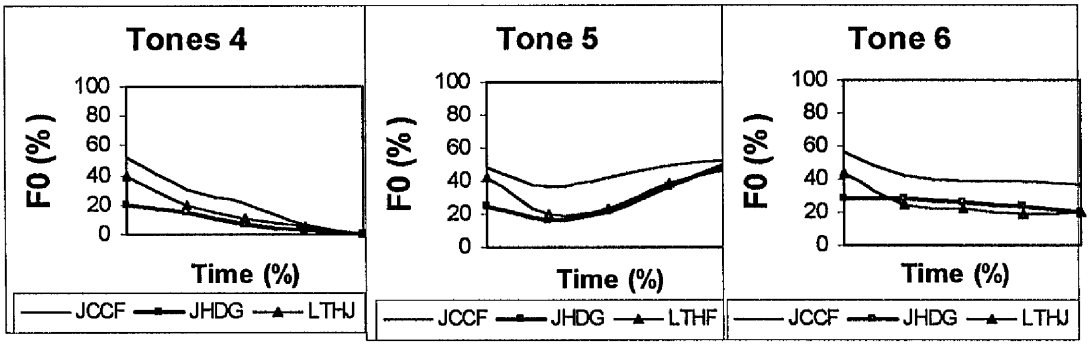


Fig. 2.3 (a-f). Mean F_0 contours of the six tones in monosyllables arranged by individual tones, on the same set of words as in figure 2.2.

2.3.a 2.3.b 2.3.c



2.3.d 2.3.e 2.3.f



T1 is a high level tone. Its shape is like a clothes-hanger: high in the centre but low at both ends. Informant JHDG has the most arched curve shape and JCCF the least. There is no falling variant among the informants. The onset is the highest compared with that of the other tones. It apparently stands out from other tones by virtue of its overall high pitch level.

T2 is a high rising tone. The rise is preceded by a slight dip. The turning point is about a quarter of the duration of the syllable from the onset. The rise starts slowly, and when it reaches half of the duration of the syllable, it suddenly jumps up sharply. The `sharp jump' appears in the last three quarters of the tonal duration for JCCF, in the last half portion of the tonal duration for JHDG and in the last 60% of the tonal duration for LTHJ. The `sharp jump' separates T2 from the slow rising T5. The offset is the highest compared with other tones.

T3 is a mid level tone. It keeps constantly level particularly for the middle portion of the tonal duration. The slope in the first and the last quarter of the tonal duration differs from the informants. Its onset is above that of the other tones apart from T1, and its offset is above that of T4 and T6.

T4 is a falling tone, and is the only falling tone. Its onset is squeezed with T2, T5 and T6 in a mid low range. It falls all the way. The fall ends as the lowest offset compared with the other tones.

T5 is a low rising tone. It drops at first, then slowly climbs up about for the last three quarters of duration of the syllable. Its offset just meets up with T3. The order of the height of the offset between T3 and T5 differs among the informants (T3 is higher than T5 for JCCF while T5 is higher than T3 for the others).

T6 is a low level tone. It keeps constantly level for the last three quarters of the tonal duration. Its onset is just below that of T3 but the offset is far below T3. It has an immediate fall after the onset for a quarter of the duration of the syllable for informants JCCF and LTHJ.

2.1.3.1.2. Hierarchy of the Height of the Onsets and the Offsets of the Six Tones

Shape and height are both important to Cantonese tones. Overall, the onsets are crowded in while the offsets spread over. The height of the onsets is particularly crowded in the lower

part of the frequency range where T6, T4, T5 and T2 are squeezed, but T1 and T3 stretch above them. Incidentally, the onset of T6 just comes above the onsets of T5, T2 and T4 and the order of the last three is not consistent across the informants. The hierarchy of the height of onsets in this study agrees with that in Ching's informants (see figures in Ching 1981: 49) - the highest onset is T1, the next is T3, then the others' orders differ from one informant to another. The hierarchy of the height of the onsets of T2, T4, T5 and T6 also displays a non-uniform pattern in the data of Fok's informants (see figures in Fok 1974: 149-154). This implies that the F_0 values in onsets are not so important as constituent factors in the identity of tones (particularly T2, T4, T5 and T6).

The hierarchy of the height of the offsets is clearly located as T2 is the highest and T4 is the lowest; T1, T5, T3 and T6 (listed in order from the highest to the lowest) are in between; T3 and T5 are very close together and sometimes in reverse order across the informants. The hierarchy of the height of the offsets in this study agrees with that in Ching's informants but agrees with Fok's only where her T1 is a high level tone (see figures in Fok 1974: 149) (her T1 has a high fall variant). This implies that the F_0 values in offsets are not very important either, at least not T3 and T5, though more important than the F_0 values in onsets.

2.1.3.1.3. Summary

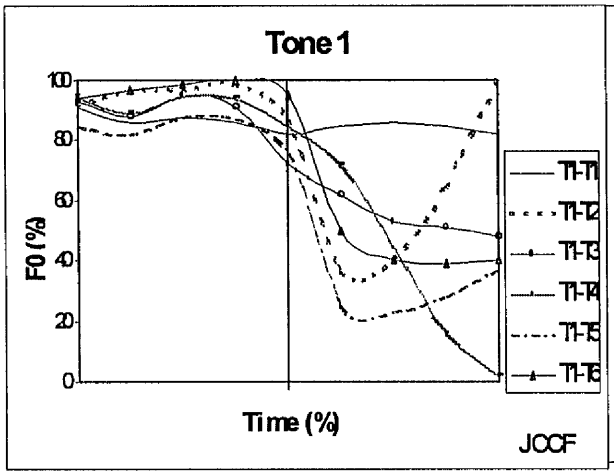
The level tones differ from each other in overall height particularly in the last three quarters of the tonal duration. The rising tones differ from each other both in height and in slope: the heights differ in the last two quarters of the tonal contour as a result of their different degrees of rising - T2 has a steep rise and T5 has a gentle rise. The falling tone differs from others both in height and in shape. The mid range tones T3 and T5 differ from each other primarily in shape - one is level, one is rising.

2.1.3.2. Disyllables

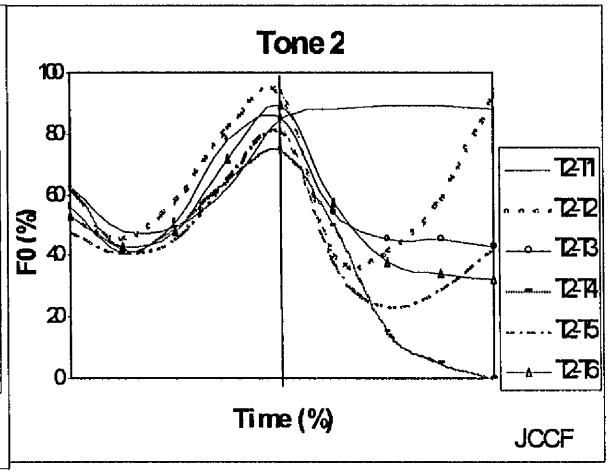
Fig. 2.4 (a-c). Mean F_0 contours of T1 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing anticipatory effects. Inside each figure, the vertical line indicates the tonal boundary.

Fig. 2.5 (a-c). Mean F_0 contours of T2 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing anticipatory effects. Inside each figure, the vertical line indicates the tonal boundary.

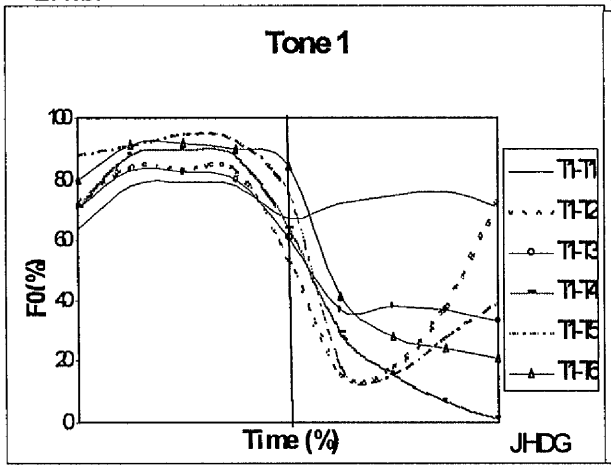
2.4.a.



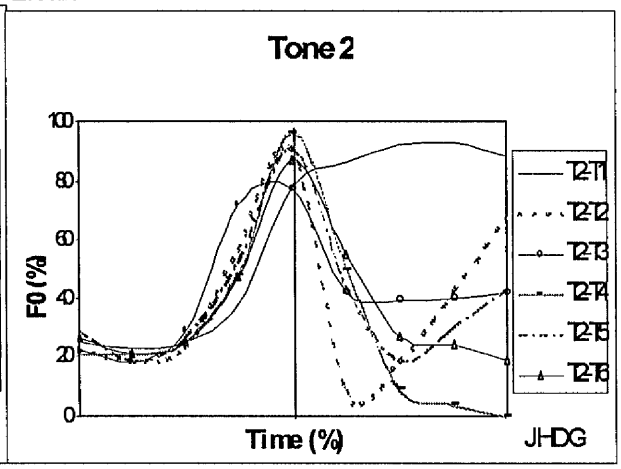
2.5.a.



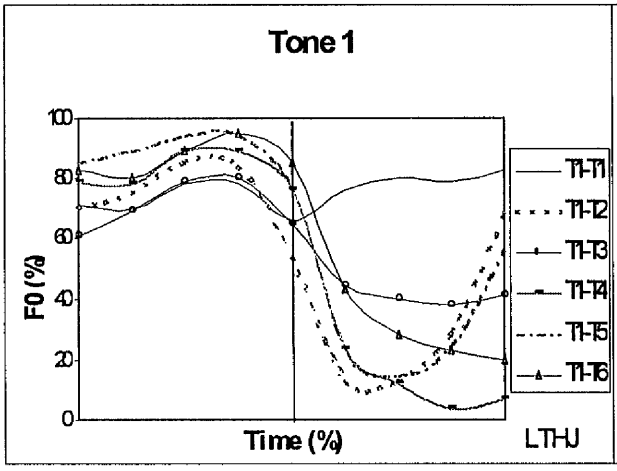
2.4.b.



2.5.b.



2.4.c.



2.5.c.

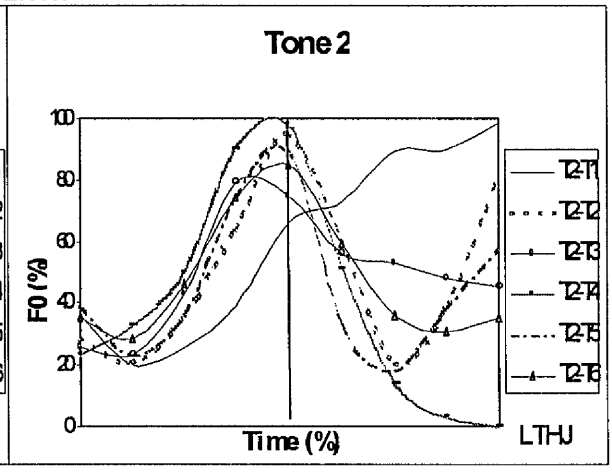
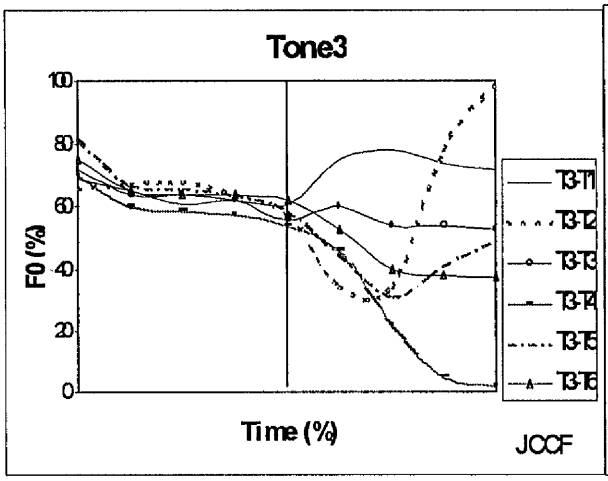


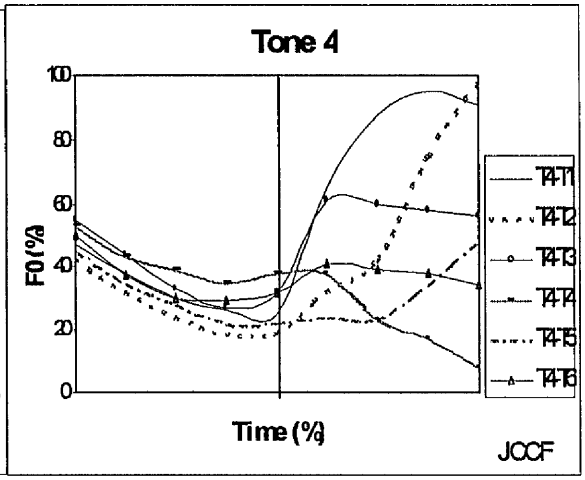
Fig. 2.6 (a-c). Mean F_0 contours of T3 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing anticipatory effects. Inside each figure, the vertical line indicates the tonal boundary.

Fig. 2.7 (a-c). Mean F_0 contours of T4 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing anticipatory effects. Inside each figure, the vertical line indicates the tonal boundary.

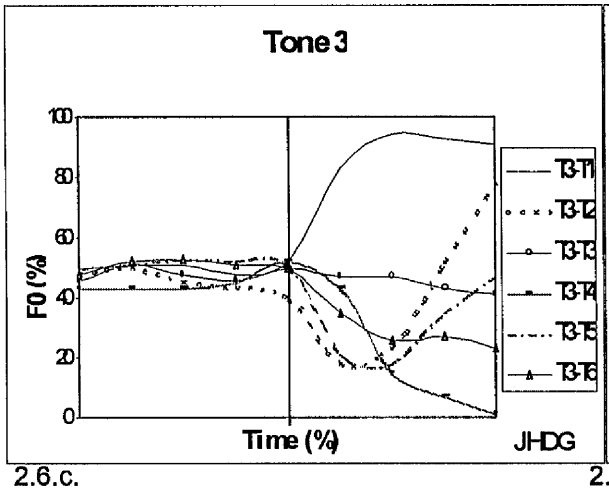
2.6.a.



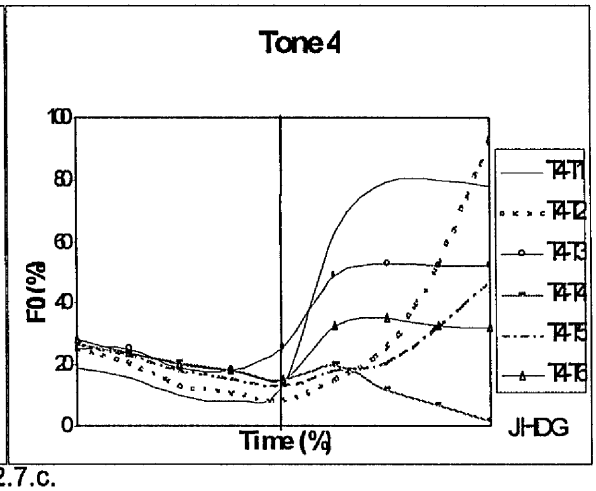
2.7.a.



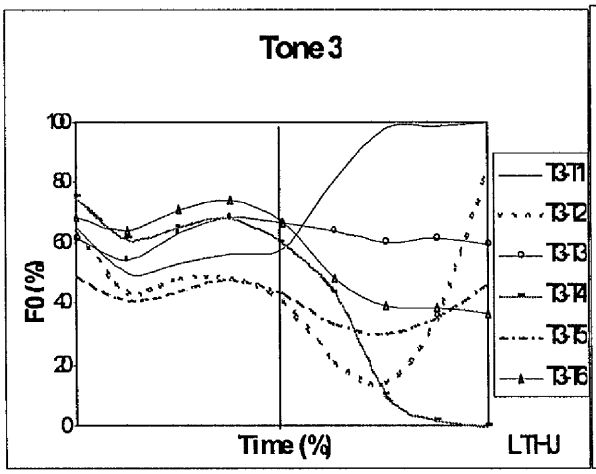
2.6.b.



2.7.b.



2.6.c.



2.7.c.

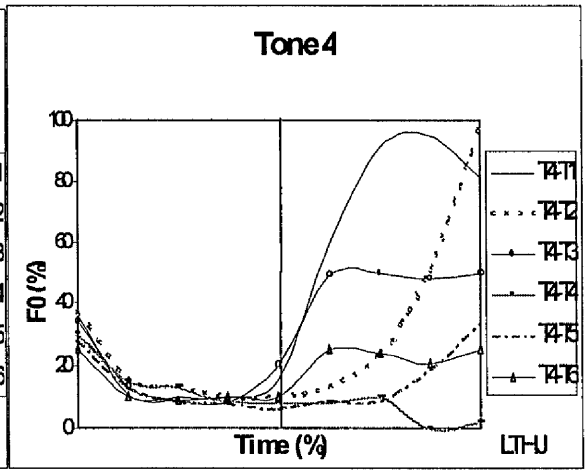
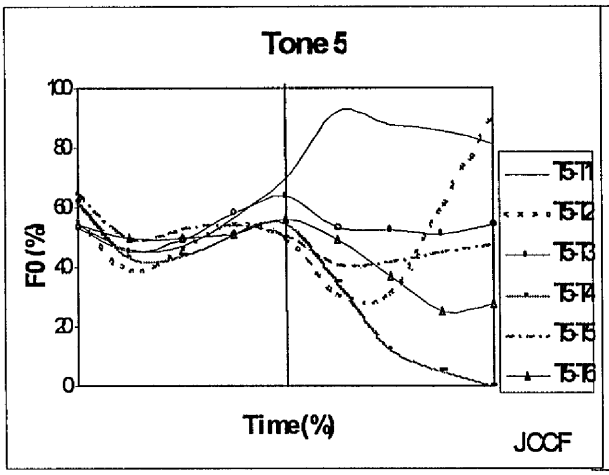


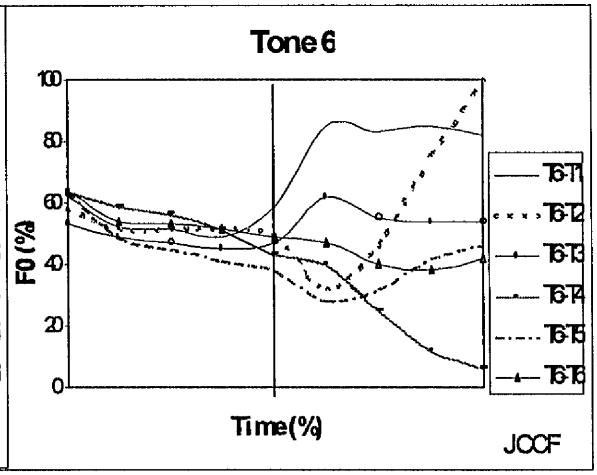
Fig. 2.8 (a-c). Mean F_0 contours of T5 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing anticipatory effects. Inside each figure, the vertical line indicates the tonal boundary.

Fig. 2.9 (a-c). Mean F_0 contours of T6 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing anticipatory effects. Inside each figure, the vertical line indicates the tonal boundary.

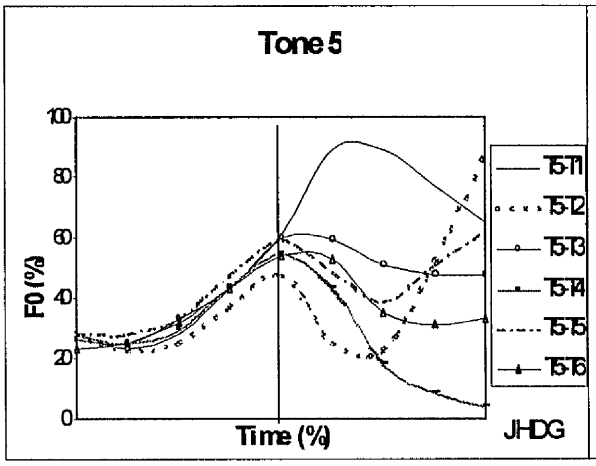
2.8.a.



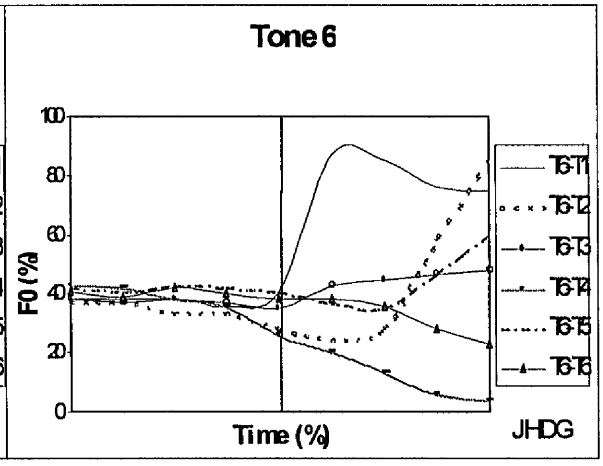
2.9.a.



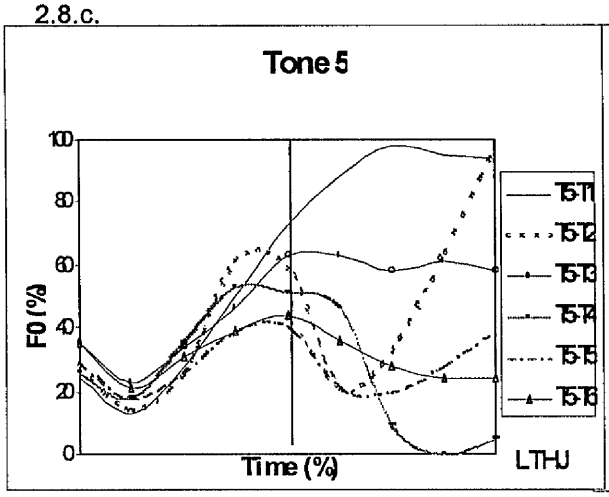
2.8.b.



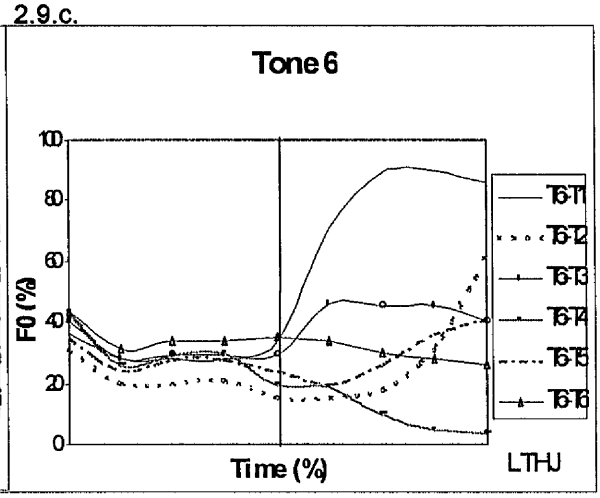
2.9.b.



2.8.c.



2.9.c.



2.1.3.2.1. Coarticulatory Effects on Individual Tones

2.1.3.2.1.1. Anticipatory Effects

The figures 2.4-2.9 illustrate the six target tones as the first tone of the disyllabic words, displaying their phonetic variations as a result of the influence of their following tones.

2.1.3.2.1.1.1. Phonetic Variation in the Height of Individual Tones

All informants show a greater height variation in T1 than in the other tones. T1 is significantly lower all the way through when followed by T1 than when followed by other tones. T1, on the other hand, remains at a higher level when followed by T5 or T6 than when followed by other tones. This can be ascribed to a dissimilatory effect on height and such an effect to a different extent can be seen in other tones.

T4 is apparently higher when followed by T4 than when followed by other tones. T2 and T5 are raised by different low onset tones across the informants. For informants LTHJ and JHDG, T2 is ostensibly higher at the last quarter of the tonal duration when followed by T4. For informant JCCF, T2 raises the preceding T2 much more than T4 does.

All informants show a smaller height variation in T5 than in the other tones. This may be because it begins in the first half as high as that of T2 but in the second half it rises less than T2, ending in between T6 and T2 in its canonical form. T5 is slightly higher when followed by T5 than when followed by other tones for the first three quarters of its duration for informants JCCF and LTHJ.

T3 and T6 have no uniform pattern of height hierarchy caused by the following tones among the informants.

Across target tones, T6, T5, T4 and T2 raise the height of the target tone compared with T1 and T3 which tend to set the target tone at a middle or low position. T1 is not found to raise any target tones. This indicates that the anticipatory effect on height is significant and constitutes a dissimilatory effect which can extend backwards to the onset of the target tone.

This finding regarding how far the anticipatory effect on height extends is similar to that in Mandarin, where "the effects are exerted on the whole tone" (Shen 1990: 284) but different from that in Thai, where "anticipatory effects extend backward to about 50% of the duration of preceding syllable" (Gandour *et al* 1994: 489). However, our finding regarding whether the anticipatory effect on height is assimilation or dissimilation agrees with that in Thai, where "a low onset of the following tone raises the height of the preceding target tones that have higher F_0 offsets whereas a high onset of the following tones lowers the height of preceding target tones with lower F_0 offsets" (Potisuk *et al* 1997: 33). This is in contrast to Mandarin, where a low onset lowers and a high onset raises the preceding offset, for example, "tonal offsets are lowered the most by tone ∅ and raised the most by tone 4, and the next most by tone1." (Shen 1990: 285)

2.1.3.2.1.1.2. Phonetic Variation in the Slope of Individual Tones

All target tones preserve their canonical shape despite some small assimilatory modifications at the last quarter of the tonal duration. T1 maintains its fairly level contour with some fluctuation. The centre point (at 50% of its duration) seems to be a fulcrum of a scale which maintains the height but allows the two sides to modify slightly. T1 displays no uniform pattern for the first quarter of the tonal duration. It dips at the last quarter of its tonal duration in order to meet all other low onset tones. The dipping phenomenon is also seen when followed by a T1. This may suggest that the change in syllable boundary has an effect on the frequency of vibration of the vocal folds.

T2 maintains its small fall followed by a big rise, but reduces its dramatic `flying-up' tail at the last quarter of its duration to a moderate rise. The bending-off slope at the last quarter of the duration of T1 and the slow-rising slope at the last quarter of the duration of T2 exhibit a little anticipatory assimilation on the slope to the following tone. T2 exhibits a more steeply rising slope from the dip (25% of its duration) to the offset when followed by T2 and T4 than when followed by other tones. On the other hand, the slope of T2 is the least steep when followed by T1. The last quarter of the tonal duration even exhibits a small fall when followed by T3.

This shows that dissimilation primarily affects the height (as discussed above) and hence also affects the slope. This finding on the behaviour of the T2 runs counter to Hashimoto's observation in Cantonese that the rising tones in context tend to have a more prominent rise if followed by higher frequency tones (1972: 93).

T3 maintains its fairly level shape but with a small downward slant. Even though there are traces of modifications at the offset, very little anticipatory change is found in T3.

The slope of T4 has different characteristics among the informants: for JHDG, T4 falls in a straight line; for LTHJ, T4 falls heavily at the beginning but levels out over the rest of the tonal body; for JCCF, T4 maintains its falling shape but dramatically reduces the extreme drop in its tail to a moderate falling or a levelling out or a moderate rising for the last half of its duration. The fall of T4 behaves in the same way when followed by any of the six tones.

T5 maintains its small fall followed by a gently rising shape. But for informant JCCF, the last quarter of the tonal duration of T5 sometimes (when followed by T2 and T5) changes its direction to a fall. The change of direction in the last quarter of the tonal duration is also found when followed by T4 and T5 for informant LTHJ. The rising of the last portion of T5 is significantly steeper when followed by T1 than when followed by other tones across the informants.

T6 maintains its fairly level shape with slightly downward slanting. The first quarter of its tonal duration maintains a gentle slant down in the data of informants JCCF and LTHJ while being level in the data of informant JHDG. At the last quarter of its duration, it turns from downward to upward only when followed by T1. This indicates that assimilation is present.

The phenomenon of changing direction of F_0 in the last portion of the target's duration is also seen in Vietnamese (Han and Kim 1974: 227), where the curve tone changes its upward tail to downward when followed by a level tone. In Mandarin, tone 1 changes from a shallow downward falling-rising contour to a falling contour when before tone Ø, while tone Ø

generally has a falling contour across tonal environments (Shen 1990: 284). [Shen distinguishes tone sandhi from tonal coarticulation: the former is attributed to language specific morphophonemic constraints while the latter is attributed to language independent biomechanial constraints.] Across the six tones, anticipatory effect on slope in Cantonese is minimal and is assimilatory in nature extending backward to a quarter of the duration.

2.1.3.2.1.2. Carryover Effects

The figures 2.10-2.15 illustrate the six target tones in the position of the second syllable of disyllabic words, displaying their phonetic variation as a result of the influence of their preceding tones.

2.1.3.2.1.2.1. Phonetic Variation in the Height of Individual Tones

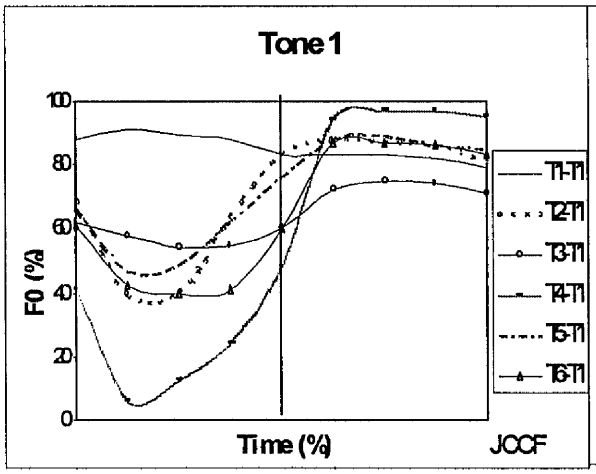
Regarding the effect of coarticulation on height, two sides of the story are to be considered: first, the beginning part of a tone contour, which is about the first quarter of its duration; second, the main body of a tone contour, which is about the last three quarters of its duration. The beginning part of a tone contour tends to exhibit an assimilatory effect, whereas the main body of a tone contour tends to exhibit a dissimilatory effect. Generally speaking, the onsets of the six target tones are raised by the preceding high offset tones - T2 and T1, and pulled down by the preceding low offset tones - T4 and T6. Whatever the target tone is, the height of its onset is ranked in order from high to low when it is preceded first by T2 or T1, then by T3 or T5 or T6, and then by T4 (with a very small number of exceptions, for example, the preceding T3 comes as the second highest onset of T1 for LTHJ). This order is as in that of the tonal offsets in monosyllabic words. This picture once again confirms that whatever modification in height or slope is made on the preceding tone, the resistant effect on the preceding tone is much stronger than that on the following tone. In other words, the carryover effect in Cantonese tones is much stronger than the anticipatory effect.

However, this assimilatory nature of the carryover effect in the beginning portion of a target tone changes in the main body of the target tone. All informants show a greater height variation in T1 than in the other tone. T1 remains at a low level when preceded by T1 (the

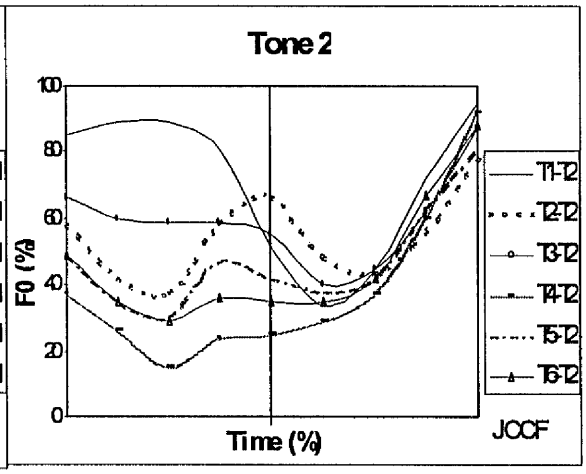
Fig. 2.10 (a-c). Mean F_0 contours of T1 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing carryover effects. Inside each figure, the vertical line indicates the tonal boundary.

Fig. 2.11 (a-c). Mean F_0 contours of T2 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing carryover effects. Inside each figure, the vertical line indicates the tonal boundary.

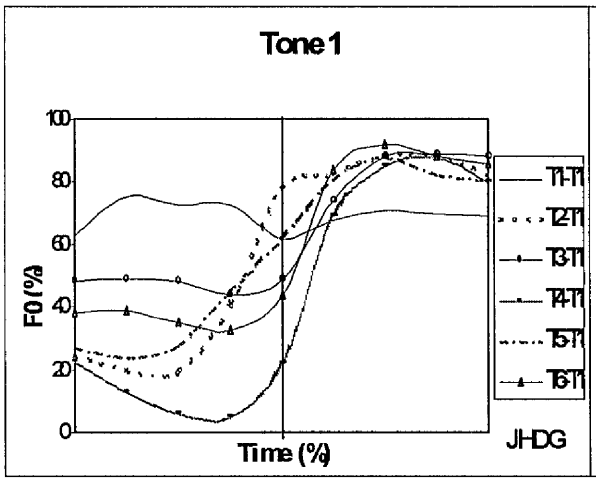
2.10.a.



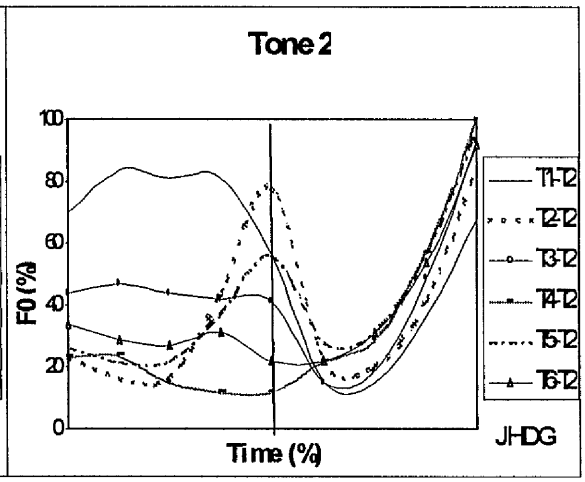
2.11.a.



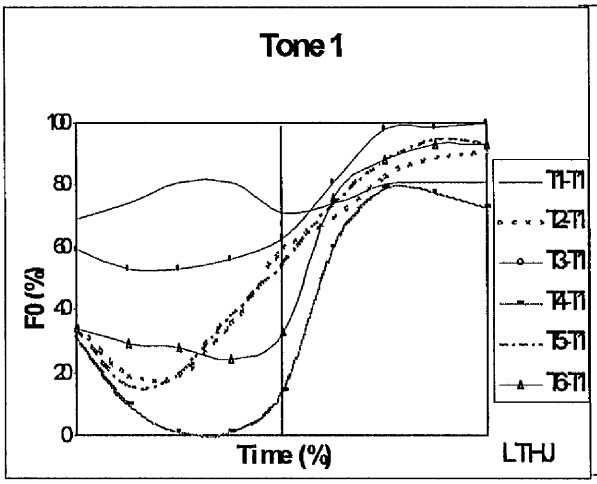
2.10.b.



2.11.b.



2.10.c.



2.11.c.

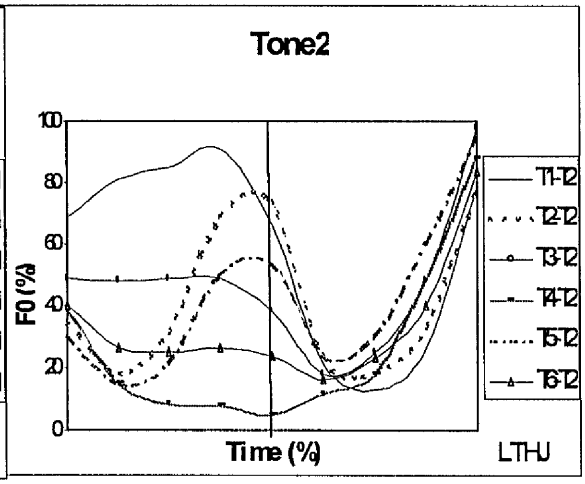
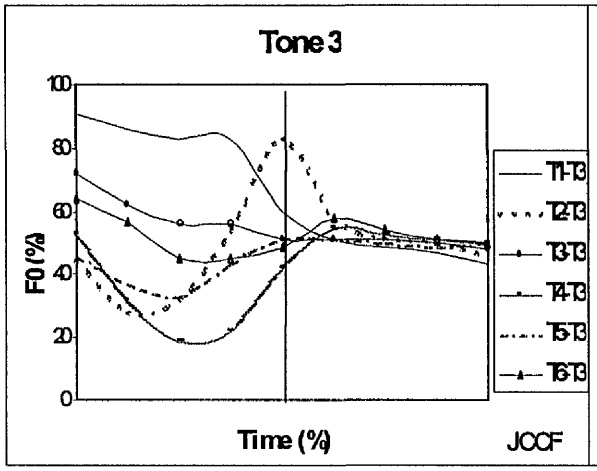


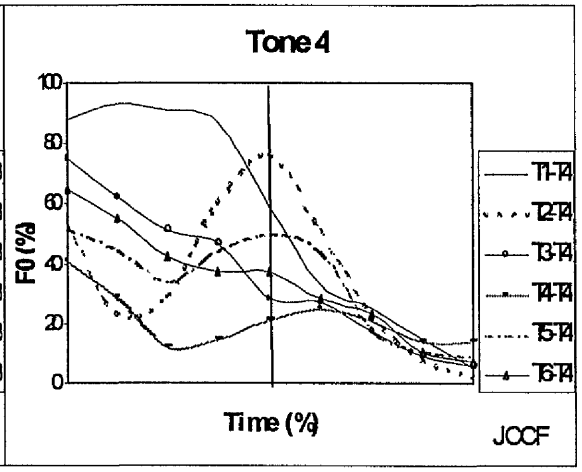
Fig. 2.12 (a-c). Mean F_0 contours of T3 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing carryover effects. Inside each figure, the vertical line indicates the tonal boundary.

Fig. 2.13 (a-c). Mean F_0 contours of T4 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing carryover effects. Inside each figure, the vertical line indicates the tonal boundary.

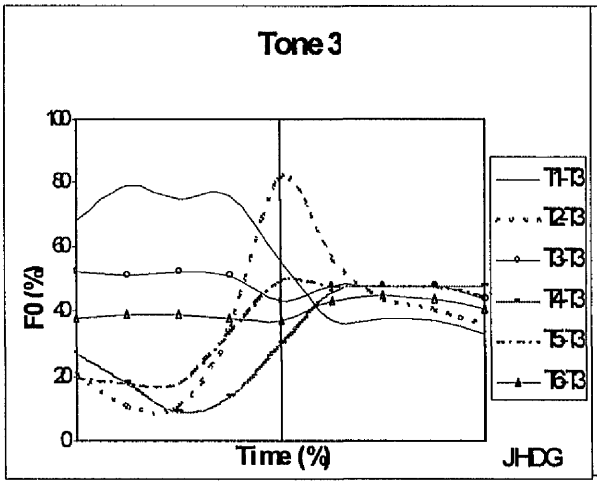
2.12.a.



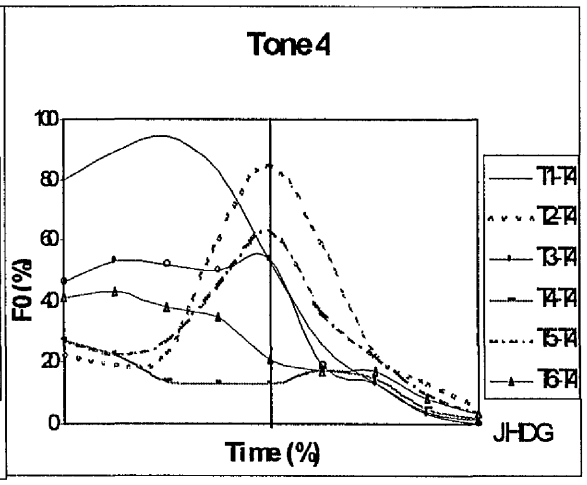
2.13.a.



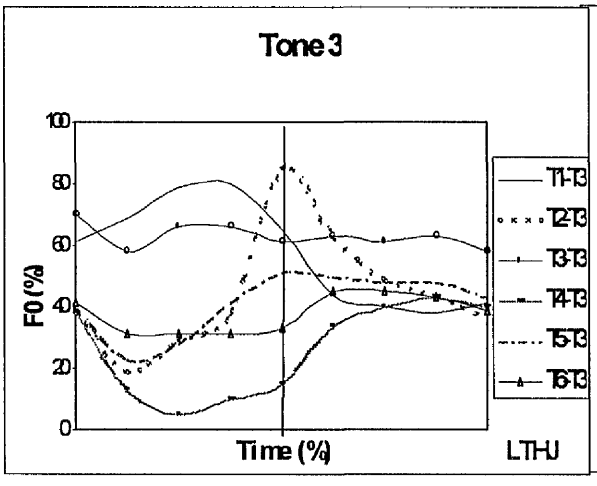
2.12.b.



2.13.b.



2.12.c.



2.13.c.

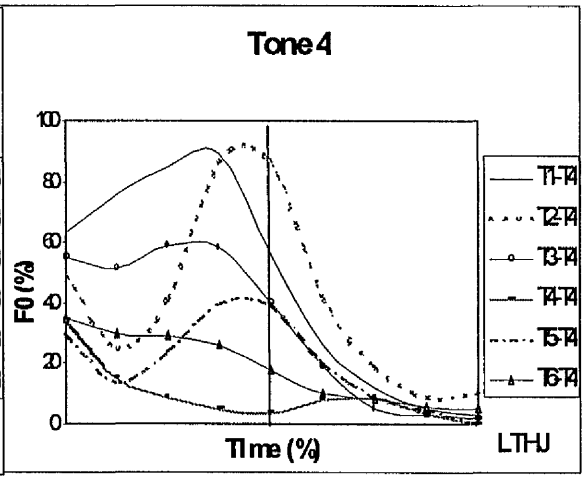
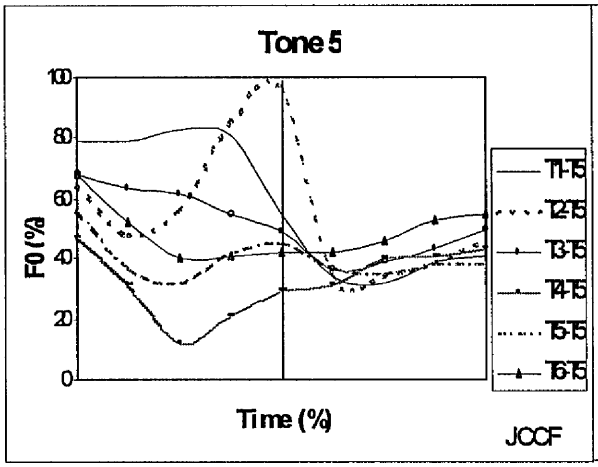


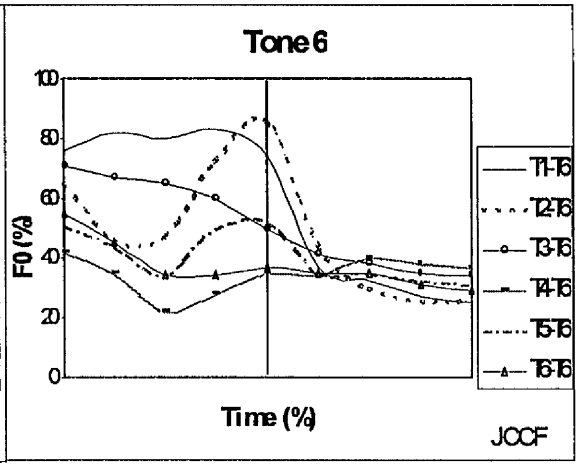
Fig. 2.14 (a-c). Mean F_0 contours of T5 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing carryover effects. Inside each figure, the vertical line indicates the tonal boundary.

Fig. 2.15 (a-c). Mean F_0 contours of T6 and its following tones produced by JCCF, JHDG and LTHJ, respectively, showing carryover effects. Inside each figure, the vertical line indicates the tonal boundary.

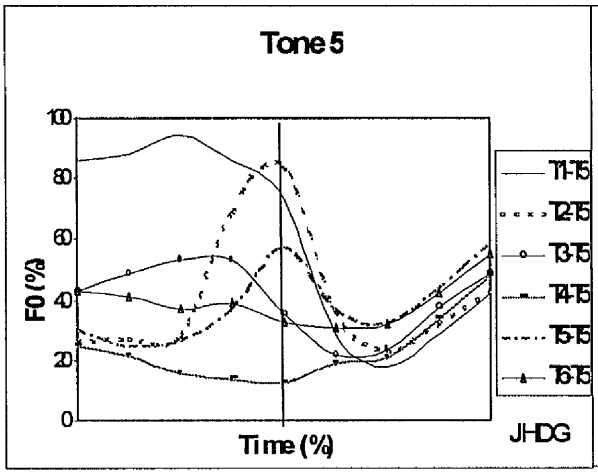
2.14.a.



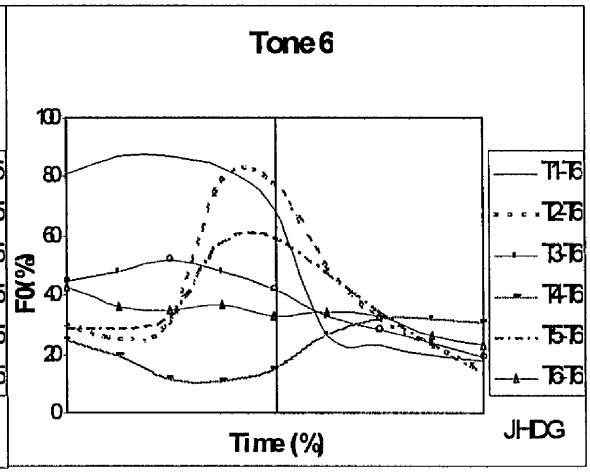
2.15.a.



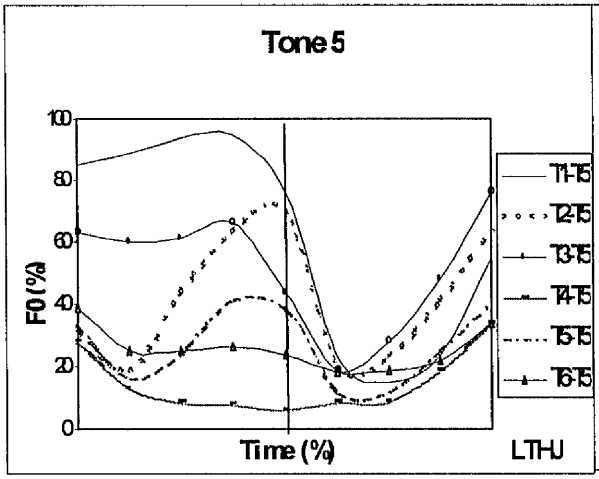
2.14.b.



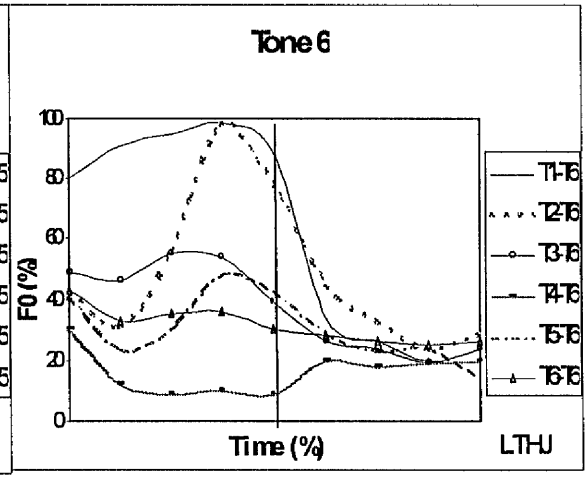
2.15.b.



2.14.c.



2.15.c.



lowest for JHDG and the penultimate lowest for JCCF and LTHJ). T1 is generally higher when preceded by tones other than T1.

T2 is apparently higher in the first half of its duration when preceded by T2 as a result of being raised by its high offset. However, it sinks down to the lowest offset soon after enjoying the highest position. Almost all onsets of the six individual target tones are the highest when preceded by T2 in comparison with other preceding tones. (Alternatively, the preceding T1 raises the following onset to the highest for T1, T5 and T6 for informant LTHJ, this may be due to tonal peak placement of the preceding T2 occurring much earlier). Even though the preceding T2 raises the onset of the following tones, it does not raise the overall contour of the following tones. What happens is that the target tone has to intersect with others in a quarter or a half of its duration when preceded by T2 but then settles in the lower level (the lowest for informant JCCF and the penultimate lowest for informants JHDG and LTHJ.) This finding is in contrast with Han and Kim's finding for Vietnamese, that "if the onset of a tone is raised under the influence of the preceding tone with a high overall pitch the end would also be raised, the amount of the rise being dependent on the amount of the rise at the onset" (1974: 228). This finding is also in contrast with Shen's finding for Mandarin that the entire tonal contours are shifted up or down by the surrounding tones (1990: 293).

T3 is lower when preceded by T1 than when preceded by other tones. T4 is higher when preceded by T2 than when preceded by other tones for informants JHDG and LTHJ. It is higher when preceded by T1 for informant JCCF. There is no uniform pattern for T5 among the three informants. T6 is apparently higher when preceded by T4 than by other tones for informants JCCF and JHDG. It is higher when preceded by T6 for informant LTHJ.

Across target tones, T6 and T4 preceding the target tones tend to raise the height of the target tones when compared to T1 and T2. All these once again confirm that the carryover effect on height manifests itself as dissimilation in the main body of a target tone, which is similar to that for the anticipatory effect. This finding is in contrast with Han and Kim's finding for Vietnamese that a tone tends to be a high variant before or after a broken, rising,

or level tone, and a low variant before or after a curving or falling tone. "Tones with a high overall pitch pull up the pitch of the tones in their immediate environment, and tones with a low overall pitch pull down the pitch of the tones adjacent to them" (1974: 226). Regarding the carryover effect on height, this finding is also in contrast with that for Thai claimed by Gandour *et al* that "carryover effects extend forward to about 75% of the duration of the following syllable" (1994: 489).

2.1.3.2.1.2.2. Phonetic Variation in the Slope of Individual Tones

The carryover effect on slope exhibits an assimilatory modification. All target tones maintain their canonical shape for the last three quarters of their duration. For the first quarter of their duration, the tone contours head up or head down from their onsets whose heights are different to their canonical position - the tonal onset of the second syllable is built on the offset of the first syllable which has small modifications towards the second syllable.

T1 remains level for the last three quarters of its duration and has a rising slope for the first quarter of its duration, even when preceded by T1 or T2. The occurrence of a small concave pattern in the juncture of the tone contour of the two-tone sequences of T1-T1 may suggest that change in syllable boundary is accompanied by change of frequency of vibration of the vocal folds.

T2 maintains its sharp rising slope in the last half of its duration. It has a falling contour in general (except when preceded by the very low offset T4) for the first quarter of its duration then starts rising gently from the beginning of the second quarter of its duration.

T3 maintains its fairly level shape with slight slanting downward for the last three quarters of its duration. It rises for the first quarter of the duration only when preceded by low offset T4 and T6.

T4 maintains its falling shape with different degrees of falling slope. The greatest steepness of fall is when T4 is preceded by the highest offset T2. It also reaches a very low offset in comparison with when preceded by other tones for informants JCCF and JHDG.

T5 maintains its gentle rising shape in the last three quarters of its duration. The beginning portion of T5 maintains a fall in general (except when following T4), the degree of the fall totally depending on the height of the offset of its preceding tone - a higher offset falls more and this results in a steeply downward slope; a lower one falls less and the result is a gentle downward slope. This partly agrees with Hashimoto's (1972: 93) observation for Cantonese that "... the rising tones in context tend to have a greater fall at the onset if preceded by higher frequency tones."

T6 exhibits a similar slope under the same conditions as T5 for the first quarter of its duration. It maintains a fairly level shape with gentle slanting downward for the last three quarters of its duration.

Across the six tones, all of them maintain their canonical slopes for the last three quarters of their duration, but may change their direction in the first quarter of the duration depending on the height of the preceding offset. In other words, the carryover effect on slope only extends to the first 25% of the duration of the target tones and the effect is assimilatory in nature. For Thai, two contrasting results are obtained by two separate experiments: Gandour *et al.* claim that in carryover assimilation, slope is relatively unaffected (1994: 489); Potisuk *et al.* claim that the carryover effect in slope extends forward about 70% of the total duration of the following syllable (1997: 32). They conclude that the different results are evidently caused by the presence or the absence of continuous voicing across the syllable boundary. Gandour *et al.* use intervening voiceless plosives at the beginning of the target syllables; Potisuk *et al* 's utterances are continuously voiced throughout. They think the effect of voiced onset on the F_0 contour manifests itself in stronger assimilatory effects. The purpose of the design of continuous voicing across the syllable boundary in this study is to enable a non-stopping F_0 contour throughout the adjacent tones.

Comparing the modification on slope in the two directions from the syllable boundary, the carryover effect is obviously greater than the anticipatory effect, even though both last for only about the first 25% of the tonal body (starting from the syllable boundary).

2.1.3.2.2. The Pattern of the Six Tone Contrast in Different Tonal Environments

The twelve figures above (2.4 –2.15) illustrate the six contrastive tones in the first or the second syllable of disyllabic words when they are adjacent to each of the six individual tones. They display the phonetic variations in the pattern of the six tone contrast despite the fact that they are modified to a great extent in terms of height and slope under the influence of either the preceding or the following tones. In particular the main body of the tone (i.e., three quarters of the duration away from the adjacent tone), clearly display the relatively constant contrastive pattern.

2.1.3.2.3. The Impact of Segments in Tonal Contours

The following figures 2.16 and 2.17 illustrate the six individual tones in identical tonal sequences of the disyllabic words carried by different segments (taking the mean of the same three informants).

a. In the figures 2.16 and 2.17, the solid line represents a tone carried by a /ji:/ segmental string for the first syllable and by other sonorants for the second syllable; the dashed line represents a tone carried by a /ji:/ segmental string for the second syllable and by other sonorants for the first syllable. The difference in height or in slope of the pairs of the identical tonal sequences in the identical tonal environment but carried by different segments strongly suggests that the difference may be caused by different segments.

Generally speaking, the tone contour carried by /ji:/ is considerably higher in F_0 and has less fluctuation than that carried by other sonorants for the same tone. This is clear because /i:/ is the highest vowel and the formants of /j/ are very similar to those of /i:/. A difference in height or in the slope of the pairs of the tonal contour may reflect the difference of segments which carry it.

Fig. 2.16. Mean F_0 contours of T1, T3, T4 and T6 occurring twice consecutively. The vertical line inside the figure indicates the tonal boundary. The "T" in the legend represents a tone with segments other than /ji:/. The segments are specified as: the solid lines represent the following syllables: /**ji:**^{1} ji:^{1}/, /**ji:**^{3} ow^{3}/, /**ji:**^{4} na:m^{4}/ and /**ji:**^{6} mej^{6}/; the dashed lines represent the following syllables: /ɐw^{1} **ji:**^{1}/, /ji:m^{3} **ji:**^{3}/, /jɐn^{4} **ji:**^{4}/ and /ma:n^{6} **ji:**^{6}/.

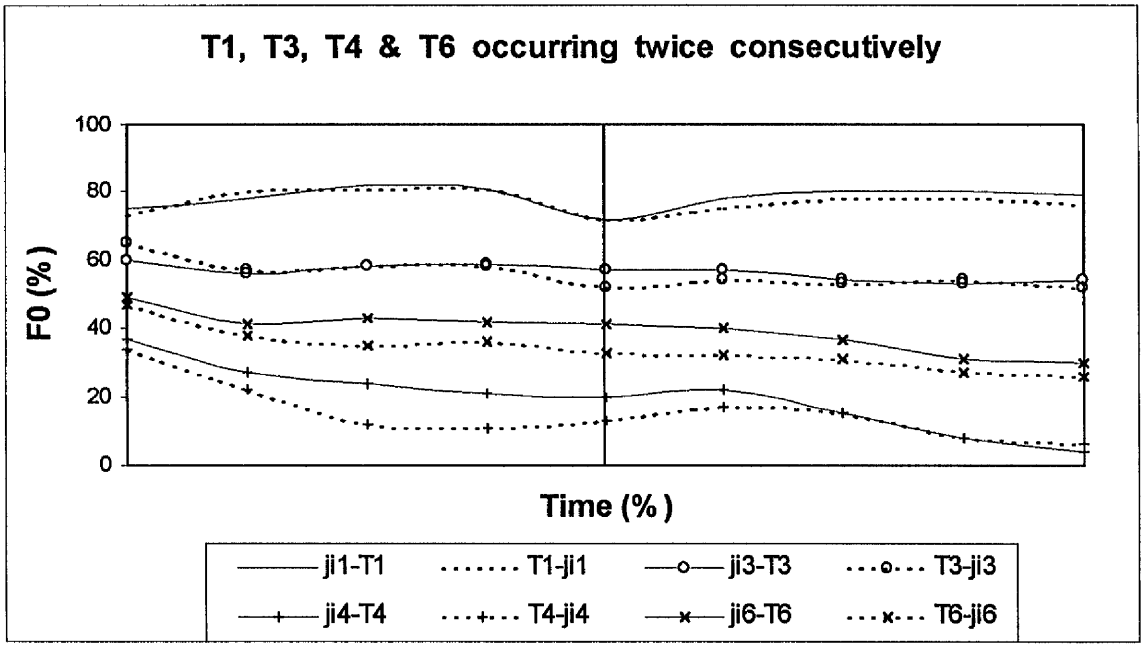
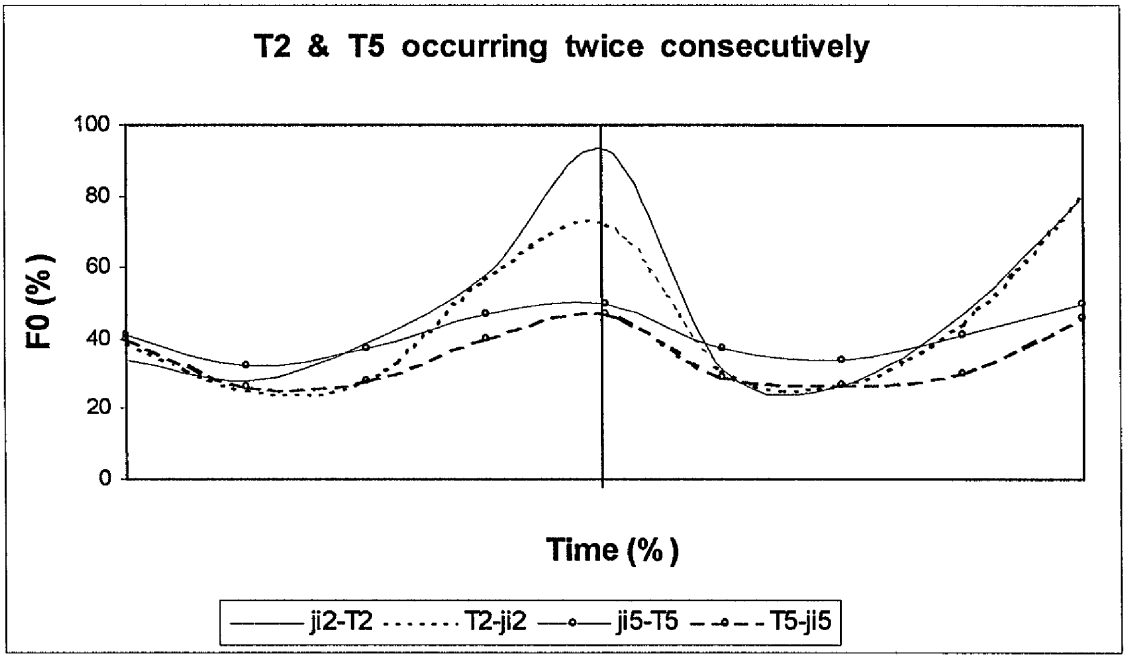


Fig. 2.17. Mean F_0 contours of T2 and T5 occurring twice consecutively. The vertical line inside the figure indicates the tonal boundary. The "T" in the legend represents a tone with segments other than /ji:/. The segments are specified as: the solid lines represent the syllables /**ji:**^{2} ŋɐj^{2}/ and /**ji:**^{5} jy:^{5}/; the dashed lines represent the syllables /ŋɐj^{2} **ji:**^{2}/ and /jɐw^{5} **ji:**^{5}/.



This enables us to explain why in T2 the two tone contours differ in height in the first syllable but almost overlap in the second syllable. The segments in T2 are /ji: ŋɐj / and /ŋɐj ji:/. /ɐ/ is a low central vowel whose F_0 is lower than that of the high vowel /i:/, and therefore displays a lower F_0 and a steeper falling tone contour than that of /i:/ in the first syllable. For the second syllable, the tone contour carrying /ɐ/ falls more than the one carrying /i:/, and the two tonal contours meet up for the rest of the duration. The initial nasal consonant /ŋ-/ appears to have influence on the F_0 of the following vowel /ɐ/, owing to most of its energy being in the low frequencies.

T4 exhibits a similar phenomenon to T2. The two tone contours of T4 are carried by /ji: na:n/ and /jɐn ji:/ and end up with /jɐn/ lower in F_0 than /ji:/ in the first syllable and almost overlap in the second syllable.

T5 is carried by /ji: jy:/ and /jɐw ji:/. The tone contour in the first syllable is similar to T2 - the tone carried by /ɐ/ has a lower F_0 and a steeper falling slope than that of /i:/. In the second syllable, the contour of /ji:/, beginning with a low onset is close to that of /jy:/ but they do not overlap because /y:/ is also a high vowel.

T6 is carried by /ji: mej/ and /ma:n ji:/. Its tonal contours have no overlapping or intersecting. The tone contour of /ji: mej/ consists of high vowels and occupies a higher F_0 position, whereas the tone contour of /ma:n ji:/, whose low vowel /a:/ keeps the tone contour in a low F_0 position, displays a deeper falling slope than that of /ji:/ in the first syllable. The tone contour of /ji:/ in the second syllable has less fall than its preceding tone contour or than the tone contour of /mej/ in the second syllable.

T3 is carried by /ji: ow/ and /ji:m ji:/. The tone contour of /ji: ow/ has slight fluctuation in the syllable of /ow/ whereas in the syllable of /ji:/ displays a fairly level line. Similarly, the tone contour of /ji:m ji:/ in the central portion of the second syllable keeps a fairly level line which

is like that in monosyllabic words, whereas the tone contour in the first syllable of /ji:m/ displays a gentle fall reflecting the influence of the nasal coda.

T1 shows a different picture in which two tone contours (carried by /ji: ji:/ and /ɐw ji:/) are very close in terms of F_0 height and slope but diverge slightly in the /ɐw/. These findings strongly suggest that if we want to compare two tonal contours, it is more reliable if the two tonal contours bear the same segments.

b. No attempt is made in this experiment to control the intonation. When two syllables carry an identical tone, the pitch pattern clearly displays a declining contour from the beginning of the first tone to the end of the second tone. The absolute F_0 value of the tone height in the first syllable is higher than that in the second syllable. This may reveal that the universal declination operates in the disyllabic utterance and interacts with the two-tone sequences.

The three level tones display a gently slanting shape, as does T4. T6 declines more than T3, T1 the least. As a result of the continuing downward slope of the two identical tones when occurring consecutively, the first syllable is higher in pitch than the second syllable. This can be observed casually in figure 2.16. The mean of aggregate F_0 values at the first syllable and the second syllable of the two T1s is 78 vs 77 of the two T3s is 58 vs 42, of the two T6s is 41vs 33, of the two T4s is 22 vs 13 (at the percentage scale of the F_0). (The aggregate F_0 values are taken from the five point measurements). As for the rising tones, the mean of the aggregate F_0 values at the first syllable is also higher than the second syllable: 77 vs 47 for the two T2s and 36 vs 35 for the two T5s (at the percentage scale of the F_0). (The aggregate F_0 values of the rising tones are taken from the last four points of the measurements on the grounds that the assimilatory effect on slope heavily affects the first quarter of the tonal duration of the second syllabie and this in turn raises the low F_0 onset of the rising tones.)

These results do not encourage a conclusion that the universal declination of intonation affects disyllabic words in reading lists more than in other environments, or that the universal terminal fall affects disyllabic words in reading lists more than in other environments. As experienced experimenters know, there is no perfect design for an experiment. Live conversation is the most real and natural speech but it is hard to control the material; with designed reading it is easier to control the material but it is not natural everyday speech; reading disyllabic words in a carrier sentence is said to be more natural than in an isolated situation; nonsense words can be used but sound artificial. I am not denying that intonation and tonal coarticulation interact with each other, but in tonal coarticulation experiments it is best to minimise such interaction. The present study is aimed at examining the tonal behaviour in two-tone sequences. A reading list of designed disyllabic words avoiding surrounding tonal influence is expected to be ideal as long as it is not affected by any strong intonational interference.

2.1.3.2.4. Timing Transition

2.1.3.2.4.1. Residual Energy vs. Energy Run-out

For the illustration of the six tones when following T1 and T2 see figures 2.4 and 2.5.

It is reported that in Mandarin the overall tonal values of following tones are higher when after tone 2 than after tone 1 even though they both have high offsets, because the F_0 direction in tone 1 is downward, while that in tone 2 is upward. The overall tonal values are pushed higher by a preceding tone with a rising F_0 than by a preceding tone with a falling Fo, although the F_0 heights of the two tones are at the same tonal register (Shen 1990: 292). A similar phenomenon is reported in Vietnamese - that every individual tone when following a rising tone is higher by approximately 4-6 semitones than its corresponding tone in a monosyllabic utterance. This "seems to be due mostly to the higher onsets after a rising tone caused by the extremely high end of the preceding rising tone" (Han and Kim 1974: 230).

However, Cantonese seems to reveal a different picture of behaviour of the overall tonal pitch following a high offset tone. As discussed above, the onset of a Cantonese tone can be

raised by a high offset tone or lowered by a low offset tone and lasts for about the first 25% of its duration before heading up or heading down to its own height and own shape. The two sets of data representing the pitch range of the main body of the contour of the six tones, i.e., when preceded by a high offset tone and by a low offset tone, are very close, because the elasticity of every individual tone is very small. This may be due to the fact that the Cantonese tones include three level tones. The three are very much alike in terms of shape and only differ in height, which does not leave much space for them to shift up and down dramatically.

Regarding the overall pitch height of the two sets of the following six tones, they are very similar when preceded by the high offset tones, T1 and T2. The only difference between these two sets is, one tends to stretch out a little for the first half of its duration when preceded by T1, whereas in the other one, tones tend to cling together for the first quarter of its duration when preceded by T2. The overall pitch values of the onset of the six tones when preceded by T2 is higher than the other set which are preceded by T1. However, the former set intersects with the latter set at the central portion of the tonal duration: the former set hangs down while the latter set stretches out. The slight stretching out in pitch range of the six tones when preceded by T1 may be due to the carry over of some residual strength from the preceding high level T1. The crowded clinging together in pitch range of the six tones when preceded by T2 may be due to no energy remaining from the preceding high rising tone T2 and this results in a `hanging down' shape. When the low rising T5 is the preceding tone compared with T3 which has a similar offset value, the hanging down shape of the following tones is not consistent among the informants. Only the high rising tone runs out its energy in the journey of steeply climbing up in order to maintain its identity and leaves almost nothing to the following tones. Moreover, it cannot resist being affected by the following tones: the tonal peak is pushed up to an earlier point before the tonal boundary.

On the other hand, a delayed peak can occur, for example, in a T1-T4 sequence. The high F_0 contour of T1 spreads over to the beginning portion of the following T4, where the F_0 reaches a peak after the offset of T1 and right in the middle of the initial consonant of T4

(see figure 2.18). Also, in a T4-T1 sequence, a creaky voice in a low F_0 T4 can spread to the initial consonant of the following high T1, or, in other words, a T1 carries over the laryngeal feature of the preceding T4 (see figure 2.19). When we talk about `spread over', we refer to a feature on the left segment or tone spreading onto the segment or tone on the right. The occurrence of a delayed peak may be due to the residual strength from the preceding T1, which is strong enough to spread over its high F_0 contour to the following low tone until its energy is finished - a delayed peak occurs.

Fig. 2.18. Delayed peak of T1. Extracted F_0 contour (Fx) of the disyllabic word /ji:^{1} la:n^4/ (lie on railing). The vertical line inside the figure indicates the tonal boundary. The highest F_0 value of T1 is 222Hz located at the offset. The highest F_0 value of the disyllable is 272Hz located after onset of the T4 (indicated by the arrow).

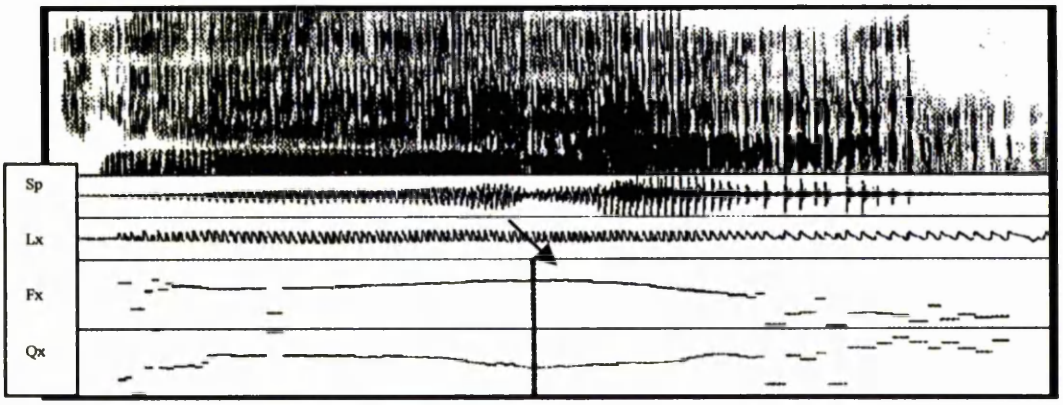
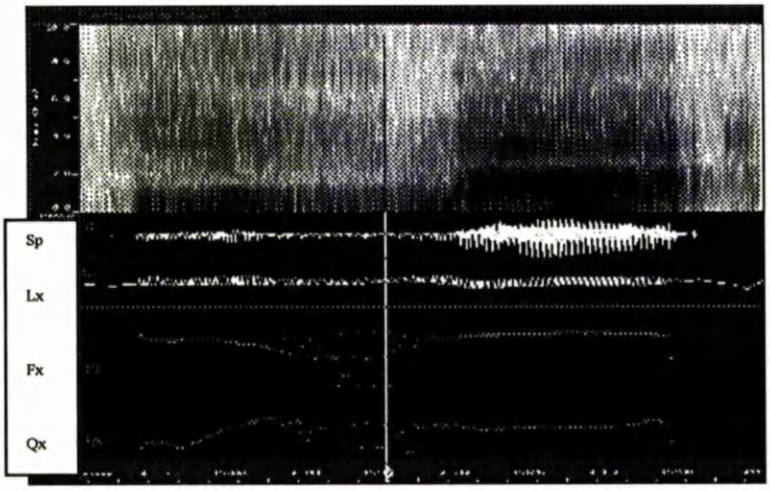


Fig 2.19. The creaky feature spreading onto the following tone. Extracted F_0 contour (Fx) of the disyllabic word /ji:^{4} ma:^{1}/ (aunt), spoken by JCCF. The F_0 contour shows that the aperiodic cycles (indicated by the arrows) reflecting the mode of vibration of the vocal folds in the low falling T4 are across the syllable boundary and spread to the beginning of the following T1.



2.1.3.2.4.2. Earlier Peak Before a Tonal Boundary

The peak of a preceding high rising T2 is not necessarily at the far end of the offset phonetically, but can be located about 10-15% of the duration backward from the offset. We call this an earlier peak (see figures 2.20 and 2.21). This is different from when T2 is in canonical form where the highest F_0 is at the offset. In the case of the earlier peak, the F_0 of the actual peak is in between both of the last two measuring points (i.e., 75% and 100% of the duration). Similar cases can be found in low rising T5. The phenomenon of earlier tonal peak in T2 and T5 is particularly evident when they precede T2, T3, T4, T5 and T6. Also, in some cases of non-rising tones, a small peak can be created about 10-15% of the duration backward from the offset.

In discussing the earlier peak placement in intonation, Bruce invokes a theory termed Gestural Overlap to explain some cases in Swedish and an alternative theory termed Tonal Repulsion to account for some cases in English. Bruce explains gestural overlap by saying that "an upcoming fall seems to overlap with the preceding rise giving an earlier peak location", whereas tonal repulsion: "implies a true temporal reorganisation of the accent gesture(s), which may result in anticipation of the first gesture and also delay of the second gesture." (Bruce 1990: 112) In the absence of understanding how the laryngeal gestures interact during tonal coarticulation, the present study does not attempt to test whether the earlier timing of a tonal peak happens because of gestural overlap, or because of tonal repulsion. But if we look at cases in non-rising tones where a peak is created, I doubt that an overlap model can explain how a small peak can be created in a fairly slanting slope.

The phenomenon of earlier tonal peak is also reported for Thai by Potisuk *et al* (1997). They see the earlier peak as a correlate of the dissimilatory nature of anticipatory effects. "[T]he lower the onset of the following tone, the earlier the point at which the peak of the rising F_0 contour, and conversely, the higher the onset of the following tone, the later the peak of the rising F_0 contour will be reached in the preceding tone." "[T]he peak F_0 of tones with rising

F_0 trajectives is realised at an earlier point within the syllable in order to maintain tonal identity as well as to allow for a less abrupt change in acceleration that is required in making transition to a following tone with a low onset." (Potisuk *et al.* 1997: 34) This analysis seems to conflict with other findings on the dissimilatory nature of anticipatory effects on height in Thai tone, reported in the same paper, when it says that only high and rising tones are affected. If a preceding rising tone adjusted itself "for a less abrupt change in acceleration", it would not be allowed to be raised "greater in height when followed by a low or rising tone". Instead, it would be lowered in height in order to meet up with the low onset tones.

The earlier peak in T2 and T5 or the created peak in non-rising tones in the preceding syllable, seems to be explained by the tonal repulsion model: an oncoming fall follows too closely, so pushes the tonal peak to an earlier point or pushes up the level to a crease - a created peak. As discussed above, the carryover effect is greater than the anticipatory effect in Cantonese. The strong counteracting force of the oncoming fall from the low onset of the following tones not only pushes up the height of the entire preceding tones, but also presses up the tonal peak to an earlier placement.

Fig. 2.20. Early peak of T2. Extracted F_0 contour (Fx) of /ji:n^{2} ji:^{6}/ (develop), spoken by JHDG. The dashed line indicates the location of the rising peak, the solid line indicates the offset of the T2. Both are labelled with arrows.

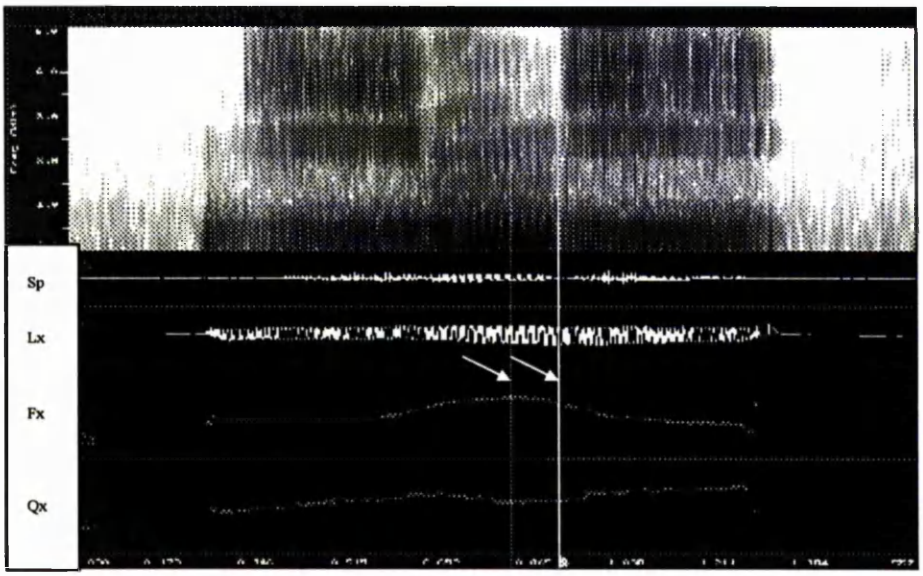
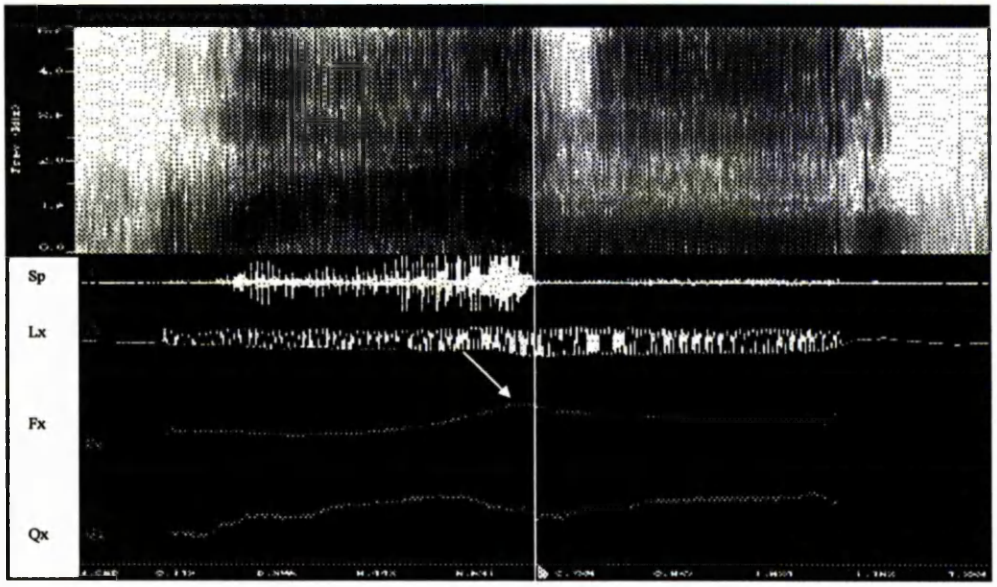


Fig. 2.21. Early peak of a T2. Extracted F_0 contour (Fx) containing the disyllabic word /wa:^{2} ji:^{3}/ (significance of a painting). The rising T2 reaches its peak (indicated by the arrow) 28ms before the syllable boundary (marked by the vertical line) and begins to fall at the syllable boundary.



2.1.3.2.5. Summary

In conclusion, the findings of this study support the view that tonal influence in two-tone sequences is bidirectional. Anticipatory and carryover effects on slope are both assimilatory in nature but on height are dissimilatory in nature. Both anticipatory and carryover effects on height encompass the whole of the duration of the target tones. An anticipatory effect on slope is minimal and extends for only the last quarter of the target tone's duration, whereas a carryover effect on slope extends for not more than the first half of the target tone's duration. The carryover effect is greater than the anticipatory effect. The contrastive pattern of the six individual tones remains despite the extensive phonetic variation in height and slope.

2.2. TONAL BEHAVIOUR IN COMPATIBLE & CONFLICTING TONAL ENVIRONMENTS (EXPERIMENT II)

2.2.1. Prediction

As shown in the result of the investigation of phonetic variation of the six individual tones when undergoing tonal coarticulation in two-tone sequences in the last section, the carryover effect is greater than the anticipatory effect and the modification of the slope is concentrated in the area of tonal transition. In respect of slope, in the domain of timing, there is a transition period which covers a quarter of the duration of the preceding tone and 25 percent of the duration of the following tone. A tone ends as an offset transition modified towards the following tone, and a tone begins with an onset transition which is building on the foundation of the preceding tone. In other words, the section of the tonal body not adjacent to another tone has found the right place to settle and is not much affected. One would wonder what the tonal contour would become if it has adjacent tones at both ends. If a tone is modified in both ends of its contour about a quarter of each end of its duration, can it maintain its contrastiveness? The present experiment attempts to examine tonal behaviour when the tone is in the middle position of three-tone sequences where the contexts are 'conflicting' and 'compatible'.

2.2.1.1. Defining Compatible and Conflicting Tonal Environments

In the case of three-tone sequences, I expect that the target tone in the middle syllable of the three tones has its onset modified by the preceding tone and its offset modified towards the following tone. I expect that the target tone to be modified more when it is in a conflicting tonal environment than in a compatible tonal environment, as discrepancies of the F_0 value between the tones coming in contact is greater in a conflicting context than in a compatible context. The terms 'conflicting' and 'compatible' tonal environment are adopted from Xu (1994: 2241): "A 'compatible' context is an environment in which adjacent phonetic units have identical or similar values along the phonetic dimension under scrutiny. A 'conflicting' context is an environment in which adjacent phonetic units have very different values along that phonetic dimension." I treat the Cantonese words, /$t^{h}i{:}n^{1}$ $ji{:}n^{4}$ $wa{:}n^{1}$/ (a natural bay), for example, as being in a conflicting tonal environment, where the tonal boundaries between

the first and the second syllable, and between the second and the third syllable are the boundaries between a high offset tone and a low onset tone, and between a low offset tone and a high onset tone, respectively. The F_0 values of those onsets and offsets are very different. In contrast, the words /tɕʰɐw¹ ji:n¹ joŋ¹/ (an old man smoking), for example, I treat as being in a compatible tonal environment, where the tonal boundaries between the first and the second syllable, and between the second and the third syllable are boundaries between onsets and offsets whose F_0 values are very close.

Fig. 2.22. Examples of tones in the middle syllable of compatible tonal environments.

COMPATIBLE	
_ / ‾	‾ ‾ ‾
/m⁴ ji:n² ŋa:m¹/ (not perform correctly)	/tɕʰɐw¹ ji:n¹ joŋ¹/ (an old man smoking)

Fig. 2.23. Examples of tones in the middle syllable in conflicting tonal environments.

CONFLICTING	
‾ / _	‾ ‾ ‾
/tɔ:¹ ji:n² jy:n⁴/ (lots of performers)	/fɔ:ŋ⁴ ji:n¹ mu:n⁴/ (smoke preventing door)

When a tone occurs in a compatible tonal environment, i.e., its onset shares a similar F_0 value to that of the offset of the preceding tone, and its offset to the onset of the following tone, I expect its canonical tonal shape to be able to stretch fully. Therefore, in the case of /tɕʰɐw¹ ji:n¹ joŋ¹/ (an old man smoking), I expect the tonal contour of /ji:n¹/ to be able to keep a fairly level shape at a high F_0 level. In contrast, when a tone occurs in a conflicting tonal environment, i.e., the F_0 value of its onset differs substantially from that of the offset of its

preceding tone and its offset from the onset of its following tone, I expect its canonical form to change at both ends. Therefore, in the case of /t^hi:n^1 ji:n^4 wa:n^1/ (a natural bay), I expect the tonal contour of /ji:n^4/ to become concave where both its low ends are pulled up by the high pitch of the adjacent tones; and in the case of /fɔ:ŋ^4 ji:n^1 mu:n^4/ (smoke preventing door), I expect the tonal contour of /ji:n^1/ to become convex where both its high ends are pulled down by the low pitch of the adjacent tones.

2.2.1.2. Coarticulation in Compatible and Conflicting Tonal Environments

As discussed before, we believe that the articulators never move from one static state to another, and never finish one sound completely before producing the next in connected speech. Instead, they create an area which is not completely characterised as the preceding sound nor the following sound, but is a transition which allows the articulators to move from one state to another. This transition, where the gestures may overlap in time, is called coarticulation. But in a string of tones in speech, the duration of one syllable can be very short and it may have a few vocal pulses only. The crucial question is whether, in such a short time, the articulators responsible for F_0 control - the vocal folds - can increase (or decrease) the rate of vibration from one tone to another without losing the phonetic characteristic of the tone which the speaker is supposed to convey - i.e., which the articulators actually intend to articulate?

After examining the rising tone and the falling tone in a conflicting and a compatible context, Xu (1994) concludes that there is a significant interaction between tone and context in Mandarin. "For the rising tone, the slope is positive in compatible context (sic.) but negative in conflicting context (sic.). For the falling tones, the slope is steeper in compatible context (sic.) than in conflicting context (sic.)" (p.2242). He also concludes that the preceding tone exerts more influence on the target tone than the following tone (p.2243). Shen (1990) reaches a different conclusion, namely that Mandarin tones maintain their contour shape despite being modified by their adjacent tones. One can regard these two observations as compatible with each other rather than conflicting if the different methods of investigation are

taken into account. Xu uses natural words, whereas Shen uses nonsense words with a special instruction to the subjects to stress the three syllables evenly. Potisuk *et al* (1997) admittedly also use three-tone sequences in their experiment in Thai but do not use the tone in the middle syllable as a target tone. Thus, no discussion is included concerning this topic. According to figure 6 (p.32) in their paper, the rising tone does not appear to have a falling contour when between a high and a low tone (i.e., in a conflicting context), neither does the falling tone appear to have a rising contour when between a low and a high tone (i.e., in a conflicting context). There seems to be no uniform answer between different experiments in tone languages for the above question - what the tonal behaviour will be when undergoing influence by the adjacent tones at both ends. Will the vocal folds respond to the tone in the middle syllable of a three-tone sequence as a transition from the left to the right as indicated in Xu's study? Or, no such special treatment as indicated in Shen's study? The present experiment attempts to find out how Cantonese tones will respond for the above question.

2.2.2. Procedures

2.2.2.1. Material

The selected target syllable is /ji:n/, which has special advantages in terms of: (a) voicing through the whole syllable in order to realise the tonal contour throughout the syllable from beginning to end; (b) like formants between /j/ and /i:/ avoiding fluctuation in tonal contours caused by different formants from different vowels; (c) being one of the few syllables which not only can carry six tones but also have the possibility of combinations of three-tone sequences as real words. Sonorants are chosen for the first and the third syllables at the syllable boundaries with the target syllable in order to ensure that voicing continues throughout the syllable boundaries.

Only the onset of T1 satisfies as the maximum high onset at the tonal boundary and only the offsets of T1 and T2 satisfy as the maximum high offset at the tonal boundary. T4 serves as a low onset and low offset at the tonal boundaries. (In order to pair up the low-high-low combination for the target T1, T5 is used as the substitute for T4 serving as the low tone onset. The offset of T6 is the penultimate lowest and the offset of T4 is the lowest, thus both

T4 and T6 are used for the first syllable in three-tone sequences.) Therefore, only T1, T2 and T4 are used as target words in the conflicting tonal environments. The combinations of three-tone sequences chosen in conflicting tonal environments are listed as below.

Fig. 2.24. Examples of T1, T2 and T4 in conflicting tonal environments.

Tone 1	Tone 2	Tone 4
T4-T1-T4 (_ ¯ _)	T2-T2-T4 (/ / _)	T1-T4-T1 (¯ _ ¯)
T4-T1-T5 (_ ¯ ,)	T1-T2-T4 (¯ / _)	T2-T4-T1 (/ _ ¯)

For their compatible counterparts, T1 and T4 can be put in identical-tone sequences of trisyllabic words, i.e., T1-T1-T1 (¯ ¯ ¯) and T4-T4-T4 (_ _ _). T2 is placed adjacent to the high onset T1 and the low offsets T4 and T6 (i.e., T4/T6-T2-T1, (_ / ¯)). Apart from these, I also investigate T2, T3, T5 and T6 in the middle syllable of the identical-tone trisyllabic words, i.e., T2-T2-T2, T3-T3-T3, T5-T5-T5 and T6-T6-T6 as a byproduct of the experiment in order to reveal the pattern of the six tone contrast in the middle syllable of three-tone sequences. The target tone which is placed in the middle syllable of a three-tone sequence is believed to be optimal for showing coarticulatory effects since tonal coarticulations appear to be restricted to contiguous tones only (Potisuk *et al* 1997 and Shen 1990). (The list of trisyllabic words used in the present experiment can be found in appendix II.)

2.2.2.1. Speakers and Recording

The trisyllabic words used in this investigation are real words and are isolated from other neighbouring words. The informants JCCF, JHDG and LTHJ used in the first experiment read the designed word lists, with three repetitions, in a sound proofed booth at SOAS. The recording was made with a laryngograph device as described in the previous section. All six tones in three identical tonal sequences were used in the investigation; and the three tones T1, T2 and T4 were used for the investigation of tones in the middle syllable of the three-tone sequences in compatible and conflicting contexts. The recording procedures and the methods for the determination of the time points of syllable boundaries, and for segmentation, are the same as in the first experiment. F_0 values are taken from five points

in the target tones: the onset point, the mid point, the offset point, and the points at a quarter and three quarters of the total duration. Each syllable is normalised for duration on a percentage scale and all raw F_0 data are normalised. By such normalisation, F_0 contours may be compared across speakers. F_0 contours were smoothed by curve-fitting for display purposes only.

2.2.3. Results and Discussion

2.2.3.1. T1, T2 and T4 in Conflicting and Compatible Tonal Environments

The following figures 2.25 - 2.29 illustrate T1, T2 and T4 in conflicting and compatible tonal environments. Figures 2.25 – 2.27 are arranged by tones and figures 2.28 – 2.29 are arranged by tonal contexts.

As expected, a tone is modified at both ends by its preceding and following adjacent tones. That confirms that anticipatory and carryover effects are both present in the tonal contour when a tone is in the middle syllable of a three-tone sequence.

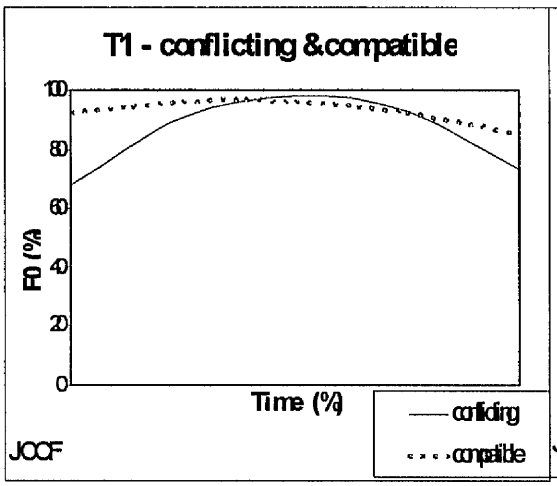
2.2.3.1.1. The Gravity

In terms of slope, T1 becomes a big arch, T4 a big fall-rise shape and T2 a right-oriented slantwise "∫ " shape in conflicting contexts; T1 keeps its canonical level shape, T4 becomes a fairly gentle curve and T2 a bending " / " line in compatible contexts. Recall that in two-tone sequences, the tonal boundary of T1-T1 always exhibits a dip, the tonal boundary of T4-T4 is always higher than the valley of the first T4 and the offset of the second T4. All three tones maintain their identity in around the central portion (i.e., in the area of the second and third quarters of the tonal duration) of their tonal contour (the central point area for T1, left of centre for T4 and right of centre for T2) while they allow a quarter of the duration from the onset and a quarter of the duration from the offset to alter their heights or directions in favour of assimilation from/to their adjacent preceding and following tones. The central portion of the contour where tones keep their identity, is the least affected by accommodation from/to the preceding or the following adjacent tones in terms of height or slope. I call it the gravity

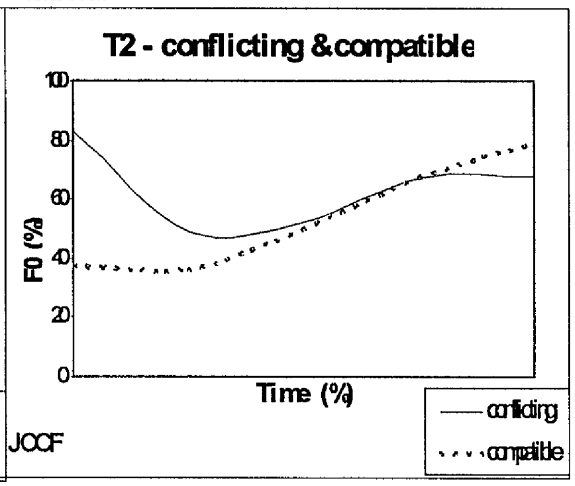
Fig. 2.25 (a-c). Mean F_0 contours of T1 in the middle syllable of 3-tone sequences produced by JCCF, JHDG and LTHJ, respectively. The solid line represents T1 in conflicting, and the dotted line in compatible, tonal context.

Fig. 2.26 (a-c). Mean F_0 contours of T2 in the middle syllable of 3-tone sequences produced by JCCF, JHDG and LTHJ, respectively. The solid line represents T2 in conflicting, and the dotted line in compatible, tonal context.

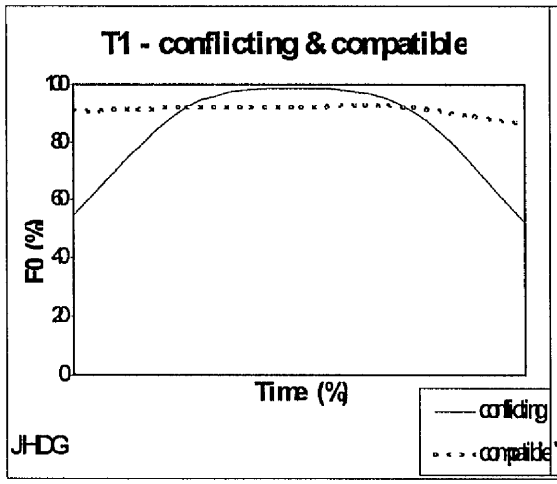
2.25.a.



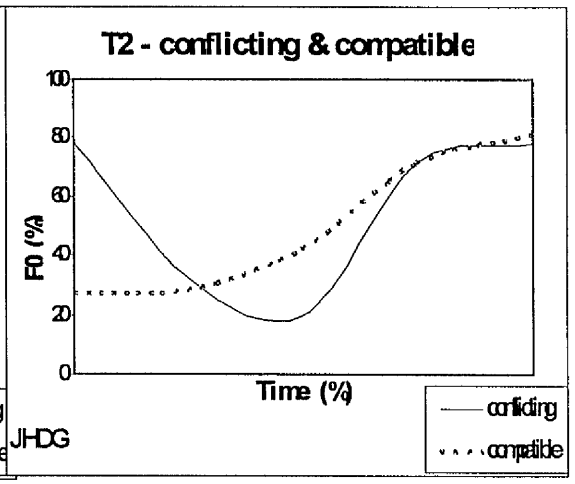
2.26.a.



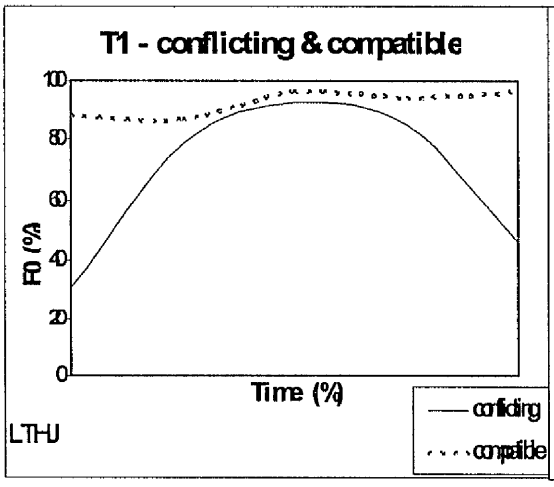
2.25.b.



2.26.b.



2.25.c.



2.26.c.

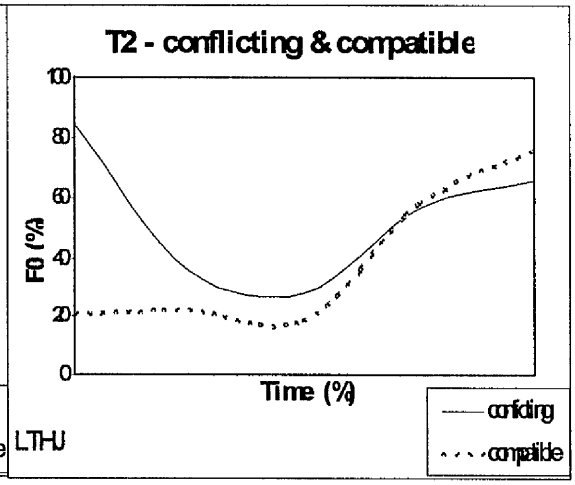
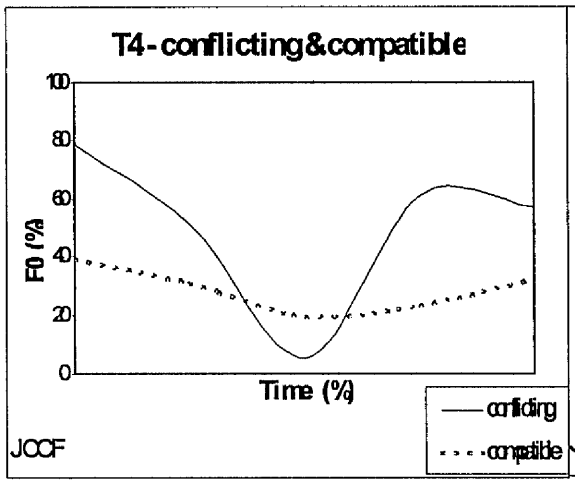
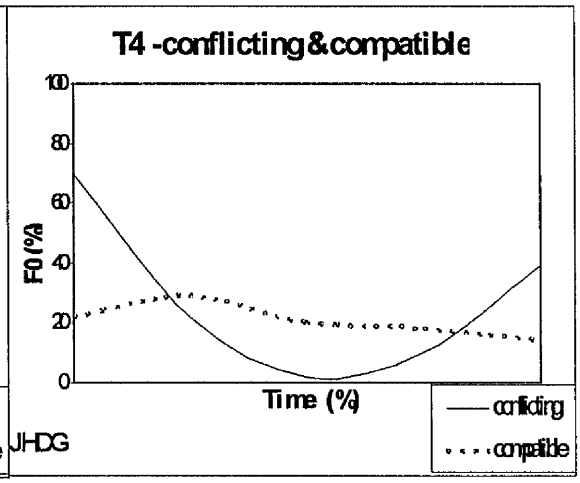


Fig. 2.27 (a-c). Mean F_0 contours of T4 in the middle syllable of 3-tone sequences produced by JCCF, JHDG and LTHJ, respectively. The solid line represents T4 in conflicting, and the dotted line in compatible, tonal context.

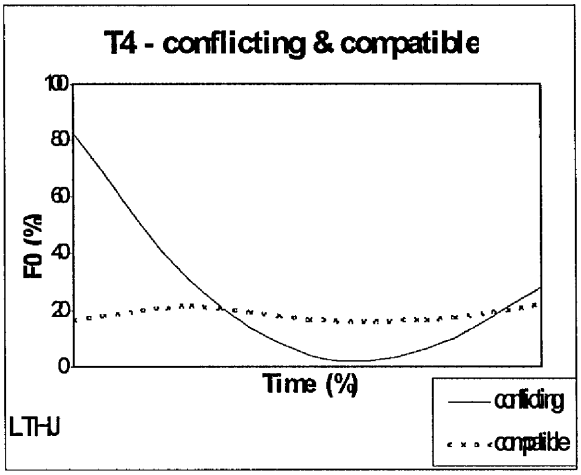
2.27.a.



2.27.b.



2.27.c.



of the tone. Every tone has its own gravity which has a strong power to resist modification from any direction and to maintain its characteristic in terms of height and slope.

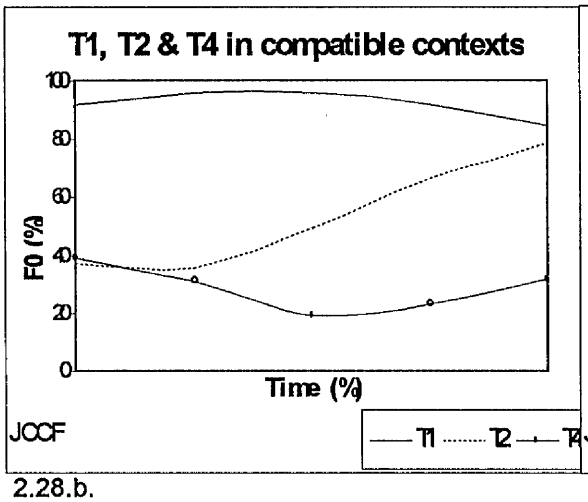
2.2.3.1.2. Tone 1

There are similarities in the central portions of the two contours of T1 when occurring in conflicting and in compatible contexts. First, T1 keeps its fairly level contour in the central portions both in conflicting and compatible contexts. Second, the height of the central portion of T1 in compatible contexts is very close to that in conflicting contexts. The two

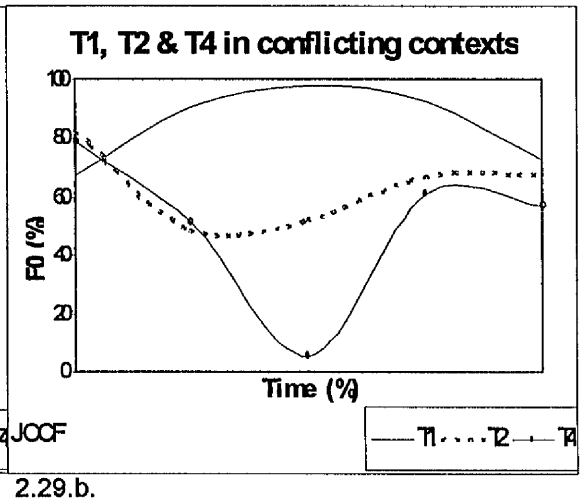
Fig. 2.28 (a-c). Mean F_0 contours of T1, T2 and T4 in the middle syllable of 3-tone sequences of compatible tonal contexts produced by JCCF, JHDG and LTHJ, respectively.

Fig. 2.29 (a-c). Mean F_0 contours of T1, T2 and T4 in the middle syllable of 3-tone sequences of conflicting tonal contexts produced by JCCF, JHDG and LTHJ, respectively.

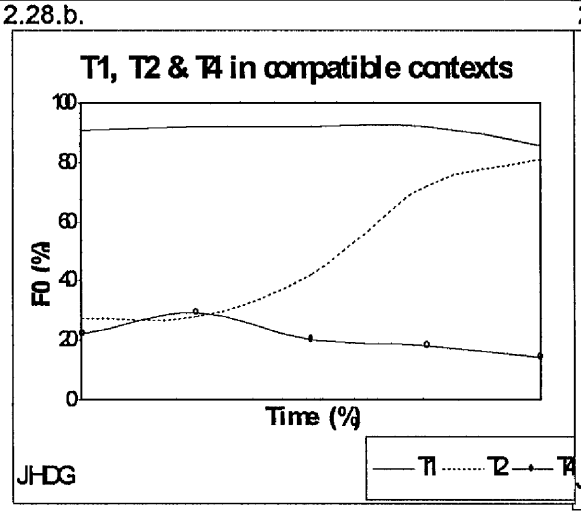
2.28.a.



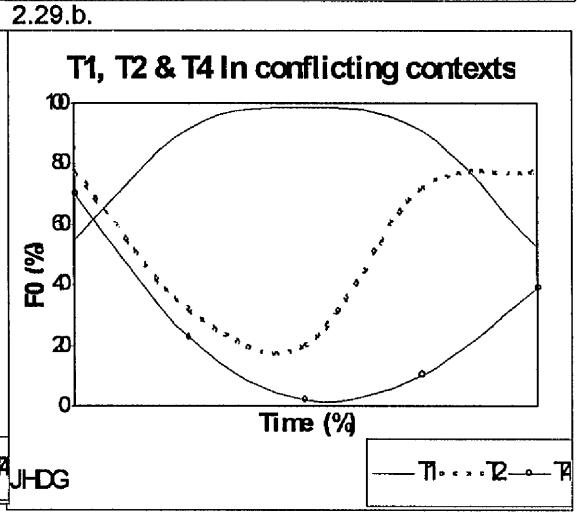
2.29.a.



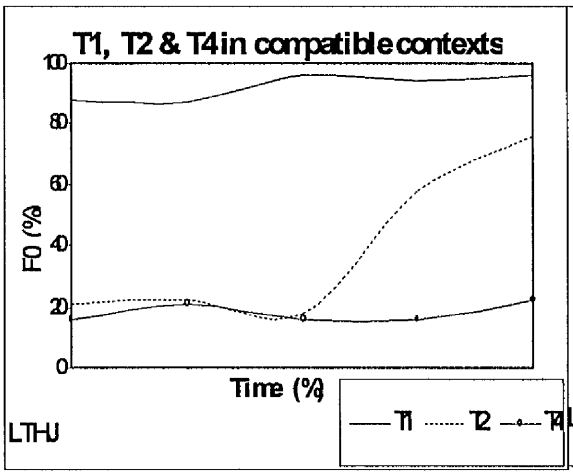
2.28.b.



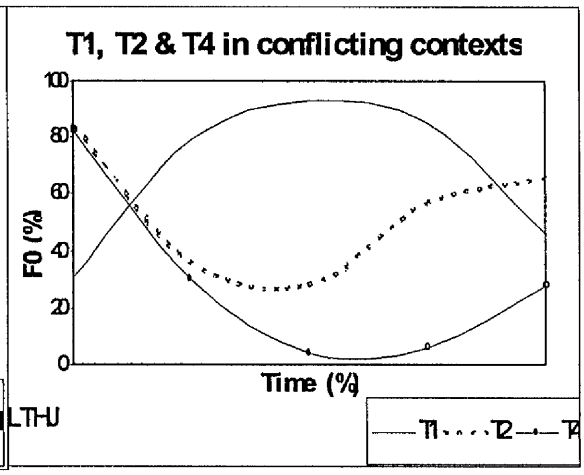
2.29.b.



2.28.c.



2.29.c.



contours almost overlap. This suggests that their tonal identifications are maintained in the central portion of their tonal bodies, regardless of the considerable modifications in the two quarters at each end.

In conflicting context, T1 bows down its head in order to connect to the low offset of the preceding tone and bends its tail in order to link up with the low onset of the following tone. Therefore, the slope appears as a convex as predicted. However, the degree of the 'bending' at both ends is not equal. The offset does not have to be lower than the onset, as one might have predicted in view of the declination effect. In fact, the offset is higher than the onset, i.e., the onset modifies more than the offset does. This may be due to the fact that the height of the offset of the preceding T4 is much lower than that of the onset of the following T4, as a result of coming at the end of a fall. Thus, the target T1 has to bow its onset more than its offset.

In compatible contexts, T1 appears as a fairly straight level line. Its offset is lower than its onset for informants JCCF and JHDG. This coincides with the tonal behaviour of T1 in the two-tone sequence - the end (i.e., either the head or the tail) of the tonal contour dips a little when it comes in contact with another preceding or following T1. Comparing the pitch range of the onsets and the offsets between the compatible and the conflicting contexts, the pitch range of the onsets is obviously wider than that of the offsets. The onsets are modified more than the offsets. The same is true of T2 and T4. This finding strongly supports the claim in the last section that the carryover effect is greater than the anticipatory effect in tonal coarticulation.

2.2.3.1.3. Tone 2

In both the tonal contexts (i.e., compatible and conflicting) T2 maintains its rising shape in the third quarter of the tonal duration. This part is called the gravity, as the tone's characteristics are concentrated here. The slopes from its onsets and extending to over a quarter of the duration, differ in direction between the two contexts. In conflicting contexts, the onset starts from high then heads down to the lowest point of the contour at the second

quarter of the duration. From the lowest point, the contour starts rising and more or less overlaps with the third quarter of the duration of the compatible-context contour. When it reaches the last quarter of the duration, the slope turns away from continuous rising to a fairly level shape in order to link up with the low onset of the following tone.

In contrast, in compatible contexts the first quarter of the duration of the contour has a fairly level shape because the onset is in contact with a low offset tone. The last quarter of the duration of the contour is a continuous gentle rising until the offset which is in contact with a high onset tone. Again, the preceding tone exerts greater influence on the target tone than the following tone does. The difference of F_0 value and slope is greater in the first quarter of the duration than in the last quarter of the duration in both conflicting and compatible contexts. The offsets of the two contours in the two contexts are close in F_0 value while the onsets of the contours are far apart. This indicates that the preservative power of the tonal nature is greater at the offset than at the onset.

In conflicting contexts, the F_0 value is higher at the onset. In fact, the peak F_0 after rising is below the F_0 value of the onset. This is also attributed to the preservative power of the tonal nature at the high offset of the preceding tone, which lifts up the onset of the target tone to a great extent. This is very different from the pictures shown in the last section in respect of tonal behaviour in two-tone sequences or in monosyllables. In those circumstances, its F_0 maximum always coincides with the rising peak and is always at (or very near) the offset whether it is the preceding tone or the following tone or is isolated, and whether it is adjacent with a high or a low tone.

There are two characteristics of T2 when occurring in a conflicting context. One is that the rising peak can occur much earlier (about 25% of the duration before the offset) when followed by a tone which has a falling movement at the beginning. Another is that the F_0 value of the onset can be higher than that of the rising peak. If we take the tonal environment into account, we would not be surprised at this new phenomenon. As I argued

in the last section, a tone is modified by its preceding tone and a tone is modified by its following tone in a two-tone sequence, and the carryover effect is greater than the anticipatory effect. When a T2 is in a conflicting environment, its low onset is pulled up by the preceding tone, however, it has to struggle to preserve its rising feature - rising from low to high. Therefore, it quickly falls down to a low position to prepare for a pre-rise. The level of this fall differs across the informants, but the level of the rise reached is uniform across informants. After the rise, the two contours (i.e., in conflicting and in compatible contexts) intersect at around the end of the third quarter of the tonal duration. The contour successfully keeps its rise at the third quarter of the tonal duration, but the rising peak is comparatively earlier due to the fact that its offset is pulled down by the oncoming fall from the following tone. It seems that the shape of a contour is triggered by the forces from both directions - from left to right and from right to left.

In discussing the intonation phenomenon in Hausa, Inkelas and Leben (1990: 22) point out that "a primary low tone triggers the lowering of the pitch register in which all following tones in the phrase are realised" and this is called downdrift. In other words, the mechanism is that the left triggers the right. In discussing the assimilatory process in intonation in Hausa, they point out that a low tone can raise the preceding high tone, or it can raise itself if no high tone follows. The former is called high raising, the latter low raising. In other words, these mechanisms work from the right to the left. In discussing the timing of prenuclear high accents in English, Silverman and Pierrehumbert (1990: 102) describe how a high accent peak can be shifted earlier by the following low accent. This also implies that it triggers from right to left. I argue in the last section that tonal assimilatory effects in Cantonese occur in both directions (though carryover effects are greater than anticipatory effects) and an earlier tonal peak is likely to be pushed up by the oncoming fall. That the F_0 maximum within a T2 is not at the rising peak but at the onset in a conflicting context, is simply because the contour of a tone is affected from both directions. The left force pulls up its onset and the right force pulls down its offset. The left force is stronger (in other words, anticipatory effect is weaker) as shown in its lifting the onset of the target T2 which is higher than the offset of the target T2 which is pulled down by the right force. Recall that the onset of a T2 is much

lower than its offset in its canonical form. If it did not meet persistence in maintaining its rising characteristic, it would be straightened and become a falling contour. When it meets persistence in rising, it is being pushed by the oncoming fall from the right force and anticipates by assimilation to the low onset of the following tone. As a result, the strength of the rise is too weak to extend to the offset and has to give way - the actual rising peak is realised much earlier before the offset.

2.2.3.1.4. Tone 4

For T4, the two contours (i.e., in conflicting and in compatible contexts) differ from each other to a great extent in respect of height and slope. However, in both their gravity keeps in the left centre and remains low. In compatible context, the contour falls slightly to the centre point before it rises gently towards the offset for informants JCCF and LTHJ. The tone cannot meet persistence in falling all the way through for the two informants because its onset is pulled down by the preceding low offset and it has to anticipate in linking up the following tone (e.g., a T4) whose onset is somewhat higher than the inherent offset of T4. For informant JHDG, the tonal contour falls continuously after the centre point.

In the case of a string of low falling tones, i.e., in compatible contexts, we can expect that the onsets are successively lower in F_0, reflecting of intonational declination or tonal coarticulation (the onset's absolute F_0 is dragged down by its preceding low offset tones). However, we cannot imagine that this could happen in a linear scale: for example, if a fall of a low falling tone is 80Hz starting at 200Hz, the end of the third low falling tone will end up below 0 Hz. Also, they do not seem to work like downdrift as in Hausa. They display a gentle concave contour in the central portion, therefore they leave lots of space in F_0 for the following low falling tones - they preserve their falling feature at the second quarter of the tonal body in spite of a low start before giving way to modification by the following tone. As a result of falling, the centre point of the target T4 is lower than the starting point of the second quarter of the duration - the inherent pitch pattern. The onset of the target T4 is lower than the onset of its preceding T4 as a result of triggering by the left force; the offset of the target T4 is lower than the offset of its preceding T4 because it falls from an already low start and is

followed by triggering by the right force. (The fall can stretch further if there is no right force triggered as in the last syllable position.) Apart from this, the gradual downward trajectory in a string of low falling T4 is attributed to the interaction of the universal declination phenomenon.

When T4 is in conflicting contexts, the contour falls steeply to the centre point or the third portion of the tonal body before rising towards the offset which is pulled up by the following high onset tone. The whole contour displays a big V shape or a hanging hammock with someone sitting right in the centre. The similarity of the two contours (i.e., in conflicting and in compatible contexts) is that their fall is bigger than their rise and their centre points are close in F_0 value.

Xu (1994) also concludes that in Mandarin the carryover effect is greater than the anticipatory effect. His conclusion is based on comparing the mean F_0 value of the target tones between the two tonal environments. He only takes the mid point of the contour as the F_0 value, therefore he finds that "the rising tone has a higher mean F_0 in conflicting context than in compatible context, the falling tone has lower mean F_0 in conflicting context than in compatible context" (p.2243). This is because "the preceding tone exerts more influence on the target tone than the following tone" (p.2244). Even though we have come to the same conclusion, the facts which the conclusion is based on seem very different. Our finding regarding the F_0 values at the central portion (centre area for T1, left of centre for T4 and right of centre for T2) of the two contours (i.e., in conflicting and in compatible contexts) of a tone being very close is in complete contrast with Xu's finding in Mandarin. (He suggests that a slope can be shifted up and down and can change its direction according to the contexts.)

2.2.3.1.5. Summary

The first and the last quarter of the tonal duration are modified towards the adjacent tones both in conflicting and in compatible tonal contexts. The modification appears heavier at the first quarter of the tonal duration and in conflicting tonal contexts. The F_0 values at the

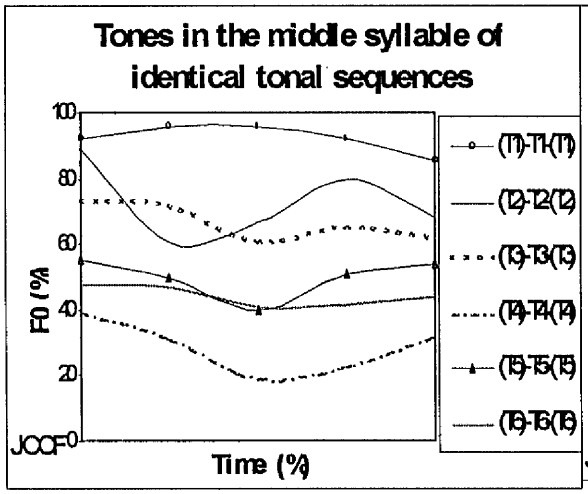
central portion (centre area for T1, centre left for T4 and centre right for T2) of the two contours (i.e., in conflicting and in compatible contexts) of a tone are very close. That is where the core of the gravity of the tone sits. It is where the tone defines its identity.

2.2.3.2. Tones in the Middle Syllable of Identical-Tone Trisyllabic Utterances

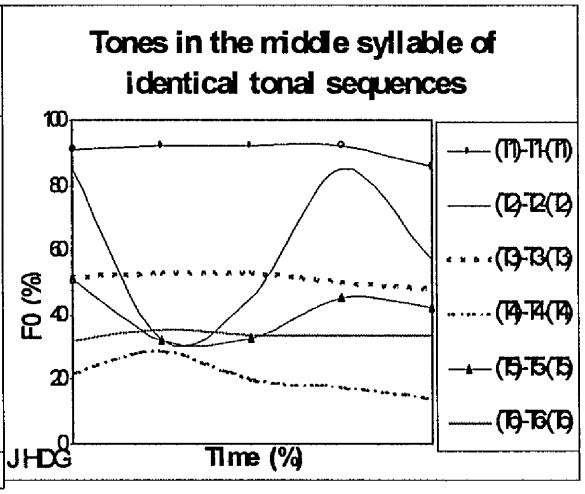
The following figures (2.30) illustrate the six tones in the middle syllable of the trisyllabic identical tonal sequences.

Fig. 2.30 (a-c) Mean F_0 contours of each of the six tones as they occur in the middle syllable of a sequence of three identical tones, produced by JCCF, JHDG and LTHJ, respectively.

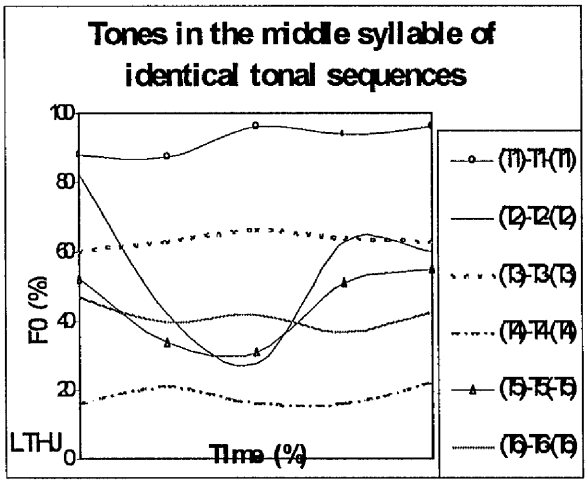
2.30.a.



2.30.b.



2.30.c.



Regarding the point about the characteristics of T1 and T4 being in the middle syllable of identical-tone trisyllabic utterances, see the discussion in 2.2.3.1.2 for the former and

2.2.3.1.4 for the latter on these two tones occurring in the middle syllable of compatible tonal environments; regarding the point about the characteristics of T2 being in the middle syllable of identical-tone trisyllabic utterances, see the discussion in 2.2.3.1.3 on this tone occurring in the middle syllable of conflicting tonal environments.

2.2.3.2.1. Tone 3 and Tone 6

The other two level tones, i.e., T3 and T6, can be taken as the tones in the middle syllable in compatible tonal environments when they are preceded and followed by the identical tones in three-tone sequences. Their horizontal level shapes decide their onsets and their offsets are similar in height. Thus, when they come in sequences of the identical tone, their onsets and offsets will come in contact with the offsets of the preceding tones and the onsets of the following tones whose heights are similar to their own onsets and offsets.

These two level tones preserve their canonical shape and the relative height. They maintain their level contours and keep contrast in height with one another and with another level tone T1. These are consistent with their tonal behaviours in two-tone sequences. In the two-tone sequences of T3-T3 and T6-T6, all tonal contours exhibit a sudden slight fall at the central portion (i.e., a quarter, half or three quarter of the tonal duration) while remaining downward as a whole. The gravity of T3 and T6 remains in the central portion of the contour. All level tones display a downward trajectory which is consistent with the canonical form and the forms in two-tone sequences.

2.2.3.2.2. Tone 5

The gradient of T5 rises gradually (in comparison with the high rising T2 with sharp rising gradient).

T5, when it is preceded and followed by identical tones in three-tone sequences is considered to be in conflicting tonal environments. When it comes in sequences of identical tones, its onset will conflict with the high offset of the preceding tone and its offset will conflict with the low onset of the following tone.

T5 maintains its small fall and gentle rise shape along with its particular height position. In terms of height, its rising contour in the last half of the duration remains somewhere between T4 and T3, which is consistent with its canonical form and its forms in two-tone sequences. Its fall extends from the onset to the midpoint of the tonal body before rising for informants JCCF and LTHJ. The steepest slope is the third quarter of the tonal duration, which is just the beginning of the rising part. The highest F_0 remains at or near the offset after rising. However, its second quarter of the tonal body changes from rising in the canonical form to continuously gently falling in the present experiment. Recall that the second quarter of the tonal duration does not always remain rising when it is in two-tone sequences. Therefore, only the third quarter of the contour is where the gravity of tone falls. The gravity shows the identification of the tone - a gentle rise after a fall, where the beginning of the midpoint of the contour is very close to the midpoint of T6 in term of height and remains rising during the third quarter of the duration.

There are some similarities between T5 and T6 both in overall height and contour in the data of informant JCCF. What makes T5 differ from T6 is the relatively steeply rising feature of T5 which makes T5 stand out from T6 particularly at the third quarter of the duration. Moreover, the pre-rise fall in the first half duration of T5 also distinguishes itself from T6 which is comparatively level, in the data of the other informants.

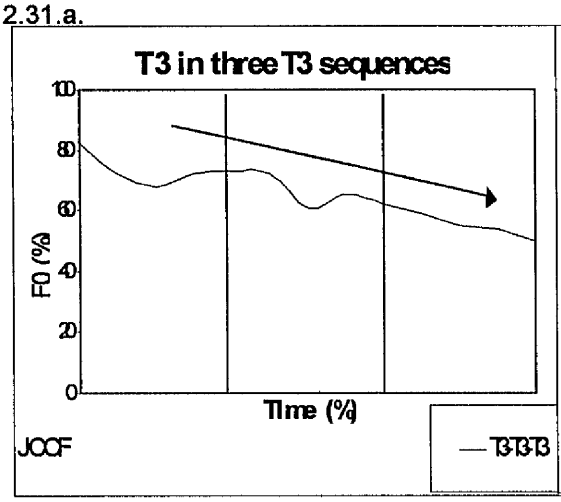
2.2.3.2.3. Declination

Across the informants, the rise in F_0 of the target T2 is much less compared with that of its preceding T2 and that of its following T2. The tonal peak of the target tone is lower than that of its preceding T2 (see figures 2.33). It is not surprising that the tonal peak of the last tone is higher than that of the target tone for informants JCCF and LTHJ, because it has more space to stretch itself as there is no force from the right to influence it. Even so, it is lower than the tonal peak of the first tone. The trend of the tonal contour, successively lower in pitch in trisyllabic identical-tone sequences, also appears in T3, T4 and T5 (see figures 2.31,

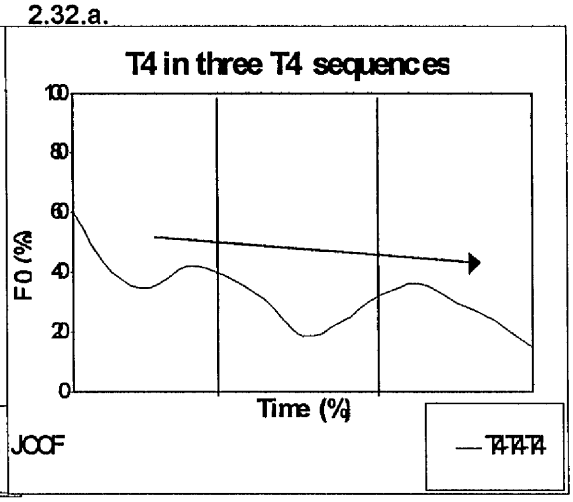
Fig. 2.31 (a-c). Mean F_0 contours of T3 occurring 3 times consecutively, produced by JCCF, JHDG and LTHJ, respectively. Two vertical lines inside the figure indicate the tonal boundaries & the arrow indicates the declining F_0 contour.

Fig. 2.32 (a-c). Mean F_0 contours of T4 occurring 3 times consecutively, produced by JCCF, JHDG and LTHJ, respectively. Two vertical lines inside the figure indicate the tonal boundaries & the arrow indicates the declining F_0 contour.

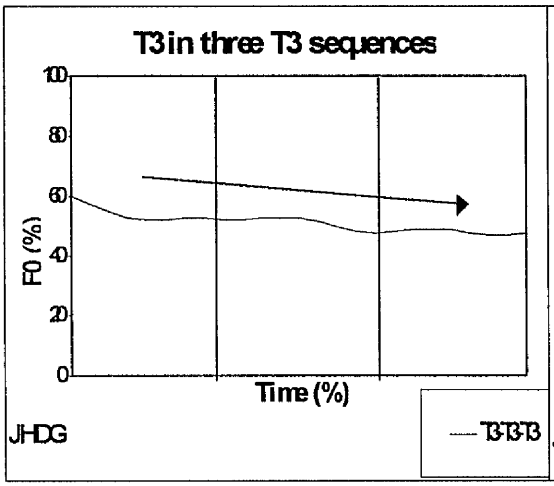
2.31.a.



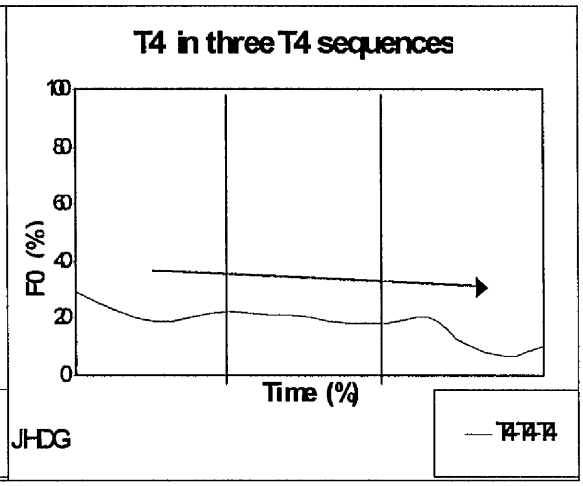
2.32.a.



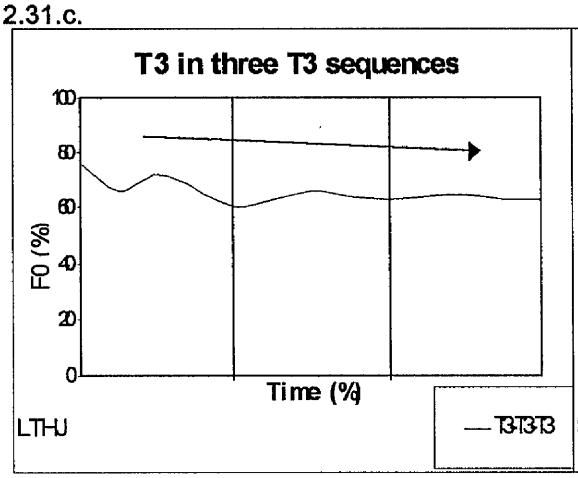
2.31.b.



2.32.b.



2.31.c.



2.32.c.

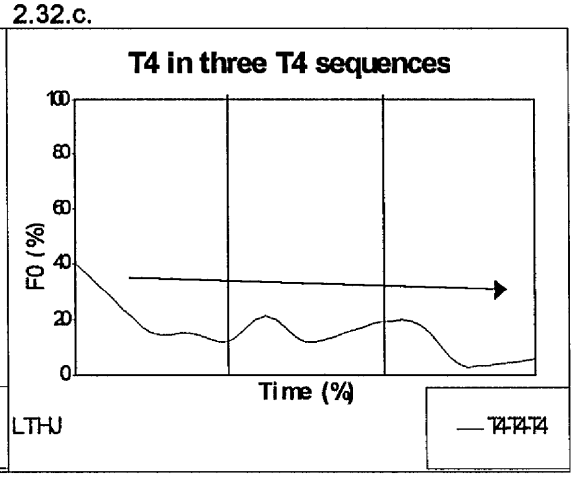
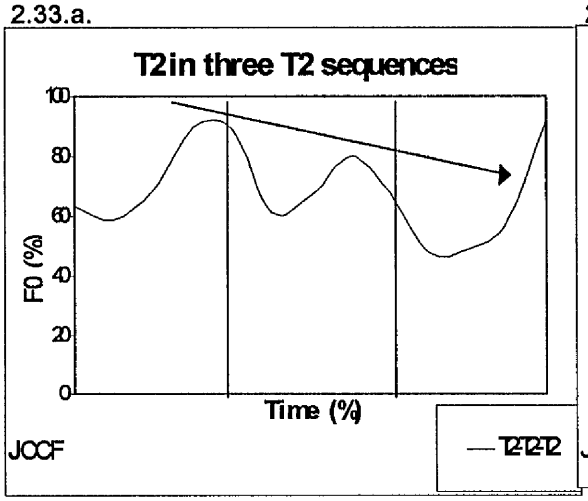


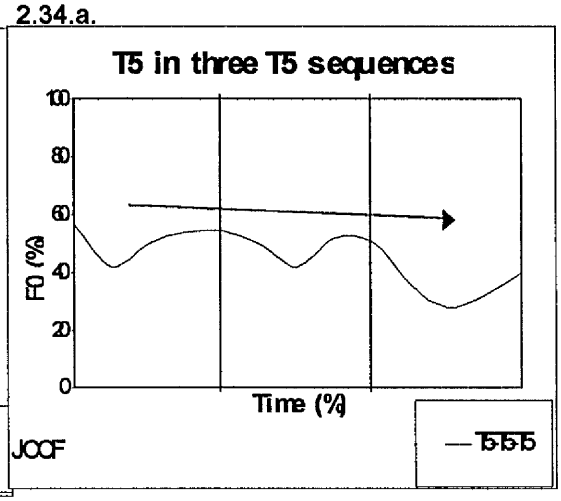
Fig. 2.33 (a-c). Mean F_0 contours of T2 occurring 3 times consecutively, produced by JCCF, JHDG and LTHJ, respectively. Two vertical lines inside the figure indicate the tonal boundaries & the arrow indicates the declining F_0 contour.

Fig. 2.34 (a-c). Mean F_0 contours of T5 occurring 3 times consecutively, produced by JCCF, JHDG and LTHJ, respectively. Two vertical lines inside the figure indicate the tonal boundaries & the arrow indicates the declining F_0 contour.

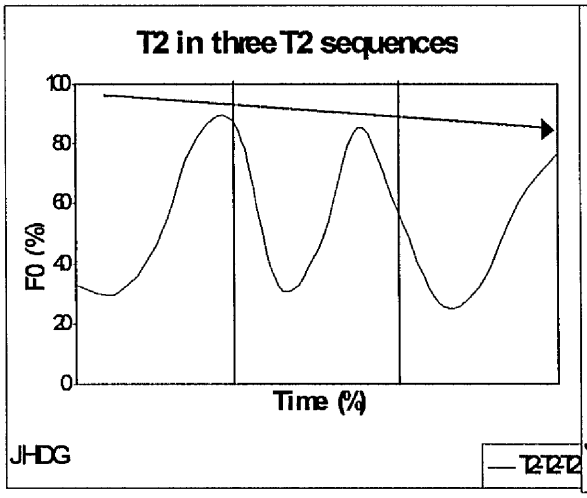
2.33.a.



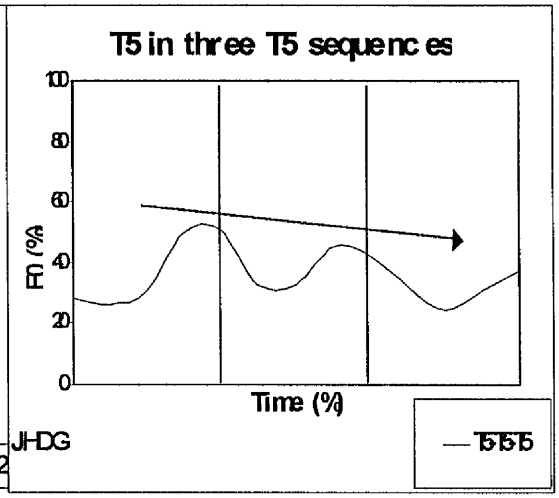
2.34.a.



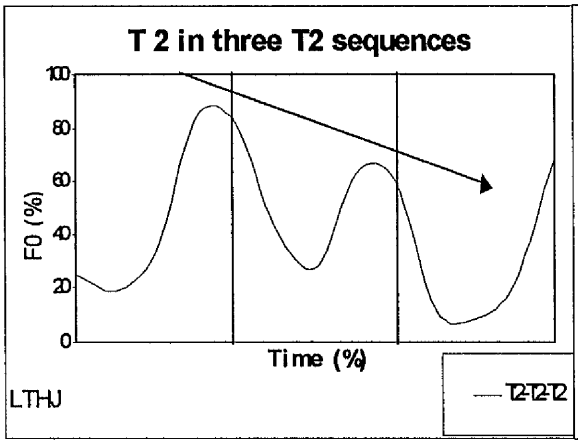
2.33.b.



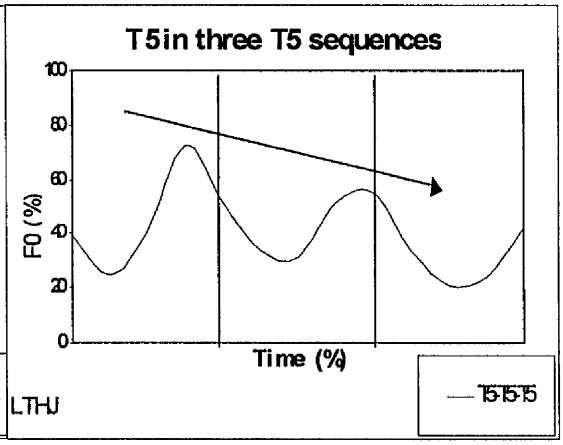
2.34.b.



2.33.c.



2.34.c.



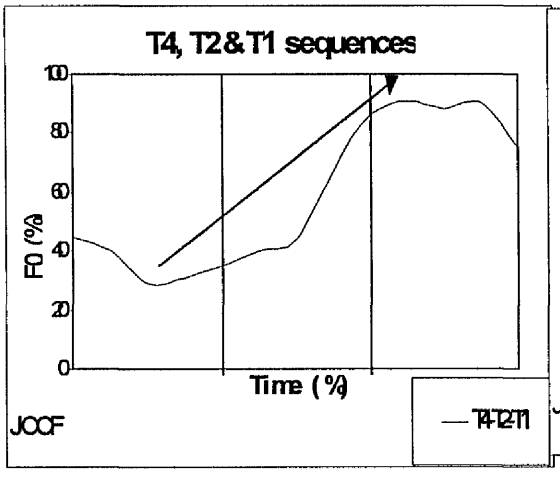
2.32 and 2.34). This may be attributed to the universal declination of intonation. Again, no intonation control attempt is made in the present experiment. The fact that a declining trajectory of the level tones, the falling tone and the rising tones occurs in all monosyllabic, disyllabic and trisyllabic words may indicate that the universal declination of intonation occurs in short utterances generally, whatever their length.

2.2.3.2.4. Sequences of T4-T2-T1

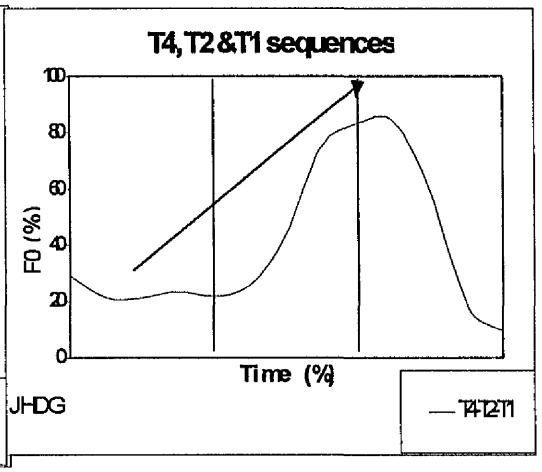
One would wonder what the contour will be if the tonal sequences are low-start but high-end, for instance, the tonal sequences of T4-T2-T1. Will the low T4 trigger the lowering of the following tones as with the downdrift or downstep in many African languages? As a by-product in the present experiment, the following figures give evidence that the low-to-high tonal sequences in Cantonese act differently. T1 does not lose its high pitch position, nor does T4 lose its low pitch position. The rising peak of T2 is at offset and the maximum F_0 value of the three tone sequences locates in the middle portion of T1. (Alternatively, it is located at the first quarter of the duration of T1 for informant JHDG as he exhibits a fall variant for T1 here.)

Fig. 2.35 (a-c). Mean F_0 contours of T4, T2 and T1 occurring consecutively, produced by JCCF, JHDG and LTHJ, respectively. Two vertical lines inside the figure indicate the tonal boundaries and the arrow indicates the rising F_0 contour.

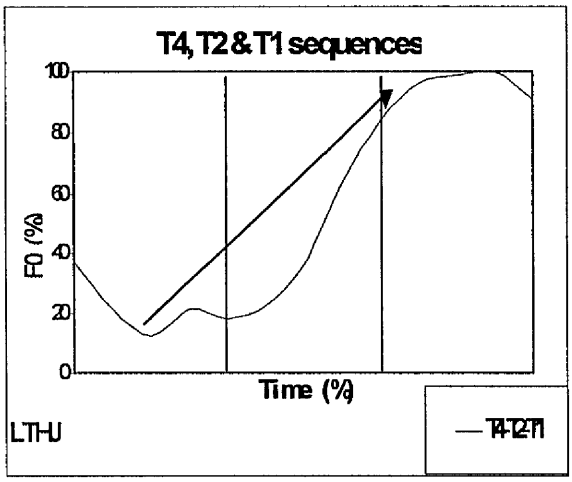
2.35.a.



2.35.b.



2.35.c.



2.2.3.2.5. Six Tone Contrast

Recall that in discussing the previous work done on Cantonese tones, I summarised the identification results of a perceptual experiment conducted by Vance (1977). Vance synthesised 64 different pitch patterns by allowing each of the 8 frequencies, as onset frequency, to combine with each of the 8 frequencies, as offset frequency using the syllable /jɐw/ to identify tones. Here I reconstruct his identification results in which the value of each case is taken as the value indicated by over fifty percent of his subjects, to draw a figure of the six tonal contours. The subjects, as native speakers, tended to tolerate variations of pitch for each tone. As a result, some tones are identified in a wider pitch range. For example, T6 is associated with between 92 to 206Hz for the onset F_0 and between 92 to 154Hz for the offset F_0. In these cases, I take the middle F_0 value 149Hz for the onset and the middle F_0 value 123Hz for the offset. First of all, for convenience, I record the identification result chart which I summarised in chapter one. Then I reconstruct the tonal contours based on his response data and the method mentioned above.

Fig. 2.36. Summarised response data from Vance (1977). Stimuli on which there was over fifty percent agreement (except T4 which was identified only nine times out of twenty). (cf. Cheung 1986)

Onset F_0	Offset F_0	92	106	119	133	154	178	206	245
	Tone identified as								
245									
206									
178									
154									
133									
119									
106									
92									

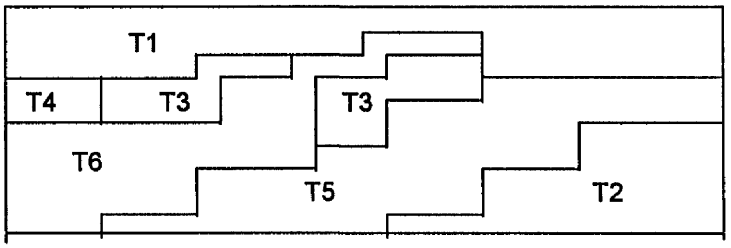
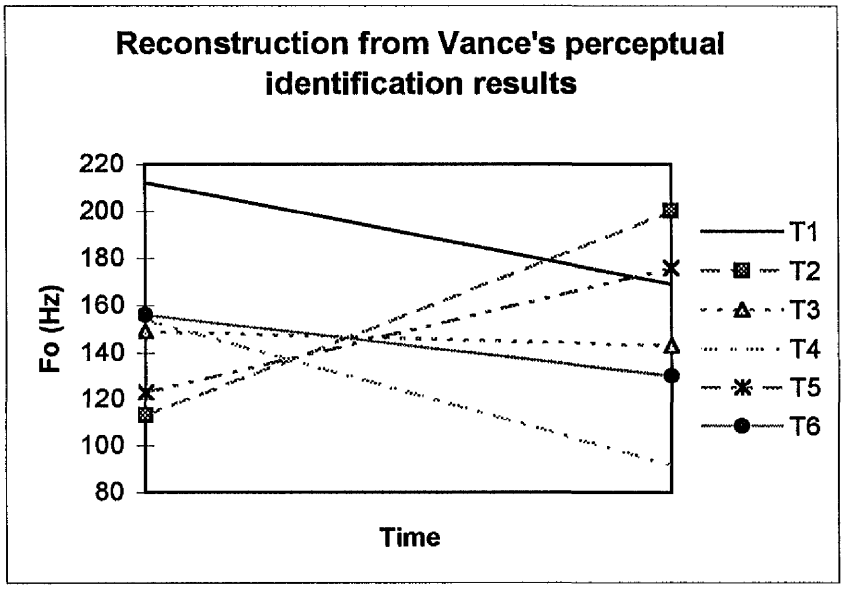


Fig. 2.37. Reconstruction from Vance's (1977) perceptual identification results. (See discussion above.)



The response results in Vance's experiment shows that the tonal onsets tend to be crowded in the low F_0 range and the tonal offsets tend to spread out. It seems that the onset F_0 value is less important than the offset F_0 value in the identification of Cantonese tones. Cheung, in discussing Vance's results, sees that the offset pitch is not the most reliable clue but the onset pitch is even less relevant (1986: 186). The result in the present experiment coincides

with Cheung's point of view drawn from Vance's perceptual experiment. The present experiment further claims that the most important element in representing the identity of tones is the gravity of tones. The gravity is never located at the beginning or the end portions of the tonal body but in the central portion. The beginning and the end portions of the tonal body need to make considerable adjustments to accommodate to the adjacent tones, and this can result in a change in their height and direction. This leads listeners to a great extent to tolerate a wide pitch range at the beginning and the ending of a tone. The reconstructed figure shows clearly that the height hierarchy and the direction of the slope coincide with those in the monosyllables shown in my last experiment. Moreover, if my hypothesis about less relevance in the beginning and ending portions turns out to be right, we will find that the height hierarchy and the direction in the central portion of the tonal contours in the reconstructed figure not only coincide with those in the monosyllables shown in my last experiment but also coincide with those in the gravity of the middle syllable in the three-tone sequences shown in the present experiment. In the central portion of the tonal contours in the reconstructed figure, in the experiments on monosyllables, disyllables and the middle syllable of trisyllables, there are lots of similarities in terms of tonal height and slope, as discussed above. Cheung, as a phonologist, chooses to use a bundle of feature values to explain why T1 tends to be perceived in such a wide range - that is, T1 is [+high, +extreme, -rising] - combined with the lack of another tone which is [+high, +falling]. He says that because of the lack of any tone that is [-extreme, +falling], a slight fall is tolerable for T3 and T6. For T2 and T5, the most important difference is one of gradient (1986:186). We can safely say that the height and slope of tonal contour in the gravity are both important in terms of tone production and perception in Cantonese.

2.2.4. Summary

A tone maintains its identity in terms of height and slope in its gravity - which is located in the central portion of the tonal body - while allowing the two ends to be modified in terms of height and slope in order to accommodate to the preceding and following adjacent tones. The gravity of the three level tones is located at the centre point area; the gravity of the falling tone is located at the central left whereas the gravity of the two rising tones is located

at the central right. The gravity is the least affected in accommodating to its adjacent tones. When a tone is in a conflicting tonal environment, its two ends (about a quarter of the duration from each end) are open to modification in terms of height and slope while the characteristic gravity is preserved. When a tone is in a compatible tonal environment, its canonical form can be preserved to a greater extent, but does not necessarily remain exactly the same. This is due to the coarticulation occurring at both ends - the preceding and the following tones both affect the target tone. In other words, a tonal contour is triggered by forces from both left and right. The central portion of the tonal body in compatible contexts draws close to and may overlap with the gravity of the tonal contour in conflicting contexts. The carryover effect appears to be greater than the anticipatory effect. The modification in the first quarter of the duration is greater than in the last quarter of the duration. The declination effect interacts with tonal contour as shown in the downward trajectory of level tones, of the falling tones, and of the rising peaks of the rising tones when in sequences. These findings are consistent with the results obtained in the experiment on tonal behaviour in two-tone sequences.

The table below lists the location and the characteristics of the gravity of the six tones when in the middle position of a three-tone sequence and coincides with the characteristics of the six tones in monosyllabic and disyllabic words.

Table 2.1. The location and the characteristics of the gravity of the six tones.

	T1	T2	T3	T4	T5	T6
location	centre point area	central right	centre point area	central left	central right	centre point area
height & direction	high & level	high & rising	mid & level	low & falling	mid-low & rising	mid-low & level

CHAPTER THREE

LEXICAL PROMINENCE

3.1. PROMINENCE AND STRESS

3.1.1. Defining Prominence and Stress

Prominence is a perceptual quantity that may be increased by means of any sound attributes, such as higher pitch, greater loudness, longer duration; whereas stress is an articulatory gesture (Lehiste 1970: 119 and 1993: 238) in which a syllable is produced by means of more energetic movements, resulting in higher fundamental frequency, greater intensity or longer duration due to the greater physical effort.

Prominence and stress are closely related due to their sharing similar acoustic properties (F_0, intensity, duration, timbre), but they are not the same thing. Jones (1940) was probably the first one to point out that stress perception also involves knowledge of the language in which the utterance is spoken in addition to the phonetic cues present in the sound waves; whereas the prominence of a syllable is its general degree of distinctness, this being the combined effect of the timbre, length, stress and (if voiced) pitch of the syllabic sound.

There are some intrinsic factors affecting intensity and pitch that we must take into account when we talk about stress and prominence. First, the higher the vowel, the higher the F_0 (Lehiste and Peterson 1961: 419, Ohala 1978: 29). The difference in F_0 between high and low vowels in the same intonational context can be as much as 25Hz. Second, low vowels are associated with higher amplitude while high vowels are associated with lower amplitude. In Lehiste and Peterson (1959: 360), listeners identified the vowels that were produced with a greater amount of effort such as /i/ and /u/ as louder than vowels having greater intrinsic amplitude, but produced with normal effort, such as /ɑ / and /ɔ / in American English. Initial

and final consonants also affect the decibel range on vowel nuclei. The essential order from strong to weak is (a) semi-vowels and voiced plosives, (b) glides, nasals and voiced fricatives, (c) voiceless fricatives and (d) voiceless plosives (Lehiste and Peterson 1959: 363). Moreover, syllables with higher pitch are generally heard as louder. This is due to the coincidence of a harmonic with the frequency of the first formant, since most of the energy of the vowel is contained in the first formant (Lehiste 1993: 238). In other words, an increase of perceived loudness can be caused by an increase in subglottal pressure which can produce an increase in the rate of vibration of the vocal folds or can result in greater amplitude of sound waves. Increased perceived loudness can also be caused by increased tension of the vocal folds.

Unlike tone or vowel quality, prominence is not a property of individual syllables. Prominence becomes evident in the prosodic organisation of syllables in a string (lexical items or utterances). A perceived prominent syllable is a syllable in a word or a phrase that is under the speaker's intended focus in the utterance (according to Bolinger (1972: 35), who goes on to say that words in utterances can be focused to signal contrast or special informativeness. He also takes the view that focused words are marked by pitch accents in English). In this chapter, I shall describe my understanding of how prominence is achieved on the lexical level in Cantonese and shall leave the discussion of contextual variations in prominence to the next chapter under the heading of tonic prominence. First of all, I would like to review how prominence is understood in other languages from some recent studies.

3.1.1.1. Prominence in Stress-accent Languages

In stress-accent languages like English, a stressed-syllable produced by a speaker is usually perceived as a prominent syllable by the listener. The classic perceptual tests claim that higher pitch and longer duration are more important than higher intensity for stress judgments in English (Fry 1955, 1958, 1965). It is well known that reduced vowels never associate with stressed syllables; stressed syllables usually associate with higher amplitude, longer duration and turning pitch; sentence intonation is ultimately an overriding factor in determining the judgement of stress. Beckman and Edwards claim that phonetically, prominence at the level

of stress foot and intonation phrase is realised primarily by the pitch contour (1994: 12). At the level of stress foot, the prominence is defined phonologically by the distribution of qualitative features of vowel quality (p.9). At the higher level of prosodic hierarchy, on the other hand, any effect of the prominence on the syllable's degree of vocal-tract opening is ancillary to the associated intonation pattern (p.30).

There are no uniform patterns of correlation between acoustic properties and stress (through which prominence is achieved) among stress-accent languages. Finnish, for example, "carries fixed word-level stress on the first syllable of a word but has contrastive length oppositions in any syllable, resulting in possible word types with stressed short syllables and unstressed long syllables" (Lehiste 1993: 239). In Hungarian, short vowels have greater intrinsic intensities than the long vowels (except for the /a/-/a:/ pair), due to the short vowels being more lax and open and long vowels being tense and close, as reported by Fonagy (1966: 231) after investigating the intensities of vowels in disyllabic words. Fixeed stress may occur on the first syllable of a Hungarian word. However, Fonagy also gives examples of many words in Hungarian in which the unstressed second syllable has longer duration, greater intensity and higher F_0 than the first stressed syllable. In Yiddish, Weinreich writes, every morpheme which consists of more than one syllable may receive a stress on one of the syllables, but not on others - the place of stress is very firmly fixed, for instance, /'mame/ (mother) and /a'her/ (hither). The distinguishing value of stress within the morpheme is almost nil: there are practically no pairs of morphemes that are distinguished by nothing except the place of stress. On the other hand, the stress is not an obligatory loudness at all, rather a place where relative loudness occurs if the morpheme is to be emphasised in the text. For example, if no emphasis falls on any morphemes of /vu 'arbet di 'mame/ (where does mother work?), then /'ar/ will not be any louder than /bet/ nor /'ma/ than /me/ (Weinreich 1954: 2). According to Ingham (1974: 56-57), Khuzistan, an Arabic dialect spoken in Southern Iran, shows prominence patterns relatable to its syllabic structure as well as to the presence of certain grammatical elements regularly associated with that prominence pattern. For instance, /wi:k/ in /yra:' wi:k/ (he will show you) is prominent because it is quantitatively long and is final, and /tab/ in /ke'tabha/ (he wrote it) is prominent

because it is the only long syllable of the form. The prominence pattern of /'s̲ara:wi:k/ (what shall I show you?) is relatable to the presence of the element /s̲-/ (what) and forms involving /'s-/ regularly show prominence on the first syllable.

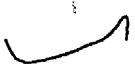
3.1.1.2. Prominence in Pitch-accent Languages

In Japanese, a pitch-accent language in which the accent is treated as a lexically linked H-tone, prominence is defined by Pierrehumbert and Beckman as a reflection of discourse focus indicated by using local expansions of pitch range. A major effect of focus is to raise the H-tone peak on the focused word. For instance, Pierrehumbert and Beckman (1988: 93, 98, 99) find in comparing *amai mame'* (an unfocused unaccented adjective followed by an unfocused accented noun), with *AMAI mame'* (a focused unaccented adjective followed by an unfocused accented noun) that the peak in the adjective when under focus was generally higher than in the sequences with no special focus on either the adjective or the noun.

3.1.1.3. Prominence and Stress in Tone Languages

In tone languages, tonal-pitch as a paradigmatic specification, is used as a property of syllables. Thus, stress and prominence are more complex in tone languages than in non-tone languages since one of their possible phonetic correlates - fundamental frequency - interweaves with the property of tone. More research is found on the topic of stress than prominence. Phoneticians are still exploring the physiological cause and perception of stress and prominence in tone languages. As in the category of stress-using languages, a prominent syllable in some tonal languages is likely to coincide with a stressed syllable (this is discussed in passing by Kratochvil (1998: 435) in relation to Mandarin and by Luksaneeyanawin (1983: 197) in relation to Thai). Relative duration is reported to be the perceptual cue in stressed syllables in Mandarin (Shen 1993: 415), and to be the predominant cue in signalling the distinction between stressed and unstressed syllables in Thai (Potisuk *et al* 1996: 200). Shen's conclusion is based on the comparison between duration and intensity only. In contrast, Coster and Kratochvil (1984: 124) in their Mandarin perceptual test find that fundamental frequency alone is a very strong discriminating dimension in respect of stress groups (yielding 93% correct rate). Amplitude appears as a

relatively good stress discriminating dimension (yielding 68% correct rate) but duration proves to be irrelevant (p.125). By testing 1390 syllables in four stress groups, they declare that the polydimensional complex signalling of stress is shared there by at least two other distinct prosodic features - tone and intonation (p.120).

Even though higher pitch is often heard as louder, increases in pitch and increases in amplitude do not necessarily go hand-in-hand in tone languages. Every individual tone has its own amplitude character, according to Kratochvil. It is well known that the tonal contour of Mandarin tone 3 is frequently identical to that of tone 2 as a result of tonal sandhi, but Kratochvil (1968: 39) finds that the tone 3 retains its original distinct amplitude contour - a falling tendency in loudness characteristic of the first two-thirds of the citation form  ; while the loudness of tone 2 rises gradually towards the end of the vowel  . There is a near perfect correlation between rates of change of amplitude and F_0 near peak F_0, according to Rose (1984: 164). He points out that both vocal fold vibration and subglottal pressure are involved in the production of F_0 features in Shanghai-Zhenhai. The F_0 of tone 3  and tone 4  has the same onset value, rises through the same range, and reaches the same peak value but with different timing for each tone. But in tone 3, this rise is produced by a burst of subglottal pressure whereas in tone 4 it is produced mainly by an increase in vocal fold tension (Rose 1984: 165).

3.1.1.4. Prominence and Stress in Cantonese

To the best of my knowledge, no study under the heading of prominence in Cantonese has been undertaken. Recall that I have mentioned in chapter one the contributions made under the heading of stress in Cantonese by Chao (1947), Barnett (1945-50), Hashimoto (1972) and Vance (1976) either by impressionistic observation or by experimental work; and the discussion by Fok (1974) under the heading of emphatic expression.

There is no general agreement concerning whether stress actually exists in Cantonese words. Chao was probably the first to open the topic of stress in Cantonese: "Stress is not a constituting element of words in Cantonese... The identity of words in Cantonese is never

affected by stress" (1947: 36). Hashimoto also holds the view that "stress does not play a distinctive role in the phonological system" of Cantonese (1972: 101), while Barnett declares that Chao's view on p.36 is an oversight as he finds at least one word with different meanings when stressed and unstressed (1945-50: 742).

In further discussion, Chao points out that there are differences in the degree of force or loudness in different words in an utterance but "the variation depends upon the sounds concerned and upon the relative prominence of words in the sentence determined by stylistic factors". Chao was also probably the first to mention the term 'prominence' in relation to Cantonese (1947: 36), but he does not discuss it. Barnett considers that there is a definite stress in every closely joined combination and there are many words whose stress seems to be dictated purely by euphony or the requirements of emphasis. He observes that the majority of disyllables have a 'stressed-unstressed' ('s-u', henceforth) pattern. Hashimoto points out that high pitch tones in Cantonese are often heard as more strongly stressed in normal speech than other tones, and in contrast with Barnett that the preferred stress pattern for nouns is to have the last syllable more heavily stressed.

Fok and Vance are not concerned with the existence or the placement of stress but test what the phonetic cues of 'emphatic expression' or 'contrastive stress' would be in Cantonese. Tones can go through a wider frequency range, and are higher in amplitude and slightly longer in duration in emphatic expressions than in neutral expressions (see the production experiments by Fok (1974: 28, 30, 34)). But correct identification of tones with emphatic expression is less frequent than of tones with neutral expression in her perceptual experiments (p.42, 87). By giving test words contrastive stress in frame sentences, Vance (1976: 381) claims that contrastive stress in Cantonese has at most a minor effect on F_0 since the increases in F_0 in the contrastive stress word may simply reflect an increase in overall pitch level. He says his impressionistic judgment is that length and intensity are affected.

3.1.2. Lexical Stress and Isochrony in Cantonese

In my view, on an impressionistic account, it is found that Cantonese has no contrastive stress pattern in words similar to that of the English word pairs `**per**mit (v.) and **per**mit (n.)', or similar to the Mandarin compound/phrase pair `dōng.xi (thing: compound) and dōng xī (east and west: phrase)', or the Thai noun-noun compound/noun-verb sequences (weak-strong vs. strong-strong) (for examples see Potisuk *et al* 1996: 213). I can give no minimal pairs in Cantonese whose lexical meaning changes depending on the stress pattern. English is well known to have variable lexical stress. On the other hand, many languages are reported to have fixed lexical stress (Hyman 1977: 52). For example, Finnish is said to have fixed first syllable stress (Lehiste 1993: 239) and Polish is said to have fixed penultimate syllable stress (Hyman 1977: 52). However, there is neither variable nor fixed lexical stress in polysyllabic words in Cantonese. In Hyman's study of the incidence and placement of stress in 444 languages (Cantonese is not included), he reports that 16 have no stress-accent or tone (1977: 66). Clearly Cantonese is not the only language in the world which does not make lexical use of stress. I take comfort from these findings in denying that prominence in Cantonese is achieved through stress.

Whoever attempts to categorise a stress word-pattern in Cantonese would have about as much success as one who attempts to categorise a lexical tonal-pattern in English. However, I am not saying that every single word in Cantonese is perceived with the same degree of prominence in utterances. There is no contrast of stressed-unstressed syllables in Cantonese, but different degrees of prominence between syllables in utterances do exist.

Regarding isochrony, the language does not exhibit stress-timed rhythm. This does not mean that the language has to fall into the type of syllable-timed rhythm, which is defined as a language which shows a tendency for every syllable to "recur at equal intervals of time" (Abercrombie 1967: 97). In my experience, only school children tend to give approximately regular intervals of time to syllables (except the final syllable of a single utterance which may tend to be lengthened very much) and to give equal force in articulation to syllables thereby manufacturing a 'machine gun' effect when they try to recite aloud a given text, a poem, or a

multiplication chart, by memory. They fail to signal the points of interest, or the points of reference, that are appropriate in telling a story. No doubt such a style was employed by the informant of Jones and Woo in reading 'literary texts', which were consequently transcribed by Jones and Woo (1912) as ♪♪♩ , ♪♪♪♩ or ♪♪♪♪♩ for three-syllable poems, four-syllable poems and five-syllable poems, respectively (1912: 87). Such a style differs from the style of everyday conversation, is not acceptable to teachers and is regarded as a wrong way of reciting. The variation of prominence of syllables in actual speech tends not to be faithfully reflected in text readings. If one gives equal prominence to every single syllable he/she speaks in the language, one will sound foreign to the natives. If one attempts to learn the language and to articulate correctly, one should be able to manage not only all the segmental elements but also the suprasegmental parameters including how prominence is assigned to syllables in utterances.

As we discussed in chapter one, there is considerable variety of syllable structure and vowel and consonant form in Cantonese. The consonants and vowels both have different sonority and different phonological length. However, the phonological length of syllables varies to a great extent in actual speech. Thus a phonologically long syllable (say, an open syllable containing a long vowel) is sometimes realised in a much shorter form than a phonologically short syllable (say, a plosive-coda syllable containing a short vowel) when they occur in sequence. Syllables in sequence in utterances are not perceived as having equal duration. Thus Cantonese cannot be described as a language exhibiting a syllable-timed rhythm (like French for example, as claimed by Abercrombie 1967). Prominence in Cantonese lands on a lengthened syllable. We tend to hear sequences of alternating lengthened and shortened syllables in actual speech as linked into groups which we call `feet' (the term is borrowed from Abercrombie (1967: 131)). A foot is a rhythmic beat. For this, I would like to borrow the term `prominence based rhythm' from Laver to describe the perceived rhythm of Cantonese. (He uses the term for describing the rhythm of languages with widely varying syllable duration in which prominence arises in syllables of heavy metrical weight, syllables receiving lexical stress, or both, with unstressed syllables being made particularly non-prominent through vowel reduction. (Laver 1994: 529)).

I shall speak of prominence rather than stress simply because the greater force of articulation (i.e., stress, as defined above) is only one part of the picture of achieving prominence in Cantonese. Prominence is achieved primarily by longer duration, which can also be accompanied by greater loudness and the exaggeration of the pitch, interweaving with the lexical tones and ultimately governed by intonation in utterances.

3.2. THE 'WORD'

3.2.1. Defining the `Word'

The first problem we encounter before starting the description of prominence in Cantonese is how to define 'word'. I must warn that there is no equivalent corresponding to the English 'word' in Chinese. In discussing stress Barnett (1945-50) and Hashimoto (1972) avoid using 'word' explicitly but use monosyllable and disyllable, etc., instead. On the other hand, Chao (1968: 136-193) has a comprehensive discussion on defining the 'word' in Mandarin Chinese. At the end he gives synoptic tables of 'word-like units' as a conclusion, in which 'morphemes' and `morpheme complexes' are listed. Bloomfield was a pioneer in defining 'word': "Forms which occur as sentences are free forms. ...A free form which is not a phrase, is a word. ...[A] word is a minimum free form. ... is the smallest unit of speech. ... [W]e have the custom of leaving spaces between words in our writing and printing" (Bloomfield 1935: 178). He uses modern Chinese as an example of a language which has free forms only and, in which each word consists of one syllable.

3.2.2. /tɕi:^{6}/ (Word)

In that sense, one syllable equals one word in Chinese. From this, we can safely say that a word for 'word' in Chinese is /tɕi:^{6}/, because in Chinese almost all monosyllables are morphemes and are written as single characters with spaces in between - those, we call /tɕi:^{6}/, such as /ɕy:^{1}/ (book) and /t^{h}ɔ:j^{2}/ (table). Morphemes which consist of more than one syllable are very rare, such as /ka:t^{6} tɕa:t$^{6/2}$/ (cockroach), /k^{h}ej^{4} lɐn$^{4/2}$/ (a mythical animal which is a symbol of joy and peace). In this case, one morpheme equals two syllables equals

two /tɕi:^{6}/. In other words, /ka:t^{6} tɕa:t$^{6/2}$/ (cockroach) is a disyllabic morpheme, and consists of two /tɕi:^{6}/. A /tɕi:^{6}/ does not only refer to a monosyllable in writing which is monographic, it is commonly translated as `character', but can also refer to a monosyllable in spoken Chinese. One can ask, "/ka:t^{6} tɕa:t$^{6/2}$ ni:^{1} lœ:ŋ^{5} kɔ:^{3} tɕi:^{6} ti:m^{2} ɕɛ:^{2} /?" (cockroach-these-two-measure-characters-how-write - How do [you] write the two characters - (cockroach)?) A Cantonese idiom describing someone who does not have clear articulation for every single word is /ka:w^{3} tɕi:^{6} pɐt^{1} tɕ^{h}eŋ^{1}/ or /ka:w^{3} tɕi:^{6} m^{4} tɕ^{h}eŋ^{1}/ (bite-word-not-clear). In English, the `four letter word' is a euphemism for a swear word. In Cantonese, the equivalent expression is /ɕa:m^{3} tɕi:^{6} keŋ^{1}/ (three-word-scripture) as in /k^{h}ɵy^{5} kɔ:ŋ^{2} ɕa:m^{3} tɕi:^{6} keŋ^{1}/ (he-utters-three-words-scripture - he utters the 'four letter word'). So far, it seems clear that a /tɕi:^{6}/ in Cantonese is not only a written unit, i.e., a character, but also refers to a spoken unit which consists of one morpheme and one syllable, which is what a child learns to speak, and what a university counts for the maximum or minimum length of dissertations.

Even though every monosyllabic morpheme can be uttered in isolation, in an utterance some of them need to combine with another morpheme in order to be freed, as /m^{4}/ (not) in /m^{4} tɕ^{h}eŋ^{1}/ (not-clear). However, in a loose sense, if we regard the English word `the' or 'is' as a word which satisfies the criterion of a `free form' in actual speech (as classified by Bloomfield - p.179), all Cantonese /tɕi:^{6}/ including particles, suffixes, measures and the negation /m^{4}/ are free and thus are words. All Cantonese /tɕi:^{6}/ also satisfy another criterion, that is interruptibility. For example, /m^{4} tɕ^{h}eŋ^{1}/ (not-clear) can be broken up by /hɐj^{6} kej^{2}/ (is-quite) into /m^{4} hɐj^{6} kej^{2} tɕ^{h}eŋ^{1}/ (not-is-quite-clear - not quite clear). Even polysyllabic monomorphemes can also be interrupted by other morphemes. For example, /ka:t^{6} tɕa:t$^{6/2}$/ (cockroach) can be broken up by /mɐt^{1} jɛ:^{5}/ (what-stuff) into /ka:t^{6} mɐt^{1} jɛ:^{5} tɕa:t$^{6/2}$/ ('cock'-what-stuff-'roach' - what cockroach are you talking about?).

3.2.3. /tɕʰi:⁴/ (Word)

The Cantonese word /tɕi:⁶/ therefore, seems to correspond to the English word `word' very well until the English term `compound word' appears. An example of a compound word in English is a lexical item such as 'tea-bag'. However, the compound /tɕʰa:⁴ pa:w¹/ (tea-bag) in Chinese consists of two /tɕi:⁶/, which is in this sense two words and is called a disyllabic /tɕʰi:⁴/. The term /tɕʰi:⁴/ is a learned term devised by modern Chinese grammarians (Chao 1947: 37). It is regarded as the smallest unit of free form and `semi-free' form in speech. It covers compound words but also monosyllabic simple words. Take the examples from above, /tɕʰa:⁴/ (tea) is a /tɕʰi:⁴/, /pa:w¹/(bag) is a /tɕʰi:⁴/, /tɕʰa:⁴ pa:w¹/ (tea-bag) is also a /tɕʰi:⁴/; /ka:t⁶ tɕa:t⁶/² / (cockroach) is a /tɕʰi:⁴/, /m⁴/ (not) is a /tɕʰi:⁴/, /tɕʰeŋ¹/ (clear) is also a /tɕʰi:⁴/. /tɕʰa:⁴/ (tea) is an example of free form and /m⁴/ (not) is an example of semi-free form. /tɕʰi:⁴/, in this sense, is a term used by linguists. It is not in general use, unlike the English word `word'. Traditionally, /tɕʰi:⁴/ refers to a particular kind of poetry with fixed rhyme schemes. Such a rhyme scheme is a so-called /tɕʰi:⁴ ti:w⁶/ (poetry-rhyme), whereas /tɕi:⁶ ti:w⁶/ (word-rhyme) refers to syllable tone. /tɕʰi:⁴/ also refers to an expression consisting of words as in /tɕʰi:⁴ pɐt¹ ta:t³ ji:³/ (expression-not-reach-meaning - the words fail to convey the idea) or /tɕʰow³ tɕʰi:⁴ pɐt¹ tɔ:ŋ³ / (arrange-words-not-appropriate - inappropriate wording). But if one wants to say `someone used a wrong word', he has to say /joŋ⁶ tɕʰɔ:³ tɕi:⁶/ (use-wrong-word) to refer to a single syllable word which is incorrectly used. As mentioned above, /tɕʰi:⁴/ can refer to an expression consisting of more than one /tɕi:⁶/, and includes compounds, such as /ɕy:¹ tʰɔ:j²/ (write-table - desk, in which both constituents are roots), and derived words such as /tʰɔ:j² tɕɐj²/ (table-nominal suffix with a meaning of diminish - small table, in which the /tɕɐj²/ is a suffix with the root /tʰɔ:j²/). A compound differs from a complex compound such as /ta:j⁶ hɔ:k⁶ kɔ:ŋ² ɕi:¹/ (big-study-talk-teacher - university lecturer), in which, /ta:j⁶ hɔ:k⁶/ (big-study - university) is a subordinate compound and /kɔ:ŋ² ɕi:¹/ (talk-teacher -

lecturer) is also a subordinate compound, /ta:j^{6} hɔ:k^{6}/ modifies /kɔ:ŋ^{2} ɕi:^{1}/, /ta:j^{6} hɔ:k^{6} kɔ:ŋ^{2} ɕi:^{1}/ is a subordinate complex compound. It is possible to have a pause in between two compounds but not within a compound. In standard romanisation - pinyin, not graphs - of Mandarin, a space is written in between the two compounds of a complex compound but not within a compound. A complex compound is called /tɕ^{h}i:^{4} jow^{2}/ (/jow^{2}/ means `group' - a group of /tɕ^{h}i:^{4}/).

3.2.4. /tɕi:^{6}/ and /tɕ^{h}i:^{4}/ in Standard Work of Scholarship

There are two types of dictionaries of Chinese which are both called 'dictionary' in English. One is /tɕi:^{6} ti:n^{2}/, one is /tɕ^{h}i:^{4} ti:n^{2}/ (/ti:n^{2}/ means standard work of scholarship). In a /tɕi:^{6} ti:n^{2}/, every entry is a /tɕi:^{6}/, compounds and phrases are given only for the purpose of exemplyfing the different meanings of the /tɕi:^{6}/, as illustrated by the "/tɕoŋ^{1} wa:^{4} ɕɐn^{1} tɕi:^{6} ti:n^{2}/" (Zhonghua Xin Zidian (China New Dictionary) 1976). In a /tɕ^{h}i:^{4} ti:n^{2}/, under the heading of every single /tɕi:^{6}/, all possible compounds, commonly used complex compounds or phrases which consist of or begin with the same entry are listed, as the "/hɔ:n^{3} jy:^{5} tɕ^{h}i:^{4} ti:n^{2}/ " (Hanyu Cidian (Chinese Dictionary) 1990). A syllabary in Chinese by no means has to be called /tɕi:^{6} wu:j^{6}/ (/wu:j^{6}/ means collection), as the "/k^{w}ɔ:k^{3} jɐm^{1} jy:t^{6} jɐm^{1} ɕɔ:k^{3} jɐm^{1} tɕi:^{6} wu:j^{6}/" (Guoyin Yueyin Suoyin Zihui (A Chinese Syllabary Pronounced According to Mandarin and Cantonese) 1987). There are some scholars who take a different view of the distinction between /tɕi:^{6}/ and /tɕ^{h}i:^{4}/ that we cannot ignore. Lü collects 800 /tɕ^{h}i:^{4}/ in his /ji:n^{6} tɔ:j^{6} hɔ:n^{3} jy:^{5} pa:t^{3} pa:k^{3} tɕ^{h}i:^{4}/ (Xiandai Hanyu Babai Ci (The Eight Hundred Modern Chinese Words) 1980), in which he defines /tɕ^{h}i:^{4}/ as including free form and semi-free form morphemes. /ɕy:^{1}/ (book) is an example given for the former and a particle /nɛ:^{1}/ for the latter (p.4). He does not define /tɕi:^{6}/ but uses it as the written form for morphemes (p.4). Thus, his 800 entries of /tɕ^{h}i:^{4}/ include morphemes such as /hɔ:j^{1}/ (open) as well as compounds such as

/hɔ:j^{1} tɕ^{h}i:^{2}/ (open-begin – beginning). In their /kɔ:ŋ^{2} ɕek^{1} k^{w}ɔ:ŋ^{2} tɕɐw^{1} wa:^{2} tɕ^{h}i:^{4} ti:n^{2}/ (Gangshi Guangzhouhua Cidian (HongKong Cantonese Dictionary) 1999), Cheung and Ni collect 7000 entries of /tɕ^{h}i:^{4}/ including monosyllables and polysyllables such as `/jɐw^{5} t^{h}ɐw^{4} wɐj^{1}/, /mow^{5} mej^{5} tɕɐn^{6}/' (have-head-brave, -no-tail-position - Fine start, poor finish).

3.2.5. Morpheme, Syllable, Single Word, Compound Word, Complex Compound Word, Phrase, Set Phrase and Expression

So far, I presume it has become clear that the word `word' in English can be translated into either /tɕi:^{6}/ or /tɕ^{h}i:^{4}/ in Cantonese: /tɕi:^{6}/ refers to a word which consists of one morpheme, one syllable, which can be uttered in isolation and is written with a single character - I call it a simple word; whereas /tɕ^{h}i:^{4}/ refers to an expression which has a lexical meaning and consists of more than one syllable, more than one morpheme, and where a pause can be possibly inserted at the right-end boundary when in a string of words - I call this a compound word. In other words, a word can refer to a simple word /tɕi:^{6}/ or a compound word /tɕ^{h}i:^{4}/. Idioms in Chinese usually consist of four syllables (but not necessarily in the last two cases below) and are commonly called /tɕ^{h}i:^{4}/, as /ɕeŋ^{4} jy:^{5} tɕ^{h}i:^{4}/ (/ɕeŋ^{4} jy:^{5}/ means 'fixed sayings'), /hɐw^{2} jy:^{5} tɕ^{h}i:^{4}/ (/hɐw^{2} jy:^{5}/ means 'verbal') and /tɕok^{6} jy:^{5} tɕ^{h}i:^{4}/ (/tɕok^{6} jy:^{5}/ means 'colloquial'). They are called /tɕ^{h}i:^{4}/. However, they are not necessarily compounds, they could be set phrases. These set phrases have a lexical meaning, they can be phrase-words and are usually included in dictionaries. /tɕ^{h}i:^{4}/ seems to have a looser sense of including compounds and derived words plus set phrases. However, I only use the English word 'word' to refer to simple words, compound words, complex compound words and derived words in Cantonese. I use the English word 'phrase' to refer to any phrases in Cantonese, including set phrases or just a phrase such as /ni:^{1} tɕœ:ŋ^{1} t^{h}ɔ:j^{2}/ (this-measure-table - this table), which are considered to be a larger free form but smaller than a clause (cf. Bloomfield 1935: 185). Sometimes it is very difficult to draw a clear line between a compound and a phrase. /ta:j^{6} hɔ:k^{6}/ (big-study - university) is an adjective-noun subordinate compound, while /ta:j^{6} t^{h}ɔ:j^{2}/

(big-table) is a phrase because it has no added meaning beyond the literal meaning as in big dog, big house, etc. But set phrases do have a lexical meaning. The term `expression' seems to fit very well to include words and phrases which consist of more than one morpheme. (cf. Bloomfield 1935: 196 and Chao 1968: 258).

3.3. PROMINENCE IN CANTONESE

3.3.1. The Loudness of Speech

From a sociolinguistic point of view, accent-community is one of the factors used to characterise a typical average loudness of the speech of its individual members. It is said that the speech of Gaelic communities in Scotland seems markedly less loud than the speech of urban Egyptian Arabic communities, for instance (Laver 1993: 500). Chinese speakers tend to give the impression of speaking loudly, including Cantonese speakers. When you hear two or more people raising their voice to each other, they are not necessarily shouting at each other as you might think, it may be just friendly chatting or relaxed joking. Chen (1991: 4), a sociolinguist, thought that in Hong Kong it might be only uneducated people who would speak with a loud voice, but later in the restaurant of a university he heard professors and lecturers talking with unnecessarily loud voices. He concludes that it is because Hong Kong is a noisy place - the decibel level of background noise is as high as 80 - that everyone is induced to raise his voice. My experience of this is not only with Hong Kong Cantonese but also with other variant accents of Cantonese or other Chinese dialects, not to mention Mandarin, outside of Hong Kong. Sometimes if I do not follow people's conversation, I have to watch in order to confirm that they are actually laughing and chatting in a friendly way rather than shouting. Regarding Cantonese, I think this may be due to the lack of dynamic stress in the language. The whole sentence or the whole utterance, rather than a single word or a single syllable, is highlighted by raising the pitch and the volume when the speaker wants to attract attention. This view can be partly supported by Vance's finding discussed above - the increases in F_0 on a contrastively stressed word may simply reflect an increase in overall pitch level (1976: 381). It is just like when Cantonese speakers hear a debate in the House of Commons by British English speakers, all they can hear is

aggressive shouting. But, behind the high pitch and loudness, we, as phoneticians, can still hear the different degrees of pitch and loudness with different words in an utterance.

Those differences are the realisation of prominence of syllables when interweaving with intonation. A prominent syllable achieved by loudness, is not obligatory in Cantonese. It is unlikely to emerge in reading a word-list since it does not give any additional meaning on top of the lexical meaning. Jones and Woo (1912) failed to point out any difference in degree of loudness in the transcription of their 'texts in colloquial style' in Cantonese, but one may think that it is simply because their informant only read the prepared texts rather than said them in live conversation. When my informants were asked to read a long list of polysyllabic words, some tended to give greater force of articulation on the first syllable, some on the second, some on the last, or some on none, etc, all the way down without any consistency, particularly when they were tired. This reveals that there is no regular prominent pattern achieved by the degree of loudness at word-level in Cantonese.

Let us take an example from Barnett (1945-50: 742) who thinks the word /tɕi:^{3}/ means `most' when stressed but means `only then' when unstressed, in the example of /t^{h}eŋ^{1} jɐt^{6} tɕi:^{3} how^{2} hɐy^{3} mɐn^{6} ha:w^{6} tɕ^{h}œ:ŋ^{2}/ (tomorrow-day-[most/only then]-good-go-ask-school-master). I take it for granted that his 'stress' refers to the effect of loudness, as he says that he does not propose to say anything about the effect of stress on duration and tonality. He proposes that if /tɕi:^{3}/ is stressed the sentence means (a) "You'd best ask the headmaster tomorrow"; if it is unstressed, the meaning is (b) "Don't on any account ask the headmaster before tomorrow". He is right in the sense that the example is given in an utterance, so that we have no way to avoid discussing the word without reference to syntactic structure and utterance-intonation. In fact, the only reason for /tɕi:^{3}/ to receive prominence is that it is an adverb modifying the adjective /how^{2}/ (good) in an utterance. /tɕi:^{3}/ (most) plus /how^{2}/ (good) to form `the best', is the focus of meaning in the given sentence. Being the focus of the sentence as well as being the prominence receiver within the focus of a combination of adverb-adjective, the double identities make the adverb /tɕi:^{3}/ (most) not only much more prominent than the adjective

/how^{2}/ (good), but also stand out from the whole sentence. However, if it is in coordination with another adverb /how^{2}/ (very) to form an adverb compound `only then' or `not until', the prominence which it receives is much less than when it is preceding an adjective in an utterance. Furthermore, being an adverb meaning `not until', it receives less focus in terms of semantics in the given sentence since there are other pieces of necessary information preceding (i.e., a time-adverb) or following (i.e., a verb). If we assume that the last few words (/hɵy^{3} mɐn^{6} ha:w^{6} tɕ^{h}œ:ŋ^{2}/ (go-ask-school-master - go to ask the headmaster)) of the sentence are the old information which can be taken out, so we shorten the sentence to a minimum size and we can find the differences of minimum information contained. When it means `the best', the shortest sentence can be /t$^{h}$eŋ$^{1}$ jɐt$^{6}$ tɕi:$^{3}$ how$^{2}$/ (tomorrow-day-most-good - the best is tomorrow), as an answer to the question `when is the best time to go to ask the headmaster', or even /tɕi:^{3} how^{2} ka:^{3} la:k^{3}/ (most-good-particles - the best, for sure), as an answer to the question `don't you think I should go to ask the headmaster tomorrow'. When it means `not until', the shortest sentence can be /t^{h}eŋ^{1} jɐt^{6} tɕi:^{3} how^{2} hɵy^{3}/ (tomorrow-day-until-good-go - go not until tomorrow), as an answer to `should I go to ask the headmaster today'. In the latter case, the verb /hɵy^{3}/ (go) is necessary information in the sentence sharing the focus of the preceding adverb. Anyway, as I discussed in Chapter I, this is an extreme case as we hardly ever hear nowadays examples such as the ones he quoted fifty years ago. /tɕɵy^{3} how^{2}/ (most-good) is the daily usage for `the best' and /ɕi:n$^{1}$ tɕi:$^{3}$/ (first-then) for `not until'. I think, even if there was any confusion arising in the examples he quoted, listeners would ask for clarification, and speakers would rephrase the sentence or use a synonym rather than try to put more `stress' on the `stressed word' or try to give `less stress' on the `unstressed word', since this is not a matter of minimal pairs of stressed and unstressed but a matter of prominence.

Certainly, a syllable which is pronounced more loudly than its neighbouring syllables when under focus, appears to listeners to be more prominent. I observe that this phenomenon is also found in other languages, for example, Japanese and Korean, since loudness is one (the

least consistent one) of the qualities of focused syllables both linguistically and paralinguistically. If /tɕi:^{3} how^{2}/ in sentences (a) and (b) above is replaced by /tɕey^{3} how^{2}/ (most-good: the best) (sentence (a)) and /ɕi:n^{1} tɕi:^{3}/ (first-then - not until) (sentence (b)), being in each case under focus, the most obvious device we can find to produce prominence is to lengthen these two adverbial combinations (particularly the first syllables); loudness of syllables does not play a significant linguistic role at lexical level in the language.

3.3.2. Rhythm

The variation of length of syllables in speech is a predominant characteristic of Cantonese. The duration of a syllable in a citation form is practically never the same as in live speech. A vowel before a plosive coda is shorter than before a nasal while the relative difference in vowel length is still maintained (see chapter one). In live speech, a syllable ending with a plosive coda is not necessarily shorter than a syllable ending with a nasal. There is a contrast of long and short vowels in the Cantonese vowel system. However, the contrast is based on differences of vowel quality (see chapter one), as, for example, the contrast between /i:/ and /e/, /a:/ and /ɐ/, etc. To lengthen any vowel will not result in confusing the vowel system. Syllables can be lengthened and shortened without affecting their morphemic or lexical meanings. When syllables appear in a string, they adjust their length, appearing to be irregular rhythmic patterns and they never maintain their duration in citation forms. The variation of duration of syllables in connected speech tends not to be faithfully reflected in text readings in Cantonese. In Cantonese text readings, speakers tend to give approximately equal duration to each syllable. This reflects the difference between Mandarin and Cantonese in terms of prominence. Mandarin speakers can group syllables into compounds or phrases from the same reading text, and assign stress and rhythm to compounds and phrases, as stress and rhythm are part of the composition of a compound or a phrase. In Mandarin, a stressed syllable has a longer duration than an unstressed syllable. For example, the stressed first syllable in /liángxin/ (good-heart: conscience) is longer than the unstressed second syllable and the stressed first syllable in /fàngchu.lai./ (release-out-come: let out) is longer than the unstressed second and third syllables. In contrast, there is no lexical stress nor regular rhythmic patterns in Cantonese. Take an example from the

discussion above: /t^{h}eŋ^{1} jɐt^{6}/ (tomorrow-day: tomorrow), __ __ would be the normal pattern in reading; but in live conversation, the rhythms can vary to ____ __ or __ ____ without destroying the lexical meaning. In a longer string of syllables, the normal rhythm involves variation in the length of syllables; for syllables to have even length is considered unnatural. If a syllable is under focus, the prominence is indicated by syllable lengthening. This kind of lengthening is a common feature of prominence in Cantonese. This may be due to the heavy functional load of pitch which does not leave much room for prominence exploitation. Thus, in the example above, if the first syllable is under focus, the rhythm will be ____ __; if the second syllable is under focus, the rhythm will be __ ____. (If we take the underlining "___" for a syllable, __ and ____ indicate relative shorter or longer duration.)

The duration of a syllable in an utterance varies in relation to different factors, primarily in relation to the formation of the word (e.g., monomorphemic word and compound word, etc.) which the syllable(s) belongs to, and under focus. By lengthening or shortening the duration of a syllable in a polysyllabic word, the prominence of the syllable is achieved.

I shall give an account of prominence achieved by the variation of duration on reduplicated words, coupling words and onomatopoeic words before going on to give an account of compounds.

I must warn that there is no single rhythmic pattern in Cantonese words. While I point out the most common rhythmic pattern in any particular word, I am not ruling out other alternative rhythmic patterns which could actually occur in utterances but do not violate the phonological structure of the word.

3.3.2.1. Reduplications

Certainly in many languages the process of reduplication is a morphological phenomenon which has a particular semantic function. For example, this process may be found in Greek, where the initial consonant of the root is reduplicated in perfective forms, e.g., /'lu:o:/ (I loose), becomes /'leluka/ (I have loosed) (see Crystal 1985: 259); in Ewe, the nouns /sasya/

(drying) and /dzadzra/ (selling) are derived by reduplication of the corresponding verb root /sya/ (to dry) and /dzra/ (to sell), respectively (see Ansre 1963: 128). These examples clearly indicate that morphological reduplication is a process of repetition whereby the form of a prefix/suffix reflects certain phonological characteristics of the root (Crystal 1985: 259) and the root (or part of the root) is reduplicated in grammatical contexts.

Turning to reduplications in Cantonese, which I am going to discuss below, we find that some do reflect certain phonological characteristics of the root (e.g., couplings) and some do not (e.g., duplicated rhythmic suffix), but they are all morphological devices for intensifying the meaning. Certainly reduplications have a special function in morphology, but I think that in Cantonese reduplication is also a process of achieving prominence, on the ground that the morphological intensification is brought about by prolonging the root to at least double the length. Since Cantonese words are basically monosyllabic, I propose that the very fact of reduplication, coupling and onomatopoeia in polysyllabic words has a prominence effect. This prominence is achieved by lengthening a word to double or several times from monosyllable-long to polysyllables-long.

3.3.2.1.1. Reduplicated Rhythmic Suffixes - Forms of (X)YY

Let us start from the duplicated rhythmic suffixes in adjectives, such as /ha:k^{1} mɐŋ^{1} mɐŋ^{1}/ (dark-duplicated rhythmic suffixes), or /ly:n^{6} la:p^{6} la:p^{6}/ (mess-duplicated rhythmic suffixes). Most of the duplicated rhythmic syllables are suffixes, prefixes are rare, as in /la:p^{6} la:p^{6} ly:n^{6}/ (duplicated rhythmic prefixes-mess). All forms of adjective in combination with duplicated rhythmic suffixes have an intensified meaning. The duplicated rhythmic suffixes do not necessarily have a meaning in the context, they are only added to prolong the adjectives. The most common rhythm of the three syllables is staccato. This means every individual syllable is chopped abruptly, that is produced by means of more energetic movements, with the pumping of more breath from the lungs. They are heard to last for approximately equal duration but do not sound prolonged because of the abrupt chopping. They have approximately equal loudness. They are louder than a non-staccato syllable because of the more energetic movements. In any event, the adjective is lengthened by adding two more

syllables. If the first syllable is lengthened, only the last two syllables are chopped (in this case, the first syllable can be accompanied with breathy voice). The possibility of lengthening either the second or the third syllable is not ruled out. Any form of reduplication in adjectives is to give prominence in utterances. Thus, it is not surprising that the adjective is usually expressed with an emphatic expression. The common rhythmic patterns can be as follows (! indicates staccato and * indicates breathy voice; the most common pattern is listed first) :

i. ! ! !
ii. *___ ___ ___ (or ___ ___ ___)
iii. ___ ___ ___

The new trisyllabic word can be one, two or three rhythmical feet, depending on the numbers of the lengthened syllables the word has. It was suggested in section 3.1.2. that a foot is normally grouped with a lengthened and a shortened syllable. A lengthened syllable is the constituting syllable of a foot. It normally accompanies a shortened syllable, occasionally two, seldom three but can also be none. In pattern (i) above, every syllable is chopped thus is heard to occupy equal beat – they are counted as three feet when one taps one's finger. In the patterns (ii) and (iii), the foot is made up of one lengthened syllable and two shortened syllables. The lengthened syllable occupies half a foot whereas the other comparatively shortened syllables together occupy another half.

As for the non-staccato pattern (ii) and (iii), it is not surprising that there is a strong likelihood that the last syllable will be lengthened as it stands at the boundary of a foot which can be the margin of an uninterrupted utterance. This kind of utterance-marginal lengthening (the term is used in Laver 1994: 532) is not considered to be part of the rhythm of the word. However, it is hard to tell whether the lengthened final syllable is part of the rhythm of the word or a result of the utterance-marginal lengthening. Therefore, the possibility for pattern (ii) becoming ____ __ ____ and for pattern (iii) becoming __ ____ ____ will not be illustrated here, even though it does exist. This will apply to the following discussion. As it stands at the boundary of a foot, the final syllable also has the possibility of exhibiting the raising of the tonal offset of the last syllable (as long as it is not a high pitch tone itself, i.e., T1 or T2). The

pitch raising is restricted to tonal offset, i.e., the tone starts from its original onset pitch and rises to a higher pitch, around the range of the offset of T5.

By adding two more syllables, the reduplicated rhythmic suffixes prolong the adjective, converting it from monosyllabic to trisyllabic. Being a trisyllabic adjective, it can vary the length of the root word and of the added syllables. This is a clever way to achieve prominence in a language like Cantonese which lacks dynamic stress and in which pitch already has a heavy functional load (for signalling lexical tones and utterance intonation).

3.3.2.1.2. Reduplicated Adjectives

3.3.2.1.2.1. Forms of XX and XX(Y)

The forms of adjective reduplications in this paragraph have long attracted linguists' attention. Unfortunately most linguists discuss them from the viewpoint of word class (e.g., Gao 1980: 60 and Rao *et al* 1982: 315) but never discuss prominence. Forms of the adjective reduplications are: /ha:k^{1} ha:k^{1}/ (dark-dark) and /ha:k^{1} ha:k^{1} tej^{2}/ (dark-dark-a bit). They all have an intensified meaning. The length of the first syllable of XX is usually twice as long as the second, and also it features greater loudness and sometimes breathy voice. This is in contrast to XX(Y), in which the second syllable is usually twice as long as the first. The most common rhythmic patterns are (' indicates a greater force of articulation):

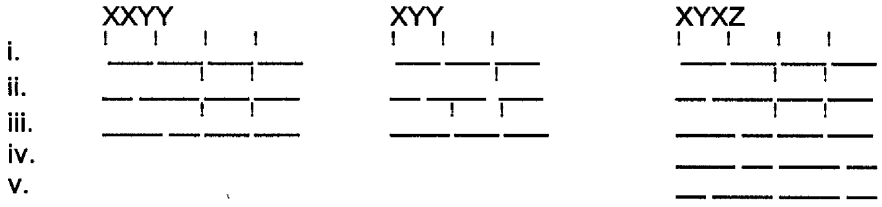
	XX	XX(Y)
i.	____ __ (or *'____ __)	__ ____ __

The staccato pattern is not applied to these adjectives as they are constrained by specific features of prominence as we set out below. The first syllable of the adjective in the form XX usually changes to a high pitch T2 (as long as it is not a high pitch T1 or T2 itself) in order to realise prominence in combination with prolonged length and greater loudness, i.e., /ly:n^{6} ly:n^{6}/ (mess-mess) is realised in /ly:n^{2} ly:n^{6}/. This feature of prominence also applies to the second syllable of the adjective in the type of the adjective preceding the suffix - form XX (Y), i.e., /ly:n^{6} ly:n^{6} tej^{2}/ (mess-mess-a bit) is realised in /ly:n^{6} ly:n^{2} tej^{2}/. Thus, prominence of

these two forms of adjective is achieved by not only prolonging the length of the word as in rhythmic reduplications, but also by heavily lengthening one particular syllable of the word and by changing the tone to higher pitch phonologically in one particular syllable of the word. (This will be discussed further in section 3.3.3.) Gao (1980: 60) points out that the sound of these particular syllables discussed above are lengthened and intensified, but does not discuss the matter further.

3.3.2.1.2.2. Forms of XXYY, XYY, XYXZ

Other forms are: the form of XXYY, such as /ha:k^{1} ha:k^{1} ɕɐt^{6} ɕɐt^{6}/ (dark-dark-solid-solid - dark and strong); the form of XYY, such as /mi:n^{6} ha:k^{1} ha:k^{1}/ (face-dark-dark - look angry); the form of XYXZ, such as /ha:k^{1} hɐw^{2} ha:k^{1} mi:n^{6}/ (dark-mouth-dark-face - look angry). (This last is called partially duplicated and is not parallel to the earlier examples.) They all have an intensified meaning. The most common rhythm for the above forms is also staccato. If a syllable is lengthened, it can be accompanied by greater loudness. The common rhythmic patterns are:



The form of XXYY is an expansion of a coordinate compound (i.e., XY - X and Y are both centres and have the same function as the whole construction). After reduplication, the Ys tend to be chopped while the Xs vary their lengths. On the other hand, the form of XYXZ is a complex compound consisting of two adjective subordinate compounds, X_1Y and X_2Z (i.e., X_1 modifies Y and X_2 modifies Z). Its patterns (iv) and (v) clearly show that by lengthening one of the syllables the second compound can easily break off from the first compound to form a separate foot. The possible different rhythms between the two types of adjective reduplication clearly indicate that patterns of rhythm not only have to do with the number of syllables in a word but also with the formation of the word. Every syllable in these examples

is a morpheme in its own right which differs from the examples given in the duplicated rhythmic suffixes.

3.3.2.1.3. Reduplication differs from Repetition

As discussed above, words can achieve prominence by extending the numbers of syllables, and by the interplay of tonal pitch, syllable length and loudness between these syllables. This is characteristic of prominence in Cantonese. Reduplication is not to be equated with repetition. In a speaker's mind, /ha:k^{1} ha:k^{1}/ (dark-dark) is a duplicated adjective which is more than a mere a repetition of adjectives. This is because the reduplication is not a repetition of the meaning of a word, it is a way of expanding the length of a word internally to achieve semantic intensification and prominence. After being duplicated, the reduplication is a word in its own right.

Thai also has different forms of reduplication. Luksaneeyanawin (1983: 103) gives a detailed account of Thai reduplications in discussing Thai intonation. Some of the forms of the reduplications are very similar to the forms of /ha:k^{1} ha:k^{1} tej^{2}/ (dark-dark-a bit) and /ha:k^{1} ha:k^{1}/ (dark-dark) in Cantonese, which she calls "simple reduplicative" and "intensifying reduplication", respectively. The accentual pattern of the Thai disyllabic simple reduplicatives is R 'O, as in /lɯ:aŋ^{r} 'lɯ:aŋ^{r}/ (yellowish) - the first /lɯ:aŋ^{r}/ is the reduplicator and unaccented while the second /'lɯ:aŋ^{r}/ (yellow) is the base word and is accented. The accentual pattern of the disyllabic intensifying reduplicatives is *R 'O, as in /*lɯ:aŋ^{r} 'lɯ:aŋ^{r}/ (very yellow) - the first /*lɯ:aŋ^{r}/ (yellow) is a reduplicator acquiring an intensifying accent (indicated by *) while the second /'lɯ:aŋ^{r}/ is the base word which is accented. By duplicating the base word and by varying the degree of accent of the reduplicator, the two syllables consisting of identical segments and tones do not mean the same thing to her - they are not repeating themselves to reach the meaning of `very' or `-ish'. The effect of the reduplication in these two cases is similar to that in Cantonese from the point of view of achieving prominence; but the processes of reduplication differ between the two languages - one varies

its degrees of accent and one varies its rhythms. Cantonese varies the rhythm of /ha:k^{1} ha:k^{1}/ (dark-dark) and /ha:k^{1} ha:k^{1} tej^{2}/ (dark-dark-a bit) to achieve prominence. However, an additional morpheme is necessary to form the meaning of `a bit' or `-ish' to the duplicated adjective in contrast with the former form.

This clearly indicates that duration is not a constituent element of a morpheme. There are no minimal pairs of words which only differ in duration. Variation of the length of the syllables in a word is an aid to achieving prominence.

3.3.2.1.4. Reduplicated Nouns, Verbs and Adverbs

Duplicated nouns and measures differ from the above duplicated adjectives. The reduplication is not to achieve intensification or prominence but to signal the plural. They do not fall into the same category of adjective reduplications which use word prolongation to achieve prominence. This is reflected in their rhythm. They do not normally use staccato rhythm but are the same as normal compounds (see discussion in 3.3.2.4.) Noun reduplication is very rare, as in /jɐn^{4} jɐn^{4}/ (person-person - everybody) or, /loŋ^{1} loŋ^{1} la:^{3} la:^{3}/ (hole-hole-gap-gap - every hole and gap). Measure reduplication is more common, as in /kɔ:^{3} kɔ:^{3}/ (measure-measure - everyone). The possibility for syllable lengthening in disyllabic noun and measure reduplications is different from adjective reduplication as discussed above: either syllable can take longer duration in an utterance. For XXYY tetrasyllabic noun reduplications, lengthening either the first or the second syllable twice as long as the other is considered the most common pattern.

Reduplicated verbs and partial reduplication of nouns or verbs have an intensificatory effect as they are not used as nouns. Adverbs are also intensified by reduplication. Tetrasyllabic verbs such as /jɐm^{2} jɐm^{2} ɕek^{6} ɕek^{6}/ (eat-eat-drink-drink - eating and drinking frequently) and tetrasyllabic adverbs such as /fa:j^{3} fa:j^{3} tɕ^{h}ej^{3} tɕ^{h}ej^{3}/ (quickly-quickly-neatly-neatly) also have the same rhythmic patterns as the XXYY tetrasyllabic adjective reduplications. Rhythmic patterns in partial reduplications of nouns, verbs and adverbs, e.g., /mi:n^{6} hoŋ^{4} mi:n^{6} lok^{6}/

(face-red-face-green - be flushed or be pale), /ka:w^{2} foŋ^{1} ka:w^{2} jy:^{5}/ (stir-wind-stir-rain - make trouble) and /t^{h}oŋ^{4} low^{1} t^{h}oŋ^{4} pow^{1}/ (together-stir-together-boil - work closely), respectively, are the same as those in XYXZ partial reduplications of adjectives.

3.3.2.2. Couplings

Recall that the constituents of a syllable are onset, rhyme and tone, as discussed in chapter one. Reduplication is not limited to whole syllables, it can involve any part of the constituents of a syllable. I call this type of reduplication - coupling, or rhyme coupling. For example, /la:w^{2} ka:w^{6}/ (untidy) and /la:t^{6} t^{h}a:t^{3}/ (dirty), are rhyme reduplicated; /la:w^{4} ɕa:w^{4}/ (careless) and /ta:p^{3} ɕa:p^{3}/ (which has to follow the negation /mow^{5}/ to mean: `careless') are both rhyme and tone duplicated. Onset reduplications are rare, as in /lɔ:ŋ^{1} lɐj^{2}/ (angry with no reason). Most examples of these reduplications are purely colloquial, they may not even have a proper written form. If one really wants to write these words down, one can either make up a new character with the same phonetic component to show the pronunciation, or just simply draw a square " [] " to indicate an unwritten syllable (even Rao *et al* use these methods in their "Dictionary of Cantonese Colloquialisms" (1982)). At least for this reason, they are not entries in dictionaries (except Rao *et al* 1982 which give phonemic transcriptions) and it is very difficult to trace their etymology. Without enough evidence, it is very difficult to draw a clear line between a disyllabic morpheme and a root word combined with a rhythmic suffix, or a compound. Whatever they are, their reduplications in onset, or in rhyme or in both rhyme and tone constitute a disyllabic rhyme-coupling word. Note that almost all these rhyme coupling words are adjectives and related to a derogatory meaning. As in the examples above, none of the syllables can be singled out to be a morpheme, they are unlikely to fall outside the category of disyllabic morphemes. Cantonese words are basically monosyllabic, the very fact of coupling in disyllabic words has a prominence effect. The most common rhythmic pattern of the rhyme couplings is staccato (as illustrated in i below), or, lengthening the first syllable (as illustrated in iii below). The couplings can also be reduplicated, such as /la:t^{6} la:t^{6} t^{h}a:t^{3} t^{h}a:t^{3}/ (very dirty) - derived from /la:t^{6} t^{h}a:t^{3}/ (dirty).

The rhythmic patterns are the same as the XXYY tetrasyllabic adjective reduplications discussed above. The common rhythmic patterns are:

Rhyme coupling in disyllabics is a special type of prolonging, as is rhyme coupling in trisyllabics and tetrasyllabics. The form of coupling in trisyllabics and tetrasyllabics differs from that in disyllabics. In trisyllabics and tetrasyllabics, there are only one or two root words and the rest are rhythmic suffixes. The rhyme coupling is the reduplication of the rhyme in the rhythmic suffixes, not the reduplication of the rhyme in the root word. The suffixes are only added to prolong the root word. For example, in the tetrasyllabics /wu:^{1} li:^{1} ta:n^{1} tow^{1}/ (dirty-rhythmic suffixes) and /wu:$^{1/4}$ li:^{1} ma:^{5} tɕ^{h}a:^{5}/ (dirty-rhythmic suffixes), the first syllables are root words, the second syllables are common rhythmic infixes, the last two syllables are the rhyme coupling suffixes - /ta:n^{1} tow^{1}/ (rhythmic suffixes) duplicates the onset and the tone whereas /ma:^{5} tɕ^{h}a:^{5}/ (rhythmic suffixes) duplicates the rhyme and the tone.

In the trisyllabics /tɕa:p:^{6} pɐŋ^{1} lɐŋ^{1}/ (miscellaneous-rhythmic suffixes) and /ha:m^{6} pa:ŋ^{6} la:ŋ^{6}/ (all-rhythmic suffixes), the first syllables are root words, the last two syllables are the rhyme coupling suffixes. Their root word usually receives longer duration. The second syllables of the trisyllabics are particularly short, they are compressed into a nasalised /pɐ^{1}/ or /pa:^{6}/ after the particularly lengthened root word and they are almost clustered with the last syllable into a /pl-/ cluster. Again, the very fact of coupling in trisyllabics and in tetrasyllabics has a prominence effect by prolonging a word with added syllables, as in disyllabics. The common rhythmic patterns are:

3.3.2.3. Onomatopoeia - Forms of XY(Z) and XYZZ

Onomatopoeia seems not to be a proper topic in linguistics. It may be simply because the imitation of a sound is just a symbolic form illustrating the meaning more immediately than ordinary speech-forms. It seems that anyone can invent an imitative sound as he likes and thus no special linguistic value exists. However, Cantonese has numerous onomatopoeic expressions which include words of different length, i.e., monosyllabics, disyllabics and tetrasyllabics, and they include nouns, verbs, adverbs and adjectives. All of them are monomorphemic. They are adopted into the phonological system and into the vocabulary of the language, and are vivid in daily conversation. They can be seen as a variety of reduplication and coupling since they also achieve prominence by means of expanding the number of syllables of a word.

Take an example from monosyllabic onomatopoeia /pa:ŋ^{4} / (to bang); by coupling the onset and rhyme, a disyllabic-long onomatopoeia is formed: /pa:ŋ^{4} pa:ŋ^{2}/ (a gun). Actually, the most common form of disyllabic onomatopoeia involves the addition to the rhyme-coupled syllable of the term /ɕɛ:ŋ^{1}/ (sound), for example: /pa:ŋ^{4} pa:ŋ^{2} ɕɛ:ŋ^{1}/ (bang or gun fire); or by coupling the onset and the coda of the first two syllables together with duplicating the monosyllable at the last two syllables, a tetrasyllabic-long onomatopoeia is formed: /pe:ŋ^{4} pe:ŋ^{1} pa:ŋ^{4} pa:ŋ^{4} / (bang or gun fire). Trisyllabic onomatopoeia is rare, as in /hi:^{1} ha:^{1} ha:^{1}/ (laugh) and I am not going to discuss it here. The common rhythmic patterns in tetrasyllabic and disyllabic onomatopoeia are:

	XYZZ	XY(Z)
i.	— — — —	— — —
ii.	⌐¬ — ⌐ — ¬	— — —
iii.	— — — —	

All tetrasyllabic examples of onomatopoeia are monomorphemic. This seems to be a good place to consider assigning what the typical rhythmic patterns of a tetrasyllabic monomorphemic word are, because the internal morpheme boundaries do not have to be taken into account. We can conclude that in a tetrasyllabic word, the most common pattern

is to vary the length of the first and second syllables so that one is twice as long as the other. Disyllabic examples of onomatopoeia are also monomorphemic and usually followed by another morpheme to form a subordinate compound. The most common pattern in disyllabic onomatopoeia is also to vary the length of the first and second syllable so that one is twice as long as the other.

3.3.2.4. Compounds

3.3.2.4.1. Simple Compounds

The majority of simple compounds are disyllables. By looking at the dimensions of syntactic structures and syntactic functions of compounds in Cantonese, we can classify compounds into several main types as listed below:

SYNTACTIC STRUCTURES	SYNTACTIC FUNCTIONS			
	Nouns	*Verbs*	*Adjectives*	*Adverbs*
Coordinate Compounds	/tɕɐj^{2} nɐy:^{2}/ son-daughter (children)	/ɕɐt^{1} pa:j^{6}/ lose-lose (lose)	/ti:w^{1} ma:n^{4}/ tricky-unreasoning (unreasonable)	/ma:n^{6} jɐt^{1}/ ten thousand-one (in case)
Subject-Predicate Compounds	/tej^{6} tɕɐn^{3}/ earth-shake (earth quake)	/low^{6} kɔ:^{3}/ road-pass (pass)	/ta:m^{2} ta:j^{6}/ gall bladder-big (audacious)	/ɕɐt^{6} pit^{1}/ circumstance-must (unavoidably)
Subordinate Compounds	/la:j^{1} nɐy:^{2}/ last-daughter (youngest daughter)	/lej^{6} joŋ^{6}/ advantage-use (make use of)	/t^{h}i:n^{1} tɕɐn^{1}/ sky-true (innocent)	/tɕɐw^{6} hɐj^{6}/ at once-is (precisely)
Verb-Object Compounds	/ti:m^{2} ɕɐm^{1}/ touch on-centre (dim sum)	/teŋ^{2} kɛ:ŋ^{2}/ punch-neck (argue defiantly)	/fa:n^{2} k^{w}ɐt^{1}/ overturn-bone (unscrupulous)	/tow^{3} tɐj^{2}/ reach-bottom (after all)
Verb-Complement Compounds		/ha:ŋ^{4} hɔ:j^{1}/ walk-far (walk away)		

The dominant feature of the rhythm in Cantonese is the free alteration of the length of the syllables of words in utterances. There is no particular rhythmic pattern for disyllabic compounds. Either syllable in the compound can alter its length freely. And this is the characteristic feature of disyllabic compounds for Cantonese. Lengthening any syllable in a

compound does not change its lexical meaning. Generally, no obvious difference in rhythmic patterns can be found between verbs, adjectives and adverbs.

3.2.4.2. Complex Compounds

Syntactic structure does not affect the rhythmic pattern in simple compounds, nor in complex compounds. The most common complex compounds are tetrasyllables. Trisyllables are much less common. There is no obvious difference in rhythmic pattern between a subordinate complex like /ly:n^{4} hɐp^{6} k^{w}ɔ:k^{3}/ (unite-join-nation - United Nations) which contains a simple coordinate compound /ly:n^{4} hɐp^{6}/ (united), and a verb-object complex like /tɕ^{h}ɛ:^{1} ta:j^{6} p^{h}a:w^{3}/ (to wheel-big-cannon - to tell a lie) which contains a simple subordinate compound /ta:j^{6} p^{h}a:w^{3}/ (big-cannon); nor between a coordinate complex like /jɐn^{4} ɕɐj^{3} k^{w}ɐj^{2} ta:j^{6}/ (person-small-ghost-big - young but smart) consisting of two simple subject-predicate compounds /jɐn^{4} ɕɐj^{3}/ (young) and /k^{w}ɐj^{2} ta:j^{6}/ (smart), and a subject-predicate complex like /ɕɐn^{4} keŋ^{1} ɕɵy:^{1} jœ:k^{6}/ (mind-longitude-decay-weak - neurasthenia) which consists of a simple subordinate compound /ɕɐn^{4} keŋ^{1}/ (nerve) and a simple coordinate compound /ɕɵy:^{1} jœ:k^{6}/ (decay). The alteration of syllable length is important in complex compounds. The possible rhythmic patterns are (not favouring one or another):

There is a possible pause following the right end of the first simple compound, or the simple word if it is not a simple compound. Chao calls this kind of possible pause the `open juncture' (1947: 36). Patterns (ii) and (iii) in the tetrasyllables show that the two consecutive compounds can easily break up into two feet by lengthening one of the syllables, as we suggested in sub-section 3.3.2.1.2.2. If the possible pause is realised, say, because a speaker hesitates, the lengthening of the preceding syllable can be much longer than any syllable lengthening within the normal rhythmic patterns.

If a complex compound is longer than four syllables, the greatly preferred rhythm is to have groups of minimal complex compounds, further divided into groups of simple compounds – groups of two. In a big group, syllables tend to further break into small groups and to alter their length in complementing for each other. I give an example below, a complex compound consisting of one simple compound followed by three complex compounds. (The boundaries between complex compounds or between a complex compound and a simple compound are indicated by '|' , while '-' indicates boundaries between simple compounds or between a simple compound and a simple word.) The most common rhythmic pattern is:

| jeŋ^{1} k^{w}ɔ:k^{3} | lən^{4} tən^{1} - ta:j^{6} hɔ:k^{3} | a:^{3} ʧɐw^{1} - fej^{1} ʧɐw^{1} | ji:n^{4} kɐw^{3} - jy:n^{2} |

(English-nation-London-big-study-Asia-continental-Africa-continental-study-investigate-institute)

(SOAS, University of London, UK)

|XX| XX-XX| XX-XX| XX-X|
|____ __| ____ __ -____ __| ____ __- ____ __| ____ __- ___|

Here we have the /jeŋ^{1} k^{w}ɔ:k^{3}/ (UK) modifying the rest of the construction, in which the /ji:n^{4} kɐw^{3} - jy:n^{2}/ (research institute), a subordinate complex including a compound of coordinates, is the head, modified by a coordinate complex /a:^{3} ʧɐw^{1} - fej^{1} ʧɐw^{1}/ (Asian and African) in subordination; the subordinate complex group is further modified by a subordinate complex /lən^{4} tən^{1} - ta:j^{6} hɔ:k^{3} / (University of London). Any syllable can be lengthened to constitute a picture of free variation in syllable lengthening. For example, if there are two syllables in a group, as in /jeŋ^{1} k^{w}ɔ:k^{3}/ (UK), the possible rhythmic patterns are: ____ __, __ ____. The rhythmic patterns of the group /lən^{4} tən^{1} - ta:j^{6} hɔ:k^{3}/ (University of London) are: ____ __ ____ __, __ ____ __ ____ , ____ __ __ ____, or __ ____ ____ __.

3.3.2.5. Other Rhythms

Here I am going to give an account of the rhythm of some common expressions which might not be lexical items nor belong to the categories above. Some are phrases.

As discussed above, the characteristic feature of rhythm in Cantonese is the alteration of the length of syllables in a rhythmic foot. This also applies to phrases. For example, in an adverb-adjective combination like /how^{2} how^{2}/ (very-good), the rhythmic pattern is to lengthen either syllable. The most common pattern is to lengthen the adverb: ____ __, unless the adjective is under focus. If the adjective consists of two syllables as in /how^{2} t^{h}i:n^{1} tɕɐn^{1}/ (very-sky-true - very innocent), the most common pattern is ____ __ __.

The noun is somewhat lengthened when following the prefix /a:^{3}/ and makes up an iambic template, according to Yip (1992: 28). She adopts examples from Whitaker (1955, 1956) and says that (with some exceptions) the prefix /a:^{3}/ is added only to monosyllabic surnames, family relations, names based on birth order and nicknames, all of which have high pitch. She claims that the initial syllable is correspondingly less stressed and this is almost the only situation in Cantonese where stress can be detected. I agree that the most common rhythmic pattern for the structure of the prefix plus a monosyllabic name is iambic (short-long beat), but I do not think the iambic pattern is a phonological issue. Other alternative rhythmic patterns of the prefix /a:^{3}/ plus name cannot be ruled out. I think her claim about stress on the second syllable is questionable. A syllable standing at the end of an utterance can be lengthened in any event. If we look at the example of /a:^{3} ma:^{4}/ (paternal grandma) or /a:^{3} p^{h}ɔ:^{4}/ (maternal grandma) in which the noun does not have high pitch, the second syllable may sound less 'stressed' than that in /a:^{3} pa:^{1}/ (daddy) or /a:^{3} ma:^{1}/ (mummy).

In a structure of verb/adjective-negation-verb/adjective like /how^{2} m^{4} how^{2}/ (good-not-good - good or not), the most common rhythmic pattern is to shorten the syllabic /m^{4}/: ____ __ ____. /m^{4}/ can also segmentally coalesce with the following syllable to a /mow^{2}/. Barnett also observes the phenomenon of the coalescence of /m^{4}/ (1945-50: 742). /m^{4}/ appears normally to be shorten when it is preceded and followed by other syllables, such as in the expression of /ka:w^{2} m^{4} ti:m^{6}/ (work-not-ok - cannot work out), particularly if the first syllable

is lengthened. In contrast, when /m^4/ is at the beginning of an expression like /m^4 hɐj^6 kej^2 ti:m^6/ (not-be-quite-ok - not quite ok) or /m^4 ti:m^6/ (not-ok), /m^4/ appears normally to be lengthened. These examples clearly indicate that no particular syllable needs to be lengthened or shortened. The position of the syllable in an expression may impact on the rhythmic pattern, following our argument that syllables tend to alter their length when in a string.

Measures usually precede a noun. Thus, in the combination of a measure-noun or a pronoun-measure-noun, the rhythm varies. For example, the rhythmic patterns for /tɕœ:ŋ^1 t^{h}ɔ:j^2/ (measure-table - the table) can be: __ ____ or ____ __ ; and for /ni:^{1} tɕœ:ŋ^1 t^{h}ɔ:j^2/ (this-measure-table - this table) can be: ____ __ __ or __ ____ __, etc. Any syllable when under focus is lengthened.

Verb-suffixes by no means have to take shorter duration than verbs, e.g., /ɕek^6 tɕɔ:^{2}/ (eat-past tense suffix - ate), the rhythmic patterns can be: __ ____ or ____ __. If the suffix is in the middle of an expression, it can be very short. The rhythmic pattern can be, e.g., ____ __ __ in the example of /ɕek^6 tɕɔ:^{2} fa:n^6/ (eat-past tense suffix-rice - had a meal), if the verb is under focus and hence lengthened. If the object of the verb is under focus and hence lengthened, the pattern can be __ __ ____. Of course, the suffix can itself be lengthened, and thus the pattern can be __ ____ __. The shortening of a syllable in the middle of a phrase when the syllable is not under focus is common, complementing the lengthening of a neighbouring word to realise two full beats. For example, in the expression of /ka:w^2 tɐk^1 ti:m^6/ (work-able-ok - do well), the most common pattern is ____ __ ____. The majority of expressions in Cantonese are a combination of two syllables involving a rhythm of long-short (or short-long). In a three-syllable expression, shortening one of the syllables is one way to break the rhythm. By lengthening the first syllable and shortening the second syllable, a long-short-short or long-short-long pattern results. It may be for this reason that Barnett

decides that there is a predominant tendency to stress the first and the third syllables in trisyllables.

3.3.2.6. Summary

The central feature of prominence in Cantonese is the alteration of the length of syllables in a string, which can be a lexical item or a phrase. Lengthening or shortening a syllable within a word may be dictated by syntactic factors (e.g., the position of the syllable in a compound, a reduplication, a polysyllabic monomorpheme or a phrase) and by euphony (e.g., in complementing each other); no regular patterns of rhythm can be found. Some words are lengthened by expanding the numbers of syllables and then interplaying the alteration of length of the syllables within the words, as with the reduplications and couplings, in order to achieve prominence. A syllable when under focus tends to be lengthened, otherwise staccato.

3.3.3. Pitch

There is much less scope for using pitch to achieve prominence in Cantonese as opposed to using duration. The reason is simple: pitch already has heavy functional load - for signalling lexical tone and utterance intonation - so that there is not much room for variation. Unlike Mandarin tones which only exploit pitch contour or most tone languages in Africa which only exploit pitch height, the Cantonese tone system has six contrastive tones which employ both pitch height and pitch contour. What little room is left for prominence manipulation in pitch arises from the wider frequency range which individual tones can exhibit when their syllable is under focus (cf. Fok 1974: 28). When the frequency range is expanded, the syllable is also lengthened (Fok, p.35). However, the 'wider frequency range' is actually very narrow: T1 and T2 can rise a little while T4 can fall a little, but the other three in the middle, T3, T5 and T6, have nowhere to stretch (see discussion in chapter two above and figures 31, 33 in Fok). Syllables with higher pitch are generally heard as louder or more prominent (e.g., Hashimoto 1972: 101, Cheung 1987: 187, Lehiste 1993: 238). The low falling tone has a louder end, according to Cheung (1986: 187). Under these phonetic and phonological conditions, syllables in Cantonese achieving prominence by using pitch do so by using the

phonologically higher pitch tones (i.e., T1 and T2) and using the contrast of higher pitch tones and the phonologically low pitch tone (i.e., T4), as outlined below.

Abercrombie suggested that "[t]he rhythm of everyday speech is the foundation of verse, in most languages" (1967: 98). This is also true in Cantonese in reciting poems in that syllables are heard to be stretched and compressed in terms of duration. However, this is not visible in the written form of poems as every individual character occupies the same space and none has fixed longer/shorter duration than others. In contrast, using T1 and T4 to bring syllables into prominence can be seen in the scheme of classical poems. Each kind of classical poetry conforms to a definite pattern: the particular number of lines contained in a poem and the particular number of syllables contained in a line, with a strict tonal pattern and a fixed rhyme scheme. In classical phonology, rhyme includes vowel, coda and tone of a syllable; in other words, tone has to be considered when it comes to selecting rhyme words. There are nine tones in classical phonology (i.e., the six tones plus the three plosive-coda tones). Only T4 and the non-plosive-coda T1 are grouped as /p^{h}eŋ^{4} ɕeŋ^{1}/ (level-tones), the other seven are /tɕɐk^{1} ɕeŋ^{1}/ (non level-tones). Every single syllable in a poem is strictly fixed as to whether /p^{h}eŋ^{4} ɕeŋ^{1}/ or /tɕɐk^{1} ɕeŋ^{1}/ should be used. For example, the fixed tonal pattern in the first line of a poem containing four lines of seven syllables is: /p^{h}eŋ^{4}-p^{h}eŋ^{4}-tɕɐk^{1}-tɕɐk^{1}-p^{h}eŋ^{4}-p^{h}eŋ^{4}-tɕɐk^{1}/, e.g., the first line in a poem by Li Bai (A.D. 701 - 762) with the title "Sang under the moon in the West Chamber in Jinling city" in present Cantonese: /kɐm^{1} leŋ^{4} jɛ:^{6} tɕek^{6} lœ:ŋ^{4} foŋ^{1} fa:t^{3}/ (Jin-Ling-night-silent-cool-wind-start). Only T1 and T4 are used to contrast with other tones in the strict rhyme scheme, clearly indicating that the extreme pitch (i.e., extreme high and extreme low) together with the relatively long duration (i.e., distinct from the plosive-coda tones which may share the same pitch) produce prominence.

We will now discuss the use of T1, T2 and T4 in achieving prominence in reduplications and onomatopoeia and the change-tones.

3.3.3.1. Onomatopoeia – Forms of XY(Z) and XYZZ

Onomatopoeia of the disyllabics and tetrasyllablcs not only uses the expansion of the length of a word but also uses the extreme pitch tones in their constituents to realise prominence. There are only two tones used out of six as tonal constituents in tetrasyllabics - T1 and T4, the two extreme pitch levels. Some of the tonal forms are either all T1s or all T4s as T1-T1-T1-T1 and T4-T4-T4-T4, but the majority combine the two in the form: T4-T1-T4-T4. In order to maximize the highest pitch of T1 and the lowest pitch of T4, only /i:/ or /e/, the highest possible vowels for the rhyme, is chosen for the first two syllables and /a:/ is chosen for the last two syllables (in some exceptional cases other vowels are used as substitutes for /a:/). For example:

Tonal Constituents of tetrasyllabic onomatopoeia:

T4-T1-T4-T4 /pe:ŋ^{4} pe:ŋ^{1} pa:ŋ^{4} pa:ŋ^{4}/ (sound of banging/gun firing - bang or gun fire)
T4-T1-T4-T4 /ji:^{4} ji:^{1} ja:^{4} ja:^{4}/ (sound of groaning - groan)

Three tones, T1, T4 and T2, are used as tonal constituents in disyllabic onomatopoeia. There are two main tonal forms - one leaves T1 on its own: T1-T1, the other pairs up T4 and T2: T4-T2. (The pairing T4-T4 occurs, but is very rare.) For example:

Tonal Constituents of disyllabic onomatopoeia:

T1-T1 /lɔ:k^{1} lɔ:k^{1}/ (/ɕɛ:ŋ^{1}/) (sound of speaking English-sound - speaking English fluently)
T4-T2 /wɛ:^{4} wɛ:^{2}/ (/ɕɛ:ŋ^{1}/) (sound of crying-sound - crying heavily)

T2 is only used as the last constituent of the disyllables. It is preceded by a low offset T4 and followed by a high onset T1. In such a tonal environment, coupled with the most common rhythmic pattern (__ ____ __, as discussed above), its extreme high offset can therefore be maximized (see discussions in chapter two) in achieving prominence.

3.3.3.2. Reduplications – Forms of (X)YY, XX and XX(Y)

Making use of the high pitch tones and the contrast of high and low pitch tones to achieve prominence in reduplications is restricted to the duplicated rhythms and two forms of

disyllabic adjective reduplications. Those forms of reduplications are listed in 3.3.2.1.1. and 3.3.2.1.2. above. In the duplicated rhythmic suffixes following an adjective, the composition of rhythmic suffixes is either T1-T1, T4-T4 or T1-T4 (with very few exceptions such as T1-T3 in /hoŋ^{4} tɔ:ŋ^{1} tɔ:ŋ^{3}/ (red-rhythmic suffixes)):

Tonal Constituents of the Rhythmic Suffixes (X)YY:

T1-T1	(/tɕ^{h}ɛ:ŋ^{1}/) /pi:w^{1} pi:w^{1}/	(green-rhythmic suffixes)
T4-T4	(/wɔ:ŋ^{4}/) /k^{h}ɐm^{4} k^{h}ɐm^{4}/	(yellow-rhythmic suffixes)

If the rhythmic suffixes are morphemes, they may change to a high pitch tone, as in /ɕɐw^{3} ma:ŋ^{2} ma:ŋ^{2}/ (thin-grass hopper-grass hopper - very thin) and /pa:k^{6} moŋ^{1} moŋ^{1}/ (white-dim-dim - hazy white). In the first example, /ma:ŋ^{2}/ is a /pi:n^{3} jɐm^{1}/ (changed tone), its original tone is T5; in the second example, the original tone of /moŋ^{1}/ is T4, it changes to T1. Some morphemic rhythmic suffixes retain their original tones, as /p^{h}ɐn^{3} p^{h}ɐn^{3}/ in /hœ:ŋ^{1} p^{h}ɐn^{3} p^{h}ɐn^{3} / (fragrant-spout-spout - fragrant).

In the form of XX and XX(Y) disyllabic duplicated adjectives, the first constituent of the disyllables of XX and the second constituent of the disyllables of XX(Y) obligatorily use T2 to realise prominence in combination with prolonged duration and greater force of articulation. T1 does not change to a T2, as it is a high tone itself. The extreme high offset of T2 can be maximized in the course of lengthening before it comes to meet any low onset tone of the following syllable. Of the three means (i.e., pitch, duration and loudness) of achieving prominence in this syllable, only higher pitch is phonologically compulsory; longer duration maximizes the high offset of T2 (for the high pitch of T1, if it is a T1); the degree of loudness is optional. Loudness becomes cumulative, after the other two means of achieving prominence are totally exhausted.

Tonal Constituents of duplicated adjectives for forms of XX and XX(Y):

XX	T2-X	/hoŋ^{2} hoŋ^{4}/	(red-red)
XX(Y)	X-T2	/hoŋ^{4} hoŋ^{2}/ (/tej^{2}/)	(red-red-a bit)

3.3.3.3. /pi:n^{3} jɐm^{1}/ (Changed-tones)

The cause of the phenomenon of /pi:n^{3} jɐm^{1}/ (Changed-tones) in Cantonese is still a puzzle. Since Chao (1947: 34) pointed out that /pi:n^{3} jɐm^{1}/ have a morphological meaning, "namely, that familiar thing (or person, less frequently action) one often speaks of", almost all discussions of /pi:n^{3} jɐm^{1}/ are morphologically oriented (e.g., Whitaker 1955-6, Hashimoto 1972: 93, Rao *et al* 1981: 275, Wong 1981 and 1982, Cheung 1986: 41, Bauer and Benedict 1997: 142). Chao considers that /pi:n^{3} jɐm^{1}/ are really very similar to the retroflex suffix /-r/ in Mandarin. Hashimoto claims that morphemes with /pi:n^{3} jɐm^{1}/ are colloquial, the same words used in literary combinations never carry /pi:n^{3} jɐm^{1}/. Moreover, the conditioning for the /pi:n^{3} jɐm^{1}/ is not predictable; in the lexical items of Cantonese, all morphemes which are subject to /pi:n^{3} jɐm^{1}/ have to be specifically marked. Thus, /t^{h}ɔ:j^{4}/ (table) is changed to /t^{h}ɔ:j^{2}/ while /t^{h}ɔ:j^{4}/ (terrace) remains unchanged; /nɐy^{2}/ (daughter) is derived from /nɐy^{5}/ (female), /a^{3} wɔ:ŋ^{2}/ (prefix-Wong: Wong (a surname)) is used rather than /a^{3} wɔ:ŋ^{4}/ , etc.

Certainly /pi:n^{3} jɐm^{1}/ have a special function in the morphology, but I think the process of /pi:n^{3} jɐm^{1}/ is a phenomenon of achieving prominence on the grounds that only the high pitch tones are allowed to be used. The high pitch tones bring prominence to the syllable, related to the familiarity that one wants to highlight. Hashimoto notes this in passing (see 1972: 182). In a speaker's mind, to lift up the pitch of the T1 or the offset of T2 is to bring prominence to the T1 or T2 syllable, not to mention the raising of the pitch of the other tones to extremely high. Almost all syllables with /pi:n^{3} jɐm^{1}/ are nouns and occur as the last syllable of a word, i.e., at the end of a foot of a lexical item; therefore they gain more time to realise their full tonal excursion. Perhaps this explains why in some studies, the /pi:n^{3} jɐm^{1}/ T1* is considered higher than the original T1 (e.g., Bauer and Benedict 1997: 142) and the /pi:n^{3} jɐm^{1}/ T2* is considered higher than the original T2 (e.g., Whitaker 1955-6: 188, who

describes the height of the /pi:n^{3} jɐm^{1}/ as sailing off like a kite); T1* and T2* are considered as separate tonal categories in their studies. For example, the T2* in /low^{5} tɕœ:ŋ^{2}*/ (old-Cheung: Cheung) is claimed to be higher than the original T2 in /tɕœ:ŋ^{2}/ (Cheung); and the T1* in /low^{5} tɕœ:ŋ^{1}*/ (old-Cheung: Cheung) is claimed to be higher than the original T1 in /tɕœ:ŋ^{1}/ (Cheung).

As I mentioned in chapter one, no experimental evidence provides support for the claim that /pi:n^{3} jɐm^{1}/ have different tonal value, therefore I do not treat the /pi:n^{3} jɐm^{1}/ as independent tones merely on account of their special morphological meanings. They have special phonological status because they bring such special morphological meanings into prominence through the phonetically most prominent tones - T1 and T2. T2 is used the most. Apart from being a noun and occurring at the end of a word most of the time, /pi:n^{3} jɐm^{1}/ can be found rarely in a monosyllabic adverb/adjective which is positioned at the beginning of an utterance. /fa:j^{3}/ (quick) in /fa:j^{2} ti:^{1}/ (quick-a bit – hurry up) is the only example I can think of. The high-pitch T2 helps to bring prominence to the adverb/adjective. Apart from T1 and T2, T4 in /pi:n^{3} jɐm^{1}/ is also found, but its use is strictly restricted to the combination with T1 for contrasting the two extreme pitch tones, as in /ma:^{4} ma:^{1}/ (mother) and /tɕɛ:^{4} tɕɛ:^{1}/ (sister). /ma:^{4}/ is derived from /ma:^{1}/ (mother), whereas /tɕɛ:^{4}/ and /tɕɛ:^{1}/ are both derived from /tɕɛ:^{2}/ (sister).

3.3.3.4. Summary

Pitch achieves prominence in Cantonese by using the phonologically higher pitch tones and by the contrast between the high pitch tones and the low pitch tone. This can be found in the tonal constituents of onomatopoeia, reduplications and /pi:n^{3} jɐm^{1}/. Since pitch is used in lexical tones and utterance intonation, there is not much scope for phonetic variations in achieving prominence at word level.

CHAPTER FOUR

DESCRIPTION OF INTONATION IN CANTONESE

4.1. INTRODUCTION

4.1.1. Defining Intonation

Intonation, like tone, involves using pitch to convey meaning but at utterance level rather than at word level. Intonation is speech melody. As well as linguistic factors (e.g., grammatical meaning), there is no doubt that paralinguistic factors such as attitudinal states also have a great impact on intonation, as speakers have one kind of emotional state or another and listeners are very sensitive to what kind of mood the speakers convey. With a single word in an utterance, an angry emotion tends to be expressed in a loud voice with higher pitch and shorter duration by Cantonese speakers, according to Fok's experiments (1974: 34). But I observe that an angry emotion can also be expressed in a deep creaky voice by a calm speaker. It is generally agreed that the variation in tone of voice signalling a particular attitudinal state for paralinguistic communication is relatively unsystematic (Crystal 1985: 220, Laver 1994: 22, Ladd 1996: 37). In actual speech in a tone language, the two systems of pitch fluctuation - tone and intonation - occur simultaneously as no one can speak without intonation or dispense with the lexical tones. The output of the co-occurrence of tone and intonation can be quite complex: if the lexical tones are lost, the lexical meanings are also lost; if the intonation is distorted, the meaning of the utterance can be lost. If one or another is not articulated accurately, the speaker does not sound right to the natives. How do the relatively fixed-pitch phonological tones cooperate with the utterance intonation when one speaks? Or, how does a tone language speaker use intonation when the utterance consists of words with relatively fixed pitch? A popular answer to these questions is: intonation is superimposed on tone (e.g., for Chengdu Sichuanese, Chang (1958: 70); for Thai, Luksaneeyanawin (1983: 171); for Mandarin, Shen (1990: 89)).

Unlike a non-tone and a stress-using language, say, English, which changes the pitch movements of stressed syllables (and unstressed syllables) to a great extent in manipulating intonation, Cantonese as a tone language employs both pitch heights and pitch orientation at the lexical level in the first place. There is not much freedom left for changing pitch movement on top of lexical tone orientation in Cantonese. This indicates very clearly that intonation manipulation in Cantonese cannot be expected to be the same as in a non-tone language like English. Consider an example from Pierrehumbert (1980: 7-8 and 255-256). She takes five different melodic enunciations of the same stress pattern word `Anna' (the first syllable is stressed and the second is unstressed) as spoken by a native speaker of English, each melody producing its own expressive force.

	Patterns of Expressive Force	Melodic Patterns
a.	`Anna' was used as the answer to a question.	H* L¯L%
b.	`Anna' was used as the answer to a question, but carried an implication that the answer is incomplete.	H* L¯H%
c.	The pattern was used for calling out to Anna.	H*+ L¯H ¯L¯%
d.	The pattern was used to convey incredulousness, and to imply that the speaker is giving only one of many possible examples.	L*+ H¯L¯H%
e.	The pattern was used for a question: "Is it Anna?"	L* H¯H%

(Notation note : `H' and `L' mean `high' and `low' tone; the asterisk and superscript hyphen mark associated and unassociated stress, respectively. The plus sign indicates that the tones before and after are bitonal; the percent sign indicates a tone associated with a boundary.)

According to Pierrehumbert, the intonation of utterance (a) is the tonal sequence of H* and L¯ L%. The contour is to be analysed as a string of tones occurring at well-defined points in the utterance. The tonal sequence itself makes up the intonation contour of the utterance (a). However, this kind of abstract description does not seem to apply to Cantonese. The melody of Cantonese heavily relies on the succession of individual syllable tones which make up the utterance. The lexical tonal sequence has a relatively bound pitch height contour before any expressive intonation applies. Therefore, with a natural intonation, a H - L -H pitch contour

can represent a tonal sequence of T1-T4-T1 whereas an inverted pitch contour of L-H-L can represent a tonal sequence of T4-T1-T4. But a pitch contour of L-H-L not only can represent the tonal sequence of T4-T1-T4, but can also represent T4-T3-T4, T4-T6-T4, T6-T3-T6, T6-T1-T6 or T3-T1-T3 because the lexical tones of T1, T3, T4 and T6 differ to a great extent in terms of relative pitch height but have a similar downward slope. The abstract terms of H and L are obviously not adequate to represent even the lexical tones of Cantonese.

Clearly, the description of intonation in Cantonese has to take into account that the lexical tones presumably set out a basic pattern of melody in utterances. Every lexical tone has its relative height and contour shape that cannot be altered significantly as in the example of 'Anna above. The melody pattern of 'Anna can be altered from H*L¯L% to L*H¯H%, but the lexical meaning is preserved. I therefore have to discard the idea of using the abstract concept of intonational tonal sequences of H and L (or H, M, L, or H, HL, LH, L, etc) to represent the intonation contour in Cantonese which is actually based on a succession of lexical tonal contours. In fact, this is a crucial point that any researcher who studies intonation in a tone language cannot afford to ignore. Research on intonation in tone languages has explored different frameworks and different approaches, and has developed in sophistication and depth. Nevertheless, research on Cantonese intonation has to go its own way to a great extent because it has its own rich tonal system and its own distinctive characteristics of prominence.

4.1.2. Intonation in Tone Languages

Chang (1958: 84) suggests that intonation modifies the lexical tones and not that the tones decide the intonation of the sentence. After analysing the actual speech of a speaker, she claims that the intonation of Chengdu Sichuanese consists of three factors: the pitch level, the pitch range and the perturbation of the final syllable. Ordinary statements and questions have a medium range of pitch and have a mid and low pitch level; but they have different perturbation shapes on the final syllable - falling for the former and rising for the latter. She posits two tunes in the language: falling tune and rising tune - solely according to the perturbation of the final syllable. Therefore, ordinary statements have a falling tune whereas

questions have a rising tune. In contrast to ordinary sentences in which syllables are more or less evenly stressed, she finds that emphatic sentences in which one particular word or syllable receives an extra stress have a wider range. This type of sentence has a falling tune and its pitch level is between mid-high and low (p.79). She concludes that it is difficult to make any general statements on the relationship between pitch level and the type of sentence but the range is at least a clue to the emotional state of the speaker: sentences containing a protest or implying dismissal of the topic have a narrow range; emphatic sentences and sentences expressing contempt have a wide range; ordinary statements and questions have a medium range (p.83). According to the perturbation of the final syllable, she finds that the rising tune is used for questions requiring answers other than `yes' or `no', sentences expressing vexation or containing a protest, and unfinished sentences; and the falling tune is used for ordinary and emphatic statements, sentences expressing emphatic approval, awe, contempt or surprise, and sentences implying dismissal of the topic.

Dung *et al* take the stress patterns and sentence-final particles (morphosyntactic markers) into account when analysing the intonation of Hanoi Vietnamese. Four out of the six tones are used in their experiments. They consider that duration and intensity are the most important cues for stress. The tones in unstressed syllables are shorter and there is less time for tonal target values to be reached (1998: 399). They claim that the intonation pattern of interrogative sentences differs from that of declarative sentences in having a rising contour. They show that the rising contour happens at the last syllable, whether or not the final syllable is a morphosyntactic marker. Preceding that rising contour is a slightly falling F_0 line from the peak of the sentence and the falling F_0 line is similar to that in declarative sentences (p.408). Interrogatives and imperatives have a high and declaratives have a mid register, whereas `attitudinal' and `expressive' sentences have an extra-high register. They conclude that different sentence types are differentiated more by their average register than by their overall intonation pattern.

In Mandarin, De Francis (1963: xxx) suggests that the difference in register (i.e., the overall pitch level) rather than in pitch curves is the difference between the two major intonation

patterns, statements and interrogatives. Question intonations are realised at a higher register than statement intonations. Shen (1990: 27) agrees with him and further claims that the starting pitch level of an identical tonal string in question intonation is higher than that in statement intonation. Tao (1996: 34) comments critically that Shen's data only use identical tonal strings and are based on isolated and controlled sentences which are not sufficient to justify this conclusion. He states that non-identical tonal strings predominate in discourse. It is quite common for an utterance beginning with a higher pitch not to be an interrogative utterance (p.35). However, he confirms from his data of natural discourse and agrees that contour alone is not sufficient to differentiate between the two utterance types. He uses the discussion of prosodic properties by Vaissière (1983) to set out the criteria for identifying intonation units in Mandarin: a) pause between the intonation units (p.37), b) declination of the intonation units, c) resetting the baseline for the ensuing intonation units (p.40), d) lengthening at the end of intonation units (p.39). He gives only one example for each criterion, and does not point out whether all the criteria apply to his data. He also gives examples to demonstrate that a continuing intonation can have a falling contour (p.46) whereas a final intonation can have a rising ending (p.45). Ho (1977: 450) demonstrates that interrogative and exclamatory intonations not only have higher Fo, but also have a rising contour in the sentence-final position, while declarative intonation has a distinctly lower F_0 and a falling contour in sentence-final position. Chao proposes that "normal intonation is simply a succession of the tones" (1968: 41). In contrast, Shen says that "[s]entence intonation does not consist of a succession of lexical tones but results from its own pitch movement, which varies in accord with modality and attitude" (1990: 78). Kratochvil, on the other hand, claims that the data of his informant from natural speech has shown that the general properties of intonation in normal Bejing dialect speech appear to be very similar to those of intonation in languages which have been studied in the western linguistic tradition (1998: 430). He demonstrates that the modification of individual tones can be more pronounced when different stress prominence patterns are assigned, and that this can result in a different location of the F_0 peak in an utterance. He concludes that a combination of downward channeling in both F_0 and amplitude, and a terminal fall in F_0 appears to characterise the intonation of declarative sentences.

Luksaneeyanawin posits four contrastive intonation tunes in Thai after an experiment of one-word sentences with two informants: Tune 1- the Falling Tune, Tune 2 - the Rising Tune, Tune 3 - the Lowering Tune and the Tune 4 - the Raising Tune (also named the Convolution Tune) (1983: 322). Thai, a five tone language, "basically is monosyllabic and non-inflectional" (1983: 50). Its "lexemes have a fixed word accent system" (1998: 376). Content monosyllabic words are accented and are realised as stressed syllables in neutral or unmarked speech situations. Grammatical monosyllabic words are unaccented and are usually realised as unstressed syllables except in marked situations where contrastiveness or emphasis is given to the word, resulting in stress (1998: 376). She concludes that "[e]ach phonological tone still keeps its phonetic features distinct from other phonological tones" when under the influence of the intonation tunes (1998: 384). She observes that Tune 2 has the highest pitch (1983: 335). Utterances with Tune 1 and Tune 3 have low pretonics (the syllable before the tonic syllable) whereas utterances with Tune 2 and Tune 4 have high pretonics (p. 334). She decides that a stressed syllable is a foot (1983: 290). A foot is a rhythmical unit of speech from one stressed syllable up to but not including the next stressed syllable (1983: 224). She defines the tonic as "the word which is realised with the phonetic prominence and is considered the phonological prominence" (1983: 290). The tonic is used not only to show the focus of information in contrastive situations but is also used to show the terminal transition (the ending of one information unit). The tonic "is indicated by a prominent stress on the final syllable of the last word (whether content or grammatical) in an information unit". An information unit can be composed of one independent meaningful unit called 'word', or a set of words which are syntactically and semantically coherent and unified. She proposes that in Thai an information unit is a tone group and a tone group has a tonic.

In Hausa, a two tone language of Nigeria, Lindau (1985: 26) proposes that intonation in statements is represented by a downward slope while in yes/no questions it is represented by a zero slope. The rate of slope of the top line and the bottom line is related at least to sentence type, sentence length, and tone pattern (p.37). However, the precise amount of slope appears to be idiosyncratic (p.30). Statements and corresponding yes/no questions

differ significantly in total duration. Questions are about 10% shorter than the corresponding statements and are not characterised by a raised register (p.34). On the other hand, Inkelas and Leben find that the high tones in the second phrase of yes/no questions are raised above the level at which they would appear in a declarative and this phenomenon is termed "global raising" (1990: 18). They agree with Selkirk (1980), Nespor and Vogel (1982) that intonational phrases typically contain a syntactic constituent or constituents which form a coherent semantic unit. They give as an example that there is a phrase boundary between the two objects of double object constructions, and point out that the final syntactic constituent, the direct object, tends to constitute its own intonational phrase. However, the intonational phrases (i.e., prosodic constituents) do not necessarily match syntactic constituents. For example, emphasized words in Hausa always begin a new intonational phrase. "Emphasis can be added to most positions in a Hausa utterance. And if every emphasised word begins a new phrase, it follows that almost any sentence can be parsed differently into intonational phrases depending on which word is emphasised. Yet the syntactic constituent structure does not vary along with the prosodic structure" (p.19).

In discussing intonation in Yoruba, Laniran claims that "tones and phonetic implementation rules are the same as intonation". Her phonetic implementation rules "include rules that assign F_0 increments (e.g., H-Raising) and decrements (e.g., downstep) to the basic F_0 values assigned to the tones" (1993: 2). H-Raising means that a H tone is raised when following a L tone; downstep is a phenomenon by which following H and L tones in a sequence are realised with lower F_0 values than preceding H and L tones (p.8). There are transitional tones whose F_0 values are the intermediate values between two specified adjacent tones (p.73). She concludes that it is possible to have an intonational system without a loss in tonal contrasts (p.209). She does not propose any contrastive tune systems as other researchers on tone languages do, as mentioned above. Instead, she works on the tonal effects from different tonal sequences. She challenges the Lund School Model (Garding 1981, Bruce 1978, etc.) and Pierrehumbert's (1980) and Pierrehumbert and Beckman's model (1986, 1988) and concludes that they cannot give an adequate account of a three tone language like Yoruba. She finds that the register model (Clements 1981) can best accord

with her data – F_0 values for each tone are subsequently calculated taking into account local increments and global shift in register (p.206).

Obviously researchers on intonation in tone languages more or less concentrate on how different types of pitch configuration in utterances represent different types of grammatical sentences (e.g., declaratives and interrogatives) and sum up by recounting how many contrastive tunes there are in the language. It seems quite common that interrogatives tend to have a rising tune and be in a higher register, whereas declaratives tend to have a falling tune and be in a lower register. Both height and direction of the pitch are taken into account in characterising a tune. Tonal contrasts remain when under influence from intonation. In stress-using tonal languages, the tones in the stressed syllables characterise the tunes. The tone in an emphatic/focused stressed syllable is prominent. The tune is realised in the final syllable of an utterance (e.g., in Chengdu Sichuanese, Chang (1958: 77)). A syllable tone when under focus becomes a tonic (e.g., in Thai, Luksaneeyanawin (1983: 294)), or can convert its static/dynamic tonal contour to a high register (e.g., in Vietnamese, Dung *et al* (1998: 414)), or can rise both in F_0 and amplitude (e.g., in Mandarin, Kratochvil (1998: 429)). Most claims are based on invented sentences in reading style (e.g., Ho 1977, Luksaneeyanawin (1983), Shen (1990), Dung *et al* (1998) and Laniran (1993)), while some are based on natural speech (Chang (1958), Tao (1996) and Kratochvil (1998)). The universal function of speech is to communicate. Using invented sentences has the advantage of concentrating on examining some particular aspects of intonation, for example - tonal interaction and patterns of declaratives and interrogatives. However, this method limits the domain of intonation. We know that the overwhelming proportion of daily speech occurs in non-reading style rather than reading style. As I consider that intonation is speech melody, I would like to use natural speech as research material and will discuss the intonation in Cantonese in these areas:

a) The foot;

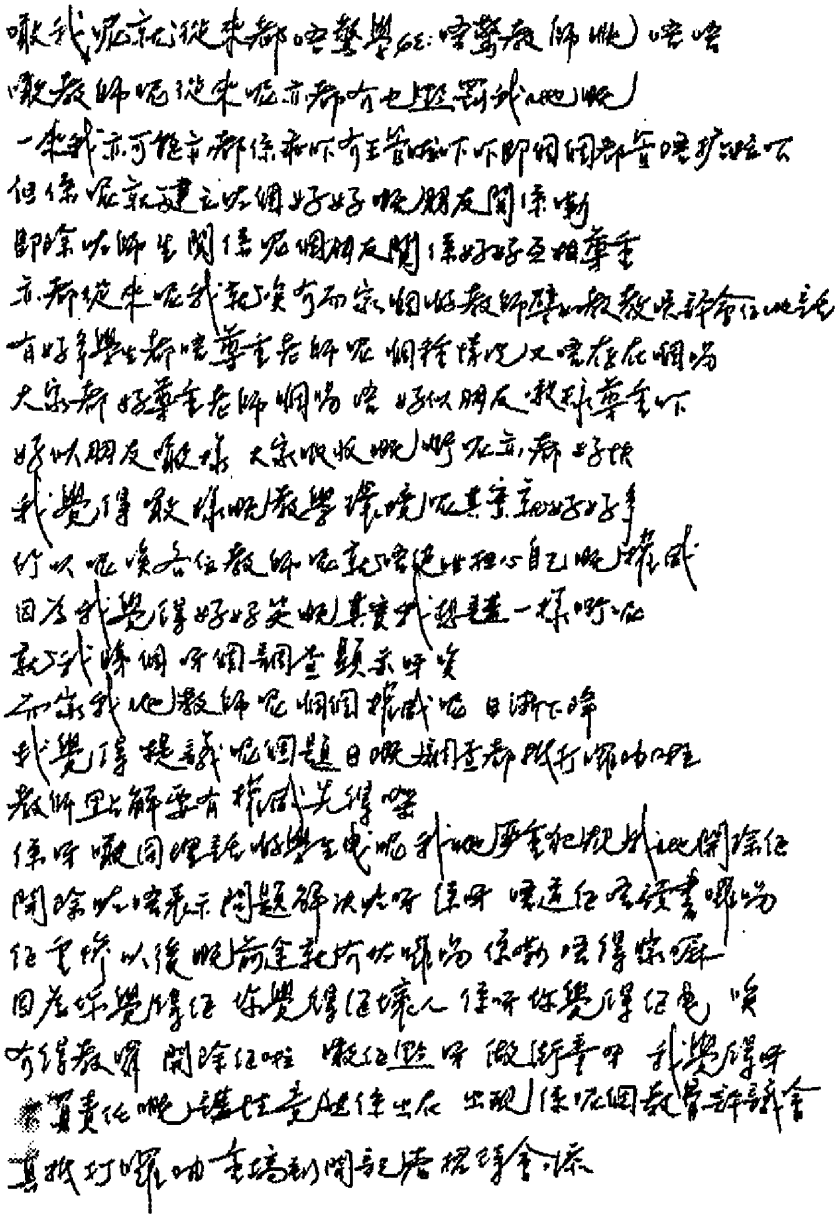
b) The intonation group;

c) The major intonation group.

4.2. THE MATERIAL

Intonation is beyond the segmental and is only present in connected speech. A genuine description of intonation cannot avoid analysing data from natural speech. I therefore recorded a natural conversation from a radio programme in Hong Kong. There are three speakers, all are young professional commentators and have very clear articulation. They broadcast in a studio, the recording conditions sound satisfactory. The conversation was taped from a radio and then the recordings were played into a Speech Workstation (SW) software package which is installed in a 486 Dx66 16M Ram PC computer. The recordings can be displayed as waveforms and spectrograms simultaneously and the F_0 contour can also be derived by the package. All syllables are measured in terms of duration and pitch. The methods of segmentation are conventional. In the case of geminate consonants occurring at syllable boundaries, for example, /-t # t-/, arbitrariness cannot be avoided. In the case of the F_0 contour failing to be derived from a few syllables, hand measures from the tenth harmonic are employed. In the case of overlapping speech, i.e., two people speaking at the same time, the piece has to be excluded from the main acoustical analysis. There are 101 syllables excluded from the main acoustical analysis for this reason but they are used for non-acoustical analysis and they are listed under the numbers of [7], [14-15] and [25-34] in section 4.5.6.1. The content of the conversation concerns education in Hong Kong and the speakers sound not emotionally involved. The speaking rate is normal. In such a short piece of conversation containing 331 syllables, some basic grammatical structures are included: declarative, interrogative, echo questions, etc. I will transcribe the recording in the Chinese script first, then give translations into English, then give phonemic transcriptions for the recording before giving a syntactic structure analysis.

4.2.1. The Material Transcribed in Chinese Script



4.2.2. The English Translation

(Speaker A): I was never frightened of stu- ... never frightened of teachers. (Speaker B): Mm Hmm. (Speaker A): And teachers never really punished us either. On the one hand, it was possibly because I was ... (Speaker B): good ... (Speaker A): ... uncontrollable. That means no one was able to be in control at all. However, we were on very friendly terms. This means, apart from the teacher-student relationship, the friendship was very strong, mutual respect. I never ... unlike those teachers nowadays, for example, the Body of Educational Assessment say that many students don't respect teachers, such a situation did not exist. Everybody respected teachers very much. (Speaker C): Right! The same as friends who respect each other. (Speaker A): The same as friends ... Everybody learned really fast. I think this sort of educational environment was in fact much better. Therefore, teachers shouldn't worry about their authority. Because I think it is silly. Actually I just want to say one thing. I see that investigations have shown that "nowadays the authority of us teachers is on the decline". I think that whoever suggested the topic of this investigation should have his bottom smacked. Why do teachers need to have authority anyway? (Speaker C): That's right! And the students were naughty. They committed a serious breach of the rules. We expelled them. Expelling them does not mean the problem has been solved. (Speaker A): That's right! Do they not study? They ... (Speaker C): Much worse! (Speaker A): Their future would have been doomed! (Speaker C): Right! (Speaker A): You cannot do that! Because you think they are ... (Speaker C): You think they are bad people ... (Speaker A): Right! You think they are naughty, unteachable, then expel them. So, what will they do then? They'll be out on the streets?! I think that is an irresponsible proposal and a very surprising one coming from the Body of Educational Assessment. They all really deserve to have their bottoms smacked – they even held a press conference!

4.2.3 The Phonemic Transcription

The literal meanings of the words in the material are given under the phonemic transcription. A dash in the literal meanings of the words indicates the boundary between the syllables. A line under sequences of two or more syllables in the phonemic transcriptions indicates that the syllables in question together form a word, i.e. a polysyllabic word or compound word. Particle-cluster and polysyllabic-monomorphemic words are also underlined. Monosyllabic words are not underlined.

(Note for abbreviations: p. – particle pl. – plural i. – interjection
m. – measure word including classifier
(2) – two syllables form a monomorphemic word)

/ kɐm^{2} ŋɔ:^{5}nɛ:^{1}tɕɐw^{6} tɕ^{h}oŋ^{4} lɔ:j^{4} tow^{1} m^{4} kɛ:ŋ^{1} hɔ:k^{6} ɕ^{6} ɛ:^{6} m^{4} kɛ:ŋ^{1} ka:w^{3} ɕi:^{1} kɛ:^{3}/
in that case-l- p.- then- from-come- also-not-afraid-study-stu- not-afraid-teach-expert-p.

/m^{4} m^{2} kɐm^{2} ka:w^{3} ɕi:^{1} nɛ:^{1} tɕ^{h}oŋ^{4} lɔ:j^{4} nɛ:^{1} jek^{6} tow^{1} mow^{5} mɐt^{1} /
mm-hmm in that case-teach-expert- p.- from-come- p.- also-also- not- any

/ti:m^{2} fɐt^{6} ŋɔ:^{5} tej^{6} kɛ:^{3} jɐt^{1} lɔ:j^{4} ŋɔ:^{5} jek^{6} hɔ:^{2} nɐŋ^{4} jek^{6} tow^{1} hɐj^{6} k^{w}a:j^{1} ha:^{2} /
how-punish-l- pl.- p. one-come- l- also- possible- can - also-also-am good- i.

/mow^{5} wɔ:ŋ^{2} k^{w}u:n^{2} la:^{1} ha:^{2} a:^{4} tɕek^{1} kɔ:^{3} kɔ:^{3} tow^{1} k^{w}u:n^{2} m^{4} ti:m^{6} la:^{1}ha:^{2}/
have not-king- control- p.- i.- oh- that is- m.- m.- all- control-not- o.k.- p.- i.

/ta:n^{6} hɐj^{6} nɛ:^{1} tɕɐw^{6} ki:n^{3} la:p^{6} tɕɔ:^{2} kɔ:^{3} how^{2} how^{2} kɛ:^{3} p^{h}ɐŋ^{4} jɐw^{5} k^{w}a:n^{1} hɐj^{6} la:k^{3}/
but- is- p.- then-build-establish-ed-m.- very-good- p-friend-friend- concern-relate-p.

/tɕek^{1} tɕ^{h}ej^{4} tɕɔ:^{2} ɕi:^{1} ɕɐŋ^{1} k^{w}a:n^{1} hɐj^{6} nɛ:^{1} kɔ:^{3} p^{h}ɐŋ^{4} jɐw^{5} k^{w}a:n^{1} hɐj^{6} how^{2} how^{2}/
i.e.-apart-p.-teacher-student-concern-relate-p.-m.-friend-friend-concern-relate-very-good

/ wu:^{6} ɕœ:ŋ^{1} tɕy:n^{1} tɕoŋ^{6} jek^{6} tow^{1} tɕ^{h}oŋ^{4} lɔ:j^{4} nɛ:^{1} ŋɔ:^{5} tɕɐw^{6} ŋ^{6} /
mutual-mutual- respect-respect- also-also- from-come- p.- l- then-hm

/mow^{5} ji:^{4}ka:^{1} kɔ:^{2} ti:^{1} ka:w^{3} ɕi:^{1} pej^{3} jy:^{4} ka:w^{3} ka:w^{3} iɛ:^{6}p^{h}eŋ^{4} wu:j^{2} k^{h}ɵy^{5}tej^{6} /
have not-now(2)- that-pl.-teach-expert-example-as-teach-teach-hm-assess-committee-he-pl-

/wa:^{6} jɐw^{5} how^{2} tɔ:^{1} hɔ:k^{6} ɕa:ŋ^{1} tow^{1} m^{4} tɕy:n^{1} tɕoŋ^{6} low^{5} ɕi:^{1} nɛ:^{1} kɔ:^{2} tɕoŋ^{2} /

say-have-very-many-study-pupil- all- not- respect-respect- old-expert- p.- that- m.

/tɕ^{h}eŋ^{4} fɔ:ŋ^{3} jɐw^{6} m^{4} tɕ^{h}yn^{4} tɕ^{h}ɔ:j^{6} kɔ:^{3} wɔ:^{3} /

situation-situation- also- not- exist- exist- p.- p.-

/ ta:j^{6} ka:^{1} tow^{1} how^{2} tɕy:n^{1} tɕoŋ^{6} low^{5} ɕi:^{1} kɔ:^{3} wɔ:^{3}/

big- family- all- very- respect-respect- old-expert- p.- p.

/m^{6} how^{2} tɕ^{h}i:^{5} p^{h}ɐŋ^{4} jɐw^{5} kɐm^{2} jœ:ŋ^{2} tɕy:n^{1} tɕoŋ^{6} ha:^{2} how^{2} tɕ^{h}i:^{5}/

mh- very- like- friend-friend- so- look- respect-respect- i. very- like

/p^{h}ɐŋ^{4} jɐw^{5} kɐm^{2} jœ:ŋ^{2} ta:j^{6} ka:^{1} k^{h}ɐp^{1} ɕɐw^{1} kɛ:^{3} jɛ:^{5} nɛ:^{1} jek^{6} tow^{1} how^{2} fa:j^{3}/

friend-friend- this- look- big-family- suck-receive- p.- things-p.- also- all- very-quick

/ŋɔ:^{5} kɔ:k^{3} tɐk^{1} kɐm^{2} jœ:ŋ^{2} kɛ:^{3} ka:w^{3} hɔ:k^{6} wa:n^{4} keŋ^{2} nɛ:^{1}/

I- feel- able- this- m.- p.- teach-learn- surrounding-environment- p.-

/k^{h}ej^{4} ɕɐt^{6} tɕɐw^{6} how^{2} how^{2} tɔ:^{1}/

that- fact- then- good- very- much

/ɕɔ:^{2} ji:^{5} nɛ:^{1} ŋɛ:^{6} kɔ:k^{3} wɐj^{2} ka:w^{3} ɕi:^{1} nɛ:^{1} tɕɐw^{6} m^{4} ɕɐj^{2} kɐm^{3} ta:m^{1}ɕɐm^{1} tɕi:^{6} kej^{2} kɛ:^{3}/

There-fore-p.-em- every-m.- teach-expert-p.- then- no-need-so- bear-heart-self-self-p.

/k^{h}y:n^{4} wɐj^{1} jɐn^{1} wɐj^{6} ŋɔ:^{5} kɔ:k^{3} tɐk^{1} how^{2} how^{2} ɕi:^{3} kɛ:^{3} k^{h}ej^{4} ɕɐt^{6} ŋɔ:^{5} ɕœ:ŋ^{2} kɔ:ŋ^{2} /

power-might-because-for- I- feel-able- very-good-laugh-p.- that-fact- I- want- say

/jɐt^{1} jœ:ŋ^{6} jɛ:^{5} nɛ:^{1} tɕɐw^{6} ŋɔ:^{5} t^{h}ɐj^{2} kɔ:^{3} a^{3} kɔ:^{3} ti:w^{6} tɕ^{h}a:^{4} hi:n^{2} ɕi:^{6} a^{3} ɛ:^{6}/

one- m.- thing-p- that- I- see- m.-ah- m.- investigate-check-apparent-show-p.-em

/ji:^{4} ka:^{1} ŋɔ:^{5} tej^{6} ka:w^{3} ɕi:^{1} nɛ:^{1} kɔ:^{2} kɔ:^{3} k^{h}y:n^{4} wej^{1} nɛ:^{1} jɐt^{6} tɕi:m^{6} ha:^{6} kɔ:ŋ^{3} /

now(2)- I- pl.- teach-expert- p.- that-m.- power-might-p.- day- gradually-down-sink

/ŋɔ:^{5} kɔ:k^{3} tɐk^{1} t^{h}ɐj^{4} ji:^{5} ni:^{1}kɔ:^{3} t^{h}ɐj^{4} mok^{6} kɛ:^{3} ti:w^{6} tɕ^{h}a:^{4} tow^{1} tɐj^{2} ta:^{2} lɔ:^{1} jɐw^{2} la:^{1}/

I- feel-able-lift-discuss-this-m.- title-topic- p.- check-check-all- deserve-beat-bottom(2)- p.

/ka:w^{3} ɕi:^{1} ti:m^{2} ka:j^{2} ji:w^{3} jɐw^{5} k^{h}y:n^{4} wɐj^{1} ɕi:n^{1} tɐk^{1} ka:^{2} hɐj^{3} a:^{3} /

teach-expert- how-explain- need-have- power-might- until-o.k.- p. right- p.

/ kɐm^{2} t^{h}oŋ^{4}ma:j^{4}wa:^{3} ti:^{1}hɔ:k^{6} ɕa:ŋ^{1} jɐj^{5} nɛ:^{1}ŋɔ:^{5} tej^{6} ji:m^{4} tɕoŋ^{6} fa:n^{6} k^{h}ɐj^{1}/

in that case-and-near-say-pl.-study-pupil-naughty-p.- I- pl.-serious-heavy-offence-regulation

/ŋɔ:^{5} tej^{6} hɔ:j^{1} tɕ^{h}ɵy^{4} k^{h}ɵy^{5} hɔ:j^{1} tɕ^{h}ɵy^{4} tɕɔ:^{2} m^{4} pi:w^{2} ɕi:^{6} mɐn^{6} t^{h}ɐj^{4} ka:j^{2} k^{h}y:t^{3} tɕɔ:^{2} a:^{3}/

l- pl-open-expel-him- open-expel-ed- not-show-show-ask- title- solve-firm- ed- p.

/hɐj^{3} a:^{3} m^{4} t^{h}oŋ^{1} k^{h}ɵy^{5} m^{4} tok^{6} ɕy:^{1} l ɔ:^{3} wɔ:^{3} k^{h}ɵy^{5} tɕoŋ^{6} tɕ^{h}a:m^{2}/

right- p.- not-through- he- not- read-book- p.- p.- he more-tragedy

/ji:^{5} hɐw^{6} kɛ:^{3} tɕ^{h}i:n^{4} tkow^{4} tɕɐw^{6} mow^{5} tɕɔ:^{2} lɔ:^{3} wɔ:^{3} hɐj^{6} la:^{3} /

in- back- p.- front-journey- then- not have- ed- p.- p. right- p.

/m^{4} tɐk^{1} ka:^{1} ma:^{3} jɐn^{1} wɐj^{6} nej^{5} kɔ:k^{3} tɐk^{1} k^{h}ɵy^{5} nej^{5} kɔ:k^{3} tɐk^{1} k^{h}ɵy^{5} wa:j^{6} jɐn^{4}/

not-o.k.- p.- p.- reason-for- you-feel- able- him - you- feel- able- him- bad-person

/hɐj^{6} a:^{3} nej^{5} kɔ:k^{3} tɐk^{1} k^{h}ɵy^{5} jɐj^{5} ŋɐj^{3} mow^{5} tɐk^{1} ka:w^{3} lɔ:^{3} hɔ:j^{1} tɕ^{h}ɵy^{4} k^{h}ɵy^{5} la:^{1} kɐm^{2}/

right- p- you- feel-able- him-naughty- eh- not- able-teach- p.- open-expel-him-p.-in that case

/k^{h}ɵy^{5} ti:m^{2} a:^{3} tɕow^{6} ka:j^{1} t^{h}oŋ^{4} a:^{4} ŋɔ:^{5} kɔ:k^{3} tɐk^{1}a:^{3} pɐt^{1} fu:^{6} tɕa:k^{3} jɐm^{6} kɛ:^{3} kɔ:ŋ^{2}fa:t^{3}/

he- how- p.- to be- street-child- p.- l- feel-able-p.- not- bear-duty-post- p.- talk- way

/keŋ^{2} ji:n^{4} hɐj^{6} tɕ^{h}ɵt^{1}tɕɔ:j^{6} tɕ^{h}ɵt^{1} ji:n^{6} hɐj^{2} ni:^{1}kɔ:^{3} ka:w^{3} jok^{6} p^{h}eŋ^{4} ji:^{5} wu:j^{2}/

unexpectedly-look-is-out-at-out-present-at-this-m.-teach-educate-assess-discuss-committee

/ tɕɐn^{1} tɐj^{2} ta:^{2} lɔ:^{1} jɐw^{2} tɕoŋ^{6} ka:w^{2} tow^{3} hɔ:j^{1} kej^{3} tɕɛ:^{2} tɕi:w^{1} tɔ:j^{6} wu:j^{2} t^{h}i:m^{1}/

really-deserve-beat-bottom(2)-still- do- reach-open-record-person-beckon-treat-meeting-p.

4.2.4 The Syntactic Structure of the Discourse

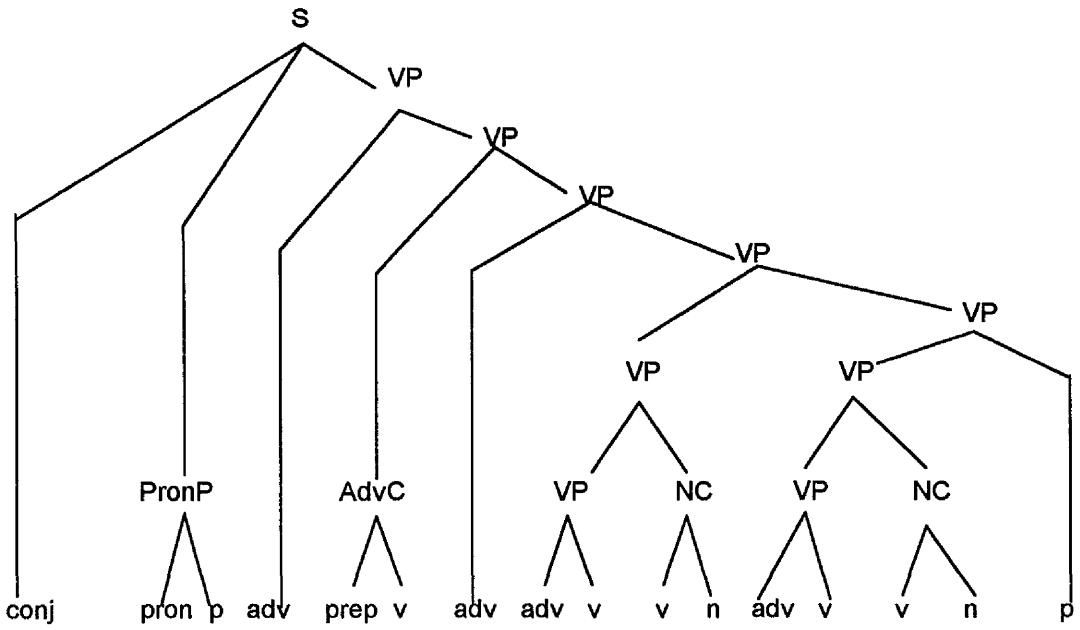
The discourse is analysed in terms of the following phrase structure constituents:

S (sentence), NP (noun phrase), VP (verb phrase), PrepP (preposition phrase), AdjP (adjective phrase), AdvP (adverbial phrase), ConjP (conjunction phrase), LocP (localiser phrase), IntP (interjection phrase), NC (noun compound), PronC (pronoun compound), VC (verb compound), PrepC (preposition compound), AdjC (adjective compound), AdvC (adverbial compound), ConjC (conjunction compound), LocC (localiser compound), PC (particle cluster), n (noun), pron (pronoun), v (verb/auxiliary verb), prep (preposition), adj (adjective), adv (adverb), p (particle), q (quantifier), m (measure), det (determinative), int (interjection), loc (localiser), pl (plural marker).

The syntactic structure of the sentences is shown in the following diagrams. Intonation groups and major intonation groups (for these items, see discussion below) are not determined by syntactic factors but correspond to different sizes of syntactic constituents which form a coherent semantic unit. A small syntactic constituent can be a monosyllabic word. A bigger syntactic constituent can be a phrase, a clause or a short sentence.

(The numbers in square bracket above the trees mark the correspondences with the major intonation groups; the numbers following a double bar under the syllables mark the correspondences with the intonation groups; the single bars under the syllables mark the correspondences with the foot.)

[1]



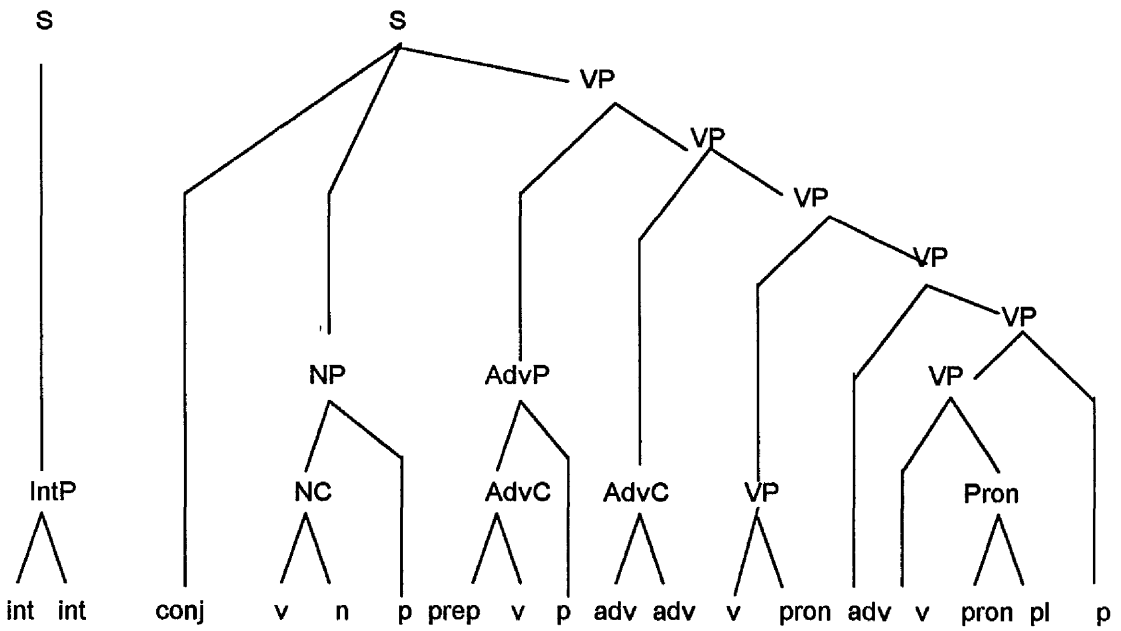
kɐm^{2} ŋɔ:^{5} nɛ:^{1} tɕɐw^{6} tɕ^{h}oŋ^{4} lɔ:j^{4} tow^{1} m^{4} kɛ:ŋ^{1} hɔ:k^{6} ɕ^{6}ɛ:^{6} m^{4} kɛ:ŋ^{1} ka:w^{3} ɕi:^{1} kɛ:^{3}

||1 | || 2 || 3 || 4 ||5 | ||

in that case-I- p.- then- from-come-also-not-afraid-study-stu- not-afraid-teach-expert-p.

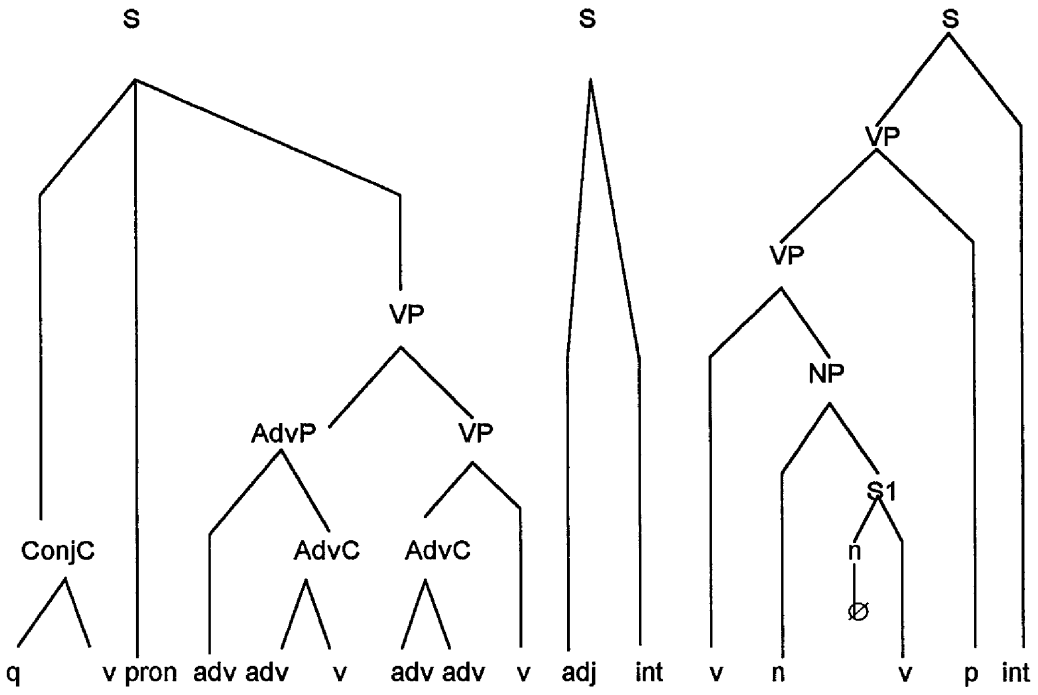
(I was never frightened of stu- ... never frightened of teachers.)

[2] [3]



(Mm Hmm. And teachers never really punished us either.)

[4] [5] [6]



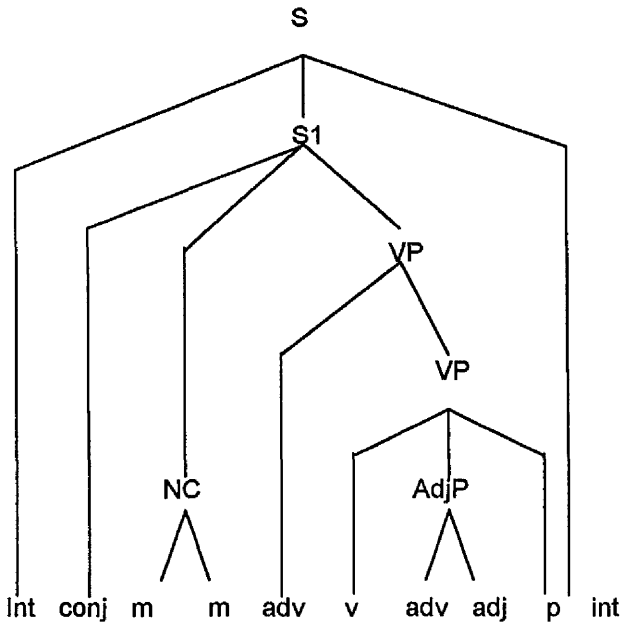
jɐt^{1} lɔ:j^{4} ŋɔ:^{5} jek^{6} hɔ:^{2} nɐŋ^{4} jek^{6} tow^{1} hɐj^{6} k^{w}a:j^{1} ha:^{2} mow^{5} wɔ:ŋ^{2} k^{w}u:n^{2} la:^{1}ha:^{2}

|| 12 || 13 | || 14 || 15 || 16 | | ||

one-come-I-also- can-possible-also-also-am-good- int.-have not-king- control- p.-int

(On the one hand, it was possibly because I was good, uncontrollable.)

[7]



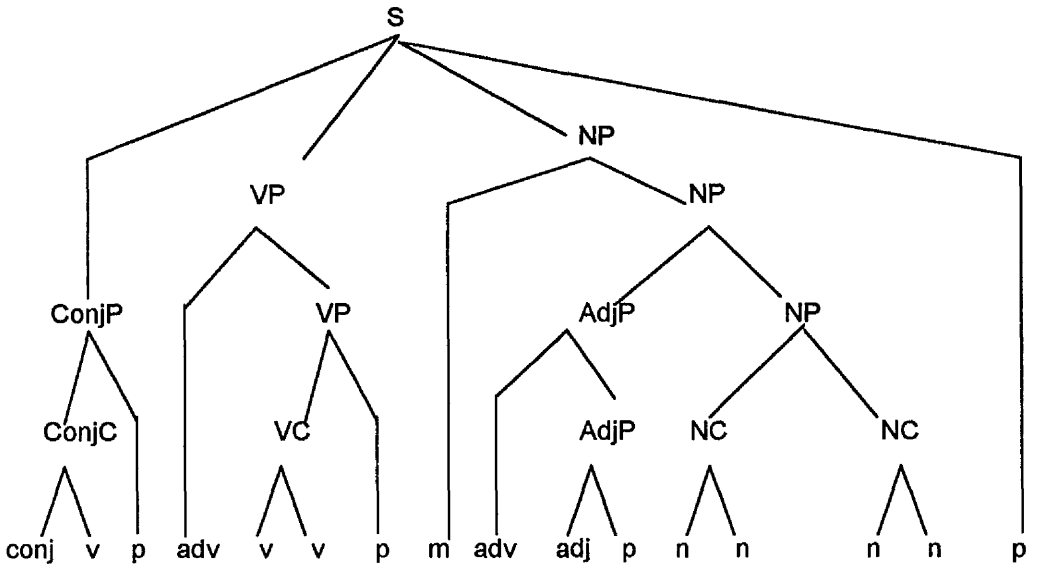
a:^{4} tɕek^{1} kɔ:^{3} kɔ:^{3} tow^{1}k^{w}u:n^{2} m^{4} ti:m^{6} la:^{1} ha:^{2}

|| 17 | | | ||

oh-that is-m.- m.- all-control- not- o.k.- p.- int.

(That means no one was able to be in control at all.)

[8]



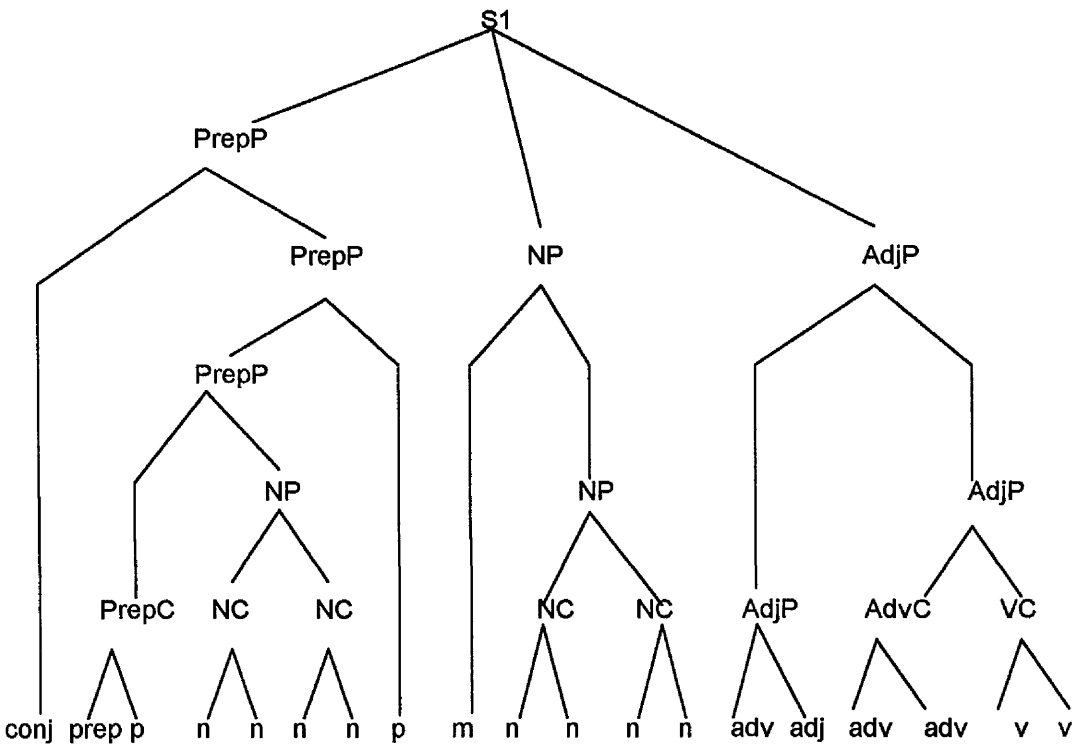
ta:n^{6} hɐj^{6}nɛ:^{1} tɕɐw^{6} ki:n^{3} la:p^{6} tɕɔ:^{2} kɔ:^{3} how^{2} how^{2} kɛ:^{3} p^{h}ɐŋ^{4} jɐw^{5} k^{w}a:n^{1}hɐj^{6} la:k^{3}

|| 18 || 19 | || 20 || 21 | ||

but-is-p.-then-build-establish-ed- m.- very- good- p- friend-friend-concern-relate-p.

(However, we were on very friendly terms.)

[9] [10]



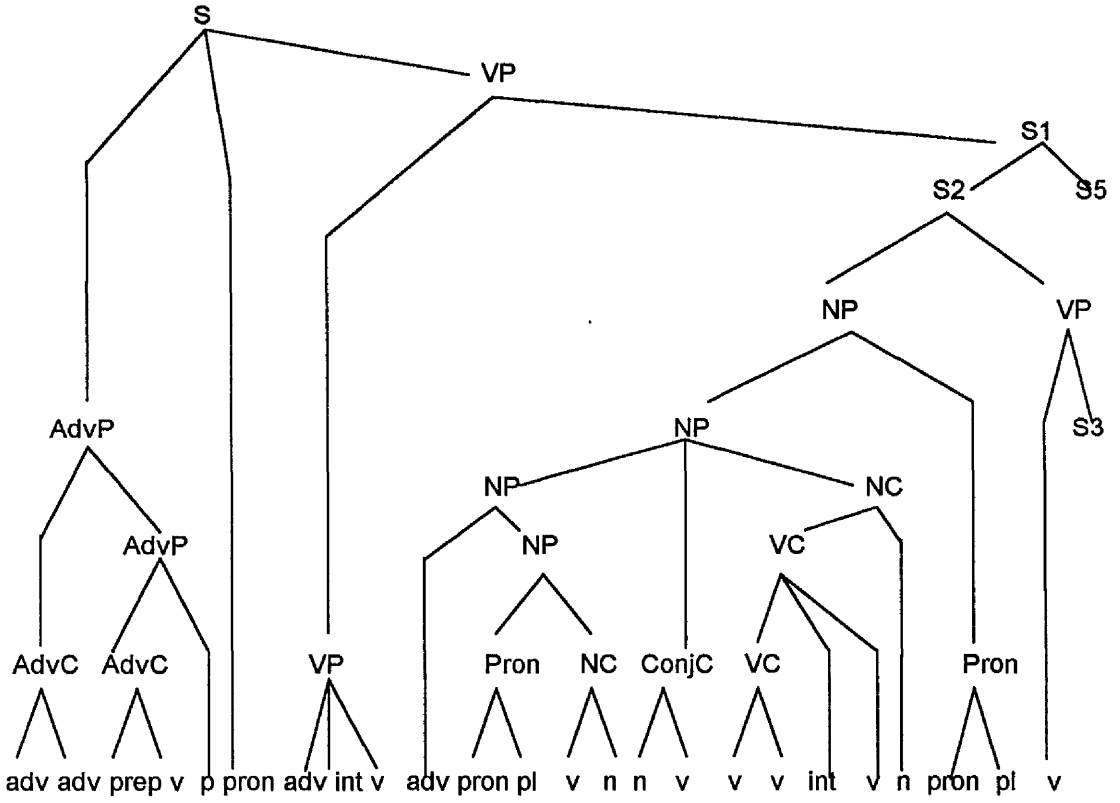
tɕek^{1}tɕ^{h}ɵj^{4}tɕɔ:^{2}ɕi:^{1}ɕɐŋ^{1}k^{w}a:n^{1}hɐj^{6} nɛ:^{1}kɔ:^{3}p^{h}ɐŋ^{4}jɐw^{5}k^{w}a:n^{1}hɐj^{6}how^{2}how^{2} wu:^{6}ɕœ:ŋ^{1}tɕy:n^{1}tɕoŋ^{6}

|| 22 || 23 | || 24 | | || 25 | ||

i.e-apart-p-teacher-student-concern-relate-p-m-friend-friend-concern-relate-very-good-mutual-mutual-respect-respect

(This means, apart from the teacher-student relationship, the friendship was very strong, mutual respect.)

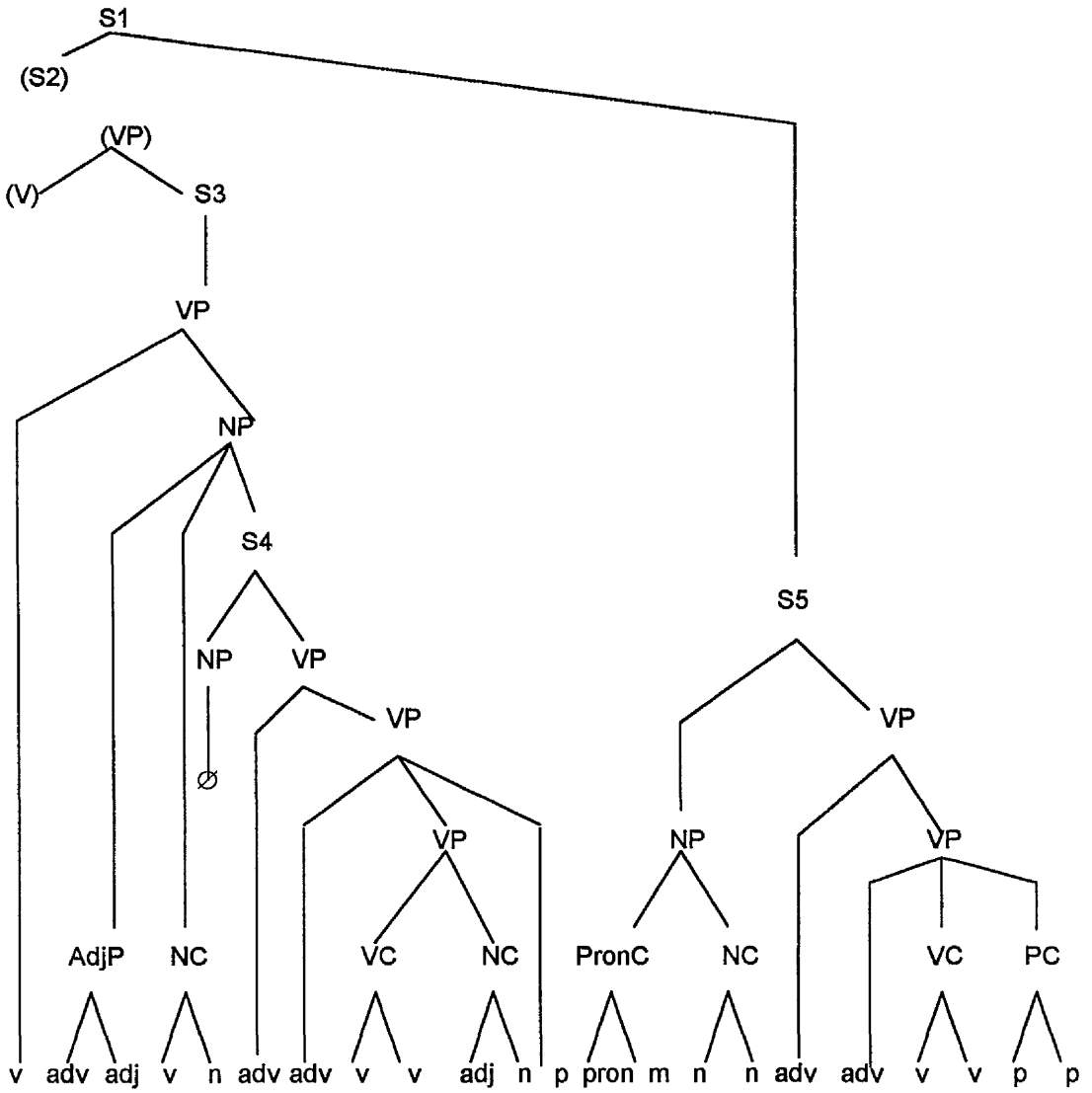
[11]



also-also-from-come-p-I-then-em-not-now(2)-that-pl-teach-expert-example-as-teach-teach-em-assess-committee-he-pl-say

(I never ... unlike those teachers nowadays, for example, the Body of Education Assessment say that)

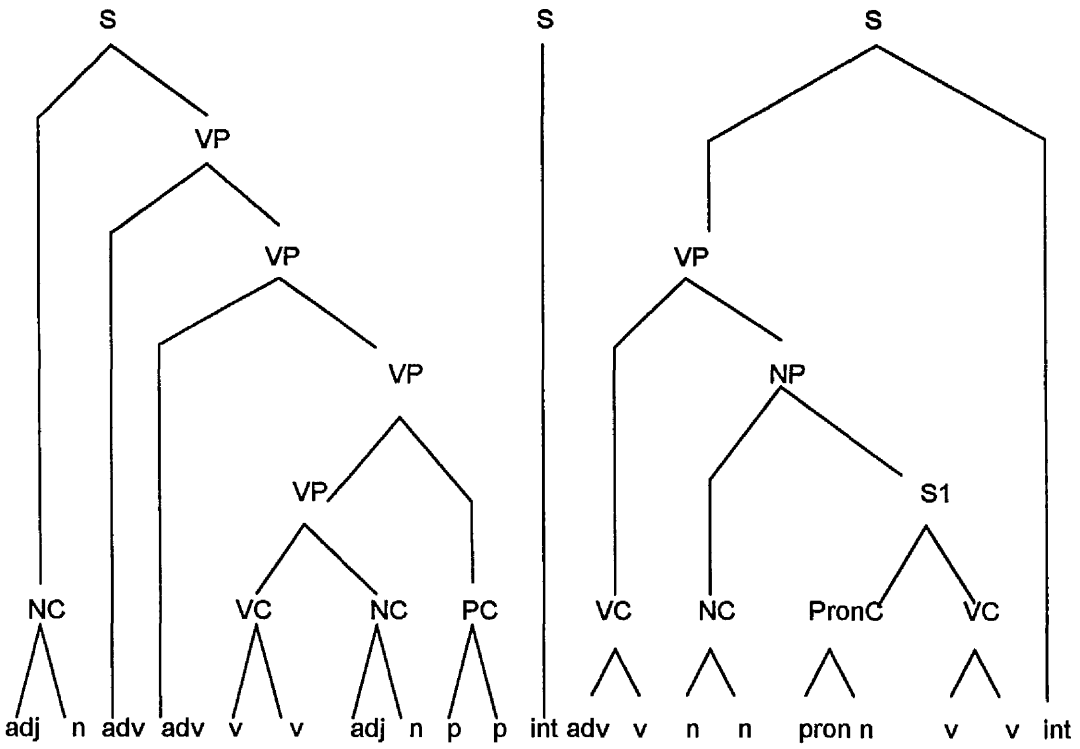
[12] [13]



have-very-many-study-pupil-all-not-respect-respect-old-expert-p-that-m-condition-situation-also-not-exist-be there-p-p

(many students don't respect teachers, such a situation did not exist.)

[14]



[15]

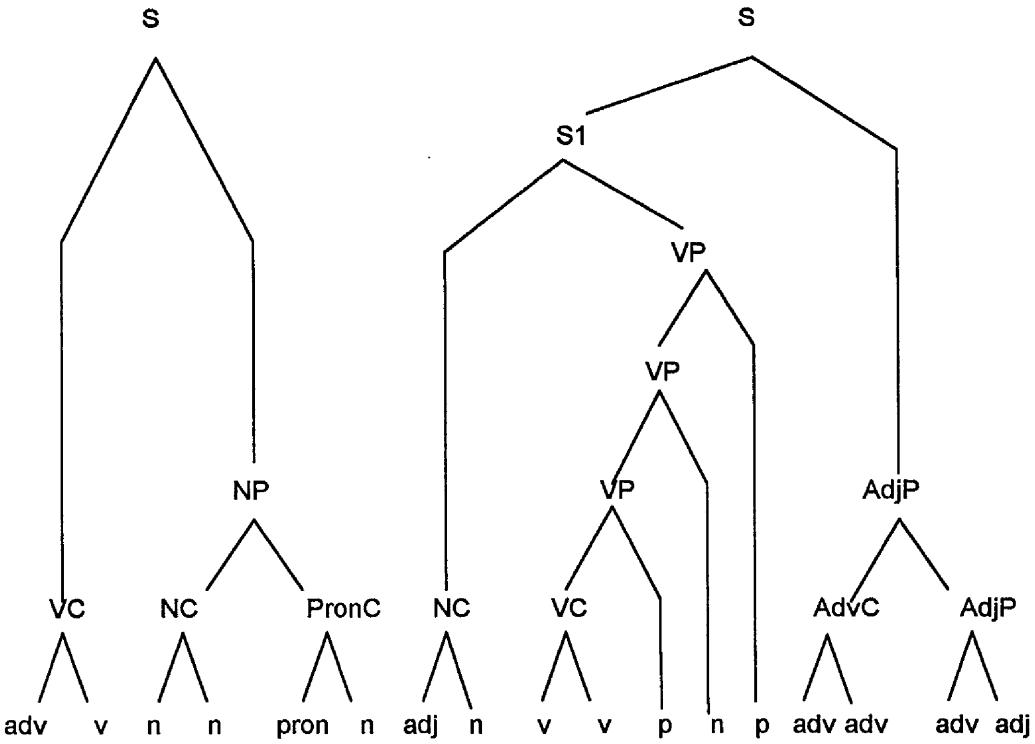
ta:j^{6}ka:^{1}tow^{1}how^{2}tɕy:n^{1}tɕoŋ^{6}low^{5}ɕi:^{1}kɔ:^{3}wɔ:^{3}m^{6}how^{2}tɕ^{h}i:^{5}p^{h}ɐŋ^{4}jɐw^{5}kɐm^{2}jœ:ŋ^{2}tɕy:n^{1}tɕoŋ^{6}ha:^{2}

|| 35 || 36 | || 37 | || 38| || 39 | | ||

big-family-all-very-respect-respect-old-expert-p-p-mh-very-like-friend-friend-this-look-respect-respect-int

(Everybody respected teachers very much. Right! The same as friends who respect each other.)

[16] [17]



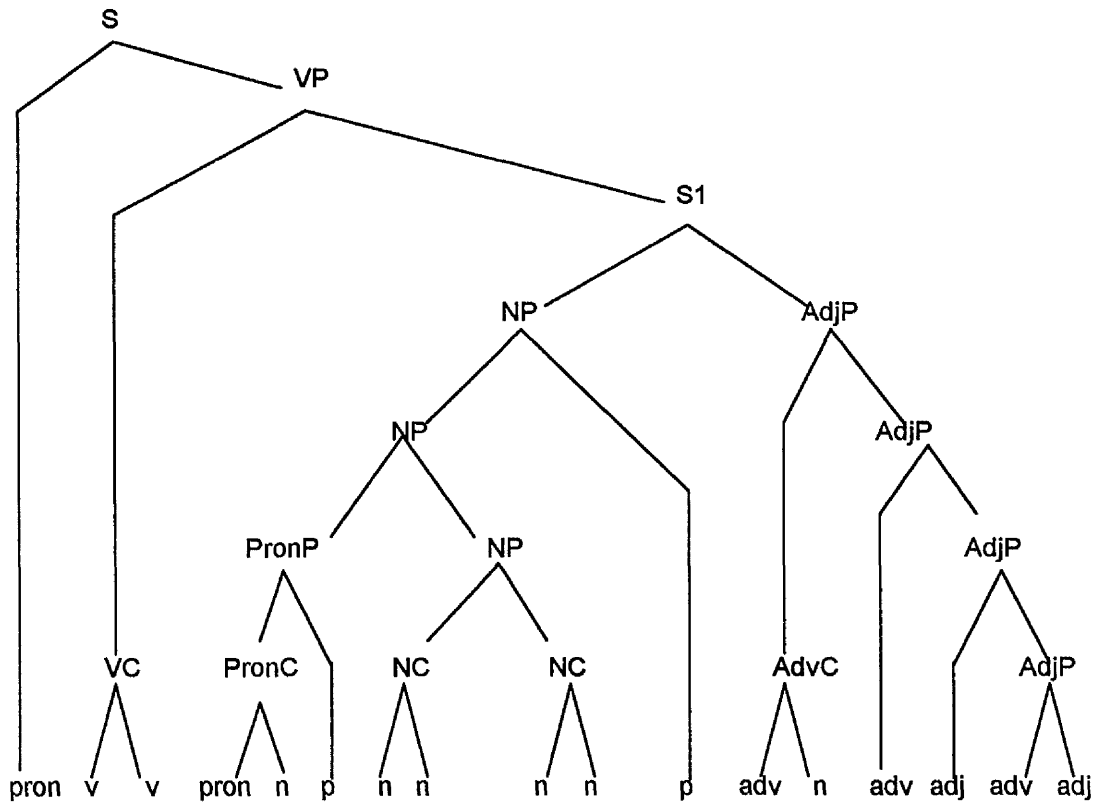
how^{2} tɕ^{h}i:^{5} p^{h}ɐŋ^{4} jɐw^{5}kɐm^{2}jœ:ŋ^{2} ta:j^{6} ka:^{1} k^{h}ɐp^{1} ɕɐw^{1} kɛ:^{3} jɛ:^{5}nɛ:^{1} jek^{6}tow^{1} how^{2} fa:j^{3}

|| 40 || 41 | || 42 || 43 | || 44 | ||

very-like-friend-friend-this-look-big-family-suck-receive-p.-things-p.-also-all-very-quick

(The same as friends ... Everybody learned really fast.)

[18]



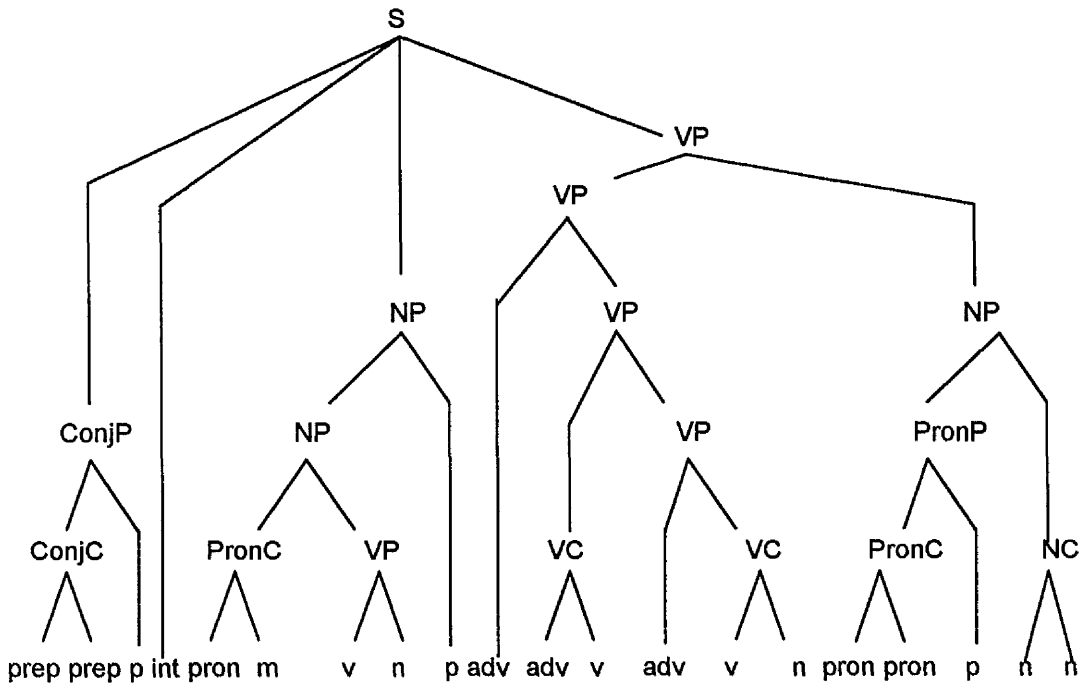
ŋɔ:^{5} kɔ:k^{3}tɐk^{1} kɐm^{2}jœ:ŋ^{2}kɛ:^{3}ka:w^{3}hɔ:k^{6} wa:n^{4} keŋ ^{2} nɛ:^{1} k^{h}ej^{4} çɐt^{6} tɕɐw^{6} how^{2} how^{2} tɔ:^{1}

|| 45 || 46 | | || 47 | ||

I- feel-able- this-look-p.- teach-learn-surrounding-environment-p.-that-fact-then-good-very-much

(I think this sort of educational environment was in fact much better.)

[19]



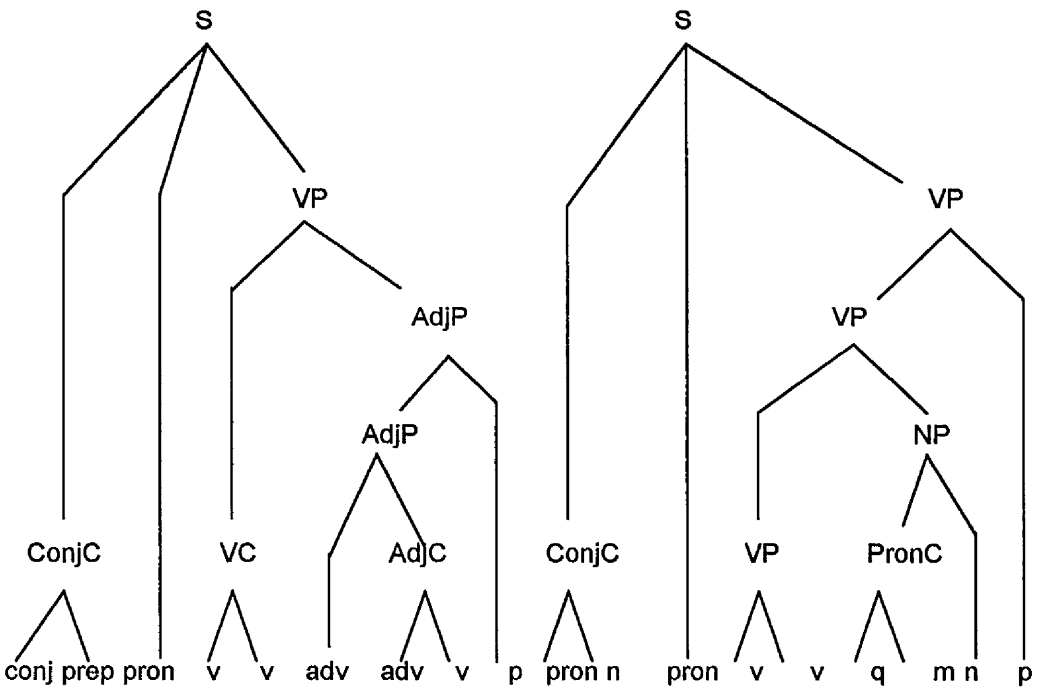
ɕo:^{2}ji:^{5} nɛ:^{1}ŋɛ:^{6}kɔ:k^{3} wɐj^{2} ka:w^{3} ɕi:^{1}nɛ:^{1}tɕɐw^{6}m^{4}ɕɐj^{2} kɐm^{3} ta:m^{1}ɕɐm^{1} tɕi:^{6}kej^{2} kɛ:^{3} k^{h}y:n^{4}wɐj^{1}

|| 48 || 49 || 50 | | || 51 | | | | ||

there-fore-p-em-every-m-teach-expert- p- then-no-need- so- bear-heart-self-self- p-power-might

(Therefore, teachers shouldn't worry about their authority.)

[20]



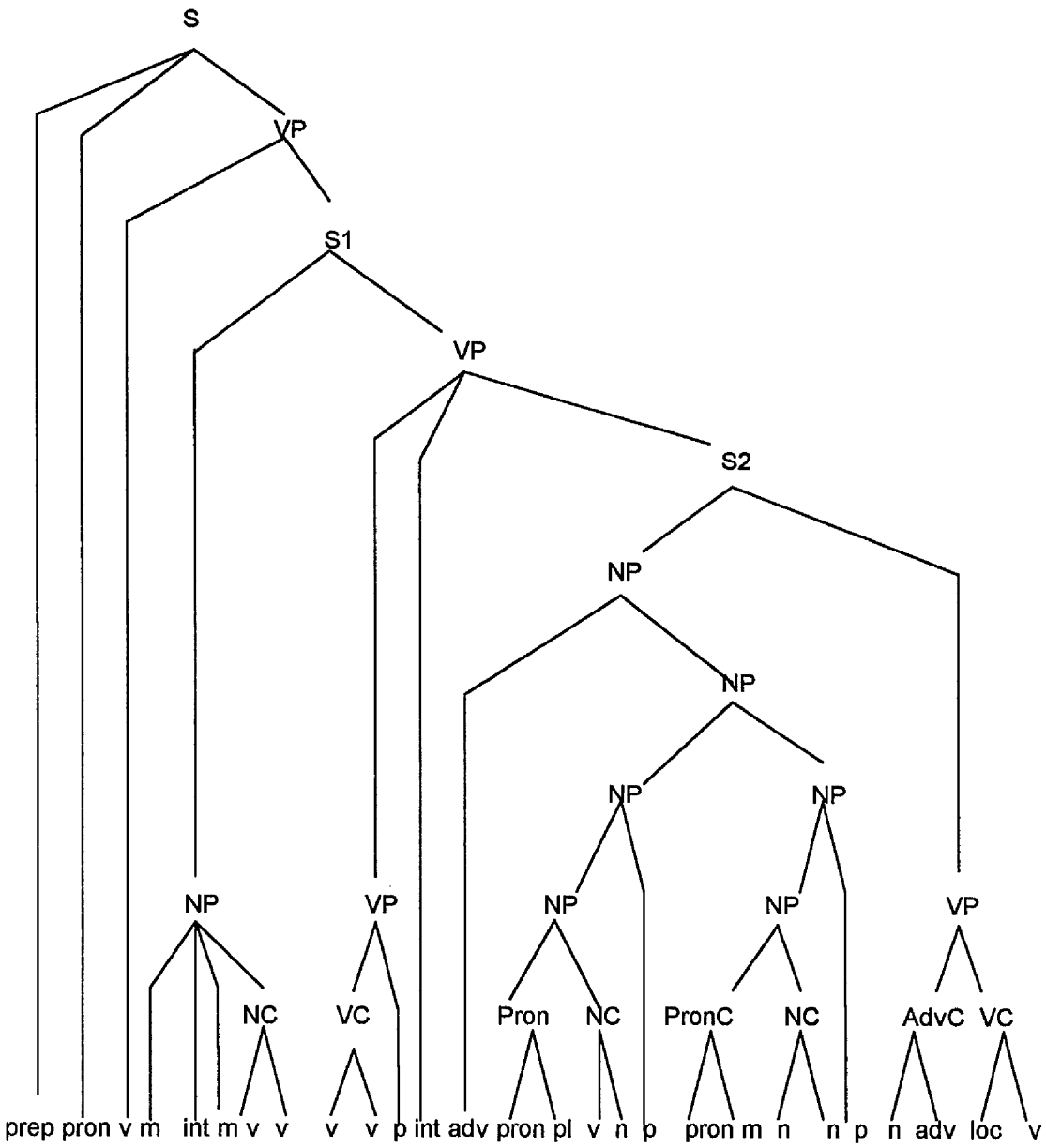
jɐn^{1} wɐj^{6} ŋɔ:^{5} kɔ:k^{3} tɐk^{1} how^{2} how^{2} çi:^{3}kɛ:^{3} k^{h}ej^{4} çɐt^{6} ŋɔ:^{5} çœ:ŋ^{2} kɔ:ŋ^{2} jɐt^{1} jœ:ŋ^{6} jɛ:^{5} nɛ:^{1}

|| 52 | || 53 | || 54 || 55 | | ||

because-for-I- feel-able- very-good- laugh-p.- that-fact- I-want- say- one-m.- thing-p.

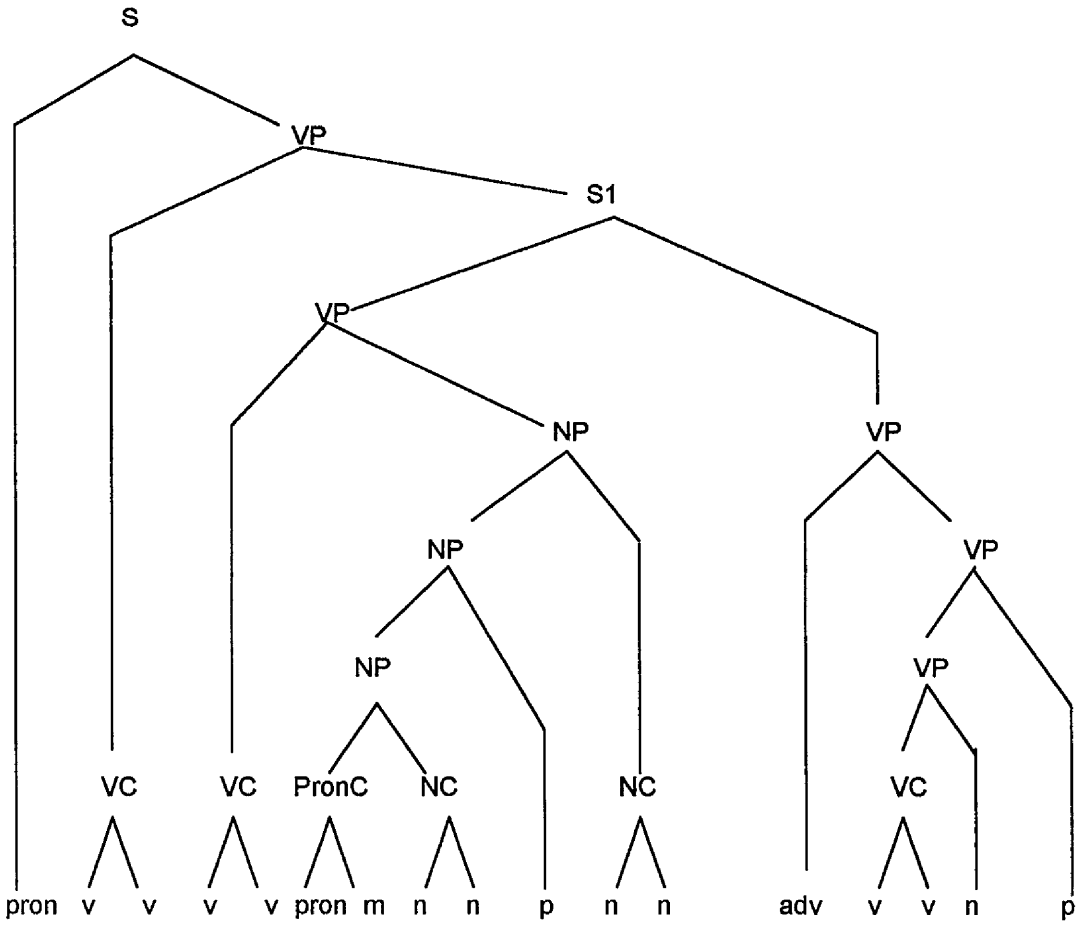
(Because I think it is silly. Actually I just want to say one thing.)

[21] [22]



(I see that investigations have shown that "nowadays the authority of us teachers is on the decline".)

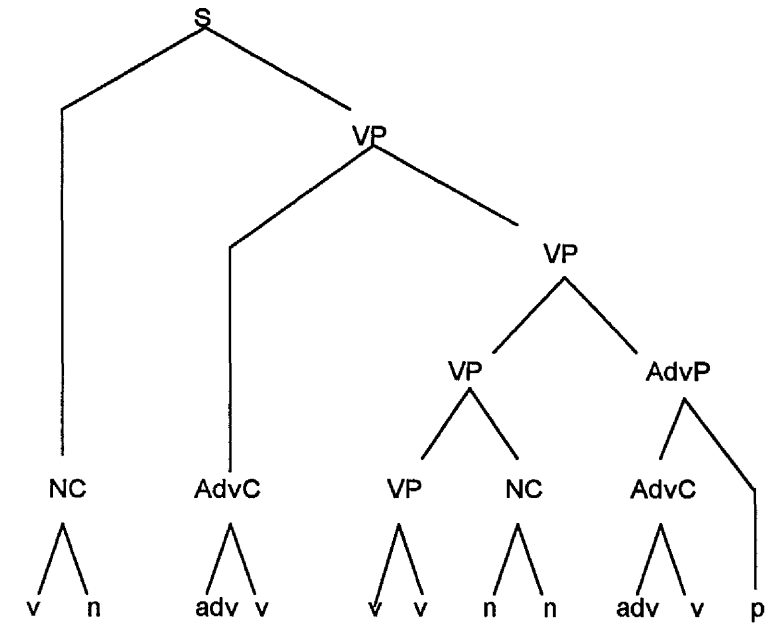
[23]



I- feel-able-lift-discuss-this-m.- title-topic-p.-investigate-check- all-deserve-beat-bottom(2)-p.

(I think that whoever suggested the topic of this investigation should have his bottom smacked.)

[24]



ka:w^{3} ɕi:^{1} ti:m^{2} ka:j^{2} ji:w^{3} jɐw^{5} k^{h}y:n^{4} wɐj^{1} ɕi:n^{1} tɐk^{1} ka:^{2}

|| 69 | | || 70 | | ||

teach-expert-how-explain-need-have-power-might-until-o.k.- p.-

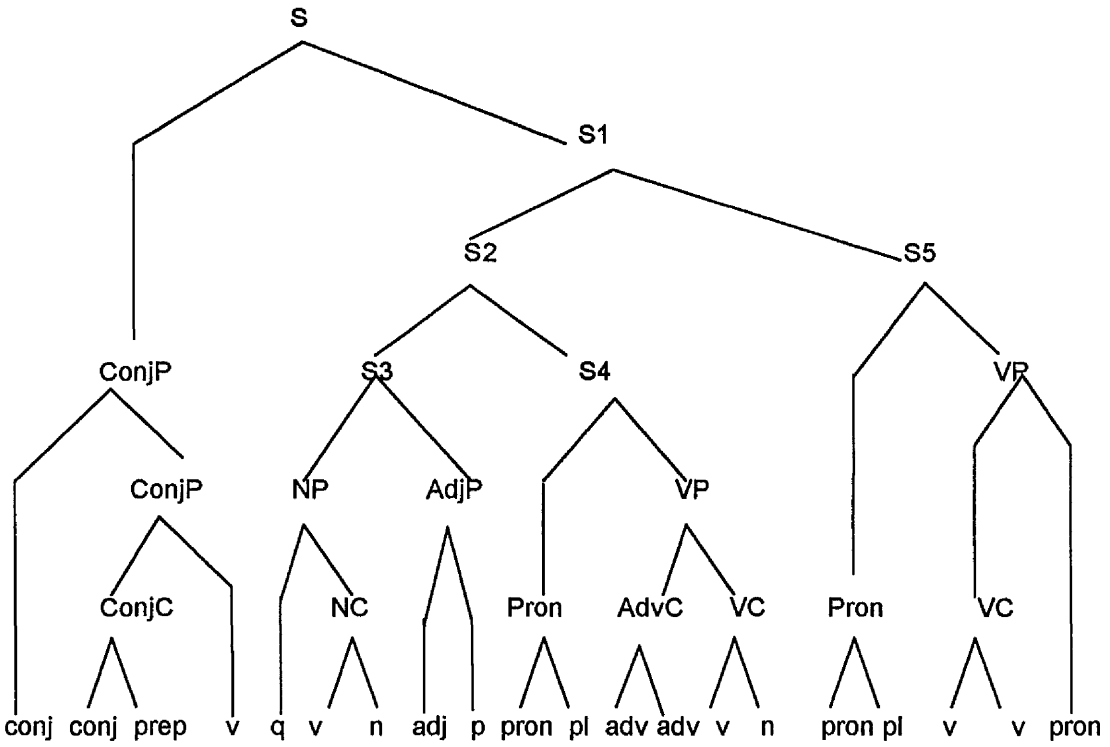
(Why do teachers need to have authority anyway? That's right!)

[25]

hɐj^{3} a:^{3}

|| 71 ||

right-p.



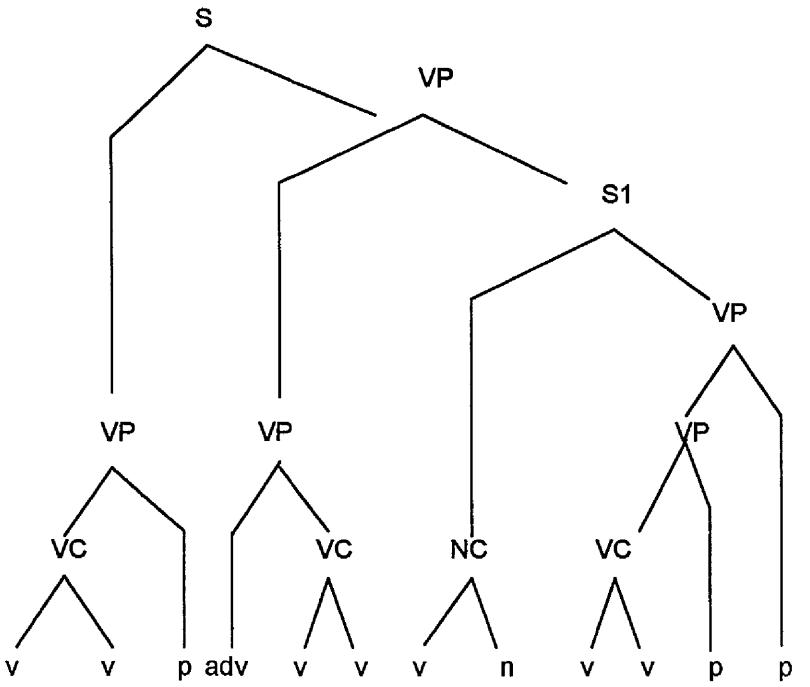
kɐm^{2} t^{h}oŋ^{4}ma:j^{4} wa:^{3} ti:^{1}hɔ:k^{6}ɕa:ŋ^{1}jɐj^{5}nɛ:^{1}ŋɔ:^{5} tej^{6}ji:m^{4}tɕoŋ^{6}fa:n^{6}k^{h}ɐj^{1}ŋɔ:^{5} tej^{6}hɔ:j^{1}tɕ^{h}ɵy^{4}k^{h}ɵy^{5}

||72| | || 73 | || 74 | | || 75 | ||

in that case-and-near-say-pl-study-pupil-naughty-p-I-pl-serious-heavy-offence-regulation-I-pl-open-expell-him

(And those students were naughty. We committed a serious breach of the rule. We expelled him.)

[27]



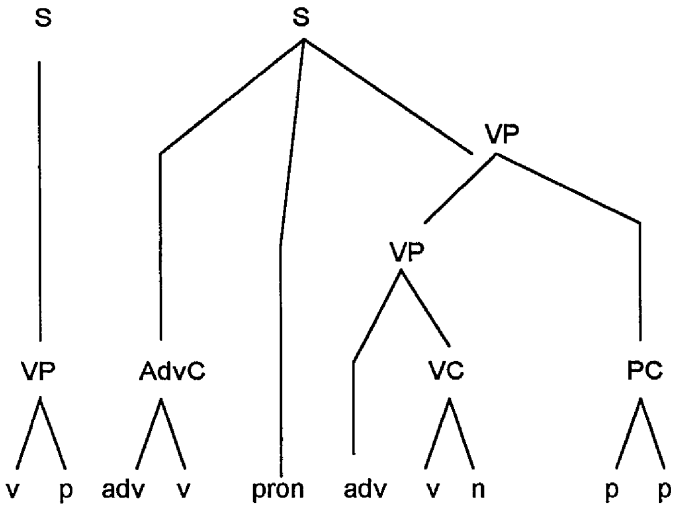
hɔ:j^{1} tɕ^{h}ɐy^{4} tɕɔ:^{2} m^{4} pi:w^{2} ɕi:^{6} mɐn^{6} thɐj^{4} ka:j^{2} k^{h}y:t^{3} tɕɔ:^{2} a:^{3}

|| 76 | | | | ||

open-expel- ed- not-show-show-ask- title-unlock- firm- ed- p

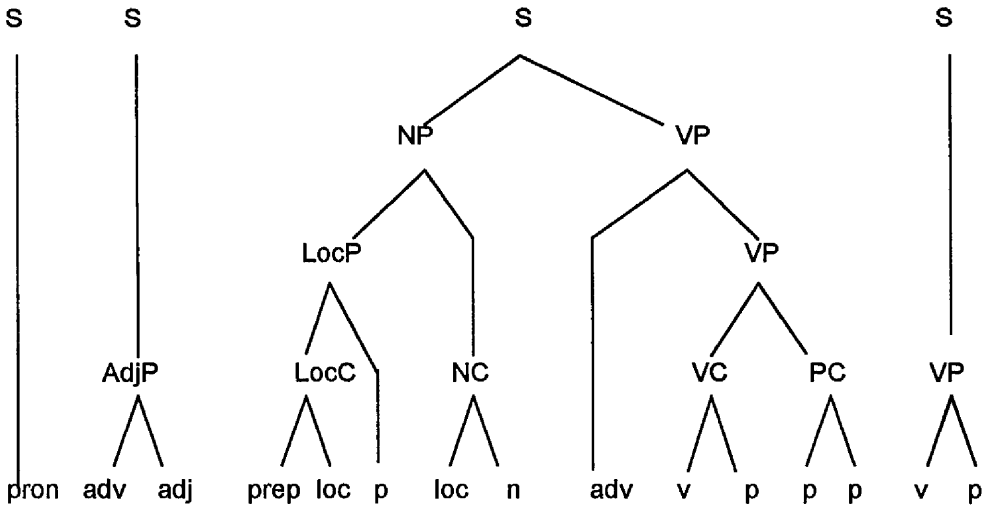
(Expelling him does not mean the problem has been solved.)

[28]



(That's right! Does he not study?)

[29] [30] [31] [32]



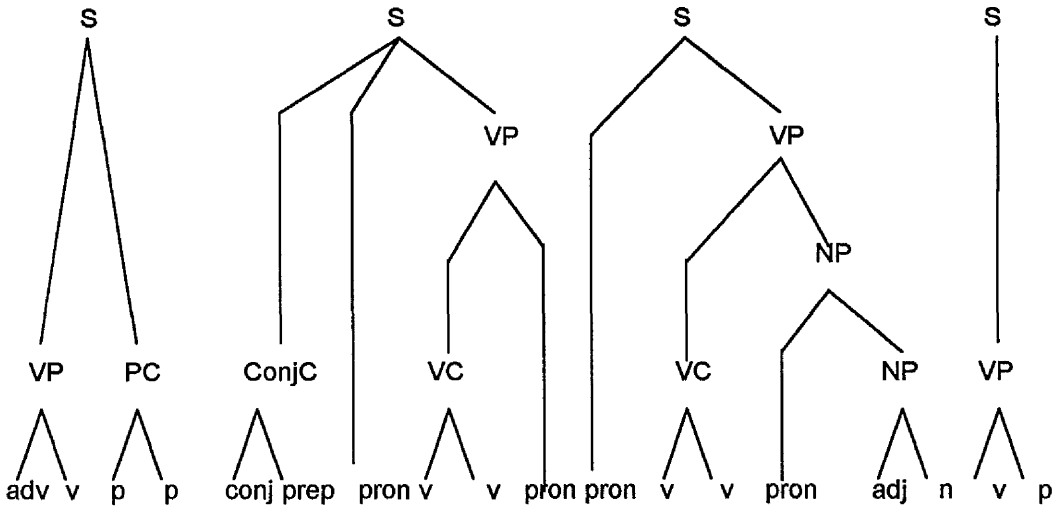
$k^{h}ɵy^{5}$ tɕoŋ^{6} tɕ^{h}a:m^{2} ji:^{5} hɐw^{6} kɛ:^{3} tɕ^{h}i:n^{4} tkow^{4} tɕɐw^{6} mow^{5} tɕɔ:^{2} lɔ:^{3} wɔ:^{3} hɐj^{6} la:^{3}

|| 79 || 80 | || 81 | | | || 82 ||

he- more-tragedy- in-back- p.- front-journey-then-not have- ed- p.- p.- right-p.

(He... Much worse! His future would have been doomed! Right!)

[33] [34] [35]



m^{4} tɐk^{1} ka:^{1} ma:^{3} jɐn^{1} wɐj^{6} nej^{5} kɔ:k^{3}tɐk^{1} k^{h}ɵy^{5} nej^{5} kɔ:k^{3}tɐk^{1} k^{h}ɵy^{5} wa:j^{6} jɐn^{4} hɐj^{6}a:^{3}

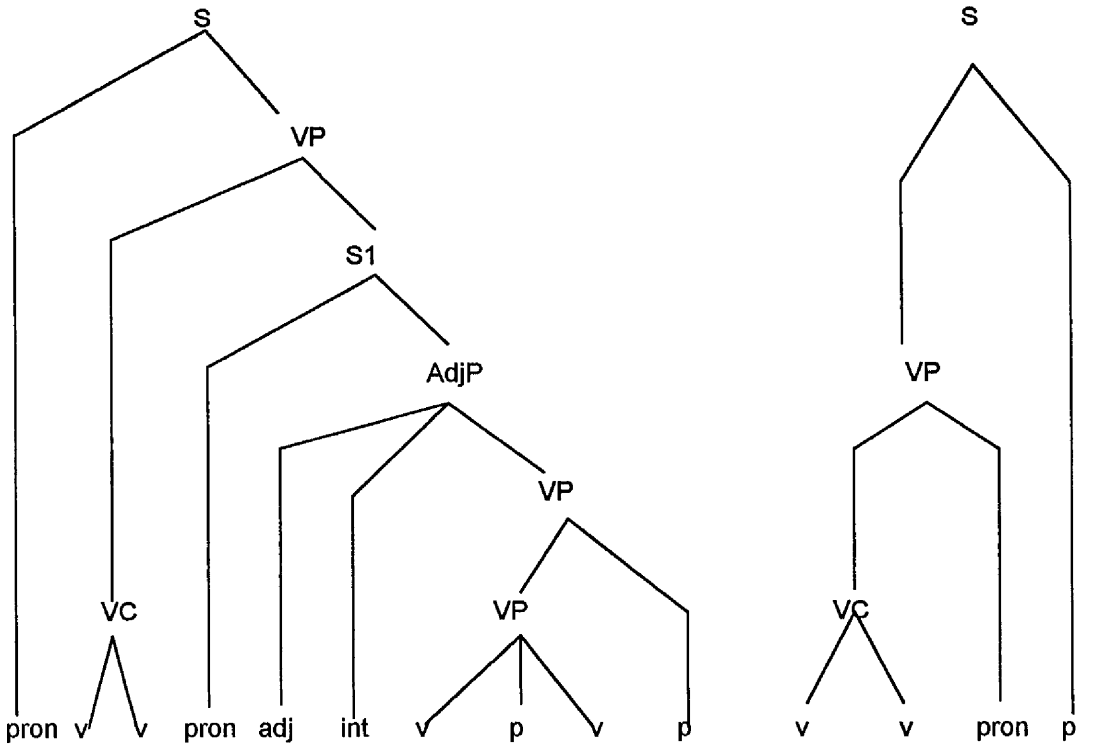
|| 83 | || 84 | || 85 | || 86 ||

not-o.k.-p.- p.-reason-for- you- feel- able- him- you-feel- able-him- bad-person-right-p.

(You cannot do that! Because you think he is ... You think he is a bad person ... Right!)

[35] continued

[36]



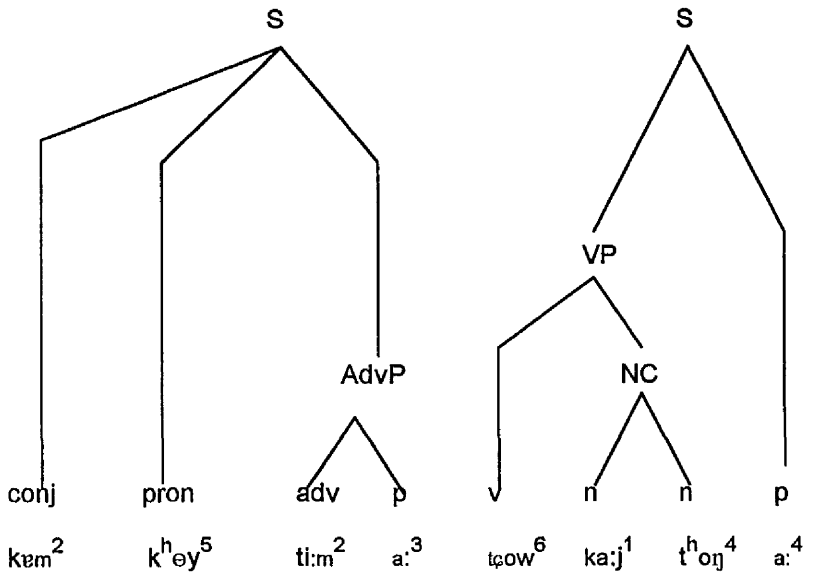
nej^{5} kɔ:k^{3} tɐk^{1} k^{h}ɵy^{5} jɐj^{5} ŋɐj^{3} mow^{5} tɐk^{1} ka:w^{3} lɔ:^{3} hɔ:j^{1} tɕ^{h}ɵy^{4} k^{h}ɵy^{5} la:^{1}

|| 87 | || 88 || 89 | || 90 | ||

you-feel- able-him-naughty-eh-have not-able- teach- p.- open- expel- him- p.

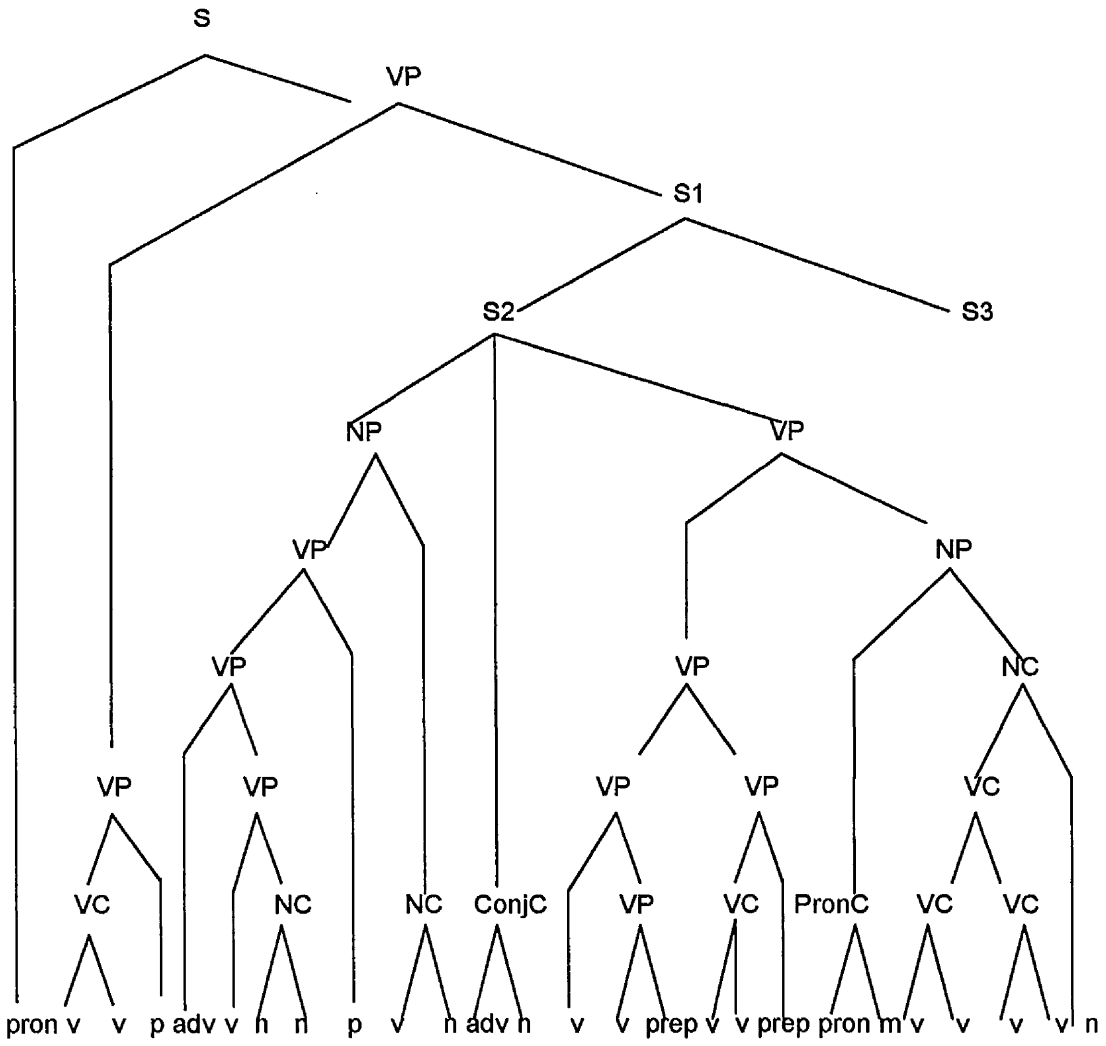
(You think he is naughty, unteachable, then expel him.)

[36] continued



(So, what will he do then? He'll be out on the streets?!)

[37] [38] [39]



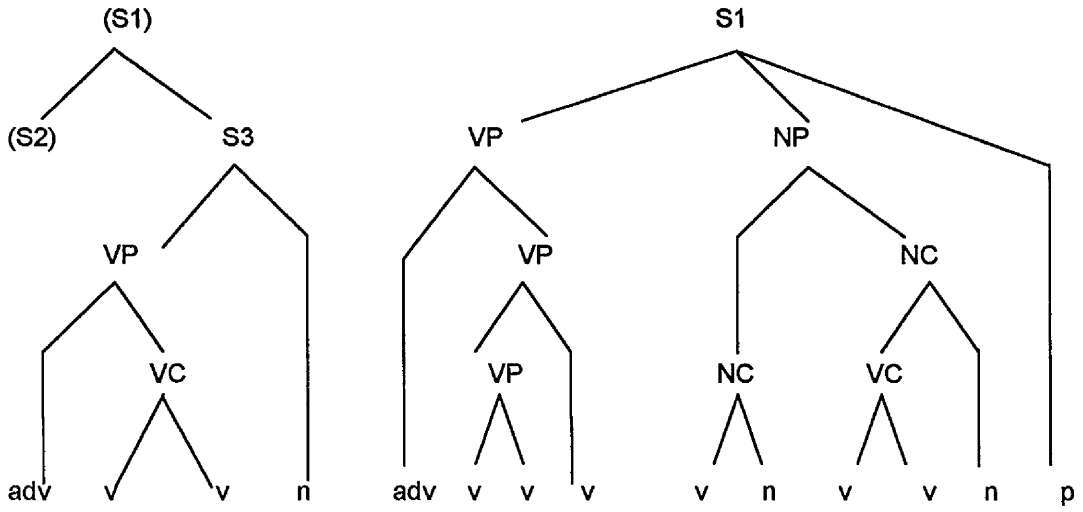
ŋɔ:^{5}kɔ:k^{3} tɐk^{1}a:^{3}pɐt^{1}fu:^{6}tɕa:k^{3}jɐm^{6}kɛ:^{3}kɔ:ŋ^{2}fa:t^{3}keŋ^{2}ji:n^{4}hɐj^{6}tɕ^{h}ɵt^{1}tɕɔ:j^{6}tɕ^{h}ɵt^{1}ji:n^{6} hɐj^{2}ni:^{1}kɔ:^{3}ka:w^{3}jok^{6}p^{h}eŋ^{4}ji:^{5}wu:j^{2}

||93| | ||94 | | | ||95 ||96 ||97 ||98 | | | ||

I-feel-able-p-not-bear-duty-post-p-talk-way-unexpectedly-look-is-out-at-out-present-at-this-m-teach-educate-assess-discuss-committee

(I think that is an irresponsible proposal and a very surprising one coming form the Body of Educational Assessment.)

[39] continued [40]



tɕɐn^{1} tɐj^{2} ta:^{2} lɔ:^{1} jɐw^{2} tɕoŋ^{6} ka:w^{2} tow^{3} hɔ:j^{1} kej^{3} tɕɛ:^{2} tɕi:w^{1} tɔ:j^{6} wu:j^{2} t^{h}i:m^{1}

||99 | | | | || 100 || 101 | | || 102 | | | 103 ||

really-deserve-beat-bottom(2)-still- do- reach-open-record-person-beckon-treat-meeting-p.

(They all really deserve to have their bottoms smacked – they even held a press conference!)

4.3. THE FOOT

4.3.1. Defining the Foot

The foot is a rhythmic unit on which the rhythm of a spoken language is based. The systems of speech melody and speech rhythms are highly interrelated. Note that Halliday defines the rhythmic unit before going on to describe melody in English as he considers the rhythmic structure in his definition of 'salient' (1970: 1). Rhythm is a topic at the edge of intonation, see my view above that intonation is speech melody. But it is a relevant domain for intonation structure because any speech melody pattern has to be based on some sort of speech rhythm. As a language which is not stress-using and in which the phonological structure of the sound is usually described in terms of monosyllabic words, Cantonese is susceptible to being described as exhibiting syllable-timed rather than stress-timed isochrony, given that "every language in the world is spoken with one kind of rhythm or with the other" (Abercrombie 1967: 97). This suggests that syllables in Cantonese tend to be approximately equal in length in utterances and therefore one syllable is one foot. However, we tend to hear the sequences of alternating lengthened and shortened syllables in actual speech in Cantonese as linked into groups, and we call these `feet'. A foot is a rhythmical beat. Prominence lands on a lengthened syllable and such a rhythmic grouping in Cantonese is described as `prominence based rhythm'. (For details, see discussion in chapter three 3.1.2.)

In Cantonese, syllable length can be affected by different factors, for example, syllable structures, tempo or emphatic stress, etc. In terms of syllable structure, the same vowel can vary its length according to the type of coda that follows (see discussion in chapter one 1.2.3.3.1.1.). The long-short vowel distinction remains whatever type of rhymes the vowel belongs to, as the long-short vowels also differ in quality (e.g., /a:/ and /ɐ/). A given vowel appears to be at its shortest if it has a plosive coda rather than a nasal coda, and at its longest if it is in an open syllable. The nasal codas vary their length according to whether they follow a long or short vowel and exhibit complementary distribution: a long nasal follows a short vowel, and vice versa (see discussion in chapter one 1.2.3.3.1.2.2.). On the other hand, with the same syllable structure, the syllable length is not very much affected by tones (see discussion in chapter one 1.2.4.1.5.). I argued in chapter one that the duration of a vowel can

be stretched or shortened but the vowel quality is maintained. I also consider that the duration of a syllable can be stretched but the lexical meaning is maintained (see discussion in chapter one 1.2.3.3.1.2 and 1.2.4.1.5). The flexibility or stretchiness of the length of a syllable gives the rhythmic grouping in Cantonese more flexibility.

When we come to decide how many syllables there are in a foot, we mean how many syllables recur at approximately equal intervals of time as we hear them in a steady speech tempo. To my perception, syllables in succession are rarely to be heard lasting for nearly equal time. There is always one syllable which is heard to be longer than the others. Lengthened and shortened syllables tend to be in pairs (although where there are two shortened syllables they can come in threes). I call them long-short or short-long pairs. Long refers to a lengthened syllable and short refers to a shortened syllable. They are not phonologically long (e.g., a syllable containing a long vowel) or phonologically short (e.g., a syllable containing a short vowel), rather their actual durations are freely stretched or compressed to accommodate alterations in rhythm and prominence (for details, see discussion in chapter three). In utterances, syllables are normally divided into long-short or short-long pairs and each disyllabic long-short or short-long pair forms a foot. Normally there is one and only one `long' in a foot, and a foot can have one or two `short'(s) or none. A pair of one long and one short is the most common structure. The `long' of a foot is a potential prominent syllable. Differences in segmental sonority and in syllable weight seem to have little to do with the long-short or short-long pattern within a foot.

4.3.2. The Foot Grouping

4.3.2.1. Description of the Foot Grouping

A disyllabic monomorpheme (as /lɔ:^{1} jɐw^{2}/ in example 67-68 below), a disyllabic compound (as in discussion of example 25 below) or two successive monosyllabic words (as in discussion of example 56 below) are normally grouped into a foot. Some words may have their preferred rhythmic pattern (for instance, a reduplicated rhythmic suffix or a reduplicated adjective as discussed in chapter three 3.3.2.1.1 and 3.3.2.1.2.1, respectively). However, the syntactic structure of a word does not determine the long-short or short-long pattern in a

string of words. A lengthened syllable if placed at the beginning of a foot is usually accompanied by a slightly greater force of articulation whereas the slightly greater force of articulation is not found when the lengthened syllable is not placed at the beginning of a foot. If it happens that a monosyllabic word occurs beside a disyllabic word, or in between two disyllabic words, there will be a number of possibilities for the monosyllabic word: (a) if it is lengthened, it stands as a foot on its own (as /nɛ:^{1}/ in the following discussion of example 50); (b) if it is shortened, it joins either the disyllabic word forming a foot (as /ŋɔ:^{5}/ in the following discussion of example 52-53); or, it combines with a silence to make a foot (as /kɐm^{2}/ in the following discussion of example 7 below).

Let us start with the following example (the numbers preceding the annotation indicate the intonation group the feet belong to; for definition of intonation group see next section; "|" indicates a foot boundary; the underlining "___" stands for a syllable, __ and ____ respectively indicate relative shortened or lengthened; for figures, see appendix ix):

52-53 | jɐn^{1} wɐj^{6} ŋɔ:^{5} | kɔ:k^{3} tɐk^{1} | how^{2} how^{2} | çi:^{3} kɛ:^{3} |
| ____ __ __ | _ ___ | ____ _ | ____ _ |
because-for-I-feel-able-very-good-laugh-p.
(Because I think it is silly)

Here we have four feet. There are three different types of syllable structure in the first foot. The second and the third syllables compress into a half beat in order to lend time to the first syllable (115ms ~ 57ms ~ 88ms, respectively). The constituents of the second foot are two plosive-coda syllables. The long beat lands on the short vowel syllable rather than the long vowel one. As mentioned above, vowels always maintain their phonological long-short contrast in any syllable structure as the long-short vowels also differ in quality. Here a long vowel syllable is realised in shorter duration than a short vowel syllable (115ms ~ 147ms, respectively). Disyllabic compounds or disyllabic monomorphemes normally coincide with a foot. This is how native speakers decide how to group sequences of syllables. The compound /jɐn^{1} wɐj^{6}/ because-for - because) is grouped in the first foot with the monosyllabic pronoun /ŋɔ:^{5}/ (I). The verb-suffix combination /kɔ:k^{3} tɐk^{1}/ (feel-able - feel) is grouped into

the second foot. The monosyllabic pronoun /ŋɔ:^{5}/ (I) in between the two could in general either combine with the previous compound or the following verb-suffix combination, or just stand on its own as a foot. The last possibility is ruled out in this instance as it is too short (88ms) in comparison with its neighbours (its previous compound lasts 172ms and its following lexical item lasts 262ms). It combines with the previous compound to form a foot almost equal in length with the following combination. The third foot consists of two identical syllables, one is heard to be longer than the other (135ms ~ 105ms, respectively). There are two long vowels in the fourth foot, the first syllable is much longer than the second syllable (345ms ~ 50ms, respectively). The first syllable is under focus, its length is seven times that of the second syllable. Every foot does not have absolutely equal length as music bars do, the last foot in the example above lasts twice as long as its individual predecessors. However, we cannot count on the physical evidence from measurements as the fact is that empirical investigations have reportedly failed to show any regularity of isochrony, according to Laver (1994: 523). Counting rhythmic patterns is not something mechanical which can be done by a metronome. It relies on something like Laver's suggestion that "listeners have a cognitive readiness to impose a regularised temporal grouping on utterances they hear" (1994: 525). The four feet in the example above seem to be grouped quite evenly, i.e., each foot equalises the temporal intervals (the last foot is just one and a half in length), but this is not always the case. An utterance may vary its tempo and this may stretch or compress a foot to some extent. A disyllabic stretching of long-short or short-long pattern is the basic constituent of a foot. If there are more than two syllables, either there will be three syllables squeezed together into a foot like the first foot of the example above or a single syllable forms a foot by itself. A disyllabic compound word normally is not split up by a foot boundary (see discussion below), but here the adjective compound /how^{2} ɕi:^{3}/ (good-laugh – laughable) divided between the last two feet is considered special. The special condition for this is that the adjective compound and the monosyllable adverb /how^{2}/ (very) form a phrase /how^{2} how^{2} ɕi:^{3}/ (very-good-laugh – so funny). The syntactic structure of the phrase is very loose. The first two syllables `very good' can be a very common phrase and for this reason they can form a foot. In that case, they leave the last syllable `laugh' on its own and form an expression

`very very funny' – a phrase consisting of a double adverb followed by an adjective. When a disyllabic adjective is under focus, either syllable is lengthened in order to achieve prominence (see discussion in chapter three 3.3.2.4.1.). Here /ɕi:^{3} / (laugh) is chosen to be lengthened, /how^{2}/ (good) is relatively compressed. Such a short-long pattern can constitute a foot by itself. However, the disyllabic compound follows an adverb /how^{2}/ (very) forming an adverb-adjective phrase (/how^{2} how^{2} ɕi:^{3}/ (very-good-laugh – so funny). The adverb /how^{2}/ (very) is positioned at the beginning of a foot. A common rhythmic pattern of an adverb-adjective phrase is to have the adverb prolonged in both disyllabic and trisyllabic phrases (see discussion in chapter three 3.3.2.5.). When the two identical syllables are adjacent, the short-long pattern of the disyllabic adjective causes their preceding adverb to take its longer duration as its usual form. As a result, the first syllable (/how^{2}/ (good)) of the disyllabic adjective is further compressed as between two lengthened syllables, and the second syllable (/ɕi:^{3}/ (laugh)) of the adjective has to start a new foot.

According to Halliday "a musical bar may also begin with a rest, a silent beat... A foot may begin with a silent beat, without the rhythm becoming disrupted or lost" (1970: 1). When this occurs in Cantonese speech the `rest' occupies the `lengthened' part of a foot defined as a rhythmic beat. The syllable which shares a foot with the silence is naturally shortened and the silence is usually accompanied by a pause. The following example is taken from the beginning of an utterance spoken by the speaker A who spoke after the speaker B (the silence is indicated by a subscript " ˄ ").

7 | kɐm^{2} | ka:w^{3} ɕi:^{1} nɛ:^{1} |
| ˄ _ | ____ __ __ |
in that case-teach-expert-p.
(And, teachers)

4.3.2.2. The Arrangement of Long-short and Short-long

The types of compound or the meaning of its individual syllables do not decide a long-short or a short-long pattern. The `long' or `short' is decided by the comparison of the length with its

neighbours. The following example consists of two coordinate compounds in two feet, but the first foot is a short-long and the second is a long-short pattern:

25 | wu:^{6} ɕœ:ŋ^{1} | tɕy:n^{1} tɕoŋ^{6} |
| _ ___ | ___ __ |
mutual-mutual-respect-respect
(mutual respect)

Even if two successive syllables do not form a compound, they are also likely to be grouped into a foot to form either a long-short or a short-long pattern. There is no regularity in how the long-short and short-long patterns arrange themselves as can be seen in the following examples, one consisting of two feet and another one consisting of three feet long:

56 | tɕʁw^{6} ŋɔ:^{5} | t^{h}ʁj^{2} kɔ:^{3} |
| ___ _ | ___ _ |
concern-I-see-m.
(I see that)

10-11 | mow^{5} mʁt^{1} | ti:m^{2} fʁt^{6} | ŋɔ:^{5} tej^{6} kɛ:^{3} |
| ___ _ | _ ___ | ___ _ _ |
have not-any-how-punish-I-pl.-p.
(never really punished us)

4.3.2.3. Final-particle

A final-particle can form a foot on its own but not necessarily, it depends on its own relative length in comparison with its preceding syllables. As a final syllable in the utterance, it can be squeezed into the preceding foot or it can stand on its own. Examples for the former can be seen from the last foot of 11 above, for the latter can be seen from 50 below:

50 | kɔ:k^{3} wʁj^{2} | ka:w^{3} ɕi:^{1} | nɛ:^{1} |
| ___ _ | ___ _ | ___ |
every-m.-teach-expert-p.
(teachers)

In the case of final-particle clusters, the actual duration of the two syllables can be very short as the two tend to get squeezed together or even squeezed with the preceding syllable(s) to

form a foot. The last syllable is always longer than the first one and can be lengthened greatly. The following two examples (33-34 and 35-37) show that the same particle cluster /kɔ:^{3} wɔ:^{3}/ in the first example is too short to be heard as an independent foot: 47ms for the first syllable whose vowel is deleted and 87.5ms for the second syllable, the particle cluster being realised as a consonant cluster segmentally. Therefore, the last foot consists of four syllables phonemically but only three syllables in realisation. In comparison with the length of its preceding syllable (227ms), the cluster is obviously to be heard to combine with its preceding syllable into one foot. In the second example, the lengths of the syllables of /kɔ:^{3}/ and /wɔ:^{3}/ are 67and 115ms, respectively. In comparison with the length of the preceding syllable (230ms), the particle cluster is heard to be a single foot on its own.

187ms 227ms 47ms 87.5ms

33-34 | kɔ:^{2} tɕoŋ^{2} | tɕ^{h}eŋ^{4} fɔ:ŋ^{3} | jɐw^{6} m^{4} | tɕ^{h}yn^{4} tɕ^{h}ɔ:j^{6} kɔ:^{3} wɔ:^{3} |
| _ ____ | ___ _ | _ ___ | _ _____ _ _ |
that-m.-condition-situation-also-not-exist-be there-p.-p.
(such a situation did not exist)

145ms 230ms 67ms 115ms

35-37 | ta:j^{6} ka:^{1} | tow^{1} how^{2} | tɕy:n^{1} tɕoŋ^{6} | low^{5} ɕi:^{1} | kɔ:^{3} wɔ:^{3} |
| ___ _ | _ _____ | ___ _ | _ ___ | _ ___ |
big-family-all-very-respect-respect-old-expert-p.-p.
(Everybody respected teachers very much)

4.3.2.4. Staccato

In the case of focus, the lexical items are heard to be significantly lengthened as in example 53 above (in which /ɕi:^{3}/ (laugh) is lengthened the most), or to be chopped: a foot is occupied by a single syllable. Such a rhythm is called staccato. The staccato syllables are produced by means of more energetic movements and are heard to last for approximately equal duration but do not sound prolonged because of the abrupt chopping. They are louder than a non-staccato syllable because of the more breath pumping from the lungs. Staccato examples are illustrated below (a chopped syllable is indicated by a superscript " $^{|}$"):

64 | jɐt^{6} tɕi:m^{6} | ha:^{6} | kɔ:ŋ^{3} |
| _ ___ | ˈ___ | ˈ___ |

day-gradually-down-sink

(on the declined)

99 | tɕɐn^{1} | tɐj^{2} | ta:^{2} | lɔ:^{1} | jɐw^{2} |
| ˈ___ | ˈ___ | ˈ___ | ˈ___ | ˈ___|

really-deserve-beat-bottom(2)

(deserve to have their bottoms smacked)

4.3.2.5. One-syllable Utterance

In the case of a one syllable utterance, a foot is naturally occupied by the single syllable. In the following example, the speaker A was interrupted by speaker C just after he uttered the first word of the utterance:

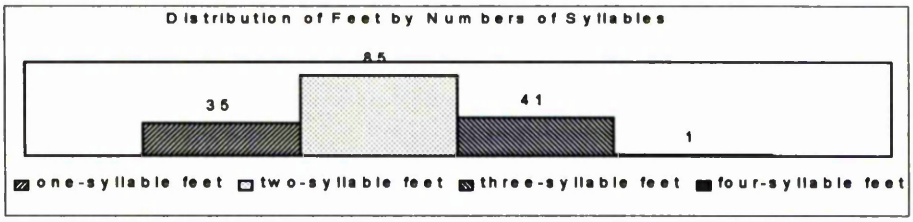
79 | k^{h}ɵy^{5} |
| ___ |

he

(he)

There are a total of 162 feet containing 331 syllables in our data for acoustical analysis. Their distribution is displayed in the figure below.

Fig. 4.1. The distribution of the 162 feet by numbers of syllables.



4.4. INTONATION GROUP

4.4.1. Defining Intonation Group

4.4.1.1. The Resetting of Pitch Span

In spoken Cantonese, the melody of a stream of speech is marked off by resetting the pitch span into small chunks. I shall call each small chunk an intonation group. The only criterion

for marking an intonation group is the resetting of the pitch span. This means that the pitch span of an intonation group differs from that of its preceding and following intonation group. Only the resetting of the pitch span makes an intonation group a discrete unit in speech prosody. The resetting of pitch span is defined as shift of the fitted lines (for this item, see discussion below) of the unit of the intonation group to a different register (i.e., different from the fitted lines in the previous intonation group). An intonation group is also a unit of information a speaker wants to convey, which coincides with a syntactical unit – a word, a phrase, a clause or a sentence. It is also unitary in semantic terms. A polysyllabic word, for instance, a simple compound or a polysyllabic-monomorphemic word, does not break off an intonation group.

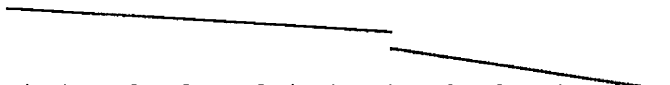
Here I illustrate two pairs of consecutive intonation groups whose pitch span is shifted to a different register. The 'shift' can be up or down. The first pair shows that the second intonation group disjoints itself from the first intonation group by shifting up, whereas the second pair shows that the second intonation group disjoints itself from the first intonation group by shifting down ("||" indicates the boundary of intonation groups and "%" indicates a pause):

24-25 || kɔ:^{3} p^{h}ɐŋ^{4} jɐw^{5} | k^{w}a:n^{1} hɐj^{6} | how^{2} how^{2} || wu:^{6} ɕœ:ŋ^{1} | tɕy:n^{1} tɕoŋ^{6} ||%

|| _ ___ _ | ___ _ | _ ___ || _ ___ | ___ _ ||

m.-friend-friend-concern-relate-very-good-mutual-mutual-respect-respect

(the friendship was very strong, mutual respect)



67-68 || t^{h}ɐj^{4} mok^{6} kɛ:^{3} | ti:w^{6} tɕ^{h}a:^{4} tow^{1} || tɐj^{2} ta:^{2} | lɔ:^{1} jɐw^{2} | la:^{1} ||

|| ___ _ _ | ___ _ _ || ___ _ | _ ___ | ___ ||

title-topic-p-investigate-check-all-deserve-beat-bottom(2)-p.

(the topic of this investigation should have his bottom smacked)

The pitch span varies from one intonation group to another in two respects: the height of the highest fitted line which covers all peaks in the F_0 contour, and the width between the highest fitted line and the lowest fitted line which covers all valleys in the F_0 contour. Taking the examples 24 and 25 above, the difference of the height of the highest lines can be easily detected: that of 25 is higher than that of 24 (T1 in 24 is 103Hz and in 25 is 169Hz and 158Hz). So can the difference of the width between the highest lines and the lowest lines: the band of F_0 value is 53Hz between T1s and T6s in 25, but only 25Hz between T1 and T6 and only 40Hz between T1 and T4 in 24. The width of the tonal space between T1 and T6 is greater in 25 than in 24, not to mention that the lowest line in 24 covers T4 which is phonologically lower than T6. The above example clearly demonstrates that only the resetting of pitch span can mark off the intonation group boundary.

The pitch span is similar to `register' as defined by Ladd (1990: 38). Ladd distinguishes register from range. In Ladd's terminology, register refers to a band of F_0 values as a subset of the full range, which varies for phrase, sentence, and paragraph boundaries, or for different degrees of local prominence or emphasis. "[R]ange is idealised as constant for a given speaker on a given occasion", it "can vary for truly paralinguistic differences of overall interest, arousal, etc" (Ladd 1990: 38). The pitch span here is defined as a band of F_0 values used to realise the tonal contrast within two fitted lines (i.e., the highest and the lowest fitted lines) varying for intonation groups only. It can also be called the tonal space. An intonation group represents a unit of information which can include paralinguistic factors. The pitch spans shift up and down, and stretch wider and narrower in breaking up utterances into intonation groups.

There can be, but is not necessarily, a pause to mark the intonation group boundary. In most cases, no pause is present. An intonation group consists of one or several feet. One or two feet in one intonation group is usual. An intonation group consisting of four or five feet is generally considered long. An intonation group consisting of four feet is not necessarily longer in actual duration in comparison with an intonation group consisting of two feet.

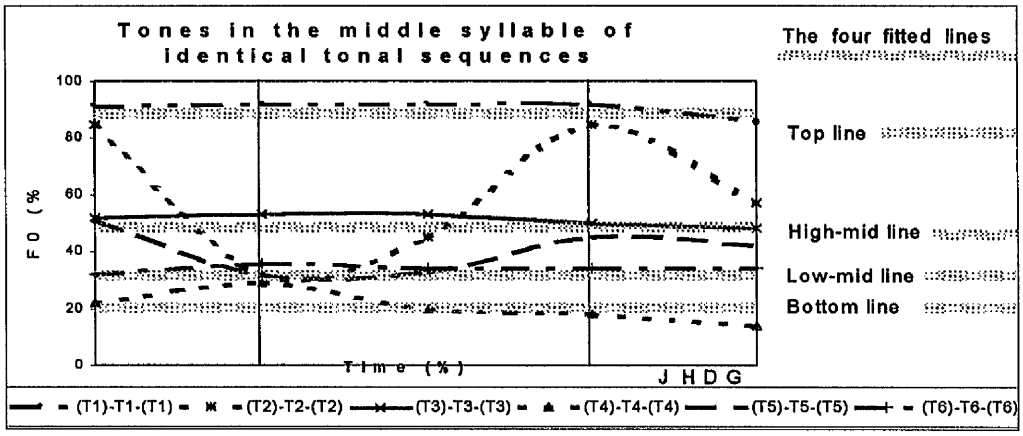
4.4.1.2. The Fitted Lines

An intonation group has a united pitch contour on its own, disjointed from its neighbouring intonation groups. The united pitch contour is a slope line which is a succession of individual lexical tones despite some broken units caused by voiceless segments. The united pitch contour can be represented by several parallel fitted lines: top line, high-mid line, low-mid line and bottom line. The disjointing from one intonation group to another is represented by the disjointing of the fitted lines. The fitted lines are drawn by linking up the salient points of the identical tones in an intonation group. The 'salient point' in the contour of a phonological tone is called the core and is located at the gravity which is usually at the central half of the contour of a phonological tone (see detailed discussion in chapter two), although sometimes it may not be in the central half (for example, when in connected speech the syllable becomes very short). The gravity is where the tones maintain their identities in terms of contour height and shape, and where they are least modified by the neighbouring tones in terms of tonal coarticulation. The peaks in the tonal contours of T1, T2 and T5, the valleys in the tonal contour of T4, and the centres in the tonal contours of T3 and T6 are the salient points. The peaks of T1 and T2 share a similar F_0 value and therefore T1 and T2 are similar tones. In an intonation group, a fitted line linking up their peaks is called the `top line'. On the other hand, as a rising tone, T2 shares a similar F_0 value in its valley to the valleys of T5 and T6. The pitch height in the centre of T5 is very similar to that of T6 whereas the pitch height at the peak of T5 is similar to the centre of T3. Thus, T5 is similar to T6 on the one hand, but is like T3 on the other hand. A fitted line connecting up the peaks of T3 and T5 is called the `high-mid line'. On the other hand, a fitted line linking up the valleys of T5 and T6 is called the `low-mid line'. A fitted line linking up the two T4s is called the `bottom line'.

Here I illustrate a figure of the six phonological tones occurring in the middle syllable of a three identical tonal sequence taken from chapter two, spoken by JHDG. For the sake of demonstration, the four fitted lines defined above are added into the figure although there is not an intonation group. Bear in mind that the F_0 values of a peak and a valley of a phonological tone will not be the same in conflicting as in compatible contexts. Also they can

be affected by other factors, such as vowel height, initial consonant and declination, etc. Therefore, some allowance for variation should be made when fitting the four lines.

Fig. 4.2. The four fitted lines. The six phonological tones occurring in the middle syllable of a three identical tonal sequence. The two vertical lines contain the area of the central portion of the syllabic tonal contours. The four horizontal thick dashed lines represent the added fitted lines.



In an intonation group, such as the above example 25 (/wu:^{6} ɕœ:ŋ^{1} tɕy:n^{1} tɕoŋ^{6}/ (mutual respect)) consisting of two T1s and two T6s, a low-mid line links up the cores of the T6s. (Alternatively, it can link up the lowest ends of the T6s. The result is not much different because T6 is a level tone with slightly sliding shape in its canonical form.) This line is parallel to the top line which links up the peaks of the two T1s. The tonal space between the low-mid line and the top line is very much the same from the beginning to the end. Thus we can use one of the lines serving as a reference line to project another by just making them parallel.

In example 24 (/kɔ:^{3} p^{h}ɐŋ^{4} jɐw^{5} k^{w}a:n^{1} hɐj^{6} how^{2} how^{2}/ (the friendship was very strong)) quoted above, we have three high tones (i.e., T1 and T2 twice) to build up the top line. Below it, the centres of the low mid tones (i.e., T5 and T6) build up a line parallel to the top line as a low-mid line. The low-mid line is right between the high mid tone (i.e., T3) and the lowest tone (i.e., T4). The hierarchy of the tonal height of the six tones is maintained within the intonation group, however, this does not cross the boundary into another intonation group.

21 || $p^{h}ɐŋ^{4}$ jɐw^{5}| k^{w}a:n^{1} hɐj^{6} la:k^{3}||
|| ____ _ | ____ _ _ ||
friend-friend-concern-relate-p.
(on very friendly terms)

The above example 21 shows that there are no identical tones in one intonation group. The centre points of T5 and T6 show that they are similar tones and can be linked up to serve as a reference line for the top line. The top line going through the peak of T1 can be projected as a line above, and parallel to, this reference line (of T5 and T6). The high mid tone (i.e., T3) sits right between the top line and the reference line, and its fitted line (i.e., the high-mid line) can be projected as a line in between, and parallel to, these two lines. The lowest tone (i.e., T4) is far below at the bottom, and its fitted line (i.e., the bottom line) can be projected as a line below, and parallel to, the reference line.

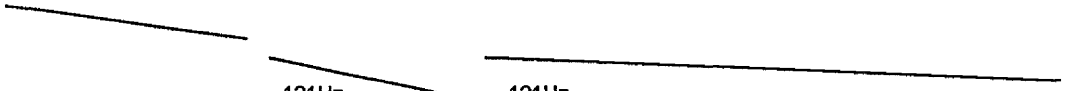
There is a maximum of four fitted lines in an intonation group, i.e., top line, high-mid line, low-mid line and bottom line. The four lines are parallel to each other. When linking up identical/similar tones, we must consider the factors which may affect the tonal value, such as a fundamental frequency difference due to the intrinsic value of vowel height, a tonal contour disturbance from the plosive onset of the syllable, a tonal contour dissimilation in height due to the effect of tonal coarticulation, etc (for details see discussion in chapter two). Therefore, a local disturbance event caused by such factors may have to be overlooked. The following shows the up and down tonal contour of the four T1s in example 23 and the three T6s in example 64. The tonal contours of the individual tones are not in a straight line at all as they might be in isolated situations, the ups and downs in the lines can be as much as 5Hz in each of them:

23 || ɕi:^{1} ɕɐŋ^{1} | k^{w}a:n^{1} hɐj^{6} nɛ:^{1} ||%
|| ___ _ | ____ _ _ ||
teacher-student-concern-relate-p.
(teacher-student relationship)

64 || jɐt^{6} tɕi:m^{6} | ha:^{6} | kɔ:ŋ^{3} || %
|| _ ___ | ! | ! ||
day-gradually-down-sink
(on the decline)

In some cases, the resetting of pitch span is not always clearly illustrated because the fitted lines are not clear. For example:

28-30



||mow^{5}|ji:^{4}ka:^{1}|kɔ:^{2} ti:^{1}|| ka:w^{3} ɕi:^{1}| pej^{3} jy:^{4}|| ka:w^{3}| ka:w^{3}|iɛ:^{6}| p^{h}eŋ^{4} wu:j^{2}| k^{h}ɵy^{5} tej^{6} wa:^{6}||
|| ___ |_ ___ | ___ _ || ____ _ |_ ___ || ____ | ___ | ___ | ____ _ | _ _ ____ ||
have not-now(2)-that-pl-teach-expert-example-as-teach-teach-em-assess-committee-he-pl-say
(Unlike those teachers nowadays, for example, the Body of Educational Assessment say)

The high tones in the three top lines of the three intonation groups are very close in the above example. The top line of the first intonation group is linked up by three high tones but that of the following intonation group is projected by paralleling the high-mid line which links up the two T3s. Likewise, the top line of the last intonation group is projected by paralleling the high-mid line which links up the two T3s. The projection in the middle intonation group is carried out by drawing a line through the peak of T1 parallel to the high-mid line which links up the two T3s. Likewise, the projection in the last intonation group is carried out by drawing a line through the peak of T2 parallel to the high-mid line which links up the two T3s and also parallel to the low-mid line which links up the three T6s. We find that the top line in the intonation group 30 disjoints from that in the intonation group 29 by rising up a bit, and the top line in the intonation group 29 disjoints from that in the intonation 28 by lowering a bit. Furthermore, the high-mid line which links up the two T3s, in the intonation group 29 is lower that that in the intonation 30. Moreover, the three bottom lines which are projected by drawing a line through the lowest point of T4 parallel to the top line of intonation group 28 and parallel to the high-mid lines of intonation group 29 and 30 are not in line with one another. As a whole, not only the high-mid line in the intonation group 30 is higher than that in the

intonation group 29, so are the top lines and the bottom lines. As we defined that only the disjointed fitted lines break up a continuous speech into small units – intonation groups, it would be a mistake to overlook the uplifting of the fitted lines in the intonation group 30. The foot /pej^{3} jy:^{4}/ seems to hover precariously between the two last intonation groups. The phonetic value of the core of the T3 /pej^{3}/ is 119Hz which is very near to that of the previous T3 (121Hz) and also that of the following T3 (121Hz). What makes this foot belong to the intonation group 29 is that its T4 /jy:^{4}/ is not in line with the following T4. The valley of /jy:^{4}/ is 80Hz, which is lower than that of the following T4 /p^{h}eŋ^{4}/ (83Hz). A foot cannot cross the boundary of an intonation group. Within an intonation group, the top line must be parallel with the bottom line and the middle line(s).

4.4.1.3. Tonal Contrast

Within an intonation group, the gravity of the individual tones maintains their tonal contrast in terms of height: T1 and T2 are the highest while T4 is the lowest, T3, T5 and T6 are in between (T3 is higher than T6, and the rising slope of T5 intersects the two). For example, in example 21 (/p^{h}ɐŋ^{4} jɐw^{5} k^{w}a:n^{1} hɐj^{6} la:k^{3}/ (on very friendly terms)) above, the F_0 values of the cores of T1, T3, T4, T5 and T6 are 120Hz, 86Hz, 78Hz, 95Hz and 83Hz, respectively.

The shape of the contour also remains despite some modification caused by local events, for example, by an aspirated plosive consonant onset; or, by tonal coarticulation. Take the intonation group 95 and 96 for example, the tonal sequence is T2-T4-T6-T1-T6. T2 raises its peak to the extreme high. The rising peak is near the offset of the syllable and is as expected (see discussion in chapter two). The following T4, whose onset is built up at the high offset of T2, quickly pulls down to its extreme low. Thus the falling slope of T4 is steeper than usual. It reaches the valley approximately 10 percent of the duration before its offset, which is the transition to the low level T6. The following T6, starting its onset from the low offset T4, quickly rises to its canonical height before falling slightly. The rise does not contaminate the identity of T6, being merely the carryover coarticulatory effect. The beginning of the tonal contour of the following T1 is broken up by its voiceless onset consonant /tɕ^{h}/ but displays a

high falling slope at the vocalic portion. This indicates that it comes back to its canonical height despite being preceded by a tone with a mid-level pitch offset. Afterwards, it falls down quickly in preparation for the transition to the following low-level T6. The voiceless affricate consonant onset /tɕ/ of the following T6 blocks the continuous slope line from the preceding T1. The contour resumes with a gentle fall.

95-96 || keŋ^{2} ji:n^{4} hɐj^{6} || tɕ^{h}ɵt^{1} tɕɔ:j^{6} ||
|| ____ _ _ || __ ___ ||
unexpectedly-look-to be-out-at
(and a very surprising one coming from)

Negative slope (i.e., a phonological rising tone appears to be a falling slope, or vice versa) is found in the present data but is extremely rare. Fok (1974) and Kong (1980) tell us that their informants find it very difficult to tell T3 and T5 apart, particularly referring to the syllable containing the segments /ɕi:/. Killingley (1985) suggests that these two phonological tones have merged in Malaysian Cantonese. Cheung (1986: 213) dismisses Killingley's suggestion for mainstream Cantonese. My informants show no difficulty in telling them apart in any tonal environments in reading texts. The occurrence of negative slope in T5 (5 out of 24) is much higher than that of T3 (2 out of 52) in the present data. They all occur in a middle position of conflicting tonal environments and the entire lexical contour appears as a transition period: the contour of T3 becomes rising and no level is found, the contour of T5 becomes falling and no rising peak is found. Only the lexical items /ŋɔ:^{5}/ (I) and /k^{h}ɵy^{5}/ (he) occur with negative slope twice. The particular tonal environments, and the low occurrence of the negative slope and of the related lexical items do not blur the distinction between the two tones nor is there enough evidence to suggest a variant form for them.

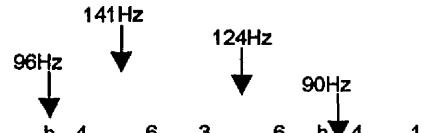
4.4.1.4. The Highest and the Lowest Lines

The highest and the lowest lines in an intonation group can be any of the four fitted lines. In the examples 28, 29 and 30 above, the highest line and the lowest line are the top lines and the bottom lines, respectively. In the example 23 above, there is only a top line and a low-mid line in the intonation group. Therefore, the two lines are the highest line and the lowest

line, respectively. Note that the term `the lowest line' has a different sense here than the `base line' in the sense in which the latter is used by some other researchers. For example, the end of the base line refers to the lowest range which a speaker's voice usually reaches. The phenomenon of parallels between the highest line and the lowest line in Cantonese differs from the general features of the range of intonation contour observed in some non-tone languages. It is claimed that there is a tendency for the F_0 range to diminish as a function of time in one breath-group in a number of languages, as summarised by Vaissière (1983: 55). Ladd (1984: 54) also illustrates figures of a string of tones in a hypothetical tone language utterance, showing that the range of pitch for tonal sequences of whatever structure, e.g., MHLHMHLM or LLMMHH, becomes gradually narrower. In discussing the delta programme in Bambara, a Mande tone language spoken in Mali, Hertz (1990: 245) shows a similar picture of a gradually narrowing span between a top line and a base line in a diagram of a sentence. In contrast, in Cantonese, the highest line and the lowest line used for the range of pitch for tonal contrasts are straight and parallel. The six Cantonese tones are constituted not only by the pitch orientation but also by the pitch height, they need enough space in between each other for the purposes of maintaining their phonological contrast. Gradually diminishing the pitch span within an intonation group would result in diminishing the tonal contrast.

4.4.1.5. F_0 Maximum and F_0 Minimum

Another feature of the intonation contour in a number of languages, as summarised by Vaissière (1983: 55), is a tendency for the maximal value of the F_0 to be located on the first prosodic word of the sentence. However, in Cantonese, the actual slope of the contour which is made up by a string of tones is the reality of the intonation contour shape. Therefore, in a tonal sequence of T4-T1, the F_0 minimum is predicated at the first syllable and the F_0 maximum at the second syllable (see discussion in chapter two). The following example illustrates that a high tone T1 is at the end of the intonation group but retains the higher F_0 value compared with other tones within the intonation group:



67 || $t^{h}ɐj^{4}$ mok^{6} kɛ:^{3} | ti:w^{6} $tɕ^{h}$a:^{4} tow^{1} ||

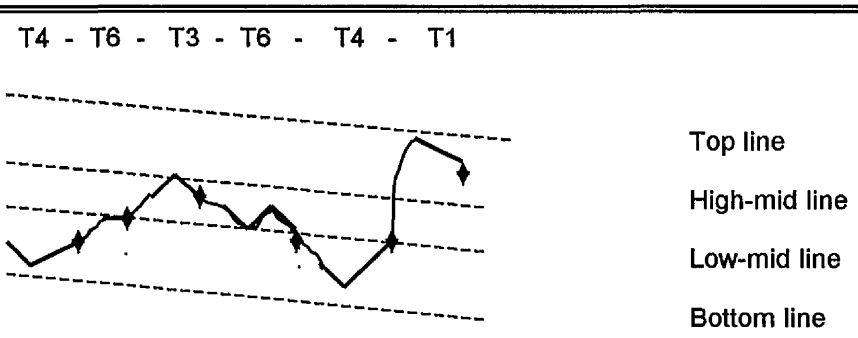
|| ____ _ _ | ____ _ _ ||

title-topic-p-investigate-check-all

(the topic of this investigation)

We can draw a figure reflecting the slope of the contour in the above example by simply drawing the tonal pitch of the single tones in succession:

Fig. 4.3. The tonal contrast in an intonation group represented by the four fitted lines.



The high end of the slope of this intonation group is a high tone T1. The arch shape of the slope preceding the high end is made up by the sequence of low falling tone T4, low-level tone T6, mid-level tone T3, then reversed to a low-level tone T6 and a low-falling tone T4. This lexical tonal contour sequence clearly indicates that the tonal contour of phonological tones in succession is the property of the intonation contour in Cantonese. The up and down movements of the slope line are formed by the phonological tonal contour in sequence. The phonological tones tend to preserve their phonological tonal value. Thus, within an intonation group, the fitted lines must be generated by linking up the phonologically equivalent tones (or similar tones). This is very different from a non-tone language, in which the top line and the base line are specified by the prominent points in the slope of the contour, usually the peaks and the valleys, for example, as in Spanish and French (e.g., Vaissière 1983: 55); or different from a level tone language like Hausa (e.g., Lindau 1985: 26), in which the top line and the base line are linked up by the peaks of the H tones and the valleys of the L tones, respectively.

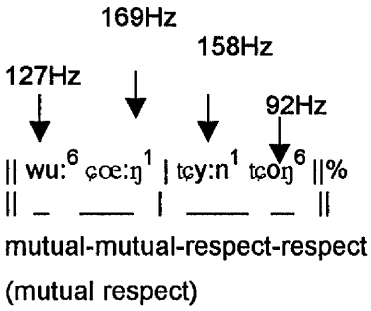
4.4.2. Declination

4.4.2.1. Defining Declination

The pitch span in an intonation group demonstrates a phenomenon of declination. Declination refers to the fundamental frequency having a tendency to decline gradually during the course of utterances (Ladd 1984: 53).

Declination is an F_0 decreasing phonetic process in Cantonese that is not subject to any particular tonal sequences. It differs from such phonological processes as downstep or downdrift in African languages. For example, Hyman (1975: 226) claims that an alternation of H-L-H sequences is subject to downdrift in a number of languages, for example, Igbo, Hausa and Twi, etc. This phenomenon, representing an automatic lowering process, applies progressively to each H preceded by a L. A sequence of H-L-H is not realised as [¯ _ ¯], but rather as [¯ _ —]. This contrasts with the phenomenon in identical tonal sequences in most African languages. For instance, a sequence of H tones is realised as [¯ ¯ ¯] rather than as [¯ — _] (p.228). On the other hand, downstep can be predicted morphophonemically in many cases, "in which a lowered H receives phonemic status when a L which `conditions' downdrift is lost (either through deletion or through assimilation)" (p.227). A standard example is cited in Twi (Fromkin 1972: 57): the realisation of /mí ɔ bú/ (my stone) (H-L-H) is /mí ˡbú /. By downdrift, the second H is lowered. At this point, a rule of vowel deletes the L /ɔ/ and the result is H-ˡH. That is, a H followed by downstepped ˡH, which is contrasted with H-H and H-L. Downstep and downdrift are phonologically or morphophonemically predictable. Pierrehumbert (1980) adopts the term `downstep' traditionally used in African studies in her description of the intonation pattern of English: any bitonal accent (indicated by a plus sign), H*+L, H+L*, L+H* and L*+H (the asterisk mark is associated with a stressed syllable), triggers the lowering of the pitch of the following H (p. 152). English downstep is conditioned by the morphological organisation of the intonation; it takes place in sequences of the form H+L H and H L+H, but not in other alternation tonal sequences (p.150). Examples are given in the utterance `an orange ballgown' with (a) H*H*LL% - standard declarative intonation; (b) H*+L H*L L% - a downstepping accent on orange; (c) L*H*LL% - surprise-redundancy contour (Beckman and

Pierrehumbert 1986: 257). Downstep only occurs in example (b) where the bitonal H*+L causes the following H* to have a lower F_0 value than the preceding H*. However, declination in Cantonese is not triggered by any phonological sequence of tones nor determined by any linguistic factors. It is an automatic lowering process present all the time during the course of speech and applies to all tonal sequences. The 'lowering' is based on the comparison of the F_0 value between identical tones or similar tones within the intonation group in question.



Phonologically equivalent tones decrease their phonetic values when they appear later in the same intonation group. Take the example from intonation group 25: the top line and the low-mid line appear as a parallel downward shape, i.e., the tonal peak of the second T1 (158Hz) is lower than that of the first T1 (169Hz) and the lowest point of the first T6 (127Hz) is higher than that of the second T6 (92Hz). The prominent point of the first syllable tone is higher than the second syllable tone if the two are phonologically equivalent tones. In example 67 which we discussed in the previous section, there are two T4s and two T6s. The F_0 value is 96Hz in the core of the first T4 and 90Hz for the second, and 141Hz for the first T6 and 124Hz for the second. The identical tones have no identical F_0 value in the same intonation group: the later a tone occurs, the lower the F_0 value it has. The fitted lines of the pitch span used for linking up the identical tones descend gradually as shown in the figure above.

The downward trend of the fitted lines representing the pitch span of the intonation group is a matter of fact, present in all intonation groups. It can easily be discerned by visual inspection, rather than requiring precise acoustic measurement. If a rising slope occurs, it perhaps indicates the F_0 maximum located on a T1 which is at the end of the slope of the contour as in example 67, but it does not indicate the absence of declination. As I argued before, the slope

line which is formed by successive single phonological tonal contours is the intonation contour. If there are three identical/similar tones in an intonation group, the identical tones A, B and C are linked up by a fitted line, where B is lower than A and C is lower than B.

We are not very sure about the cause of the universal declination. The presence of declination in every single intonation group in Cantonese strongly supports the view of an automatic process -- it is due to interaction between the larynx and the respiratory system (Lieberman 1967, Collier 1975 and 1983, Cohen *et al* 1982). The physiological explanation suggests that the drop in F_0 is a result of the natural falling off in subglottal pressure that accompanies expiration in speech. Subglottal pressure and laryngeal tension are the principal components of output F_0 (as discussed in chapter two and three). Obviously speakers can control the resetting of the pitch if they want to, as they can raise their voice or can adjust which phonological tone they want to produce. Speakers can also control the falling or rising of tonal movement as they like, as when they produce a falling tone or a rising tone. Speakers can also drop their voice at the end of an utterance as if they have lost interest. However, speakers do not seem to be able to control the declination.

4.4.2.2. Identical Tonal Sequences

Despite the present data clearly showing that the phenomenon of declination exists in all intonation groups which consist of a variety of tonal combinations, a small experiment was conducted to strengthen the argument. The experiment was designed to investigate whether declination exists in identical tonal utterances. Each tone involved three or four utterances of different length, each utterance consisting of four to eleven syllables. Details of the utterance data can be found in appendix three. Two of the informants (JHDG and LTHJ) used in the experiments of tonal coarticulation read the designed utterances, with three repetitions, in a sound proofed booth at SOAS. The recording was made with a laryngograph device as described in chapter two. The recording procedures and the method of segmentation were the same as that used in experiments of tonal coarticulation. The informants were free to choose how to group the syllables into feet and where to define the intonation groups. Measurements of the F_0 values were taken only from the centre point of T1, T3 and T6, and

the valley of T4. The two rising tones, T2 and T5 were measured at two points: the peak and the valley. Both time and the F_0 data are normalised. The results strongly support the claim made in this section: declination exists everywhere in Cantonese (as displayed in figures 4.4 - 4.9).

4.4.2.3. Long and Short Intonation Group

It is said that in longer utterances a slower declination rate is found invariably (Maeda 1976, quoted by Ladd 1984: 57). Hertz (1990: 245) suggests that in Bambara, a two tone language, the starting and ending frequencies of the baseline are relatively independent of the sentence duration while the starting and ending frequencies of the topline are a function of sentence duration. Laniran (1993: 203) presents evidence that declination in Yoruba is dependent on tone type: L tones have the steepest slopes and H tones the shallowest. On the other hand, it is reported that "calculation of the exact rate of declination is a difficult task" (Vaissière 1983: 56). Not only can the general tendency of declination of an intonation group be easily detected by eye, but also a steep declination line can easily be distinguished from a gentle declination line by visual inspection. Certainly it is difficult to find a long intonation group with a very steep slope, but it is not uncommon for short intonation groups to be presented with a steep or gentle slope. Data shown in this study clearly indicate that the slope of the declination line is not simply decided by the length of the intonation group. Here are two pairs of examples, one short and one long, each pair lasting for a similar length of time – 510ms ~ 487ms for the short pair (examples 89 and 91, respectively); and 1193ms ~ 1187ms for the long pair (examples 61 and 98, respectively). The short pair both consist of four syllables in two feet; whereas the long pair both consist of 7 syllables one in three and one in four feet. The different slope of the declination line in each pair can be easily caught by eye in terms of relative gentleness or steepness.

Short intonation groups:

89 || mow^{5} tɐk^{1} | ka:w^{3} lɔ:^{3} || ⟶ (510ms)
|| ___ _ | ___ _ ||
have not-able-teach-p.
(unteachable)

Fig.4.4. Mean F_0 contours of T1 occurring consecutively showing the pitch of identical tones descending gradually.

Fig. 4.5. Mean F_0 contours of T3 occurring consecutively showing the pitch of identical tones descending gradually.

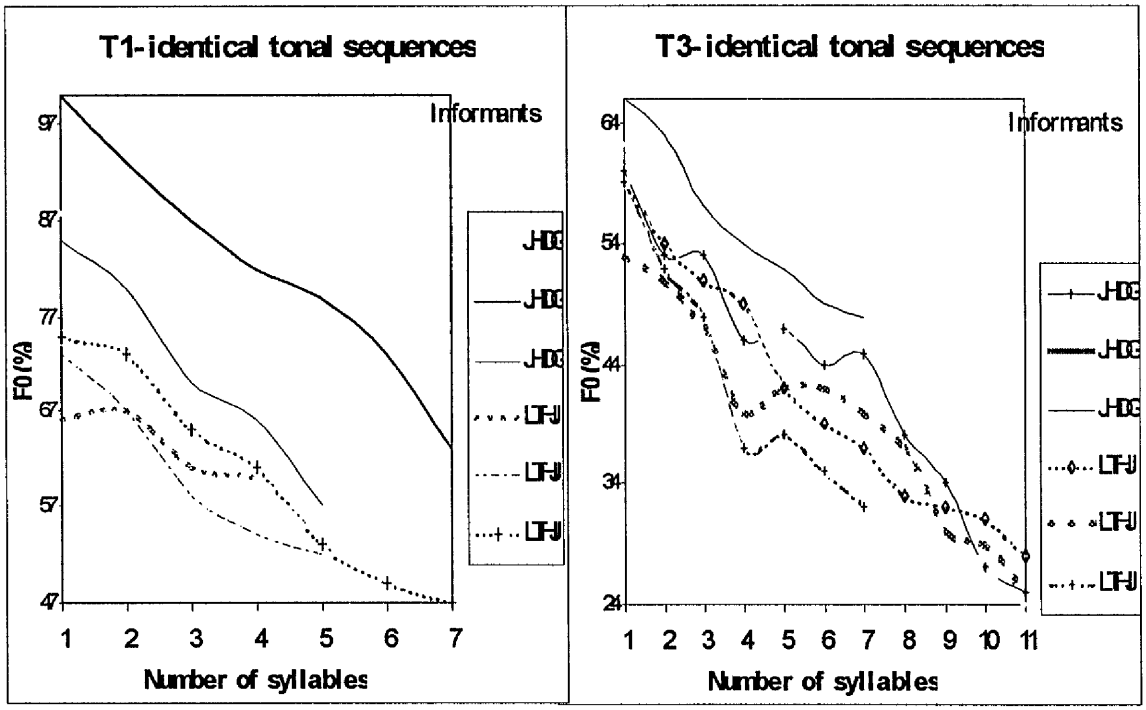


Fig.4.6. Mean F_0 contours of T4 occurring consecutively showing the pitch of identical tones descending gradually.

Fig. 4.7. Mean F_0 contours of T6 occurring consecutively showing the pitch of identical tones descending gradually.

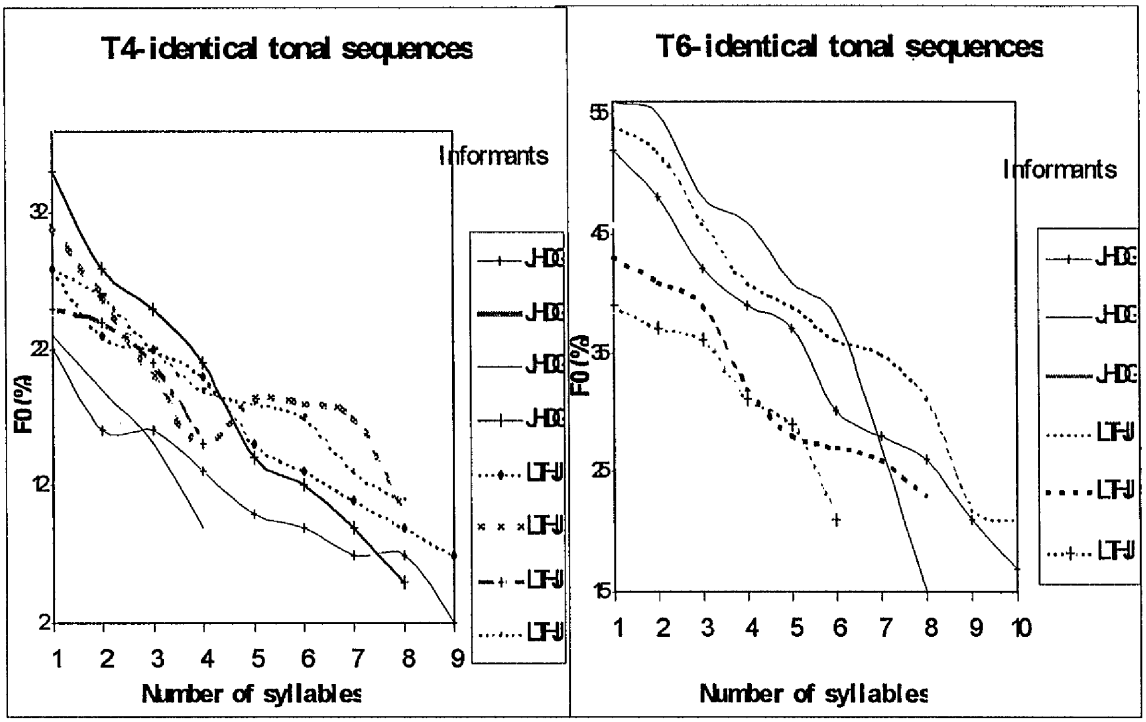


Fig.4.8. Mean F_0 contours of T2 occurring consecutively showing the pitch of identical tones descending gradually.

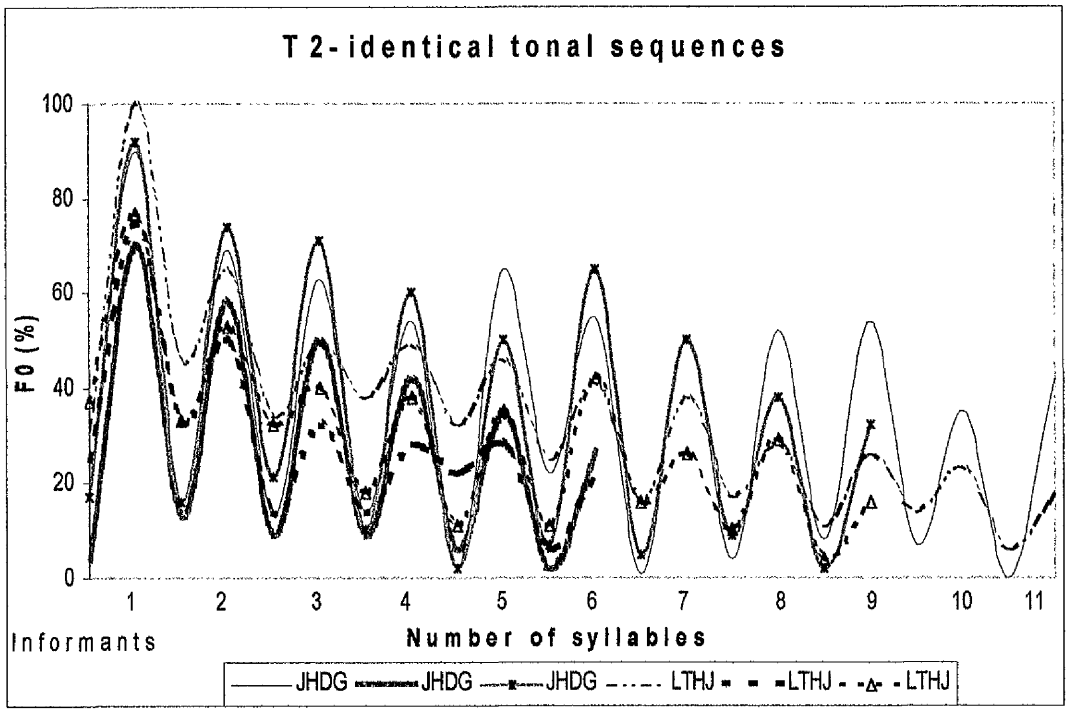
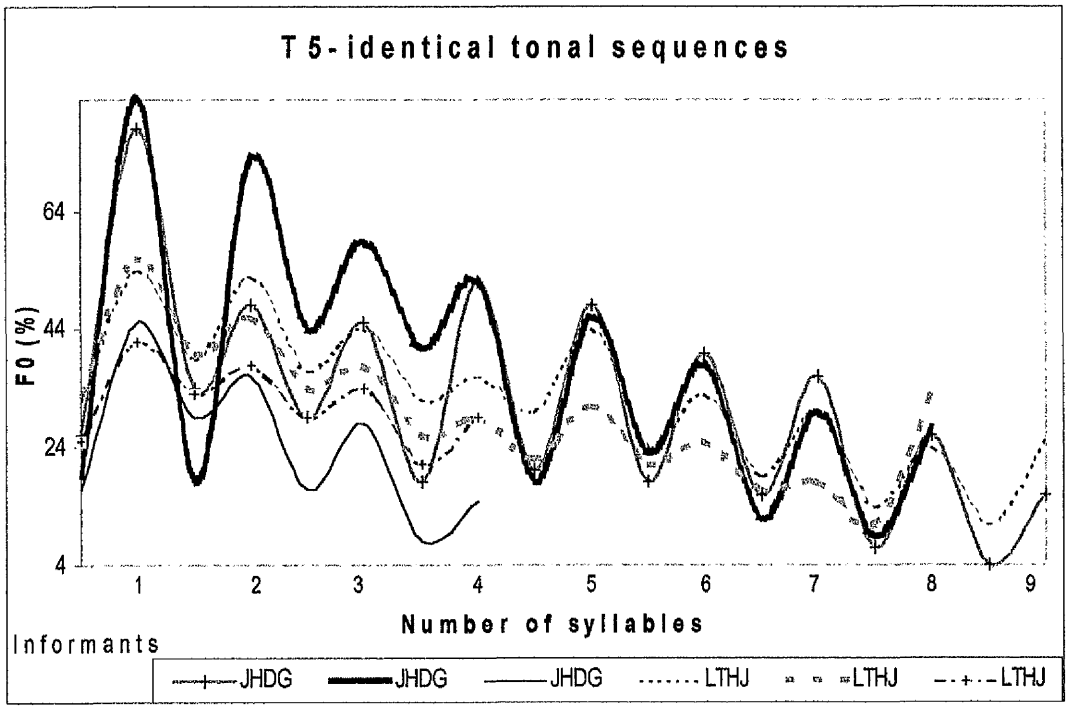


Fig.4.9. Mean F_0 contours of T5 occurring consecutively showing the pitch of identical tones descending gradually.



91 || kɐm^{2} k^{h}ɵy^{5} | ti:m^{2} a:^{3} ||% ⟶ (487ms)
|| _ ___ |_ ____ ||
in that case-he-how-p.
(So, what will he do then)

Long intonation groups:

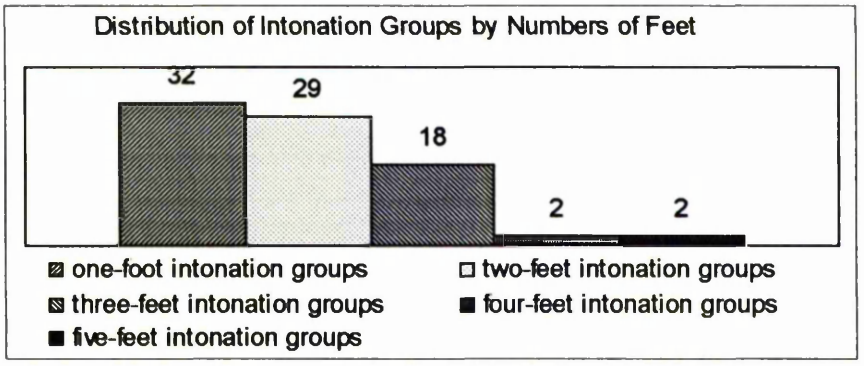
61 || ji:^{4} ka:^{1} | ŋɔ:^{5} tej^{6} | ka:w^{3} ɕi:^{1} nɛ:^{1} || ⟶ (1193ms)
|| ___ _ | _ ___ | ___ _ _ ||
now(2)-I-pl.-p. -teach-expert-p.
(nowadays (the authority of) us teachers)

98 || ni:^{1} kɔ:^{3} | ka:w^{3} jok^{6} | p^{h}eŋ^{4} ji:^{5} | wu:j^{2} || ⟶ (1187ms)
|| _ __ | ___ _ | ___ _| ___ ||
this-m.-teach-cultivate-assess-discuss-committee
(the Body of Educational Assessment)

I have made no attempt to calculate the declination rate in each intonation group. The slope of the declination line seems to be affected by many factors, for example, the length of the intonation group, the position of the intonation group in the whole utterance, the type of sentence (command, request, etc), the tempo, etc.

There are 83 intonation groups in our data for acoustical analysis containing 162 feet whose distribution is displayed in the following figure.

Fig. 4.10. The 83 intonation groups distributed by numbers of feet.



4.4.2.4. Declaratives and Interrogatives

It is reported that declination can be suppressed in making interrogatives for a number of languages (Vaissière 1983: 57), and questions requiring answers other than `yes' or `no' may even form a raising tune in Chengdu (Chang 1958: 78). Lindau (1985: 37) uses yes-no questions and question-word questions in Hausa for experiments, and reports that "all questions are specified with zero slope" (1985: 37). However, the interrogatives in the present data all manifest declination as do declaratives. Here are examples of `wh-' interrogatives which contain a question indicator `why' (/ti:m^{2} ka:j^{2}/) or `how' (/ti:m$^{2}$/) followed by either final-particles `/ka:^{2}/" or "/a:^{3}/':

69-70 || ka:w^{3} ɕi:^{1} | ti:m^{2} ka:j^{2} | ji:w^{3} jɐw^{5} || kʰy:n^{4} wej^{1} | ɕi:n^{1} tɐk^{1} | ka:^{2} ||
|| ____ _ |___ _____ |____ __ || ____ __ |____ _ |___ ||
teach-expert-how-explain-need-have-power-might-until-o.k.-p.
(Why do teachers need to have authority anyway?)

91 || kɐm^{2} kʰɵy^{5} | ti:m^{2} a:^{3} ||%
|| _ ____ |__ ____ ||%
in that case-he-how-p.
(So, what will he do then?)

One can argue that, it is common for an interrogative to display a falling trend if it contains a question word (i.e., a `wh-' word), as in English. The following example is a rhetorical interrogative without a question word but with a final-particle /a:^{4}/. It also displays the phenomenon of declination:

92 || tɕow^{6} | ka:j^{1} tʰoŋ^{4} | a:^{4} ||%
|| ^ __ |____ __ |__ ^||%
to be-street-child-p.
(He'll be out on the streets?!)

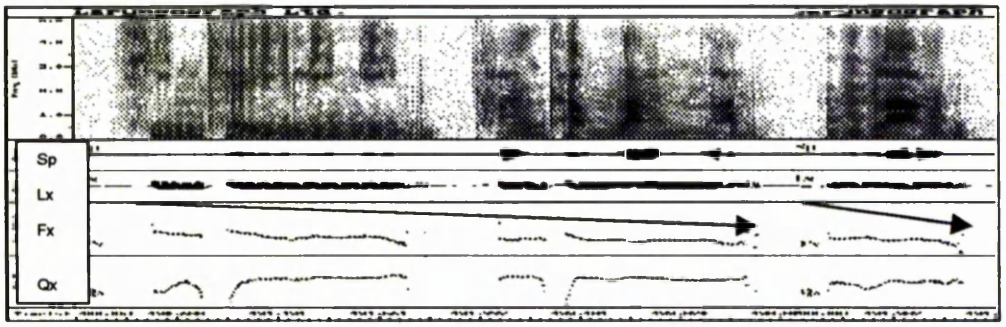
4.4.2.4.1. Yes-No and Echo Questions

The following is a constructed yes/no question consisting of a string of T4s. Yes/no questions in Cantonese are choice type questions giving a choice of `yes' or `not yes'. A final-particle /a:^{3}/ is often attached at the end of the utterance but not obligatory. The invented utterance

employs T4 only for the sake of using identical tones with the negation /m^4/. No final-particle is used; this is also for the sake of not distorting an identical tone sequence. Two informants (JHDG and LTHJ) read out the yes/no questions and their corresponding answers which are illustrated below in figures 4.11a and 4.11b, respectively.

Fig. 4.11 (a -b). Extracted F_0 contours (Fx) of the question and answer (previous-period-year-year-bound-piano-difficult-not-difficult, not-difficult `Did you find it difficult to play the piano at the previous stage? –No'), spoken by JHDG for (a) and LTHJ for (b). The arrows indicate the declining F_0 over the utterance. The F_0 values are taken from the valley of the tones.

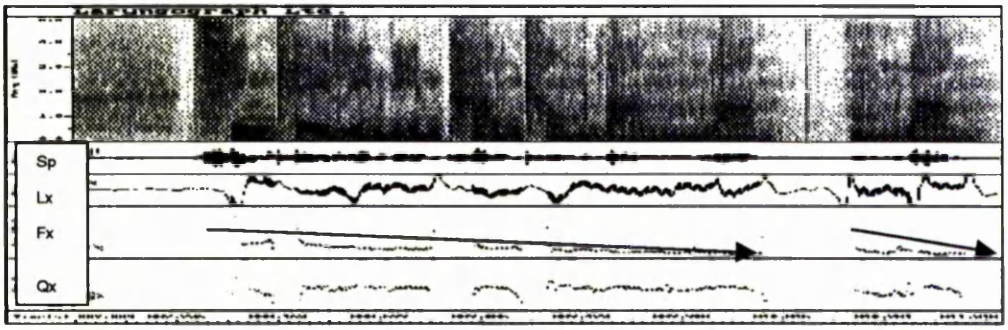
4.11.a.



/tɕ^{h}i:n^4 k^{h}ej^4 ni:n^4 ni:n^4 t^{h}a:n^4 k^{h}ɐm^4 na:n^4 m^4 na:n^4/ /m^4 na:n^4/

F_0 (Hz): 139- 137- 129- 127- 119- 125- 111- 113- 110 122- 120

4.11.b.



/tɕ^{h}i:n^4 k^{h}ej^4 ni:n^4 ni:n^4 t^{h}a:n^4 k^{h}ɐm^4 na:n^4 m^4 na:n^4/ /m^4 na:n^4/

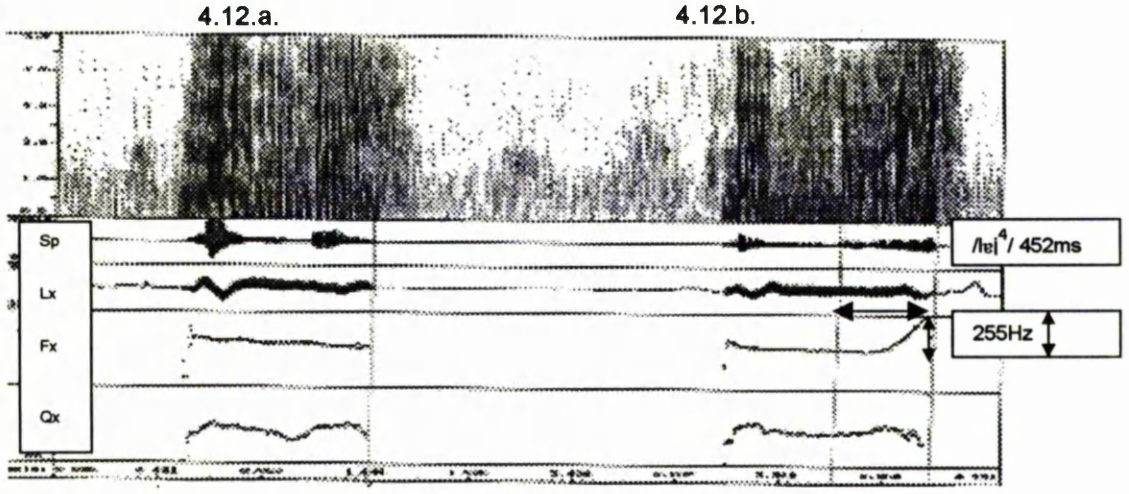
F_0 (Hz): 115- 112- 103- 103- 103- 96- 93- 91- 90 98- 93

The downward trend line in these questions and answers is easily detected. In the utterance spoken by JHDG, there is a 29Hz drop from the beginning to the end of the question and only a 2Hz drop in the answer; in the utterance spoken by LTHJ, there is a 25Hz drop in the question and a 5Hz drop in the answer. The evidence from the two informants shows that the amount of declination does not favour the view of suppression of declination in interrogatives.

This provides evidence to support the claim above that declination operates in both declaratives and interrogatives in Cantonese.

However, it is often said that a declarative can be turned into a question simply by using a rising pitch, for example, when the sentence is spoken in doubt or suspicion. According to this view, no final-particle is used to indicate an interrogative, but rather the distinction between an interrogative and a declarative is totally dependent on the tonal contour, i.e., one is rising and the other is falling. This view would also imply that declination may be absent in this type of interrogative. It would lead one to expect that declination in Cantonese is not an automatic phonetic property but depends on a linguistic choice between the different linguistic functions (i.e., interrogative and declarative). In order to investigate whether declination is present in interrogatives which are heard to be rising, I constructed a pair of utterances and had informant JCCF read them out. The result is displayed in figure 4.12. Each one of the pair of utterances consists of identical words, but one is expected to be falling and one is expected to be rising. The utterances consist of phonologically equivalent tones and no other tones interfere.

Fig. 4.12 (a-b). Extracted F_0 contours (Fx) of a Yes/No interrogative and an echo interrogative respectively, spoken by JCCF. The words of each of the two utterances are /lɐj^4 m^4 lɐj^4/ (come-not-come) which is interpreted as `[Are you] coming?' for (a) and `[Did you ask me if I am] coming or not?' for (b). In (b), the horizontal double-headed arrow indicates the duration of the last syllable and the vertical double-headed arrows indicate the rise from the valley to the peak of the syllable. The vertical lines mark the boundaries of the last syllable.



Downward trend in F_0 contours is clearly displayed in different length and different types of utterances in figures 4.4–4.12. The F_0 contours of yes-no interrogatives in figures 4.11a-b and 4.12a read out by different speakers display a uniform downward trend. The downtrend line in the figure 4.12b clearly demonstrates that declination is also present in an echo interrogative which is normally expected to be rising. A rising contour can be present but it is a local event affecting only the last syllable of the utterance. This finding sustains Fok's claim. Fok (1974) conducted an experiment using one word interrogative sentences and claimed that all tones end with a rising tail, and end at about the same frequency region – they become less distinct; however, they still "maintain their initial frequency distinction". "The starting frequency of T1 is the highest, then comes T3 followed by T2, T4, T5 and T6" (p.29). I am not sure whether her term `initial frequency' is the same thing as the `starting frequency' or refers to the frequency at a certain length of time from the starting point. By examining the figures she offers, the pattern of relationships among the tones is shown as maintaining the contrast in the first half (or less than three quarters) of the tonal duration. If this observation is valid, her term `rising tail' makes sense: the tail is the last half (or slightly more than one quarter) of the tonal duration. She conducted another experiment by putting the test word in the penultimate syllable of an invented sentence and claims that "[t]he tones said in isolation all give an upward turn as the rising intonation in this case can be passed on to nothing else. Tones extracted from sentences conform more to their basic structure as the rising tail is shifted to the last word" (p.31). In my view, a word when expressing query, doubt or suspicion ending with a rising tail does not necessarily induce a `rising intonation' in the system of Cantonese intonation. A rising tune is proposed by some researchers in some East Asian tone languages: Chang (1958) for Chengdu Sichuanese, Ho (1977) for Mandarin, Luksaneeyanawin (1983) for Thai, and Dung *et al* (1998) for Vietnamese (for detailed discussion see above). Chang, Ho and Dung's rising tune is realised as the perturbation of the final syllable. Luksaneeyanawin's Tune 2 (rising tune) affects the rising pitch contour of the tonic syllable, and the tonic syllable is usually located at the end of a tone group. If I am not mistaken, the rising intonation contour in those languages suggests that almost nothing happens until the final syllable. Certainly there are various modifications of the final tone of

an utterance. A rising tail attaching to the end of some forms of interrogatives is probably not enough to justify the conclusion that there is a rising tune in Cantonese or that declination is suppressed in interrogatives in Cantonese. My data show that the rising tail with various modifications occur on the last syllable of utterances. Apart from signalling some forms of interrogatives (e.g., echo interrogatives as in figure 4.12b), it can also signal declaratives with other attitudes, such as ironic or impatient, etc. I shall demonstrate this in the following section 4.4.2.4.2.

4.4.2.4.2. Rising Tail

Here are examples of declaratives with different attitudes: confirmative, ironic and impatient confirmation in figures 4.13, 4.14 and 4.15, respectively (spoken by JHDG).

Fig. 4.13. Extracted F_0 contours (Fx) of a confirmative declarative '/tɕ^{h}i:n^4 k^{h}ej^4 ni:n^4 ni:n^4 t^{h}a:n^4 k^{h}ɐm^4 m^4 na:n^4/ (previous-period-year-year-bound-piano-not-difficult - It was not difficult to play the piano in the previous stage)', spoken by JHDG. The arrow indicates the declining trend of the F_0 contour.

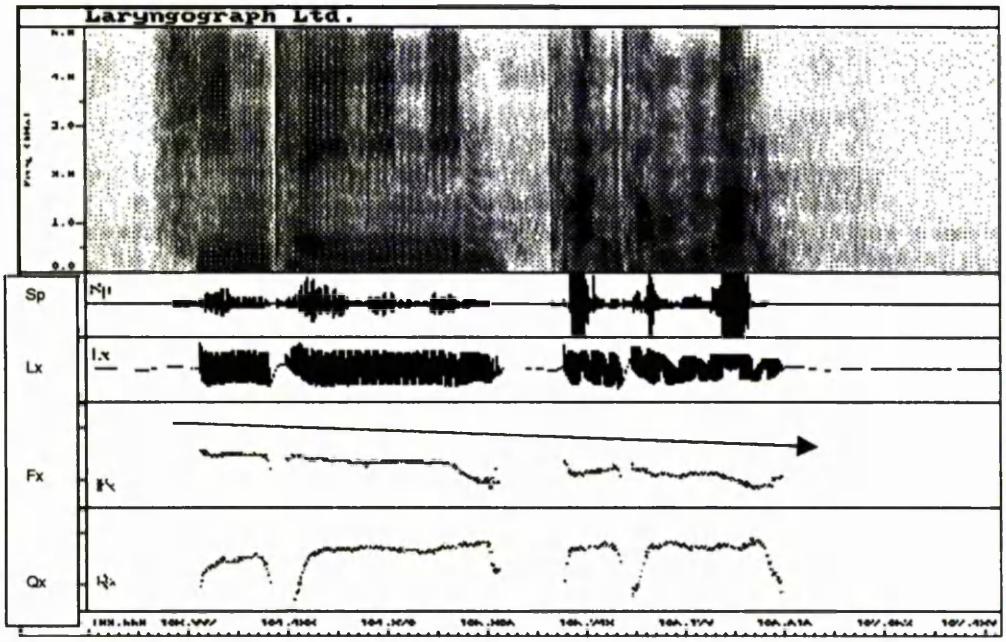


Fig. 4.14. Extracted F_0 contours (Fx) of a declarative of the same words as in figure 4.13 but with an ironic expression, spoken by JHDG. The arrow indicates the declining trend of the F_0 contour. The horizontal double-headed arrow indicates the duration of the last syllable. The upward arrows indicate the rise from the valley to the peak of the syllable and the downward arrows indicate the fall from the peak to the offset of the syllable.The two vertical lines mark the boundaries of the last syllable.

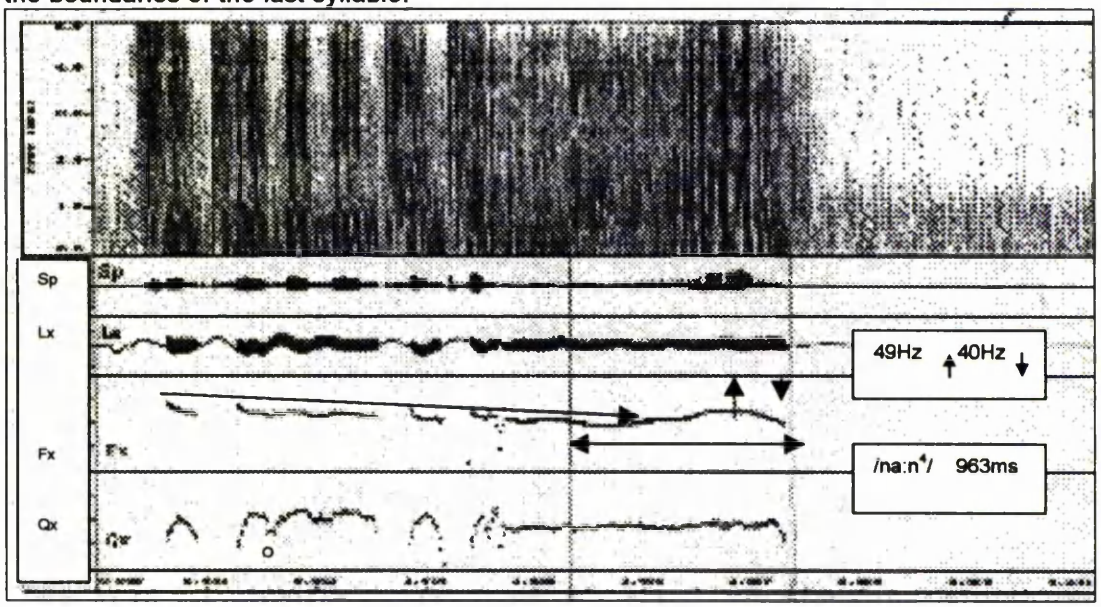
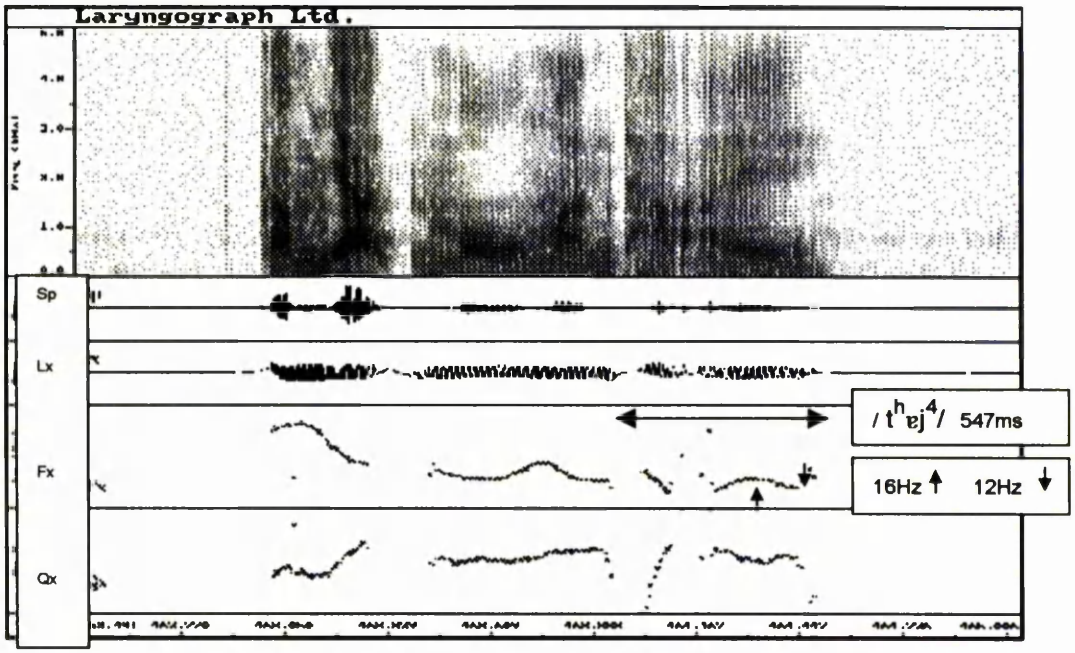


Fig. 4.15. Extracted F_0 contours (Fx) of a declarative with impatient confirmation '/kɐm^{1} jɐt^{6} mow^{5} mɐn^{6} t^{h}ɐj^{4}/ (to-day-not-ask-topic - there is no problem today)', spoken by JHDG. The horizontal double-headed arrow indicates the length of the final syllable. The upward arrows indicate the rising peak from the valley of the syllable and the downward arrows indicate the fall from the peak of the syllable.



In contrast to that in ironic expression in figure 4.14, the rising slope of the last syllable in impatient confirmative declarative in figure 4.15 does not rise as much. The rise soars to 49Hz for the former, but only 16Hz for the latter. Besides the noticeable difference in the amount of the rising between the two, the difference in duration of the two attracts more attention: the mean duration of the preceding syllables is similar (269ms for the former and 261ms for the latter), but the final syllable in the former is more than double that in the latter (963ms for the former and 547ms for the latter). The last syllable in the former shows that its duration (963ms) is nearly four times more than the mean duration (269ms) of its preceding syllables whereas the last syllable in the latter shows that its duration (547ms) is just about double the mean duration (261ms) of its preceding syllables. The amount of the lengthening is much smaller for the latter.

As for the declaratives with ironic and impatient confirmation, as shown in figure 4.14 and 4.15 respectively, the last syllable T4 displays a rise after the canonical fall and the rise is followed by another fall. Comparing this rise with that in echo interrogative (in figure 4.12b), the difference in placement of the rising peak is obvious – the rising in declaratives is near the centre between the onset and the offset of the vocalic portion whereas the rising peak is at the offset for the echo interrogative.

The greatly lengthened last-syllables in the examples above play a role of tonic prominence (for definition, see next section); they have a particular modified tonal contour by attaching a sharp rising tail, or a moderate rising tail followed by a small fall, or a gentle rising followed by a gentle fall, etc. They use a great amount of prolongation, different F_0 peak placements and the various pitch modification to the last half duration of the tonal contour to convey various linguistic and paralinguistic meaning, such as echo interrogatives, ironic declaratives and impatient declaratives etc. Certainly, a larger experiment needs to be conducted before sustaining this claim.

Any pitch modification in the final syllable is treated as a local event, which does not contaminate the downward trend of the whole utterance. We have not found that declination

suffers from different linguistic functions (e.g., declaratives vs. interrogatives) or from different paralinguistic functions (e.g., ironic, impatient, surprised or suspicious, etc.)

The downward slope present in all utterances clearly indicates that declination is an acoustic output and is a phonetic property of utterances. The presence of declination does not signal grammatical functions or paralinguistic functions. The declination lines operate within an intonation group and determine that the phonologically identical tones decreases their phonetic pitch value going from left to right.

.

4.4.3. Tonic Prominence

4.4.3.1. Defining Tonic Prominence

An intonation group can coincide with any grammatical unit: a word, a phrase, a clause or a sentence, for the reason that the grammatical unit is not the criterion for marking off an intonation group. Each intonation group has its own information to convey which is syntactically and semantically coherent. The unit of information distinguishes itself from the neighbouring information units by shifting the pitch span. In an intonation group, there may be a single piece of information or several pieces of information, one of which is the most important and is then realised as predominant in the intonation group. The predominant piece in an intonation group is the tonic. This is adopted from Halliday. Halliday has an integrative system to analyse intonation in English (1967, 1970): tonality, tonicity and tone. His tonality involves the division of an utterance into units called tone groups, which is the origin of my `intonation group'. His tonicity involves singling out the most salient syllable in a tone group, which is called the tonic syllable, I adopt this and shall discuss it in the following. However, his tone, which is the type of pitch movement chosen by the tonic syllable, cannot fit into the system of Cantonese tonic prominence except for some special cases which can be referred to 4.4.2.4.2 and 4.4.2.4.1.

Following Halliday, within the intonation group (`tone group' in his term) there is always some part that is especially prominent – this is the part that the speaker wants to show to be the most important in the message. The part that is prominent is called the tonic, and

prominence of this kind in the intonation group is called tonic prominence. Unlike English, in which the tonic is realised in tone and the tone makes extensive usage of the pitch contour movements, the tonic in Cantonese is realised in terms of prolongation of the tonic syllable.

As I argued in chapter three, the most crucial point in talking about prominence at lexical level in Cantonese is that prominence is realised as prolongation of a syllable. The syllable can be a monosyllabic word, a constituent of a polysyllabic compound or of a polysyllabic monomorphemic word. There is no phonological type of `stress' at word level. No syllable in a word is required to be more prominent than the other. Any syllable in a word or in a phrase is free to be lengthened or shortened in accordance with the need of the rhythmic structure or of prominence. The freedom to lengthen and shorten a syllable in a word or in a phrase for the sake of rhythmic structure has been discussed above. We return to prominence at the level of the intonation group, which is realised in the most lengthened syllable which we call the tonic. We call this the tonic prominence. Prominence at word level is not necessarily realised as a tonic in an intonation group, as a prominent syllable of a word is not necessarily the most important information the speaker wants to convey. The tonic prominence takes advantage of the prolongation of a syllable, the pitch orientation already having a heavy functional load in terms of lexical tones and therefore not having much room for any other linguistic function (with the exception of the last tone of an utterance, see discussion in the last section 4.4.2.4.2). Thus, prolongation of a syllable is nearly the sole means of achieving prominence, other than the use of staccato rhythm. Raising intensity may also be used, but this is optional. As every intonation group is a unit of information itself, in which the most important part the speaker wants to convey is realised in longer duration, that part can also be accompanied by higher intensity, thus heightening the prominence. If the important part is a disyllabic word, either or both syllables is/are prolonged. If both syllables are prolonged, one must be much longer and this is the tonic syllable. In the following example, the syllables /k^hy:n^4 wɐj^1/ are both prolonged as the most important part of the intonation group, but only /k^hy:n^4/ is the tonic as its duration is more than double that of /wɐj^1/ (405ms ~ 200ms). The tonic syllable is bold and italic:

405ms 200ms

51 || tɕɐw^{6} m^{4} ɕɐj^{2} |kɐm^{3} ta:m^{1} ɕɐm^{1} | tɕi:^{6} kej^{2} kɛ:^{3} | ***k^{h}y:n^{4}*** wɐj^{1} ||
|| _ _ ____ | _ ____ _ | _ ____ _ | ____ _ ||
then-no-need-so-bear-heart-self-self-p-power-might
(you shouldn't worry about your authority)

4.4.3.2. Location of a Tonic

The tonic can fall on any part of an intonation group. The above example shows that the tonic is located on the penultimate syllable of the last foot. The following examples show that the tonic can be located on the first foot of the intonation group (example 94) or a syllable in the middle of an intonation group (example 16):

94 || ***pɐt^{1}*** fu:^{6} | tɕa:k^{3} jɐm^{6} kɛ:^{3} | kɔ:ŋ^{2} fa:t^{3} ||%
|| ____ _ | ____ _ _ | ____ _ ||%
not-bear-duty-post-p-talk-way
(that is an irresponsible proposal)

16 || mow^{5} wɔ:ŋ^{2} | ***k^{w}u:n^{2}*** | la:^{1} ha:^{2} ||
|| _ ____ | ____ | ____ _ ||
have not-king-control-p.-int.
(uncontrollable)

The tonic always falls on the `long' of a foot. Where the tonic falls on the first syllable of a foot, it can be accompanied with a little greater force of articulation. A little greater force of articulation can reinforce the tonic. This can be seen by comparing the intensity of the syllable /tɕ^{h}ɵt^{1}/ which appears as a tonic in the example 97 and a non-tonic in example 96 below:

96-97 || tɕ^{h}ɵt^{1} ***tɕɔ:j^{6}*** || ***tɕ^{h}ɵt^{1}*** ji:n^{6} hɐj^{2} ||
|| _ ____ || ____ _ _ ||
out-at-out-present-at
(coming from, coming from)

Voice quality can also be used to enrich the prominence, for example, breathy voice in /pɐt^{1}/ in example 94 above.

An alternative way of showing prominence is by means of staccato rhythm. The prominence is realised on every single syllable that is heard to be chopped abruptly by means of more energetic movements, with the pumping of more breath from the lungs. It remains the case that only the most lengthened syllable is the tonic syllable. The chopped syllables are produced with relatively equal force and equal beat (not in absolute sense), and therefore are heard as staccato. We can compare the two examples below, one is staccato (indicated as superscript " $^{\prime}$ ") and one is not:

68 || tɐj^{2} ta:^{2} | lɔ:^{1} jɐw^{2} | ***la:^{1}*** ||
|| ____ _ | __ ____ | ____ ||
deserve-beat-bottom(2)-p.
(should have his bottom smacked)

99 || tɕɐn^{1} | ***tɐj^{2}*** | ta:^{2} | lɔ:^{1} | jɐw^{2} || %
|| $^{\prime}$____ | $^{\prime}$____ | $^{\prime}$____ | $^{\prime}$____ | $^{\prime}$____ ||%
really-deserve-beat-bottom(2)
(They all really deserve to have their bottoms smacked)

Comparing the intensity between syllables in the two examples above, (99) which is heard to be staccato shows to have more even intensity than that in (68) which is heard not be have staccato rhythm.

Where the tonic does not fall on the first syllable of a foot, this higher intensity is usually not found, as can be seen in the example 96 /tɕɔ:j^{6}/ above and in the example 64 /tɕi:m^{6}/ below.

64 || jɐt^{6} ***tɕi:m^{6}*** | ha:^{6} | kɔ:ŋ^{3} ||%
|| __ ____ | $^{\prime}$____ | $^{\prime}$____ ||%
day-gradually-down-sink
(is on the decline)

The intonation group 64 shows only the last syllables /ha:^{6} kɔ:ŋ^{3}/ are heard to have equally great force and thus staccato. They show greater intensity than that in the preceding syllables and recur at even intervals (209ms and 214ms, respectively). However, the second syllable of the intonation group, /tɕi:m^{6}/, the most lengthened syllable, is heard to be more prominent than the others and thus is the tonic. The syllable boundary between /jɐt^{6}/ and /tɕi:m^{6}/ is a bit arbitrary, as a result of the difficulty of drawing a clear line in the middle of the geminate /-tt-/(121.6ms). However, even if the plosive onset /t/ is not included, the duration of the rest segments of the syllable /tɕi:m^{6}/ (226.5ms – starting from VOT) is in fact longer than the duration of the preceding syllable /jɐt^{6}/ excluding /t/ (121ms).

4.4.3. Final-Particle

A tonic can fall on any content word or function word, even on a final-particle. The particle /a:k^{3}/ in the following example conveys a sense of certainty and emphasis, adding an exclamation mark to the previous syllable /hɐj^{6}/ when the tonic falls on it. This can be translated into English as `You are right!' with a falling pitch on `right' preceding a mid even pretonic [-- -- |╲] (tone 1 in Halliday's terminology).

86 || hɐj^{6} ***a:k^{3}*** ||
|| _ ___ ||
right-p.
(Right!)

Final-particles have exclusive space to stretch out the syllable duration as they are at the end of an intonation group. They are not sentence-final particles as some researchers (e.g., Cheung 1986 and Kwok 1984) name them, because they occur not necessarily at the end of a sentence. They can occur anywhere in the middle of a sentence, attached to the end of any constituent unit of a sentence. The constituent unit can be a noun, a verb, an adjective or an adverb, a phrase or a clause. Chao (1968: 795) claims that particles in Mandarin only belong to phrases or sentences whereas Gao (1980: 204) takes the view that Cantonese particles

can attach not only to sentences but also to subjects or to predicates. In my view, it may often happen that the end of a grammatical unit, e.g., sentence or any constituent-within-a-sentence, coincides with the end of an intonation group. Final-particles can be lengthened but, are not necessarily always lengthened as is sometimes thought. In the present data, out of the 83 measured intonation groups, only 15 tonics (i.e., the most lengthened syllable) fall on the last syllable. Of the 15 final-syllable tonics, 5 are final-particles. This leaves, out of a total of 31 final-particles, 26 which are non-tonic syllables. The occurrence of final-particles is quite high in natural discourse. More than two out five intonation groups contain final-particles in the present data (31 out of 83). 13 of the 31 final-particles (clusters are counted as one) are followed by a pause, but only one of them is the tonic. Given that the stretching and shortening of a syllable duration is to function for the alteration of rhythmic patterns of the language and the prolongation of a syllable duration is to function for prominence purposes, the so-called universal automatic utterance-marginal lengthening theory seems to be out of place in the Cantonese intonation system. I do not deny that there is a tendency here and there in Cantonese speech to lengthen the final syllable (examples can be seen in the invented utterances under the section of `declination'). My point is that it would be a mistake to think that this tendency obscures the intrinsic functions of lengthening and shortening in Cantonese. The present data provides strong confirmation of this. Cantonese speakers seem not to take advantage of the universal utterance-marginal lengthening, but would rather choose carefully which syllable to lengthen when they choose which piece of important information to convey.

In her study of sentence-final particles, Kwok observes that with the same sentence-final particle, "a slightly different pronunciation could be more often linked to another meaning" (1984: 12). She compares the two sentences: /k^hɵy^5 hɐw^6 ɕaːŋ^1 nɛː^1/ (she-late-born-p. – she is young.) and /k^hɵy^5 nɛː^1/ (she-p. – what about her?). Kwok perceives that the particle /nɛː^1/ in the first sentence is longer and louder while in the second sentence it is shorter but higher in pitch. Then she concludes that the second is less prominent in terms of loudness. "The first /nɛː^1/ is used to offer an explanation, and the sentence may be glossed as 'It's because she's young.' The second /nɛː^1/ has the function of relating what is said to something else, so

that the second sentence is translatable as 'What about her?' " (p.13). I think that her observation illustrates that when the final-particle is short, it only conveys a matter of lexical meaning, `what about' as in the second sentence she gives; when the final-particle is long, it conveys another meaning [`it is because'] which is outside its lexical meaning as in the first sentence she gives. Note that loudness is only to enrich the prominence in my view. In the first sentence, /nɛ:^{1}/ is lengthened and is thus a tonic, its lexical meaning `obviously' is overwhelming. Kwok interprets it as `it's because'. I would rather think it enhance the lexical meaning `obviously' to `you ought to have known'. If we shorten the first syllable /k^{h}ɵy^{5}/ in her second sentence we will find that the tonic /nɛ:^{1}/ enhances the lexical meaning `what about' to `you should to have thought of'. If we lengthen any of the first three syllables in her first sentence, the result would be `she is young (implying: not he or other persons)' if /k$^{h}$ɵy$^{5}$/ is the most lengthened; or `she is young (imply: not old)' if either /hɐw^{6}/ or /ɕa:ŋ^{1}/ is lengthened. In these two cases, the final-particle /nɛ:^{1}/ is omitted because the sense of `obviously' can be more easily achieved by a lengthening of any of the three syllables in the utterance. Bear in mind, tonic is a property of intonation and only conveys the meaning of intonation.

As in the example 26 above, /jek^{6} tow^{1} tɕ^{h}oŋ^{4} lɔ:j^{4} nɛ:^{1}/ (also-also-from-come-p. – always), an adverbial phrase, which is the beginning of a sentence, the final-particle /nɛ:^{1}/ is merely to bridge the beginning and the rest of the sentence. Now that we have an intonation group, it is much easier to show that the final-particles only occur at the end of an intonation group. We can name them `intonation-group final-particles' if we want to. They can be lengthened to several times as long as any syllable in the intonation group but this seldom happens. It follows that they are seldom the tonic. The following example 70 shows that the duration of the final-particle /ka:^{2}/ (614ms) is more than double the length of the second lengthened syllable /k^{h}y:n^{4}/ (286.6ms), whereas the example 92 shows that the most common case is that the duration of the final-particle is not lengthened even though the final-particle (i.e., /a:^{4}/ (72ms)) occurs before a pause. (70) is a part of the sentence `Why should teachers need to

have authority anyway?'. The lengthened final-particle /ka:^{2}/ is to strengthen the lexical meaning of the adverbial compound /ɕi:n^{1} tɐk^{1}/ (anyway) by adding a tone of `it can't be helped'. (92) is an independent sentence. The final-particle functions as a question-marker turning the statement `to be a street-child' into a rhetorical question by adding up a tone as a question-mark plus an exclamation-mark `?!'. However, the final-particle /a:$^{4}$/ adds nothing to the intonation beyond its lexical meaning. /ka:j$^{1}$/ (street) is the tonic. The sentence can be translated into English with a high falling tonic on `street' following a low even pretonic in the sentence `He'll be out on the streets?!'.

70 || k^{h}y:n^{4} wɐj^{1} | ɕi:n^{1} tɐk^{1} | ***ka:^{2}*** ||%
|| ____ __ |____ _ |___ ||
power-might-until-o.k.-p.
(authority anyway)

92 || tɕow^{6} | ***ka:j^{1}*** t^{h}oŋ^{4} | a:^{4} ||%
|| ^ __ |_____ __ |__ ^||%
to be-street-child-p.
(He'll be out on the streets)

The final-particle can, rarely, occur as the penultimate syllable in an intonation group, as in the intonation group 1. In such a case, the last syllable has no substantial meaning but is just a filling word, and functions like a particle to connect the previous pronoun /ŋɔ:^{5} / (I) to the following adverb /tɕ^{h}oŋ^{4} lɔ:j^{4}/ (from-come - always).

4.5. MAJOR INTONATION GROUP

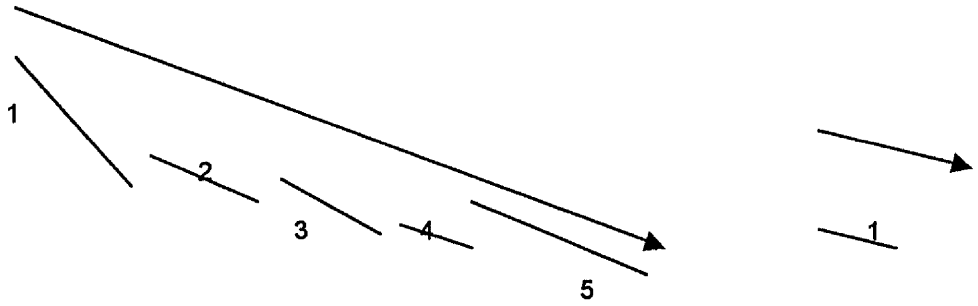
4.5.1. Defining Major Intonation Group

The resetting of pitch span divides an utterance into intonation groups. The pitch movement of intonation groups is downward, but the resetting is upward. A series of intonation groups are in progressive descending order, the starting point of each intonation group being lower than the starting point of the previous one, but higher than the end point of the previous one. I call such a series of intonation groups a `major intonation group'. A major intonation group

displays a downward trend. Together with the downward trend of the constituent intonation groups it displays a picture of 'downtrend of downtrend'. Giving the first major intonation group as an example in the following:

Major Intonation Group 1: Major Intonation Group 2:

(The downtrend line of the first major intonation group covers the downtrend lines of the intonation groups 1-5)



(Intonation groups 1-5)

||***kɐm*²** ŋɔ:⁵ |nɛ:¹ʨɐw⁶||***ʨʰoŋ*⁴** lɔ:j⁴ tow¹ ||m⁴ ***kɛ:ŋ*¹** hɔ:k⁶ ||***ɕ*⁶**ɛ:⁶|| m⁴ kɛ:ŋ¹| ka:w³ ***ɕi:*¹** kɛ:³||
|| ___ _ | _ ___ || ___ _ _ || _ ___ _ || ___ || _ ___ | _ ___ _ ||
in that case-I-p.–then-from-come-also–not–afraid–study-stu-not–afraid–teach–expert-p.
(I was never frightened of stu- ... never frightened of teachers)

4.5.1.1. Plateau Line

A major intonation group is the largest prosodic domain into which utterances are divided. It can consist of one or more intonation groups. Five is considered long. A major intonation group is a united intonation unit represented by a united fitted downtrend line over all the top lines of the constituting intonation groups, until the next major intonation group comes up which blocks the continuation of the fitted downtrend line by resetting the pitch span to a comparatively much higher level. This line, I call a `plateau line' which differs from the top line of an intonation group. It is a property of the major intonation group that contains it and it covers a bigger size of the resetting of the pitch span. It is expected not to coincide with the top line of all of the constituent intonation groups. Recall that the four fitted lines in intonation groups are lexical-tone linked. If there are no high tones in an intonation group, the highest fitted line is expected not to be a top line. This can be seen from the major intonation group 1 above, where only the constituting intonation groups 1, 2, 3 and 5 provide a top line but the

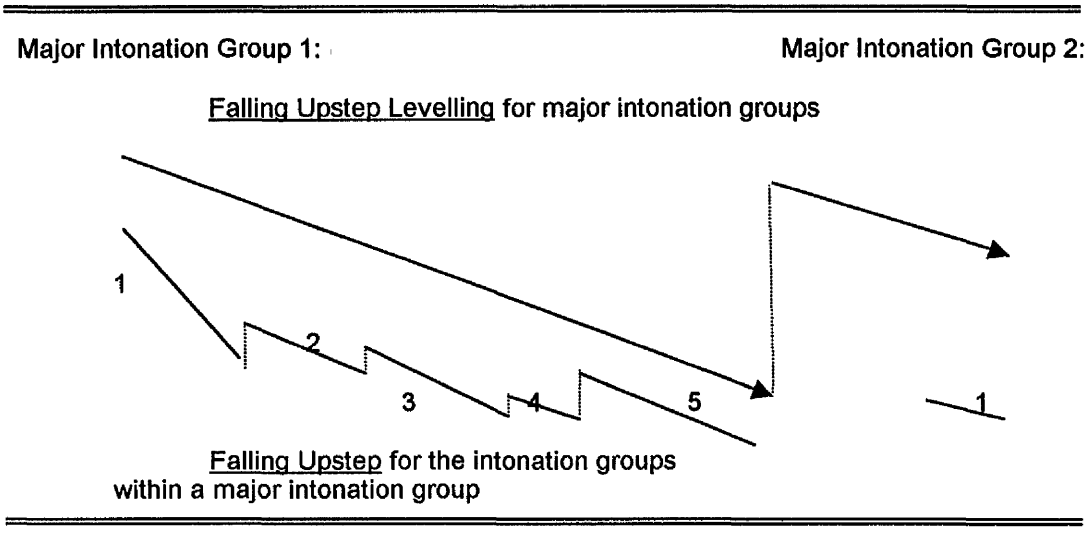
intonation group 4 does not. The highest fitted line in the intonation group 4 (consisting of two T6s) is a low-mid line. In such a case, no comparison can be made between the highest fitted lines of the intonation group 4 and the intonation group 1. When we say the fitted plateau line of the major intonation group 1, it must cover all the top lines (i.e., not only the highest lines) of the constituent intonation groups from 1 to 5, therefore, some allowance of the tonal space to the top of the intonation group 4 must be given. The only criterion for dividing a major intonation group from one to another is the pitch span resetting to a much higher level than the end of its previous intonation group. This bigger size of resetting blocks the continuation of the downtrend line, the plateau line, which covers all the top lines of the previous intonation groups.

4.5.1.2. Falling Upstep

The internal relationship between the intonation groups in a major intonation group is the upward setting of the pitch span. As intonation groups within a major intonation group are in a downward trend, but the resetting is upward, I shall term the pitch movement in the major intonation group the `falling upstep'.

The term `falling upstep' can also describe the pitch movement between the major intonation groups. A major intonation group displays a downtrend line, the starting point of a major intonation group not necessarily being higher than the starting point of the previous one, but definitely striking higher than the end point of the previous one. Thus, the falling upstep has to be modified to the falling upstep levelling when applying to the relationship between the major intonation groups.

Fig. 4.16. Falling upstep and falling upstep levelling.



4.5.2. Resetting

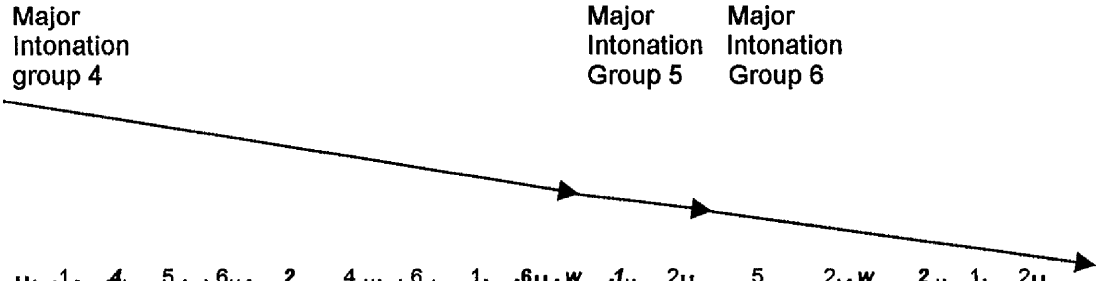
The major intonation group may correspond to any grammatical unit, but is not determined by the syntactic structure of an utterance. The grammatical unit is not the criterion for the plateau line resetting. A major intonation group represents a unit of information which is syntactically and semantically coherent. This can be seen from the major intonation group 1 quoted above. A major intonation group differs from intonation groups in that its boundary usually ends with a larger grammatical unit, in many cases a short sentence or a clause. Or, it can consist of a couple of short sentences. If the boundary of a major intonation group does not correspond to a complete short sentence or a clause, there must be a `good reason' (the term is borrowed from Halliday 1970: 3) for it. The `good reason' for an uncompleted short sentence corresponding to a major intonation unit can be, for example, that interruption occurs during a speaker's speech, an utterance drags too long or an error occurs, etc.

4.5.2.1. Interruption

An example of interruption occurs during a speaker's speech can be given as in the major intonation group 4, 5 and 6.

Major Intonation group	Speaker	English Equivalent
[4]	A:	On the one hand, it was possibly because I was ...
[5]	B:	good, ...
[6]	A:	uncontrollable.

The major intonation group 5 shows that when speaker B jumps into the conversation with only one foot anticipating what he thinks the speaker A means while speaker A is talking through the major intonation group 4, speaker B does not reset the pitch span to a new level. Instead, he fits his voice to the end of the major intonation group 4 of the speaker A in `completing' speaker A's sentence. Despite the unexpected interruption from speaker B, speaker A is able to carry on his parroting of speaker B's view first in the major intonation group 6, before getting into his own view in the major intonation group 7. He does the same as speaker B, in that he does not reset the pitch span to a new level but rather fits his voice to the end of the major intonation group 5 of the speaker B. As a result, the major intonation groups 4, 5 and 6, produced by speaker A, B and A respectively, are in a continuous downward trend, despite the fact that the very nature of the voice of every individual speaker is different. The reason these three pieces of continuous conversation are divided into three major intonation groups is simply because they are produced by different speakers. A major intonation group also shows that it is a united unit not only in terms of pitch contour but also in semantic and syntactic terms.



||jɐt^{1} ***lɔ:j^{4}***|ŋɔ:^{5} jek^{6}||***hɔ:^{2}*** nɐŋ^{4} ||jek^{6} tow^{1} ***hɐj^{6}***||***k^{w}a:j^{1}***|ha:^{2}||mow^{5} wɔ:ŋ^{2}|***k^{w}u:n^{2}*** |la:^{1}ha:^{2}||

|| _ __ | __ - || __ _ || _ _ __ || __ | __ || _ __ | __ | __ __ ||

one-come- I-also-possible-can-also-also-am- good-int.-have-not-king-control-p.-int

(Speaker A) (Speaker B) (Speaker A)

|| On the one hand, it was possibly because I was || good, || uncontrollable. ||

(The bold double bar stands for the boundary of the major intonation groups.)

4.5.2.2. Long Utterances

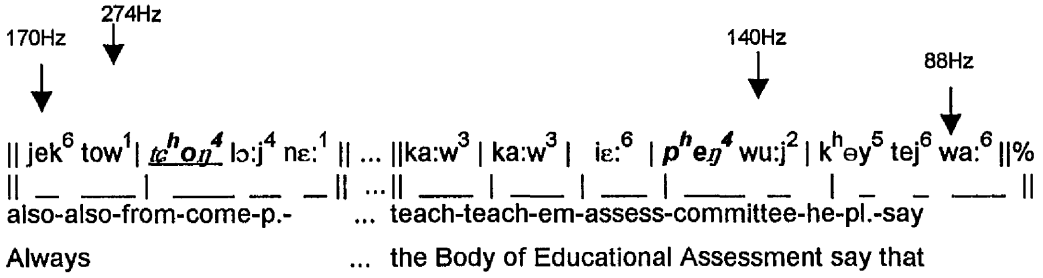
If an utterance is dragged on for too long a bigger size of resetting is required, as in the major intonation group 11, 12 and 13.

Major Intonation Group: (English Equivalent)

[11] `I never, ... unlike teachers nowadays, for example, the Body of Educational Assessment say that, ... '
[12] `...'many students don't respect teachers'....'
[13] `... such a situation did not exist.'

The major intonation group 11 contains four intonation groups and lasts for 4557ms. The F_0 of the first high tone which is the second syllable /tow^1/ of the major intonation group is 274Hz but the F_0 of the last high tone, which is the second antepenultimate syllable /wu:j^2/ of the major intonation group is 140Hz. The first and the last syllables of the major intonation group both are T6, whose F_0 is 170Hz for the first (i.e., /jek^6/) and 88Hz for the last (i.e., /wa:^{6}/).

The major intonation group 11:



This clearly demonstrates that the end of the major intonation group reaches the near bottom of the speaker's pitch range and the bigger size of pitch resetting makes it obvious that another unit of information is coming up. However, the bigger size of resetting (i.e., the starting of the major intonation group 12) is carefully chosen not to break up a clause but is in between two clauses. It occurs at the quotation of what people say in the previous major intonation group. On the other hand, this also clearly shows that the speaker divides a long utterance into units of information by resetting the pitch span, corresponding to the major intonation groups. Another bigger size of resetting, i.e., the start of the major intonation group 13 occurs at the concluding point of the example in the previous major intonation group. The three major intonation groups form a complete sentence, and each of them forms a complete clause. The uncompleted utterances in each of them are considered as a complete

information unit as well as a grammatical unit as they are clauses. In contrast, some intonation groups can correspond to a small grammatical unit, for example, a determinative-measured compound /kɔ:^{2} tɕoŋ^{2}/ (that m. – such a) at the first intonation group of the major intonation group 13.

4.5.2.3. Errors

If an error in choosing a lexical item occurs, the correction can be made within the same major intonation group. Errors can occur in the middle of a major intonation group, as in the major intonation group 1, in which the speaker wants to say the lexical item /ka:w^{3} ɕi:^{1}/ (teach-expert – teacher). However, he makes a semantic error in selecting semantically related vocabulary item /hɔ:k^{6} ɕa:ŋ^{1}/ (study-pupil - student) and utters the first half of it /hɔ:k^{6}/ (study). Before he manages to correct himself properly, the onset of the last half lexical item, /ɕ:/, has come out from his mouth already. He manages to stop himself from producing the whole syllable /ɕa:ŋ^{1}/, but he still utters a filling item /ɛ:^{6}/. Then he quickly repeats the verb phrase /m^{4} kɛ:ŋ^{1}/ (not-frighten) before producing the correct lexical item. All this happens in a major intonation group within three intonation groups, demonstrating that the speaker considers that the correct lexical item together with its previous intonation groups forms a complete unit of information.

Speakers seem not to intend to reset the pitch to a new major intonation group after making an error as discussed above. They cannot even reset the pitch to a new intonation group. They correct the error within the same intonation group. There is a hesitation error made in the fourth intonation group of the major intonation group 11, in which the speaker may be not very familiar with the trisyllabic lexical item /ka:w^{3} p^{h}eŋ^{4} wu:j^{2}/ (teach-assess-committee – the Body of Educational Assessment) which is an abbreviation of /ka:w^{3} jok^{6} p^{h}eŋ^{4} ji:^{5} wu:j^{2}/. He repeats the first syllable and still stumbles over the word, then has to utter a filling item /iɛ:^{6}/ as a transition until the right word comes to his mind. The speaker, in this case, does not say again the complete lexical item (all three syllables) but rather is happy with pronouncing the

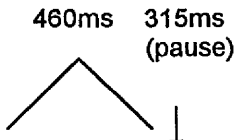
last two syllables /$p^h eŋ^4$ wu:j^2/ of the trisyllabic word separated from the first syllable /ka:w^3/ with the filling item /iɛ:^{6}/ in between. Moreover, he puts all of this into one intonation group. He is presenting units of information in such a way that if forgetfulness causes interruption in some word or phrase, he still trying to complete the same information unit. Thus, the trisyllabic compound is interrupted, but when resumed it is still a unit of information, represented in an intonation group. In both hesitation errors and errors of choosing lexical items, the mistaken syllable or the filling item displays a relatively long duration. This may be due to the `thinking time' the speaker needs to take. The /ç^6 ɛ:^{6}/ in the major intonation group 1 lasts for 250ms, it is rather long in comparison with the duration of the previous syllables /m^4 kɛ:ŋ^1 hɔ:k^6/: 140ms, 149ms and 141ms, respectively and the following syllables /m^4 kɛ:ŋ^1 ka:w^3 çi:^{1} kɛ:^{3}/: 151ms, 172ms, 165ms, 178ms and 152ms, respectively. The duration of /ka:w^3 ka:w^3 iɛ:^{6} $p^h eŋ^4$ wu:j^2/ in the major intonation group 11 is 290ms, 388ms, 400ms, 231ms and 182ms respectively. The first three syllables are relatively long in comparison with the duration of the previous syllables /pej^3 jy:^{4}/ (for instance): 98ms and 160ms and the following syllables /$p^h eŋ^4$ wu:j^2/ and the syllables in the following foot /$k^h ɵy^5$ tej^6 wa:^{6}/ (they said): 130ms, 120ms and 159ms, respectively. Hesitation naturally causes a longer time for a syllable to be completed. In such a case, the lengthened syllable cannot be considered as a tonic syllable as it does not represent a piece of important information unit in a speaker's mind.

4.5.2.4. Pausing

Pause is definitely not a reliable element to mark off a major intonation group. Pause may or may not occur at the end of an intonation group, or a major intonation group. In this sense, the major intonation group is not an audible intonation domain. Certainly, a pause usually occurs at the end of a grammatical unit. But the length of a pause seems to have nothing to do with the length of the intonation unit. For example, a long pause of 315ms follows the first intonation group of 460ms in the major intonation group 8, whereas a short pause of 192ms follows the first intonation group of 409ms within the major intonation group 23. The major

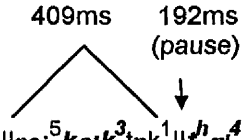
intonation group 36 lasts 1907ms followed by a pause of 368ms, whereas the major intonation group 11 lasts 4557ms followed by a pause of 336ms.

The major intonation group 8:



|| ***ta:n*⁶**hɐj⁶nɛ:¹||tɕɐw⁶ ***ki:n*⁶** la:p⁶ | tɕɔ:²kɔ:³||***how*²**how²|kɛ:³||***pʰɐŋ*⁴** jɐw⁵|kʷa:n¹hɐj⁶ la:k³||
|| ___ _ _ || _ ___ _ | ___ _ || ___ _ | ___ || ___ _ | ___ _ _ ||
but-is-p.-then-build-establish-ed -m.-very-good-p-friend-friend-concern-relate-p.
(However, we were on very friendly terms)

The major intonation group 23:



||ŋɔ:⁵***kɔ:k*³**tɐk¹||***tʰɐj*⁴** ji:⁵|ni:¹kɔ:³||***tʰɐj*⁴** mok⁶kɛ:³|ti:w⁶tɕʰa:⁴tow¹||tɐj² ta:²| lɔ:¹ jɐw² |***la:*¹** ||
|| _ ___ _ || ___ _ | _ ___|| ___ _ _ | ___ _ _ || ___ _ | _ ___ | ___||
I-feel-able-lift-discuss-this-m.-title-topic-p-investigate-check-all-deserve-beat-bottom(2)-p.
(I think that whoever suggested the topic of this investigation should have his bottom smacked)

According to Black *et al* (1966: 240), there is no significant difference in the use or location of pauses by speakers of English, Hindi, Japanese and Spanish. And Fodor *et al* (1974: 425) suggest that breathing is dependent on syntax: speakers will only breathe when allowed to do so, as it were, by the constituent structure of the utterance. Pauses in Cantonese seem to have another function which is to call attention. I shall call this the `highlight pause'. A highlight pause is usually inserted in the major intonation group, perhaps after the first intonation group of a major intonation group or in between any constituent intonation groups. Only half of the major intonation groups are followed by a pause but a quarter of the major intonation groups are interrupted by a pause in the present data. The major intonation groups 8 and 19 give evidence of the highlight pause. In these two examples, the highlight pause is inserted after the first intonation group. In the major intonation group 8, the highlight pause is after the lexical items /ta:n^{6} hɐj^{6} nɛ:^{1}/ (however); while in the major intonation group 19, the

highlight pause is after the lexical items /ɕɔ:^{2} ji:^{5} nɛ:^{1}/ (therefore). The highlight pause in the former case is calling attention to the argument while in the latter case it is calling attention to the conclusion. The major intonation group 8 follows the fact mentioned in the previous major intonation groups that the speaker was uncontrollable and the teachers never really punished him. The pause after the conjunction `however', is to call attention to the argument of `we were on very friendly terms'. The major intonation group 19 follows the argument in the previous major intonation groups that an educational environment where teachers and students have strong friendship is far better than the situation where the Body of Educational Assessment complaints that students do not respect teachers. The pause after the conjunction `therefore', is to call attention to the conclusion of `teachers, you shouldn't worry about your authority'. The highlight pause in the major intonation group 9 gives evidence that it is calling attention to another side of the story of `apart from', by being inserted between the clauses: `i.e., apart from the teacher-student relationship' and `the friendship was very strong'.

The major intonation group 19:

(pause)

↓

||***ɕɔ:^{2}*** ji:^{5} nɛ:^{1}|| ***ŋɛ:^{6}*** ||kɔ:k^{3} wɐj^{2}|ka:w^{3}ɕi:^{1}|***nɛ:^{1}***||tɕʁw^{6} m^{4}ɕɐj^{2}|kʁm^{3}ta:m^{1}ɕʁm^{1}|tɕi:^{6} kej^{2} kɛ:^{3}|***k^{h}y:n^{4}*** wɐj^{1}||

||___ _ _||__||___ __ | ___ _ |___|| __ __ ___ |_ ___ _ |__ ___ _ |___ __||

there-fore-p.-em- every-m.-teach-expert-p.-then-no-need-so-bear-heart-self-self-p.-power-might

(Therefore, teachers, you shouldn't worry about your authority.)

The major intonation group 9:

(pause)

↓

|| tɕek^{1} ***tɕ^{h}ej^{4}*** tɕɔ:^{2} || ***ɕi:^{1}*** ɕɐŋ^{1}| k^{w}a:n^{1} hɐj^{6} nɛ:^{1}|| kɔ:^{3} ***p^{h}ɐŋ^{4}*** jʁw^{5} | k^{w}a:n^{1} hɐj^{6} | how^{2} how^{2} ||

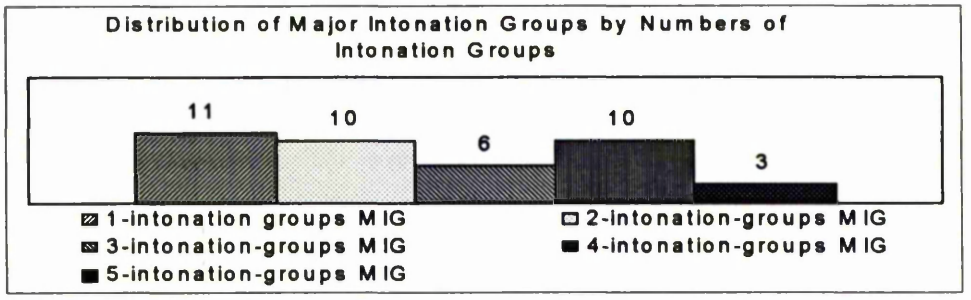
||_ ____ _ || ___ _ | ____ _ _ || _ ____ _ | ___ _ | _ ___ ||

ie-apart-p-teacher-student-concern-relate-p-m-friend-friend-concern-relate-very-good

(This means, apart from the teacher-student relationship, the friendship was very strong)

There are total 27 major intonation groups containing 83 intonation groups in our data for acoustical analysis whose distribution is displayed in the following figure.

Fig. 4.17. The distribution of the 27 major intonation groups by numbers of intonation groups.



4.5.3. Tonic Prominence

4.5.3.1. Defining Tonic Prominence

There is a piece of important information which a speaker wants to convey in the major intonation group. This is the tonic prominence. I shall call it the major tonic prominence, the syllable which carries this prominence is the major tonic syllable. The major tonic is realised in the most lengthened syllable. At the same time, it is also the tonic in the intonation group it belongs to. A major tonic can fall on any constituent intonation group, usually the last constituent intonation group. The Major intonation group 14 shows that the major tonic falls on the tonic of the last intonation group whereas the major intonation group 1 shows that the major tonic falls on the tonic of the second intonation group. (Tonics are bold and italic, the major tonic is underlined.)

The major intonation group 1:

||***kʁm***^{2} ŋɔ:^{5} |nɛ:^{1} tɕʁw^{6}||***tɕ^{h}oŋ***^{4} lɔ:j^{4} tow^{1} ||m^{4} ***kɛ:ŋ***^{1} hɔ:k^{6} ||***c***^{6}ɛ:^{6}|| m^{4} kɛ:ŋ^{1}| ka:w^{3} ***ɕi:***^{1} kɛ:^{3}||
|| ___ _ | _ ____ || ____ _ _ || _ ____ _ || ___ _ || _ ___ | _ ___ _ ||
in that case–I-p.–then-from-come-also–not–afraid–study-stu-not–afraid–teach–expert-p.
(I was never frightened of stu- ... never frightened of teachers)

The major intonation group 14:

|| ***ta:j***^{6} ka:^{1} ||tow^{1} how^{2} | ***tɕy:n***^{1} tɕoŋ^{6} ||low^{5} ***ɕi:***^{1} | kɔ:^{3} wɔ:^{3} ||%
|| ___ _ || _ ____ | ____ _ || _ ___ | _ __ ||%
big-family- all-very-respect-respect-old-expert-p.-p.
(Everybody respected teachers very much)

4.5.3.2. Tonics and the Major Tonic

We know that a major intonation group consists of one or more intonation groups and each constituent intonation group consists of a tonic. A major tonic is the most lengthened syllable among the constituent tonics. The second most lengthened tonic among the constituent tonics in a major intonation group is the secondary tonic. The secondary tonic prominence is realised by the lengthened syllable which carries it, but the lengthening is comparatively shorter than the major tonic. The secondary tonic is also a tonic in the intonation group it belongs to. The major and the secondary tonic are in different intonation groups. Thus, no secondary tonic can be found in a major intonation group if it consists of only one intonation group.

In the major intonation group 4, the major tonic /hɔ:^{2}/ (probable) is the most important information the speaker wants to convey in the utterance `On the one hand, I probably was; on the other hand, I also was ...'. In this utterance, he prolongs the major tonic to 310ms to emphasise the probability of his suggestion. Note that the tonic /hɐj^{6}/ (am) in the last intonation group is to provide the positive side of his suggestion, but only lasts for 220ms. The tonic /lɔ:j^{4}/ (come) is the half of the lexical item /jɐt^{1} lɔ:j^{4}/ (one-come – on the one hand) in the first intonation group is to tell listeners there will be two sides of the story (probability and positive) that they are going to hear, which is only 201ms. The major tonic falling on the first syllable of the lexical item `probably' is to emphasise the implication that the `probability' the speaker ascribes is greater than usual and he then lets listeners make up their mind what side of the story they like to pick up.

The major intonation group 4:

201ms 310ms 220ms

↓ ↓ ↓

|| jɐt^{1} ***lɔ:j^{4}*** || ŋɔ:^{5} jek^{6} | ***hɔ:^{2}*** nɐŋ^{4} || jek^{6} tow^{1} ***hɐj^{6}*** ||

|| _ ___ || ___ _ | ___ _ || _ _ ___ ||

one-come-I-also-possible-can-also-also-am

(On the one hand, it was possibly because I was)

In the major intonation group 8, the major tonic falls on the syllable /how^{2}/ (very) which describes the level of the good relationship of the teachers and students that the speaker is going to talk about. The secondary tonic falls on the syllable /p^{h}ɐŋ^{4}/ (friend) which is the description of the relationship of the teachers and students he is talking about. Two pieces of information are singled out by means of syllabic prolongation, and as a result the major (322ms) and the secondary (290ms) tonics are relatively longer than other tonics. The tonic in the first intonation group, /ta:n^{6}/ (but) (125ms) is comparatively shorter than other tonics, indicating that `but' merely connects the preceding major intonation group to the following units of information. The tonic in the second intonation group, /ki:n^{3}/ (to build) (250ms) is comparatively longer than the tonics in the first intonation group but comparatively shorter than the major and the secondary tonics. The speaker does not emphasise the lexical item `to build'. The literal translation for this sentence is `so, however, (we) built a very good friendship' (However, we were on very friendly terms.) The major tonic falls on `very'. The topic `we' is not present, the major tonic falls on the middle part of the comment and the secondary tonic falls on the last part. The whole comment contains no given information.

The major intonation group 8:

125ms 250ms 322ms 290ms

||***ta:n***^{6}hɐj^{6}nɛ:^{1}||tɕɐw^{6}***ki:n***^{3}la:p^{6}|tɕɔ:^{2}kɔ:^{3}||***how***^{2}how^{2}|kɛ:^{3}||***p***^{h}***ɐŋ***^{4}jɐw^{5}|k^{w}a:n^{1} hɐj^{6}la:k^{3}||

but-is-p.-then-build-establish-ed-m.-very-good-p-friend-friend-concern-relate-p.
(However, we were on very friendly terms)

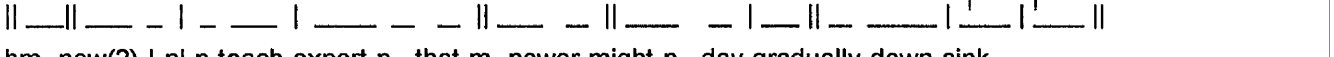
4.5.3.3. Pitch Span

In a major intonation group the major tonic, as well as being the longest, is also realised in covering the widest pitch span. Thus, a major tonic stretches its contour to a great extent. This can be more easily understood if the major tonic is in a conflicting tonal environment. A major tonic T2 demonstrates a sharper rise in a conflicting tonal environment than in any other circumstance, as shown in the major intonation group 4. The major tonic /hɔ:^{2}/ has a similar contour to the similar tone /jɐt^{1}/ which is in the preceding intonation group as both are

in conflicting tonal environments. The T1 is preceded by a T3 and followed by a T4 whereas the major tonic T2 is preceded by a T6 and followed by a T4. The rising of the T2 exerts 89Hz to reach its peak while the T1 exerts 51Hz. The pitch span represented by a top line and a bottom line of the intonation group which contains the T2 /hɔ:^{2}/ is 111Hz, while the pitch span represented by a top line and a bottom line of the intonation group which contains the T1 /jɐt^{1}/ is 86Hz. A major tonic T4 also demonstrates a greater fall in a conflicting tonal environment than in any other circumstance, as shown in the major intonation group 22. The major tonic /k^{h}y:n^{4}/ follows a T3 and precedes a T1. The offset of the T3 is 194Hz. The T4 falls to 99Hz before going up to the following T1 whose peak is 194Hz. The pitch span represented by a top line and a bottom line in the intonation group is 95Hz whereas that in the one before the preceding intonation group is 88Hz. The widest pitch span of a major tonic has to be understood in terms of the fitted lines of the intonation group it belongs to as compared with the same fitted lines of the neighbouring intonation groups within the same major intonation group.

The major intonation group 22:

|| ***ɛ:***^{6}|| ji:^{4} ka:^{1}| ŋɔ:^{5} tej^{6} | ***ka:w***^{3} ɕi:^{1} nɛ:^{1}|| ***kɔ:***^{2} kɔ:^{3}|| ***k^{h}y:n***^{4} wɐj^{1}|nɛ:^{1}|| jɐt^{6} ***ʨi:m***^{6}| ha:^{6}| kɔ:ŋ^{3}||



hm- now(2)-l-pl-p-teach-expert-p.- that-m.-power-might-p.- day-gradually-down-sink
(that `nowadays the authority of us teachers is on the decline')

4.5.4. Declination

4.5.4.1. Defining Declination

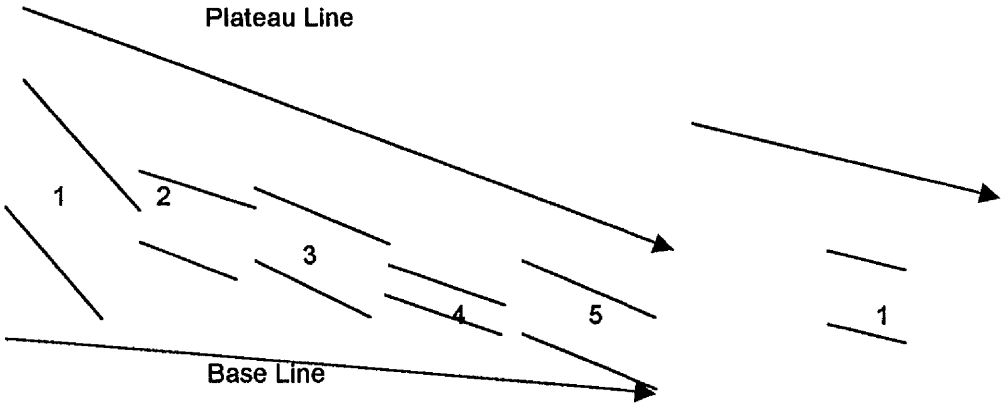
The declination phenomenon is present in all intonation groups as well as in all major intonation groups. The declination in intonation groups is within the declination in a major intonation group. The amount of declination is constant in an intonation group while it decreases along a major intonation group. The term `amount of declination' refers to the difference in F_0 value between the beginning and the end of the fitted lines, which is adopted from Ladd (1984: 57).

4.5.4.2. Plateau Line and Base Line

The fitted lines here refer to the plateau line and the base line. The plateau line of a major intonation group covers the top lines of the constituent intonation groups and the base line of a major intonation group underlays all the bottom lines of the constituent intonation groups. Recall that all fitted lines in an intonation group are lexical-tone related. If there is no low tone in an intonation group, no bottom line is found, as in the intonation group 1. Some allowance to the tonal space below the low-mid line in the intonation group 1 should be given. The plateau line and the base line generated in the major intonation group 1 are shown in the following figure. The two lines show that the range of pitch is getting gradually narrower along the course of the speech time. The plateau line falls steeply while the base line falls gently. On the other hand, the parallel lines in each intonation group which are lexical-tones related show the amount of declination as constant.

Fig. 4.18. Plateau line and base line generated in major intonation group 1.

Major intonation Group 1: Major intonation Group 1:



194Hz 185Hz 160Hz 165Hz 132Hz

||*kɐm*^{2}ŋɔ:^{5}|nɛ:^{1}tɕɐw^{6}||*t̪ɕ^{h}o ŋ*^{4}lɔ:j^{4}tow^{1}||m^{4}*kɛ:ŋ*^{1}hɔ:k^{6}||ɕ^{6}ɛ:^{6}||m^{4}kɛ:ŋ^{1}|ka:w^{3}*ɕi:*^{1}kɛ:^{3}||*m*^{4}m^{2}||

in that case-I-p-then-from-come-also-not-afraid-study-stu-not-afraid-teach-expert-p-mh-mh

('I was never frightened of stu- ... never frightened of teachers.' 'Mm Hmm.')

In general, the end of the base line of a major intonation group can be seen as reaching the bottom of a speaker's voice range. Thus, we can say that the pitch span in an intonation group refers to the `pitch register' in Ladd's terminology (1990: 38) of a given speaker, which can shift up and down and can stretch wider or shrink narrower or can vary between grammatical units or for local prominence. While the pitch span in a major intonation group refers to the neutral voice of a given speaker in a particular communication circumstance, which is usually below falsetto and above creaky phonation, varies for truly paralinguistic differences of overall interest, arousal, etc, and is usually described as the `pitch range' (as in Ladd's terminology (1990: 38)). In many cases, the beginning of the plateau line of a major intonation group reaches the `ceiling' of a speaker's pitch range and the end of the base line of the major intonation group reaches `bottom' of the speaker's pitch range. But the terms `ceiling' and `bottom' are not used in an absolute sense. It is common knowledge that speakers usually use the lower two thirds of their pitch potential. The `ceiling' and the `bottom' of their pitch range can therefore be taken as referring to the lower two thirds only. The data show that the highest F_0 at the beginning of a major intonation group for speaker A is approximately 250Hz, and the lowest F_0 at the end of a major intonation group for the same speaker is approximately 95Hz. In some extreme cases, the F_0 maximum (realised on T1 or T2) for the same speaker could be as high as 274Hz (e.g., /tow^1/ and /tɐk^1/ in the major intonation groups 11 and 37, respectively) or could be as low as 176Hz (e.g., /tow^1/ and /k^{h}ɐp^1/ in the major intonation groups 14 and 17, respectively), whereas the F_0 minimum (realised on T4) could be as low as 70Hz (e.g., /wa:n^4/ in the major intonation group 18).

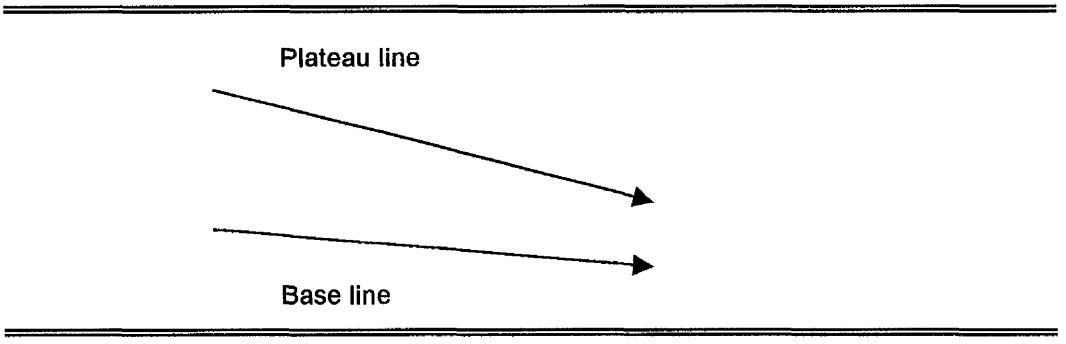
4.5.4.3. Even Fall and Uneven Fall

The slow declining base line which is constantly presented in major intonation groups can be used as a basic line representing the bottom or near the bottom of a given speaker's voice range, whereas the steeply declining plateau line which varies its slope in major intonation groups can be represented the rate of declination. The amount of declination between the beginning and the end of the major intonation group can be either evenly falling down through the course of time or falling down sharply at the beginning before getting back to a gradual

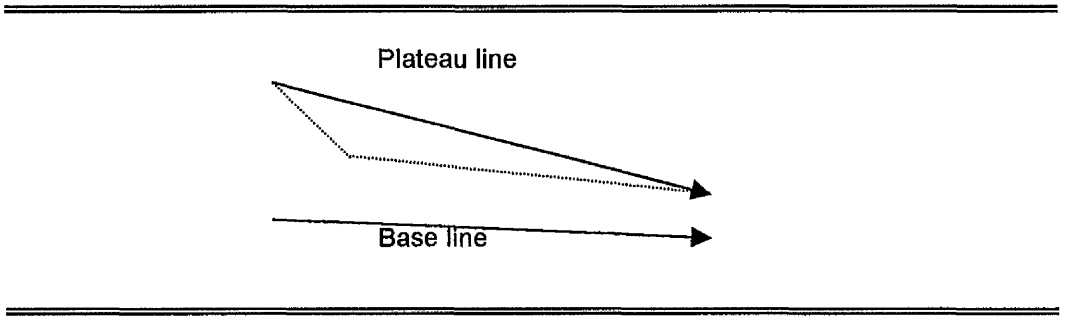
fall, demonstrating an overall picture of an uneven fall. Most cases are evenly falling. As I argued elsewhere, declination in Cantonese is systematic and present in all intonation groups. The uneven fall speaks of the sharply shifting down the setting of the intonation groups, not the rate of declination. That is the internal relationship between the first constituent intonation group(s) and the following intonation groups. An even fall matches with the plateau line, whereas an uneven fall displays a downward bending below the plateau line.

The even fall and uneven fall of the plateau line are shown as follows.

The even fall: the amount of declination decreases evenly from the beginning to the end:



The uneven fall: the first part of the dotted line shows that the amount of declination is dramatic at the beginning of a major intonation group:



The major intonation group 1 illustrated above is a common uneven fall, which shows that the amount of declination is greater at the beginning of a major intonation group. The major intonation group 35 gives evidence of an even fall, in which the speaker raises his voice all the way through in order to distinguish his mimicking voice from his `real' voice in the

following major intonation group 36. This even fall in this case is on a high register throughout.

The major intonation group 35:

|| hɐj^{6} ***ak:^{3}*** || nej^{5} kɔ:k^{3} tɐk^{1} | k^{h}ey^{5} ***jɐj^{5}*** || ***ŋɐj^{3}*** || mow^{5} tɐk^{1} | ***ka:w^{3}*** lɔ:^{3} ||
|| _ ___ || _ ___ _ | _ ___ || ___ || ___ _ | ___ _ ||
right-p.-you-feel-able-him-naughty-oh-not-able-teach-p.
(Right! You think he is naughty, unteachable)

The major intonation group 37, on the other hand, gives evidence of an uneven fall, in which the speaker makes a conclusion after raising the pitch in the first constituent intonation group for the purpose of calling for attention. After raising the pitch span of the first constituent intonation group, he dramatically shifts down the pitch span of the second constituent intonation group.

The major intonation group 37:

|| ***ŋɔ:^{5}*** |kɔ:k^{3} tɐk^{1} | a:^{3} || ***pɐt^{1}*** fu:^{6} | tɕa:k^{3} jɐm^{6} kɛ:^{3} | kɔ:ŋ^{2} fa:t^{3} ||
|| ___ | _ ___ |__ || ____ _ |____ _ _ | _____ __ ||
I-feel-able-p.-not-bear-duty-post-p.-talk-way
(I think that is an irresponsible proposal)

4.5.4.4. Tonal Values

If there are two identical/similar tones at the beginning of each of the constituent intonation groups, the absolute F_0 values of the identical/similar tones usually decrease one after another. The comparison may not be reliable if made between the identical/similar tones which are not both at the same position (i.e., both at the beginning or both at the end) of the successive intonation groups. The comparison may also not be reliable even if it takes place between identical/similar tones which are both at the end of successive intonation groups. As we know, a tone located at the end of an intonation group can vary its F_0 value heavily as a result of suffering from a steep declination. In the example above, the major intonation group 1 has identical T1 at the antepenultimate/last syllable of the intonation groups 1, 2, 3

and 5. The F_0 peak value of the identical tones (i.e., the syllables /nɛ:^{1}/, /tow^1/, /kɛ:ŋ^1/ and /ɕi:^{1}/) in the four intonation groups is 194Hz, 185Hz, 160Hz and 132Hz, respectively, in descending order. However, the comparison cannot be made between the two T1s - /kɛ:ŋ^1/ which occur twice in the major intonation group 1. Its F_0 value is 160Hz when occurring at the middle syllable of the three syllables of the third intonation group and 165Hz occurring at the second syllable of the five syllables of the fifth intonation group. The following is an example of the major intonation group 18 used in the comparison of the pitch of identical tones in the last syllables of each of the three intonation groups. The last syllables in each intonation group is T1, but the absolute values are progressively lower: 210Hz /tɐk^1/ for the first intonation group, 124Hz /nɛ:^{1}/ for the second and 92Hz /tɔ:^{1}/ for the last.

The major intonation group 18:

210Hz 124Hz 92Hz

||ŋɔ:^{5}kɔ:k^{3}tɐk^1||***kɐm***^{2}jœ:ŋ^{2}kɛ:^{3}|ka:w^{3}hɔ:k^6|wa:n^{4}keŋ^{2}nɛ:^{1}||k^{h}ej^{4}ɕɐt^6 tɕɐw^6|***how***^{2}how^{2}tɔ:^{1}||

I-feel-able-this-m-p-teach-learn-surounding-environment-p-that-fact-then-good-very-much

(I think this sort of educational environment was in fact much better)

4.5.5. Anacrusis

4.5.5.1. Defining Anacrusis

When a major intonation group consists of several intonation groups, the first intonation group usually has the highest pitch setting regardless what lexical tones it contains. This intonation group, which differs from the following intonation groups by its outstanding high pitch setting, is called the anacrusis. The term is adopted from Chao (1932: 108). In discussing American English, Chao distinguishes the unstressed syllables occurring before the first stressed syllable from the rest of the syllables owing to their different pitch. He calls this chunk the anacrusis. The anacrusis in Cantonese is usually short, only consisting of one or two feet. The most striking high pitch setting distinguishes it from the rest of the following intonation groups. Whether in an even fall or an uneven fall as discussed above, the first constituent

intonation group occupies the highest pitch setting. We can see from the example above, the major intonation group 18, that the ratio of the F_0 value (210Hz, 124Hz and 92Hz) between the identical tones at the last syllables of the three intonation groups (/tɐk^1/, /nɛ:^{1}/, and /tɔ:^{1}/) is 2.3 : 1.3 : 1. The F_0 difference of the identical tones between the second and the last intonation groups is not much, but the difference is great between the anacrusis and the last two intonation groups.

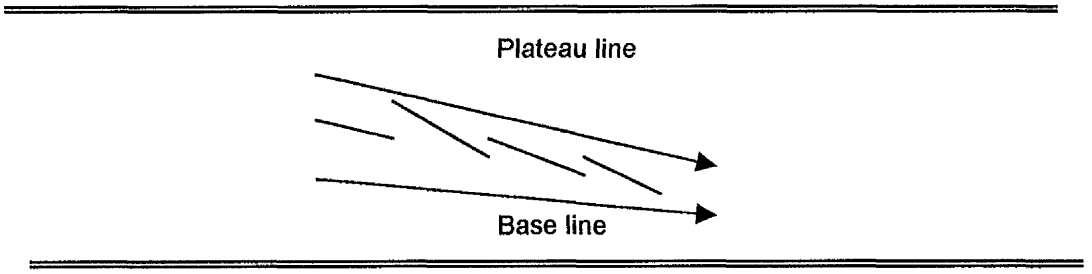
4.5.5.2. Crippling Anacrusis

Sometimes the anacrusis displays a pitch movement of 'crippling up'. In this case, the anacrusis consists of two intonation groups, the second one is higher than the first one - the F_0 maximum falls in the second intonation group. If a major intonation group has a crippling anacrusis, the intonation group immediately following the anacrusis is likely to contain the major tonic. For example:

The major intonation group 8:

||***ta:n***^{6}hɐj^{6}nɛ:^{1}||tɕɐw^6***ki:n***^{3}la:p^6|tɕɔ:^{2}kɔ:^{3}||***how***^{2}how^2|kɛ:^{3}||***p***^{h}***ɐŋ***^{4} jɐw^5|k^{w}a:n^1 hɐj^{6}la:k^3||
|| ___ _ _ || _ ___ _ | __ _ || __ _ | __ || ___ _ | ___ _ _ ||
but-is-p.-then-build-establish-ed-m.-very-good-p-friend-friend-concern-relate-p.
(However, we were on very friendly terms)

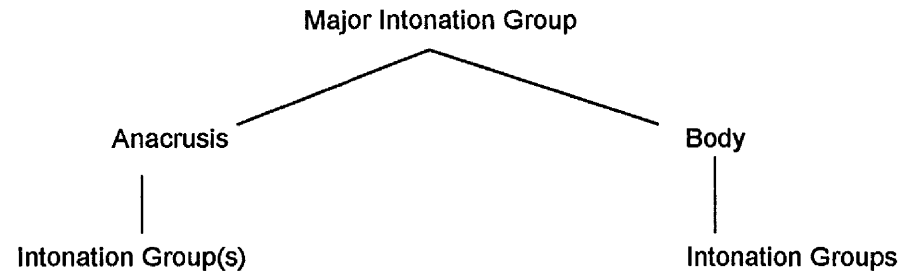
The following figure demonstrating the F_0 maximum falls in the second constituent of the crippling anacrusis.



4.5.6. Structure of Major Intonation Group

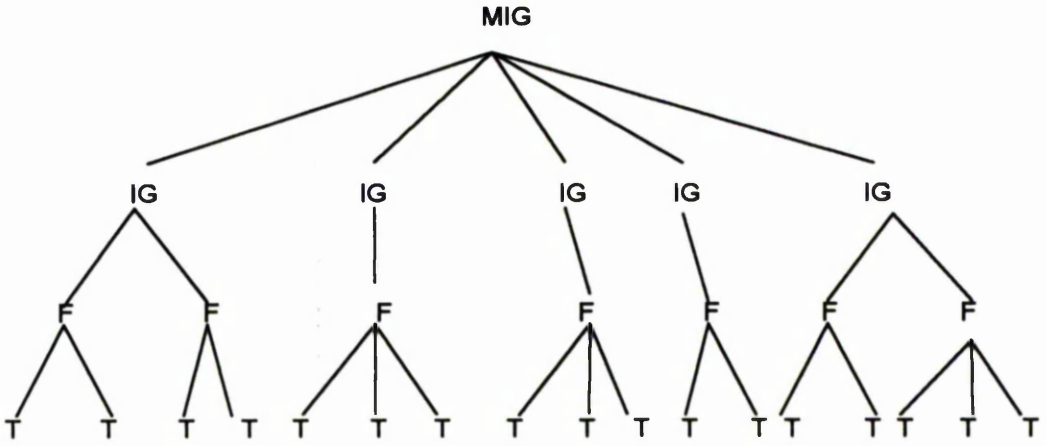
In discussing intonation structure in English, Palmer (1922: 92) divides an utterance into head, nucleus and tail. Chao (1932: 109) modifies that into anacrusis, main head, and body. His body includes nucleus and tail in Palmer's terms, and Palmer's head includes anacrusis and main head in Chao's terms. Halliday (1970: 5) suggests that a tonic extends from the tonic syllable right up to the end of the tone group, and this may be preceded by a pretonic. His pretonic is similar to Palmer's head, and his tonic is similar to Chao's body. I shall divide a major intonation group in Cantonese into anacrusis and body only. Each anacrusis includes one intonation group, or a crippling anacrusis with two intonation groups, which occupies the outstanding by high pitch setting. The rest of the intonation groups make up the body. The major tonic usually falls in the body. Within the body, the pitch of each intonation group is in descending order. In the case of a major intonation group which consists of only one intonation group, there is no anacrusis.

Fig.4.19. Structure of a major intonation group



We can speak of hierarchical structure of a major intonation group because it consists of one or more intonation groups. The hierarchy of prosodic categories is: the major intonation group (MIG), the intonation group (IG), the foot (F) and the syllabic tone (T). The common structure for a foot is a lengthened and a shortened syllable. There must be a tonic syllable in an intonation group and a major tonic syllable in a major intonation group. The example above from the major intonation group 1 appears to be like this:

Fig. 4.20. Structure of the prosodic hierarchy of a major intonation group.



||***kɐm***^{2} ŋɔ:^{5} | nɛ:^{1}tɕɐw^{6}|| ***<u>tɕ^{h}oŋ^{4}</u>*** lɔ:j^{4} tow^{1} ||m^{4} ***kɛ:ŋ^{1}*** hɔ:k^{6}||***ɕ***^{6} ɛ:^{6}|| m^{4} kɛ:ŋ^{1}| ka:w^{3} ***ɕi:***^{1} kɛ:^{3}||

in that case-I-p.-then-from-come-also-not-afraid-study-em-not-afraid-teach-expert-p.

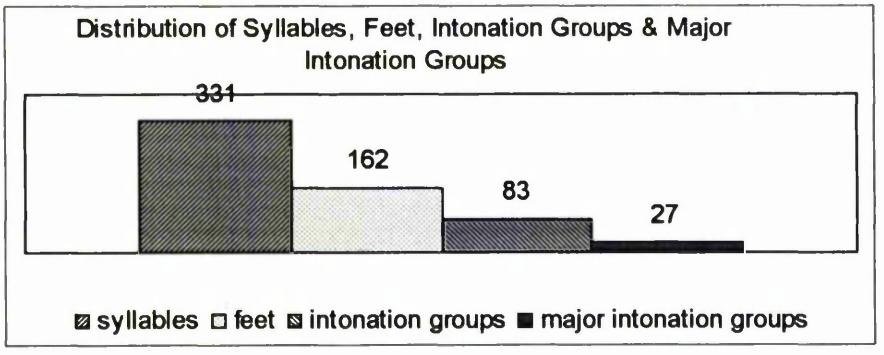
(I was never frightened of stu- ... never frightened of teachers)

4.5.6.1. The Phonological Divisions of the Utterances

(Note: | - foot boundary; || - intonation group boundary; % - pause;
bold and italic - tonic; underlined – major tonic;
the underlining "__" - a syllable, __ - shortened syllable, _____ - lengthened syllable;
numbers in brackets - the numbers of major intonation groups;
numbers in italic - the numbers of intonation groups;
A, B and C - the three speakers)

The following list by phonological division of the utterances corresponds to the material in section 4.2. The distribution of syllables, feet, intonation groups and major intonation groups in our data used for acoustical analysis are listed in the following figure.

Fig. 4.21. The distribution of syllables, feet, intonation groups and major intonation groups used for acoustical analysis.



[1] A:

1 || ***kɐm²*** ŋɔ:⁵ | nɛ:¹ tɕɐw⁶ ||
|| ____ _ | _ ____ ||
in that case–l-p.–then
(I)

2 || ***tɕʰoŋ⁴*** lɔ:j⁴ tow¹ ||
|| ____ __ __ ||
from-come-also
(always)

3 || m⁴ ***kɛ:ŋ¹*** hɔ:k⁶ ||
|| __ ____ __ ||
not–afraid–study
(not frightened of stu-)

4 || ***ɕ⁶*** ɛ:⁶ ||
|| ____ __ ||
stu-
(stu-)

5 || m⁴ kɛ:ŋ¹ | ka:w³ ***ɕi:¹*** kɛ:³ ||%
|| __ ____ | __ ____ _ || %
not–afraid–teach–expert-p.
(not frightened of teachers)

[2] B:

6 || ***m⁴*** m² ||
|| ____ _ ||
mh-mh
(Mm, hmm)

[3] A:

7 || kɐm² | ***ka:w³*** ɕi:¹ nɛ:¹ ||
|| ^ _ | ____ __ __ ||
in that case-teach-expert-p.
(And, teachers)

8 || ***tɕʰoŋ⁴*** lɔ:j⁴ nɛ:¹ ||
|| ____ __ __ ||
from-come-p.
(always)

9 || jek⁶ ***tow¹*** ||
|| __ ____ ||
also-also
(also)

10 || ***mow⁵*** mɐt¹ | ti:m² fɐt⁶ ||
|| ____ _ | __ ____ ||
have not-any-how-punish
(never really punished)

11 || ***ŋɔ:⁵*** tej⁶ kɛ:³ ||
|| ____ __ __ ||
I-pl.-p.
(us)

[4]

12 || jɐt[1] ***lɔ:j[4]*** ||
|| _ ___ |
one-come
(In the first place)

13 || ŋɔ:[5] jek[6] | ***hɔ:[2]*** nɐŋ[4] ||
|| ____ _| ___ _ ||
I-also-possible-can
(it was possibly)

14 || jek[6] tow[1] ***hɐj[6]*** ||
|| _ _ ___ ||
also-also-am
(because I was)

[5] B:

15 || ***k[w]a:j[1]*** ha:[2] ||
|| ___ _ ||
good-int.
(good)

[6] A:

16 || mow[5] wɔ:ŋ[2] | ***k[w]u:n[2]*** | la:[1] ha:[2] ||
|| _ ____ | ____ | ___ _||
have not-king-control-p.-int
(uncontrollable)

[7]

17 || a:[4] tɕek[1] | kɔ:[3] kɔ:[3] tow[1] | ***k[w]u:n[2]*** m[4] ti:m[6] | la:[1] ha:[2]||
|| ____ _| ____ _ _ | _____ _ ___ |_ ____||
oh-i.e.-m.-m.-all-control-not-o.k.-p.-int.
(That means no one was able to be in control at all)

[8]

18 || ***ta:n[6]*** hɐj[6] nɛ:[1]||%
|| ___ _ _ ||
but-is-p.
(However)

19 || tɕɐw[6] ***ki:n[3]*** la:p[6] | tɕɔ:[2] kɔ:[3]||
|| __ ____ __ |____ _ ||
then-build-establish-ed- m.
(we were on)

20 || ***how[2]*** how[2] | kɛ:[3]||
|| ___ _ | ___||
very-good-p.
(very)

21 || ***p[h]ɐŋ[4]*** jɐw[5]| k[w]a:n[1] hɐj[6] la:k[3]||
|| ____ _ | ____ _ _ ||
friend-friend-concern-relate-p.
(friendly terms)

[9]

22 || tɕek^{1} ***tɕ^{h}ɵj^{4}*** tɕɔ:^{2} ||
|| _ ____ _ ||
i.e.-apart-p.
(This means, apart from)

23 || ***ɕi:^{1}*** ɕɐŋ^{1} | k^{w}a:n^{1} hɐj^{6} nɛ:^{1} ||%
|| ___ _ | ____ _ _ ||
teacher-student-concern-relate-p.
(the teacher-student relationship)

24 || kɔ:^{3} ***p^{h}ɐŋ^{4}*** jɐw^{5} | k^{w}a:n^{1} hɐj^{6} | how^{2} how^{2} ||
|| _ ____ _ | ___ _ | _ ___ ||
m.-friend-friend-concern-relate-very-good
(the friendship was very strong)

[10]

25 || wu:^{6} ***ɕɶ:ŋ^{1}*** | tɕy:n^{1} tɕoŋ^{6} ||%
|| _ ____ | ____ _ ||
mutual-mutual-respect-respect
(mutual respect)

[11]

26 || jek^{6} tow^{1} | ***tɕ^{h}oŋ^{4}*** lɔ:j^{4} nɛ:^{1} ||
|| _ ____ | ____ _ _ ||
also-also-from-come-p.
(never)

27 || ŋɔ:^{5} ***tɕɐw^{6}*** ŋ^{6} ||
|| _ ____ _ ||
I-then-hm
(I)

28 || ***mow^{5}*** | ji:^{4} ka:^{1} | kɔ:^{2} ti:^{1} ||
|| ____ | _ ___ | ___ _ ||
have not-now(2)-that-pl.
(unlike nowadays)

29 || ***ka:w^{3}*** ɕi:^{1} | pej^{3} jy:^{4} ||
|| ____ _ | _ ___ ||
teach-expert-example-as
(those teachers, for example,)

30 || ka:w^{3} | ka:w^{3} | iɛ:^{6} | ***p^{h}eŋ^{4}*** wu:j^{2} | k^{h}ɵy^{5} tej^{6} wa:^{6} ||%
|| ____ | ____ | ____ | _____ _ | _ _ ____ || %
Teach-teach-em-assess-committee-he-pl.-say
(the Body of Educational Assessment say)

[12]

31 || jɐw^{5} how^{2} tɔ:^{1} | hɔ:k^{6} ***ɕa:ŋ^{1}*** ||
|| ___ _ _ | _ ____ ||
have-very-many-study-pupil
(many students)

32 || tow^{1} m^{4} | ***tɕy:n^{1}*** tɕoŋ^{6} | low^{5} ɕi:^{1} nɛ:^{1} ||%
|| _ ___ | ___ _ | _ ___ _ ||%
all-not-respect-respect-old-expert-p.
(don't respect teachers)

[13]

33 || kɔ:^{2} ***tɕoŋ^{2}*** ||
|| _ ___ ||
that-m.
(such a)

34 || tɕ^{h}eŋ^{4} fɔ:ŋ^{3}| jɐw^{6} m^{4}| tɕ^{h}yn^{4} ***tɕ^{h}ɔ:j^{6}*** kɔ:^{3} wɔ:^{3} ||%
|| ____ _ | _ ___ | _ ____ _ _ ||
condition-situation-also-not-exist-be there-p.-p.
(situation did not exist)

[14]

35 || ***ta:j^{6}*** ka:^{1} ||
|| ___ _ ||
big-family
(Everybody)

36 ||tow^{1} how^{2} | ***tɕy:n^{1}*** tɕoŋ^{6} ||
|| _ ____ | ____ _ ||
all-very-respect-respect
(respected)

37 || low^{5} ***ɕi:^{1}*** | kɔ:^{3} wɔ:^{3} ||%
|| _ ___ | _ __ ||%
old-expert-p.-p.
(teachers very much)

[15] C:

38 || m^{6} | ***how^{2}*** tɕ^{h}i:^{5} ||
||__ | ____ _ ||
mh-very-like
(Right! The same as)

39 || p^{h}ɐŋ^{4} jɐw^{5} | kɐm^{2} jœ:ŋ^{2} | ***tɕy:n^{1}*** tɕoŋ^{6} ha:^{2} ||
|| ___ _ | ____ _ | ____ _ _ ||
friend-friend-this-look-respect-respect-int
(friends who respect each other)

[16] A:

40 || ***how^{2}*** tɕ^{h}i:^{5} ||
|| ____ _ ||
very-like
(The same as)

41 || ***p^{h}ɐŋ^{4}*** jɐw^{5} | kɐm^{2} jœ:ŋ^{2} ||
|| ___ _ | _ ____ ||
friend-friend-this-look
(friends)

[17]

42 || ***ta:j^{6}*** ka:^{1}||
|| ___ _ ||
big-family
(Everybody)

43 || k^hɐp^1 ɕɐ***w^1*** | kɛ:^{3} jɛ:^{5} nɛ:^{1} ||
|| _ _____ | _ ____ _ ||
suck-receive-p-things-p.
(learned)

44 || jek^6 tow^1 | how^2 ***fa:j^3*** ||
|| _ ___ | _ ____ ||
also-all-very-quick
(really fast)

[18]

45 || ***ŋɔ:^{5}*** kɔ:k^3 tɐk^1 ||
|| ___ _ _ ||
I-feel-able
(I think)

46 || ***kɐm^2*** jœ:ŋ^2 kɛ:^{3} | ka:w^3 hɔ:k^6 | wa:n^4 keŋ ^{2} nɛ:^{1} ||
|| ____ _ _ | ____ _ | ____ _ _ ||
this-m.-p.-teach-learn-surrounding-environment-p.
(this sort of educational environment)

47 || k^hej^4 ɕɐt^6 tɕɐw^6 | ***how^2*** how^2 tɔ:^{1} ||%
|| _ _ ____ | ____ _ _ ||
that-fact-then-good-very-much
(was in fact much better)

[19]

48 || ***ɕɔ:^{2}*** ji:^{5} nɛ:^{1} ||%
|| ___ _ _ ||%
There-fore-p.
(Therefore)

49 || ***ŋɛ:^{6}*** ||
|| ___ ||
em
(em)

50 || kɔ:k^3 wɐj^2 | ka:w^3 ɕi:^{1} | ***nɛ:^{1}*** ||
|| ____ _ | ____ _ | ____ ||
every-m.-teach-expert-p.
(teachers)

51 || tɕɐw^6 m^4 ɕɐj^2 | kɐm^3 ta:m^1 ɕɐm^1 | tɕi:^{6} kej^2 kɛ:^{3} | ***k^hy:n^4*** wɐj^1 ||
|| _ _ ___ | _ ____ _ | _ ____ _ | ____ _ ||
then-no-need-so-bear-heart-self-self-p.-power-might
(you shouldn't worry about your authority)

[20]

52 || jɐn^1 wɐj^6 ŋɔ:^{5} | kɔ:k^3 ***tɐk^1*** ||
|| ___ _ _ | _ ___ ||
because-for-I-feel-able
(Because I think)

53 || how^2 how^2 | ***ɕi:^{3}*** kɛ:^{3} ||
|| ____ _ | ____ _ ||
very-good-laugh-p.
(it is silly)

54 || k^hej^4 ɕɐt^6 ŋɔ:^{5} ||
|| _ _ ___ ||
That-fact-I
(Actually I)

55 || ɕœ:ŋ^2 kɔ:ŋ^2 | jɐt^1 jœ:ŋ^6 | jɛ:^{5} nɛ:^{1} ||%
|| ____ _ | _ ____ | ___ _ ||%
want-say-one-m.-thing-p
(just want to say one thing)

[21]

56 || tɕɐw^6 ŋɔ:^{5} | $t^hɐj^2$ kɔ:^{3} ||%
|| ___ _ | ____ _ ||%
concern-I-see-m.
(I see)

57 || a^3 ||
|| ___||
ah
(that)

58 || kɔ:^{3} ti:w^6 $tɕ^ha:^4$ ||
|| _ _ ____ ||
m.-investigate-check
(investigations)

59 || hi:n^2 ɕi:^{6} | a^3 ||%
|| ____ __ | ___ ||%
appear-show-p.
(have shown)

[22]

60 || ɛ:^{6} ||
|| ___||
hm
(that)

61 || ji:^{4} ka:^{1} | ŋɔ:^{5} tej^6 | ka:w^3 ɕi:^{1} nɛ:^{1}||
|| ___ _ | _ ___ | ____ __ __ ||
now(2)-I-pl.-p. -teach-expert-p.
("nowadays (the authority of) us teachers)

62 || kɔ:^{2} kɔ:^{3} ||
|| ___ _ ||
that-m.
(that)

63 || $k^hy:n^4$ wɐj^1 | nɛ:^{1}||
|| ____ _ | ___ ||
power-might-p.
(authority)

64 || jɐt^6 $tɕi:m^6$ | ha:^{6} | kɔ:ŋ^3 ||%
|| _ ____ | ____ | ____ ||%
day-gradually-down-sink
(is on the decline)

[23]

65 || ŋɔ:^{5} ***kɔ:k^{3}*** tɐk^{1}||%
|| _ ____ _ ||%
I-feel-able
(I think that)

66 || ***t^{h}ɐj^{4}*** ji:^{5} | ni:^{1} kɔ:^{3} ||
|| ____ _| _ ____ ||
lift-discuss-this-m.
(whoever suggested the)

67 || ***t^{h}ɐj^{4}*** mok^{6} kɛ:^{3} | ti:w^{6} tɕ^{h}a:^{4} tow^{1}||
|| ____ __ __ |____ __ __ ||
title-topic-p.-investigate-check-all
(topic of this investigation)

68 || tɐj^{2} ta:^{2} | lɔ:^{1} jɐw^{2} | ***la:^{1}*** ||
|| ____ _| _ ____ | ___ ||
deserve-beat-bottom(2)-p.
(should have his bottom smacked)

[24]

69 || ***ka:w^{3}*** ɕi:^{1} | ti:m^{2} ka:j^{2} | ji:w^{3} jɐw^{5} ||
|| ____ _ | __ _____ | ____ __ ||
teach-expert-how-explain-need-have
(Why do teachers need to have)

70 || k^{h}y:n^{4} wɐj^{1} | ɕi:n^{1} tɐk^{1} | ***ka:^{2}*** ||%
|| ____ __ | ____ _ | ___ ||
power-might-until-o.k.-p.
(authority anyway)

[25] C:

71 || hɐj^{3} ***a:3*** ||
|| _ ___ ||
right-p.
(That's right)

72 || ***kɐm^{2}*** | t^{h}oŋ^{4} ma:j^{4} |wa:^{3} ||
|| ___ | ___ _ | ___ ||
in that case-and-near-say
(And)

73 || ***ti:^{1}*** hɔ:k^{6} ɕa:ŋ^{1} | jɐj^{5} nɛ:^{1} ||
|| ___ _ _ | _ ___ ||
pl.-study-pupil-naughty-p.
(those students were naughty)

74 || ŋɔ:^{5} tej^{6} | ***ji:m^{4}*** tɕoŋ^{6} | fa:n^{6} k^{h}ɐj^{1} ||
|| _ ___ |___ _ |_ ___ ||
I-pl.-serious-heavy-offence-regulation
(We committed a serious breach of the rule)

[26]

75 || ŋɔ:^{5} tej^{6} | ***hɔ: ɾ^{1}*** tɕ^{h}ɵy^{4} k^{h}ɵy^{5} ||
|| ___ _ | ___ __ __ ||
I-pl.-open-expel-him
(We expelled him)

[27]

76 || hɔ:j^{1} tɕ^{h}ɵy^{4} tɕɔ:^{2} | m^{4} **pi:w^{2}** ɕi:^{6} | mɐn^{6} thɐj^{4} | ka:j^{2} k^{h}y:t^{3} | tɕɔ:^{2} a:^{3} ||
|| ____ _ __ | _ ____ _ | ____ _ | ____ _ | ____ _ ||
open-expel-ed-not-show-show-ask-title-unlock-firm-ed-p.
(Expelling him does not mean the problem has been solved)

[28] A:

77 || hɐj^{3} ***a:^{3}*** ||
|| _ ___ ||
right-p.
(That's right)

78 || m^{4} t^{h}oŋ^{1} | k^{h}ɵy^{5} m^{4} | tok^{6} ***ɕy:^{1}*** | lɔ:^{3} wɔ:^{3} ||
|| _ ___ | _ ___ | _ ___ | _ ___ ||
not-through-he-not-read-book-p.-p.
(Does he not study)

[29]

79 || ***k^{h}ɵy^{5}*** ||
|| ____ ||
he
(He)

[30] C:

80 || ***tɕoŋ^{6}*** | tɕ^{h}a:m^{2} ||
|| ____ | ____ ||
more-tragedy
(Much worse)

[31] A:

81 || ji:^{5} hɐw^{6} kɛ:^{3} | tɕ^{h}i:n^{4} tkow^{4} | tɕɐw^{6} ***mow^{5}*** | tɕɔ:^{2} lɔ:^{3} wɔ:^{3} ||
|| _ _ ____ | ____ _ | _ ____ | ____ _ _ ||
in-back-p.-front-journey-then-not have-ed-p.-p.
(His future would have been doomed)

[32] C:

82 || ***hɐj^{6}*** la:^{3} ||
|| ___ _ ||
right-p.
(Right)

[33] A:

83 || ***m^{4}*** tɐk^{1} | ka:^{1} ma:^{3} ||
|| ____ _ | _ ____ ||
not-o.k.-p.-p.
(You cannot do that)

84 || jɐn^{1} wɐj^{6} ***nej^{5}*** | kɔ:k^{3} tɐk^{1} k^{h}ɵy^{5} ||
|| _ _ ____ | _ _ ____ ||
because-for-you-feel-able-him
(Because you think he)

[34] C:

85 || nej^{5} ***kɔ:k^{3}*** tɐk^{1} |k^{h}ɵy^{5} wa:j^{6} jɐn^{4} ||
|| _ ____ _ | _ _ ____ ||
you-feel-able-him-bad-person
(You think he is a bad person)

[35] A:

86 || hɐj^{6} ***a:k^{3}*** ||
|| _ ____ ||
right-p.
(Right)

87 || nej^{5} kɔ:k^{3} tɐk^{1} | k^{h}ɵy^{5} ***jɐj^{5}*** ||
|| __ _____ __ | __ ____ ||
you-feel-able-him-naughty
(You think he is naughty)

88 || ***ŋɐj^{3}*** ||
|| ____ ||
eh
(eh)

89 || mow^{5} tɐk^{1} | ***ka:w^{3}*** lɔ:^{3} ||
|| ____ _ | ____ _ ||
not-able-teach-p.
(unteachable)

[36]

90 || ***hɔ:j^{1}*** tɕ^{h}ɵy^{4} | k^{h}ɵy^{5} la:^{1} ||
|| _____ __ | __ _____ ||
open-expel-him-p.
(then expel him)

91 || kɐm^{2} ***k^{h}ɵy^{5}*** | ti:m^{2} a:^{3} ||%
|| _ ____ | __ ____ ||%
in that case-he-how-p.
(So, what will he do then)

92 || tɕow^{6} | ***ka:j^{1}*** t^{h}oŋ^{4} | a:^{4} ||%
|| ^ __ | _____ __ | __ ^ ||
to be-street-child-p.
(He'll be out on the streets)

[37]

93 || ***ŋɔ:^{5}*** | kɔ:k^{3} tɐk^{1} | a:^{3} ||
|| ___ | _ ____ | ___ ||
I-feel-able-p.
(I think)

94 || ***pɐt^{1}*** fu:^{6} | tɕa:k^{3} jɐm^{6} kɛ:^{3} | kɔ:ŋ^{2} fa:t^{3} ||%
|| ____ _ | ____ _ _ | _____ __ ||%
not-bear-duty-post-p-talk-way
(that is an irresponsible proposal)

[38]

95 || ***keŋ^{2}*** ji:n^{4} hɐj^{6} ||
|| ____ _ _ ||
unexpectedly-look-to be
(and a very surprising one)

96 || tɕ^{h}ɐt^{1} ***tɕɔ:j^{6}*** ||
|| _ ___ ||
out-at
(coming from)

[39]

97 || ***tɕ^{h}ɐt^{1}*** ji:n^{6} hɐj^{2}||
|| ___ _ _ ||
out-present-at
(coming from)

98 || ni:^{1} kɔ:^{3} | ka:w^{3} jok^{6} | ***p^{h}eŋ^{4}*** ji:^{5} | wu:j^{2} ||
|| _ __ | ___ _ | ___ _ | ___ ||
this-m.-teach-cultivate-assess-discuss-committee
(the Body of Educational Assessment)

99 || tɕɐn^{1} | ***tɐj^{2}*** | ta:^{2} | lɔ:^{1} | jɐw^{2} || %
|| ¹___ | ¹___ | ¹___ | ¹___ | ¹___||%
really-deserve-beat-bottom(2)
(They all really deserve to have their bottoms smacked)

[40]

100 || tɕoŋ^{6} ***ka:w^{2}*** tow^{3} ||
|| _ ___ _ ||
still-do-reach
(they even)

101 || ***hɔ:j^{1}*** | kej^{3} | tɕɛ:^{2} ||
|| ¹___ | ¹___ | ¹___ ||
open-record-person
(held a press)

102 || tɕi:w^{1} | tɔ:j^{6} | ***wu:j^{2}*** ||
|| ¹___ | ¹___ | ¹___ ||
beckon-treat-meeting
(conference)

103 || ***t^{h}i:m^{1}*** ||%
|| ¹___ ||%
p
(!)

CHAPTER FIVE

CONCLUSION

In this thesis we have offered a phonetic account of Cantonese intonation and a characterisation of inter-tone influences and of prominence which both contribute to the characterisation of the phonetic account. To my knowledge no such account has ever been undertaken before in relation to Cantonese. The phonetic account is based on the framework established in Halliday (1967 and 1970), which chunks off connected speech into variable size constituents. There are three different constituent sizes which are the elements of phonetic account in our discussion: the foot, the intonation group and the major intonation group.

The first of the constituents is the foot which divides utterances into isochronous periods, commonly made up of a lengthened syllable accompanied by a shortened syllable or syllables. The sequences of alternating lengthened and shortened syllables in actual speech are linked into pairs which we call "feet". The lengthened syllable is where the prominence lands, and the rhythm is called the "prominence-based rhythm". We argue that in utterances syllables are freely stretched and compressed to accommodate alterations in rhythm and prominence, regardless of their phonological longness or shortness, resulting in pairs of alternating lengthened and shortened syllables. The phonologically long and short syllables are never realised in actual speech and syllables in train never occupy equal time.

The second of the constituents is the intonation group which is composed of one or more feet, in which all syllabic tones keep the same melodic "key". Within a "key", the pitch heights of identical phonological tones are found in descending order. Shifting the melodic "key" up and down is the demarcation between intonation groups. Different "keys" have different pitch span. An intonation group is also a piece of information which is coherent in terms of

semantics and syntax. A tonic prominence is realised by the most lengthened syllable and is the most important piece of information in the intonation group.

The third of the constituents is the major intonation group which is the largest prosodic constituent in an utterance, composed of one or more intonation groups, in which the changes in melodic "key" are in descending order. A larger scale of pitch resetting is the demarcation between major intonation groups. A major intonation group usually represents a piece of complete information, usually corresponding to a clause or a short sentence, but it is not determined by syntactic factors. In a major intonation group, a major tonic is the prominence realised in the most lengthened syllable. Within a major intonation group, the phonetic value of phonologically identical tones at the same position (e.g., at the beginning) of the constituent intonation groups decreases.

The division of the intonation contour of the utterance in Cantonese into foot, intonation group and major intonation group; defining what criteria those 'chunks' meet and how they relate to each other; assigning the tonic prominence of those 'chunks' and defining the criteria applied to this; - these are the contributions of this thesis to describing intonation in Cantonese.

Prominence, which contributes to the characterisation of the phonetic representation, is realised in terms of the duration rather than the pitch in Cantonese, due to the fact that phonological tones already have heavy function load in pitch to maintain their identity. Prominence variance at lexical level and utterance level is examined. It can fall on any syllable in a word and in an utterance. Lexical prominence is concerned with the formation of the word (e.g., the position of the syllable in a compound, a reduplication, a polysyllabic monomorpheme or a phrase) and the tonal constituents of the word (e.g., the extreme high and low pitch tones in onomatopoeia, reduplications and /pi:n^{3} jɐm^{1}/ (changed-tone)), and it is ultimately governed by contextual prominence which is realised as tonic prominence. Prominence marks the most important piece of information in an utterance. It is achieved by primary longer duration, which can also be accompanied by greater loudness and the exaggeration of the pitch, interweaving with the lexical tones.

Tones influence each other when they are adjacent, which contributes to the characterisation of the phonetic representation, resulting in modification of the tonal contour and pitch height. The modification in height is dissimilatory and encompasses the whole of the duration of the syllable, while the modification in slope is assimilatory and is concentrated in the area of tonal transition – between a quarter or a half of the duration of the preceding tone and between a quarter or a half of the following tone. Carryover coarticulatory effects are greater than anticipatory effects. The most stable area which is not in the domain of tonal transition and is less affected by modification is around the central portion of the tonal contour, called the gravity, which is where a tone maintains its identity. The phonological tones do not lose their contrast when they are in a string of identical or different tones.

An intonation contour in Cantonese is composed of a succession of phonological tones which are grouped into feet, intonation groups and major intonation groups. The phonological tones can be shifted up or down, and can be lengthened or shortened, and can be modified in the light of tonal coarticulation, prominence variation, the position in the intonation group and the major intonation group.

BIBLIOGRAPHY

ABERCROMBIE, DAVID 1967 *Elements of General Phonetics*. Edinburgh University Press.

ANSRE, GILBERT 1963 Reduplication in Ewe. *Journal of African Languages* 2: 128-132.

BAUER, ROBERT S 1979 Alveolarization in Cantonese: A Case of Lexical Diffusion. *Journal of Chinese Linguistics* 7: 132-43.

1982 *Cantonese Sociolinguistic Patterns: Correlating Social Characteristics of Speakers with Phonological Variables in Hong Kong Cantonese.* Ann Arbor: University Microfilms International.

1983 Cantonese Sound Change across Subgroups of the Hong Kong Cantonese. *Journal of Chinese Linguistics.* 11. 2: 301-354.

1984 The Expanding Syllabary of Hong Kong Cantonese. Paper in *The xvith International Conference on Sino-Tibeta: 1-16.*

BAUER, R S & P K BENEDICT 1997 *Modern Cantonese Phonology*. Berlin: Mouton de Gruyter.

BARNETT, K M A 1949-1950 A Transcription for Cantonese Notes on Mr Yuen Ren Chao's 'Cantonese Primer'. *Bulletin of SOAS* XIII: 724-745.

BECKMAN, M & J EDWARDS 1994 Articulatory Evidence for Differentiating Stress Categories. In *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form:* 7-33. Ed. by Beckman M and J Kingston. . Cambridge University Press.

BECKMAN, M & J PIERREHUMBERT 1986 Intonation Structure in English and Japanese. *Phonology Yearbook* 3: 255-310.

BIRD, CHARLES S 1966 *Aspects of Bambara Syntax*. Unpublished Ph.D Thesis, UCLA.

BLACK J W, O TOSI, S SINGH & Y TAKEFUTA 1966 A Study of Pauses in the Reading of One's Native Language and in English. *Language Speech* 9: 237-241.

BLOOMFIELD, LEONARD 1935 *Language*. (reprinted 1973) George Allen & Unwin Ltd.

BOLINGER, DWIGHT 1972 Accent is predictable (if you're a mind-reader). *Language* 48: 633-44.

BROWMAN, C & L GOLDSTEIN 1986 Gest: Towards an Articulatory Phonology. *Phonological Yearbook* 3: 219-252

BRUCE, GOSTA 1977 *Swedish Word Accents in Sentence Perspective*. Lund: Gleerup.

1990 Alignment and Composition of Tonal Accents: Comments on Silverman and Pierrehumbert's Paper. In *Papers in Laboratory Phonology 1:* 107-114. Ed. by Kingston and Beckman. Cambridge University Press.

CEN, QIXIANG 1946 Guangzhou Yinxi Gaishuo (Outline of the Phonetic System of Cantonese). *Guangdong Jianshe* 1-2: 1-9, Guangzhou.

CHAN, K M K 1987 Tone and Melody in Cantonese. Paper to appear in the *Thirteenth Annual Meeting of the Berkeley Linguistics Society.*

CHANG, CHOIYEUNG 1987 Xianggong Yueyu Ziyin Cunzaide Yixie Wenti (Problems in Word Pronunciation in Hong Kong Cantonese). *Proceedings from the 1st International Conference on Yue Dialect 1987.* Ed. by Sin, Chow-Yiu. Hong Kong: Modern Education Research Society. (1994: 97-102.)

1992 Prescriptive or Descriptive? Attitude in Cantonese Dictionaries. *Departmental Journal - Dept of Chinese and Linguistics, City Polytechnic of Hong Kong:* 93-100.

1992 Zhi Yueyin Yidude Yizhong Changjian Xianxiang – Youbian Dubian (Divergences of the 'You Bian Du Bian' Method). *Chinese Language Advancement Bulletin 37:* 51-54. The Chinese Language Society of Hong Kong.

CHANG, CHOIYEUNG & LAIYIN CHEUNG 1987 *Guoyin Yueyin Tongyin Zihui (A Chinese Syllabary Pronounced According to Cantonese and Mandarin).* Hong Kong: Chunghwa Press.

CHANG, NIENCHUANG 1958 Tones and Intonation in the Chengdu Dialect. *Phonetica* 2: 59-84.

CHAO, YUANREN 1926 Beijing Suzhou Changzhou Yuzhucide Yan Jiu (Study of Helping Word in Dialects of Beijing, Suzhou and Changzhou). *Qinghua Journal* III: 2: 865-98.

1930 A System of Tone-letters. *Le Maitre Phonetique* 30: 24-27.

1932 A Preliminary Study of English Intonation (With American Variants) and Its Chinese Equivalents. T'sai Yuan Pei Anniversary Volume, supplementary vol. 1 of *Bulletin of the Institute of History and Philology of the Academica Sinica.* Peiping.

1947 *Cantonese Primer.* Harvard University Press.

1968 *A Grammar of Spoken Chinese.* (1985 3rd ed.) University of California Press.

CHEN, YUAN 1991 Bianyi Yu Yingbian - Tan Dangqian Zhongguo Yuyan Wenti (Language Change - Some Concerns about Contemporary Chinese Language). *Zhongguo Yuyan Tongxin* 20: 3-13, Hong Kong.

CHENG, TING'AU 1993 *Guangzhouhua Yanjiu Lunzhu Suoyin (A Selective Bibliography of Cantonese Linguistics: 1900-1993).* Hong Kong: Educational and Cultural Press.

CHEUNG, HUNGNIN S 1972 Xianggang Yueyu Yufan de Yanjiu (*The Study of Grammar in Hong Kong Cantonese).* The Chinese University of Hong Kong Press.

CHEUNG, KWANHIN 1986 *The Phonology of Present-Day Cantonese.* Unpublished Ph.D Thesis, University College London.

CHEUNG, LAIYIN 1983 Youyin Wuzi Yinjie Zongji (Tally of Syllables with Sound but No Written Form in Cantonese). *Yuwen Zazhi* 10: 28-35, Hong Kong.

1986 Guangzhouhua Gaosheng Biandiao Guilu Chuta (Preliminary Study of the Rule of Biandiao in Cantonese). *Yuwen Zazhi* 13: 41-45, Hong Kong.

CHEUNG, LAIYIN & CHOIYEUNG, CHANG 1987 *Guoyin Yueyin Tongyin Zihui (A Chinese Syllabary Pronounced According to Cantonese and Mandarin).* Chunghwa Press, Hong Kong.

CHEUNG, LAIYIN & LIEHUAI, NI 1999 *Gongzhi Guangzhouhua Cidian (Hong Kong Cantonese Dictionary).* Wanli Press, Hong Kong.

CHING, CHINGYUK T 1981 *Communication of Lexical Tone Patterns in Cantonese.* Unpublished Ph.D Thesis, University College London.

CLEMENTS, GEORGE N 1983 The Hierarchical Representation of Tone Features. *Current Approaches to African Linguistics* 1: 145-176, ed. by Ivan R. Dihoff.

CLEMENTS, G N & K C FORD 1979 Kikuyu Tone Shift and its Synchronic Consequences. *Linguistic Inquiry* 10: 2: 179-210.

CLUMECK H, D BARTON, M MACKEN & D HUNTINGTON 1981 The Aspiration Contrast in Cantonese Word-initial Stops: Data from Children and Adults. *Journal of Chinese Linguistics* 9: 211-25.

COHEN, A ET AL 1982 Declination: Construct or Intrinsic Feature of Speech Pitch? *Phonetica* 39: 254-273.

COLLIER, RENE 1975 Physiological Correlates of Intonation Patterns. *Journal of Acoustical Society of America* 58: 249-255.

1983 Physiological Explanations of Fo Declination. In A Cohen & M P R van den Broecke (eds.) *Abstracts of the 10th Intonational Congress of Phonetic Sciences:* 440, Dordrecht: Foris.

COSTER, D & P KRATOCHVIL 1984 Tone and Stress Discrimination in Normal Beijing Dialect speech. In *New Papers on Chinese Language Use:* 119-131. Ed. by B Hong, Australia National University Press.

CRYSTAL, D 1969 *Prosodic Systems and Intonation in English.* Cambridge University Press.

1985 *A Dictionary of Linguistics and Phonetics* (2nd ed). Oxford: Basil Blackwell.

CRUTTENDEN, A 1969 Review of Halliday 1967. *Journal of Linguistics* 5: 309-315.

1986 *Intonation.* Cambridge University Press.

DAVIES P, LINDSEY G A, FULLER H & A F FOURCIN 1986 Variation in Glottal Open and Closed Phase for Speakers of English. *Proceedings of the Institute of Acoustics* 8: 539-546.

DE FRANCIS, J 1963 *Beginning Chinese.* New Haven; London: Yale University Press.

DOW, D-M F 1972 *An Introduction to the Pronunciation of Chinese.* Britain: F D-M Dow.

DUNG DOTHE, THIENHUONG TRAN & GEORGES BOULAKIA 1998 Intonation in Vietnamese. In *Intonation Systems:* 395-416. Ed. by Hirst, D and A Di Cristo, Cambridge University Press.

FOURCIN A J & ERM ABBERTON 1971 First Applications of a New Laryngograph. *Medical and Biological Illustrated* 21: 172-182.

FLYNN, CHOI-YEUNG-CHANG see CHANG, CHOI-YEUNG

FODOR, J A, T G BEVER & M F GARRETT 1974 *The Psychology of Language:* 420-428. McGraw Hill, New York.

FOK, C Y-Y 1974 *A Perceptual Study of Tones in Cantonese.* University of Hong Kong.

FONAGY, IVAN 1966 Electro-physiological and Acoustic Correlates of Stress and Stress Perception. *Journal of Speech and Hearing Research* 9: 231-244.

FROMKIN, VIRGINIA A 1972 Tone Feature and Tone Rules. *Studies in African Linguistics* 3:47-76.

1978 *Tone: A Linguistic Survey.* (Ed) Academic Press, Inc.

FRY, D B 1955 Duration and Intensity as Physical Correlates of Linguistic Stress. *Journal of the Acoustical Society of America* 27: 765-768.

1958 Experiments in the Perception of Stress. *Language and Speech* 1: 126-152.

1965 The Dependence of Stress Judgements on Vowel Formant Structure. *Proceedings of the 5th International Congress of Phonetic Sciences:* 306-311, Karger.

1976 *Acoustic Phonetics.* (Ed) Cambridge University Press.

GANDOUR, JACK 1981 Perceptual Dimensions of Tone: Evidence From Cantonese. *Journal of Chinese Linguistics* 9:20-36.

GANDOUR, J, S POTISUK , W DECHONGKIT & W PONGIORPISIT 1992a Tonal Coarticulation in Thai Disyllabic Utterances: A Preliminary Study. *Linguistics of the Tibeto-Burman Area* 15: 93-110.

1992b Anticipatory Tonal Coarticulation in Thai Noun Compounds. *Linguistics of the Tibeto-Burman Area* 15: 111-124.

GANDOUR, J, S POTISUK & S DECHONGKIT 1994 Tonal Coarticulation in Thai. *Journal of Phonetics* 22: 447-492.

GARDING, E 1981 Contrastive Prosody: A Model and Its Applications. *Studia Linguistica* 35, 1-2: 146-165.

GAO, HUANIAN 1980 *Gguangzhou Fangyan Yanjiu (Studies on the Guangzhou Dialect).* Hong Kong: Commercial Press.

GUOYIN YUEYIN TONGYIN ZIHUI 1987 *(A Chinese Syllabary Pronounced According to Cantonese and Mandarin).* ed. by Chang, Choiyeung & Laiyin Cheung. Hong Kong: Chunghwa Press.

HALLIDAY, M A K 1967 *Intonation and Grammar in British English.* Mouton.

1970 *A Course in Spoken English.* Oxford University Press.

HAN, S & K-O KIM 1974 Phonetic Variation of Vietnamese Tones in Disyllabic Utterances. *Journal of Phonetics* 2: 223-232.

HASHIMOTO, O-K Y 1972 *Phonology of Cantonese.* Cambridge: Cambridge University Press.

HERTZ, SUSAN R 1990 The Delta Programming Language. In *Papers in Laboratory Phonology 1.* Ed. By Kingston, J and M Beckman, Cambridge University Press.

HO, AICHEN T 1977 Intonation Variation in a Mandarin Sentence for Three Expressions: Interrogative, Exclamatory and Declarative. *Phonetics* 34: 446-457.

HOMBERT, JEAN-MARIE 1978 Consonant Types, Vowel Quality and Tone. In *Tone: A Linguistic Survey:* 77-111. Ed by V Fromkin, Academic Press.

HOWARD, DAVID M 1995 Variation of Electrolaryngographically Derived Closed Quotient for Trained and Untrained Adult Female Singers. *Journal of Voice*: 9 (2) 163-172.

HYMAN, LARRY 1975 *Phonology: Theory and Analysis.* Holt, Rinehart & Winston.

1977 Studies in Stress and Accent. *South California Occasional Papers in Linguistics* IX.

INGHAM, BRUCE 1974 *The Phonology and Morphology of the Verbal Piece in an Arabic Dialect of Khuzistan.* Unpublished Ph.D Thesis, SOAS.

INKELAS, S & W LEBEN 1990 Where Phonology and Phonetics Intersect: The Case of Hausa Intonation. In *Papers in Laboratory Phonology 1 - Between the Grammar and Physics of Speech:* 17-34. Ed. by J Kingston & Beckman M, Cambridge University Press,

JOHNSON, K 1987 Intonation in Cantonese. *Working Papers in Linguistics – The Ohio State University*: 1-15.

JONES, Daniel 1940 *An Outline of English Phonetics* (6th ed.) New York: Dutton.

JONES, D & K-T WOO 1912 *A Cantonese Phonetic Reader.* London: University of London.

KAO, D L 1971 Structure of the Syllable in Cantonese. The Hague: Mouton.

KARLGREN, B 1915-26 *Etudes sur la Phonologie Chinoise*. Stockholm: Gotembourg.
1923 *Analytical Dictionary of Chinese and Sino-Japanese*. New York: Dover (1974).

KEATING, PATRICIA A 1990 The Window Model of Coarticulation: Articulatory Evidence. *Papers in Laboratory Phonology 1:*. 451-470. Ed. by Kingston J and M Beckman, Cambridge University Press.

KELLINGLEY, S Y 1985 *A new look at Cantonese tones: five or six?* Newcastle Upon Tyne: S Y Killingley.

KONG, PAULINE 1993 A Palatographic and Linguagraphic Study of [+coronal] Obstruents. In Hong Kong Cantonese. *Abstracts: the Fourth International Conference on Yue Dialects:* 80, Hong Kong: Linguistic Society of Hong Kong.

KONG, QINGMING 1987 Influence of Tones upon Vowel Duration in Cantonese. *Language and Speech:* 30, Part 4: 387-399.

KONG, YINGYEE 1993 VOT in Cantonese: Acoustical Measurements. *Abstracts: the Fourth International Conference on Yue Dialects:* 81, Linguistic Society of Hong Kong.

KRATOCHVIL, PAUL 1968 *The Chinese Language Today*. Hutchinson University Library London.
1998 Intonation in Beijing Chinese. *Intonation Systems:* 417-431. Ed. by Hirst D and A Di Cristo. Cambridge University Press.

KWOK, C H L 1984 *Sentence Particles in Cantonese*. Hong Kong: University of Hong Kong.

KWOK, C H L & K K LUKE 1986 Yueyu Yudiao Chutan (Preliminary Study in Cantonese Intonation). Hong Kong: *Yuwen Zazhi* 13: 32-40.

LADD, ROBERT D 1984 Declination: A Review and Some Hypotheses. *Phonology Yearbook* 1:53-74.
1996 *Intonational Phonology*. Cambridge University Press.

LADEFOGED, PETER 1975,1982 & 1993 *A Course in Phonetics* (1st, 2nd and 3rd ed). New York: Harcourt Brace Jovanovich.

LANIRAN, YETUNDE OLABISI 1993 *Intonation in Tone Language: The Phonetic Implementation of Tones in Yoruba*. Cornell University, Ithaca.

LASS, NORMAN J 1993 *Principles of Experimental Phonetics*. St Louis: Mosby-Year Book.

LAVER, JOHN 1994 *Principles of Phonetics*. Cambridge University Press.

LEA, WAYNE 1977 Acoustic Correlates of Stress and Juncture. In Hyman: *Studies in Stress and Accent:* 83-119.

LEE, XINGDE T 1983 The Vowel System in Two Varieties of Cantonese. *UCLA Working Papers in Phonetics* 57: 97-113.
1985 Guangzhouhua Yuanyinzhi Yinzhiji Changduan Duili (Opposition of Vowel Quality and Length in Cantonese). *Fangyan* 1 28-38.

LEHISTE, ILSE 1970 *Suprasegmentals.* The MIT Press.

LEHISTE, I & G PETRSON 1959 Vowel Amplitude and Phonemic Stress in American English. *Journal of the Acoustical Society of America* 31: 428-35. Reprinted in Fry 1976: 355-368.

1961 Some Basic Considerations in Analysis of Intonation. *Journal of the Acoustical Society of America* 33: 419-425.

LI, XINKUI & JIAJIAO HUANG ET AL 1995 *Guangzhou Fangyan Yanjiu (Study of the Cantonese Dialect)*. Shaogan: Guangdong Renmin Press.

LIEBERMAN, P 1967 *Intonation, Perception and Language*. Cambridge, MA: MIT Press.

LINDAU, MONA 1985 Testing a Model of Intonation in a Tone Language. *UCLA Working Papers in Phonetics* 61: 26-39.

LISKER, L & A S ABRAMASON 1964 A Cross-language Study of Voicing in Initial Stops: Acoustic Measurement. *Word* 20: 384-422.

LOCKHART, J H S 1882 Canton Syllabary. *China Review* X: 312-326.

LÜ, SHUXIANG 1980 *Xiandai Hanyu Babai Ci (Eight Hundred Modern Chinese Words)*. Commercial Press, Hong Hong, 3rd ed., 1990.

LUKE, KENGKWONG 1983 Xianggang Yueyuli [u:] [y:] de Yiwei Wenti (On the Phonemic Status of [u:] [y:] in Hong Kong Cantonese). *Yuwen Zazhi* 11: 37-43.

LUCAS, A 1985 *Chinese Linguistics: A Bibliography: 1975-1982*. Paris: Langages Croises.

LUKSANEEYANAWIN, S 1983 *Intonation in Thai*. PH.D Thesis, University of Edinburgh.
1998 Intonation in Thai. In *Intonation Systems:* 376-394. Ed. by Hirst D and A Di Cristo. Cambridge University Press.

MAEDA, SHINJI 1976 A Characterisation of American English Intonation. Ph.D Thesis, MIT.

NESPOR, M & I VOGEL 1982 Prosodic Domain of External Sandhi Rules. In ven der Hulst H & N Smith (eds). *The Structure of Phonological Representations*: Part 1: 225-255, Dordrecht: Foris.

OHALA, JOHN 1978 Production of Tone. In Fromkin: *Tone: A Linguistic Survey*: 5-39. Academic Press.

PALMER, H E 1922 *English Intonation with Systematic Exercises*. Cambridge: Heffer.

PARKER, E H 1880a Canton Syllabary. *China Review* VIII: 363-82.
1880b Characterless Chinese Words. *China Review* IX: 85-88.

PIERREHUMBERT, J 1980 *The Phonology and Phonetics of English Intonation*. Ph.D thesis, MIT.

PIERREHUMBERT, J & M BECKMAN 1988 *Japanese Tone Structure*. The MIT Press.

POTISUK, S, J GANDOUR & M HARPER 1996 Acoustic Correlates of Stress in Thai. *Phonetica* 53: 200-220.
1997 Contextual Variations in Trisyllabic Sequences of Thai Tones. *Phonetica* 54: 22-42.

RAO, BINGCAI & WUJI ZHAO 1982 *Guangzhouhua Fangyan Cidian (Cantonese Dialect Dictionary)*. Hong Kong: Commercial Press.

RAO, BINGCAI 1995 Xiuding Guangzhouhua Pingyin Fang'ande Xinshexiang Tentative Ideas for Revising the Cantonese Romanization Scheme). *Paper Presented at Fifth International Yue Dialects Conference*. Guangzhou: Jinan University.

Robins, R H 1964 *General Linguistics: an Introductory Survey*. London: Longman.

ROSE, PHIL 1982 *An Acoustically Based Phonetic Description of the Syllable in the Zhenhai Dialect.* Ph.D. Theses, University of Cambridge.
1984 The Role of Subglottal Pressure and Vocal Cord Tension in the Production of Tones in a Chinese Dialect. In *New Papers on Chinese Language Use:* 113-168. Ed. by B Hong, Australia National University Press.

SCHACHTER, PAUL 1965 Some Comments on J.M. Stewart's 'The Typology of the Twi Tone System', *Bulletin of the Institute of African Studies* 1:28-42, Legon University, Ghana.

SELKIRK, E 1980 On the Prosodic Structure and its Relation to Syntactic Structure. In T Fretheim (ed) *Nordic Prosody II.* Trondheim: TAPIR.

SHEN, XIOANAN S 1990 Tonal Coarticulation in Mandarin. *Journal of Phonetics* 18: 281-295.
1993 Relative Duration as a Perceptual Cue to Stress in Mandarin. *Language and Speech* 36(4): 415-433.

SILVERMAN, K & PIERREHUMBERT J 1990 The Timing of Prenuclear High Accents in English. In *Papers in Laboratory Phonology 1 - Between the Grammar and Physics of Speech:* 72-106. Ed. by J Kingston & Beckman M, Cambridge University Press.

TAO, HONGYIN 1996 *Units in Mandarin Conversation: and Prosody, Discourse and Grammar.* John Benjamin Publishing Co.

VAISSIÈRE, J 1983 Language-Independent Prosodic Features. In *Prosody: Models and Measurements:* 53-66. Ed. by Cutler A and D B Ladd, Springer-Verlag Berlin Heidelberg.

VANCE, T J 1976 An Experimental Investigation of Tone and Intonation in Cantonese. *Phonetica* 33: 368-92.
1977 Tonal Distinctions in Cantonese. *Phonetica* 34: 93-107.

VANCE, T H & C A WALKER 1976 Tone and Intonation in Cantonese. *Canadian Linguistic Society* 12: 640-655.

WANG, LI 1936-37 *Zhongguo Yinyunxue (Chinese Phonology).* Shanghai: Shangwu.

WEINREICH, URIEL 1954 Stress and Word Structure in Yiddish. In *The Field of Yiddish: Studies in Yiddish Language, Folklore and Literature:* 1-27. Ed. by U Weinreich, New York: Linguistic Circle of New York.

WHITAKER, K P K 1955-56 A Study on the Modified Tones in Spoken Cantonese. *Asia Major* 5: 9-36, 184-207.

WILLIAMS, S W 1856 *Tonic Dictionary of the Chinese Language in the Canton Dialect.* Canton: Chinese Repository.

WONG, K S M 1981 A Quantitative Study of Tone Change in Cantonese: Some Preliminary Findings. *Chicago Linguistic Society* 17: 419-30.
1982 *Tone Change in Cantonese.* Ph.D Thesis, University of Illinois at Urbana-Champaign.

WONG, S-L 1941 *A Chinese Syllabary Pronounced According to the Dialect of Canton.* Hong Kong: Chung Hwa.

XIANDAI HANYU CIDIAN 1990 *Modern Chinese Dictionary.* Beijing: Commercial Press.

XU, YI 1994 Production and Perception of Coarticulated Tones. *Journal of Acoustical Society America*: 2240-2253.

YANG, FUMIAN P 1974 *Chinese Linguistics: A Selected and Classified Bibliography.* Hong Kong: The Chinese University Press.
1981 *Chinese Dialectology: a Selected and Classified Bibliography.* Hong Kong: The Chinese University Press.

YAU, SHUNCHIU 1980 Sentential Connotation in Cantonese. *Fangyan* 1: 35-52.

YIP, MOIRA 1992 Prosodic Morphology in Four Chinese Dialects. *Journal of East Asian Linguistics*: 1.1: 1-32.

YUAN, JIA-HUA ET AL 1960 *Hanyu Fangyin Gaiyao (Survey of Chinese Dialects).* Beijing: Wenzi Gaige.

YUYAN YANJIUSUO 1978 *Zhongguo Yuyanxue Lunwen Suoyin (A Bibliography of Chinese Linguistics Papers):* 1, Beijing: Shangwu.
1983 *Zhongguo Yuyanxui Lunwen Suoyin (A Bibliography of Chinese Linguistics Papers):* 1, Beijing: Shangwu.

ZEE, YUNYANG E 1980 *Phonetic Studies of Chinese Tones.* Ph.D Thesis, UCLA.
1993 A Phonetic and Phonological Analysis of Cantonese Vowels and Diphthongs. *Abstracts, the Fourth International Conference on Yue Dialects:* 98, Hong Kong: Linguistic Society of Hong Kong.

ZHAN, BOHUI & YATSHING CHEUNG 1990 *Zhujiang Sanjiaozhou Fangyan Zongshu (A Survey of Dialects in the Pearl River Delta, A Synthetic Review.* Vol. 3. Hong Kong: New Century Publishing House.

ZHONGHUA XIN ZIDIAN 1976 *China New Dictionary.* Hong Kong: Chung Hwa Books.

APPENDIX I

LIST OF MONOSYLLABIC AND DISYLLABIC UTTERANCES

1. /ji:^{1}/ medical/ cloth/ aunt
2. /ji:^{2}/ chair/ rely on/ lean on
3. /ji:^{3}/ meaning/ Italy/ thought
4. /ji:^{4}/ son/ and/ move
5. /ji:^{5}/ ear/ already/ discuss
6. /ji:^{6}/ two/ righteousness/ easy

7. /ji:^{1} ji:^{1}/ aunt-aunt - Auntie
8. /ɐw^{1} ji:^{1}/ Aw-aunt - Auntie Aw
9. /ji:^{1} jy:n^{2}/ medical-building - hospital
10. /ma:w^{1} ji:^{2}/ cat-chair - cat-chair
11. /ji:^{1} wɐj^{3}/ lean-lean - lean on
12. /wa:j^{1} ji:^{3}/ bad-idea - bad thought
13. /ji:^{1} li:w^{4}/ medical-treatment - medical treatment
14. /wɐj^{1} ji:^{4}/ mighty-appearance - imposing appearance
15. /ji:^{1} leŋ^{5}/ cloth-collar - cloth collar
16. /mɐn^{1} ji:^{5}/ mosquito-ear - mosquito ear
17. /ji:^{1} mi:n^{6}/ E-noodle – a kind of noodle
18. /ma:^{1} ji:^{6}/ twin-two - two pairs

19. /ji:^{2} ji:^{1}/ Ji-aunt - Auntie Ji
20. /jɐw^{2} ji:^{1}/ pomelo-skin - pomelo skin
21. /ji:^{2} ŋɐj^{2}/ chair-short - short chair
22. /ŋɐj^{2} ji:^{2}/ short-chair - short chair
23. /ji:^{2} jɐn^{3}/ chair-mark - chairmark (on floor)
24. /wa:^{2} ji:^{3}/ painting-meaning - meaning of a painting
25. /ji:^{2} la:n^{4}/ lean-railing - lean on railing
26. /wɐj^{2} ji:^{4}/ zig-zag - zigzag
27. /ji:^{2} low^{5}/ rely on-old - old age arrogance
28. /ji:m^{2} ji:^{5}/ cover-ear - cover ear (with fingers)
29. /ji:^{2} la:j^{6}/ rely-rely - rely on
30. /ji:n^{2} ji:^{6}/ perform-righteousness - develop (human relationship)

31. /ji:^{3} ji:^{1}/ Ji-aunt - Auntie Ji
32. /a:^{3} ji:^{1}/ (prefix)-aunt - Aunt
33. /ji:^{3} jeŋ^{2}/ idea-reflect - reflection (thought)
34. /a:^{3} ji:^{2}/ (prefix)-two - servant
35. /ji:^{3} ow^{3}/ Italy-Austria
36. /ji:m^{3} ji:^{3}/ bore-idea - bored-feeling
37. /ji:^{3} mɐn^{4}/ Italy-language - Italian language

38. /jɐw^{3} ji:^{4}/ young-child - little child
39. /ji:^{3} mɐj^{5}/ barley-rise - barley
40. /ji:n^{3} ji:^{5}/ swallow-look - swallow dance
41. /ji:^{3} mej^{6}/ idea-taste - imply
42. /ji:w^{3} ji:^{6}/ need-easy - take things easy

43. /ji:^{4} ma:^{1}/ aunt-mother - maternal aunt
44. /meŋ^{4} ji:^{1}/ famous-medical - famous doctor
45. /ji:^{4} ji:n^{2}/ suit-perform - suitable to perform
46. /lən^{4} ji:^{2}/ wheel-chair - wheel chair
47. /ji:^{4} ji:n^{3}/ suit-feast - suitable to feast
48. /mɐn^{4} ji:^{3}/ essay-idea - context meaning (of written pieces)
49. /ji:^{4} na:m^{4}/ query-difficult - difficulty
50. /jɐn^{4} ji:^{4}/ person-(suffix) - little person
51. /ji:^{4} nəy^{5}/ son-daughter - son and daughter
52. /mow^{4} ji:^{5}/ model-mock - mock
53. /ji:^{4} mɐn^{6}/ query-ask - doubt
54. /lej^{4} ji:^{6}/ separate-different -separate

55. /ji:^{5} loŋ^{1}/ ear-hole – ear-hole
56. /jy:^{5} ji:^{1}/ rain-coat - raincoat
57. /ji:^{5} wu:j^{2}/ discuss-meeting - assembly
58. /low^{5} ji:^{2}/ old-two - number two (of brothers)
59. /ji:^{5} ɔ:n^{3}/ discuss-case - agenda
60. /jɐw^{5} ji:^{3}/ have-idea - to desire
61. /ji:^{5} jy:n^{4}/ discuss-person – member of parliament
62. /nəy^{5} ji:^{4}/ daughter-(suffix) - daughter
63. /ji:^{5} jy:^{5}/ ear-speak - whisper (in ear)
64. /jɐw^{5} ji:^{5}/ have-ear - listen
65. /ji:^{5} nɔ:j^{6}/ from-inside - within
66. /joŋ^{5} ji:^{6}/ brave-righteousness - righteousness

67. /ji:^{6} na:j^{1}/ two-grand - mistress
68. /ji:^{6} ji:^{1}/ easy-medical - easily curable
69. /ji:^{6} wu:^{2}/ two-Wu - two-stringed musical instrument
70. /la:n^{6} ji:^{2}/ broken-chair - broken chair
71. /ji:^{6} ji:m^{3}/ easy-bore - easily bored
72. /wu:j^{6} ji:^{3}/ meet-idea - understand (take the hint)
73. /ji:^{6} ji:n^{4}/ easy-set on fire - inflammable
74. /nɔ:j^{6} ji:^{4}/ inside-move - move in
75. /ji:^{6} joŋ^{5}/ righteousness-brave - upright and brave
76. /wu:j^{6} ji:^{5}/ meeting-discuss - conference
77. /ji:^{6} mej^{6}/ different-smell - strange smell
78. /ma:n^{6} ji:^{6}/ ten thousand-two - twelve thousand

APPENDIX II

LIST OF TRISYLLABIC UTTERANCES

1.	/tɕhɐw^{1} ji:n^{1} joŋ^{1}/	smoke-old-man - old man smoking
2.	/pi:w^{2} ji:n^{2} ji:^{2}/	perform-perform-chair – chair used for performance
3.	/ɕɐw^{3} ji:n^{3} jɐn^{3}/	thin-swallow-print - print of thin swallow
4.	/tɕhœ:ŋ^{4} ji:n^{4} le:ŋ^{4}/	long-prolong-age - prolong the life
5.	/jɐw^{5} ji:n^{5} jɛ:^{5}/	have-pot-thing – a pot of (food)
6.	/ta:j^{6} ji:n^{6} low^{6}/	big-expose-expose - expose to view
7.	/m^{4} ji:n^{2} ŋa:m^{1}/	not-perform-correct - not perform correctly
8.	/ta:j^{6} ji:n^{2} wa:n^{1}/	big-perform-bay - Big Performance Bay (place name)
9.	/pi:w^{2} ji:n^{2} jɐn^{4}/	perform-perform-people - performer
10.	/tɔ:^{1}ji:n^{2} jy:n^{4}/	many-perform-person - many performers
11.	/hi:n^{2} ji:n^{4} wa:n^{1}/	obvious-look-bend - obviously bending
12.	/thi:n^{1} ji:n^{4} wa:n^{1}/	sky-look-bay - natural bay
13.	/fɔ:ŋ^{4} ji:n^{1}mu:n^{4}/	prevent-smoke-door – smoke blocking door
14.	/fɔ:ŋ^{4} ji:n^{1}jɐn^{5}/	prevent-smoke-addict - preventing smoking-addiction

APPENDIX III

LIST OF UTTERANCES CONSISTING OF IDENTICAL TONAL SEQUENCES

1. /tɕy:^{1} fej^{1} tɕhoŋ^{1} thi:n^{1}/
pig-fly-punch-sky
(Pigs fly sharply to the sky.)

2. /hɔ:j^{1} tɕhœ:ŋ^{1} fa:^{1} hɔ:j^{1} hœ:ŋ^{1}/
open-window-flower-open-fragrant
(Once the window is open, you can smell the fragrance from the flowers.)

3. /tɕhen^{1} thi:n^{1} toŋ^{1} foŋ^{1} tɕhɐy^{1} tɕy:^{1} kɔ:ŋ^{1}/
spring-sky-east-wind-blow-pearl-river
(In Spring, the wind blows to the River Pearl from the east.)

4. /ti:m^{2} ka:j^{2} pow^{2} ɕɐw^{2} tɔ:ŋ^{2} hɐj^{2} ɕa:ŋ^{2} kɔ:ŋ^{2} jeŋ^{2} hœ:ŋ^{2} how^{2}/
how-explain-protect-defend-party-at-province-harbour-in-reflect-sound-good
(Why do the Conservative Party have a good influence in the Province and the Harbour?)

5. /tɐŋ^{2} tɕɐw^{2} fɔ:^{2} keŋ^{2} tɐŋ^{2} tɕɔ:^{2} how^{2} kɐw^{2}/
wait-run-fire-alarm-wait-ed-very-long
(Have been waiting for the fire rehearsal for a long time.)

6. /ka:w^{2} ka:w^{2} ha:^{2} how^{2} how^{2} thɐj^{2}/
stir-stir-a bit-very-good-look
(It is fun to make trouble.)

7. /hɐj^{2} kɐw^{2} lɐw^{2} jɐm^{2} tɕɐw^{2} kɔ:ŋ^{2} kɐj^{2} how^{2} ɕeŋ^{2}/
at-nine-floor-drink-wine-talk-chat-very-smart
(It is very smart to have a chat and a drink on the ninth floor.)

8. /fɔ:ŋ^{3} ka:^{3} fa:j^{3} tɕhɵj^{3} ki:t^{3} tɕœ:ŋ^{3} ɕoŋ^{3} fɔ:^{3} hɵy^{3} fa:t^{3} kwɔ:k^{3}/
relief-leave-quick-fun-knot-account-deliver-goods-go-France-nation
(Quickly go to settle the account and deliver the goods to France.)

9. /ɕej^{3} kɔ:^{3} kwa:j^{3} ha:k^{3} ha:k^{3} ha:k^{3} hej^{3} hej^{3} kɐw^{3} ɕa:j^{3} kwa:j^{3}/
four-m.-strange-guest-guest-guest-air-air-enough-all-strange
(The four strange guests standing on ceremony are strange enough.)

10. /kɔ:^{3} pow^{3} kow^{3} ji:^{3} ki:n^{3} tɕeŋ^{3} khɔ:k^{3}/
m.-report-report-idea-see-right-right
(The opinion in the report is correct.)

11. /tɕhi:n^{4} khej^{4} ni:n^{4} ni:n^{4} tha:n^{4} khɐm^{4} na:n^{4} m^{4} na:n^{4}/
previous-period-hit-piano-difficult-not-difficult
(Was it difficult to play piano at the previous stage?)

12. /wɔ:ŋ^{4} hɔ:^{4} tɕhœ:ŋ^{4} lɐw^{4} ŋɐw^{4} jœ:ŋ^{4} ɕeŋ^{4} khwɐn^{4}/
yellow-river-long-flow-cow-sheep-accumulate-group
(The Yellow River flows forever and there are lots of cattle and sheep.)

13. /jɐn^{4} m^{4} ɕeŋ^{4} jɐn^{4}/
human-not-turn into-human
(Human not appearing human.)

14. /mej^{5} fu:^{5} nɵj^{5} ma:n^{5} ma:n^{5} mi:n^{5} khœ:ŋ^{5} tɕhɔ^{5}: ha:^{5}/
beautiful-woman-lady-evening-evening-reluctant-force-sit-a bit
(The beautiful woman sits reluctantly for a while every evening.)

15. /low^{5} jɐw^{5} ma:j^{5} jɛ:^{5} weŋ^{5} jy:n^{5} jɐw^{5} lɐj^{5}/
old-friend-buy-thing-ever-long-have-polite
(The old friend who comes to shop is always polite.)

16. /ŋɔ:^{5} mow^{5} lej^{5} kɵy^{5}/
I-not-pay attention-him
(I ignore him.)

17. /hɔ:k^{6} ha:w^{6} ti:n^{6} ɕi:^{6} wu:t^{6} toŋ^{6} ta:j^{6} lek^{6} jy:^{6} pej^{6}/
study-school-electric-watch-active-move-big-strength-prepare-prepare
(Prepare well for the television activity in school.)

18. /ta:n^{6} hɐj^{6} ha:^{6} pi:n^{6} mej^{6} tɕɵn^{6} li:k^{6} pa:n^{6} ɕi:^{6}/
but-is-below-side-not-to the best-strength-work-thing
(But the people at the lower level do not do their best in the work.)

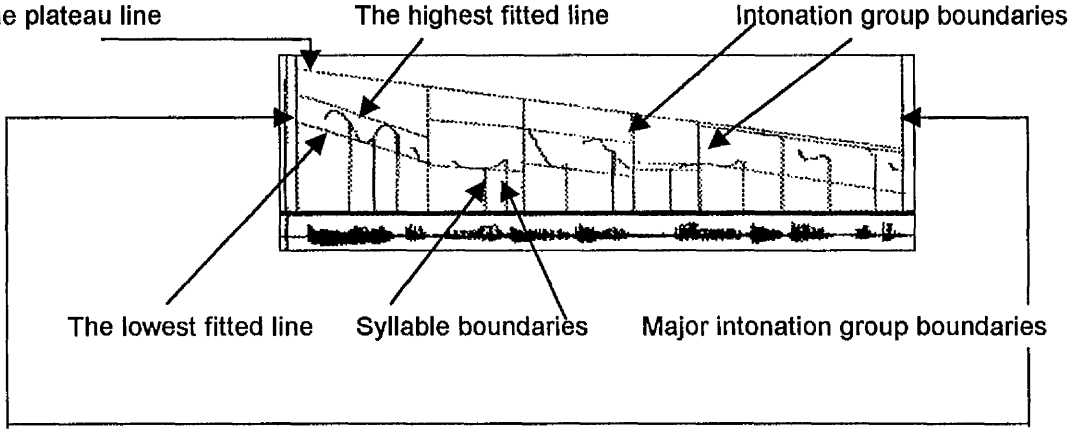
19. /tɐj^{6} ji:^{6} jɐt^{6} tɕow^{6} wa:j^{6} ɕi:^{6}/
the-second-day-do-bad-thing
(Committing an offence on the next day.)

Appendix IV

THE ACOUSTIC RECORDS (DIAGRAMS OF F0 CONTOURS EXTRACTED FROM THE SPEECH WORKSTATION) USED IN THE ANALSIS OF THE PHONOLOGICAL DIVISIONS OF THE UTTERANCES – CHAPTER FOUR

Note:

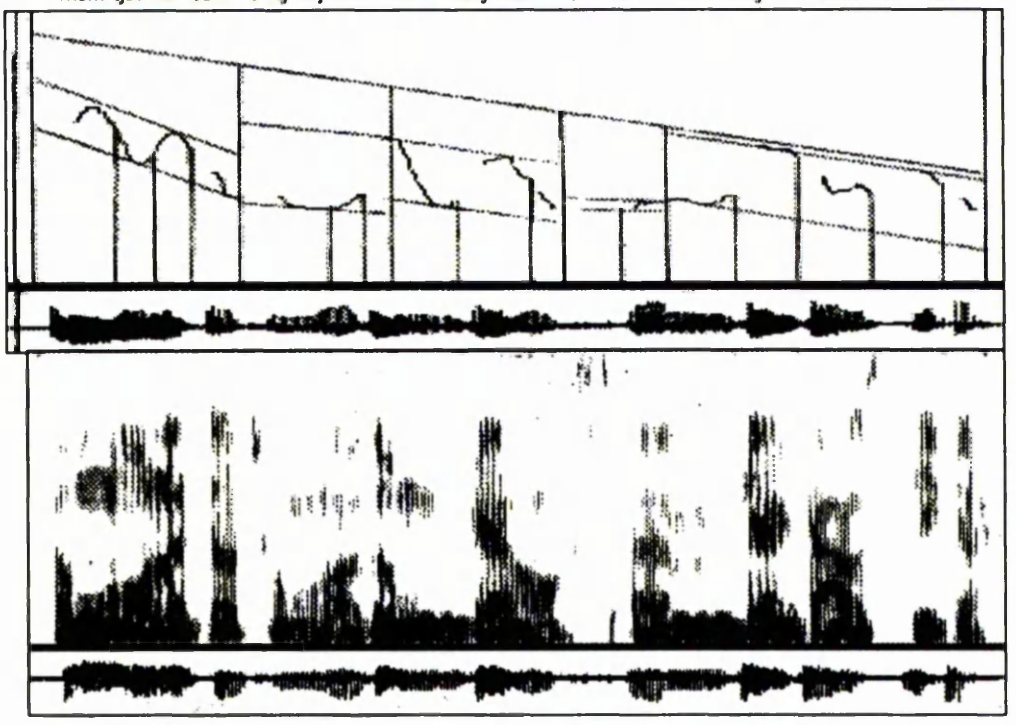
In the diagram of F_0 contours, the numbers in square bracket above the diagram mark the correspondences with the major intonation groups and the numbers immediately after mark the correspondences with the intonation groups. The added vertical lines under the F_0 contours mark the syllable boundaries. The added vertical lines linked with the top and bottom of the diagram mark the boundaries of the major intonation groups; the added sliding line between the major intonation groups represents the plateau line. The added vertical lines linked with the plateau line and the bottom of the diagram mark the boundaries of the intonation groups; the added sliding lines between the boundaries of the intonation groups represent the fitted lines (i.e., highest and lowest) of the intonation group.



[1]

1 2 3 4 5

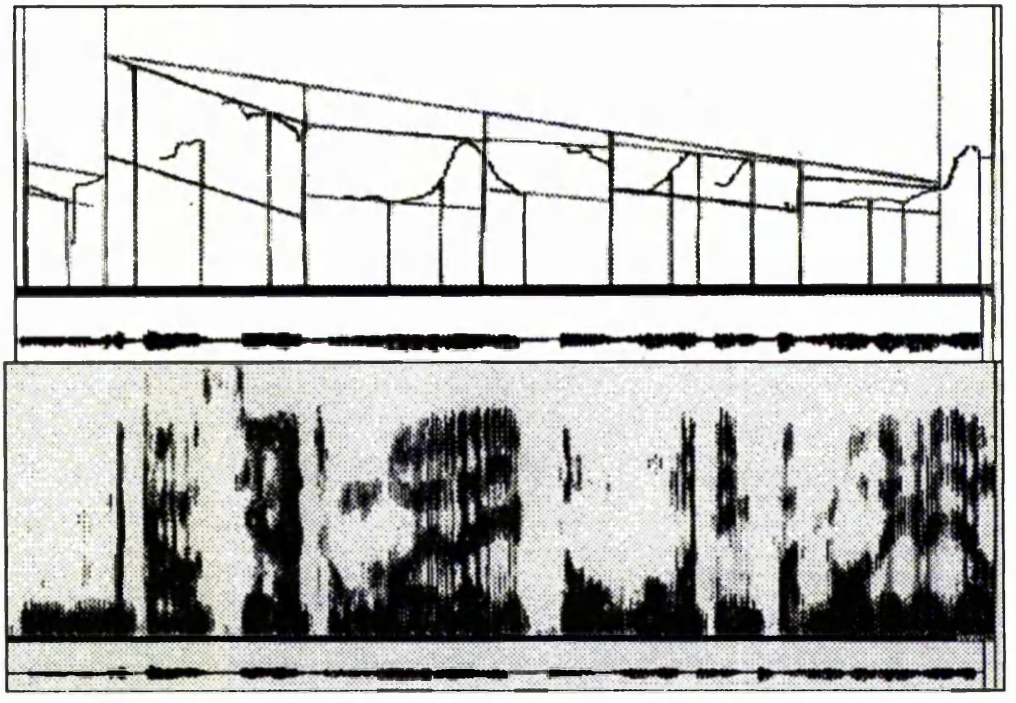
/kɐm^{2}ŋɔ:^{5}nɛ:^{1}tɕɐw^{6} tɕ^{h}oŋ^{4}lɔ:j^{4}tow^{1} m^{4} kɛ:ŋ^{1} hɔ:k^{6} ɕ^{6} ɛ:^{6} m^{4} kɛ:ŋ^{1} ka:w^{3} ɕi:^{1} kɛ:^{3}/



[2] [3]

6 7 8 9 10 11

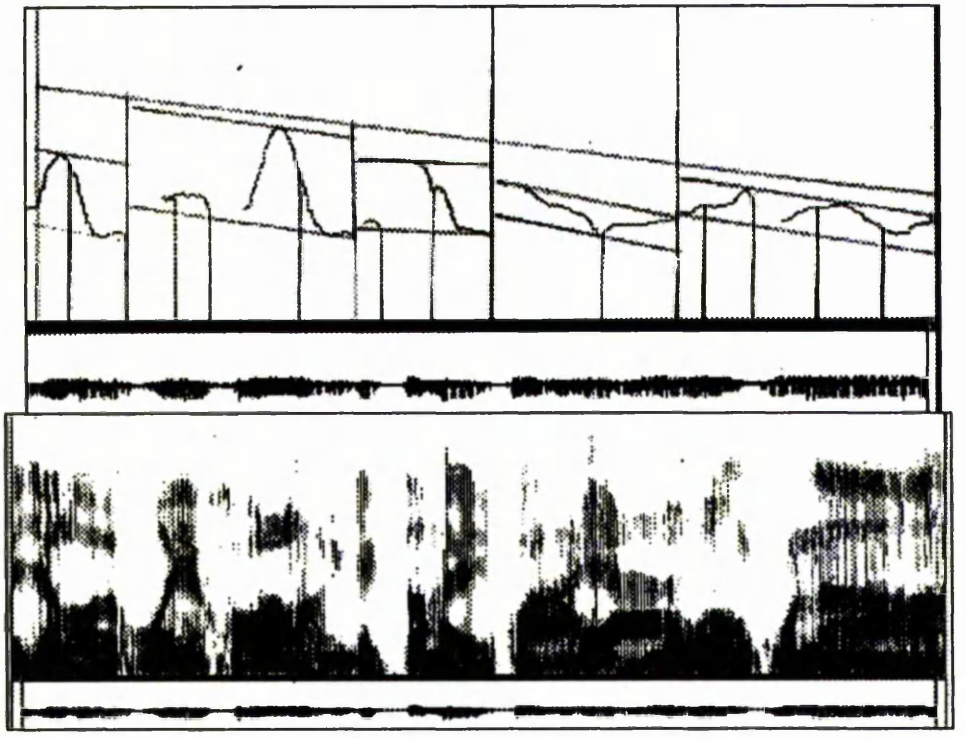
/m^{4} m^{2} kɐm^{2} ka:w^{3} ɕi:^{1}nɛ:^{1} tɕ^{h}oŋ^{4} lɔ:j^{4} nɛ:^{1} jek^{6} tow^{1} mow^{5} mɐt^{1} ti:m^{2}fɐt^{6} ŋɔ:^{5} tej^{6}kɛ:^{3}/



[4] [5] [6]

12 13 14 15 16

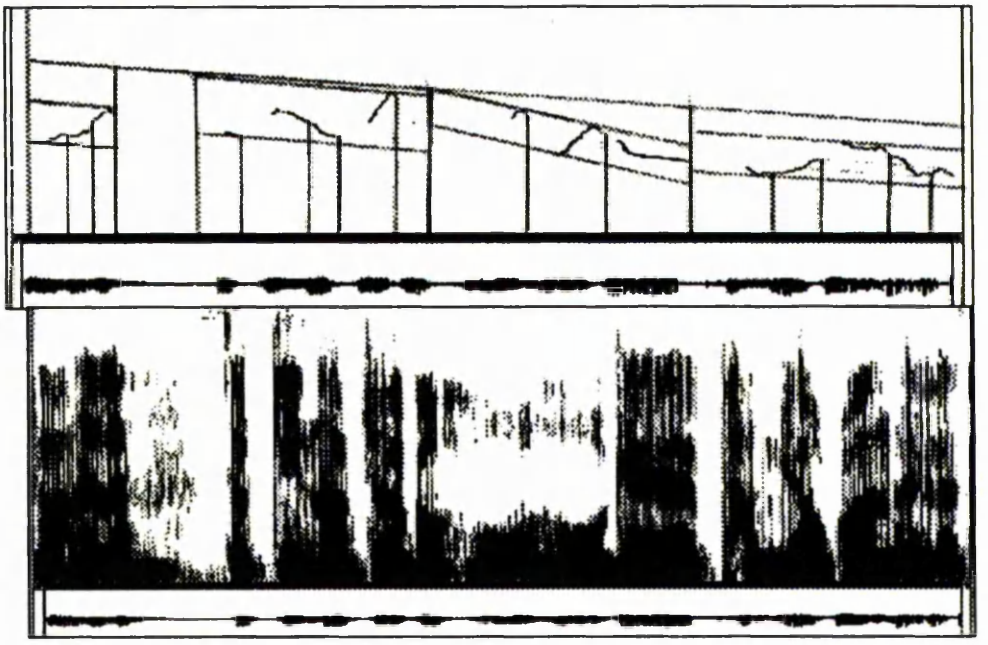
/jɐt^{1} lɔ:j^{4} ŋɔ:^{5} jek^{6} hɔ:^{2} nɐŋ^{4} jek^{6} tow^{1} hɐj^{6} k^{w}a:j^{1} ha:^{2} mowˆwɔ:ŋ^{2} k^{w}u:n^{2} la:^{1}ha:^{2}/



[8]

18 19 20 21

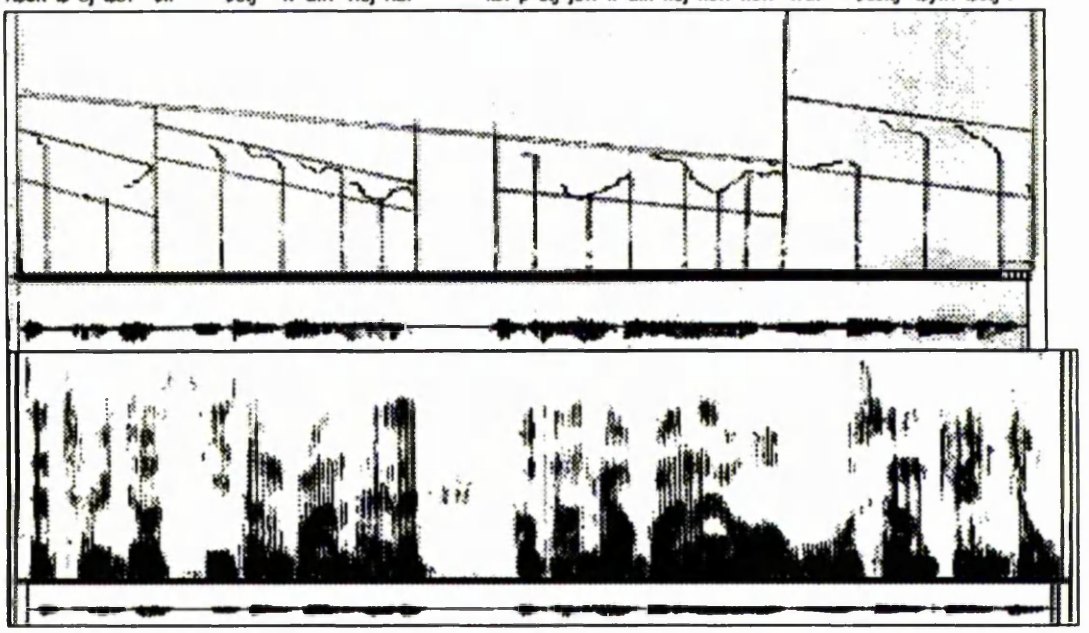
/ta:n^{6}hɐj^{6}nɛ:^{1} tœw^{6}ki:n^{3}la:p^{6}tɕɔ:^{2}kɔ:^{3} how^{2} how^{2} kɛ:^{3} p^{h}ɐŋ^{4} jɐw^{5} k^{w}a:n^{1}hɐj^{6}la:k^{3}/



[9] [10]

22 23 24 25

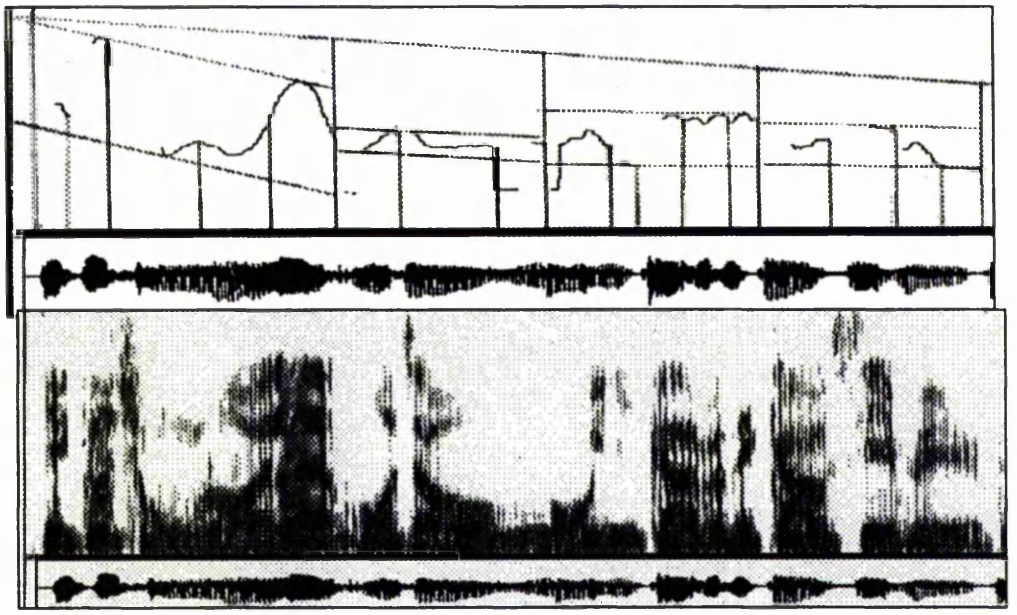
/tɕek^{1}tɕ^{h}ej^{4}tɕɔ:^{2} ɕi:^{1} ɕɐŋ^{1} k^{w}a:n^{1} hɐj^{6}nɛ:^{1} kɔ:^{3}p^{h}ɐŋ^{4}jew^{5}k^{w}a:n^{1}hɐj^{6}how^{2}how^{2}wu:^{6} ɕœ:ŋ^{1} tɕy:n^{1}tɕoŋ^{6}/



[11]

26 27 28 29

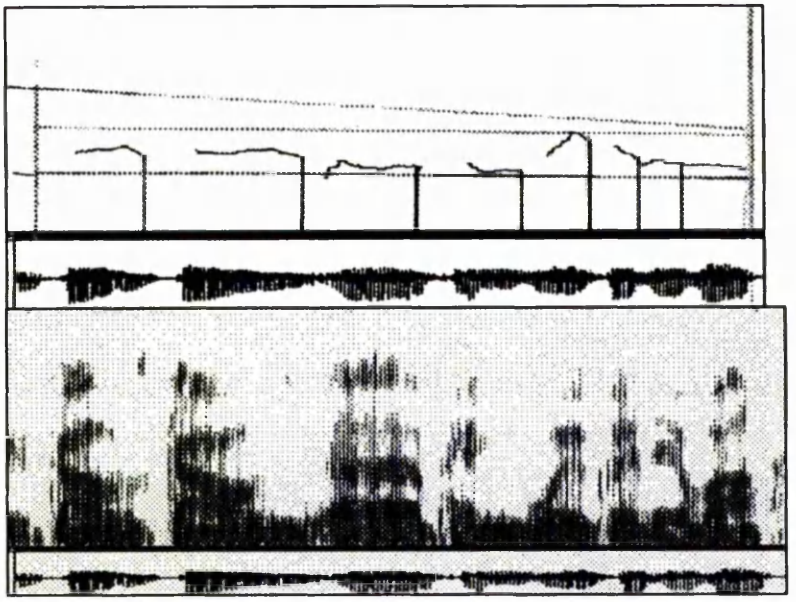
/jek^{6}tow^{1}tɕhoŋ^{4} lɔ:j^{4} nɛ:^{1} ŋɔ:^{5} tɕɐw^{6} ŋ^{6} mow^{5} ji:^{4} ka:^{1} kɔ:^{2} ti:^{1} ka:w^{3} ɕi:^{1}pej^{3}jy:^{4}/



[11] continued

30

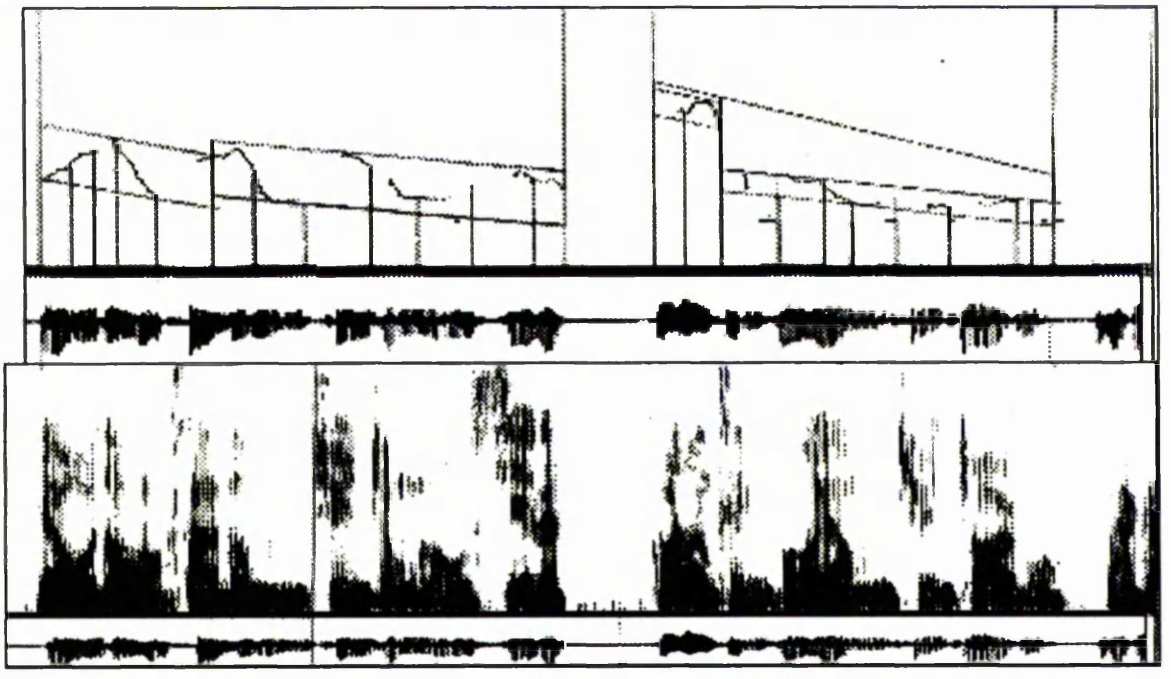
/ka:w^{3} ka:w^{3} ɕ:^{6} p^{h}eŋ^{4} wu:j^{2} k^{h}ɵy^{5} tej^{6} wa:^{6}/



[12] [13]

31 32 33 34

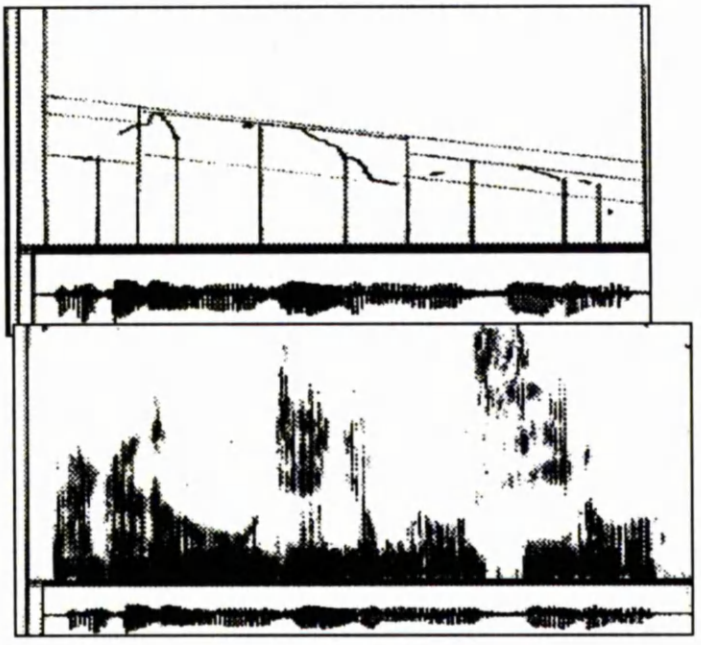
/jɐw^{5} how^{2} tɔ:^{1} hɔ:k^{ɐ}ɑ:ŋ^{1} tow^{1} m^{4} tɕy:n^{1} tɕoŋ^{6} low^{5} ɕi:^{1} nɛ:^{1} kɔ:^{2} tɕoŋ^{2} tɕ^{h}eŋ^{4} fɔ:ŋ^{3} jɐw^{6} m^{4} tɕ^{h}yn^{4} tɕ^{h}ɔ:j^{6} kɔ:^{3} wɔ:^{3}/



[14]

35 36 37

/ta:j^{6} ka:^{1} tow^{1} how^{2} ɪɕy:n^{1} ɪɕoŋ^{6} low^{5} ɕi:^{1} kɔ:^{3} wɔ:^{3}/



[17]

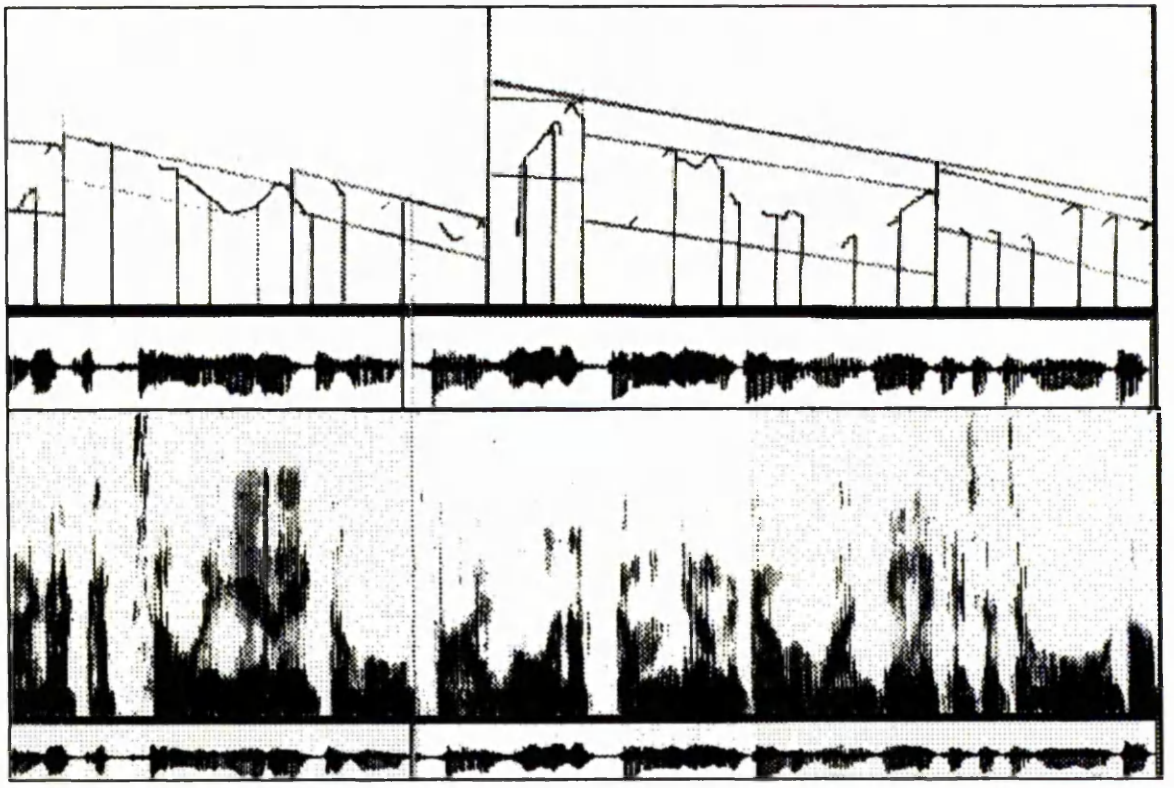
42 43 44

/ta:j^{6}ka:^{1}k^{h}ɐp^{1}ɕɐw^{1}kɛ:^{3}je:^{5}nɛ:^{1}jek^{6}tow^{1}how^{2} fa:j^{3}

[18]

45 46 47

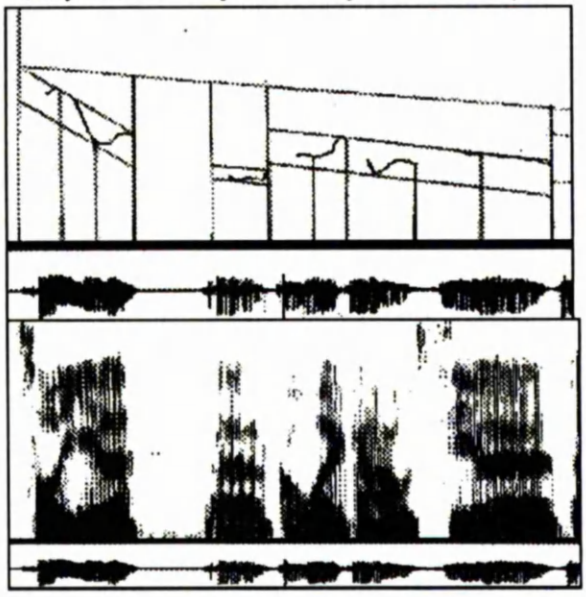
ŋɔ:^{5}kɔ:k^{3}bak^{1} kɐm^{2} jɑɛ:ŋ^{2}kɐ:^{3}kɐ:w^{3}hɔ:k^{6} wɐ:n^{4} kɐŋ^{2}nɛ:^{1} k^{h}ej^{4}ɕɐt^{6}ɪɕɐw^{6}how^{2}how^{2}tɔ:^{1}/



[19]

48 49 50

/cɔ:^{2} ji:^{5} nɛ:^{1} ŋɛ:^{6} kɔ:k^{3}wɐj^{2} ka:w^{3} ci:^{1} nɛ:^{1}/



[19]continued

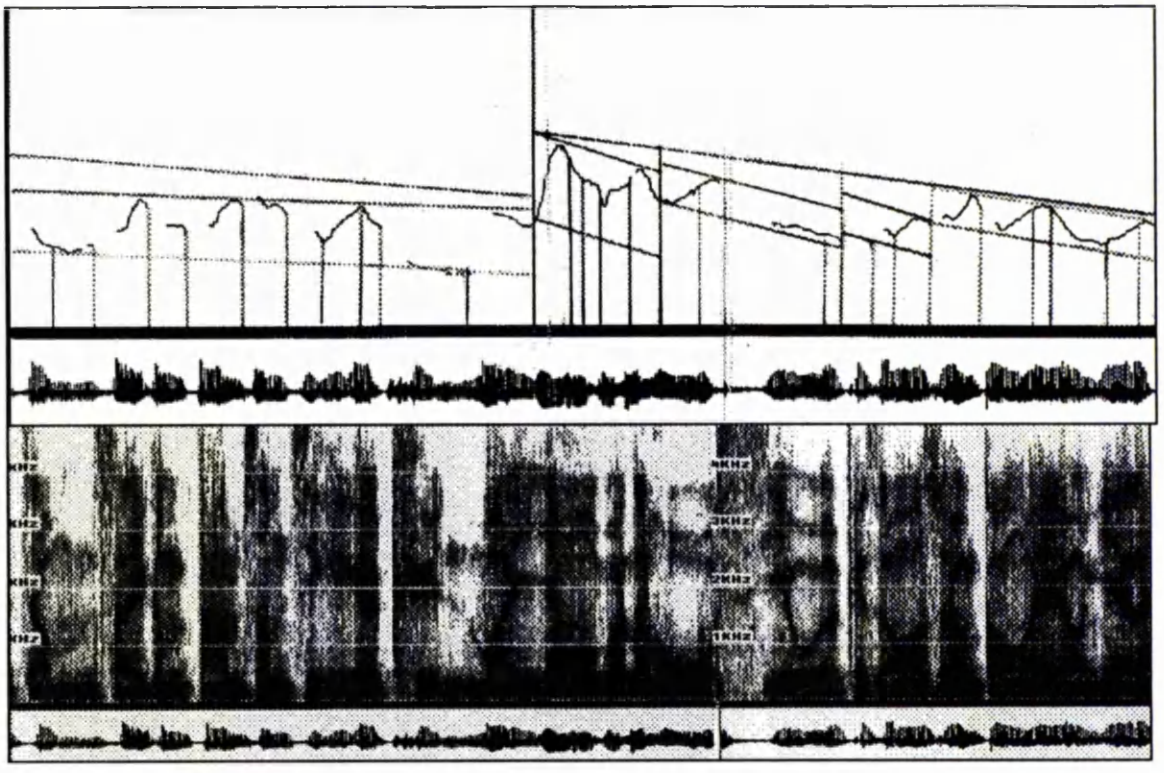
51

/tcɐw^{6}m^{4} cɐj^{2} kɐm^{3}ta:m^{1}cɐm^{1}tci:^{6}kej^{2}kɛ:^{3} kʰy:n^{4} wɐj^{1}

[20]

52 53 54 55

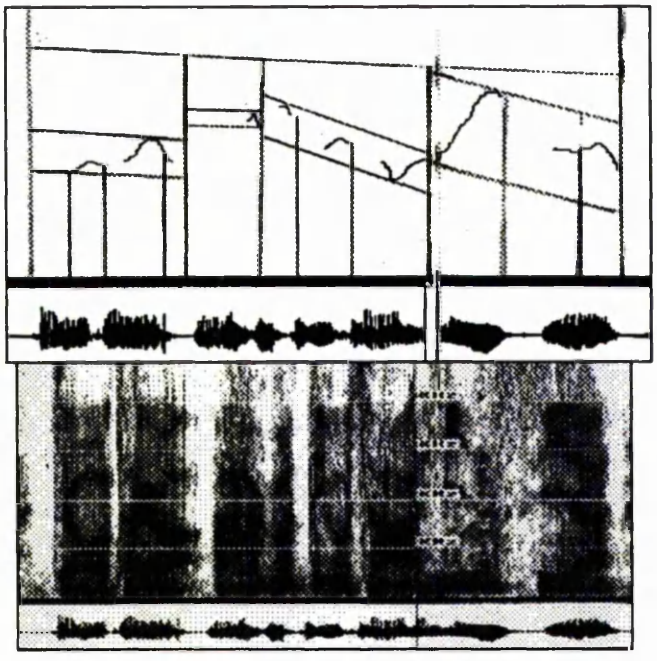
jɐn\wɐ[illegible]how^{2}how^{2} ci:^{3} kɛ:^{3} kʰej^{4}cɐt^{6}ŋɔ:^{5}cœ:ŋ^{2} kɔ:ŋ^{2} jɐt^{1} jœ:ŋ^{6} jɛ:^{5}nɛ:^{1}/



[21]

56 57 58 59

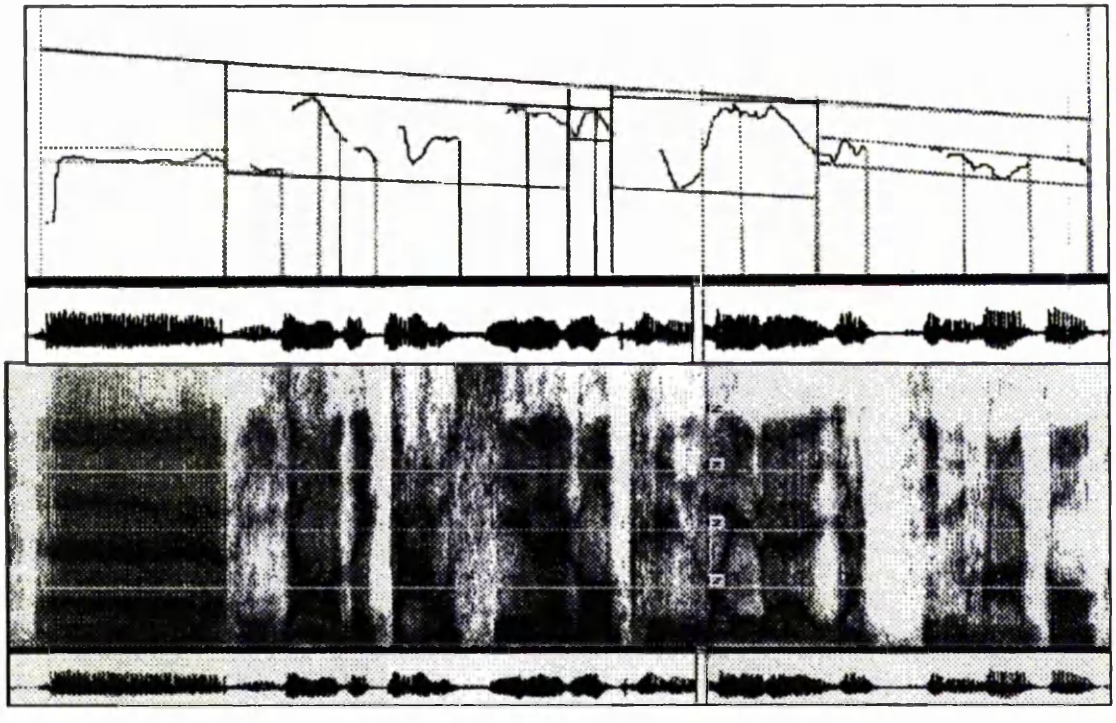
/tɕæw^{6}ŋɔ:^{5}thej^{2}kɔ:^{3} a^{3} kɔ:^{3} ti:w^{6} ʈʂha:^{4} hi:n^{2} ʂi:^{6} a^{3}/



[22]

60 61 62 63 64

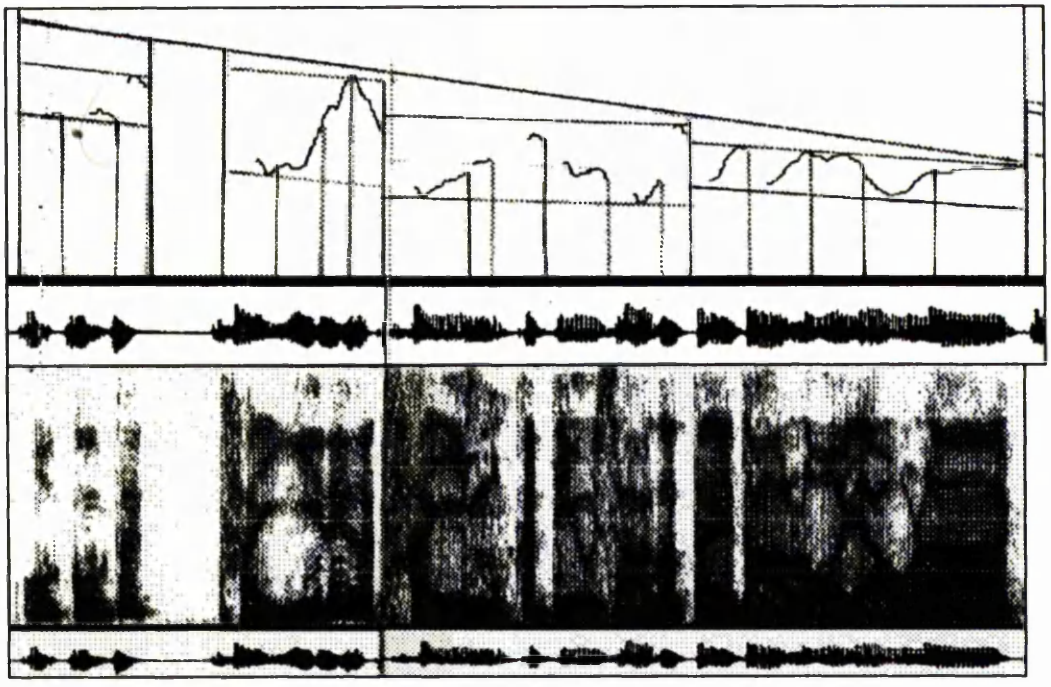
/ɛ:^{6} ji:^{4} ka:^{1}ŋɔ:^{5}tej^{6} ka:w^{3} ʂi:^{1} nɛ:^{1}kɔ:^{2}kɔ:^{3} k^{h}y:n^{4} wej^{1} nɛ:^{1} jɐt^{6} tɕi:m^{6} ha:^{6} kɔ:ŋ^{3}/



[23]

65 66 67 68

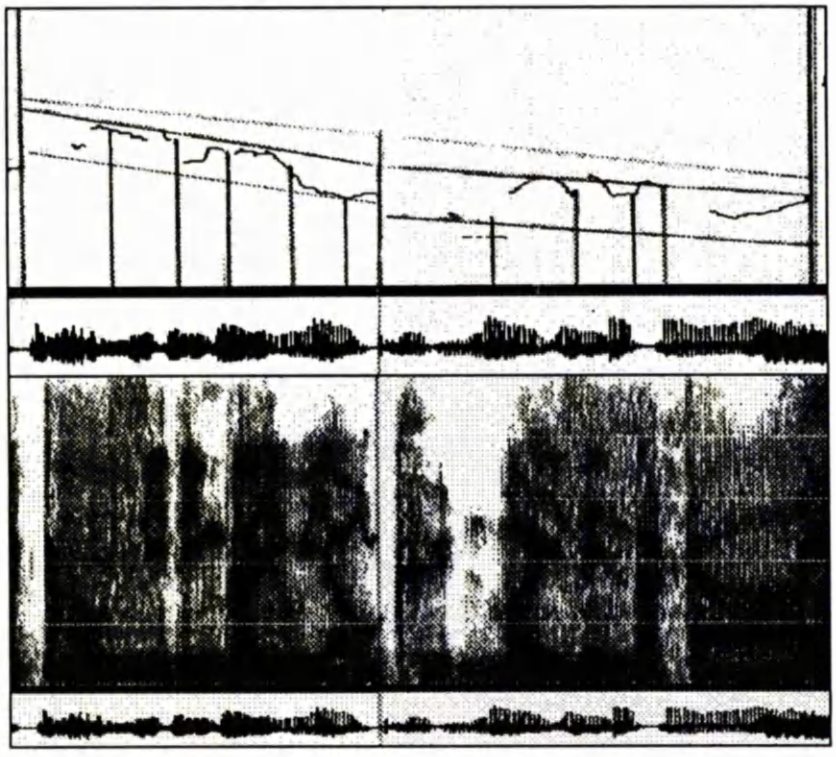
/ŋɔ:^{5}kɔ:k^{3}tɐk^{1} tʰɐj^{4} ji:^{5} ni:^{1}kɔ:^{3} tʰɐj^{4} mokˢ kɛ:^{3} ti:w^{6} tɕʰa:^{4}tow^{1} tɐj^{2} ta:^{2} lɔ:^{1} jɐw^{2} la:^{1}/



[24]

69 70

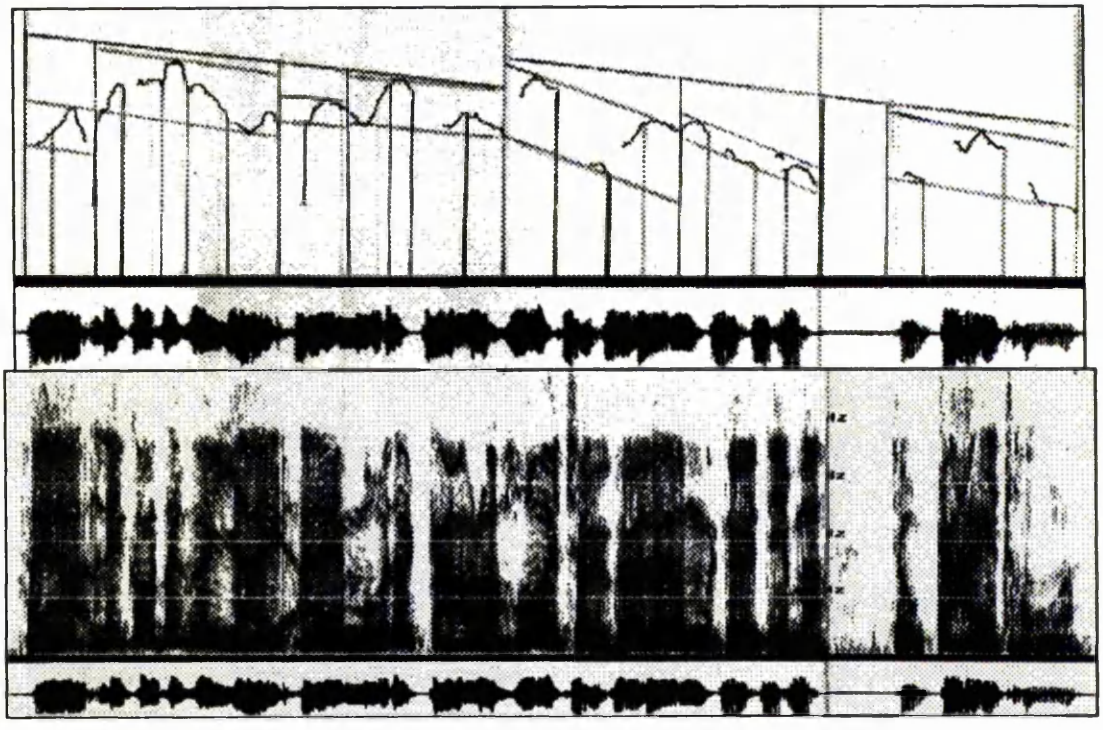
/ka:w^{3} cl:^{1} ti:m^{2} ka:j^{2} ji:w^{3}jɐw^{5} kʰy:n^{4} wɐj^{1} cl:n^{1} tɐk^{1} ka:^{2}/



[35] [36]

86 87 88 89 90 91

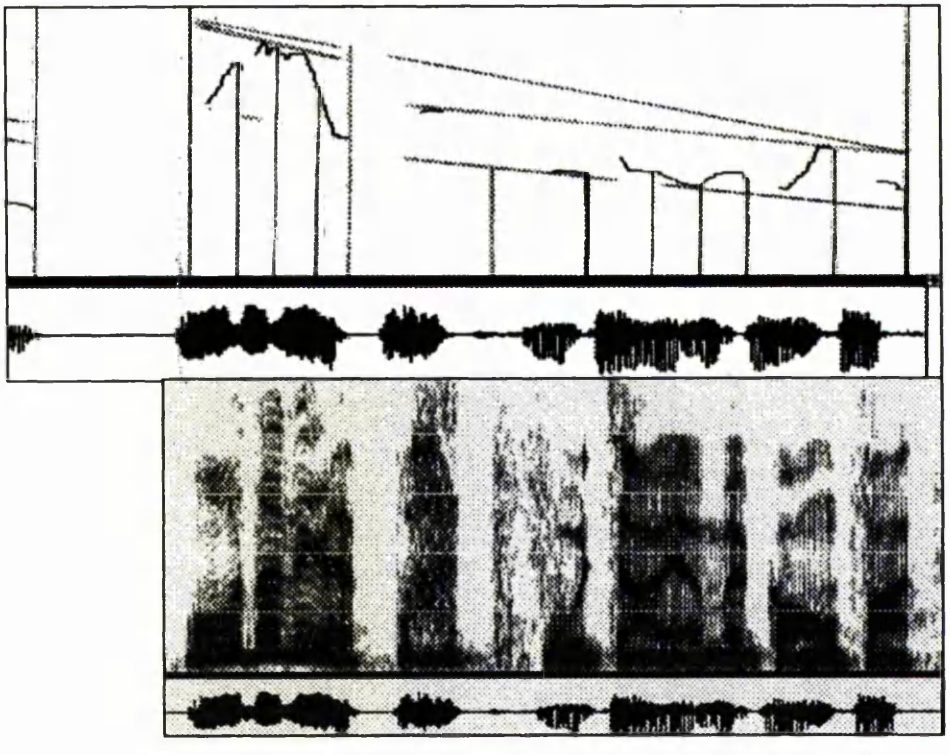
/hɐj^{6}a:k^{3} nɐj^{5}kɔ:k^{3}tɐk^{1}kʰɵy^{6} jɐj^{5} ŋɐj^{3} mow^{5}tɐk^{1}ka:w^{3} lɔ:^{3} hɔ:j^{1} tɕʰɵy^{4}kʰɵy^{5}la:^{1} kɐm^{2}kʰɵy^{5}li:m^{2}a:^{3} tɕow^{6} ka:j^{1} tʰoŋ^{4}a:^{4}/



[37]

93 94

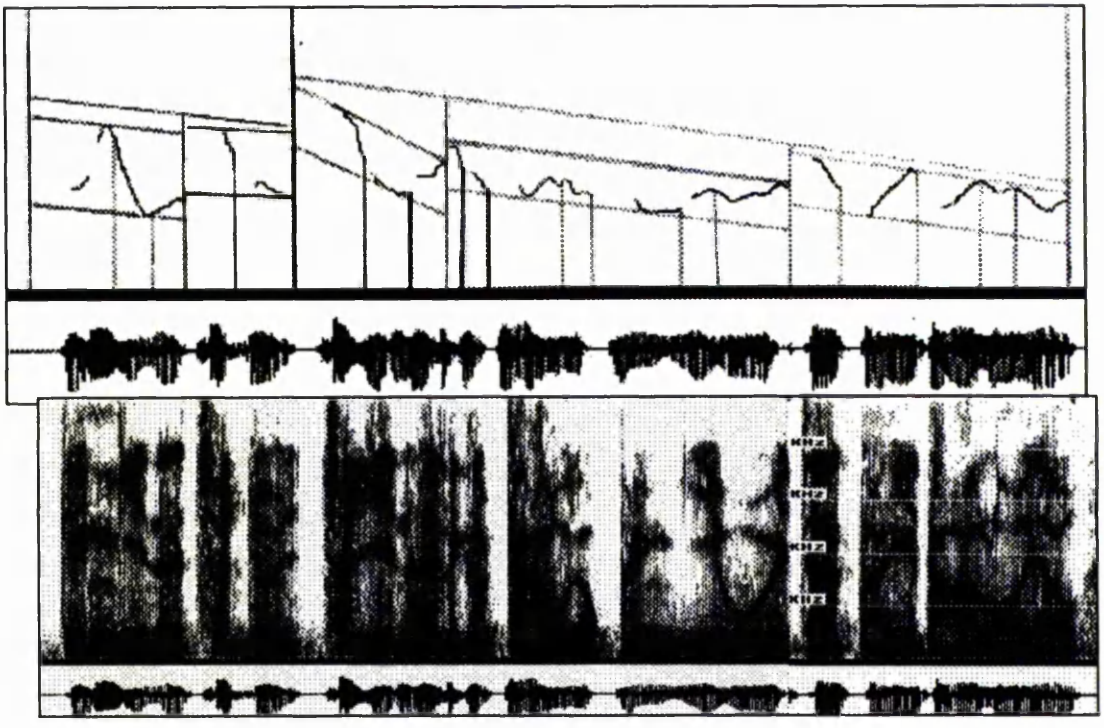
/ŋɔ:^{5} kɔ:k^{3} tɐk^{1} a:^{3} pɐt^{1} fu:^{6} tɕa:k^{3} jɐm^{6} kɛ:^{3} kɔ:ŋ^{2} fa:t^{3}/



[38] [39]

95 96 97 98 99

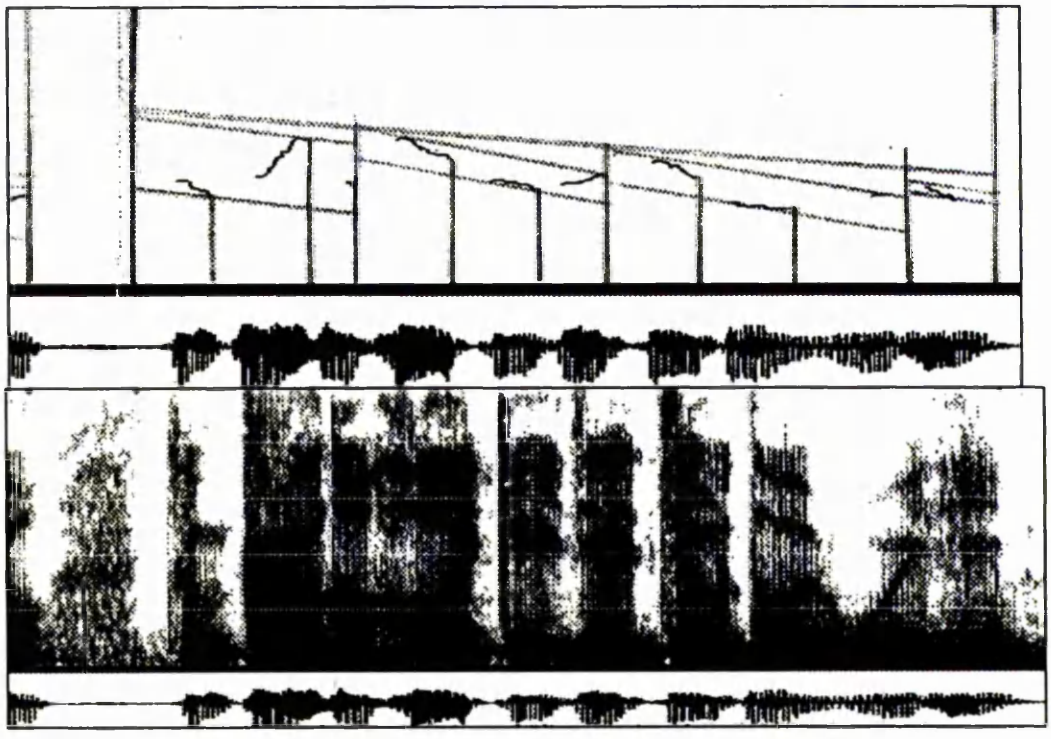
/keŋ^{2} ji:n^{4} hæj^{6} tchet^{1} tɕɔ:j^{6} tchet^{1} ji:n^{6} hæj^{2} ni:^{1} kɔ:^{3} ka:w^{3} jok^{6} peŋ^{4} ji:^{5} wu:j^{2} tɕɛn^{1} tɛj^{2} ta:^{2} lɔ:^{1} jew^{2}/



[40]

100 101 102 103

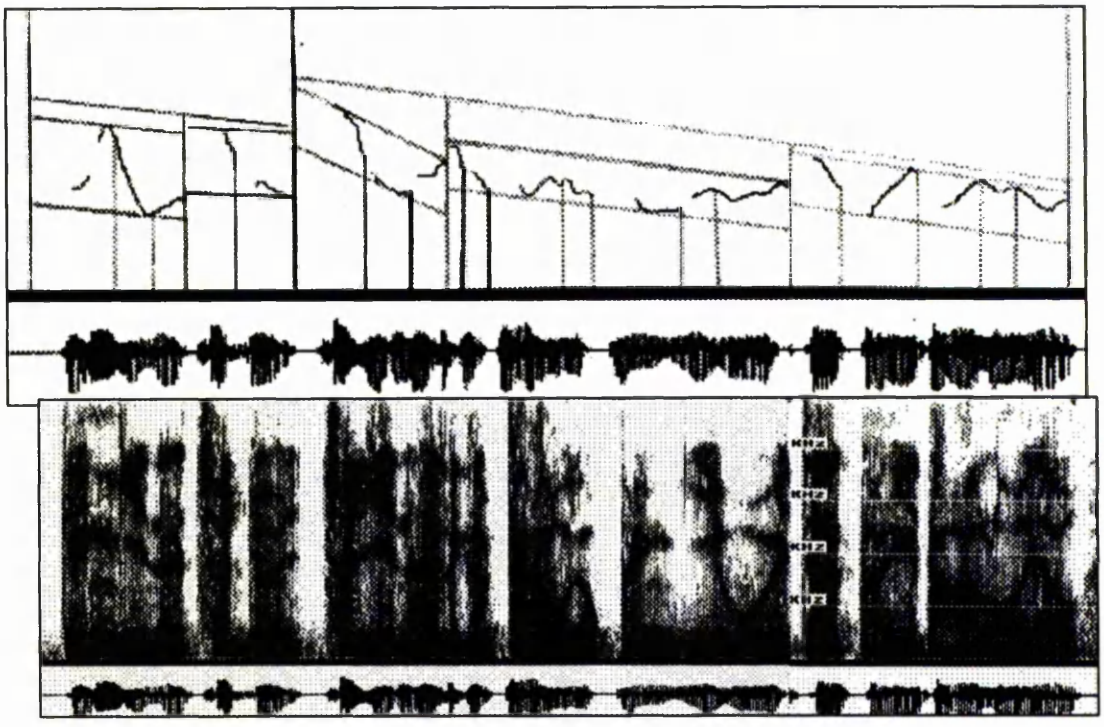
/tɕoŋ^{6} ka:w^{2} tow^{3} hɔ:j^{1} kej^{3} tɕɛ:^{2} tɕi:w^{1} tɔ:j^{6} wu:j^{2} t^{h}i:m^{1}/



[38] [39]

95 96 97 98 99

/keŋ^{2} ji:n^{4}hɐj^{6}tchet^{1}tɕɔ:j^{6} tchet^{1} ji:n^{6}hɐj^{2} ni:^{1}kɔ:^{3} ka:w^{3}jok^{6} peŋ^{4} ji:^{5} wu:j^{2} tɕɐn^{1} tɐj^{2} ta:^{2} lɔ:^{1}jɐw^{2}/



[40]

100 101 102 103

/tɕoŋ^{6} ka:w^{2} tow^{3} hɔ:j^{1} kej^{3} tɕɛ:^{2} tɕi:w^{1} tɔ:j^{6} wu:j^{2} t^{h}i:m^{1}/