

Tone samples of Cantonese and English in Hong Kong

September 2013

Table of Contents

1. Introduction
2. Design of data collection
3. Reading the Excel files

1. Introduction

This database was established in an attempt to study the tonal properties of Hong Kong Cantonese and Hong Kong English speech. Towards this end, elicitation is made from representative local speakers of these two languages using stimuli that would adequately cover all the known tone types.

Since F0 is normally assumed to be the main acoustic property that relates to tones, it is this measurement that is being reflected in this study. However, all our recordings are provided here in case the user would like to look into other acoustic properties.

This database is distributed without compensation to anyone who asks for it. If you find it useful for your research or study, please cite it as:

WEE, Lian-Hee, Queenie K.Y. Chan, Winnie H.Y. Cheung and Suki S.Y. Yiu (2013) *Tone samples of Cantonese and English in Hong Kong*. HKBU Phonology Laboratory, Hong Kong Baptist University [Recordings].

We would be much encouraged to hear from you if you use this work, as well as any queries or comments. These may be sent to:

Caretaker, HKBU Phonology Lab

HKBUPhonologyLab@gmail.com or lianhee@hkbu.edu.hk

We are grateful to the Hong Kong Baptist University Faculty Research Grant FRG2/11-12/135 for funding and to the participants who must remain anonymous.

2. Design of data collection

2.1 Participants

Six Hong Kong locals are identified, balanced for gender. Subjects fall within the age range of 20-35 years, and minimally have tertiary education. This is the group most likely to use both English and Cantonese.

Code	Gender	Age
F1	female	21
F2	female	25
F3	female	28
M1	male	22
M2	male	24
M3	male	32

To ensure representativeness, these subjects first provide a recording of their speech (approx. 3 minutes). These are presented to a random selection of other Hong Kongers to judge if the samples are typical of Hong Kong speech. These recordings contain information that may reveal the identities of the participants, and hence are not included in this database for distribution.

2.2 Stimuli

There are two sets of words: Cantonese and English. The Cantonese word list contains 18 target characters and 18 fillers. The 18 characters are comprised of three CV triplets, each triplet corresponding to one of the six tonal contours in the Cantonese inventory. Each item in the triplet is varied for vowels [i], [a], [u] to cover the cardinal points of tongue body positions so that normalization may be made for impact of vowel quality on F0 values. The characters and their *Jyutping* romanizations are given below:

	[a] type	[i] type	[u] type
Tone 1	巴 baa1	資 zi1	姑 gu1
Tone 2	寡 gwaa2	指 zi2	古 gu2
Tone 3	霸 baa3	志 zi3	故 gu3
Tone 4	爬 paa4	磁 ci4	扶(扶持) fu4
Tone 5	棒 pang5	市 ci5	婦 fu5
Tone 6	罷(罷工) baa6	字 zi6	負 fu6

The English wordlist contains 20 target words and eight disyllabic fillers, all words familiar to the typical HKE speaker. These 20 are divided into two sets: monosyllabic words and polysyllabic words. The words are chosen such that the five known tones of HKE (H, L, F, M, Mf) are fully represented in all the positions where they may occur: word-initially, word-medially and word-finally. To the extent possible, each tone is also varied for the three cardinal vowels. For example L, there are no [u] ending words that carry that tone, hence impossible to get any readings for the F0 properties that relate to that vowel. Such cases do not threaten the value of this experiment.

[i]-type	[a]-type	[u]-type
tree	bar	shoe
mean	arm	moon

* all monosyllabic words have F tone in HKE.

tone				
H-Mf-L	criminal	laughably	purify	popular
M-H-L	elicit	unfasten	exclusion	erotic
M-M-F	disagree	disregard	introduce	
H-Mf-Mf-L	noticeable	regularly	necessary	

** H and M are never word-final; L and F always word-final; Mf always penult or antepenult.

The HKE stimuli as constructed above would adequately cover the distribution of HKE tones in terms of where each tone may occur in a word balanced for three main tongue body positions corresponding to [i, a, u].

2.3 Data Elicitation

The elicitation procedure takes into account the following factors:

- (i) the position of a target in an utterance
Is the target in the initial, medial or final position of an utterance?
- (ii) interference from syntax
Is the target in the middle or boundary of a syntactic phrase?
- (iii) articulation type
Is the articulation made as if reading from a list or a sentence?
- (iv) priming effects
Is there any information that would prime the subject to be inclined to a certain mode of utterance?

To deal with these factors, subjects are to repeat each target as a triplet in a single utterance, thus providing recordings of the target utterance-initial, medial and final positions. This triplet approach factors out all syntactic influence since the triplet will not be placed within any syntactic construction. To avoid a list reading, a prompt will be used so that the subjects will have to make their articulations within the context of a single utterance. Use of a prompt might interfere with the subjects' articulatory habits. For example, an English prompt might nudge the subject towards a certain accent of English. To ensure neutrality, an English prompt will be used to elicit Cantonese utterances and vice versa.

Recording for each item will be elicited twice. All items will be mixed with fillers to randomize them. (Click [here](#) for details on elicitation).

2.4 Technical details of recordings and measurements

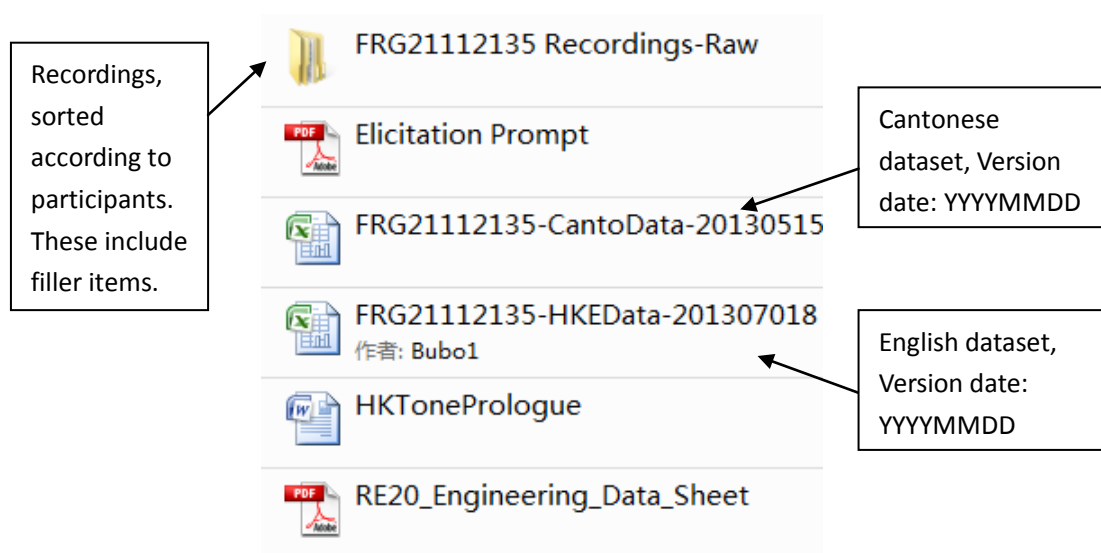
Recordings are made at the HKBU phonetics lab using Praat (ver. 5.3.39, Boersma and Weenink 2013) with a sampling frequency of 44100Hz. This is done with a microphone Electro-voice RE20 (details [here](#)). F0 contours are extracted and

time-normalized using Xu’s (2012) script by identifying the rime of each syllable at 10% intervals. Measurement is not taken at the 0% (start)point to allow for articulatory transitions between onset and rime. The starting point is marked at the onset of voicing of the immediately following vowel. For Cantonese, since all stimuli begin with a plosive and end with a sonorant, the end point is marked at either the closure of the following syllable’s plosive onset or in utterance-final cases, at the end of the utterance. For English cases, starting points are marked at the onset of voicing of the immediately following vowel if the onset is an obstruent; and if the onset is a sonorant, at the point where the wave-form shows a sharp rise in amplitude. Endpoints are marked at the point of closure if immediately followed by plosives; at the point where the waveform amplitude diminishes sharply if immediately followed by non-plosive consonants. In cases where there might be an ambisyllabic consonant, the length of the consonant to used to determine if there was germination. Geminate cases are dealt with by marking the syllable boundary at the midpoint of the consonant. These are all hand-checked in consultation with both the waveforms and wide-band spectrograms. Though grey areas and possible margins of errors are impossible to ignore, we believe that for the purposes of obtaining F0 contours corresponding to the various tone types, the errors are within tolerable limits.

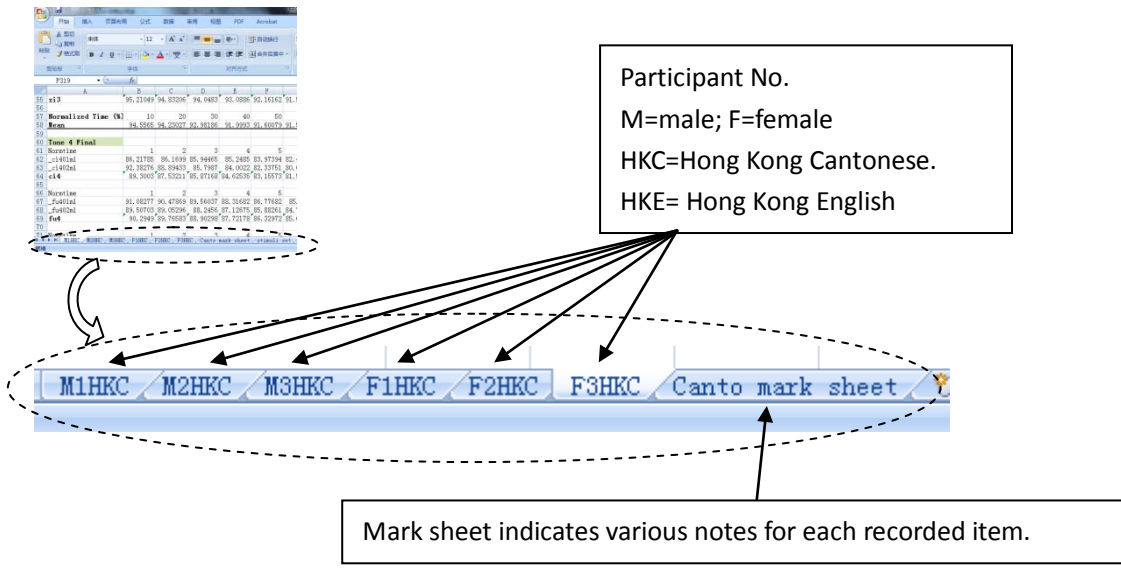
3. Reading the Excel files

3.1 General information

The folder contains the following items:



The two Excel spreadsheet files contain our measurements from the recordings. At the bottom of the spreadsheet, the measurements taken from each participant is listed from left-to-right, shown below.



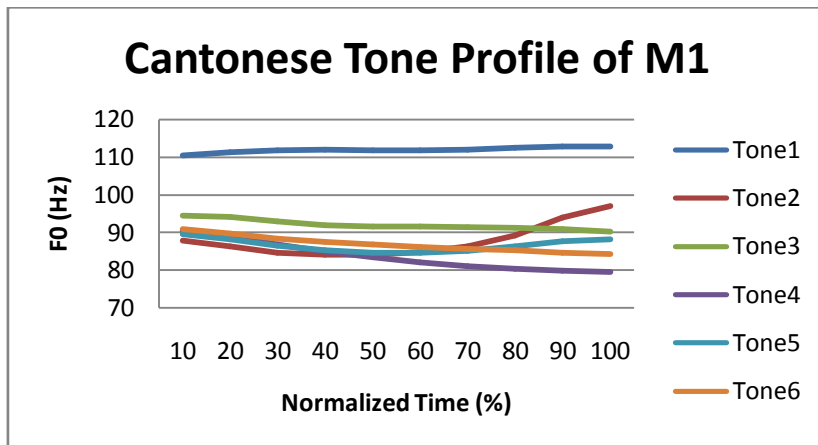
3.2 Reading the spreadsheet: Cantonese

The spreadsheet containing our measurements of the recordings for Cantonese tones are organized as illustrated below.

	A	B	C	D	E	F	G	H	I	J	K
M1: Cantonese Tone Profile											
Tone 1 Final											
Normtime		1	2	3	4	5	6	7	8	9	10
_baa101ml		106.421	106.3512	106.4633	106.92	107.3391	107.6071	107.8206	108.0422	107.8282	107.58
_baa102ml		107.8687	107.9834	108.5616	109.4302	109.4379	109.0818	109.0153	109.0726	109.2129	109.3262
baa1		107.1448	107.1673	107.5124	108.1751	108.3885	108.3445	108.418	108.5574	108.5205	108.4531
Normtime		1	2	3	4	5	6	7	8	9	10
_gul01ml		117.5773	116.8579	116.3785	116.1588	115.4878	114.5801	114.2534	114.7623	115.6456	115.9916
_gul02ml		106.0399	110.7003	111.9507	111.2289	110.7603	110.6602	110.8905	111.4011	112.219	112.0573
gul		111.8086	113.7793	114.1646	113.6938	113.124	112.6201	112.572	113.0817	113.9323	114.0244
Normtime		1	2	3	4	5	6	7	8	9	10
_zil101ml		111.7264	112.2176	112.746	112.9109	112.8049	112.7939	113.2036	114.1495	115.12	115.4212
_zil102ml		113.2305	114.0252	114.7425	115.1501	115.58	116.1791	116.7624	117.2077	117.2043	117.1081
zil		112.4785	113.1214	113.7443	114.0305	114.1924	114.4865	114.983	115.6786	116.1622	116.2647
Normalized Time (%)		10	20	30	40	50	60	70	80	90	100
Mean		110.4773	111.356	111.8071	111.9665	111.9016	111.817	111.991	112.4392	112.8717	112.9141

From these the F0 profiles may be generated. Below is an example using the average F0 values for each tone type articulated as utterance-final syllables by M1. The user may consider generating other graphs depending on what is being queried. S.D.

values can also be calculated for each interval point to determine the degree of overlap between contours of different tone types.



In analyzing all the recordings, it is inevitable that there will be unclear cases or technical problems. These are indicated in the Mark Sheet, illustrated below.

Participant **Utterance position:**
initial, medial, final

	A	B	C	D	E	F	G	H	I	J
		M1			M2		M3			F
		In	Me	Fi	In	Me	Fi	In	Me	Fi
1	baa101	x	x	x	x	x	x	x	x	x
2	baa102	x	x	x	x	x	x	x	x	x
3	gwaa201	x	x	x	x	x	x	x	x	x
4	gwaa202	x	x	x	x	x	x	x	x	x
5	gu301	x	x	x	MU	MU	MU	x	x	x
6	gu302	x	x	x	x	x	x	MU	MU	MU
7	paa401	x	x	x	x	x	x	x	x	x
8	paa402	x	x	x	x	x	x	x	x	x
9										
10	Remarks									
11	x									
12	x									
13										
14	MU									
15	FD									

Recorded item. Final number "1" indicates first recording. "2" the second recording.

X = interval point selected and measured

Colors and codes indicate notes

Legend for colors and codes may be read here.

- x** used item
- x** The syllable has normal-creaky phon
- MU** discarded item
- MU** Mistaken Utterance
- FD** The informant mis F0 distortion

3.3 Reading the spreadsheet: Hong Kong English

Like the Excel spreadsheets for the Cantonese data, the measurements for Hong Kong English are also laid out as a set of spreadsheets. Because English words may be polysyllabic, there are some additional indications. These are illustrated below.

Words with number of syllables

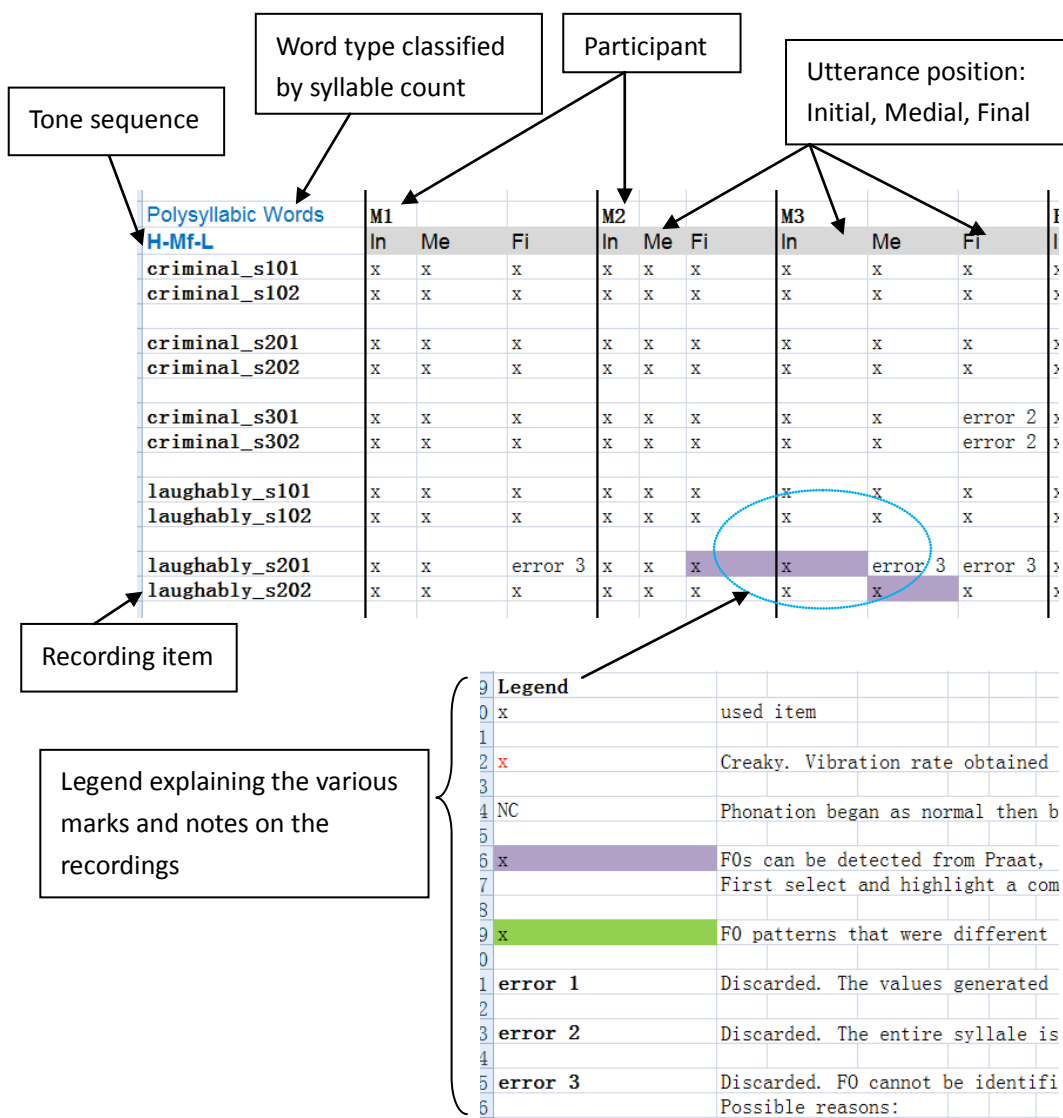
Utterance position: initial, medial or final

Syllable measured

Trisyllabic-Final			
H-Mf-L (Final)			
1st syllable-Final			
Normtime	1	2	3
_criminal_s101ml	113.0639	112.7351	112.2391
_criminal_s102ml	124.3936	122.9118	121.1275
criminal	118.7287	117.8234	116.6833
Normtime	1	2	3
_laughably_s101ml	101.7011	101.5936	101.5848
_laughably_s102ml	104.1823	104.2573	104.4075
laughably	102.9417	102.9254	102.9961
Normtime	1	2	3
_popular_s101ml	115.133	114.3922	113.5917
_popular_s102ml	115.8402	115	114.1904
popular	115.4866	114.6961	113.891
Normtime	1	2	3
_purify_s101ml	110.7686	110.5516	110.3019
_purify_s102ml	113.9973	113.1708	112.
purify	112.3829	111.8612	111.3
Normalized Time (%)	10	20	
Mean	111.3553	110.7193	109.6
3rd syllable-Final			
Normtime	1	2	3
_criminal_s301ml	87.38492	86.50539	85.43273
_criminal_s302ml	83.75009	82.41747	81.0195
criminal	85.56751	84.46143	83.22612
Normtime	1	2	3
_laughably_s301ml	64.18364	60.87407	56.65496
_laughably_s302ml	error 2		
Normtime	1	2	3

Some cases contained creakiness or errors, which are indicated. Details are provided in the marksheet.

For some recordings, additional remarks are also indicated at the extreme right of the spreadsheet corresponding to each participant. The various indications or creakiness and other possible technical errors are also explained in the marksheet.



References

Boersma, Paul and David Weenink (2013) Praat: doing phonetics by computer [Computer program]. Version 5.3.39, retrieved Jan 2013 from <http://www.praat.org/>

Xu, Yi (2012) ProsodyPro. Version 4.3, retrieved Aug 2012 from <http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/>