廣州話元音的底層形式

The Underlying Representation of Cantonese Vowels

Dominic Yu
Undergraduate Honors Thesis
Department of Linguistics
Stanford University
2000 May 19

ABSTRACT

In a traditional approach to phonology, segments in a language are grouped into phonemes, where allophones of the same phoneme, though phonetically distinct, have the same underlying form and "sound the same" to a native speaker of that language. However, for Cantonese, there are no vowel alternations that indicate which allophones should belong to which phonemes. Furthermore, relying on distributional data yields multiple, conflicting phonemic analyses. A survey was given to native Cantonese speakers asking for their judgements on vowel similarity. Results suggest that the underlying forms of the vowels are not particularly abstract, i.e., they are relatively similar to the surface forms.

A M D G

ACKNOWLEDGEMENTS

I have had the fortune of working on this honors thesis over the last year. It has been quite an experience. First and foremost, I would like to thank Edward Flemming for being such a good advisor through thick and thin. I would also like to thank Elizabeth Traugott and Will Leben, whose advising and support throughout this endeavor has been of immense value; Chao Fen Sun, always a source of assistance and inspiration; and all my professors and instructors who have collectively convinced me that linguistics is the place for me.

I would like to thank all the people who have accompanied me on this journey, in rough chronological order: Sam Hu, since way back when (four years, who'd have thunk?); Juno Nakamura, since way back at the beginning of the year; all the Honors College people, including the symbolic systems people we hung out with; the linguistics staff; and all my fellow linguistics undergraduate and graduate friends and acquaintances. Thank you to Stacey Woo, Vivian Gee, Klara Elteto, and Casey Anne Sirl for their presence, helping me in times of need and keeping me sane. And thank you to Okada House, past, present, and future.

I would also like to thank my subjects—AF, AC, JP, JC, CW, JW, JY, EW, EY, HW, TK, VH, CY, AY, CC, AL, WC, VG, SN, and OK—who graciously gave of their time. Thanks also to the Human Subjects Protocol, which protects the anonymity of said subjects.

Most importantly, I need to thank my family—especially my mother—for their patience and support. I cannot express my gratitude in words.

And . . . last but not least, I would like to thank the whole crew of Ling 206. You guys definitely put the [fʌn] in [fʌnɑlədʒi].

A NOTE ON ROMANIZATION

LSHK romanization for Cantonese (粵拼/jyut6 ping3) is used in this paper. A brief layout of this system is presented below.

Onsets

b [p]	p [p^h]	m [m]	f [f]
d [t̪]	t [$t̪^h$]	n [n̪]	l [l̪]
g [k]	k [k^h]	ng [ŋ]	h [h]
gw [k^w]	kw [k^{hw}]	w [w]	
z [ts]	c [ts^h]	s [s]	j [j]

Nuclei

aa [aː]	i [iː, ɪ]	u [uː, ʊ]	e [ɛː, e]	o [ɔː, o]
yu [yː]	oe [œː]			
a [ɐ]	eo [ɵ]			

Codas

p [p˺]	t [t̪˺]	k [k˺]
m [m]	n [n̪]	ng [ŋ]
i [i, y]	u [u]	

Syllabic nasals

m [m̩]	ng [ŋ̍]

Tones (placed at the end of the syllable)

1 [˥]	4 [˩˩]
2 [˧˥]	5 [˨˩]
3 [˧]	6 [˨]

Examples

fan1 分	fat1 忽
fan2 粉	
fan3 訓	faat3 發
fan4 墳	
fan5 奮	
fan6 份	fat6 佛

More details are available from the Linguistic Society of Hong Kong. At the time of this writing, the web site is available at <`http://www.hku.hk/linguist/lshk/`>.

Introduction

Classical phonemics tells us that the many and varied sounds of a language are grouped into a small, finite number of phonemes. Under this view, phonemes of a language have the same underlying representation and "sound the same" to native speakers of that language. Sapir states that "what is felt by the speakers to be the 'same' sound has perceptibly different forms as [phonetic] conditions vary" (1925, p. 21). Twadell, though skeptical of the phoneme defined as a psychological reality, summarizes this viewpoint quite succinctly:

> [One] way of formulating the contrast would be this: Although the speakers produce and the hearers experience objectively different sounds, they are not aware of that difference; the speaker intends to produce the same sound, and the hearer has the impression of hearing the same sound. (1935, p. 56)

One question to be asked, then, is whether or not it is true that phonemes of a language will sound "the same" to native speakers of that language.

This study takes a look at this question in the case of the Cantonese vowels. A survey was given to native speakers of Cantonese asking for their intuitions on vowel similarity. The study tested the hypothesis that speakers would respond using phonological criteria rather than mere phonetic similarity to determine vowel equivalence. This hypothesis is complicated by the extent to which phonetic similarity affects the phonological system, especially for Cantonese, which lacks alternations, and whose vowels are in multiple complementation. Results indicate that speakers are using more than phonetic criteria to make vowel similarity judgements; however, they tend to make more distinctions than what one would expect from a phonemic analysis.

Phonemes: Psychological Reality and Other Desiderata

The concept of the phoneme is based on the existence of surface contrasts. In

English, the existence of a contrast between [kʰɑt] and [gɑt], for example, motivates positing two phonemes, /k/ and /g/. However, not all contrasts show up in all environments, e.g., there is no /k, g/ distinction after initial /s/ in English. Thus, we have [skɑt] but not *[skʰɑt] or *[sgɑt]. How, then, does one decide whether this voiceless unaspirated velar stop [k] groups phonemically with [kʰ] or [g]?

When grouping sounds into phonemes, two criteria generally have to be met:

1. complementary distribution
2. phonetic similarity

Taking the classic example from English, [ŋ] and [h] fulfill the first criterion, as [h] appears only at the beginning of a syllable, but [ŋ] appears only at the end. Nevertheless, [ŋ] and [h], failing the second criterion, are not considered to be the same phoneme. If we apply the same criteria to the above example of [k, kʰ] and [k, g], we find that both groupings satisfy both criteria; if anything, the second group is more phonetically similar, since initial /g/ is often realized as voiceless unaspirated [k].

In such cases, where both above conditions are met, some other factors may be appealed to:

1. the phoneme as psychological reality
2. existence of alternations
3. phoneme parsimony

The first is the idea of the phoneme as a psychological unit. Jaeger (1986) conducted a psychological experiment to ascertain the correct psychological grouping of [k] after [s]. The results supported "the analysis that in English the [k] after initial [s] belongs to the /k/ phoneme."

A second factor that influences the grouping of sounds into phonemes is the

existence of alternations. In English, [t^h, ɾ, t¬] are considered allophones of /t/ because of alternations such as “notation” [t^h], “notable” [ɾ], and “note” [t¬].

Often a third desideratum is that of phoneme parsimony—minimizing the number of phonemes. Taking again the example of [k] after [s] in English, this factor is one of the reasons we group [k] with /k/, rather than positing a three-way distinction among /k^h, k, g/.

Complications with Cantonese

Overview of the Cantonese Syllable

Before we apply these considerations to Cantonese, let us first take a brief look at the vowels of Cantonese in the context of the phonological system. In Chinese phonology, syllables are usually divided into two parts: the initial (聲母/shēngmǔ), or onset, and the final (韻母/yùnmǔ), or rime. The structure of the Cantonese syllable is illustrated below (adapted from Li 1985; syllabic [m, ŋ] are omitted):

(1) initials	p, p^h	t, t^h	ts, ts^h	k, k^h	kw, kw^h
	m	n		ŋ	
		l	j		w
	f		s	h	

(2) vowels	iː		yː		uː
		ɪ			ʊ
		e		ɵ	o
		ɛː	œː		ɔː
				ɐ	
				aː	

The finals chart below is arranged such that each row contains one vowel and each column contains one final consonant/offglide.

(3) finals	open	diphthongs		nasal			checked		
	aː	aːi	aːu	aːm	aːn	aːŋ	aːp	aːt	aːk
		ɐi	ɐu	ɐm	ɐn	ɐŋ	ɐp	ɐt	ɐk
	ɛː					ɛːŋ			ɛːk
		ei							
						ɪŋ			ɪk
	iː		iːu	iːm	iːn		iːp	iːt	
	œː					œːŋ			œːk
			ɵy		ɵn			ɵt	
	ɔː	ɔːi			ɔːn	ɔːŋ		ɔːt	ɔːk
			ou						
						ʊŋ			ʊk
	uː	uːi			uːn			uːt	
	yː				yːn			yːt	

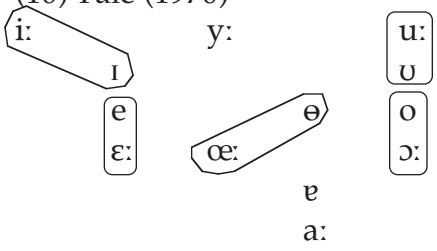
The distribution facts of the vowels can be succinctly described as follows:

(4) [i, u] do not occur before velars.
(5) [ɪ, ʊ] only occur before velars [k, ŋ].
(6) [e, o] only occur in the diphthongs [ei, ou].
(7) [œ] only occurs in open syllables and before velars.
(8) [ɵ] only occurs between coronals, and in the diphthong [ɵy].
(9) [u] cannot occur after coronals.

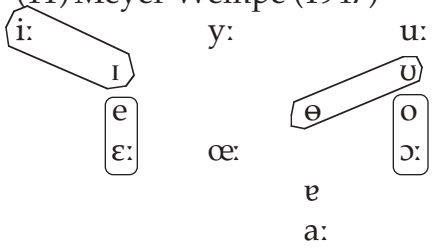
Phonemic Analyses of the Cantonese Vowels

Under a traditional framework, such multiple complementation begs for some kind of phonemic grouping. However, there are several different phonemic analyses, and they are not all compatible with each other. Over the years, several different romanization systems have been developed for Cantonese, three of which are illustrated below:

(10) Yale (1970)



(11) Meyer-Wempe (1947)



(12) Bauer (1997)

iː		yː		uː
ɪ				ʊ
e			ɵ	o
ɛː		œː		ɔː
			ɐ	
			aː	

Chao (1934, p. 46) notes that "given the sounds of a language, there are usually more than one possible way of reducing them to a system of phonemes. . . . As we emphasize this or that motive, we should arrive at a different system of organization of elements into phonemes. . . . these motives are often conflicting." To illustrate, following are several very different phonemic analyses of the Cantonese vowels.

First, there is a striking pattern of complementary distribution which has often been described as a tense-lax distinction (Hashimoto 1972, p. 153, and Yip 1996): [a] vs. [ɐ], [ɛ] vs. [e], [i] vs. [ɪ], [ɔ] vs. [o], [ʊ] vs. [u], and [œ] vs. [ɵ]. Noting that, with the exception of [a, ɐ], all of these pairs are in complementary distribution and are phonetically similar to each other, we can group them as follows, with a total of 8 phonemes:

(13)

iː	yː		uː
ɪ			ʊ
e	ɵ		o
ɛː	œː		ɔː
	ɐ		
	aː		

Here are the corresponding rules (we consider [ɵ] to be tense, following Yip 1995):

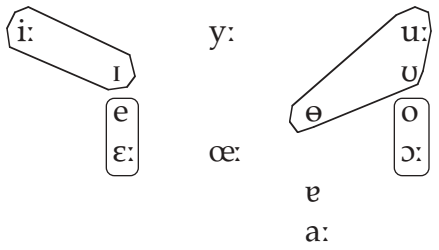
(14) V[+hi, +tense] —> [-tense] / ___ [velar]
(/i, u/ change to [ɪ, ʊ] before /k, ŋ/)

(15) V[-hi, -low, -tense] —> [+tense] / ___ V[+hi]
(/ɛ, œ, ɔ/ change to [e, ɵ, o] in diphthongs)

(16) œ —> ɵ / ___ [coronal]

Of course, there are variations on this analysis that also yield 8 phonemes. Noting the limited environments in which [ɵ] occurs, we can have this:

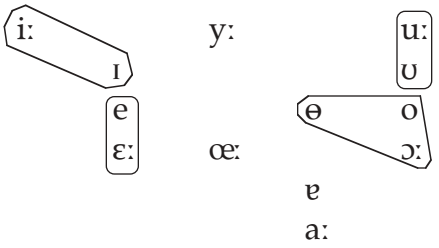
(17)



(18) u —> ʊ / ___ [velar]
(19) u —> ɵ / [coronal] ___ [coronal]

Or this:

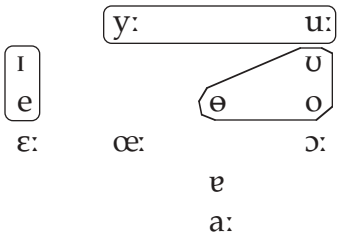
(20)



(21) ɔ —> o / ___ u
—> ɵ / [coronal] ___ [coronal]

A second approach is to utilize the short/long distinction. First, we recognize [ɪ] and [e] as the short counterparts to [ɛː], and [ʊ, o] as the short counterparts to [ɛː] (Bauer 1997, p. 45). Next, taking advantage of the limited distribution of [ɵ], we group it with the short back vowels. Finally, we treat [yː] as an allophone of /uː/, noting that the only place where they contrast is after the plain velars [k, k^h]. By positing /kw/ or /kw^h/ as the underlying onset before /uː/ for these cases, we can treat [uː, yː] as noncontrastive. In this way, the surface distinction between [kyːn] and [kuːn] is due to an underlying distinction between /kuːn/ and /kwuːn/. Now we can come up with the following analysis with 9 vowel phonemes (this is Chao's analysis, Hashimoto, p. 154):

(22)



(23) V[-hi, -low, +tense] —> [+hi, -tense] / ___ [velar]
(/e, o/ change to [ɪ, ʊ] before velars)

(24) o —> ɵ / [coronal] ___ [coronal]

(25) uː —> uː / [labial] ___
—> yː / elsewhere

We have thus far gone through four possible phonemic analyses under a traditional framework, relying only on the distributional facts of the vowels. Many more analyses are possible, including ones with only seven vowel phonemes. But without vowel alternations, and without information regarding the psychological reality of these phonemicizations, there is little basis on which to choose among them. In light of this, we can posit another hypothesis: speakers will group the vowels according to their distribution facts, and will either all agree on one phonemic analysis, or will all have different phonemic analyses which fall into the set of possibilities outlined above.

Implications of a Surface-Oriented Framework

An alternative view of underlying representations is provided by an output-based framework such as Optimality Theory (OT). With such framework, questions about the "correct" phonemic grouping or the "correct" underlying form become moot. Under OT, any reasonable input will yield the correct surface form as the output, given an appropriately ranked set of constraints. What is important to our discussion here is the notion of Lexicon Optimization, developed by Prince and

Smolensky (1993, p. 192). Yip summarizes this as follows:

> . . . for a given phonetic form, the chosen UR will be the one that maps onto the surface form with the fewest violations of high-ranked constraints. Since FAITHFULNESS is a set of constraints, the chosen UR will, *ceteris paribus,* be the one with the fewest FAITHFULNESS violations, the one most similar to the surface form. (1996)

Yip states that unlike rule-based grammars, whose basic architecture encourages abstract underlying forms, a surface-oriented framework favors more concrete forms. Yip argues against the principle of PARSIMONY, which includes economy of phoneme inventory:

> The usual rationale for assuming that PARSIMONY exists is that lexical storage is expensive, and must be minimized. Recent work has thrown doubt on this notion. . . . In cases like [Cantonese], this rationale is certainly unconvincing, because the numbers are so small. . . . I conclude that the null hypothesis is that learners of languages like Cantonese, with few or no alternations, learn something very close to what they hear, and that PARSIMONY is overridden by FAITHFULNESS.

Thus, for Cantonese, we might expect that the underlying forms are quite similar to the surface forms, since there are no alternations that would force the native speaker to consider otherwise. This leads us to posit a third hypothesis: speakers' underlying representations will tend to correspond to surface forms, i.e., speakers will group vowels together which are transcribed with the same IPA symbol.

Methodology

There are two parts to this study. One is the survey on vowel similarity judgements given to native Cantonese speakers; the other involves making recordings and measuring the formants of the vowels.

The survey presents pairs of words[†] to the subject and directly asks them to judge

[†] In this paper, we will use "word" as a shorthand for "character/morpheme-syllable."

whether the vowels in those words are the same or different. By directly eliciting judgements on vowel equivalence, and by forcing subjects to decide on "same" or "different" (with no in-between), this survey tests the above hypotheses. Depending on how subjects respond, the results will support one or another hypothesis on what the underlying representations will be.

The survey contains 77 pairs of characters. In order to ensure that subjects understand the task and do not confuse vowels with, say, rimes, the instructions contain two examples of pairs whose vowels are uncontroversially the same. Let us define vowels to be uncontroversially the same (u-same) if all phonemic analyses agree that the vowels are the same. There are 13 u-same vowels in Cantonese, as listed in chart (2). Let us define vowels to be uncontroversially different (u-different) if no phonemic analysis groups them together as the same phoneme. The list of word pairs is constructed such that half of the pairs are u-same/u-different (the control cases), and the other half contains the vowel pairs whose categorizations are in question (the test cases). The list is randomized. The test cases test the following pairs: [i, ɪ], [ɪ, ei], [ei, ɛ], [u, ʊ], [ʊ, ou], [ou, ɔ], [œ, ɵ], [œ, ɵy], [ɵ, ɵy] [ɵ, ʊ], [ɵ, ou], [ɵ, u], [ou, u]. These are the pairs in complementary distribution that are described in the phonemic analyses above, along with all possible combinations of [ɵ, u, ʊ, ou]. The surveys were given to 20 subjects. Subjects were mostly native Cantonese speakers who grew up in Hong Kong. Four subjects were native Cantonese speakers who grew up in California.

If the hypothesis that speakers will group vowels according to distribution facts is true, we would expect subjects' responses to correspond to some valid phonemic analysis that groups together vowels in complementary distribution. If, on the other

hand, the hypothesis based on Lexicon Optimization is true, we would expect that subjects will tend to judge only u-same vowels as the same.

Following are the directions from the survey (romanization for the characters has been added):

> **Background**
> This is a study of vowel similarity in Cantonese.
> Some words clearly contain the same vowel. For example,
>
> 八 [baat3] and 鴨 [ngaap3]
>
> both contain the vowel in 「呀」[aa3]. Likewise,
>
> 閃 [sim2] and 必 [bit1]
>
> both contain the vowel in 「衣」[ji1].
> Other words, however, are less straightforward. This survey asks for your judgements on vowel similarity.
>
> **Directions**
> For each pair of words below, indicate whether you think their vowels are the same, or different.
>
> 龍 [lung4] 工 [gung1] same different
> 接 [zip3] 央 [joeng1] same different
> 比 [bei2] 車 [ce1] same different
> .
> .
> .

The methodology was revised after preliminary results indicated that subjects were having difficulty with the control items. The first seven subjects were given only the survey, with no explanation or training. Since there were many incorrect answers to the u-same control items, the remaining subjects were trained on some examples, given an explanation, or both. Training consisted of listing about 20 u-same pairs, and then testing the subjects with more word pairs until they felt comfortable with the exercise. Subjects who were given an explanation were told that all syllables can be separated into an initial and a final (i.e., onset and rime); that the final can be further separated into the vowel and the final consonant; and that the vowel was the part that was of

interest in this study. Two subjects (EW, EY) were given training only, four subjects (HW, TK, VH, CY) were given only an explanation, and three subjects (AY, CC, AL) were given both training and the explanation. The four subjects from California were given only the explanation. Two of the five subjects given the training did not successfully complete training, but were moved on to the survey after ten minutes of training (each session was 30 minutes, with 20 minutes allotted for the survey).

Recordings and measurements of the vowels were made to provide a basis on which to assess phonetic similarity. The word list used for recording is the same as the words used for the test cases above, with additional words added so that each vowel is exemplified by at least one word in a non-nasal environment (nasality could make the formants difficult to measure). The word list is randomized. Recordings were made of 2 of the subjects who were surveyed. Each word is located inside a carrier phrase:

讀【X】這個字。
duk6 — ze2 go3 zi6
'read' — 'this' CL 'word'

The list given to the subjects contains two iterations of the word list, once forward and once backward, so that two instances of each word are recorded.

Recordings were made using a DAT recorder at 44.1 kHz, and subsequently downsampled to 22 kHz. Vowel formants were measured using LPC analysis. Vowels were measured at a point where it appeared to be at a steady state. Diphthongs were measured for their first part, at a steady state if possible.

Results

(26) Data from the survey

					JP	EW	AL	CW	JY	VH	WC	VG	SN	OK	total		
1	著	zoek6	術	seot6	1	1	0	1	1	0	1	1	0	1	7	˘	
2	律	leot6	靴	hoe1	0	0	0	1	0	1	1	0	0	0	3	~	
3	香	hoeng1	春	ceon1	0	1	0	1	0	1	0	0	0	0	3	˜	[ɵ, œ]
4	律	leot6	略	loek6	1	0	0	1	1	1	0	1	0	1	6	°	19/40
5	論	leon6	老	lou2	0	1	0	1	0	0	0	0	0	0	2	˘	
6	術	seot6	刀	dou1	0	0	0	0	0	0	0	0	0	0	0	~	[ɵ, o]
7	好	hou2	出	ceot1	0	0	0	0	0	0	0	0	0	0	0	°	2/30
8	空	hung1	卒	zeot1	0	0	0	1	0	0	0	0	0	0	1	˘	
9	福	fuk1	春	ceon1	0	0	0	1	0	0	0	0	0	1	2	~	[ɵ, ʊ]
10	論	leon6	龍	lung4	1	1	0	0	0	0	0	0	0	1	3	°	6/30
11	括	kut3	律	leot6	0	1	0	1	0	0	0	0	0	1	3	˘	
12	進	zeon3	觀	gun1	0	1	0	0	0	0	0	0	0	0	1	~	[ɵ, u]
13	卒	zeot1	夫	fu1	1	0	0	0	0	0	0	0	0	0	1	°	5/30
14	工	gung1	冇	mou5	1	0	0	0	0	0	1	0	0	0	2	˘	
15	屋	uk1	好	hou2	1	1	0	1	0	1	1	0	0	1	6	~	[o, ʊ]
16	刀	dou1	空	hung1	1	1	0	0	1	1	0	0	0	1	5	°	13/30
17	高	gou1	官	gun1	0	1	0	0	0	0	1	0	0	1	3	˘	
18	老	lou2	夫	fu1	0	1	0	0	0	0	1	0	0	0	2	~	[o, u]
19	闊	fut3	好	hou2	0	0	0	0	0	0	0	0	0	0	0	°	5/30
20	空	hung1	烏	wu1	0	1	0	1	0	1	1	0	0	1	5	˘	
21	括	kut3	福	fuk1	0	0	0	1	0	0	0	1	0	0	2	~	[ʊ, u]
22	夫	fu1	龍	lung4	0	0	0	1	0	0	0	1	0	0	2	°	9/30
23	虛	heoi1	靴	hoe1	0	1	0	1	0	1	0	0	0	1	4	˘	
24	堆	deoi1	腳	goek3	0	0	0	1	0	1	0	0	0	1	3	~	[ɵ(y), œ]
25	央	joeng1	區	keoi1	1	0	0	1	1	0	1	0	0	0	4	°	11/30
26	雖	seoi1	卒	zeot1	0	1	0	1	1	1	0	1	1	0	6	˘	
27	春	ceon1	區	keoi1	0	0	0	1	0	1	0	1	1	0	4	~	[ɵ(y), ɵ]
28	雷	leoi4	術	seot6	0	1	0	1	0	0	0	1	0	0	3	°	13/30
29	息	sik1	知	zi1	0	0	1	1	0	0	0	0	0	1	3	˘	[i, ɪ]
30	平	ping4	連	lin4	1	1	0	1	0	0	1	0	0	1	5	°	8/20
31	息	sik1	四	sei3	1	0	0	0	0	1	0	0	0	1	3	˘	[ɪ, e]
32	非	fei1	令	ling6	1	0	0	0	0	0	1	0	0	0	2	°	5/20
33	比	bei2	車	ce1	0	1	0	1	0	0	1	0	0	0	3	˘	[e, ɛ]
34	未	mei6	吃	hek3	0	0	0	0	0	1	0	0	0	0	1	°	4/20
37	高	gou1	河	ho4	0	0	0	0	1	0	1	0	0	0	2	˘	[o, ɔ]
38	刀	dou1	割	got3	0	0	0	0	0	1	1	0	0	1	3	°	5/20
				total	11	16	1	21	6	13	13	7	2	15			

					JP	EW	AL	CW	JY	VH	WC	VG	SN	OK	
40	連	lin4	知	zi1	0	0	1	1	1	1	1	0	1	0	ˇ
41	平	ping4	亦	jik6	0	0	1	1	1	1	0	1	1	1	~
42	非	fei1	利	lei6	1	1	1	1	1	1	1	1	1	1	~
43	車	ce1	吃	hek3	1	1	1	1	1	1	1	1	1	1	~
44	金	gam1	失	sat1	0	0	0	1	1	0	1	1	1	1	~
45	沙	saa1	亞	aa3	1	1	1	1	1	1	1	1	1	1	~
46	香	hoeng1	腳	goek3	1	1	1	1	1	1	1	1	1	1	~
47	春	ceon1	律	leot6	1	0	0	1	0	0	0	1	1	1	~
48	卒	zeot1	律	leot6	1	1	1	1	1	1	1	1	1	1	~
49	於	jyu1	決	kyut3	1	1	0	1	1	1	1	1	1	0	~
50	可	ho2	火	fo2	1	1	1	1	1	1	1	1	1	1	~
51	好	hou2	刀	dou1	1	1	1	1	1	1	1	1	1	1	~
52	龍	lung4	工	gung1	1	1	1	1	1	1	1	1	1	1	~
53	烏	wu1	括	kut3	0	1	0	1	1	1	1	1	1	1	~
54	些	se1	借	ze3	1	1	1	1	1	1	1	1	1	1	~
55	割	got3	作	zok3	1	1	1	1	1	1	1	1	1	1	~
56	亞	aa3	喊	haam3	1	1	1	1	1	1	1	1	1	1	~
57	書	syu1	於	jyu1	1	1	1	1	1	1	1	1	1	1	~
58	連	lin4	詩	si1	0	0	1	1	1	1	1	1	1	1	°
59	凡	faan4	信	seon3	0	0	0	0	0	0	0	0	0	0	ˇ
60	闊	fut3	車	ce1	0	0	0	0	0	0	0	0	0	0	~
61	比	bei2	割	got3	0	0	0	0	0	0	0	0	0	0	~
62	喊	haam3	血	hyut3	0	0	0	0	0	0	0	0	0	0	~
63	丁	ding1	北	bak1	0	0	0	0	0	0	0	0	0	0	~
64	接	zip3	央	joeng1	0	0	0	0	0	0	0	0	0	0	~
65	個	go3	令	ling6	0	0	0	0	0	0	0	0	0	0	~
66	連	lin4	利	lei6	0	0	0	0	0	0	0	0	0	0	~
67	冇	mou5	書	syu1	0	0	0	0	0	0	0	0	0	0	~
68	官	gun1	區	keoi1	0	0	0	0	0	0	0	0	0	0	~
69	風	fung1	下	haa6	0	0	0	0	0	0	0	0	0	0	~
70	好	hou2	金	gam1	0	0	0	0	0	0	0	0	0	0	~
71	著	zoek6	火	fo2	0	1	0	0	0	0	0	0	0	0	~
72	南	naam4	龍	lung4	0	0	0	0	0	0	0	0	0	0	~
73	比	bei2	金	gam1	0	0	0	0	0	0	0	0	0	0	~
74	非	fei1	高	gou1	0	0	0	0	0	0	0	0	0	0	~
75	沙	saa1	夫	fu1	0	1	0	0	0	0	0	0	0	0	~
76	決	kyut3	息	sik1	0	0	1	0	0	0	0	0	0	0	~
77	吃	hek3	信	seon3	0	0	0	0	0	0	0	0	0	0	°

uncontroversially the same

uncontroversially different

The data presented above was considered usable (i.e., passed the control cases) if they fulfilled two conditions: the subject must not have marked any pairs with the same rime as “different,” and the subject must not have marked more than two u-different pairs as “same.” The first condition ensures that there was no confusion with tone. The second is a somewhat arbitrary cutoff, made because relatively few subjects scored

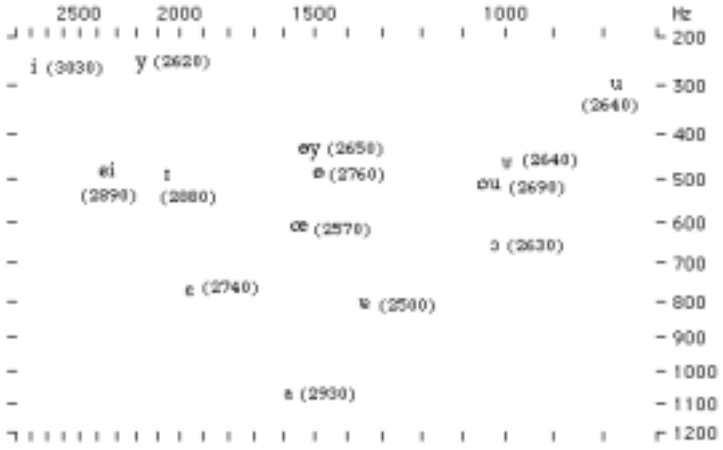
perfectly on the control items; it allows for some noise in the data. See appendix for complete results.

(27) Recording 1 (EW)

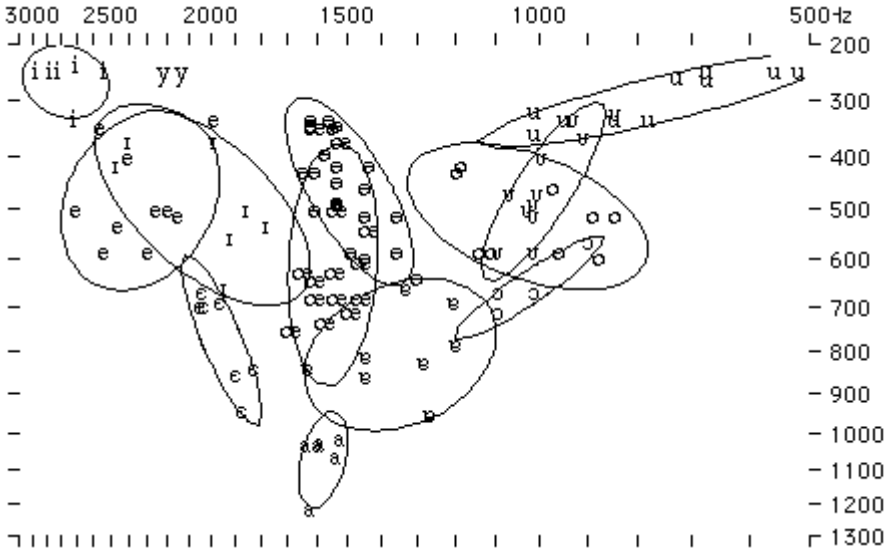
Average values of F1, F2, F3

	F1	F2	F3
y	253.2	2170.5	2622.6
i	265.6	2779.0	3032.5
ɪ	493.7	2056.4	2878.6
e	485.5	2350.2	2892.0
ɛ	775.8	1958.7	2738.6
œ	613.6	1548.6	2565.7
ɵ	491.2	1486.5	2762.0
ɵy	437.3	1512.1	2652.8
ɔ	659.5	1026.3	2626.9
o	515.7	1034.3	2686.5
ʊ	467.0	995.2	2641.6
u	299.4	776.1	2636.3
ɐ	808.2	1350.9	2502.4
a	1072.3	1578.9	2931.0

Plot of average F1, F2, (F3 in parentheses)



F1, F2 data for EW

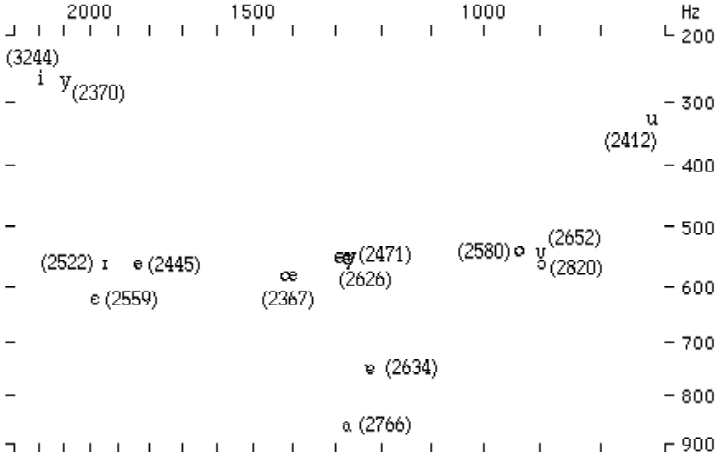


(28) Recording 2 (JP)

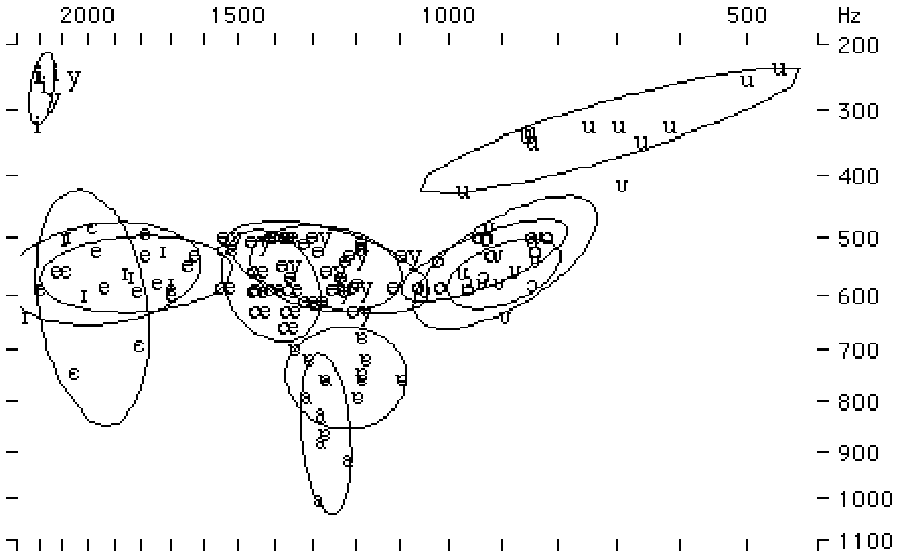
Average values of F1, F2, F3

	F1	F2	F3
y	284.7	2229.8	3244.0
i	252.7	2200.4	2200.4
ɪ	563.8	1912.8	2530.1
e	567.1	1785.6	2419.2
ɛ	617.4	2017.3	2508.3
œ	563.9	1398.3	2312.2
ɵ	534.7	1301.4	2606.9
ɵy	554.1	1283.3	2407.2
ɔ	579.7	889.2	2964.9
o	530.8	917.8	2539.4
ʊ	523.3	876.6	2670.3
u	324.7	672.7	2401.4
ɐ	764.1	1203.6	2654.2
a	855.5	1286.5	2748.1

Plot of average F1, F2, (F3 in parentheses)



F1, F2 data for JP



Discussion

Many of the subjects had great difficulty with the survey. Out of the 16 subjects from Hong Kong, only 6 passed the control cases. Contrary to expectations, subjects found the task to be quite difficult; given two syllables, speakers cannot readily tell whether the vowels are the same or not. Such an ability may well be influenced by an

alphabetic writing system; interestingly, the four subjects from California all passed the control cases.

First, we will discuss those subjects who did poorly on the control items. Results from subjects who did not pass the control cases may have been affected by other factors, such as tone. 5 people marked #51 (hou2, dou1), which has the same finals but different tones, as different. Curiously enough, 5 people marked #72 (naam4, lung4), which has the same tone, as the same, even though the vowels are completely different. In this case, they might have been comparing the general shape of the syllable as well, as both syllables start with a coronal (n and l, which have merged for some speakers), and end with nasals.

Another factor that might have contributed to noise in the data was people misreading characters. Subjects had only the characters to work off of, and with characters taken out of context there is the danger that they will be misread.

Those subjects who did pass the control cases often made more distinctions than one would expect from any phonemic analysis; also, subjects were not consistent with each other. For example, JP and VH respectively marked 11 and 13 of the test cases to be the same. Out of these, only 4 of the pairs overlap. There does not seem to be any obvious continuum of strictness in judging vowel similarity, i.e., it was not the case that a pair marked "same" by a very strict subject (one that tended to mark "different") would also be marked "same" by a less strict subject. The following chart lists the number of disagreements per pair of subjects:

(29)	JP	17	12	22	9	16	12	14	13	12
		EW	17	15	16	17	15	17	16	13
			AL	20	7	14	14	8	3	14
				CW	19	16	20	14	19	16
					JY	13	13	7	6	15
						VH	18	14	11	12
							WC	18	15	16
								VG	5	18
									SN	17
										OK

Subjects were also not consistent with themselves. In fact, none of the subjects were completely consistent for the test cases. Vowels pairs that were consistently judged to be the same by some speakers are listed below:

JP: [ɪ, e], [ʊ, o]
CW: [i, ɪ], [u, ʊ], [œ, ɵ], [œ, ɵ(y)], [ɵ, ɵ(y)]
WC: [o, ɔ]
VG: [ɵ, ɵ(y)]
OK: [i, ɪ]

The results, then, are inconsistent; furthermore, they do not agree with our second hypothesis that speakers will group vowels as traditional phonemic analysis would predict, using the distribution facts. One speaker, CW, came close, grouping many of the vowels in complementary distribution, but ended up grouping one too many; CW not only grouped [œ, ɵ] together, but also grouped two out of the three [ɵ, ʊ] items. While each individual pair is in complementary distribution, it is not possible to group all three vowels as the same phoneme, because there is a contrast between [œ, ʊ], e.g., [hʊŋ], [hœŋ].

Given these results, there are two possibilities with respect to our main hypothesis. One is that the subjects do have underlying representations of the vowels, but they were not accessing them, using instead phonetic criteria to make similarity judgements. However, this would not explain why EW, for example, decided that #53

(wu1, kut3) was the same, but #4 (leot6, loek6) was different, even though the latter pair for this speaker was acoustically more similar. Similarly, even though #40 (lin4, zi1) are acoustically similar, 4 speakers (JP, EW, VG, OK) judged them to be different.

(30)

	EW			JP		
average	F1	F2	F3	F1	F2	F3
wu1	259	702	2714	250	478	2245
kut3	330	981	2519	345	847	2502
leot6	561	1387	2798	577	1213	2586
loek6	528	1482	2714	614	1379	2423
lin4	296	2853	3048	312	2263	3164
zi1	253	2714	2962	280	2195	3132

Also, it appears that those subjects who did pass the control cases were able to isolate the vowel, rather than comparing rimes or tones or some part of the syllable large than the vowel, since all subjects agreed that #43 (ce1, hek3), #46 (hoeng1, goek3), #55 (got3, zok3), and #56 (aa3, haam3) were the same.

The other possibility is that the subjects were being more or less consistent with their underlying representations, but their URs are more finely differentiated than expected. Of the control items that were u-same, many of the ones judged to be different were pairs in which one was a checked syllable (one which ends in an unreleased [p, t, k], traditionally known as the entering tone, or 入聲/rùshēng). In the test cases as well, subjects tended to treat checked syllables differently from other syllables. In the set of items testing [ɵ] and [œ], 7 speakers judged #1 (zoek6, seot6) to be the same, and 6 speakers judged #4 (leot6, loek6) to be the same, whereas only 3 people judged #2 (leot6, hoe1) and #3 (hoeng1, ceon1) the same. Vowels in checked syllables are shorter than their counterparts in open or nasal coda syllables (Li 1985). Thus, for these subjects, vowel quality (as indicated by formants) may not be as salient a feature

of vowels as, say, length, or glottalization.

By reordering our test items and u-same items, we can see that checked vs. non-checked pairs tend to be judged as different, relative to pairs which compared the same type of syllable. The exceptions to this pattern are #15 (uk1, hou2) and #26 (seoi1, zeot1). Each of the checked vs. non-checked test items were judged to be the same by 2.3 subjects on average; non-checked vs. non-checked by 3.0 subjects; and checked vs. checked by 4.5 subjects. For the u-same items, almost all of them that were marked "different" were comparing checked vs. non-checked. On the other hand, u-same items which were checked vs. checked or non-checked vs. non-checked were all judged to be the same, with the exception of #40 (lin4, zi1) and #58 (lin4, si1).

					JP	EW	AL	CW	JY	VH	WC	VG	SN	OK	total	
2	律	leot6	靴	hoe1	0	0	0	1	0	1	1	0	0	0	3	˘
6	術	seot6	刀	dou1	0	0	0	0	0	0	0	0	0	0	0	~
7	好	hou2	出	ceot1	0	0	0	0	0	0	0	0	0	0	0	~
8	空	hung1	卒	zeot1	0	0	0	1	0	0	0	0	0	0	1	~
9	福	fuk1	春	ceon1	0	0	0	1	0	0	0	0	0	1	2	~
13	卒	zeot1	夫	fu1	1	0	0	0	0	0	0	0	0	0	1	~
15	屋	uk1	好	hou2	1	1	0	1	0	1	1	0	0	1	6	~
19	闊	fut3	好	hou2	0	0	0	0	0	0	0	0	0	0	0	~
26	雖	seoi1	卒	zeot1	0	1	0	1	1	1	0	1	1	0	6	~
28	雷	leoi4	術	seot6	0	1	0	1	0	0	0	1	0	0	3	~
29	息	sik1	知	zi1	0	0	1	1	0	0	0	0	0	1	3	~
31	息	sik1	四	sei3	1	0	0	0	0	1	0	0	0	1	3	~
34	未	mei6	吃	hek3	0	0	0	0	0	1	0	0	0	0	1	~
38	刀	dou1	割	got3	0	0	0	0	0	1	1	0	0	1	3	~
41	平	ping4	亦	jik6	0	0	1	1	1	1	0	1	1	1		~
43	車	ce1	吃	hek3	1	1	1	1	1	1	1	1	1	1		~
44	金	gam1	失	sat1	0	0	0	1	1	0	1	1	1	1		~
46	香	hoeng1	腳	goek3	1	1	1	1	1	1	1	1	1	1		~
47	春	ceon1	律	leot6	1	0	0	1	0	0	0	1	1	1		~
49	於	jyu1	決	kyut3	1	1	0	1	1	1	1	1	1	0		~
53	烏	wu1	括	kut3	0	1	0	1	1	1	1	1	1	1		°
1	著	zoek6	術	seot6	1	1	0	1	1	0	1	1	0	1	7	˘
4	律	leot6	略	loek6	1	0	0	1	1	1	0	1	0	1	6	~
11	括	kut3	律	leot6	0	1	0	1	0	0	0	0	0	1	3	~
21	括	kut3	福	fuk1	0	0	0	1	0	0	0	1	0	0	2	~
48	卒	zeot1	律	leot6	1	1	1	1	1	1	1	1	1	1		~
55	割	got3	作	zok3	1	1	1	1	1	1	1	1	1	1		°

checked vs. non-checked

checked vs. checked

3	香	hoeng1	春	ceon1	0	1	0	1	0	1	0	0	0	0	3	˜
5	論	leon6	老	lou2	0	1	0	1	0	0	0	0	0	0	2	˜
10	論	leon6	龍	lung4	1	1	0	0	0	0	0	0	0	1	3	˜
12	進	zeon3	觀	gun1	0	1	0	0	0	0	0	0	0	0	1	˜
14	工	gung1	冇	mou5	1	0	0	0	0	0	1	0	0	0	2	˜
16	刀	dou1	空	hung1	1	1	0	0	1	1	0	0	0	1	5	˜
17	高	gou1	官	gun1	0	1	0	0	0	0	1	0	0	1	3	˜
18	老	lou2	夫	fu1	0	1	0	0	0	0	1	0	0	0	2	˜
20	空	hung1	烏	wu1	0	1	0	1	0	1	1	0	0	1	5	˜
22	夫	fu1	龍	lung4	0	0	0	1	0	0	0	1	0	0	2	˜
23	虛	heoi1	靴	hoe1	0	1	0	1	0	1	0	0	0	1	4	˜
24	堆	deoi1	腳	goek3	0	0	0	1	0	1	0	0	0	1	3	˜
25	央	joeng1	區	keoi1	1	0	0	1	1	0	1	0	0	0	4	˜
27	春	ceon1	區	keoi1	0	0	0	1	0	1	0	1	1	0	4	˜
30	平	ping4	連	lin4	1	1	0	1	0	0	1	0	0	1	5	˜
32	非	fei1	令	ling6	1	0	0	0	0	0	1	0	0	0	2	˜
33	比	bei2	車	ce1	0	1	0	1	0	0	1	0	0	0	3	˜
37	高	gou1	河	ho4	0	0	0	0	1	0	1	0	0	0	2	˜
40	連	lin4	知	zi1	0	0	1	1	1	1	1	0	1	0		˜
42	非	fei1	利	lei6	1	1	1	1	1	1	1	1	1	1		˜
45	沙	saa1	亞	aa3	1	1	1	1	1	1	1	1	1	1		˜
50	可	ho2	火	fo2	1	1	1	1	1	1	1	1	1	1		˜
51	好	hou2	刀	dou1	1	1	1	1	1	1	1	1	1	1		˜
52	龍	lung4	工	gung1	1	1	1	1	1	1	1	1	1	1		˜
54	些	se1	借	ze3	1	1	1	1	1	1	1	1	1	1		˜
56	亞	aa3	喊	haam3	1	1	1	1	1	1	1	1	1	1		˜
57	書	syu1	於	jyu1	1	1	1	1	1	1	1	1	1	1		˜
58	連	lin4	詩	si1	0	0	1	1	1	1	1	1	1	1		°

non-checked
vs.
non-checked

#40 (lin4, zi1), which 4 subjects judged to be different, and #58 (lin4, si1), which 2 subjects judged to be different, is more puzzling, since the acoustic data for lin4 and zi1 suggest that the vowels are very similar. It may be that these subjects were judging based on nasality, but that would not explain why all of them marked #56 (aa3, haam3) as the same. It may also be that tone and coda consonant both have an effect on the judgement of vowel similarity. So, although subjects would have no problem with different tones if the coda consonants were the same, as in #52 (lung4, gung1), they might have trouble if the coda consonants were different, as in #40 and #58.

If anything, the results support the hypothesis suggested by Lexicon Optimization in a surface-oriented framework such as OT. Assuming that the survey

was able to access the underlying forms of the vowels for some speakers, these underlying forms appear to be quite similar to the surface forms.

Conclusion

This study attempted to access the underlying representations of Cantonese vowels by asking native speakers for their judgements on which vowels "sound the same." The results suggest that the underlying forms may indeed be quite close to the surface forms, supporting the idea that "learners acquire something close to what they hear" (Yip), absent alternations that would give evidence for phonemic groupings.

One may be tempted to throw their hands up and agree with Twaddell that "(1) we have no right to guess about the linguistic workings of an inaccessible 'mind', and (2) we can secure no advantage from such guesses" (p. 56). Nevertheless, this study did yield some interesting results. Further experiments, such as Jaeger's concept formation tests, may prove worthwhile. In this method, the training phase is more elaborate and figures more prominently in the experiment than in the present study. Furthermore, rather than asking directly for vowel equivalence, this method asks the subject to listen for "a certain sound" (Jaeger, p. 215) which is to be deduced by the subject from the training tokens. It is hoped that these more elaborate and indirect procedures would yield more definite results.

REFERENCES

Bauer, Robert S., and Paul K. Benedict. 1997. *Modern Cantonese Phonology.*

Chao, Yuen-Ren. 1934. The Non-Uniqueness of Phonemic Solutions of Phonetic Systems. In Joos (ed.) 1958, *Readings in Linguistics.*

Chao, Yuen-Ren. 1947. *Cantonese Primer.*

Hashimoto, Oi-kan Yue. 1972. *Phonology of Cantonese.*

Jaeger, Jeri J. 1986. Concept Formation as a Tool for Linguistic Research. In Ohala (ed.), *Experimental Phonology.*

李行德 [Li, Xíngdé/Lee, Thomas]. 1985. 廣州話元音的音值及長短對立 [The phonetic quality and long/short distinction of Cantonese vowels]. 方言 [Fangyan/Dialect], 7.1: 28-38.

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: constraint interaction in generative grammar.

Sapir, Edward. 1925. Sound Patterns in Language. In Joos (ed.) 1958, *Readings in Linguistics*, 19-25.

Twaddell, W. Freeman. 1935. On Defining the Phoneme. In Joos (ed.) 1958, *Readings in Linguistics*, 55-80.

Yip, Moira. 1996. Lexicon Optimization in Languages without Alternations. In Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>, ROA-135-0996.

APPENDICES

A: Survey on Vowel Equivalence

B: Complete Survey Results

C: Complete Acoustic Data

Background

This is a study of vowel similarity in Cantonese.

Some words clearly contain the same vowel. For example,

八 and 鴨

both contain the vowel in「呀」. Likewise,

閃 and 必

both contain the vowel in「衣」.

Other words, however, are less straightforward. This survey asks for your judgements on vowel similarity.

Directions

For each pair of words below, indicate whether you think their vowels are the same, or different.

龍	工	same	different	好	出	same	different
接	央	same	different	非	利	same	different
比	車	same	different	可	火	same	different
沙	亞	same	different	好	刀	same	different
進	觀	same	different	夫	龍	same	different
非	令	same	different	亞	喊	same	different
堆	腳	same	different	術	刀	same	different
空	卒	same	different	春	區	same	different
南	龍	same	different	金	失	same	different
吃	信	same	different	些	借	same	different
福	春	same	different	喊	血	same	different
著	術	same	different	比	金	same	different
比	割	same	different	個	令	same	different
屋	好	same	different	老	夫	same	different
非	高	same	different	車	吃	same	different
連	利	same	different	息	四	same	different
央	區	same	different	春	律	same	different
烏	括	same	different	雖	卒	same	different
闊	好	same	different	連	知	same	different

著 火 same different
息 知 same different
丁 北 same different
凡 信 same different
括 律 same different
卒 夫 same different
官 區 same different
書 於 same different
虛 靴 same different
沙 夫 same different
空 烏 same different
高 河 same different
高 官 same different
平 連 same different
工 冇 same different
刀 空 same different
好 金 same different
論 老 same different
律 略 same different
闊 車 same different

刀 割 same different
金 沙 same different
風 下 same different
論 龍 same different
冇 書 same different
香 腳 same different
未 吃 same different
卒 律 same different
括 福 same different
割 作 same different
律 靴 same different
平 亦 same different
於 決 same different
杏 喊 same different
香 春 same different
敢 咁 same different
決 息 same different
雷 術 same different
連 詩 same different

					AF	AC	JP	JC	CW	JW	JY	EW	EY	HW	TK	VH	CY	AY	CC	AL	WC	VG	SN	OK
1	著	zoek6	術	seot6	0	1	1	0	1	1	1	1	1	1	1	0	0	0	1	0	1	1	0	1
2	律	leot6	靴	hoe1	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0
3	香	hoeng1	春	ceon1	1	0	0	0	1	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0
4	律	leot6	略	loek6	0	1	1	0	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	1
5	論	leon6	老	lou2	0	0	0	0	1	0	0	1	1	1	1	0	0	1	1	0	0	0	0	0
6	術	seot6	刀	dou1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
7	好	hou2	出	ceot1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	空	hung1	卒	zeot1	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	0	0	0	0	0
9	福	fuk1	春	ceon1	0	1	0	0	1	1	0	0	1	0	0	0	1	1	0	0	0	0	0	1
10	論	leon6	龍	lung4	1	0	1	0	0	0	0	1	1	0	1	0	1	1	1	0	0	0	0	1
11	括	kut3	律	leot6	0	1	0	0	1	1	0	1	1	1	1	0	0	1	0	0	0	0	0	1
12	進	zeon3	觀	gun1	0	0	0	0	0	1	0	1	0	1	1	0	1	0	0	0	0	0	0	0
13	卒	zeot1	夫	fu1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	工	gung1	冇	mou5	1	1	1	0	0	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0
15	屋	uk1	好	hou2	0	1	1	0	1	0	0	1	1	1	0	1	0	1	1	0	1	0	0	1
16	刀	dou1	空	hung1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	1
17	高	gou1	官	gun1	0	1	0	0	0	0	0	1	1	0	1	0	0	1	1	0	1	0	0	1
18	老	lou2	夫	fu1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
19	闊	fut3	好	hou2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
20	空	hung1	烏	wu1	0	1	0	0	1	0	0	1	1	1	1	1	0	1	1	0	1	0	0	1
21	括	kut3	福	fuk1	1	1	0	0	1	1	0	0	1	0	0	0	0	0	1	0	0	1	0	0
22	夫	fu1	龍	lung4	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
23	虛	heoi1	靴	hoe1	0	1	0	0	1	0	0	1	1	0	0	1	1	1	1	0	0	0	0	1
24	堆	deoi1	腳	goek3	0	0	0	0	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	1
25	央	joeng1	區	keoi1	0	0	1	0	1	0	1	0	1	0	0	0	0	1	0	0	1	0	0	0
26	雖	seoi1	卒	zeot1	0	1	0	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0
27	春	ceon1	區	keoi1	0	0	0	0	1	0	0	0	0	1	1	1	1	0	1	0	0	1	1	0
28	雷	leoi4	術	seot6	1	1	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
29	息	sik1	知	zi1	0	1	0	0	1	0	0	0	1	0	1	0	0	1	0	1	0	0	0	1
30	平	ping4	連	lin4	0	0	1	0	1	1	0	1	1	0	1	0	0	0	0	0	1	0	0	1
31	息	sik1	四	sei3	1	1	1	0	0	1	0	0	1	1	1	1	1	1	1	0	0	0	0	1
32	非	fei1	令	ling6	1	1	1	0	0	0	0	0	1	1	0	0	1	0	0	0	1	0	0	0
33	比	bei2	車	ce1	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0
34	未	mei6	吃	hek3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
35	金	gam1	沙	saa1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	杏	hang6	喊	haam3	0	1	0	0	1	0	0	0	1	1	0	1	1	1	0	0	0	0	0	0
37	高	gou1	河	ho4	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0
38	刀	dou1	割	got3	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1
39	敢	gam2	咁	gam3	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

					AF	AC	JP	JC	CW	JW	JY	EW	EY	HW	TK	VH	CY	AY	CC	AL	WC	VG	SN	OK
40	連	lin4	知	zi1	1	1	0	0	1	1	1	0	1	0	0	1	1	0	0	1	1	0	1	0
41	平	ping4	亦	jik6	0	1	0	0	1	0	1	0	1	0	1	1	0	0	0	1	0	1	1	1
42	非	fei1	利	lei6	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1
43	車	ce1	吃	hek3	0	0	1	0	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1
44	金	gam1	失	sat1	0	0	0	0	1	1	1	0	1	0	0	0	0	1	0	0	1	1	1	1
45	沙	saa1	亞	aa3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
46	香	hoeng1	腳	goek3	0	1	1	0	1	0	1	1	1	0	0	1	0	1	1	1	1	1	1	1
47	春	ceon1	律	leot6	1	0	1	0	1	0	0	0	1	0	0	0	1	1	1	0	0	1	1	1
48	卒	zeot1	律	leot6	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
49	於	jyu1	決	kyut3	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	0
50	可	ho2	火	fo2	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
51	好	hou2	刀	dou1	0	1	1	0	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1
52	龍	lung4	工	gung1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
53	烏	wu1	括	kut3	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
54	些	se1	借	ze3	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
55	割	got3	作	zok3	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
56	亞	aa3	喊	haam3	0	1	1	0	1	0	1	1	1	0	1	1	0	1	1	1	1	1	1	1
57	書	syu1	於	jyu1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
58	連	lin4	詩	si1	1	1	0	0	1	1	1	0	1	0	1	1	1	0	1	1	1	1	1	1
59	凡	faan4	信	seon3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
60	闊	fut3	車	ce1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
61	比	bei2	割	got3	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
62	喊	haam3	血	hyut3	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
63	丁	ding1	北	bak1	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	0	0	0	0	0
64	接	zip3	央	joeng1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
65	個	go3	令	ling6	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
66	連	lin4	利	lei6	1	1	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0
67	冇	mou5	書	syu1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
68	官	gun1	區	keoi1	1	1	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	0	0	0
69	風	fung1	下	haa6	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
70	好	hou2	金	gam1	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0
71	著	zoek6	火	fo2	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0
72	南	naam4	龍	lung4	0	1	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0
73	比	bei2	金	gam1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
74	非	fei1	高	gou1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0
75	沙	saa1	夫	fu1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0
76	決	kyut3	息	sik1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0
77	吃	hek3	信	seon3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

EW - iteration 1

#	word	F1	F2	F3
1	jyu1	253.1	2123.4	2531.2
13	lin4	253.1	2981.2	2981.2
18	zi1	253.1	2882.8	2882.8
33	ji1	253.1	2826.5	2826.5
25	ling6	562.5	1926.5	2995.3
44	sik1	506.2	1856.2	2812.5
49	jik6	421.8	2475.0	2882.8
16	bei2	590.6	2531.2	3107.8
30	hei3	351.5	2559.3	3093.7
46	sei3	506.2	2250.0	2826.5
50	mei6	506.2	2700.0	3107.8
60	fei1	337.5	1982.8	2784.3
11	ce1	703.1	2053.1	2868.7
31	hek3	843.7	1828.1	2643.7
37	peng4	857.8	1898.4	2545.3
51	be1	703.1	2039.0	2854.6
2	joeng1	759.3	1687.5	2615.6
24	hoe1	379.6	1518.7	2517.1
26	goek3	689.0	1532.8	2460.9
36	hoeng1	351.5	1603.1	2545.3
39	zoek6	632.8	1645.3	2531.2
53	hoe1	646.8	1603.1	2615.6
54	loek6	548.4	1434.3	2714.0
3	ceot1	435.9	1645.3	2714.0
5	seot6	590.6	1490.6	2685.9
7	zeon3	492.1	1532.8	2615.6
27	zeot1	506.2	1603.1	2784.3
29	leon6	520.3	1350.0	2798.4
34	ceon1	337.5	1560.9	2629.6
43	leot6	604.6	1448.4	2798.4
4	keoi1	421.8	1532.8	2545.3
41	seoi1	421.8	1532.8	2629.6
48	deoi1	351.5	1546.8	2714.0
56	heoi1	337.5	1617.1	2545.3
59	leoi4	590.6	1350.0	2896.8
10	ho4	675.0	1012.5	2629.6
55	got3	717.1	1096.8	2545.3
9	lou2	590.6	1139.0	2798.4
21	mou5	590.6	956.2	2629.6
38	hou2	520.3	843.7	2784.3
47	gou1	435.9	1195.3	2643.7
57	dou1	421.8	1181.2	2629.6
6	gung1	337.5	928.1	2460.9
12	lung4	590.6	1096.8	2812.5
19	hung1	407.8	998.4	2460.9
22	huk1	478.1	1068.7	2671.8
28	fuk1	506.2	1026.5	2629.6
35	uk1	506.2	1026.5	2896.8
8	fu1	253.1	520.3	2629.6
14	wu1	253.1	675.0	2714.0
40	fut3	365.6	1012.5	2587.5
52	kut3	323.4	1012.5	2503.1
58	gun1	267.1	675.0	2615.6
15	gam3	843.7	1631.2	2460.9
17	gam1	815.6	1448.4	2489.0
20	hap6	829.6	1279.6	2587.5
32	hang6	787.5	1199.4	2269.9
23	haa6	1040.6	1589.0	2587.5
42	saa1	1040.6	1631.2	2756.2
45	haam3	1068.7	1532.8	2896.8

EW - iteration 2

#	word	F1	F2	F3
60	jyu1	253.4	2217.6	2714.0
48	lin4	337.9	2724.5	3115.3
43	zi1	253.4	2545.0	3041.3
28	ji1	242.8	2714.0	3347.6
36	ling6	538.5	1784.6	2661.2
24	ping4	666.9	1948.6	3042.8
17	sik1	380.1	1995.9	2819.6
12	jik6	380.1	2407.7	2935.7
45	bei2	591.3	2291.5	2714.0
31	hei3	538.5	2450.0	2766.8
15	sei3	506.8	2196.5	2882.9
11	mei6	517.4	2143.7	2661.2
1	fei1	408.8	2396.7	2974.7
50	ce1	675.8	2048.7	2840.7
30	hek3	950.4	1879.7	2566.1
10	be1	696.9	1964.2	2851.2
59	joeng1	686.4	1605.1	2629.5
37	hoe1	633.6	1541.8	2576.7
35	goek3	686.4	1457.3	2450.0
25	hoeng1	718.1	1489.0	2597.8
22	zoek3	739.2	1573.4	2460.5
8	hoe1	612.5	1457.3	2492.2
7	loek6	506.8	1531.2	2714.0
58	ceot1	496.3	1531.2	2809.0
56	seot6	517.4	1446.7	3104.7
54	zeon3	464.6	1446.7	3030.8
34	zeot1	401.2	1573.4	2819.6
32	leon6	644.1	1298.9	2629.5
27	ceon1	348.4	1531.2	2450.0
18	leot6	517.4	1351.7	2798.4
57	keoi1	432.9	1605.1	2566.1
20	seoi1	454.0	1531.2	2629.5
13	deoi1	422.4	1436.2	2661.2
5	heoi1	348.4	1615.7	2618.9
2	leoi4	592.1	1353.4	2721.0
51	ho4	570.2	897.6	2904.0
6	got3	675.8	1098.2	2428.8
52	lou2	591.3	1108.8	2946.3
40	mou5	601.9	876.5	2566.1
23	hou2	517.4	887.0	2629.5
14	gou1	464.6	971.5	2629.5
4	dou1	422.9	1184.2	2608.2
55	gung1	327.3	844.8	2450.0
49	lung4	591.3	1013.7	2513.3
42	hung1	475.2	1003.2	2481.6
39	huk1	496.3	1013.7	2724.5
33	fuk1	517.4	1013.7	2787.9
26	uk1	369.6	908.1	2809.0
53	fu1	253.4	559.6	2714.0
47	wu1	264.0	728.6	2714.0
21	fut3	337.9	781.4	2587.2
9	kut3	337.9	950.4	2534.4
3	gun1	338.3	845.9	2763.2
46	gam3	865.9	1446.7	2629.5
44	gam1	665.3	1330.6	2481.6
41	hap6	960.9	1267.2	2851.2
29	hang6	696.9	1203.8	2249.3
38	haa6	1024.3	1520.6	2999.1
19	saa1	1034.9	1584.0	3083.6
16	haam3	1225.0	1615.7	3263.1

JP - iteration 1

#	word	F1	F2	F3
1	jyu1	252.7	2041.6	2304.1
13	lin4	252.0	2250.1	3371.5
18	zi1	279.4	2206.3	3215.4
33	ji6	242.9	2226.2	3332.5
25	ling6	583.3	1701.3	2411.1
37	ping4	602.7	2012.5	2488.8
44	sik1	563.8	1856.9	2527.7
49	jik6	505.5	2080.5	2693.0
51	bei1	534.7	1623.6	2401.3
16	bei2	593.0	1691.6	2430.5
30	hei3	593.0	2187.5	2547.2
46	sei3	495.8	1779.1	2625.0
50	mei6	602.7	1691.6	2333.3
60	fei1	583.3	1740.2	2177.7
11	ce1	486.1	1983.3	2459.7
31	hek3	748.6	2051.3	2556.9
2	joeng1	525.0	1526.3	2216.6
24	hoe1	505.5	1390.2	2362.5
26	goek3	505.5	1361.1	2236.1
36	hoeng3	661.1	1361.1	2313.8
39	zoek6	593.0	1351.3	2284.7
54	loek6	593.0	1400.0	2459.7
3	ceot1	505.5	1409.7	2625.0
5	seot6	515.2	1186.1	2508.3
7	zeon3	515.2	1322.2	2576.3
27	zeot1	505.5	1361.1	2634.7
29	leon6	525.0	1186.1	2663.8
34	ceon1	593.0	1448.6	2673.6
43	leot6	583.3	1195.8	2566.6
4	keoi1	505.5	1283.3	2381.9
41	seoi1	505.5	1516.6	2459.7
48	deoi1	593.0	1234.7	2333.3
56	heoi3	631.9	1186.1	2459.7
59	leoi4	534.7	1195.8	2401.3
10	ho4	585.8	845.1	3207.6
55	got3	573.6	933.3	2722.2
9	lou2	544.4	1020.8	2605.5
21	mou5	505.5	816.6	2459.7
38	hou2	505.5	845.8	2498.6
47	gou1	505.5	943.1	2586.1
57	dou1	593.0	962.5	2547.2
6	gung1	486.1	923.6	2780.5
12	lung4	563.8	972.2	2547.2
19	hung1	583.3	904.1	2625.0
22	huk1	505.5	836.1	2780.5
28	fuk1	583.3	933.3	2586.1
35	uk1	418.0	690.2	2702.7
8	fu3	350.0	661.1	2352.7
14	wu6	259.0	500.3	2287.1
40	fut3	340.2	855.5	2381.9
52	kut3	350.0	845.8	2537.5
58	gun1	330.5	748.6	2430.5
15	gam3	758.3	1098.6	2712.5
17	gam1	865.2	1273.6	2537.5
20	hap6	758.3	1186.1	2722.2
32	hang6	758.3	1273.6	2848.6
53	bak1	680.5	1186.1	2450.0
23	haa6	758.3	1263.8	2800.0
42	saa1	1011.1	1283.3	2663.8
45	haam3	797.2	1312.5	2780.5

JP - iteration 2

#	word	F1	F2	F3
60	jyu1	285.4	2131.8	2435.1
48	lin4	372.1	2276.7	2955.6
43	zi1	279.8	2184.4	3048.8
28	ji6	281.9	2235.3	3540.5
36	ling6	570.8	1837.5	2613.5
24	ping4	642.2	2265.6	2461.8
17	sik1	526.2	1730.4	2461.8
12	jik6	508.4	2096.1	2524.3
45	bei2	597.6	1810.7	2444.0
31	hei3	561.9	2122.9	2506.4
15	sei3	588.7	1926.6	2658.1
11	mei6	553.0	1641.2	2212.1
10	bei1	535.1	1783.9	2283.4
1	fei1	527.1	1956.6	2716.0
50	ce1	561.9	2087.2	2497.5
30	hek3	695.7	1810.7	2720.5
59	joeng1	588.7	1534.2	2461.8
37	hoe1	561.9	1445.0	2372.6
35	goek3	597.6	1445.0	2426.2
25	hoeng3	633.3	1436.1	2417.2
22	zoek6	615.4	1311.2	2470.8
7	loek6	634.3	1358.0	2385.4
58	ceot1	597.6	1222.0	2613.5
56	seot6	526.2	1284.4	2640.2
54	zeon3	544.1	1222.0	2488.6
34	zeot1	570.8	1355.8	2711.6
32	leon6	588.7	1114.9	2738.4
27	ceon1	615.4	1275.5	2720.5
18	leot6	570.8	1230.9	2604.6
57	keoi1	561.9	1248.7	2426.2
20	seoi1	553.0	1355.8	2533.2
13	deoi1	588.7	1186.3	2435.1
5	heoi3	509.2	1438.4	2242.5
2	leoi4	536.0	1081.0	3037.6
51	ho4	511.5	846.7	2760.6
6	got3	580.7	964.9	2590.9
52	lou2	588.7	1061.4	2622.4
40	mou5	526.2	838.4	2622.4
23	hou2	535.1	918.7	2622.4
14	gou1	508.4	927.6	2524.3
4	dou1	589.6	1018.5	2707.1
55	gung1	508.4	927.6	2577.8
49	lung4	597.6	1052.5	2488.6
42	hung1	642.2	891.9	2791.9
39	huk1	561.9	874.1	2640.2
33	fuk1	535.1	909.8	2631.3
26	uk1	544.1	838.4	2675.9
53	fu3	330.0	615.4	2372.6
47	wu6	240.8	454.9	2203.2
21	fut3	428.1	972.2	2408.3
9	kut3	339.5	848.7	2465.8
3	gun1	330.5	696.8	2680.3
46	gam3	722.5	1177.4	2631.3
44	gam1	722.5	1311.2	2417.2
41	hap6	793.8	1195.2	2711.6
29	hang6	704.6	1346.9	2800.8
8	bak1	750.4	1188.2	2510.5
38	haa6	918.7	1213.1	2720.5
19	saa1	883.0	1275.5	2720.5
16	haam3	829.5	1275.5	2907.8