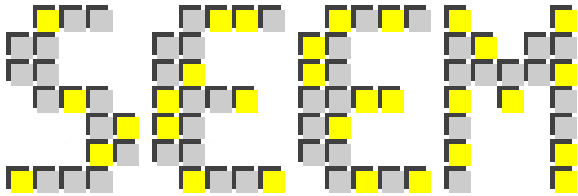# The Chinese University of Hong Kong

Department of
Systems Engineering &
Engineering
Management

Technical Report

A Contrastive Phonetic Study between Cantonese and English To Predict Salient Mispronunciations by Cantonese Learners of English

Helen Meng
Eric Zee &
Wai Sum Lee

A Contrastive Phonetic Study between Cantonese and English To Predict Salient Mispronunciations by Cantonese Learners of English

Helen Meng[1], Eric Zee[2] and Wai-Sum Lee[2]
[1]Human-Computer Communications Laboratory
Department of Systems Engineering & Engineering Management
The Chinese University of Hong Kong
hmmeng@se.cuhk.edu.hk

[2]Phonetics Laboratory
Department of Chinese, Translation and Linguistics
City University of Hong Kong
ctlzee@cityu.edu.hk, wslee@cityu.edu.hk

ABSTRACT

This work aims to derive salient mispronunciations made by Cantonese (L1) learners of English (L2) in order to support the design pedagogical and remedial instructions. Our approach is grounded on the theory of language transfer and involves systematic phonological comparisons between the two languages at the phonetic, phonotactic and prosodic levels. Major disparities across the language pair are identified to focus on phonological contexts where transfer effects are prominent. This methodology enables us to propose salient pronunciation errors that are cross-validated with anecdotal examples observed from daily conversational interactions.

Keywords: salient mispronunciations, second language acquisition

1. INTRODUCTION

The objective of this work is to derive salient mispronunciations made by Cantonese (L1) learners of English (L2). Our long-term goal is to design effective pedagogical and remedial instructions for pronunciation improvement. The target learners are adults (university students) who are native Cantonese speakers seeking to improve their pronunciation of English. In this work, the primary language (L1) is Cantonese, a major dialect spoken by over 60 million people in Hong Kong, Macau, South China and many oversees communities. The secondary language (L2) is American English.

Pronunciation errors may be due to a diversity of factors, such as an imperfect understanding of semantics, syntax, morphology, phonology, coarticulatory effects and letter-to-sound rules. As an initial step, we focus on phonology. Our proposed approach is grounded on the theory of language transfer [2,5] and involves phonological comparisons between L1 and L2 across the phonetic, phonotactic and prosodic levels. We identify major disparities across the language pair, such as missing phones and violation of phonotactic constraints, in order to focus on phonological contexts where perceived interferences of transfer features are prominent. This procedure enables us to devise a methodology for proposing salient mispronunciations that are cross-validated with anecdotal examples observed from daily conversational interactions with university students over many years. The linguistic discrepancies may also offer an explanatory model for us to understand the cause of errors [1]. In the following, we present the results of our comparison at different phonological levels and provide illustrate examples of mispronunciations.

2. PHONETIC COMPARISON

2.1. Vowels and Diphthongs

Figure 1 illustrates the Cantonese vowel charts containing the 4 short vowels, 7 long vowels and 10 diphthongs. Appendix A presents some Chinese characters whose syllable pronunciations contain these vowels and diphthongs. Meanings of the characters are parenthesized. Figure 2 illustrates the American English vowel charts containing the 13 vowels and 3 diphthongs in American English. The reduced vowel

/ə/ is excluded because its quality varies considerably based on coarticulatory context [4]. Appendix B presents English words containing these vowels and diphthongs [8].
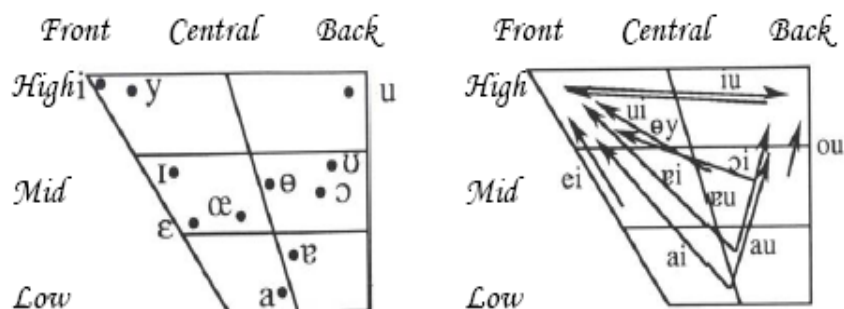

Figure 1. Cantonese vowels and diphthongs, based on [6].
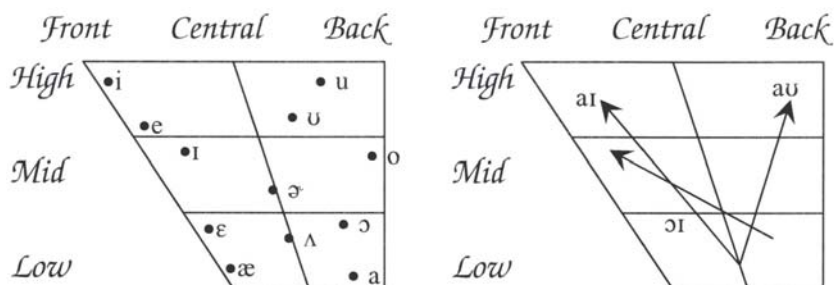Tongue positions (front, central, back, high, mid, low) are labeled.

Figure 2. American English vowels & diphthongs, based on [4].

Comparison between the vowel charts in Tables 1 and 2 helps organize our observations on common mispronunciations due to English vowels that are *missing* in the Cantonese phonetic inventory. This missing set includes /e, æ, o, ɚ ʌ, ɑ/. Hence when Cantonese native speakers enunciate these English vowels, they replace with Cantonese vowels that are close in terms of production and perception. Depending on the degree of resemblance, a *subset* of these vowels may be perceived as mispronunciations, due to prominent transfer effects from Cantonese (L1) to English (L2). Table 3 offers illustrative examples. Common substitutions for the remaining vowels in the missing set are shown in Table 4 and these are often deemed acceptable.

Table 3. Salient vowel mispronunciations of Cantonese speakers learning English.

Target English Vowel	Replacement from Cantonese	Examples of mispronunciations
æ	ɛ	"had" /hæd/ is often pronounced inaccurately as "head" /hɛd/.
ɚ	œ	Retroflexion is replaced with lip-rounding, e.g., "her" /h ɚ/ is replaced by /h œ /, which sounds like 靴 (boot).
ɑ	ɐ	The English vowel /ɑ/ (e.g. in the pronunciation of "dodge") typically has lower first and second formant frequencies than the Cantonese /ɐ/ [7] (e.g. in the pronunciation of 打 'hit'). The perceived difference, however, may be subtle.
ə	variable	The chosen substitute for the English reduced vowel /ə/ varies based on the articulatory context as Cantonese does not have a reduced vowel. For example, "about" /əbaʊt/ may be mispronounced as /abaʊt/.

3

Table 4. Substitutions deemed acceptable between target English vowels and their Cantonese counterparts.

Target English Vowel	Replacement from Cantonese	Examples of mispronunciations
e	ei	"say" /se/ versus 四 /sɐi/
o	ou	"sew" /so/ versus 蘇 /sou/
ʌ	ɐ	"gut" /gʌt/ versus 吉 /gɐt/

Additionally, there are vowels that are present in both Cantonese and English, namely /ɪ, i, ʊ. u, ɔ/. Cantonese speakers tend to substitute the English vowels with their close L1 neighbors. Examples are shown in Table 5. The first two involve mispronunciations in terms of tense-lax confusions. We conjecture that a possible reason is due to phonotactic constraints, as will be explained in a later section.

Table 5. Substitutions among long and short vowels

Target English Vowel	Replacement from Cantonese	Examples of frequent mispronunciations
ɪ	i	"sit" /sɪt/ mispronounced as "seat" /sit/
ʊ	u	"full" /fʊl/ mispronounced as "fool" /ful/
ɔ	ɔ	"caught" /kɔt/ mispronounced and sounds like "cot" (as the Cantonese /ɔ/ has a shorter duration).

2.2 Consonants

Tables 2 and 3 shows the consonants in Cantonese and American English respectively, organized according to the manner and place of articulation. Comparison between the two tables helps structure our observations in common mispronunciations for Cantonese learners of English.   We refer specifically to English consonants that are *missing* from the Cantonese inventory, including voiced plosives, fricatives and affricates. Cantonese learners often substitute for these missing English consonants with Cantonese consonants that have similar place and/or manner of articulation. We present details in the following subsections.

2.2.1 *Missing Voiced Plosives*

The voiced plosives /b, d, g/ are present in English but absent in Cantonese. In the *prevocalic* position, these are often substituted with the voiceless, unaspirated Cantonese plosives /p, t, k/ which serve as good approximations. However, in the postvocalic position, voiced plosives may be unaspirated and voicing may be realized as durational lengthening of the preceding syllable nucleus. This leads to the durational difference within the contrastive word pairs: "cab" versus "cap" (/kæb/ versus /kæp/), or "pad" versus "pat" (/pæd/ vs /pæt/). These word pairs are often not clearly distinguished by Cantonese learners, e.g.:

- "feed" /fid/ is pronounced as "feet" /fit/
- "bag" /bæg/ is pronounced as "back" /bæk/, etc.

2.2.2 *Missing Affricates*

English affricates are post-alveolar and include unvoiced and voiced tokens, namely, /tʃ, dʒ/. These are non-existent in Cantonese and are often replaced respectively with the aspirated and unaspirated alveolar affricates /tsʰ, ts/. These have close resemblance in the place of articulation, e.g.

- "charge" /tʃ a r dʒ/ is often mispronounced like 义住 / tsʰ a ts y/, which a transliteration and not a word

4

In this example, the voiced pre-palatal affricate /dʒ/ in English occurs in a postvocalic position and is realized as the alveolar affricate /ts/ in Cantonese. However, since Cantonese affricates must be prevocalic due to phonotactic constraints, the mispronunciation adds a syllable nucleus at the word ending. Suitable candidates may be the palatal vowels /y, i/, which have matching tongue position during articulation. The former may be preferable as it involves lip rounding, which resembles lip protrusion during articulation of /dʒ/.

2.2.3 Missing Fricatives

This subsection addresses English fricatives that are missing from the Cantonese inventory. We describe the common substitutions performed by Cantonese learners of English.

	Bilabial	Labio-dental	Labio-velar	Dental	Alveolar	Pre-palatal	Palatal	Velar	Glottal
Plosive	p, pʰ				t, tʰ			k, kʰ kʷ, kʷʰ	
Affricate					ts, tsʰ				
Nasal	m				n			ŋ	
Fricative		f			s				h
Approximant	w						j	(w)	
Lateral Approximant					l				

based on [Zee, 1991]

Figure 3. Consonants in Cantonese, organized according to the manner and place of articulation.

	Bilabial	Labio-dental	Labio-velar	Dental	Alveolar	Pre-palatal	Palatal	Velar	Glottal
Plosive	p,pʰ,b				t,tʰ,d			k,kʰ,g	
Affricate						tʃ, dʒ			
Nasal	m				n			ŋ	
Fricative		f, v		θ, ð	s, z	ʃ, ʒ			h
Approximant	w				r		j	w	
Lateral Approximant					l				

based on [Ladefoged, 1999]

Figure 4. Consonants in American English.

The English /v/, a voiced, labiodental fricative, is often mispronounced either as the voiceless labiodental fricative /f/ or the sonorant bilabial approximant /w/. Examples include:
- "vast" /væst/ versus "fast" /fæst/
- "vest" /vɛst/ versus "west" /wɛst/.

There are two English dental fricatives. /θ/ is voiceless and is often mispronounced as the voiceless Cantonese labiodental /f/. /ð/ is voiced and is often mispronounced as the voiced alveolar plosive in Cantonese /t/. Examples include:
- "three" /θri/ versus "free" /fri/
- "there" /ðɛr/ versus "dare" /tɛr/

5

The English alveolar, voiced fricative /z/ is often mispronounced as the voiceless /s/, e.g.:

- "seize" /siz/ versus "sees" /sis/
- "zinc" /zɪŋk/ versus "sink" /sɪŋk/

The English unvoiced post-alveolar and voiced fricatives /ʃ ʒ / are frequently substituted with the voiceless alveolar /s/. Examples include:

- "show" /ʃ o/ versus "so" /so/;
- "social" /soʃ əl) versus "soso" /soʃ əl)

2.2.4 Missing Approximants

Articulation of the English approximant /r/ involves lip rounding and retroflexion. /r/ is absent from Cantonese and is often substituted with /w/ (rounded approximant) or /l/ (lateral approximant), e.g.:

- "rate" /ret/ versus "wait" /wet/
- "very" /vɛri/ as /vɛli/ or /wɛli/

2.2.5 Confusion among /n/ and /l/

This subsection is different from the others as it describes a specific confusion between the alveolar nasal /n/ and the lateral approximant /l/. In colloquial Cantonese, /n/ is often substituted with /l/. This substitution is generally deemed acceptable, e.g.

- 你 (you) /nei/ pronounced as 理 (logic) /lei/

A similar substitution frequently occurs for the prevocalic nasal /n/ in English words uttered by Cantonese learners. This substitution is perceived as a prominent mispronunciation, e.g.

- "nine" /n ay n/ as "line" /l ay n/

3. PHONOTACTIC COMPARISON

The Cantonese syllable has a simple [C]V[C] structure, where the optional syllable onset contains one consonant (except for /kw/ or /kʰw/) and the optional syllable coda also contains one consonant. Similarly, the syllable onset and coda are optional in the English syllable, but each may contain up to three consonants, such as in the word "strengths" /strɛŋθs/. Cantonese learners frequently mispronounce English consonant clusters, either with *vowel addition* in the word-final position, e.g.

- "kissed" /kɪst/ becomes /kɪstɐ/

or *consonant deletion*, e.g.

- "exact" /ɪkzækt/ becomes /ɪk-sɛk/
- "professor" /prəfɛsə/ becomes /pou-fɛ-sa/

Both processes attempt to generate syllables that are compatible the Cantonese syllable structure.

Cantonese phonotactic constraints [5] may also be a possible cause for other phonetic substitutions in mispronunciations made by Cantonese learners, e.g.

- "tone" /t o n/ pronounced as 通 (reach) /t ou ŋ/
- "bill" /bɪl/ pronounced as 標 (sign) /b iu/
- "in" /ɪn/ pronounced as 烟 (smoke) /jɪn/

The first example word "tone" contains the English vowel /o/, whose closest Cantonese counterpart may be diphthong /ou/. However, this diphthong typically occurs in an open syllable. We may substitute with the vowel /ʊ/ since it has similar quality, but if /ʊ/ should close with a nasalized coda, it must be the velar /ŋ/. These factors may cause "tone" to be mispronounced as /t ou ŋ/. In the second example, the word "bill" contains a postvocalic /l/. This consonant never closes a syllable in Cantonese but its manifestation bears similarity to lip rounding. Hence the Cantonese diphthong /iu/ may serve as a perceptually close substitute for /ɪl/ and cause "bill" to be mispronounced as /b iu/. In the third example, the word "in" contains the vowel /ɪ/. This vowel must be preceded with a syllable onset in Cantonese and insertion of /j/ offers the closest matching place of articulation. Hence "in" becomes /jin/.

4. PROSODIC COMPARISON

English and Cantonese have significant differences in prosodic features such as loudness, duration, and pitch. A major contributor to such differences is syllable stress. English syllables may be unstressed (reduced), or they may carry primary or secondary stress. For example, the word "aeroplane" has three syllables carrying primary stress, no stress (unstressed) and secondary stress respectively. Stressed syllables generally have higher energies, longer durations and higher pitch frequencies. The strong and weak stress patterns generate rhythm in English, to which we refer as a stress-timed language.

Cantonese syllables have approximately equal durations, i.e., it is a syllable-timed language. Stressed and unstressed syllables in English are often pronounced by Cantonese learners with largely uniform durations, but that stressed syllables often have higher tones. These tone frequencies are further influenced by the lexical (nine-)tone system that is characteristic of Cantonese. Hence, Cantonese learners tend to speak English with a different rhythm.

5. CONCLUSIONS

This paper describes an initial effort to derive a list of salient pronunciation errors made by Cantonese learners of English. The proposed approach is grounded on the theory of language transfer and involves phonological comparisons between Hong Kong Cantonese (L1) and American English (L2) across the phonetic, phonotactic and prosodic levels. We identify major disparities across the two languages, which are believed to heighten the perceived phonological interference of transfer features and cause mispronunciations. Systematic phonological comparisons across the various linguistic levels enable us to devise a methodology for deriving salient mispronunciations in second language acquisition, organize anecdotal examples observed from daily conversational interactions, as well as understand the cause of errors in order to design pedagogical and remedial instructions. Future work involves the design and collection of speech corpora in order to empirically verify the mispronunciations derived above. We will also generalize the proposed methodology to derive mispronunciations made by Mandarin (L1) learners of American English (L2).

6. REFERENCES

[1] Chan, A. & Li, D. 2000. "English and Cantonese Phonology in Contrast: Explaining Cantonese ESL Learner's English Pronunciation Problems," Language, Culture & Curriculum, 13:1.
[2] Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford University Press.
[3] Ho, R. 2005. 粵音自學提綱. 香港教育圖書公司.
[4] Ladefoged, P. 1999. "American English," *Handbook of the IPA*. Cambridge University Press.
[5] Lado, R. 1964. *Linguistics Across Cultures: Applied Linguistics for Langauge Teachers.* University of Michigan Press.
[6] Zee, E. 1991. Chinese (Hong Kong Cantonese), Journal of International Phonetic Association, 21:1.
[7] Zee, E. 2003. "Frequency analysis of the vowels in Cantonese from 50 male and 50 female speakers," Proc. ICPhS, Barcelona.
[8] Zue, V., Notes on Speech Spectrogram Reading, MIT Summer Course, 1991.

Appendix A. Cantonese vowels and diphthongs in the context of syllable pronunciations of Chinese characters. Meanings of the characters are in parentheses.

Four short vowels			Seven long vowels		
ɪ	/sɪk/	色 (color)	i	/si/	絲 (silk)
ɐ	/sɐp/	濕 (wet)	y	/sy/	書 (book)
ʊ	/sʊk/	叔 (uncle)	ɛ	/sɛ/	借 (borrow)
θ	/sθt/	恤 (shirt)	œ	/hœ/	靴 (boot)
			a	/sa/	沙 (sand)
			ɔ	/sɔ/	梳 (comb)
			u	/fu/	夫 (husband)
Ten diphthongs					
ai	/gai/	佳 (good)	θy	/sθy/	稅 (tax)
ɐi	/sɐi/	西 (west)	ɔi	/hɔi/	開 (open)
au	/gau/	交 (give)	ui	/fui/	灰 (gray)
ɐu	/sɐu/	收 (receive)	iu	/siu/	燒 (burn)
ei	/sei/	四 (four)	ou	/dou/	刀 (knife)

Appendix B. American English vowels and diphthongs in the context of an English word.

i	beat	o	boat
ɪ	bit	ʊ	book
e	bait	u	boot
ɛ	bet	ɚ	Burt
æ	bat	aɪ	bite
a	Bob	ɔɪ	Boyd
ɔ	bought	aʊ	bout
ʌ	but	ə	about