Project Title	**:**	"Speaking your language"- An investigation of Cantonese child directed speech by South Asian, immigrant Chinese, and local Hong Kong mothers
Grantee	**:**	The University of Hong Kong
Principal Investigator	**:**	TO Kit-sum, Carol Faculty of Education The University of Hong Kong
Co-investigators	**:**	YAO Yao Department of Chinese and Bilingual Studies The Hong Kong Polytechnic University Alan C L YU Department of Linguistics University of Chicago

Final Report

by

Principal Investigator

(a) Title

"Speaking Your Language" – An Investigation of Cantonese Child Directed Speech by South Asian, Immigrant Chinese, and Local Hong Kong Mothers

(b) Abstract

Hong Kong has many children born to families with South Asian (SA) heritage and immigrant families from the mainland. These children usually speak a minority language as their L1. However, mastering the majority language of Cantonese is a key to be integrated into the community. This project investigated the quantity and quality of Cantonese used by their mothers by constructing an annotated speech-corpus. Participants included mothers who speak native (1) Cantonese (HKC), (2) Putonghua (PTH) and (3) a SA language. Twenty-nine, 27, and 12 mother-child dyads were recruited for the 3 groups respectively. Language samples were collected in the laboratory where the mothers interacted with their child (child-directed speech, CDS) and the experimenter (adult-directed speech, ADS). In Study 1, we found that SA group showed significantly lower MLU and TTR while those of the PTH group were comparable to the HKC group. The limited amount of Cantonese from the SA mothers may be the major obscuring factor of Chinese learning of their children. Given that adequate amount of language input is the essence of language learning and literacy development, increasing access to Cantonese input to SA children by all means is an indispensable catalyst for their oral and written Chinese learning. Study 2 examined the phonological transfer patterns among the PTH mothers. Results revealed more errors in Cantonese phonemes with phonological features that do not occur in PTH. Those sounds were substituted by alike PTH phonemes. The patterns identified can assist teachers or speech-language-pathologists to differentiate language differences versus disorders.

(c) Keyword(s)

Child-directed speech, Cantonese, MLU, vocabulary diversity, language input, phonological transfer

(d) Introduction

In the past few decades, Hong Kong has become increasingly diversified demographically, culturally, and linguistically. Although HKC remains the dominant language in Hong Kong—spoken by more than 94% of the population according to the most recent by-census in 2016—there is a rapidly growing population of non-native Cantonese speakers, most of whom are non-Chinese speaking (NCS) ethnic minorities of South Asian descent or new immigrants from the mainland, who came to Hong Kong for family reunion. Children with non-native Cantonese speaking-parents often face linguistic challenges in local schools, as the knowledge of Cantonese/Chinese — especially spoken Cantonese—is extremely important for the students to thrive in the local education system.

In addition to language education, speech and language pathology is another field that is struggling to provide effective services for children from NCS and new immigrant families in Hong Kong. The growth trajectory of the speech and language skills of the bilingual children is quantitatively and qualitatively different from monolingual children (Li, Miller, Dodd, & Zhu, 2005; Lam & To, 2018). Despite the significant studies of language acquisition by Cantonese-English bilinguals in Hong Kong, there is a decided lacuna in terms of reliable information regarding how Cantonese, the majority language, is acquired in other types of linguistic minority households. Without the understanding of typical speech and language profiles of the bilingual individuals in these minority groups, teachers and speech-language pathologists always face considerable challenges in differentiating normal developmental patterns resulting from the speaker's diverse linguistic backgrounds and real language learning difficulties that need clinical concern. Consequently, children with an NCS or new immigrant background may not receive accurate diagnosis (i.e. either under- or over-

diagnosed)—let alone proper treatment—when it comes to speech and language pathological support.

While there exist various recent efforts to improve Cantonese proficiency in minority students in primary and secondary schools (e.g., the Chinese Language Curriculum Second Language Learning Framework launched by the Education Bureau; Education Bureau, 2015), little attention is paid to the early Cantonese learning of these bilingual children during the toddler years. Beyond the education setting, an equally important, but often overlooked, factor is the linguistic input at home. In this project, we aim to fill the gap by examining the quantity and quality of the Cantonese input in NCS and new immigrant households. Our main research questions are: How much Cantonese input do children receive in NCS and new immigrant families? What is the nature of the Cantonese spoken in these households in terms of mean length of utterance and vocabulary diversity, as compared to the Cantonese spoken in native Cantonese-speaking households?

(e) Review of literature of the project

After the establishment of the corpora, the current project particularly examined two issues related to the non-Chinese speaking (NCS) mothers and the newly immigrated mothers who learned Cantonese from their local context. The literature review part below mainly focused on these two issues.

Study 1. Importance of Language-Learning Environment for Non-Chinese speaking Children

In her landmark study of language rights of in ethnic minorities in education in Hong Kong, Carmichael (2009) also noted that, in order for NCS linguistic minority children to be able to thrive in the local education system in Hong Kong, it is of utmost importance that they are functional in Cantonese/Chinese, the main language of education in Hong Kong. In

particular, they must master not only the reading and writing of written Standard Chinese, but also spoken Cantonese. Failure to achieve a level of general competence in Chinese/Cantonese can severely limit their educational achievements and their ability to take part in the workforce of the Hong Kong society. In a recent study by Oxfam (2014), which surveyed 469 South Asian families with at least one child studying in kindergartens in Hong Kong, about three-fifths (58.0%) of the parents reported that they were illiterate in Chinese, although many of them (73.4%) claimed that they could speak Cantonese. It is unclear the nature and the extent of the Cantonese heard by their children at home. Regardless of the proficiency of their own Cantonese, many of the parents surveyed deemed that learning Cantonese/Chinese is paramount to their children's educational advancement, career prospect, and integration into the local community given the high-status of Cantonese in society. To bolster their Chinese proficiency levels, the Education Bureau has launched the Chinese Language Curriculum Second Language Learning Framework for non-Chinese speaking (NCS) ethnic minority students in primary and secondary schools in 2015 (Education Bureau, 2015). Although there is no outcome study examining the effectiveness of the curriculum, the measures appeared to be assistive to the NCS students to adapt to the local education system. However, such supportive measures only apply to NCS students from Primary One to Secondary Six. Younger NCS children (before school age) were not catered toward in the policy planning, and the role of parental involvement in nurturing Chinese/Cantonese proficiency has generally been ignored. The first objective was therefore to compare the amount of Cantonese input produced by local native Cantonese mothers, NCS mothers and new immigrant mothers.

Study 2. L2 (Cantonese) speech produced by PTH mothers

In this part, we briefly review the literature on non-native (L2) speech (as compared to

native, L1 speech). Non-native speakers produce speech sounds differently from native speakers. When a speaker has access to more than one sound system, interaction between the systems will be inevitable (Best & Tyler 2007, Flege 1995, 1998). According to the Speech and Learning Model (Flege 1995, 1998), L1 sounds and L2 sounds with similar phonetic features may be perceptually linked in a bilingual speaker's mind, which allows the representations of the sounds in one language to drift toward the representations of "similar" sounds in the other language. By way of the perceptual linkage, cross-language influence from L1 may lead to shifted production of sounds in L2, resulting in perceivable "non-native accents". A voluminous body of second language acquisition literature has been devoted to the documentation of L2 accents in all levels of phonetic representations (consonant, vowel, tone, and intonation, etc.) in different L1-L2 situations. To this end, previous research regarding Cantonese and L2 acquisition has mostly focused on the acquisition of L2 sounds (in English, Mandarin, etc.) by Cantonese-L1 speakers (e.g., Chen, Ng, & Li, 2012; Holm & Dodd, 1999; Lam & To, 2017; Law, 2006; Law & So, 2006; Li, To, & Ng, 2017); however, little is known about the phonetic features of non-native Cantonese speech, as produced new immigrant or NCS adults living in Hong Kong. Only recently did researchers begin to examine NCS speakers' production of Cantonese (Mok et al, 2018), with a focus on younger speakers. On the other hand, the status of new immigrant speakers' Cantonese production is still largely unknown. It should be noted that even within the new immigrant population, there is a high degree of variation and individual differences in terms of dialectal background, linguistic heritage, socioeconomic status, communicative patterns in the family and the community, etc., all of which could lead to different bilingual (multilingual) situations. In this proposed study, we focused on the speech among mothers who are speakers of Putonghua as their L1.

(f) Theoretical and/or conceptual framework of the project

Study 1. Language-learning environment

Much research evidence attested that a crucial factor that contributes to bilinguals' L1 and L2 growth is their language-learning environments (Bedore et al., 2016; Cheung et al., 2019; Paradis, 2016). During the preschool years, language-learning environments involve dynamic interactions between communicative partners (e.g., parent-child communication) in their daily life. A number of factors could affect such interactions and the amount of L1 and L2 used across minority families. In home environment, the use of L1 and L2 depends on family members' L1 and L2 proficiency, cultural beliefs, and community/ cultural influences (Burchinal et al., 2012, Lee et al., 2015, Luo & Wiseman, 2000; Velazquez, 2008). The first part of the study examined the quantity and quality of Cantonese CDS produced by new immigrant mothers and NCS mothers when they interacted with their preschool child, as compared with the native Cantonese mothers.

Study 2. Phonological Transfer

Nonstandard productions by L2 speakers are not random. Certain regularities may be explained by the influence of L1 and this influence is called phonological transfer. In other words, phonological transfer happens when L2 phonological acquisition is influenced by L1 phonological features. Multiple factors can affect the types and outcomes of phonological transfer, including the age of acquisition, the length of exposure, and L1 and L2 acoustic similarity (Baker & Trofimovich, 2005; Major, 2008). Weinreich (1953) put forward seven categories of phonological transfer, namely, (1) sound substitution (learners use the nearest L1 phoneme as an equivalent in L2 sound system; e.g., L1-Putonghua/L2-Cantonese speakers may use [a] to substitute Cantonese /ɐ/), (2) phonological processes (learners use the L1 allophonic variant which does not occur in the same phoneme of L2; e.g., L1-French/L2-English speakers may use clear [l] in word final position to substitute velarized /r/), (3) under-

differentiation (when the two phonemes in L2 are allophones in L1; e.g., L1-Spanish/L2-English speakers generalise English /d/ and /ð/ into [d]), (4) over-differentiation (when the two phonemes in L1 are allophones in L2; e.g., L1-English/L2-Spanish speakers over-differentiate Spanish /d/ into [d] and [ð]), (5) reinterpretation of distinctions (learners misinterpret concomitant features as distinctive features; e.g., L1-English/L2-German speakers misinterpret German long/short distinctions as English tense/lax distinctions), (6) phonotactic interference (learners reconstruct L2 syllable structure to conform to L1 phonological rules on syllable level; e.g., L1-Portuguese/L2-English may produce pic[i]nic[i] for picnic) and (7) prosodic interference. Subsequent to Weinreich, various theories were posited to explain and predict phonological transfer across languages. Lado (1957) proposed Contrastive Analysis (CA), which aims to predict potential errors in L2 acquisition by comparing language systems. CA can be applied in multiple domains, yet is more widely used in phonology (Bugarski, 1991; Richards, 1971; Ringbom, 1994). According to CA theory, when the phonological systems are dissimilar, learners find it difficult to categorize the new phonemes in L2, and thus errors are likely to be the phonological features unique in L2 (Moradi & Chen, 2018; Navehebrahim, 2012; Seddighi, 2012). Meanwhile, Oller and Ziahosseiny (1970) posited a different viewpoint in CA. They claimed that when the two languages are similar, the higher acoustic similarity creates minimal and unnoticed phonetic differences, resulting in persistent non-learning (Major, 2008; Oller & Ziahosseiny, 1970). However, this hypothesis was rooted in the investigation on the different English spelling performance between the L1-Roman-alphabet group and the L1-non-Roman-alphabet group, no solid empirical evidence regarding phonological transfer in speech has been reported. CA clearly explained how similar or dissimilar phonological systems would contribute to phonological transfer. The predicting power of the approach appears to be limited. Best

(1994) posited the Perceptual Assimilation Model (PAM) to further explain the importance of perceptual ability in learning L2 and its contribution to production ability and phonological transfer. Phonological categories in L1 are developed through perceptual learning from infancy, and phonological categories in L1 may influence ones' ability to discriminate and learn phonological categories in L2. L2 learners, therefore, tend to assimilate non-native phonemes to native phonological categories that are most similar to the non-native sounds (Best, 1994; Best, McRoberts & Goodell, 2001; Tyler et al., 2014). Best and colleagues further discussed different patterns of assimilation which are summarized in Table 1. Researchers proposed that if the L2 phoneme and its L1 counterpart are in Single-Category pattern, the discrimination performance would be poor (Escudero & Boersma, 2002; Sebastián-Gallés & Bosch, 2005). Conversely, if the L2 phoneme and its L1 counterpart are in Two-Category or Uncategorized-Categorized pattern, they would be distinguished at ease. The PAM framework has provided a fruitful explanation for different realisations of the L2 phonemes produced by L1 speakers based on the perceptual differences among the phonemes in the two systems. Based on the realisations of the L1-Putonghua/ L2-Cantonese speakers, the current study explored how well the PAM can explain the realisations.

Table 1

Different Assimilation Patterns in the PAM (Best et al., 2001; Tyler et al., 2014)

Pattern	Description
Single-Category (SC)	The L2 phoneme and its L1 counterpart are assimilated to the same L1 category and they fit equally well. The discrimination between the two is difficult.
Two-Category (TC)	The L2 phoneme and its L1 counterpart are assimilated to two distinct L1 categories. The discrimination between the two is easy.
Category Goodness (CG)	Although the L2 phoneme and its L1 counterpart are assimilated to the same L1 category, one fits better, i.e. with a higher goodness rating. The discrimination performance depends on the differences in the goodness ratings between the two phonemes.
Uncategorized-Categorized (UC)	The L2 phoneme could not be assimilated to the L1 category. The discrimination between the L2 phoneme and its L1

	counterpart is easy.
Uncategorized (UU)	Two L2 phonemes are not assimilated to any L1 categories. The discrimination performance depends on their auditory similarity, regardless L1 phonological system.

Phonological Systems of Cantonese and Putonghua

Both Cantonese (Yue) and Putonghua belong to the Chinese Languages. However, there are various differences in the two phonological systems making speakers of the two languages mutually unintelligible. The International Phonetic Alphabet (IPA) system for Cantonese used in To, Cheung and McLeod (2013) and the one for Putonghua used in Lee and Zee (2003) are adopted.

Segmental Differences. The basic syllabic structures of Cantonese and Putonghua are similar: (C)V(C). Cantonese has 19 initial consonants while Putonghua has 23 initial consonants (Table 2) (Lee & Zee, 2003; Zeng, 1994). Comparing the two systems, 13 initial consonants overlap, while six initial consonants are unique in Cantonese, and 10 initial consonants are unique to Putonghua. The manners of articulation shared between Cantonese and Putonghua include plosives, affricates, fricatives, and approximants. However, Putonghua contains fricatives and affricates in dental-alveolar and alveolo-palatal places of articulation, which Cantonese lacks. Cantonese and Putonghua share two final consonants, and the remaining four final consonants are only present in Cantonese (Zeng, 1994).

There are 11 monophthongs and 11 diphthongs in Cantonese. In combination with final consonants, there are a total of 53 possible rimes in Cantonese (Chen et al., 2004). For Putonghua, vowel combinations are more diverse, with six monophthongs, nine diphthongs and four triphthongs. However, due to the limited number of final consonants, i.e. only /n/ and /ŋ/, there are only 36 possible rime combinations in Putonghua (Chen et al., 2004). Other unique features of Putonghua are the rhotic vowel /ɚ/ and two syllabic approximants /ɹ̩/ and /ɻ̩/. Table 3 summarize the differences and similarities of the two vowel systems.

Table 2. Comparison between Cantonese and Putonghua Consonants

	Shared by Cantonese & Putonghua	Only in Cantonese	Only in Putonghua
Initial consonants	p-, p^h-, t-, t^{h-}, k-, k^{h-}	k^w-, k^{wh}-	
		ts-, ts^{h-}	t͡ʂ-, t͡ʂʰ-, t͡ɕ-, t͡ɕʰ-, t͡ʂ-, t͡ʂʰ-
	f-, s-	h-	ɕ-, ʂ-, x-
	m-, n-	ŋ-	
	l-, j-, w-		
			ɻ-
Final consonants		**-p, -t, -k**	
	-n, -ŋ	**-m**	

Table 3. Comparison between Cantonese and Putonghua Vowels

Shared by Cantonese & Putonghua	Only in Cantonese	Only in Putonghua
a, ai, au	**ɐ**, **ɐi**, **ɐu**	ja, wa, waɪ, jɑʊ
ou	**ɔ**, **ɔi**	wɔ, joʊ
ei	**œ**, **ɵ**, **ɵy**	ɤ
i	ɛ, ɛu	ə, jɛ, weɪ, ɥe
u	iu, ui	ɨ, ɻ̩
y	ʊ(k)	

Suprasegmental Differences

There are six lexical tones in Cantonese, which are high-level, high-rising, mid-level, low-falling, low-rising and low-level tone (Figure 1) (Bauer & Benedict, 2011) where there are four lexical tones in Putonghua, including high, rising, low, falling, and neutral (Figure 2) (Lee and Zee, 2003; Zhu & Dodd, 2000).

Figure 1

Cantonse Tone System

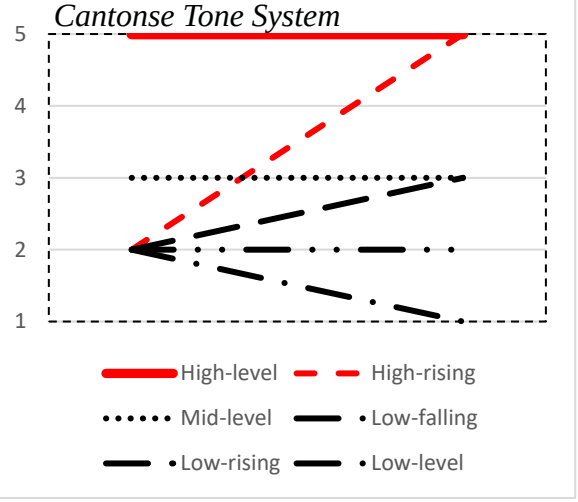
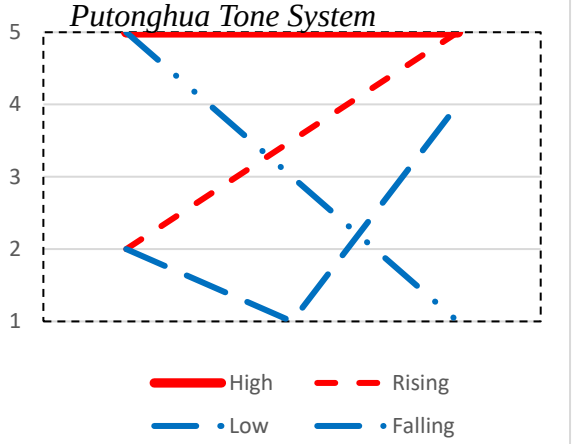


Figure 2

Putonghua Tone System



(g) METHODOLOGY

Participants

Mother speakers. We recruited three groups of mothers representing different language backgrounds, native in (1) HKC, (2) Putonghua, and (3) South-Asian languages and lived in Hong Kong during the time of the study. Participants were recruited through flyers, advertisements in website, and social media targeted for mothers living in Hong Kong with specified language backgrounds. For the second and the third groups, participants were also recruited via non-governmental organizations (NGOs) in Hong Kong. Written informed consents were obtained before the study. All mothers and/or fathers consented to release the audio recording and accompanying transcripts for the construction of a publicly accessible corpus. Each family was paid HK$800 (i.e., ~US$100) for the participation in the study.

All the mothers were interviewed verbally by Cantonese-English-Putonghua trilingual research assistants with the use of a written questionnaire in either Chinese or English regarding their language backgrounds and language use. All the three groups of mothers reported themselves as multilinguals with difference levels of proficiency.

Table 4. Demographic Background of the Mothers

Gp	Native languages	Dyad * (n)	Age	Boys : girls	Residence in HK (yrs)					Education		
					< 2	2-4	5-10	11-20	> 20	Sec.	Dipl.	Degree/ above
HKC	HKC	29		14: 15	0	0	0	1	26	3	5	19
PTH	Putonghua	27		13: 14	1	5	8	11	1	4	1	21
SA	Urdu, Punjabi, or Tamil	12		4: 8	0	0	1	1	9	5	0	6

* Four mothers did not provide demographic information (2 in Group of HKC 1, 1 in PTH Group and 3 in SA Group)

The first group consisted of 29 local mothers in Hong Kong who speak HKC since birth. All of them considered their Cantonese proficiency as native level and reported that Cantonese was used as the major language in their daily life (e.g., watching news on TV, talking to friends).

The second group was composed of 27 mothers who reported that Putonghua was their strongest language. All except three of them acquired Putonghua before the age of 5 years old. The three mothers reported that they acquire Putonghua after the preschool years and their first languages are languages of their home towns (湖南話，惠東方言，潮汕). All these mothers spoke Putonghua to their friends, colleagues or family members at home. All except seven had been staying in Hong Kong for 5 to 20 years. The remaining seven had been stayed in Hong Kong for two to four years. Their own ratings of the Cantonese proficiency ranged from very good to fair.

During participant recruitment, ten South Asian mothers who signed up was dropped out and did not take part in any of the data collection at HKU. This is because the mothers claimed that they did not know any Cantonese, or just at a minimum level which cannot even afford a simple conversation. These participants were contacted again after the outbreak of COVID-19. However, they declined to participate again. For the final sample, the third group

of mothers consisted of 12 South Asian mothers who spoke a South Asian language (Urdu 8/12, Punjabi 2/12, Tamil 2/12) as their first language. Nine of them had been staying in Hong Kong for more than 20 years while the three had been in Hong Kong for less than 20 years. The native languages of the South Asian mothers included, Urdu, Punjabi, or Tamil. Their own ratings of the Cantonese proficiency within the group varied substantially from highly proficient to poor.

Child speakers. Child participants were typically developing children at the age of 14 to 43 months old at the time of study. There were a total of 31 boys and 37 girls. Their mothers were their main caregiver. All except three were born full-term. None of the children have any diagnosed developmental disorders.

(h) Data collection and analysis

Speech samples were collected between 2019 and 2021 in Hong Kong. Each mother participated in two parts of recording for CDS and ADS. CDS of each dyad was collected individually at The University of Hong Kong. The mother was requested to interact with their child in a mock living room at The Faculty of Education as they would normally do at home for 45 minutes to 1 hour. Fathers of a few cases in the third group also came along. They were allowed to stay inside the mock living room but were reminded to wait quietly. Age-appropriate toys and books were provided for interaction. Speech samples produced by the mothers, along with the child's productions, were collected by a Zoom H5 digital recorder (holding in a shoulder bag carried by the mother) via a Sennheiser MKE 2 clip-on microphone clipped at mother's collar level.

ADS was collected within the same day or in a second visit to the University within 2-week time. The samples comprised both dialogues and monologues in HKC. Dialogues included face-to-face interviews with the experimenter which consisted of questions and

answers regarding child's developmental and social history, daily routine and mother's background, their job or their daily routine. Three mothers in the third group cannot carry out a dialogue solely in Cantonese. The conversation was then accompanied by English. Monologues were elicited via 4 tasks: a film description task, a map description task, a story retelling task and a single-word picture naming task. The monologue tasks provided a common basis for analysis while the dialogues allowed more rooms for improvisation.

CORPUS DEVELOPMENT

Transcription and forced alignment

First, within each recording, the temporal boundaries of each utterance were demarcated using the software Phon (Hedlund & Rose, 2020) by a team of trained, native Cantonese-speaking student research assistants who were proficient at transcribing Cantonese in Chinese and romanised script. Each utterance was transcribed orthographically using written conventions for Cantonese (in traditional Chinese characters), with word items that lack a common standardised form being represented in Jyutping (粵拼) phonetic romanisation. A team of research assistants who had extensive experience in transcription conducted a first-pass verification to assure accuracy of all the transcriptions. Where needed, novel lexical items were appended to the word- and/or syllable-based reference dictionaries utilized by SPPAS in Cantonese forced alignment.

Then orthographic annotations were submitted to SPPAS, which parsed utterances into both words and syllables. For each identified word and syllable, candidate phonetic forms were supplied from the reference dictionaries at this stage. Segmental units and their boundaries were then fit to each word-level annotation according to the Cantonese acoustic model in Lee et al. (2002), as implemented by SPPAS which is an open-source Python-based

software package (Bigi, 2015, 2018). Given the set of known phonetic candidate forms for a given transcription, SPPAS will identify the best segmental representation with respect to the phonetic properties of phones and phonemes contained within in the Cantonese acoustic model. While the Cantonese acoustic model performs relatively well in identifying actual phonetic segments such as consonants and vowels produced during audio recordings, it cannot identify Cantonese tonal units in its SPPAS implementation. Using a separate Cantonese reference dictionary, canonical orthographic (Chinese script) and phonetic (Jyutping and IPA) representations of each word and syllable were then assigned to each annotation so that actual phonetic transcriptions from the forced alignment procedure could be compared against their corresponding dictionary citation forms. Finally, the annotations at the utterance, lexical, syllabic, and segmental levels were combined as separate annotation tiers using PRAAT software (Boersma & Weenink, 2020).

Subsequent to the forced alignment process, acoustic data such as temporal onset and offset times and acoustic/spectral formant data were extracted using custom-made code in Praat, and ultimately, data on all phonetic segments, along with their acoustic properties, were combined using coding in R statistical computing software (R Core Team, 2020) for later analysis.

Coding and Tagging

Each transcript consisted of 19 tiers in the PRAAT file as shown in Table 5.

Table 5. Coding Used in Each of the PRAAT file

1.	WORD-Ortho-1	The orthographic transcription (in traditional Chinese characters) for each word produced by Talker 1 (mother).
2.	WORD-Jyutping-1	The Jyutping romanisation for the citation form(s) of each word produced by Talker 1 (mother).
3.	WORD-CitIPA-1	The citation phonetic form(s) (in IPA) of each word produced by Talker 1 (mother).
4.	WORD-ActIPA-	The phonetic form (in IPA) for each word that was actually

	1	produced by Talker 1 (mother).
5.	SYLL-Ortho-1	The orthographic transcription (in traditional Chinese characters) of each syllable produced by Talker 1 (mother).
6.	SYLL-Jyutping-1	The Jyutping romanisation for the citation form(s) of each syllable produced by Talker 1 (mother).
7.	SYLL-CitIPA-1	The citation phonetic form(s) (in IPA) of each syllable produced by Talker 1 (mother).
8.	SYLL-ActIPA-1	The phonetic form (in IPA) for each syllable that was actually produced by Talker 1 (mother).
9.	PHON-Phoneme-1	The actual phonetic segments (in IPA) produced by Talker 1 (mother).
10.	PHON-Phoneme-PreCor-1	The phonetic segments (in IPA) produced by Talker 1 (mother) as determined by a first-pass forced-alignment analysis procedure (Groups 2 and 3 only). Any post-hoc corrections to this tier are reflected in the tier PHON-Phoneme-1.
11.	WORD-Ortho-2	The orthographic transcription (in Traditional Chinese characters) for each word produced by Talker 2 (child).
12.	WORD-Jyutping-2	The Jyutping romanisation for the citation form(s) of each word produced by Talker 2 (child).
13.	WORD-CitIPA-2	The citation phonetic form(s) (in IPA) of each word produced by Talker 2 (child).
14.	WORD-ActIPA-2	The phonetic form (in IPA) for each word that was actually produced by Talker 2 (child).
15.	SYLL-Ortho-2	The orthographic transcription (in Traditional Chinese characters) of each syllable produced by Talker 2 (child).
16.	SYLL-Jyutping-2	The Jyutping romanisation for the citation form(s) of each syllable produced by Talker 2 (child).
17.	SYLL-CitIPA-2	The citation phonetic form(s) (in IPA) of each syllable produced by Talker 2 (child).
18.	SYLL-ActIPA-2	The phonetic form (in IPA) for each syllable that was actually produced by Talker 2 (child).
19.	PHON-Phoneme-2	The actual phonetic segments (in IPA) produced by Talker 2 (child).

It is noted that tiers that contain citation phonetic forms in Jyutping or IPA, paired curly braces ({ }) encompassing two or more items that are separated by a vertical bar (|) indicate the set of all possible citation forms that correspond to a particular item (word or

syllable), e.g.,

(Ortho) – (Jyutping) – (CitIPA)
蛋 – {daan2|daan6} – taːn
牛奶 – {au4.laai5|ngau4.naai5} – {ɐu.laːi|ŋɐu.naːi}.
呢 – {li1|nei1|nei4|ne1|ni1} – {li|nei|nɛ|ni}
其實 – kei4.sat9 – kʰei.sɐt (word-based analysis)
其 – {gei1|kei4} – {kei|kʰei} (syllable-based analysis)

The annotated files were then formatted according to the Codes for the Human Analysis of Transcripts (CHAT, MacWhinney, 2000) for morphological tagging. We follow the major parts of speech and convention used in the Cantonese Aphasia Bank (Kong & Law, 2019) which is a corpus of conversational speech produced by Cantonese-speaking aphasia speakers. To ensure anonymity, we replaced child participants' names with CHD and silenced out whole utterance with CHD, using SILENCE command of CLAN. After the morphological tagging, the files were further processed to automatically disambiguate morphemes that possess more than one parts of speech or meanings.

(i) Results and Discussion

By the end of August 2021, speech samples were analysed from 29 local mothers, 18 PTH mothers and 12 SA mothers. In total, the corpus contained 86 hours and 14 minutes of interaction. Mothers' production in the CDS and ADS consisted of 211,567 word tokens in 44,354 utterances and 166,024 tokens in 24,084 utterances respectively. The corpus is uploaded onto the website hosted in the HKU server https://ccds.edu.hku.hk/. We will make this website available to the general public. Table 6 summarizes the content of the corpus.

Table 6: Details of the Corpus

Group	Native languages	Dyad (n)	Genre	Speaker	Utterances	Word token	Duration (hrs: mins)
HKC	Cantonese	29	CDS	Mothers	23,244	118,816	28 : 29
				Children	7,176	14,231	
			ADS	Mothers	11,466	82,900	13 : 18

PTH	Putonghua	18	CDS	Mothers	14,361	64,879	15 : 57
				Children	4,345	10,156	
			ADS	Mothers	7,216	52,337	10 : 10
3	Urdu, Punjabi, Sindhi or Tamil	12	CDS	Mothers	6,749	27,872	10 : 22
				Children	2,870	8,617	
			ADS	Mothers	5,402	30,787	7 : 58

Study 1

Regarding the quantity and quality of Cantonese input, the analyses is mainly based on the CDS. The measures of (1) number of word tokens, (2) number of types, and (3) type-token ratio by the type of corpus (CDS vs. ADS) and group (CAN, PTH vs SA), were calculated by using the "freq" command in CLAN. ANOVA analyses revealed significant differences in all the measures: (1) number of word token [$F(2, 55)=7.881$, $p=.001$], (2) number of word types [$F(2, 55)=17.187$, $p<.001$], and (3) Type-token ratio (TTR) [$F(2, 55)=3.773$, $p=.029$]. Post-hoc analyses using Sheffe Test indicated that the SA group showed significantly lower word token and word type than the PTH and HKC groups while the two latter groups were comparable. A closer look indicated that TTR was the highest in the SA group. However, it should be noted that the high TTR value maybe inflated by the small value of word type. Therefore, the measure of D was computed. D is developed by Malvern and Richards (1997) as an index that measures lexical diversity through a process of curve fitting. It has been argued that D is more informative than TTR, because as opposed to the single value of TTR, it represents how TTR varies over a range of token sizes from each speaker. The measure D also has two advantages: (1) because it is not a function of the number of words in the language samples, so it is not necessary to standardize the sample length; and (2) it works well with short text (50 tokens are the minimal requirement), which is especially relevant when working with less proficient speakers [$F(2, 55)=9.565$, $p<.001$]. Post-hoc analyses indicated significant difference between the SA group and HKC group but

not other group comparisons.

Utterance-Level Characteristics

Mean length of Utterance (MLU) is an index of syntactic complexity. The higher the value of MLU, the more complex the syntactic structures produced by the mother. ANOVA results indicated that SA group produced significantly shorter MLU than the HKC group while the PTH group was similar to the HKC group.

Table 7. Lexical and Utterance-Level Characteristics of the CDS

Measures	Group	Mean	SD	Min.	Max.
Type	HKC	460.83	89.916	311	660
	PTH	405.63	116.973	215	625
	SA	227.09	155.447	14	556
Token	HKC	4148.55	1297.602	2459	7406
	PTH	3851.75	1515.962	1192	6060
	SA	2156.27	1650.431	26	5024
Type Token Ratio (TTR)	HKC	0.11559	0.018734	0.084	0.16
	PTH	0.1155	0.031801	0.08	0.18
	SA	0.17745	0.145893	0.056	0.538
D	HKC	64.8914	12.47	42.26	97.3
	PTH	51.8587	8.82813	37.14	62.76
	SA	43.841	24.10885	10.67	96.55
MLU	HKC	5.205	0.944333	3.881	7.173
	PTH	4.439	0.716536	2.588	5.409
	SA	3.665	1.346738	1.563	6.331

Study 2: Phonological Transfer in the PTH Group

Ten participants in the PTH group were selected as these mothers represent bilingual with PTH dominant. They were born and raised in mainland China. They started learning Cantonese after the age of 18 who have lived in Hong Kong fewer than 20 years. Ten matched participants from the HKC group were selected and they have been living in Hong Kong for more than 20 years (see Table 8).

Table 8. Demographic Information of Selected Participants from HKC and PTH groups

Subject	Group	Age	Hometown	Years of living in Hong Kong	Age of Acquisition (Cantonese)
a103	1	30 - 39	Hong Kong	>20	Since birth
a104	1	30 - 39	Hong Kong	>20	Since birth
a105	1	30 - 39	Hong Kong	>20	Since birth
a106	1	30 - 39	Hong Kong	>20	Since birth
a107	1	30 - 39	Hong Kong	>20	Since birth
a109	1	Over 40	Hong Kong	>20	Since birth
a110	1	30 - 39	Hong Kong	>20	Since birth
a113	1	30 - 39	Hong Kong	>20	Since birth
a114	1	30 - 39	Hong Kong	>20	Since birth
a115	1	30 - 39	Hong Kong	>20	Since birth
a202	2	30 - 39	Hunan	5-10	After 18
a203	2	30 - 39	Beijing	5-10	After 18
a204	2	30 - 39	Hunan	11 - 20	After 18
a209	2	Over 40	Hunan	<2	After 18
a213	2	30 - 39	Beijing	11 - 20	After 18
a214	2	25 - 29	Wenzhou	5-10	After 18
a215	2	30 - 39	Sichuan	11 - 20	After 18
a216	2	30 - 39	Henan	< 2	After 18
a217	2	Over 40	An'hui	5-10	After 18
a218	2	30 - 39	Hubei	11-20	After 18

Target Sound Selection

Based on the similarities of the phonemes in the two systems, four sets of target phonemes were selected for examination (Table 9): Target Set 1 included vowels that only exist in Cantonese but do not occur in Putonghua (i.e., /ɐ, ɐi, ɐu, ɔ, ɔi, œ, ɵ, ɵy/); Target Set 2 included Cantonese consonants (i.e., /ts, ts^h/) with only one distinctive feature different to the counterparts in Putonghua (i.e. /t͡s̪, t͡s̪^{h}/ or /t͡ɕ, t͡ɕ^h/); Target Set 3 included Cantonese consonants that also exist in Putonghua but in a different syllable-position (i.e., /-m, -p, -t, -k/); and Target 4 included consonants that occur in both Cantonese and Putonghua and in the same syllable-position (i.e., /f, s, m/). A total of 16 target phonemes were selected and the frequency of occurrence of these phonemes was summarized in Table 10.

Table 9. Target Phonemes and Their Presence in the Two Systems

Target Set	Target Phonemes	Cantonese	Putonghua
1	/ɐ, ɐi, ɐu, ɔ, ɔi, œ, ɵ, ɵy/	+	-

2	/ts, ts^h/	+	-
3	/-m, -p, -t, -k/	+	-
4	/f, s, m/	+	+

Note. '+' indicates the presence of the phonemes in the language while '-' indicates the absence.

Table 10. Frequency of Occurrence of Target Phonemes in the Word Reading Aloud Task

Target Set	Phoneme	Token
1	ɐ	18
	ɐi	7
	ɐu	11
	ɔ	20
	ɔi	1
	ɵ	3
	ɵy	6
	œ	6

Target Set	Phoneme	Token
2	ts/ts^h	35
3	-k	12
	-t	8
	-m	6
	-p	3
4	s	24
	m	3
	f	8

The production accuracy between the two groups for each of the target phonemes is summarized in Figure 3. The phonetic production by the two group were compared as in Table 11. The comparison between predicted errors and actual errors is shown in Table 12. Error patterns with significant results were analysed with reference to the PAM, the results are listed in Tables 13 (Target Set 1) and 14 (Target Sets 2 to 4).

Target Set 1. The accuracy across all Target Set 1 phonemes was almost perfect in HKC 1 (ranged from 98.5% to 100%) while the accuracy ranged from 47.14% to 85.2% in PTH. The accuracy of /ɐ, ɐi, ɐu/ was between 47.14% and 78.73% in PTH group which is significantly lower than the HKC group. The phoneme /ɐ/ presented with four error realisations ([ə, a, i, ai]), while /ɐi/ and /ɐu/ presented with three ([ai, ei, ui]) and two realisations ([ou, iu]) respectively. The most frequent phonological transfer patterns presented were /ɐ/→[ə] (*M*=9.6%), /ɐi/→[ai] (*M*=22.3%) and /ɐu/→[ou] (*M*=19.8%).

The accuracy of /ɔ, ɔi/ was between 50% and 85.2% in PTH group which was significantly lower than the HKC group. The phoneme /ɔ/ presented with three error

realisations ([a, ʊ, ə]) while /ɔi/ presented with one ([ai]). The most frequent phonological transfer pattern presented in each phoneme were /ɔ/→[a] (*M*=13.4%) and /ɔi/→[ai] (*M*=50%).

The accuracy of [ɵ, ɵy] was between 63.3% and 76.7% in PTH group which was also significantly lower than the HKC group. The phoneme /ɵ/ was produced as two sounds, [ə, wə] while /ɵy/ was produced as [weɪ]. The most frequent phonological transfer patterns were /ɵ/→[ə] (*M*=10%), /ɵ/→[wə] (*M*=10%) and /ɵy/→[weɪ] (*M*=36.7%). The accuracy of /œ/ was 85.0% being similar to HKC group and three erroneous realisations were [ɔ, y, a].

Target Set 2. The accuracy of /ts, ts^h/ in HKC (*M*= 98.3%) was significantly higher than the PTH Group (*M*= 85.8%). Error realisations included [tɕ/$tɕ^h$, w]. The most frequent phonological transfer pattern in the PTH group was /ts, ts^h/→[tɕ, $tɕ^h$] (*M*=12.7%).

Target Set 3. The accuracy across all Target Set 3 phonemes was almost perfect in the HKC group (ranged from 95.6% to 100%). The accuracy of /-p, -t, -k / in PTH group (*M*= 56.7%, 64.2%, 76.7%) was significantly lower than HKC group respectively. Deletion with /-p, -t, -k / was the error pattern found being statistically significant (*M*=43.3%, 34.7%, 16.7%).

PTH group showed a very high error rate in the production of /-m/ with a mean accuracy of 25%. Three error patterns were found, including final consonant deletion and substitution with /-n/ or /-ŋ/. The most frequent phonological transfer pattern in PTH Group was /-m/→[-n] (*M*=70.0%).

Target Set 4. The accuracy of /s/ was 100% and 79.4% in the HKC and PTH groups respectively and the difference was significant. The only phonological transfer pattern in PTH group was /s/→[ɛ] (*M*=20.6%). For /m, f/, both groups achieved 100% production accuracy.

Table 11. Mean Percentage of the Target Phonemes in HKC and PTH Groups

Target Set	Target	Realisation	HKC Group		PTH Group		*F* value	*p*-value
			M	*SD*	*M*	*SD*		

		[ɐ]	**99.5%**	**0.02**	**78.7%**	**0.11**	**32.72*** **	**< .001**
		[ə]	0.5%	0.02	9.6%	0.08	14.01***	.001
	/ɐ/	[a]	0.0%	0.00	7.9%	0.07	13.82**	.002
		[i]	0.0%	0.00	3.3%	0.05	5.11*	.036
		[ai]	0.0%	0.00	0.6%	0.02	1.00	.331
		[ɐi]	**100.0%**	**0.00**	**47.1%**	**0.28**	**35.03*** **	**< .001**
	/ɐi/	[ai]	0.0%	0.00	22.3%	0.18	15.10***	.001
		[ei]	0.0%	0.00	21.8%	0.18	14.95***	.001
		[ui]	0.0%	0.00	8.8%	0.06	20.38***	< .001
		[ɐu]	**100.0%**	**0.00**	**76.9%**	**0.21**	**12.16** **	**.003**
	/ɐu/	[ou]	0.0%	0.00	19.8%	0.20	9.96**	.005
		[iu]	0.0%	0.00	0.8%	0.03	1.00	.331
		[ɔ]	**98.5%**	**0.02**	**85.2%**	**0.14**	**9.19** **	**.007**
1	/ɔ/	[a]	0.0%	0.00	13.4%	0.14	9.76**	.006
		[ʊ]	0.0%	0.00	1.0%	0.02	2.25	.151
		[ə]	1.5%	0.02	0.5%	0.02	1.20	.288
		[ɔi]	**100.0%**	**0.00**	**50.0%**	**0.53**	**9.00** **	**.008**
	/ɔi/	[ai]	0.0%	0.00	50.0%	0.53	9.00**	.008
		[ɵ]	**100.0%**	**0.00**	**76.7%**	**0.22**	**10.76** **	**.004**
	/ɵ/	[ə]	0.0%	0.00	10.0%	0.16	3.86	.065
		[wə]	0.0%	0.00	10.0%	0.16	3.86	.065
		[ɵy]	**100.0%**	**0.00**	**63.3%**	**0.38**	**9.15** **	**.007**
	/ɵy/	[weɪ]	0.0%	0.00	36.7%	0.38	9.15**	.007
		[œ]	**100.0%**	**0.00**	**85.0%**	**0.32**	**2.22**	**.154**
	/œ/	[ɔ]	0.0%	0.00	5.00%	0.22	1.98	.177
		[y]	0.0%	0.00	5.00%	0.07	2.25	.151
		[a]	0.0%	0.00	5.00%	0.05	1.00	.331
		[ts, ts^{h}]	**98.3%**	**0.02**	**85.8%**	**0.08**	**23.30*** **	**< .001**
2	/ts, ts^{h}/	[tɕ, tɕ^{h}]	1.7%	0.02	12.7%	0.07	20.31***	< .001
		[w]	0.0%	0.00	0.8%	0.01	3.85	.065
		[-k]	**95.8%**	**0.04**	**76.7%**	**0.21**	**8.18** **	**.010**
3	/-k/	∅	0.0%	0.00	16.7%	0.19	7.50*	.013

		[-t]	4.2%	0.04	6.7%	0.04	1.94	.180
	/-t/	**[-t]**	**95.6%**	**0.06**	**64.2%**	**0.16**	**33.48*** **	**< .001**
		∅	0%	0.00	34.7%	0.17	43.07***	< .001
		[-k]	4.4%	0.06	1.1%	0.04	2.46	.135
	/-p/	**[-p]**	**100.0%**	**0.00**	**56.7%**	**0.39**	**12.57** **	**.002**
		∅	0.0%	0.00	43.3%	0.39	12.57**	.002
	/-m/	**[-m]**	**100.0%**	**0.00**	**25.0%**	**0.20**	**145.80*** **	**< .001**
		[-n]	0.0%	0.00	70.0%	0.15	208.90***	< .001
		[ŋ]	0.0%	0.00	3.3%	0.07	2.25	.151
		∅	0.0%	0.00	1.7%	0.05	1.00	.331
4	/s/	**[s]**	**100.0%**	**0.00**	**79.4%**	**0.15**	**19.76*** **	**< .001**
		[ɕ]	**0.0%**	**0.00**	**20.6%**	**0.15**	**19.76*** **	**< .001**
	/m/	**[m]**	**100.0%**	**0.00**	**100.0%**	**0.00**	**-**	**-**
	/f/	**[f]**	**100.0%**	**0.00**	**100.0%**	**0.00**		**--**

Note. *p< .05. **p< .01. ***p< .001.

Table 12. Comparison between Prediction Error and Actual Error

Target Set	Target	Predicted error	Actual error
1	/ɐ/	[a]	[a, **ə**]
	/ɐi/	[ai]	[ai, **ei, ui**]
	/ɐu/	[au]	[ɐu, **ou**]
	/ɔ/	[wɔ]	[**a**]
	/ɔi/	[wɔi]	[**ai**]
	/œ/	[ə]	-
	/ɵ/	[ə]	-
	/ɵy/	[ei]	[**weɪ**]
2	/ts, tsʰ/	[t͡ɕ, t͡ɕʰ]	[t͡ɕ, t͡ɕʰ]
3	/-p, -t, -k/	[-p, -t, -k]	deletion
	/-m/	[-m]	[-n]
4	/f/	[f]	-
	/s/	[s]	[**ɕ**]
	/m/	[m]	-

Note. Bold and blue fonts refer to error patterns that are not predicted by the SC pattern in the PAM.

Table 13. Phonological Transfer Pattern Analysis for Target Set 1

	PAM Pattern	Phoneme realised		Tongue height	Tongue position	Roundness	Length
Target		/ɐ/		low	central	unrounded	short
Realisation 1	TC	[ə]		**mid**	central	unrounded	short
Realisation 2	SC	[a]		low	central	unrounded	**long**
Realisation 3	TC	[i]		**high**	**front**	unrounded	**long**
Target		/ɐi/	V1	low	central	unrounded	short
Realisation 1	SC	[ai]	V1	low	central	unrounded	**long**
Realisation 2	TC	[ei]	V1	**mid**	central	unrounded	**long**
Realisation 3	TC	[ui]	V1	**high**	**back**	**rounded**	**long**
Target		/ɐu/	V1	low	central	unrounded	short
Realisation 1	TC	[ou]	V1	**mid**	**back**	**rounded**	**long**
Target		/ɔ/		mid	back	rounded	long
Realisation 1	UC	[a]		**low**	**central**	**unrounded**	long
Target		/ɔi/	V1	mid	back	rounded	long
Realisation 1	UC	[ai]	V1	**low**	**central**	**unrounded**	long
Target	TC	/ɵy/	V1	low	central	unrounded	short
			V2	high	back	rounded	long
Realisation 1		[weɪ]	V0	high	back	rounded	long
			V1	**mid**	**central**	**unrounded**	long
			V2	high	**front**	**unrounded**	**short**

Note. V0, V1, V2 refer to the first, second and third vowel in diphthongs or triphthongs. Bold distinctive features are the ones different from the distinctive features of the target phonemes.

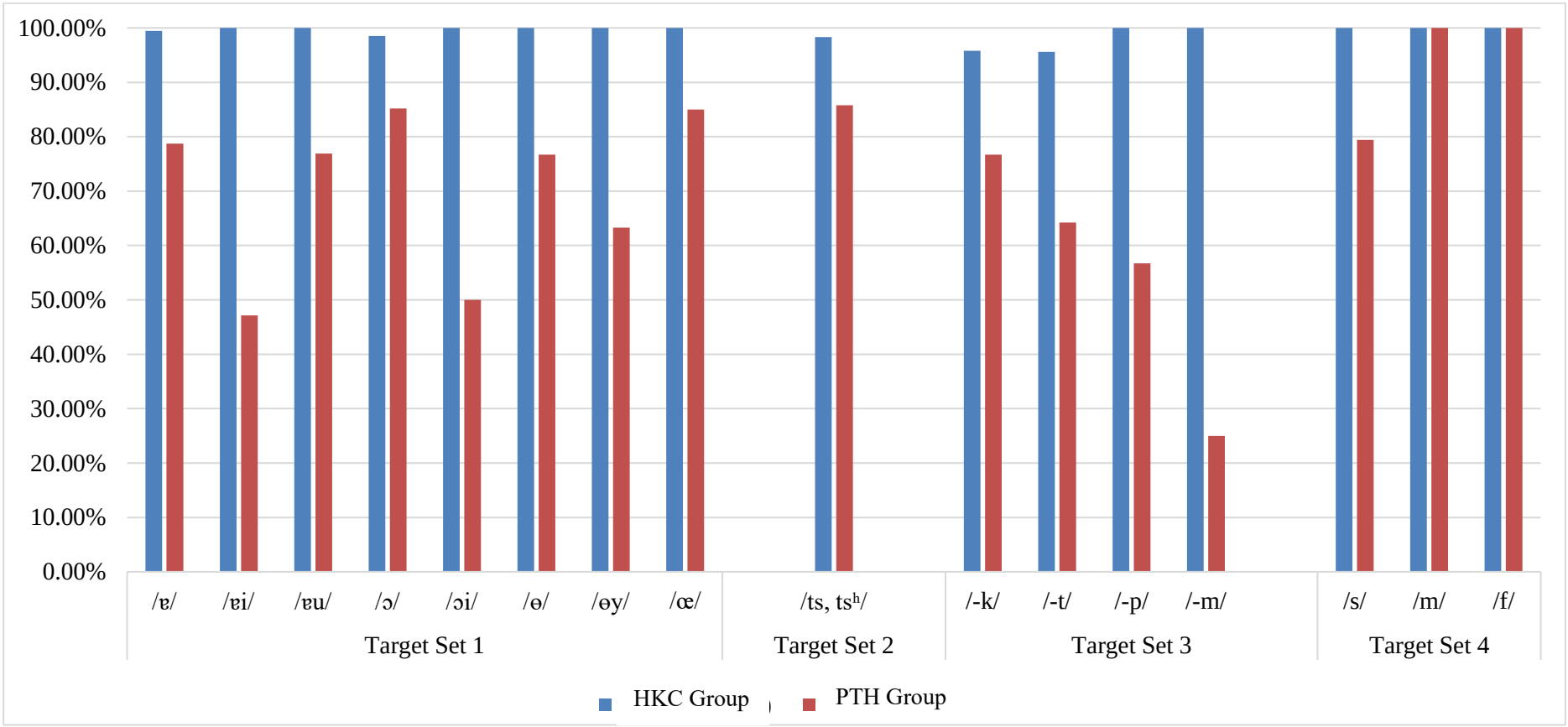
Table 14 Phonological Transfer Pattern Analysis for Target Sets 2, 3 and 4

Target Set		PAM	Phoneme	Distinctive Features
2	Target		/ts, ts^{h}/	alveolar affricate
	Realisation	SC	[tɕ, tɕ^{h}]	**alveolo-palatal** affricate
3	Target		/-k/	final velar stop
	Realisation	UC	∅	deletion
	Target		/-t/	final alveolar stop
	Realisation	UC	∅	deletion

	Target		/-p/	final bilabial stop
	Realisation	UC	∅	deletion
	Target		/-m/	final bilabial nasal
	Realisation	TC	[-n]	final alveolar nasal
4	Target		/s/	alveolar fricative
	Realisation	SC	[ɕ]	alveolo-palatal fricative

Figure 3

Comparison of Phonetic Accuracy between HKC and PTH Groups



In **Target Set 1**, realisations belonging to SC patterns as predicted (/ɐ/→[a] and /ɐi/→[ai]), UC patterns (/ɔ/→[a], /ɔi/→[ai]) and TC patterns (/ɐ/→[ə], /ɐi/→[ei]) were observed, which did not fully echo with the study that SC pattern was reported in difficult discrimination and poor production accuracy (Best et al., 2001; Tyler et al., 2014). The fact that Group 2 participants also demonstrated difficulties in producing TC and UC patterns indicates that early acquisition of dissimilar phonemes is challenging without possible direct transfer from acquired phonemes, according to Flege's Speech Learning Model (1995).

The pattern of **Target Set 2** targets (/ts/→[t͡ɕ], /ts^{h}/→[t͡ɕ^{h}]) belongs to SC, in which these two pairs are with identical manner and minimal differences in placement (i.e. alveolar affricates /ts, ts^{h}/ vs alveolo-palatal affricates /t͡ɕ, t͡ɕ^{h}/). The high acoustic similarities between these two pairs may contribute to the poor discrimination by L1-Putonghua/L2-Cantonese speakers and phonological transfer resulted (Best, 1994; 2001; Oller and Ziahosseiny, 1970). Surprisingly, the transfer pattern from Cantonese alveolar affricates /ts, ts^{h}/ to Putonghua dental-alveolar affricates /t͡s̪, t͡s̪^{h}/ was not observed, even though they appear to contrast minimally in theory. Conceivably, the perceptual and production differences between alveolo-palatal affricates (/t͡ɕ, t͡ɕ^{h}/) and alveolar affricates (/ts, ts^{h}/) for speakers would be closer than that with /t͡s̪, t͡s̪^{h}/, and hence the former pattern is more likely to happen.

Deletion and substitution (/-p/→∅, /-t/→∅, /-k/→∅, /-m/→[-n]) in **Target Set 3** were interpreted as UC patterns. When phonemes in L2 fall into different categories in L1, Putonghua speakers would prefer to drop the phoneme at the final position altogether, possibly because the constraints of phonotactic rules in L1 may override the changes on the segmental level (Weinreich, 1953). In addition, the unreleased nature of final stops in Cantonese may result in the subtle acoustic cues in perception. The negligence of unreleased final stops, as observed in merging of /-k/ and /-t/ in both Group 1 and 2, for example 一 (/jɐt/→[jɐk]), 突 (/tɐt/→[tɐk]) and 隻 (/tsɛk/→[tsɛt]), may be a result of high perceptual similarities between alveolar and velar final stops reported by Khouw and Ciocca (2006).

In **Target Set 4** phonological transfer pattern was found in /s/→[ɕ] unexpectedly, while production of /m/ and /f/ achieved 100% accurate as predicted. Given the pattern /s/→[ɕ] occurred in items with retroflex fricative /ʂ/ in Putonghua (e.g. 水 /ʂweɪ/, 書 /ʂu/ and 上 /ʂaŋ/), it is hypothesised that when

Group 2 participants articulated items with /s/ in Cantonese which has a corresponding /ʂ/ pronunciation in Putonghua, they tended to assimilate /ʂ/ to /ɕ/, given that /ɕ/ is more similar to /s/ in articulation placement and perception.

(j) Conclusions and Recommendations

Study 1 explored the quantity and quality of the Cantonese input provided by the three groups of mothers. Regardless of the phonological accuracy of the production, the CDS produced by the SA group was significantly lower than the other two groups in terms of total amount of words, vocabulary diversity, and syntactic complexity. One major distinction between the SA and the PTH groups is the nature of their L1s. Putonghua and Cantonese are related in a number of ways as they belong to the same language family. Speakers of the two languages share a similar ideographic writing system (one is traditional script and the other is the simplified script) and culture. Discrepancies exist in the phonological systems of the two Chinese languages making the speakers not mutually and completely intelligible to each other. However, the similarities between their lexicon and syntactic structures can also support the PTH mothers in the L2 Cantonese learning. The language (Cantonese) input provided by the PTH mothers was generally therefore not significantly different from the HKC mothers, implying that the amount would be of sufficient amount to support their children's Cantonese learning and Chinese learning in the long run. However, the native languages of the SA mothers, including Urdu, Punjabi and Tamil, is linguistically more distant from Cantonese when compare with the distance between Cantonese and Putonghua. Along with the cultural differences, Cantonese is therefore even more difficult to learn for these L1s. It was therefore not surprising that their Cantonese input provided by these mothers was significantly fewer in quantity and less diverse in quality. This is in line with the findings from a survey of South Asian (Pakistani and Nepalese) parents in Hong Kong (Tsung & Gao, 2012) that although these parents expressed their desires to learn Chinese and high aspirations for education of their children in Hong Kong, they lacked the linguistic and cultural capital valued by the mainstream society. It was also worth noting that quite a substantial number of SA mothers did not join the study at the end. The participant drop may indicate that they may have an even lower level of Cantonese proficiency than those SA mothers in our participant pool who had stayed in Hong Kong for more than 20 years. The SA mothers in

our participant pool may have already developed a relatively higher Cantonese proficiency after such a long stay in Hong Kong. Along this line of logic, we can imagine that the rest of the South Asian mothers in Hong Kong may provide even less Cantonese input to their children during the early years of language development. Frequency of input is a highly important factor in determining the language proficiency of children who learn two languages as their first language (De Hourwer, 2007, 2009). Given the importance of language input in language development, such a gap during toddler years may pose more dramatic differences in later language development in Cantonese and reading and writing in Chinese in general. Even with additional support in primary school years, their Chinese language learning may still lag behind those children in Chinese families and cannot achieve native-like literacy level. These young children must be given opportunity to hear Cantonese very frequently. When parents of South Asian background are keen on raising children bilingually as reviewed in various surveys and our Government aims to provide an equitable environment to all their citizens, we have to develop proactive ways to actively increase the Cantonese input to the children with the SA family background during the early toddler years and preschool years, in addition to engaging South Asian parents to learn Cantonese as a second language. In this way, the outcome of the current support provided during the school years to the children would be even more productive.

In terms of support for very young children, we recommend ways that can increase Cantonese input to young children in South Asian families. These means have to cost-effective and assessible in society. If limited input is related to parents' restricted level of language proficiency, relying solely on parents such as improving their Cantonese language ability may not an ideal way to mediate Cantonese learning of their young children, who are even more active and fast-growing learners during the critical language acquisition period. The substantially improved technology these days may provide valuable support to provide Cantonese language input in an enjoyable way to children and their parents. According to Census, mobile device or smart phone penetration among individuals at the age range between 15 to 40 years old has almost reached 100% in Hong Kong in 2021. Government may consider supporting the production of series of videos in Cantonese for these young children. However, there is in fact an extensive bank of high-quality Cantonese (original or translated) cartoons or musical (e.g., 兒歌) videos that are readily and freely available on the internet. For example, educational programs that label objects, talking directly

to the child and provide opportunities (i.e., short pauses) to respond verbally would be particularly beneficial to younger children (Linebarger et al., 2005). Programs geared towards early preschool-children can also include a coherent and integrative narrative and age-appropriate language are also beneficial to children and assist their language learning (e.g., 粉紅豬一家親 *Peppa Pig* https://www.youtube.com/watch?v=P-ExF_bFUDM). These programs can also enhance early language learners who learn more than one first languages.

New development a series of video programs supported by the Government can be an option to help. Another cost-effective option is to provide advice on the selection of these relevant videos to South Asian families (after the consideration of any potential issues of copyrights). Selection of which types of videos or cartoons can be given by language experts based on the age-appropriateness of the language (in Cantonese, cognitive level) and design. Many well-made cartoons are available. It should be noted that although there appeared to have an ongoing debate regarding the benefits and risks of screen use on child's language, a recent meta-analysis and systematic review revealed that although great ***quantity*** of screen use (screen time and background television) was linked to lower child language skills, ***better quality*** of screen (educational and co-viewing) was ***positively*** associated with more advanced child language skills (Madigan et al., 2020). It is therefore in addition to the recommendation of which high-quality and age-appropriate Cantonese channels that the South Asian parents can select to ensure the quality, it is also very paramount to highlight the limit of ***one hour per day*** for toddlers at 2 years of age and the importance of co-viewing with their children. According to the American Academy of Pediatrics (AAP) (2016), caregivers are advised to avoid use of screen media other than video-chatting and it is recommended that no screen exposure to children ***before age 18 months***. For caregivers of children 18 to 24 months of age who want to introduce digital media to their children should select high-quality programming and watch it with their children (i.e., co-viewing) to help them make sense of what they are watching. AAP also recommend screen limit to one hour per day of high-quality programs and co-viewing for children age 2 to 5 years. These recommendations should also apply to children learning more than one language as time should also be reserved for young children to practice other developmental milestones such as motor and social skills. Prolonged screen time may also limit or hinder important parent-child interactions which is also critical of child language development.

In preschools or child-care centres, these South Asian children must be given the opportunity to hear Cantonese very frequent. Teachers can create multitude of various settings in which children can hear the language. This may include constructing situations where the South Asian children can interact freely with children who speak native Cantonese. Through these peer interactions, children can learn a great deal in addition to teachers' teaching in the classroom.

Study 2 documented the phonological transfer patterns in L1-Putonghua/L2-Cantonese late bilinguals preliminarily by the use of selected target sets. The results revealed that certain vowels appeared to be more prone to be substituted than the others. The most prominent substitution was the final bilabials which do not exist in Putonghua. Substitution of Cantonese phonemes with their Putonghua counterparts with minimal contrasts may be a major error pattern in L1-Putonghua/L2-Cantonese speakers. In addition, the phonological transfer is rather consistent across PTH speakers and because of the closeness of the substitutions to the target, listeners can resolve the mismatches from the context and the mismatches did not lead to significant communication breakdown or remarkable reduction in speech intelligibility. In previous studies, members of this group encounter difficulties to integrate into mainstream HK society due to their non-standard Cantonese (Derwing & Waugh, 2012; Esser, 2006). When PTH speakers are very keen on reducing the accent, support can be sought to reduce the accent by focusing on the predominant sound patterns.

Final Remarks

The current project has developed a corpus of Cantonese child directed speech and corresponding adult directed speech produced by three groups of mothers with different language background. The language samples were collected during 2019/2020 when was an exceptional year in Hong Kong with the social events and pandemic outbreak. Participant recruitment and data collection had been a challenge. The corpus not only allows us to explore the above issues aimed in the original proposal, but the process involved had improved our knowledge in many research areas. For example, the work pipeline of constructing of a phonetically annotated corpus such as transcription, reliability check and the forced-alignment procedures for coding Chinese speech samples. The established work pipeline facilitates development of future oral language corpus in Chinese. Based on the corpus data, further analyses between the comparison of CDS and ADS in different mother groups are in progress. All relevant

publications along with the corpus will be made available to the general public for further research development.

(k) Bibliography

Baker, W., & Trofimovich, P. (2005). Interaction of native-and second-language vowel system (s) in early and late bilinguals. Language and speech, 48(1), 1-27.Gruyter.

Bedore, L. M., Peña, E. D., Griffin, Z. M., & Hixon, J. G. (2016). Effects of age of English exposure, current input/output, and grade on bilingual language performance. *Journal of Child Language*, *43*(3), 687-706.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The development of speech perception: The transition from speech sounds to spoken words*, *167*(224), 233-277.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, *109*(2), 775-794.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech. *Language experience in second language speech learning: In honor of James Emil Flege*, *17*, 13.

Bigi, B. (2015). SPPAS-multi-lingual approaches to the automatic annotation of speech. *The Phonetician*, *111*(112), 54-69.

Bigi, B. (2018). Annotation representation and File conversion tool. Contributi del Centro Linceo Interdisciplinare 'Beniamino Segre', 137, 99-116.

Boersma, P., & Weenink, D. (2020). Praat: Doing phonetics by computer, version 6.1. 16. *Amsterdam: Phonetic Sciences, University of Amsterdam*.

Bugarski, R. (1991). Contrastive analysis of terminology and terminology of contrastive analysis. In Ivir, V., & Kalogjera, D. (Eds.), *Languages in contact and contrast: Essays in contact linguistics* (Vol. 54) (pp. 73-82). Walter de Gruyter.

Burchinal, M., Field, S., López, M. L., Howes, C., & Pianta, R. (2012). Instruction in Spanish in pre-kindergarten classrooms and child outcomes for English language learners. *Early Childhood Research Quarterly*, *27*(2), 188-197.

Carmichael, S. (2009). Language rights in education: a study of Hong Kong's linguistic

minorities. *Occasional Paper*, *19*.

Chen, X., Anderson, R. C., Li, W., Hao, M., Wu, X., & Shu, H. (2004). Phonological awareness of bilingual and monolingual Chinese children. *Journal of Educational Psychology*, *96*(1), 142.

Chen, Y., Ng, M. L., & Li, T. S. (2012). English vowels produced by Cantonese–English bilingual speakers. *International journal of speech-language pathology*, *14*(6), 557-568.

Cheung, S., Kan, P. F., Winicour, E., & Yang, J. (2019). Effects of home language input on the vocabulary knowledge of sequential bilingual children. *Bilingualism: Language and Cognition*, *22*(5), 986-1004.

De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*, *28*(3). 411–424.

De Houwer, A. (2009). Bilingual First Language Acquisition. Clevedon: Multilingual Matters.

Derwing, T. M., & Waugh, E. (2012). Language skills and the social integration of Canada's adult immigrants. *IRPP study*, (31), 1.

Esser, H. (2006). *Migration, language and integration*. Berlin: WZB.

Escudero, P., & Boersma, P. (2002). The subset problem in L2 perceptual development: Multiple-category assimilation by Dutch learners of Spanish. In *Proceedings of the 26th annual Boston University conference on language development* (pp. 208-219). Somerville, MA: Cascadilla.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, *92*, 233-277.

Flege, J. E. (1998). The phonetic study of bilingualism. *Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies*, (35), 017-026.

Hedlund, Gregory & Yvan Rose. 2020. Phon 3.1 [Computer Software]. Retrieved fromhttps://phon.ca.

Tsung, L., & Gao, F. (2012). What accounts for the underachievement of South Asians in Hong Kong? The voices of Pakistani and Nepalese parents. *Educational Research*, 54(1), 51-63.

Holm, A., & Dodd, B. (1999). An intervention case study of a bilingual child with phonological disorder. *Child Language Teaching and Therapy*, *15*(2), 139-158.

Hong Kong Education Bureau. (2016). *Chinese Language Curriculum Second Language Learning Framework*. Retrieved from https://www.edb.gov.hk/tc/curriculum-development/curriculum-area/special-educational-needs/curriculum-sec-lang-alframework/index.html.

Khouw, E. & Ciocca, V. (2006). An acoustic and perceptual study of final stops produced by

profoundly hearing-impaired adolescents. *Journal of Speech, Language, and Hearing Research*, *49*, 172-185.

Kong, A. P. H., & Law, S. P. (2019). Cantonese AphasiaBank: An annotated database of spoken discourse and co-verbal gestures by healthy and language-impaired native Cantonese speakers. *Behavior research methods*, *51*(3), 1131-1144.

Lado, R. (1957) *Linguistics across cultures.* Applied linguistics for language teachers. University of Michigan Press.

Lam, K. K., & To, C. K. (2017). Speech sound disorders or differences: Insights from bilingual children speaking two Chinese languages. *Journal of communication disorders*, *70*, 35-48.

Law, C. W. (2006). Tonal characteristics of early English-Cantonese bilinguals. *HKU Theses Online (HKUTO)*.

Law, Z. W., & So, L. K. (2006). Phonological abilities of hearing-impaired Cantonese-speaking children with cochlear implants or hearing aids.

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, *36*(3), 345-366.

Lee, T., Lo, W. K., Ching, P. C., & Meng, H. (2002). Spoken language resources for Cantonese speech processing. *Speech Communication*, *36*(3-4), 327-342.

Lee, W. S., & Zee, E. (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association*, *33*(1), 109-112.

Li, X., To, C. K. S., & Ng, M. L. (2017). Effects of L1 tone on perception of L2 tone-a study of Mandarin tone learning by native Cantonese children. *Bilingualism: Language and cognition*, *20*(3), 549-560.

Luo, S. H., & Wiseman, R. L. (2000). Ethnic language maintenance among Chinese immigrant children in the United States. *International Journal of Intercultural Relations*, *24*(3), 307-324.

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk: Volume I: Transcription format and programs, volume II: The database.

Madigan, S., McArthur, B. A., Anhorn, C., Eirich, R., & Christakis, D. A. (2020). Associations between screen use and child language skills: a systematic review and meta-analysis. JAMA pediatrics, 174(7), 665-675.

Major, R. C. (2008). Transfer in second language phonology. *Phonology and second language*

acquisition, *36*, 63-94.

Malvern, D. D., & Richards, B. J. (1997). A new measure of lexical diversity. *British Studies in Applied Linguistics*, *12*, 58-71.

Mok, P., Lee, C. W., & Yu, A. C. (2018). Perception and production of Cantonese tones by South Asians in Hong Kong. In *Proc. 9th International Conference on Speech Prosody 2018* (pp. 458-462).

Moradi, Hamzeh & Chen, Jianbo. (2018). A contrastive analysis of Persian and English vowels and consonants. *Lege Artis, 3*, 105-131.

Navehebrahim, M. (2012). RETRACTED: An Investigation on Pronunciation of Language Learners of English in Persian Background: Deviation Forms from the Target Language Norms.

Oller Jr, J. W., & Ziahosseiny, S. M. (1970). The contrastive analysis hypothesis and spelling errors. *Language learning*, *20*(2), 183-189.

Oxfam Hong Kong. (2014). Survey on the Chinese learning challenges South Asian ethnic minority kindergarten students from low-income families face.

Paradis, J. (2016). The development of English as a second language with and without specific language impairment: Clinical implications. *Journal of Speech, Language, and Hearing Research*, *59*(1), 171-182.

Ringbom, H. (1994). Contrastive analysis. The Encyclopedia of language and linguistics, 2, 737-742.

Sebastián-Gallés, N., & Bosch, L. (2005). *Phonology and bilingualism*. Oxford University Press.

Seddighi, S. (2012). An account of Iranian EFL pronunciation errors through L1 transfer. *Iranian Journal of Applied Language Studies*, *2*(2), 197-214.

Team, R. C. (2013). R: A language and environment for statistical computing.

To, C. K., Cheung, P. S., & McLeod, S. (2013). A population study of children's acquisition of Hong Kong Cantonese consonants, vowels, and tones. *Journal of Speech, Language, and Hearing Research*, *56*(1), 103-122.

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, *71*(1), 4-21.

Velazquez, M. I. (2008). *Intergenerational Spanish language transmission: Attitudes, motivations and linguistic practices in 2 Mexican American communities*. University of Illinois at Urbana-Champaign.

Wei, L., Miller, N., Dodd, B., & Hua, Z. (2005). Childhood bilingualism: Distinguishing difference

from disorder. *LANGUAGE IN SOCIETY-OXFORD-*, *36*, 193.

Weinreich, U. (1953). *Languages in contact*. New York: Linguistic Circle of New York.

Zeng, Z. (1994). *Comparison between and instruction of Cantonese and Mandarin*. Hong Kong, China: Joint Publishing.

Zhu, H., & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard Chinese). *Journal of Child Language*, *27*(1), 3-42.