

# Distinct acoustic properties pattern differently in a sound change in progress in Hong Kong Cantonese

John M. Culnan 

University of Arizona, USA  
[jmculnan@email.arizona.edu](mailto:jmculnan@email.arizona.edu)

Suki S.Y. Yiu 

University of Amsterdam, The Netherlands  
[s.y.yiu@uva.nl](mailto:s.y.yiu@uva.nl)

Previous work on Hong Kong Cantonese has reported evidence of an incomplete neutralization between the labialized velar /k<sup>w</sup>/ and plain velar /k/, where both may be produced as a plain velar [k] before /ɔ/, although non-neutralized forms are also produced to some extent for some speakers (Bauer 1982, Zee 1999, To, Mcleod & Cheung 2015). The present study examines this phenomenon acoustically, and evaluates to what extent the properties of labialization are present in the neutralizing environment. Fourteen native speakers of Hong Kong Cantonese were recorded in a production experiment; these recordings were analyzed by examining formant values, duration, and rise time of intensity for vowels following target segments. Results reveal that while many acoustic differences between /k<sup>w</sup>/ and /k/ are maintained in the neutralizing environment, these speakers neutralize F3 and duration measures. Patterns of variation furthermore suggest that speakers utilize these acoustic properties differently, indicating greater degrees of (incomplete) neutralization for some speakers than others.

## 1 Introduction

Sound changes are a popular area of phonetic and phonological research, with work focusing on the theory of sound change (Labov 1981) and the acoustic and articulatory properties of sounds undergoing change (Rutter 2011). Among the most attested types of sound changes documented in languages of the world are phonological mergers (e.g. Wedel, Kaplan & Jackson 2013) and contextual neutralization (Chitoran 2002, Kochetov 2004, Warner et al. 2004, Lee & Jongman 2016, Wade 2017, Bongiovanni 2021, Conrad published online 15 April 2021). Included among these are a handful of cases in Hong Kong Cantonese (Yu 2007, 2016; Mok, Zuo & Wong 2013; Fung & Lee 2019), where contextually neutralized forms may coexist with non-neutralized forms for several generations (Bauer 1982, Newman 1987). In this study, we examine a case of contextual neutralization, where /k<sup>w</sup>/ and /k/ may be pronounced equivalently when produced before the rounded back vowel /ɔ/, such that the syllable [kɔk] may either refer to either /k<sup>w</sup>ɔk3/ ‘country’ or /kɔk3/ ‘corner’. We suggest that, for at least one group of speakers of Hong Kong Cantonese, numerous acoustic properties are preserved in the neutralizing environment, with different speakers either maintaining a clear distinction between /k<sup>w</sup>/ and /k/ or showing incomplete neutralization.

### 1.1 Contextual neutralization as sound change in progress

Acoustic studies of contextual neutralizations have demonstrated that additional care must be taken when evaluating their status: acoustic components of a sound are not always impacted equally. For example, Slowiaczek & Dinnsen (1985) examined word-final devoicing in Polish, finding that devoicing is a case of incomplete neutralization, although the way in which the distinction between underlying voiced and voiceless obstruents is preserved differs based on several factors. Kharlamov (2014) examined word-final obstruents in Russian, finding that neutralized underlyingly voiced obstruents maintain a distinction in terms of both consonant duration and glottal pulsing and that these acoustic measures interact with each other differently depending upon orthography, number of syllables in a word, and speech style. In other words, whereas a basket of acoustic measures may be subject to change in a neutralizing environment, these acoustic measures may not all be impacted or may not be impacted equally.

Previous studies have also examined the role of social group in sound change (Lippi-Green 1989), and have indicated that high levels of individual variation may be present with respect to sound change coming from sociolinguistic factors such as age or generation, gender, social group and status, and ethnicity (Bauer 1982, Milroy & Milroy 1985, Hall-Lew 2013). Dinnsen & Charles-Luce (1984), for example, examined the incomplete neutralization of word-final obstruent devoicing in Catalan, finding that while vowel duration, consonant closure duration and voicing into consonant closure were all neutralized overall, individuals showed different patterns of change to closure duration and voicing into closure in neutralizing environments.

Of interest to this paper is the incomplete neutralization of the labialized velar stop /k<sup>w</sup>/ and its plain counterpart /k/ when preceding rounded vowels in Hong Kong Cantonese, where multiple acoustic properties may be at play to various extents. While /k/ preceding rounded vowels was reported to show a significant degree of anticipatory labialization comparable to that of its contrastive /k<sup>w</sup>/ preceding unrounded vowels in Nawuri (a Kwa language spoken in Ghana), the secondary labialization with /w/ preceding rounded vowels has been described as generally not very audible as a principle of speech perception (Casali 1990). Despite the inaudibility of the secondary labialization, in languages like Marshallese (an Austronesian language spoken on the Marshall Islands), the phonetic realization of vowels were found to be systematically determined by the secondary articulations in their surrounding consonants (Choi 1995). The situation gets more complicated when there is an asymmetry between labial-initial and non-labial-initial sequences. In Romanian, different representations were necessary for [ea] vs. [ja] whereas phonetic neutralization was found for [oa] vs. [wa]. Chitoran (2002) explained such findings with language-specific frequency differences and contrast maintenance, urging for detailed phonetic studies for better understanding of the phonology of the language.

The /k<sup>w</sup>/–/k/ neutralization has been described as one of a handful of ongoing sound changes over the last century in Cantonese (Bauer 1982, Newman 1987), together with tone merger (Yu 2007, Mok et al. 2013, Fung & Lee 2019) and individual variation in the coarticulation of /s/ in different vowel contexts (Yu 2016), for example. The /k<sup>w</sup>ɔ̄/–/kɔ̄/ neutralization has been recognized since 1972 (To, Mcleod & Cheung 2015) and shown to be nearly complete in rapid speech in the early 1980s (Bauer 1982). Bauer (1982) observed that /k<sup>w</sup>ɔ̄/ sequences were produced without labialization (as [kɔ̄]) a majority of the time by most subjects at varying socioeconomic statuses and ages for numerous lexical items. Yet a more recent study found a higher degree of maintenance of the distinction between the two stops for a single lexical item, /p<sup>h</sup>ɪŋ21.k<sup>w</sup>ɔ̄25/ ‘apple’ (To et al. 2015). While these previous studies have focused on perceptual categorization of /k<sup>w</sup>ɔ̄/ productions as [k<sup>w</sup>ɔ̄] or [kɔ̄], this paper aims at offering a detailed examination of the acoustic properties involved in the realization of /k<sup>w</sup>/ when put in rounded versus unrounded environments. Worth noting is that such neutralization also exists with the aspirated labialized velar /k<sup>w</sup>h/, although there is a higher level

of uncertainty regarding the phonemic representation of individual words with /k<sup>wh</sup>/ and /k<sup>h</sup>/ as evidenced by dictionary entries Cheung et al. 2016, Sheik 2017). As such, a full discussion of the aspirated phonemes is outside the scope of the current study.

## 1.2 The status of /k<sup>w</sup>/ in the Cantonese phonological inventory

The labialized velar onset /k<sup>w</sup>/ in Cantonese is considered a complex onset segment instead of a sequence composed of [k] followed by [w], and contrasts phonemically with the plain velar stop /k/ (Bauer & Benedict 1997). In terms of articulation, the velar stop [k] and labiovelar [w] are pronounced simultaneously or nearly simultaneously. The back of the tongue is held against the velum for [k] while the lips are tightly rounded for [w]. In terms of syllable structure, Cantonese syllables are considered to have maximally one consonant in the onset position (and at most one consonant in coda position). Because this type of labialization occurs solely with velar stops, Cantonese is frequently interpreted as having a labialized velar phoneme rather than a velar–glide sequence (Bauer 1982, Newman 1987, Yip 1993, Matthews & Yip 2011, To et al. 2015). Thus /k<sup>w</sup>ɔk/ would be a valid syllable, but /kwɔk/ would not. Identifying this as a single consonant with a secondary articulation therefore fits with Cantonese phonotactics.

In terms of the Cantonese syllabary (Bauer & Benedict 1997), /k<sup>w</sup>/ tends not to combine with rimes ending in labial consonants (i.e. /w m p/) whereas /k/ combines freely with them. Furthermore, while /k<sup>w</sup>/ occurs with four vowels, the environments in which /k<sup>w</sup>/ occurs demonstrate a meaningful contrast with /k/, supporting the claim that the segments in question are two distinct phonemes, and that the process of interest is indeed neutralization of the two in certain environments.

The labialized velar stop /k<sup>w</sup>/ occurs in onset position before four vowels: /a ɔ i ɐ/ (Kao 1971). Examples of these, as well as the vowels with which /k<sup>w</sup>/ cannot occur, are shown in (1): below.

(1) a. [k <sup>w</sup> a55]	melon	[ka55]	family
b. [k <sup>w</sup> ɔ33] or [kɔ33]	pass, across	[kɔ33]	classifier
c. [k <sup>w</sup> ɪŋ25]	bright	[kɪŋ25]	view
d. [k <sup>w</sup> ɐŋ55]	bombard	[kɐŋ55]	spoon
e. *[k <sup>w</sup> øŋ]	unattested	[køŋ55]	ginger
f. *[k <sup>w</sup> eŋ]	unattested	[keŋ25]	neck
g. *[k <sup>w</sup> u]	unattested	[ku25]	ancient
h. *[k <sup>w</sup> y]	unattested	*[ky]	unattested

Note that /i/ in the context of interest is produced as [ɪ], as in (1c). It is important to understand that the number of phonemes in the vowel inventory of Hong Kong Cantonese is subject to a certain amount of debate, as there is significant allophonic variation in production (Bauer & Benedict 1997). As a result, researchers may come to different conclusions about the phonemic inventory (Zee 1991, Bauer & Benedict 1997). We follow Kao (1971) in selecting an inventory of eight vowels based upon patterns of distribution rather than the fourteen described by (Bauer & Benedict 1997) on the basis of position-based phonetic quality. This

is done for the sake of simplicity, as the vowels examined in this study are the same in both inventories.

In all except (1b), the acoustic distinction between the labialized velar stop and its velar counterpart is consistently produced. The fact that this loss of labialization takes place before the only rounded vowel with which /k<sup>w</sup>/ occurs is unsurprising, as other languages are also shown to disprefer labialized stops before back rounded vowels (Ohala & Kawasaki-Fukumori 1997). Kim (2010) demonstrates that English avoids labiovelar glides before high rounded vowels, and Korean does not allow any labialized segment before a rounded back vowel; Stonham & Kim (2008) likewise observe the loss of labialization on phonemically labialized stops before /u/ in Nuuchahnulth. This is further supported by Bauer & Benedict (1997), who observe that of 31 of 34 possible Cantonese rimes containing rounded vowels are unattested after labialized velar stops; only rimes /ɔŋ/, /ɔk/ and /ɔ/ are attested.

### 1.3 Acoustic properties of labialization

In order to evaluate how the acoustic properties of labialization are altered in the neutralizing environment, those properties must be identified and evaluated. Labialization as a secondary articulation tends to affect both the labialized consonants and the following vowels, especially portions of the vowel immediately after consonant release (Flemming, Ladefoged & Thomason 2008, Beeley 2015). These differences in acoustic measurements are a direct result of the lip rounding that is a crucial characteristic of both labialized segments and rounded vowels. Specific acoustic properties of interest in productions involving lip rounding are vowel duration, the first three formants, and rise time of intensity.

In the case of Hong Kong Cantonese, the initial consonant /k<sup>w</sup>/ is a coarticulated labialized-velar stop, as often indicated by the raised [w] in its narrow phonetic transcription, its incompatibility with rimes ending in labial consonants, and the absence of consonant clusters (Bauer & Benedict 1997). In order to pronounce the voiceless velar stop [k] and the voiced labial approximant [w] simultaneously, the lips are rounded for [w] while the back of the tongue is held against the velum for [k]. The labialized-velar stop /k<sup>w</sup>/ is often tightly rounded (with lip corner compression) so as to contrast with its plain counterpart /k/. When this tight lip rounding occurs, lips are further extended, increasing the length of the vocal apparatus and further lowering formant values.

Differences in duration have been documented when comparing a velar stop–labiovelar glide sequence with a labialized velar phoneme (Suh 2007). Catford (1988) furthermore demonstrates that, even within the realm of what is considered phonetically to be a labialized velar stop, there is a possible range of differences in the points of maximal stricture for the velar closure and the labial constriction. In other words, an articulation that is described as a production of a single labialized velar stop may have maximal labial constriction occur slightly after the velar stop is produced. When this happens, the release of the labial constriction may lengthen the sequence, similar to how a labial–velar glide sequence lengthens the following vowel.

Previous work has highlighted that labialization results in lower values for F1, F2, and F3 due to lip protrusion, which slightly increases the length of the vocal tract (Stevens 1998, Sturman et al. 2016). Stevens indicates that it may take 50–100 ms to fully return the lips to a neutral position after rounding, which means that rounding on an onset consonant will necessarily affect values on at least the first portion of the following vowel.

A further complication arises when considering the interplay of a labialized onset consonant and a rounded vowel. It has been reported that the labialized-velar stop /k<sup>w</sup>/ has lost its lip labial friction for [w] when preceding rounded vowels (Bauer & Benedict 1997), as a result of the co-occurrence of two similar labial sounds in sequence. This neutralization of /k/ and /k<sup>w</sup>/, however, is still expected to leave room for some lip-rounding on the

**Table 1** Summary of effects of lip rounding on relevant acoustic measures.

Measure	Effect of labialization
Duration	Increase
F1	Decrease
F2	Decrease
F3	Decrease
Rise time of intensity	Increase

neutralized /k<sup>w</sup>/ as an anticipatory cue of the following rounded vowels. In view of the above gestural coordination on formants, according to Stevens (1998), when lip rounding is added, F1 is lowered more for low vowels than for high vowels, and is relatively stable for /i/. Back rounded vowels, such as /ɔ/, tend to have lower F2 values than front vowels. Because labialization on consonants also corresponds to lower F2 at the vowel onset, a labialized stop followed by an unrounded vowel, especially a front vowel, will demonstrate a large increase in F2 at the start of the vowel, whereas a labialized stop followed by a rounded back vowel will have a less dramatic increase in F2 (Flemming et al. 2008). Finally, Sturman et al. (2016) reported that labialization as a secondary articulation frequently results in lower F3 values than non-labialized counterparts. Beeley (2015) further determined that F3 at vowel onset was a better measure of labialization than F2 in a dialect of British English.

Rise time of intensity, the amount of time it takes for amplitude to increase from the start of a vowel towards its maximum, may additionally be lengthened by onset labialization. This is because labialization as a secondary articulation on an onset consonant is phonetically similar to a glide–vowel sequence (Flemming et al. 2008). Because this labialization lowers amplitude (Stevens 1998) and, if it occurs slightly after the velar closure, extends the length of the following vowel, it may take longer for amplitude to increase after the labialized velar /k<sup>w</sup>/ than after /k/.

Table 1 summarizes the anticipated effects of labialization upon each of the acoustic properties examined in this study. Vowel duration should increase in the labialized condition. Each of the formants is expected to be lower in labialized conditions than in their corresponding plain velar conditions, while rise time of intensity is expected to be longer after labialized onsets than plain velar onsets.

## 1.4 Predictions

Based on the previous research into incomplete neutralization as sound change in progress and its relevance to /k<sup>w</sup>ɔ/ neutralized as [k] in Cantonese, we predict the following:

- (i) Some acoustic properties associated with labialization will be preserved in the /k<sup>w</sup>ɔ/ environment, while others will be neutralized.
- (ii) Speakers vary in the extent to which they produce different acoustic properties associated with labialization. That is, some speakers will preserve the distinction of fewer properties than other speakers.

In order to evaluate prediction (i), we examine the acoustic properties associated with labialization in the neutralization environment, and compare these with the same properties in non-neutralizing environments. Acoustic properties preserved in the /k<sup>w</sup>ɔ/ environment are expected to differ significantly from those in the /kɔ/ environment. To confirm the relevance of these acoustic properties in onset labialization in Hong Kong Cantonese, there should also

be a significant difference in non-neutralizing environments. When the values of an acoustic property do not differ significantly between /k<sup>w</sup>ɔ/ and /kɔ/, there is evidence of contextual neutralization of this property. These properties may be impacted by differences in the extent to which individuals round their lips in the /k<sup>w</sup>ɔ/ environment, whether or not they use lip compression during lip rounding, or by the timing of the maximal labial constriction relative to the velar closure.

If speakers do not engage in as much lip rounding in the /k<sup>w</sup>ɔ/ environment as they do in other /k<sup>w</sup>/ environments, the coarticulatory rounding from the vowel /ɔ/ will result in the same amount of lowering to F1, F2, and F3 values in /k<sup>w</sup>ɔ/ and /kɔ/ conditions. Likewise, unless labial constriction peaks at the same time for both the /k<sup>w</sup>ɔ/ and /kɔ/ sequences, there should be a difference in vowel duration for the two. Other vowel contexts, meanwhile, should have slightly longer duration after labialized velars than after plain velar stops.

Prediction (ii) highlights the importance of individual difference by predicting that some speakers will show greater preservation of the contrast, while others will show more complete neutralization. A speaker showing greater preservation of contrast will produce significantly different values for all or most acoustic properties associated with labialization between /k<sup>w</sup>ɔ/ and /kɔ/ sequences, whereas a speaker showing greater neutralization will show fewer significant differences among these properties. To do this, some speakers may produce /k<sup>w</sup>ɔ/ sequences with a smaller degree of lip rounding than other /k<sup>w</sup>V/ sequences, while others may alter the timing of maximal labial constriction relative to velar closure in /k<sup>w</sup>ɔ/ relative to other /k<sup>w</sup>V/ conditions.

## 2 Method

### 2.1 Participants

Fourteen native speakers of Hong Kong Cantonese (eight male, six female) participated in the study. Participants were recruited through a public university in Hong Kong and were between the ages of 18 years and 29 years at the time of the study. All participants indicated that their primary language was Cantonese, that they grew up in Hong Kong, and that they felt comfortable to complete the consent process and a language background questionnaire in English. All participants were fluent readers of traditional Chinese, and all had prior exposure to Mandarin. Furthermore, all participants indicated that they had no known history of hearing impairments or speech disorders. Responses to the language background questionnaire indicate that all participants use Cantonese regularly with their peers, and all but one use Cantonese with their parents and grandparents.

Overall articulation rates (calculated as number of syllables per second of phonation) were automatically calculated for each participant using a Praat script (De Jong & Wempe 2009). The participants had a mean articulation rate of 4.41 syllables per second, with a standard deviation of 0.35 syllables per second. Participant m4 had the lowest articulation rate at 3.96 syllables per second, while participant m3 had the highest articulation rate at 4.98 syllables per second.

### 2.2 Stimuli

The stimuli consisted of 132 target words and 129 fillers, all of which are real words and morphemes of Cantonese. Participants produced one token with /k<sup>wh</sup>/ rather than /k<sup>w</sup>/, as perceived by a native Cantonese-speaking author, so it was not considered for analysis. Each target word began with either /k<sup>w</sup>/ or /k/ followed by /a/, /ɛ/, /i/, or /ɔ/ with one of three coda conditions in the target syllable: /ŋ/, /k/ or no coda. The distribution of target syllables is shown below in Table 2.

**Table 2** Frequency of each syllable condition in the stimuli. Where possible, ten stimuli were included per condition; eleven stimuli were included in the /a/ condition as there were limited other /k<sup>w</sup>a/ stimuli. Unattested syllables are presented with a dash. Note that the word included in the stimulus list for condition /k<sup>w</sup>aj/ was produced as [k<sup>w</sup>h<sup>a</sup>aj], and was therefore removed from analysis. For the complete list of target and filler stimuli, please see the appendix.

Syllable	/ɔ/	/a/	/ɐ/	/i/
/k <sup>w</sup> V/	10	11	-	-
/k <sup>w</sup> Vŋ/	10	1	3	3
/k <sup>w</sup> Vk/	10	1	-	-
All /k <sup>w</sup> /	30	13	3	3
/kV/	10	11	-	-
/kVŋ/	11	1	10	10
/kVk/	10	10	-	10
All /k/	31	22	10	20

Targets were matched as closely as possible across onset and vowel conditions, although this was frequently not possible as only real morphemes of Cantonese were used. Thus, targets included such real words as /k<sup>w</sup>ək3.jən21/ ‘countryman’ and /kək3.jən21/ ‘everyone, each person’, but not the corresponding non-words \*/k<sup>w</sup>ik.jən/ or \*/kik.jən/. The uneven number of items in Table 2 alludes to the difficulties both of finding acceptable tokens and of matching stimuli across conditions. For example, sequences /k<sup>w</sup>ɐ/ and /k<sup>w</sup>i/ are rare in the language, occurring only in closed syllables and primarily only with coda /ŋ/; /k<sup>w</sup>ak/ is similarly an uncommon syllable in the language. In total, 1,834 target tokens were analyzed, with 131 tokens analyzed per participant.

Fillers were also real words and morphemes of Cantonese (such as /pək5/ ‘north’ and /lœŋ21.ts<sup>h</sup>a21/ ‘herbal tea’, although they began with non-target syllables. One hundred and fourteen fillers ended in /t/ or /k/ in order to serve as distractors for participants who were looking for patterns in the stimuli. A full list of target and filler stimuli appears in the appendix.

### 2.3 Procedures

Participants were recorded in a sound-proof booth in the Laboratory of Phonetics and Phonology at The University of Hong Kong using a digital speech recorder with a high-quality head-mounted microphone at a sampling rate of 44.1 kHz. Participants were seated at a desk with a Macbook laptop, which displayed one written stimulus at a time on a slide viewer. Each stimulus was produced in the carrier phrase /k<sup>h</sup>øy23.kœŋ25.tsə25/ TARGET /jət5.ts<sup>h</sup>i33/ ‘s/he said TARGET one time’, which was written at the bottom of each slide in traditional Chinese characters in case they forgot it. The carrier phrase was intended to provide a neutral semantic context in which to produce each word, although it is possible that the inclusion of the /kɔ/ sequence in /kœŋ/ might have a subtle impact upon productions of target stimuli.

Stimuli were presented a single time in a pseudo-random order with no more than two consecutive target words with the same onset–vowel sequence and no more than three total consecutive target words, as well as at most four consecutive fillers. Each participant was presented with the stimuli in the same pseudo-random order. Participants were instructed to speak at a normal pace, as though talking to their friends.

## 2.4 Measurements

Following are descriptions of the acoustic measures calculated on all vowels following either /k<sup>w</sup>/ or /k/. These measurements were: vowel duration, F1, F2, and F3 values at voicing onset, F1, F2, and F3 values at vowel midpoint, and rise time in intensity. All boundaries in target vowels were marked by identifying onset and offset of visible F2 in a spectrogram, and acoustic measurements were automatically extracted from the recordings using Praat (Boersma & Weenink 2018). The automatically extracted values were then carefully examined for accuracy by comparing them with visible formant structures in a spectrogram. The Praat script used for formant extraction was modified from Kroos, Bundgaard-Nielsen & Tyler (2010), and the script for rise time was altered from the intensity extraction script from Kawahara (2010).

### 2.4.1 Duration

Vowel duration was measured by identifying the start and end of the visible F2 using spectrograms generated by Praat. Identification of the start and end of F2 was completed manually, with vowel boundaries indicated through a textgrid tier and the interboundary length extracted automatically to provide the duration measurement.

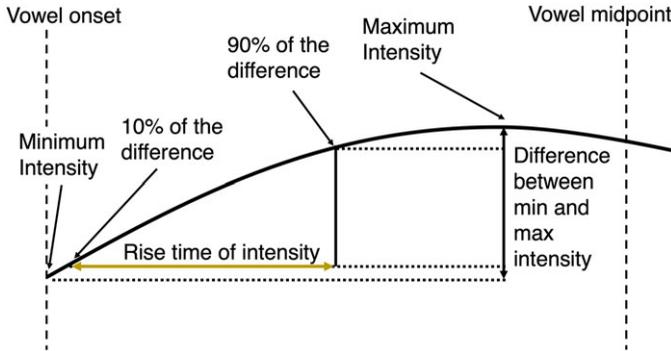
### 2.4.2 Formant measurements

F1, F2 and F3 values were extracted at 5 ms post-onset of voicing and at the vowel midpoint. These time points were selected to show the labialization during formant transition from the onset consonant to the vowel and the steady state of the vowel. Measurements were checked for accuracy. And, in places where errors were noted (e.g. where the value selected as F2 was in fact taken from F3, and where F3 was taken from F4), values were altered by manual formant measurement using their Praat spectrograms. These corrections resulted in calculations that could be reliably analyzed in linear mixed effects models.

Formant measurements were normalized in order to create a consistent comparison across speakers. This was done with the R Vowels package (Kendall & Thomas 2018) by using Labov normalization, which uses the method presented in Labov, Ash & Boberg (2008) by scaling individual speakers' formant productions to normalized Hertz values.

### 2.4.3 Rise time of intensity

Rise time of the intensity of the vowel (often called Amplitude Rise Time) refers to the increase in the amplitude of a syllable or sound sequence (Nittrouer, Lowenstein & Tarr 2014). To calculate rise time, intensity was first measured at 40 ms intervals over the first half of the vowel, the standard sampling interval in Praat. Minimum intensity near vowel onset and maximum intensity over the first half of the vowel were calculated, and their difference was taken. Parabolic interpolation was applied to the intensity measurements in order to produce continuous estimates over the vowel. Parabolic interpolation was selected also due to its higher accuracy relative to linear interpolation and the relatively long intervals between intensity measurements (Boersma & Weenink 2018). Points on this continuous curve where the minimum intensity passed 10% and 90% of the difference between minimum and maximum intensity were extracted. The time from 10% to 90% of the change in intensity was recorded as rise time in intensity. The calculation for rise time of intensity is shown in Figure 1.



**Figure 1** (Colour online) A schematic representation of rise time calculation. The rise time (yellow double-ended arrow) is calculated by identifying the points at 10% and 90% of the difference between minimum and maximum intensity along the intensity curve and calculating the length of time between them.

## 2.5 Statistical analysis

Linear mixed effects models were used to examine significance of the different acoustic properties measured. All acoustic measurements were examined as dependent variables in these models, with phonemic labialization (labialized, plain) or vowel (/ɔ v a i/) as fixed effects, depending on the model, with random slopes and intercepts for participant and random intercepts for items; random effects structure was simplified only when required based on non-convergence or singular fit. This combination of random slopes and intercepts for random effects was selected due to the within-subjects but between-items nature of the experiment. Model selection was based on the keep-it-maximal principle (Barr et al. 2013), with random slopes and intercepts selected to represent the maximal model expected based on predictions. Labialized velars and following vowel /ɔ/ were set as the reference levels to directly examine how measurements in the neutralizing environment pattern relative to other environments.

All statistical models were created using R Studio (Rstudio Team 2020, R Core Team 2021). *P*-values were calculated in the package *lmerTest* in R (Kuznetsova, Brockhoff & Christensen 2017) using the Satterthwaite approximation, which Luke (2017) has shown to be among the most accurate methods of *p*-value calculation for linear mixed effects models, especially with a smaller number of subjects. While larger LMEs were initially calculated, interactions of labialization and vowel caused difficulty of interpretation of effects, so models were simplified to compare vowel contexts for labialized velars only and to compare labialized and plan velar onset conditions for each vowel separately. Results were visualized using *ggplot2* (Wickham 2016).

## 3 Results

The results of this experiment provide information about the extent to which productions of /k<sup>w</sup>ɔ/ sequences were neutralized by these speakers, both in terms of the data overall and in terms of the differences among participants. The following sections outline first the results as a whole, then provide an examination of the data across participants in order to better understand their variability.

### 3.1 Overall results

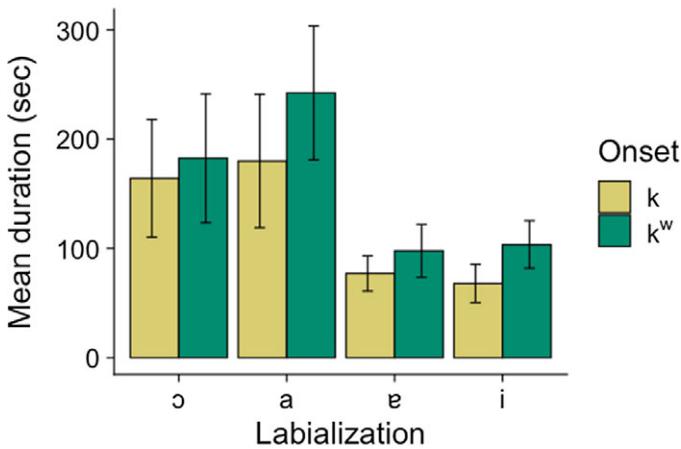
The overall results were used to examine which among the basket of relevant acoustic properties show significant differences in all environments, suggesting maintenance of contrast, and which were significantly different in all environments except the neutralizing environment.

This was done by comparing /k<sup>w</sup>/ to /k/ in each vowel environment (/ɔ̄ ɐ a i/) to determine whether there were significant differences in each environment. Further tests compared productions of /k<sup>w</sup>/ in each vowel environment to determine differences between environments.

### 3.1.1 Duration measurements

The differences in vowel duration for /k<sup>w</sup>/ and /k/ before each of the four vowels are shown in Figure 2. Duration is longer for vowels /ɔ̄/ and /a/ than for vowels /ɐ/ and /i/. Labialized velar onsets have numerically longer duration than plain velar onsets for all four vowels. There is higher variability for vowels /ɔ̄/ and /a/ than for vowels /ɐ/ and /i/.

The results of linear mixed effects models for duration, in Table 3, show a significant difference in values between labialized and plain velar onsets for vowels /a ɐ i/, but no significant



**Figure 2** (Colour online) Mean vowel duration for plain and labialized velar stops by following vowel.

**Table 3** Linear mixed effects model results for duration (ms); /k<sup>w</sup>/ and /ɔ̄/ are set as reference levels.

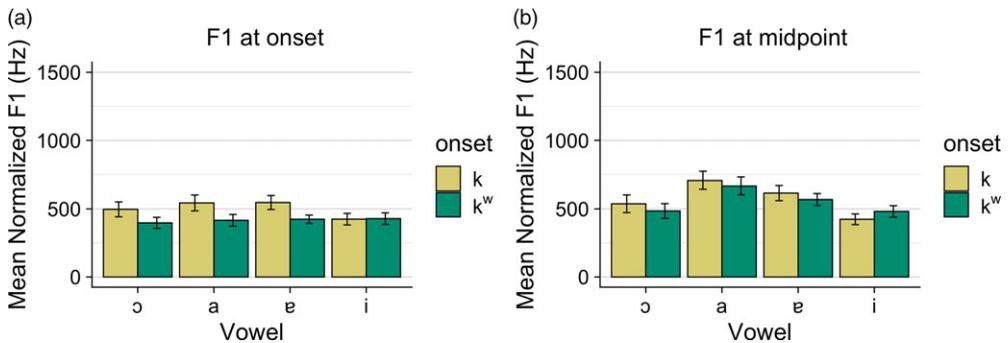
Vowel /ɔ̄/			
	est.	<i>t</i>	<i>p</i>
Intercept	182.762	18.994	<i>p</i> < .001***
/k/	-18.742	-1.657	<i>p</i> = .103
Vowel /a/			
Intercept	242.31	15.939	<i>p</i> < .001***
/k/	-62.36	-3.928	<i>p</i> < .001***
Vowel /ɐ/			
Intercept	97.810	15.053	<i>p</i> < .001***
/k/	-20.564	-3.122	<i>p</i> = .006**
Vowel /i/			
Intercept	103.858	15.316	<i>p</i> < .001***
/k/	-35.826	-5.431	<i>p</i> < .001***
All vowels, /k <sup>w</sup> / only			
Intercept	167.914	23.956	<i>p</i> < .001***
/a/	29.223	3.354	<i>p</i> = .001**
/ɐ/	-90.235	-7.250	<i>p</i> < .001***
/i/	-99.290	-9.932	<i>p</i> < .001***

\*\* *p* < .01; \*\*\* *p* < .001

difference between duration of the vowel /ɔ/ between /k<sup>w</sup>/ and /k/. There are furthermore significant differences in duration for stimuli in the /k<sup>w</sup>ɔ/ condition relative to labialized stimuli in each of the other vowel conditions. These results indicate that the speakers do not maintain the length distinction between /k<sup>w</sup>/ and /k/, instead showing signs of neutralization with the vowel /ɔ/.

### 3.1.2 F1 measurements

Figure 3 shows the mean F1 values at onset and midpoint of the following vowel (a and b, respectively). F1 values at vowel onset are higher for /k/ with vowels /ɔ/, /a/, and /ɐ/ than for /k<sup>w</sup>/, although F1 values for /i/ are the same for both onset consonants. In the /ɔ/ environment,



**Figure 3** (Colour online) Mean normalized F1 values (Hz) at onset (a) and midpoint (b) of the following vowel. Mean F1 values are higher for vowels /ɔ/, /a/, and /ɐ/ that follow plain velars at both onset and midpoint, though values for /i/ are lower after plain velar at midpoint. The value of F1 is fairly stable at onset regardless of the following vowel, although this trend has disappeared by vowel midpoint.

**Table 4** Linear mixed effects model results for F1 values at vowel onset (left) and midpoint (right); /k<sup>w</sup>/ and /ɔ/ are set as reference levels.

	F1 at vowel onset			F1 at vowel midpoint		
	Vowel /ɔ/					
	est.	<i>t</i>	<i>p</i>	est.	<i>t</i>	<i>p</i>
Intercept	397.45	63.51	<i>p</i> < .001***	484.68	47.773	<i>p</i> < .001***
/k/	99.228	10.51	<i>p</i> < .001***	52.90	5.186	<i>p</i> < .001***
	Vowel /a/					
Intercept	416.136	64.47	<i>p</i> < .001***	667.726	63.282	<i>p</i> < .001***
/k/	127.678	9.49	<i>p</i> < .001***	40.788	4.839	<i>p</i> < .001***
	Vowel /ɐ/					
Intercept	424.157	44.26	<i>p</i> < .001***	568.65	47.376	<i>p</i> < .001***
/k/	122.868	18.60	<i>p</i> < .001***	47.73	3.445	<i>p</i> = .003**
	Vowel /i/					
Intercept	428.257	55.194	<i>p</i> < .001***	482.00	39.120	<i>p</i> < .001***
/k/	-3.658	-0.581	<i>p</i> > .1	-58.34	-4.726	<i>p</i> < .001***
	All vowels, /k <sup>w</sup> / only					
Intercept	430.908	87.598	<i>p</i> < .001***	522.532	79.899	<i>p</i> < .001***
/a/	37.452	6.926	<i>p</i> < .001***	178.734	27.474	<i>p</i> < .001***
/ɐ/	45.348	5.842	<i>p</i> < .001***	86.281	9.253	<i>p</i> < .001***
/i/	-57.680	-9.056	<i>p</i> < .001***	-90.259	-11.871	<i>p</i> < .001***

\*\**p* < .01; \*\*\**p* < .001

mean F1 is slightly lower after both types of velars than values for /a/ and /ɐ/. F1 values for all four vowels in both velar conditions are higher at vowel midpoint, although there is the least change for /ki/ sequences. At vowel midpoint, there is a clear difference between /k<sup>w</sup>/ and /k/ for vowel /ɔ/. To determine whether the F1 values for vowels following /k<sup>w</sup>/ and /k/ are significantly different in each vowel condition, linear mixed results models were calculated for each.

Linear mixed effect models for F1 values are shown in Table 4. Results of these models reveal significant differences between labialized and plain velar onsets in both the neutralization environment and other vowel environments. At vowel onset, there is no significant difference in F1 values between labialized and plain velar onsets preceding vowel /i/, although there is a significant difference for vowels /ɐ/, /a/, and /ɔ/. This is consistent with the observation by Stevens (1998) that F1 is not affected by lip rounding for a high vowel such as /i/. There is furthermore a significant difference between F1 values for words beginning with /k<sup>w</sup>ɔ/ and words beginning with /k<sup>w</sup>/ in each of the other vowel conditions. At vowel midpoint, mean F1 values are significantly different between plain and labialized onsets followed by all four vowels. As at vowel onset, there is again a significant difference between F1 values for /k<sup>w</sup>ɔ/ items and items in other /k<sup>w</sup>V/ conditions. As all vowels show significant differences in F1 values between labialized and plain velar onset conditions, these results suggest that this group of speakers maintains the F1 distinction between /k<sup>w</sup>ɔ/ and /kɔ/.

### 3.1.3 F2 measurements

Figure 4 shows the mean F2 values at onset and midpoint of the following vowel (a and b, respectively). At onset of the following vowel, F2 after labialized velars is consistently lower than after plain velars, which is consistent with literature on the effects of rounding on F2 (Stevens 1998). Plain velars show varied F2 values according to vowel, which is expected and reflects the difference in quality between these vowels.

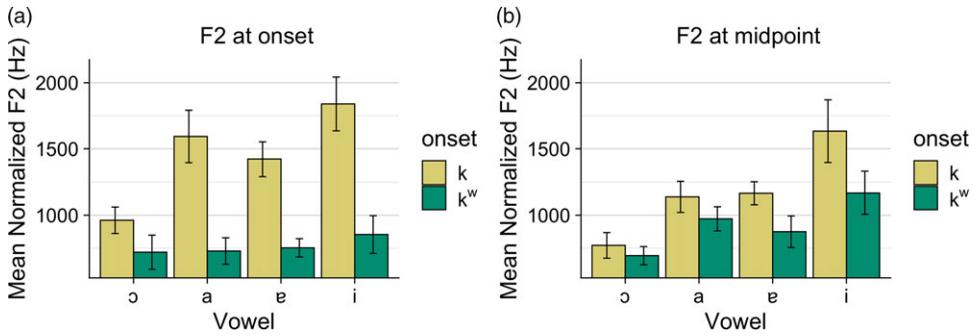
F2 values in the sequence /k<sup>w</sup>i/ are slightly higher than for the other three vowels. The difference in F2 after plain and labialized velars is much smaller for vowel /ɔ/ than for the other vowels. By vowel midpoint, F2 has increased for all vowels following labialized velars and has decreased for all vowels following plain velars. At vowel midpoint, the neutralizing environment still shows the smallest difference in F2 relative to the same following vowel after a plain velar.

Linear mixed effects models for F2 values at onset and midpoint (in Table 5) of the following vowel show that, despite the smallest differences in F2 values being between /kɔ/ and /k<sup>w</sup>ɔ/, these differences are still significant at both vowel onset and vowel midpoint. This observation is true before all unrounded vowels, as well. There is furthermore a significant difference among F2 values in each vowel condition with onset /k<sup>w</sup>/. As with F1, results of these models suggest that this group of speakers as a whole maintains the F2 distinction between labialized and plain velar onsets in the rounded vowel /ɔ/ environment. The preservation of this distinction may be due to the use of F2 as a major cue to labialization in languages of the world (Stevens 1998, Sturman et al. 2016).

### 3.1.4 F3 measurements

Sturman et al. (2016) indicate that labialization as a secondary articulation frequently results in lower F3 values than for non-labialized counterparts. However, the opposite appears true for the neutralizing environment in Hong Kong Cantonese: Figure 5 demonstrates that, with the exception of vowel /i/, F3 values at onset are numerically slightly higher for vowels following labialized velars than for those following plain velars. The difference between the means is smallest for vowel /ɔ/, however.

Table 6 reveals that at vowel onset, F3 values for /k<sup>w</sup>ɔ/ and /kɔ/ are not significantly different; however, F3 values between labialized and plain onsets are significantly different



**Figure 4** (Colour online) Mean normalized F2 values (Hz) at onset (a) and midpoint (b) of the following vowel. Mean F2 values at onset are similar for all four vowels after a labialized velar, although the decrease in F2 is smallest in the neutralizing environment. By vowel midpoint, the difference between F2 values between velar phonemes is still smallest in the neutralizing environment.

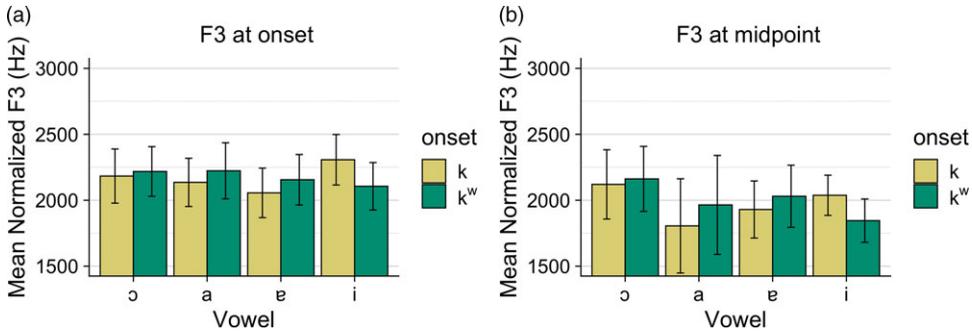
**Table 5** Linear mixed effects model results for F2 values at vowel onset (left) and midpoint (right); /kʷ/ and /ɔ/ are set as reference levels.

	F2 at vowel onset			F2 at vowel midpoint		
	est.	<i>t</i>	<i>p</i>	est.	<i>t</i>	<i>p</i>
Vowel /ɔ/						
Intercept	719.76	59.65	<i>p</i> < .001***	693.86	50.763	<i>p</i> < .001***
/k/	241.79	10.58	<i>p</i> < .001***	77.70	4.916	<i>p</i> < .001***
Vowel /a/						
Intercept	728.52	56.02	<i>p</i> < .001***	972.72	62.65	<i>p</i> < .001***
/k/	864.71	33.05	<i>p</i> < .001***	166.33	9.916	<i>p</i> < .001***
Vowel /e/						
Intercept	752.90	33.68	<i>p</i> < .001***	874.99	42.941	<i>p</i> < .001***
/k/	669.53	37.09	<i>p</i> < .001***	291.49	9.916	<i>p</i> < .001***
Vowel /i/						
Intercept	853.86	24.76	<i>p</i> < .001***	1166.40	29.99	<i>p</i> < .001***
/k/	985.46	22.44	<i>p</i> < .001***	466.63	12.83	<i>p</i> < .001***
All vowels, /kʷ/ only						
Intercept	640.13	26.74	<i>p</i> < .001***	728.09	64.51	<i>p</i> < .001***
/a/	376.25	11.51	<i>p</i> < .001***	327.28	22.59	<i>p</i> < .001***
/e/	295.56	6.30	<i>p</i> < .001***	328.48	15.80	<i>p</i> < .001***
/i/	688.00	17.90	<i>p</i> < .001***	788.31	46.27	<i>p</i> < .001***

\*\*\**p* < .001

in all other vowel conditions. As before, there is a significant difference in values between each of the four vowels in the labialized velar onset condition. These model results provide the first evidence that these speakers engage in neutralization of some acoustic properties examined here, as the difference in F3 is neutralized in the /ɔ/ condition but not in other vowel environments.

The pattern in F3 values remains the same at vowel midpoint. Labialized onsets are not associated with a significantly different F3 value from plain velar onsets in the /ɔ/ vowel condition, although they are significantly different in all other vowel environments. This result indicates that speakers show signs of neutralization in F3 values at both vowel onset and vowel midpoint.



**Figure 5** (Colour online) Mean normalized F3 values (Hz) at onset (a) and midpoint (b) of the following vowel. F3 values for vowels /ɔ/, /a/, and /ɐ/ following labialized velars are slightly higher than for the same vowels following plain velars at both onset and midpoint. F3 values pattern similarly for these three vowels at onset, although values are higher for /ɔ/ by vowel midpoint.

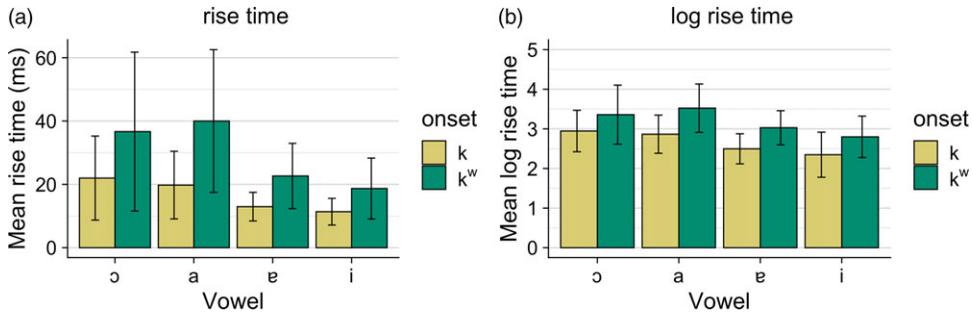
**Table 6** Linear mixed effects model results for F3 values at vowel onset (left) and midpoint (right); /k<sup>w</sup>/ and /ɔ/ are set as reference levels.

	F3 at vowel onset			F3 at vowel midpoint		
	Vowel /ɔ/					
	est.	<i>t</i>	<i>p</i>	est.	<i>t</i>	<i>p</i>
Intercept	2217.47	65.252	<i>p</i> < .001***	2160.43	41.00	<i>p</i> < .001***
/k/	-34.18	-1.088	<i>p</i> = 0.294	-40.08	-1.75	<i>p</i> = 0.090
	Vowel /a/					
	est.	<i>t</i>	<i>p</i>	est.	<i>t</i>	<i>p</i>
Intercept	2223.25	48.944	<i>p</i> < .001***	1964.56	28.11	<i>p</i> < .001***
/k/	-87.71	-2.241	<i>p</i> = 0.042*	-159.17	-2.36	<i>p</i> = 0.033*
	Vowel /ɐ/					
	est.	<i>t</i>	<i>p</i>	est.	<i>t</i>	<i>p</i>
Intercept	2155.43	55.576	<i>p</i> < .001***	2030.25	38.273	<i>p</i> < .001***
/k/	202.24	-2.703	<i>p</i> = 0.017*	-99.673	-2.267	<i>p</i> = 0.047*
	Vowel /i/					
	est.	<i>t</i>	<i>p</i>	est.	<i>t</i>	<i>p</i>
Intercept	2104.52	52.232	<i>p</i> < .001***	1844.15	40.556	<i>p</i> < .001***
/k/	202.24	4.548	<i>p</i> < .001***	183.41	4.632	<i>p</i> < .001***
	All vowels, /k <sup>w</sup> / only					
	est.	<i>t</i>	<i>p</i>	est.	<i>t</i>	<i>p</i>
Intercept	2183.25	95.778	<i>p</i> < .001***	2104.57	61.999	<i>p</i> < .001***
/a/	-29.65	-2.495	<i>p</i> = 0.014*	-272.12	-15.685	<i>p</i> < .001***
/ɐ/	-112.79	-6.611	<i>p</i> < .001***	-174.71	-7.017	<i>p</i> < .001***
/i/	91.07	6.521	<i>p</i> < .001***	-111.67	-5.484	<i>p</i> < .001***

\**p* < .05; \*\**p* < .01; \*\*\**p* < .001

### 3.1.5 Rise time of intensity

Rise time of intensity was measured across the first half of each vowel in target stimuli. Figure 6 shows bar charts of non-normalized rise time in milliseconds (a) and log rise times (b) for vowels after labialized and plain velars. In (a), mean rise time is longer after labialized velars than plain velars, with a large amount of variability for vowels /ɔ/ and /a/. The variability in /k<sup>w</sup>ɔ/ may be indicative that rise time is altered in instances of incomplete neutralization, where tokens with short rise times are more likely to be produced as plain velars. The log-transformed rise time measures in (b) demonstrate that this transformation has greatly reduced variability. Comparing the pattern of rise times across conditions for



**Figure 6** (Colour online) Mean rise time (a) and log rise time (b) of intensity on vowels following plain and labialized velars. Rise time is longer after labialized velars than plain velars, with much more variability in rise time for vowels /ɔ/ and /a/. Higher values in (b) correspond to longer rise time, while numbers closer to zero indicate shorter rise time.

**Table 7** Linear mixed effects model results for rise time (log); /kʷ/ and /ɔ/ are set as reference levels.

Vowel /ɔ/			
	est.	t	p
Intercept	3.3576	32.257	p < .001***
/k/	-0.412	-3.566	p = .002**
Vowel /a/			
Intercept	3.522	30.597	p < .001***
/k/	-0.657	-5.861	p < .001***
Vowel /e/			
Intercept	3.029	32.313	p < .001***
/k/	-0.530	-6.276	p < .001***
Vowel /i/			
Intercept	2.797	21.247	p < .001***
/k/	-0.449	-3.359	p = .004**
All vowels, /kʷ/ only			
Intercept	2.959	56.482	p < .001***
/a/	0.013	0.289	p = .773
/e/	-0.407	-6.269	p < .001***
/i/	-0.576	-10.834	p < .001***

\*\*p < .01; \*\*\*p < .001

each vowel in (b) to those in (a), the same patterns are evident in the log-transformed data that existed in the original data.

A linear mixed effects model was used to analyze the effects of vowel and labialization on log rise time of intensity. Log rise time was used because the residuals for untransformed rise time demonstrated violations of heteroskedasticity and clear deviations from normal distribution on a quantile-quantile plot. Log-transforming the data resulted in residuals that no longer violated the assumption of normal distribution of residuals.

Results of the log-transformed linear mixed effects model, presented in Table 7, demonstrate that there is a significant difference in rise time values for /kʷ/ and /k/ across all vowel conditions. For items beginning with /kʷ/, there is no significant difference between the /ɔ/ and /a/ conditions, although /ɔ/ differs significantly from /i/ and /e/. As with the results of F1 and F2 values at onset and midpoint, these results suggest that this group of speakers maintains the distinction in rise time in the /ɔ/ environment. Because there is a large amount of

variability in rise time in the /ɔ/ environment, however, there may be individual differences in this measure. As such, it is important to consider the systematic variability across speakers in the study.

### 3.1.6 Summary of overall results

The overall results demonstrate that this group of speakers overall continues to maintain distinction between /k<sup>w</sup>ɔ/ and /kɔ/ in terms of F1 values, F2 values, and rise time; however, there are signs of neutralization in terms of vowel duration and F3 values. There is furthermore a large degree of variability in both duration and rise time, indicating that there may be substantial individual differences in the maintenance or collapsing of these acoustic measures in the neutralization environment.

## 3.2 Results by participants

While the overall results showed overall maintenance of the /k<sup>w</sup>ɔ/–/kɔ/ contrast in terms of F1, F2 and rise time and loss of distinction for vowel duration and F3, there may be salient individual differences in the data. This section examines the interspeaker differences in maintenance and loss of distinction for each of the individual acoustic measures seen above. This examination allows for further insight into the contrast maintenance or incomplete neutralization by this group of speakers. To statistically examine these results, two-sample Welch's *t*-tests were performed to compare acoustic measure values for /k<sup>w</sup>ɔ/–/kɔ/ for each speaker separately. Results of these tests appear in Table 8.

### 3.2.1 Vowel duration

Measures of vowel duration by participant for /kɔ/ and /k<sup>w</sup>ɔ/ can be seen in Figure 7. Duration measurements differed significantly between the two onset conditions for speakers f1, f4, f5, and m5. For each of these speakers, duration of /kɔ/ was significantly shorter than duration of /k<sup>w</sup>ɔ/. For the remaining ten speakers, there was no significant difference between duration of the vowel in /kɔ/ and /k<sup>w</sup>ɔ/.

Interestingly, while half of the female speakers maintained the distinction in duration between /kɔ/ and /k<sup>w</sup>ɔ/, only one out of eight male speakers did so. Taken together, the results of duration by participant suggest that, although overall results indicate a loss of distinction in duration between /kɔ/ and /k<sup>w</sup>ɔ/, this contrast may be maintained by a subset of speakers.

### 3.2.2 F1 at vowel onset and midpoint

F1 values show the clearest signs of distinction of maintenance out of all the acoustic measures analyzed. Every speaker realized the distinction between /kɔ/ and /k<sup>w</sup>ɔ/ through F1. The normalized F1 values are shown in Figure 8. This figure demonstrates that at vowel onset, normalized F1 values for /kɔ/ are significantly higher than those of /k<sup>w</sup>ɔ/, with a very clear distinction between these values. Speaker m6 shows the closest mean normalized values for F1 between /kɔ/ and /k<sup>w</sup>ɔ/, although there is still a distinct separation between the two. At vowel midpoint, mean normalized F1 values are still significantly higher for /kɔ/ than for /k<sup>w</sup>ɔ/ for 13 of the 14 speakers, although by the point F1 values are much closer together. At vowel midpoint, speaker m1 shows the closest mean F1 values between the two onset conditions, and indeed is the only speaker whose normalized F1 values for /kɔ/ are not significantly higher than those of /k<sup>w</sup>ɔ/.

**Table 8** Results of two sample Welch's *t*-tests comparing acoustic measure values between /kʷɔ/ and /kɔ/ items for each speaker. Bold marks significant results.

Speaker	F1 onset	F2 onset	F3 onset	F1 midpoint	F2 midpoint	F3 midpoint	Rise time	Duration
f1	<i>t</i> (52.88) = -5.26 <b><i>p</i> &lt; .001***</b>	<i>t</i> (50.65) = -16.50 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.40) = -0.85 <i>p</i> = .4	<i>t</i> (56.48) = -1.92 <i>p</i> = .06	<i>t</i> (52.88) = -5.26 <b><i>p</i> &lt; .001***</b>	<i>t</i> (51.32) = 0.36 <i>p</i> = .72	<i>t</i> (50.81) = -1.57 <i>p</i> = .12	<i>t</i> (56.39) = 2.01 <b><i>p</i> = .0495*</b>
f2	<i>t</i> (57.91) = -15.49 <b><i>p</i> &lt; .001***</b>	<i>t</i> (55.64) = -22.89 <b><i>p</i> &lt; .001***</b>	<i>t</i> (56.10) = 9.61 <b><i>p</i> &lt; .001***</b>	<i>t</i> (54.41) = -5.68 <b><i>p</i> &lt; .001***</b>	<i>t</i> (53.52) = -5.49 <b><i>p</i> &lt; .001***</b>	<i>t</i> (54.13) = 3.50 <b><i>p</i> &lt; .001***</b>	<i>t</i> (43.35) = 0.88 <i>p</i> = .39	<i>t</i> (55.81) = 0.28 <i>p</i> = .78
f3	<i>t</i> (52.62) = -9.43 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.59) = -15.15 <b><i>p</i> &lt; .001***</b>	<i>t</i> (58.84) = 4.57 <b><i>p</i> &lt; .001***</b>	<i>t</i> (58.84) = -3.71 <b><i>p</i> &lt; .001***</b>	<i>t</i> (55.43) = -4.10 <b><i>p</i> &lt; .001***</b>	<i>t</i> (54.52) = 1.67 <i>p</i> = .10	<i>t</i> (49.79) = 1.99 <i>p</i> = .05	<i>t</i> (55.71) = 0.78 <i>p</i> = .44
f4	<i>t</i> (57.37) = -9.56 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.73) = -21.19 <b><i>p</i> &lt; .001***</b>	<i>t</i> (56.71) = 2.64 <i>p</i> = .01*	<i>t</i> (53.81) = -5.27 <b><i>p</i> &lt; .001***</b>	<i>t</i> (56.65) = -7.66 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.87) = 2.42 <b><i>p</i> = .02*</b>	<i>t</i> (46.73) = 2.03 <b><i>p</i> = .048*</b>	<i>t</i> (55.61) = 2.69 <b><i>p</i> = .009**</b>
f5	<i>t</i> (58.40) = -10.54 <b><i>p</i> &lt; .001***</b>	<i>t</i> (48.38) = -10.89 <b><i>p</i> &lt; .001***</b>	<i>t</i> (58.74) = -2.33 <b><i>p</i> = .02*</b>	<i>t</i> (58.85) = -4.72 <b><i>p</i> &lt; .001***</b>	<i>t</i> (55.32) = -4.79 <b><i>p</i> &lt; .001***</b>	<i>t</i> (58.99) = 0.60 <i>p</i> = .55	<i>t</i> (58.35) = 1.23 <i>p</i> = .23	<i>t</i> (56.48) = 2.41 <b><i>p</i> = .02*</b>
f6	<i>t</i> (58.96) = -9.87 <b><i>p</i> &lt; .001***</b>	<i>t</i> (51.76) = -11.00 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.82) = 0.20 <i>p</i> = .84	<i>t</i> (57.67) = -3.37 <b><i>p</i> = .001**</b>	<i>t</i> (58.31) = -3.82 <b><i>p</i> &lt; .001***</b>	<i>t</i> (52.67) = -0.52 <i>p</i> = .61	<i>t</i> (54.24) = -2.06 <b><i>p</i> = .04*</b>	<i>t</i> (54.15) = 0.06 <i>p</i> = .96
m1	<i>t</i> (52.89) = -4.03 <b><i>p</i> &lt; .001***</b>	<i>t</i> (33.08) = -1.74 <i>p</i> = .09	<i>t</i> (45.43) = -1.60 <i>p</i> = .12	<i>t</i> (45.67) = -1.09 <i>p</i> = .28	<i>t</i> (46.81) = -1.93 <i>p</i> = .06	<i>t</i> (57.44) = -0.60 <i>p</i> = .55	<i>t</i> (49.10) = 5.11 <b><i>p</i> &lt; .001***</b>	<i>t</i> (56.04) = 0.84 <i>p</i> = .40
m2	<i>t</i> (37.54) = -9.62 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.40) = -18.79 <b><i>p</i> &lt; .001***</b>	<i>t</i> (40.89) = 0.63 <i>p</i> = .54	<i>t</i> (48.45) = -6.73 <b><i>p</i> &lt; .001***</b>	<i>t</i> (52.57) = -7.05 <b><i>p</i> &lt; .001***</b>	<i>t</i> (59.00) = 2.18 <b><i>p</i> = .03*</b>	<i>t</i> (58.60) = 4.95 <b><i>p</i> &lt; .001***</b>	<i>t</i> (55.30) = 1.06 <i>p</i> = .30
m3	<i>t</i> (55.19) = -10.72 <b><i>p</i> &lt; .001***</b>	<i>t</i> (52.16) = -12.84 <b><i>p</i> &lt; .001***</b>	<i>t</i> (55.66) = 2.98 <b><i>p</i> = .004**</b>	<i>t</i> (46.78) = -3.02 <b><i>p</i> = .004**</b>	<i>t</i> (56.73) = -2.90 <b><i>p</i> = .005**</b>	<i>t</i> (57.62) = -0.37 <i>p</i> = .71	<i>t</i> (55.54) = 7.25 <b><i>p</i> &lt; .001***</b>	<i>t</i> (53.33) = 1.99 <i>p</i> = .05
m4	<i>t</i> (35.14) = -8.20 <b><i>p</i> &lt; .001***</b>	<i>t</i> (34.82) = -4.23 <b><i>p</i> &lt; .001***</b>	<i>t</i> (38.49) = 5.36 <b><i>p</i> &lt; .001***</b>	<i>t</i> (47.85) = -6.40 <b><i>p</i> &lt; .001***</b>	<i>t</i> (37.45) = -6.10 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.15) = 4.47 <b><i>p</i> &lt; .001***</b>	<i>t</i> (56.56) = 1.98 <i>p</i> = .05	<i>t</i> (56.07) = 0.92 <i>p</i> = .36
m5	<i>t</i> (46.14) = -7.98 <b><i>p</i> &lt; .001***</b>	<i>t</i> (36.72) = -3.76 <b><i>p</i> &lt; .001***</b>	<i>t</i> (58.91) = -1.32 <i>p</i> = .19	<i>t</i> (58.52) = -3.27 <b><i>p</i> = .002**</b>	<i>t</i> (58.92) = -2.00 <b><i>p</i> = .0496*</b>	<i>t</i> (58.95) = 0.09 <i>p</i> = .93	<i>t</i> (46.84) = 6.32 <b><i>p</i> &lt; .001***</b>	<i>t</i> (55.80) = 2.60 <b><i>p</i> = .01*</b>
m6	<i>t</i> (50.84) = -8.64 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.54) = -15.14 <b><i>p</i> &lt; .001***</b>	<i>t</i> (52.69) = 3.53 <b><i>p</i> &lt; .001***</b>	<i>t</i> (54.85) = -3.75 <b><i>p</i> &lt; .001***</b>	<i>t</i> (58.43) = -2.60 <b><i>p</i> = .01*</b>	<i>t</i> (45.64) = 1.87 <i>p</i> = .07	<i>t</i> (57.70) = 3.48 <b><i>p</i> &lt; .001***</b>	<i>t</i> (58.82) = 1.07 <i>p</i> = .29
m7	<i>t</i> (43.69) = -21.02 <b><i>p</i> &lt; .001***</b>	<i>t</i> (36.70) = -14.06 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.98) = 2.68 <b><i>p</i> = .0096**</b>	<i>t</i> (47.46) = -4.98 <b><i>p</i> &lt; .001***</b>	<i>t</i> (40.35) = -4.23 <b><i>p</i> &lt; .001***</b>	<i>t</i> (49.13) = -0.25 <i>p</i> = .80	<i>t</i> (44.00) = 6.02 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.14) = 1.36 <i>p</i> = .18
m8	<i>t</i> (52.68) = -7.01 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.62) = -7.58 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.14) = -3.12 <b><i>p</i> = .003**</b>	<i>t</i> (57.93) = -3.96 <b><i>p</i> &lt; .001***</b>	<i>t</i> (56.36) = -2.80 <b><i>p</i> = .007**</b>	<i>t</i> (56.11) = -0.66 <i>p</i> = .52	<i>t</i> (54.26) = 3.67 <b><i>p</i> &lt; .001***</b>	<i>t</i> (57.60) = 1.76 <i>p</i> = .08

\**p* < .05; \*\**p* < .01; \*\*\**p* < .001

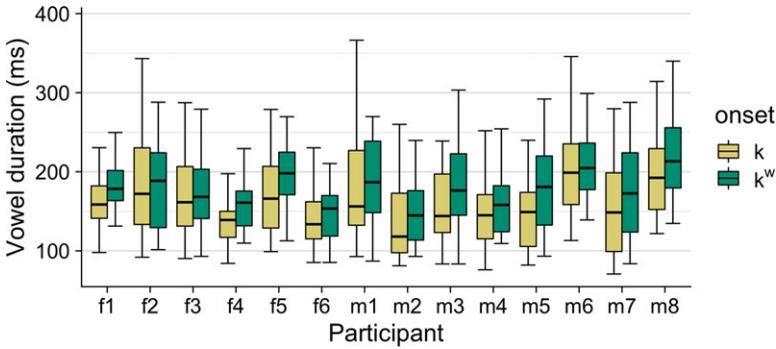


Figure 7 (Colour online) Vowel duration by participant for vowel /ɔ/.

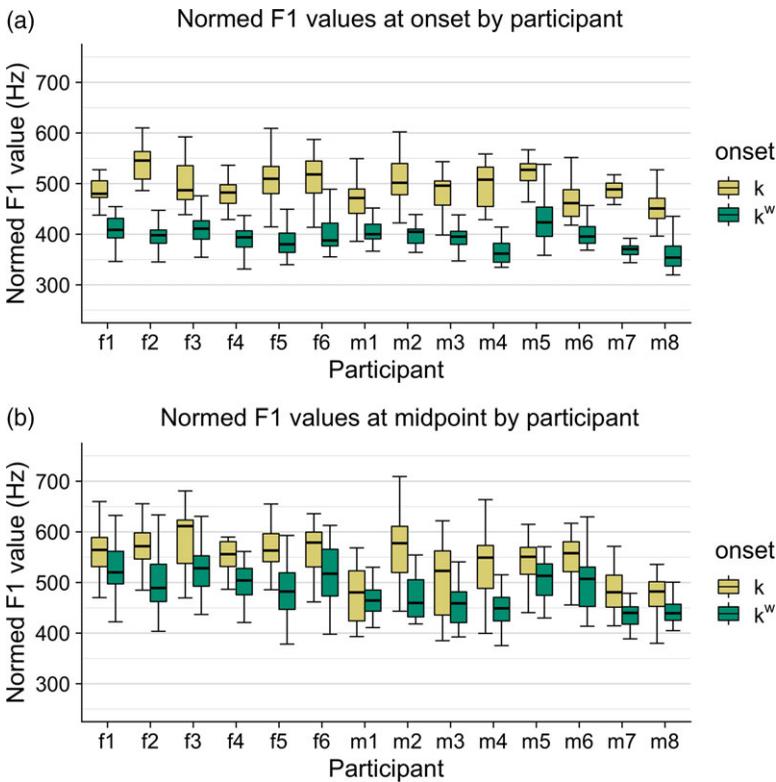
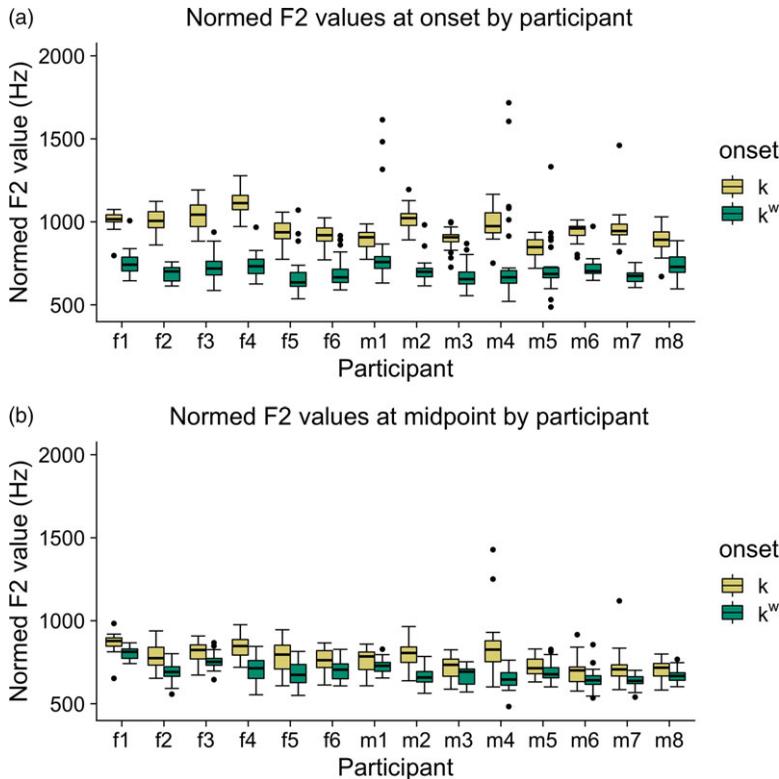


Figure 8 (Colour online) Normalized F1 values at vowel onset and midpoint by participant for vowel /ɔ/.

### 3.2.3 F2 at vowel onset and midpoint

Normalized F2 values at vowel onset and midpoint show a similar pattern to F1 values, as can be seen in Figure 9. At vowel onset, F2 values for /kɔ/ are significantly higher than those of /kʷɔ/ for 13 of the 14 speakers. As above, speaker m1 is the only participant whose values do not differ significantly between onset conditions. At vowel midpoint, F2 values for /kɔ/ and /kʷɔ/ are much closer together, although still significantly higher for /kɔ/ for all participants except m1.



**Figure 9** (Colour online) Normalized F2 values at vowel onset and midpoint by participant for vowel /ɔ/.

### 3.2.4 F3 at vowel onset and midpoint

In Section 3.1.4 above, an overall loss of distinction was shown for F3 values at both vowel onset and vowel midpoint. The comparisons of this measure across participants are shown in Figure 10. At vowel onset, normalized F3 values are significantly lower in /kɔ/ than in /kʷɔ/ for speakers f2, f3, f4, m3, m4, m6, and m7, significantly higher for speakers f5 and m8, and not significantly different for speakers f1, f6, m1, m2, and m5. These results suggest that, while there is loss of distinction of F3 at vowel onset for some speakers others maintain this contrast.

At vowel midpoint, results are likewise mixed. Speakers f2, f4, m2, and m4 have significantly lower normalized F3 values in /kɔ/ than in /kʷɔ/, while the other ten speakers have no significant difference in F3 values. As with F3 at vowel onset, these results suggest that some speakers maintain the F3 contrast between /kɔ/ and /kʷɔ/, while others show loss of distinction for F3 at vowel midpoint.

### 3.2.5 Rise time of intensity

Results by participant for rise time show that, once again, some speakers maintain the rise time contrast between /kɔ/ and /kʷɔ/ whereas others demonstrate a loss of this contrast, as shown in Figure 11. For log rise time, participants f4, f6, m1, m2, m3, m5, m6, m7, and m8 have a significant difference between rise times for /kɔ/ and /kʷɔ/. The other speakers do not show a difference in rise time values. Interestingly, rise time shows a pattern opposite to that of vowel duration in terms of the gender of speakers who show loss of distinction: four out of

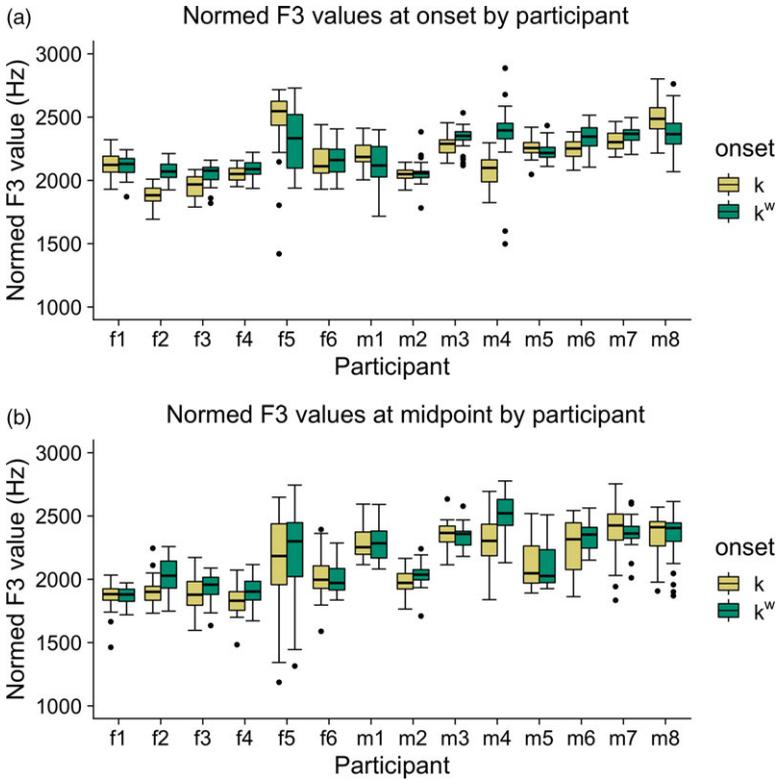


Figure 10 (Colour online) Normalized F3 values at vowel onset and midpoint by participant for vowel /ɔ/.

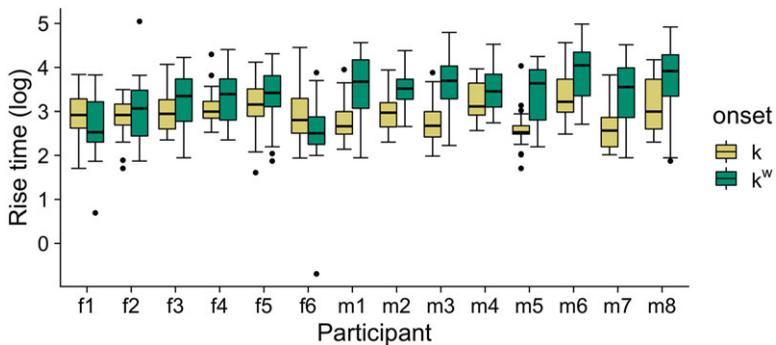


Figure 11 (Colour online) Log rise time values by participant for vowel /ɔ/.

six female participants show no significant difference in rise time values between onset conditions, whereas the same pattern holds for only one of eight male participants. Nevertheless, the contrast in these two properties of labialization does not appear in a complementary distribution, so this observation may be an artefact of the data rather than representative of a systematic difference in the realization of the /kɔ/–/k<sup>w</sup>ɔ/ contrast.

### 3.2.6 Summary of results by participants

Examining the results of acoustic measures by participants for /kɔ/ and /k<sup>w</sup>ɔ/ demonstrates that there is a large degree of maintenance of the contrast between the two in terms of acoustic

properties, although the only distinction maintained universally by these speakers was for F1 at vowel onset. While one speaker, f4, maintained the distinction between /kɔ/ and /k<sup>w</sup>ɔ/ for all acoustic measures, speaker m1 showed a loss of distinction in six of the eight acoustic measures examined. The variation in contrast maintenance suggests that, within this group of speakers, there are signs of incomplete neutralization.

Interestingly, there did not appear to be a relationship between articulation rate and maintenance or collapsing of contrast in specific acoustic measures. Both m4, who had the lowest articulation rate, and m3, who had the highest, preserved the distinction between /k<sup>w</sup>ɔ/ and /kɔ/ for six of the eight acoustic properties examined. M1, on the other hand, who only maintained the distinction between /k<sup>w</sup>ɔ/ and /kɔ/ in F1 at onset of vowel and rise time, had an articulation rate of 4.13 syllables per second of phonation, just below the average of 4.41 syllables per second. Overall, this suggests that the participants were speaking at a similar enough rate to not have a different impact of articulation rate on production of the phonetic properties examined here.

## 4 Discussion

This study examined the extent to which speakers neutralize /k<sup>w</sup>/ and /k/ before the rounded back vowel /ɔ/ in Hong Kong Cantonese. The results demonstrate that speakers of Hong Kong Cantonese show maintenance of the contrast between the two to different extents with regard to different acoustic measures. Speakers demonstrate an overall loss of distinction in F3 values at vowel onset and vowel midpoint, as well as vowel duration, but maintenance of the contrast in F1 and F2 values at vowel onset and vowel midpoint, as well as rise time of intensity. These results suggest that speakers produce /k<sup>w</sup>/ with less labialization in the /ɔ/ environment, and equally that the point of maximum constriction on the labial articulation in the /k<sup>w</sup>ɔ/ environment is equivalent to the point of maximum constriction from coarticulatory labialization present on the /k/ in the /kɔ/ environment. The results furthermore demonstrate that, while some speakers maintain differences in all or most of the acoustic properties of the onset velar, others produce incompletely neutralized versions of /k<sup>w</sup>/, where many or most acoustic measures are not significantly different between /k<sup>w</sup>/ and /k/.

While F1 and F2 values are consistent with a distinction between labialized and plain velars characterized by lip rounding, F3 values do not show the same pattern. The loss of distinction in F3 values is indicative of a lower degree of articulatory difference between /k<sup>w</sup>ɔ/ and /kɔ/ than other /k<sup>w</sup>/–/k/ pairs, likely due to the lip rounding required for the vowel /ɔ/. Further work would benefit from examining this distinction in terms of articulation through analysis of labial positioning during production of each /k<sup>w</sup>/–/k/ pair.

Analysis of rise time of intensity and vowel duration measures reveals that this group of speakers does not differentiate /k<sup>w</sup>ɔ/ from /kɔ/ in terms of vowel duration, and that many individual speakers do not differentiate the two sounds in terms of rise time of intensity, although these measures for the /k<sup>w</sup>/–/k/ pair are significantly different in other vowel environments. This may be caused by an articulatory difference in the /k<sup>w</sup>ɔ/ environment, where the point of maximal labial constriction occurs closer to simultaneously with the point of velar closure than in other /k<sup>w</sup>V/ environments. As with the above observations, an articulatory study examining the timing of labial and velar constriction would shed light on whether this possibility accounts for the differences in acoustic properties identified in this study.

The results of this study provide evidence that many speakers of Hong Kong Cantonese may still maintain the distinction between /k<sup>w</sup>/ and /k/ before /ɔ/ in production, a finding that is somewhat different from previous work on the topic (Bauer 1982). As in previous work, there is individual variation in production, although in this study there are signs that neutralization is incomplete even for the speakers who maintain the fewest phonetic contrasts between the two phonemes.

For future research, the use of lexical frequency as a manipulated factor would be beneficial. If F3 and vowel length demonstrate consistently different patterns of production in the neutralizing environment across groups, speech styles, and levels of frequency, while others do not, it is likely that the /k<sup>w</sup>ɔ̄/–/kɔ̄/ distinction is maintained more frequently than shown by previous studies. If, on the other hand, other conditions lead to complete neutralization of the distinction between /k<sup>w</sup>ɔ̄/ and /kɔ̄/, then this distinction would continue to be made only by certain speakers in specific circumstances, which is consistent with (Bauer 1982).

Additional acoustic analysis of this sound change could examine the patterns of these acoustic properties across sociolinguistic groups and registers in Hong Kong. As we have seen differences among participants in terms of their productions of the /k<sup>w</sup>ɔ̄/–/kɔ̄/ distinction, future research may wish to further explore variation in this phenomenon. One good way of examining these differences is through an understanding of how the contextual neutralization of /k<sup>w</sup>ɔ̄/ diffuses through the population, including a mapping of the phenomenon through social groups.

## Acknowledgements

We would like to acknowledge the help and support of Natasha Warner, Mike Hammond, Jonathan Yip, Diana Archangeli, and Paul Boersma, as well as *JIPA* Associate Editor Alexei Kochetov and two anonymous reviewers for their helpful feedback.

## Appendix. List of target and filler stimuli

	Target		Filler	
	Cantonese	English	Cantonese	English
1	/k <sup>w</sup> ɔ̄55/ 戈	lance	/tiŋ23.pət2/ 挺拔	tall and straight
2	/k <sup>w</sup> ɔ̄25/ 裹	bind	/jyn21.pət5/ 完畢	finish
3	/k <sup>w</sup> ɔ̄25/ 果	fruit	/fɛn25.pət5/ 粉筆	chalk
4	/k <sup>w</sup> ɔ̄33/ 過	pass	/lap2.pət5/ 蠟筆	crayon
5	/k <sup>w</sup> ɔ̄33.paw25/ 過飽	overeat	/mɛk2.sɔ̄y25.pət5/ 墨水筆	ink pen
6	/k <sup>w</sup> ɔ̄33.taj22/ 過大	oversized	/tɛj22.pət2/ 提拔	promote
7	/k <sup>w</sup> ɔ̄33.tɔ̄w22/ 過度	excessive	/pət5/ 筆	pen
8	/k <sup>w</sup> ɔ̄33.tuŋ55/ 過冬	hibernation	/pət5/ 不	not
9	/k <sup>w</sup> ɔ̄33.sɛn55/ 過身	pass away	/pət5/ 畢	finish
10	/k <sup>w</sup> ɔ̄25.jɛn21/ 果仁	nut	/pət2/ 拔	pull out
11	/k <sup>w</sup> ɔ̄k3/ 國	country	/hɔ̄w25.tət2/ 好突	very abrupt
12	/k <sup>w</sup> ɔ̄k3/ 幘	feminine	/sɛj25.tət2/ 使凸	overspend
13	/k <sup>w</sup> ɔ̄k3/ 郭	surname Kwok	/ts <sup>h</sup> uŋ55.tət2/ 衝突	conflict
14	/k <sup>w</sup> ɔ̄k3.pɔ̄w25/ 國寶	national treasures	/t <sup>h</sup> ɔ̄ŋ21.tət2/ 唐突	rude
15	/k <sup>w</sup> ɔ̄k3.taj22/ 國大	national university	/lɛp5.tət2/ 凹凸	bumpy
16	/k <sup>w</sup> ɔ̄k3.tɔ̄w22/ 國度	nation	/hɔ̄w25.wɛt2.tət2/ 好核突	really ugly
17	/k <sup>w</sup> ɔ̄k3.sin55.saj55/ 郭先生	Mr. Kwok	/wɛt2.tət2/ 核突	hideous
18	/k <sup>w</sup> ɔ̄k3.siw25.tse25/ 郭小姐	Ms. Kwok	/jɛw23.tət2/ 有凸	more than
19	/k <sup>w</sup> ɔ̄k3.tsɛŋ22/ 郭靖	Guo Jing (character)	/tət2/ 突	abruptly
20	/k <sup>w</sup> ɔ̄k3.jɛn21/ 國人	countryman	/tət2/ 凸	protruding

	Target		Filler	
	Cantonese	English	Cantonese	English
21	/k <sup>w</sup> ɔŋ25/ 廣	wide	/pɔw33.kət5/ 布吉	Phuket
22	/k <sup>w</sup> ɔŋ55/ 光	bright	/wɛn22.kət5/ 混吉	waste time
23	/k <sup>w</sup> ɔŋ55/ 胱	bladder	/t <sup>h</sup> im21.kət5/ 甜桔	sweet tangerine
24	/k <sup>w</sup> ɔŋ25.pɔ33/ 廣播	broadcast	/kaw55.kət5/ 交吉	vacate premises
25	/k <sup>w</sup> ɔŋ25.taj22/ 廣大	vast	/k <sup>h</sup> ɛm55.kət5/ 柑桔	kumquat
26	/k <sup>w</sup> ɔŋ55.tɔw22/ 光度	radiance	/daj22.kət5/ 大吉	very lucky
27	/k <sup>w</sup> ɔŋ25.tuŋ55/ 廣東	Guangdong	/sɛj33.kət5/ 細桔	small tangerine
28	/k <sup>w</sup> ɔŋ55.sin33/ 光線	optical cable	/sɛ55.t <sup>h</sup> ɔŋ21.kət5/ 沙糖桔	sugar tangerine
29	/k <sup>w</sup> ɔŋ25.ts <sup>h</sup> ɔŋ21/ 廣場	plaza	/kət5/ 桔	tangerine
30	/k <sup>w</sup> ɔŋ55.jyn21/ 光源	light source	/kət5/ 吉	auspicious
31	/kɔ55/ 哥	older brother	/pɛk5/ 北	north
32	/kɔ33/ 個	classifier	/hɔ21.pɛk5/ 河北	Hebei
33	/kɔ25/ 嗰	that	/sɛj55.pɛk5/ 西北	northwest
34	/kɔ55/ 歌	song	/taj33.pɛk5/ 太北	Tai Bei (writer)
35	/kɔ33.pit2/ 個別	individual	/tɔj21.pɛk5/ 台北	Taipei
36	/kɔ55.t <sup>h</sup> an21/ 歌壇	music world	/hɔŋ33.pɛk5/ 向北	northward
37	/kɔ33.t <sup>h</sup> ɛj25/ 個體	an individual	/p <sup>h</sup> in55.pɛk5/ 偏北	towards north
38	/kɔ25.tɔw22/ 嗰度	there	/tin55.nam21.tej22.pɛk5/ 天南地北	far apart
39	/kɔ55.sɛw25/ 歌手	singer	/tuŋ55.pɛk5/ 東北	Dongbei
40	/kɔ33.jɛn21/ 個人	the person	/jin22.tɛk5/ 認得	recognize
41	/kɔk3/ 角	horn	/tuk2.tɛk2/ 獨特	unique
42	/kɔk3/ 攞	put	/ŋa23.tɛk2/ 瓦特	watt
43	/kɔk3/ 各	each	/kej21.tɛk2/ 奇特	strange
44	/kɔk3/ 閣	pavilion	/tɔw22.tɛk5/ 道德	virtue
45	/kɔk3.pan25/ 攞板	shelf	/paj22.tɛk5/ 敗德	evil conduct
46	/kɔk3.tɛk5/ 覺得	feel	/tɛk5/ 德	virtue
47	/kɔk3.tej22/ 各地	all over (a country)	/tɛk2/ 特	special
48	/kɔk3.tɔw22/ 角度	angle	/tɛk5/ 得	able to
49	/kɔk3.sin25/ 覺醒	realize	/sɛj33.tat3/ 細筍	small patch
50	/kɔk3.jɛn21/ 各人	everyone	/ja22.pat3/ 廿八	twenty eight
51	/kɔŋ25/ 講	speak	/fat3.tat2/ 發達	developed
52	/kɔŋ55/ 崗	ridge	/jɛt5.tat3.tat3/ 一筍筍	patches
53	/kɔŋ25/ 港	harbor	/ts <sup>h</sup> yn21.tat2/ 傳達	convey
54	/kɔŋ55/ 缸	jar	/piw25.tat2/ 表達	express
55	/kɔŋ55.pɛk5/ 江北	Jiangbei	/taj22.tat3/ 大筍	big patch
56	/kɔŋ25.pan25/ 港版	Hong Kong edition	/tat3.tat3/ 筍筍	patches
57	/kɔŋ25.tɛk5.faj33/ 講得快	speak quickly	/ha2.tat2/ 下達	convey to subordinates
58	/kɔŋ25.taj22.wa22/ 講大話	lie	/tat2/ 達	extend
59	/kɔŋ55.tɔw55/ 江都	Jiangdu	/tat3/ 筍	patch

Target		Filler		
Cantonese	English	Cantonese	English	
60	/kɔŋ55.sɪŋ33/ 剛性	rigidity	/ma55.pat3/ 孖八	double eights
61	/kɔŋ55.jiw21.ts <sup>h</sup> y23/ 江瑤柱	dried scallop	/tej22.pat3/ 第八	eighth
62	/k <sup>w</sup> a25/ 寡	tasteless	/kap3.ts <sup>h</sup> ət5.kap3.pat3/ 夾七夾八	incoherent
63	/k <sup>w</sup> a33/ 卦	divinatory symbols	/kɛw25.m21.tap3.pat3/ 九唔搭八	nonsensical
64	/k <sup>w</sup> a33/ 掛	hang up	/sɛp2.pat3/ 十八	eighteen
65	/k <sup>w</sup> a55/ 瓜	melon	/daj22.kək3.pat3/ 大腳八	bigfoot
66	/k <sup>w</sup> a25.p <sup>h</sup> ɔ25/ 寡婆	widow	/pat3.sɛp2.pat3/ 八十八	eighty eight
67	/k <sup>w</sup> a33.tan55/ 掛單	house in a temple	/ts <sup>h</sup> ət5.ts <sup>h</sup> ət5.pat3.pat3/ 七七八八	mostly complete
68	/k <sup>w</sup> a55.tək5/ 瓜得	terrible	/pat3/ 八	eight
69	/k <sup>w</sup> a55.tej22.ma23.laj55/ 危地馬拉	Guatemala	/lɔw23.pak3/ 老伯	old uncle
70	/k <sup>w</sup> a55.kət3/ 瓜葛	association	/huŋ21.lɔ21.pak2/ 紅蘿蔔	carrot
71	/k <sup>w</sup> a33.səm55/ 掛心	worry	/miŋ21.pak2/ 明白	understand
72	/k <sup>w</sup> a55.jən21/ 瓜仁	melon seed	/sam55.pak3/ 三百	three hundred
73	/k <sup>w</sup> ak3/ 搵	slap someone's face	/sa55.tak2.a33.laj55.pak3/ 沙特阿拉伯	Saudi Arabia
74	/ka33/ 假	holiday	/sɛj33.pak3/ 世伯	uncle
75	/ka55/ 嘉	praise	/pən25.kak3/ 品格	personality
76	/ka55/ 家	home	/kan33.kak3/ 間隔	compartment
77	/ka25/ 架	frame	/p <sup>h</sup> ɔ33.kak3/ 破格	make an exception
78	/ka55.p <sup>h</sup> ɔ25/ 家婆	mother-in-law	/sɔŋ55.kak3/ 相隔	be separated
79	/ka55.tɛj25/ 加底	increase food bulk	/ts <sup>h</sup> a21.kak3/ 茶隔	tea strainer
80	/ka55.tək5/ 加得	can be added	/tsɔ25.kak3/ 阻隔	obstruct
81	/ka55.tan55/ 假單	leave permit	/tsi55.kak3/ 資格	qualifications
82	/ka33.kak3/ 價格	price range	/pak2/ 白	white
83	/ka55.səm55/ 加深	deepen	/pak2/ 帛	silk
84	/ka55.jən21/ 家人	family	/pak2/ 蔔	radish
85	/kak3/ 格	square	/pak3/ 佰	hundred
86	/kak3/ 隔	divide	/tsuŋ55.k <sup>w</sup> ək3/ 中國	China
87	/kak3.pik5/ 隔壁	next door	/sɔŋ55.jip2/ 商業	business
88	/kak3.tɛw3/ 格鬥	wrestle	/mat5.tɔw55/ 也都	everything
89	/kak3.sik5/ 格式	format	/p <sup>h</sup> iŋ21.k <sup>w</sup> ɔ25/ 蘋果	apple
90	/kak3.sip3/ 隔涉	remote	/t <sup>h</sup> an25.sɔŋ55.nej21.a33/ 坦桑尼亞	Tanzania
91	/kak3.tsa5/ 隔渣	filter	/ts <sup>h</sup> a21/ 茶	tea
92	/kak3.tsyt2/ 隔絕	be isolated	/maj23.je23/ 買嘢	shop

	Target		Filler	
	Cantonese	English	Cantonese	English
93	/kək3.jin21/ 格言	maxim	/ma23.laj55.səj55.a33/ 馬來西亞	Malaysia
94	/kək3.je25/ 隔夜	overnight	/lɔŋ21.ts <sup>h</sup> a21/ 涼茶	herbal tea
95	/kaŋ55.tej22/ 耕地	arable land	/fan22/ 飯	rice
96	/k <sup>w</sup> iŋ25/ 罔	shining	/lɔj23.haŋ21/ 旅行	travel
97	/k <sup>w</sup> iŋ25/ 炯	bright	/taj22.pan25/ 大阪	Osaka
98	/k <sup>w</sup> iŋ25.jɔŋ25/ 罔樣	square-faced	/jɔŋ21.tsuŋ55/ 洋蔥	onion
99	/kik5/ 擊	beat	/tsət5.tim25/ 七點	seven o'clock
100	/kik5/ 激	irritate	/jət2.pun25/ 日本	Japan
101	/kik5/ 極	extreme	/t <sup>h</sup> ɔŋ55.jyn25/ 湯圓	glutinous rice ball
102	/kik5.pəj22/ 擊斃	strike dead	/fɛj25.pək2/ 菲薄	humble
103	/kik2.taj22/ 極大	enormous	/hɔj25.kək3/ 海角	cape
104	/kik2.han22/ 極限	limit	/lɛw21.kək3/ 樓閣	pavilion
105	/kik2.tɔw2/ 極度	extremely	/sam55.kək3/ 三角	triangle
106	/kik5.tuŋ2/ 激動	excite	/tsi55.kək3/ 知覺	perception
107	/kik5.sɛw3/ 激瘦	lose a lot of weight	/tsuk5.kək3/ 觸覺	touch
108	/kik2.jiŋ21/ 極刑	capital punishment	/ŋɛw21.kək3/ 牛角	ox horn
109	/kiŋ55/ 京	capital city	/ts <sup>h</sup> aj55.tək2/ 猜度	surmise
110	/kiŋ25/ 竟	finish	/ts <sup>h</sup> yn25.tək2/ 揣度	guess
111	/kiŋ25/ 景	view	/lɔŋ21.tək2/ 量度	measure
112	/kiŋ33/ 徑	footpath	/tɔw25.pək3/ 賭博	gamble
113	/kiŋ25.tej22/ 境地	circumstances	/fan25.pək3/ 反駁	refute
114	/kiŋ55.tɔw55/ 京都	Kyoto	/hɛw23.pək2/ 厚薄	thickness
115	/kiŋ55.tuŋ22/ 驚動	alarm	/sɛm55.pək3/ 心搏	heartbeat
116	/kiŋ55.sɛw22/ 經受	endure	/t <sup>h</sup> ɛw55.pək2/ 偷薄	mean
117	/kiŋ55.jən21/ 驚人	frightening	/pək3/ 博	extensive
118	/kiŋ25.jɔŋ23/ 景仰	admire	/pək2/ 薄	thin
119	/k <sup>w</sup> ɛŋ55/ 轟	bombard	/pək2/ 雹	hail
120	/k <sup>w</sup> ɛŋ55.tuŋ22/ 轟動	cause a stir	/pək2/ 箔	foil
121	/k <sup>w</sup> ɛŋ55.tsa33/ 轟炸	bomb	/puj55.kət3/ 杯葛	boycott
122	/kɛŋ55/ 羹	spoon	/fɛn55.kət3/ 分割	cut up
123	/kɛŋ55/ 庚	age	/fɛn25.kət3/ 粉葛	segment
124	/kɛŋ25/ 耿	honest	/sa55.kət3/ 沙葛	yam bean
125	/kɛŋ25/ 梗	stiff	/sɛw55.kət3/ 收割	reap
126	/kɛŋ33.tɔ55/ 更多	more	/tsɔj25.kət3/ 宰割	slaughter
127	/kɛŋ33.taj22/ 更大	bigger	/jim55.kət3/ 閹割	castrate
128	/kɛŋ55.t <sup>h</sup> ɛj33/ 更替	replace	/kət3/ 葛	kudzu
129	/kɛŋ55.kɔj25/ 更改	alter	/kət3/ 割	cut
130	/kɛŋ55.tsin33/ 更正	make corrections		
131	/kɛŋ33.jɛw23/ 更有	certainly have		
132	/k <sup>w</sup> ɛŋ33/ 逛	stroll		

## References

- Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3), 1–43.
- Bauer, Robert S. 1982. *Cantonese sociolinguistic patterns: Correlating social characteristics of speakers with phonological variables in Hong Kong*. Ph.D. dissertation, University of California, Berkeley.
- Bauer, Robert S. & Paul K. Benedict. 1997. *Modern Cantonese phonology*. New York: Mouton de Gruyter.
- Beeley, Helena. 2015. British English [kw], [k], and [w] distinction in back round vowel contexts. *Proceedings of the 18<sup>th</sup> International Congress of Phonetic Sciences (ICPhS XVIII)*.
- Boersma, Paul & David Weenink. 2018. Praat: Doing phonetics by computer. <https://www.fon.hum.uva.nl/praat/> (accessed April 2018).
- Bongiovanni, Silvina. 2021. An acoustical analysis of the merger of /ɲ/ and /nj/ in Buenos Aires Spanish. *Journal of the International Phonetic Association* 51(2), 1770–201.
- Casali, Roderic F. 1990. Contextual labialization in Nawuri. *Studies in African Linguistics* 21(3), 319–346.
- Catford, J. C. 1988. *A practical introduction to phonetics*. Oxford: Clarendon Press.
- Cheung, Kwan Hin, Andy Chin, Peter Chung, Cream Lee, Caesar Lun, Vicky Man & Sze Wing Tang. 2016. *Linguistic Society of Hong Kong Jyutping Word List*. <https://www.lshk.org/jyutping-work-group> (30 January 2018).
- Chitoran, Ioana. 2002. A perception-production study of Romanian diphthongs and glide–vowel sequences. *Journal of the International Phonetic Association* 32(2), 203–222.
- Choi, John D. 1995. An acoustic-phonetic underspecification account of Marshallese vowel allophony. *Journal of Phonetics* 23(3), 323–347.
- Conrad, François. 2021. Regional differences in the evolution of the merger of /ʃ/ and /n/ in Luxembourgish. *Journal of the International Phonetic Association* doi: [10.1017/S0025100320000407](https://doi.org/10.1017/S0025100320000407). Published online by Cambridge University Press, 15 April 2021.
- De Jong, Nivja H. & Ton Wempe. 2009. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior research methods* 41(2), 385–390.
- Dinnsen, Daniel A. & Jan Charles-Luce. 1984. Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics* 12(1), 49–60.
- Fleming, Edward, Peter Ladefoged & Sarah Thomason. 2008. Phonetic structures of Montana Salish. *Journal of Phonetics* 36(3), 465–491.
- Fung, Roxana S. Y. & Chris K. C. Lee. 2019. Tone mergers in Hong Kong Cantonese: An asymmetry of production and perception. *The Journal of the Acoustical Society of America* 146(5), EL424–EL430.
- Hall-Lew, Lauren. 2013. ‘Flip-flop’ and mergers-in-progress. *English Language and Linguistics* 17(2), 359–390.
- Kao, Diana L. 1971. *Structure of the syllable in Cantonese*. The Hague: Mouton.
- Kawahara, S. 2010. get\_intensity\_minmax.praat. [http://user.keio.ac.jp/kawahara/scripts/get\\_intensity\\_minmax.praat](http://user.keio.ac.jp/kawahara/scripts/get_intensity_minmax.praat).
- Kendall, Tyler & Erik R. Thomas. 2018. Vowels: Vowel manipulation, normalization, and plotting in R (version 1.2.2). <https://cran.r-project.org/web/packages/vowels/vowels.pdf>.
- Kharlamov, Viktor. 2014. Incomplete neutralization of the voicing contrast in word-final obstruents in Russian: Phonological, lexical, and methodological influences. *Journal of Phonetics* 43, 47–56.
- Kim, Eun-Sook 2010. The phonological and phonetic aspects of labialization: a crosslinguistic study. *Studies in Phonetics, Phonology, and Morphology* 16(1), 41–64.
- Kochetov, Alexei. 2004. Perception of place and secondary articulation contrasts in different syllable positions: Language-particular and language-independent asymmetries. *Language and Speech* 47(4), 351–382.
- Kroos, Christian, Rikke Bundgaard-Nielson & Michael Tyler. 2010. get\_measurements.praat. [http://web.mit.edu/zqi/www/uploads/1/4/8/9/14891652/get\\_measurements.praat](http://web.mit.edu/zqi/www/uploads/1/4/8/9/14891652/get_measurements.praat).
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. *lmerTest* package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13), 1–26.
- Labov, William. 1981. Resolving the neogrammarian controversy. *Language* 57(2), 267–308.

- Labov, William, Sharon Ash & Charles Boberg. 2008. *The atlas of North American English: Phonetics, phonology, and sound change*. Berlin: Walter de Gruyter.
- Lee, Hyunjung & Allard Jongman. 2016. A diachronic investigation of the vowels and fricatives in Korean: An acoustic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association* 46(2), 157–184.
- Lippi-Green, Rosina L. 1989. Social network integration and language change in progress in a rural alpine village. *Language in Society* 18(2), 213–234.
- Luke, Steven G. 2017. Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods* 49(4), 1494–1502.
- Matthews, Stephen & Virginia Yip. 2011. *Cantonese: A comprehensive grammar*. New York: Routledge.
- Milroy, James & Lesley Milroy. 1985. Linguistic change, social network and speaker innovation. *Journal of linguistics* 21(2), 339–384.
- Mok, Peggy P. K., Donghui Zuo & Peggy W. Y. Wong. 2013. Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese. *Language Variation and Change* 25(3), 341–370.
- Newman, John. 1987. The evolution of a Cantonese phonotactic constraint. *Australian Journal of Linguistics* 7(1), 43–72.
- Nittrouer, Susan, Joanna H. Lowenstein & Eric Tarr. 2014. Amplitude rise time does not cue the /ba/–/wa/ contrast for adults or children. *Journal of Speech Language and Hearing Research* 56, 427–440.
- Ohala, John J. & H. Kawasaki-Fukumori. 1997. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In Stig Eliasson & E. H. Jahr (eds.), *Language and its ecology: Essays in memory of Einar Haugen*, 343–365. New York: Mouton de Gruyter.
- R Core Team. 2021. R: A language and environment for statistical computing. <https://www.R-project.org/>.
- RStudio Team. 2020. RStudio: Integrated development environment for R. <http://www.rstudio.com/>.
- Rutter, Ben. 2011. Acoustic analysis of a sound change in progress: The consonant cluster /stʌ/ in English. *Journal of the International Phonetic Association* 41(1), 27–40.
- Sheik, Adam. 2017. CantoDict Chinese dictionary. <http://www.cantonese.sheik.co.uk/scripts/wordsearch.php>.
- Slowiaczek, Louisa M. & Daniel A. Dinnsen. 1985. On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13(3), 325–341.
- Stevens, Kenneth N. 1998. *Acoustic phonetics*. Cambridge, MA: MIT Press.
- Stonham, John & Eun-Sook Kim. 2008. Labialization in Nuuchahnulth. *Journal of the International Phonetic Association* 38(1), 25–50.
- Sturman, Heather Wilson, Wendy Baker-Smemoe, Sofia Carreño & Bradley B. Miller. 2016. Learning the Marshallese phonological system: The role of cross-language similarity on the perception and production of secondary articulations. *Language and Speech* 59(4), 462–487.
- Suh, Yunju. 2007. Consonant-labiovelar glide combinations in Spanish and Korean. *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS XVI)*, 1773–1776.
- To, Carol K. S., Sharynne McLeod & Pamela S. P. Cheung. 2015. Phonetic variations and sound changes in Hong Kong Cantonese: Diachronic review, synchronic study and implications for speech sound assessment. *Clinical Linguistics and Phonetics* 29(5), 333–353.
- Wade, Lacey. 2017. The role of duration in the perception of vowel merger. *Laboratory Phonology* 8(1), 1–30.
- Warner, Natasha, Allard Jongman, Joan Sereno & Rachel Kemps. 2004. Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics* 32(2), 251–276.
- Wedel, Andrew, Abby Kaplan & Scott Jackson. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128(2), 179–186.
- Wickham, Hadley. 2016. ggplot2: Elegant graphics for data analysis. <https://ggplot2.tidyverse.org>.
- Yip, Moira. 1993. Cantonese loanword phonology and Optimality Theory. *Journal of East Asian Linguistics* 2(3), 261–291.
- Yu, Alan C. L. 2007. Understanding near mergers: The case of morphological tone in Cantonese. *Phonology* 24(1), 187–214.

- Yu, Alan C. L. 2016. Vowel-dependent variation in Cantonese /s/ from an individual-difference perspective. *The Journal of the Acoustical Society of America* 139(4), 1672–1690.
- Zee, Eric. 1991. Chinese (Hong Kong Cantonese). *Journal of the International Phonetic Association* 21(1), 46–48.
- Zee, Eric. 1999. Change and variation in the syllable-initial and syllable-final consonants in Hong Kong Cantonese. *Journal of Chinese Linguistics* 27(1), 120–167.