Proceedings of the 30th North American Conference on Chinese Linguistics (NACCL-30). 2019. Volume 1. Edited by Yuhong Zhu, Skylor E. Gomes, Junyu Ruan, Qian Wang, Seo-Jin Yang, Jinwei Ye, Wei Zhou. Columbus, Ohio: The Ohio State University. Pages 16-33.

Voice Quality and Identity: The Case of Hong Kong Cantonese

Roxana S. Y. Fung and Chris K. C. Lee
Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University

This study is an investigation of the voice quality of Hong Kong Cantonese (HKC) speakers separated by a series of physiological and social factors including gender, age group, education level, and English proficiency. Acoustic analyses were conducted on the speech output of 60 HKC speakers of three age groups to gauge the pitch dynamics, spectral tilt, and periodicity. Results show that the higher-educated senior male speakers speak with a breathier voice than lower-educated peers; whereas higher-educated female speakers speak with a very low (male-like) pitch floor, showing a creaky voice quality. We argue that the male speakers are influenced by RP English, whereas the female by American English. The study does not only develop a normative profile of the voice quality of HKC, but new discovery regarding crosslinguistic influence: a non-dominant language can influence one's dominant language, both at the individual and societal level.

1. Introduction

1.1. Voice quality and Indexical information

Voice quality, a less studied domain in sociophonetics, refers to the process by which the vocal folds and other muscles in the pharyngeal cavity produce sounds of different audio impressions through different movements. Voice quality is regarded as the quasi-permanent quality of a speaker’s voice (Abercombie, 1967, p. 91) and a significant component of one’s accent. It is determined by two types of factors: firstly, the anatomical makeup of a speaker’s vocal tract; and secondly, the long-term muscular settings of speaker’s larynx and supralaryngeal vocal tract, which are once acquired idiosyncratically or by social imitation, become an unconsciously habitual behaviour (Laver, 1968). In other words, the physiology of a speaker determines the range of possible operation of any voice quality feature, but the speaker may habitually select a more limited range within the possible extremes. Hence, voice quality can transmit various indexical information, such as biological, psychological, and social information.

1.1.1. Voice quality and biological information

Voice quality, determined partly by physiology, necessarily differs between male and female speakers, and speakers of different ages. Due to the difference in thickness of the vocal folds, and muscle tone, female voices are an octave higher, more dramatic in pitch change, and breathier than male voices (Hanson, Stevens, Kuo, Chen, & Slifka, 2001; Kreiman & Sidtis, 2011). Aging, on the other hand, has a rather complicated effect on a person's voice. Due to a weakened ability of control throughout the respiratory and phonatory system (Kreiman & Sidtis, 2011), aging and a speaker's voice pitch has a non-linear relationship: as a person aged (from teens until their 50s-60s), they sounded lower and lower in pitch, but after that their pitch will rise slightly until their 90s (Stathopoulos, Huber, & Sussman, 2011). Aging will also make a person's voice become hoarser, or breathier (Kreiman & Sidtis, 2011).

1.1.2. Voice quality and social information

Since voice quality can function as a social index, different languages, which transmit different social and cultural values, will have different preferred voice quality settings. Previous research even revealed that speakers of different varieties of the same language will have different habitual settings. For example, Henton and Bladon (1988) revealed that RP English male speakers tend to be breathier than male speakers of some northern British English varieties. Esling (2000) found that a particular phonation type could distinguish certain English dialects in Scotland. Newman and Wu (2011) identified breathier voice as a factor separating the speech of Asian Americans from the non-Asian Americans. Szakay (2012) evidenced that phonation is a marker of ethnicity in two varieties of New Zealand English. Fung (2015) revealed that voice quality setting is one of the principal differences between Guangzhou and Hong Kong varieties of Cantonese.

On the other hand, the preferred voice quality setting of a language can also change over time. For example, Klatt and Klatt (1990) revealed that female speakers of American English at that time spoke breathier than male speakers. Two decades later, Yuasa (2010) reported that young professional females in California tend to use a lot of vocal fries in their speech. Podesva (2013) observed a similar trend in professional females residing in Washington DC. Both studies suggest that creaky voice quality is a tool for American women to project an educated urban upwardly mobile image. All these point to the dynamic nature of the preferred voice quality settings of a language community, due to a change in social and cultural ideologies.

1.2. Voice quality and bilingualism

Each language has its own voice quality settings and previous research reported that bilingual speakers will adopt different settings when speaking the two languages. For example, Bahmanbiglu, Mojiri, and Abnavi (2017) studied Farsi-Qashqai bilinguals, and Ng, Chen, and Chan (2012) studied Cantonese-English bilinguals. Some other research suggested that the voice quality setting of one language will transfer to another. For example, Esling and Wong (1983) studied Persian-English bilinguals and reported that these speakers transferred their L1-Persian voice quality to their L2-English. Recent studies, on the other hand, revealed that voice settings of a speaker's L2 can actually affect their L1. For example, Kim (2017) the use of phrase-final creak in Spanish by L1 Spanish- L2 English bilinguals. Note that the participants in these two studies are heritage speakers of L1 living in the US, and English is actually the dominant language in the speakers' growing environment.

1.3. An acoustic investigation into voice quality

In general, the acoustic study of voice quality concerns the habitual pitch profile (which reflects the frequency of vocal fold vibration), spectral profile, and relative strengths of harmonic vs. inharmonic components of sound waves (which reflects the manner of vocal fold vibration).

1.3.1. Pitch profile

The rate of vibrations of the vocal folds causes variation in F0, which is perceived as pitch. A number of studies have compared the pitch level (the mean or median of F0) and the pitch span (the maximum minus the minimum F0) of different linguistic systems and the differences between them. For instance, Japanese vs. English (Loveday, 1981; Todaka, 1993); Polish vs. English (Majewski, Hollien, & Zalewski, 1972); British English vs. German (Mennen, Schaeffler, & Docherty, 2007); Mandarin and English (Eady, 1982). Even different dialects of the same language may differ. For instance, Torgerson (2005) found out that Taiwan Mandarin speakers have a lower pitch than Beijing Mandarin speakers. Thomas (2010) reported that in the southern region of the United States, many Caucasian females seem to use a wider pitch span than African American females or males of either ethnic group (p.226).

1.3.2. Phonation types

Table 1.1. Characteristics of different voices (Gobl & Chasaide, 2010).

	Breathy Voice	**Modal Voice**	**Creaky Voice**
Physiology[1]	Low adductive tension (AT); very weak medial compression (MC); low longitudinal tension (LT).	Moderate AT; moderate MC; moderate LT.	High AT; very high MC; little LT; thick, compressed vocal folds.
Acoustics	Lowest harmonic (H1) relatively boosted; aspiration noise.	Source spectrum similar to the ideal spectrum, i.e. a spectral tilt of 6dB per octave.	Low and irregular f0; higher harmonics relatively boosted; modulation noise.
Perception	Audible aspiration noise due to turbulent airflow through glottis.	The neutral mode of phonation, no audible noise.	Low-pitched, irregular pitch, constricted-sounding.

Phonation refers to the process by which the vocal folds and other muscles in the pharyngeal cavity. Commonly found phonation types include modal voice, creaky voice,

[1] Note that (1) AT refers to the force which draws the arytenoids together; (2) MC refers to the force by which the ligamental glottis is closed; and (3) LT is the tension of the vocal folds, mediated primarily by the thyroarytenoid muscle (Gobl & Chasaide, 2010, pp. 395–396)

and breathy voice. Table 1.1 summarises the perceptual, physiological, and acoustic characteristics of these three common phonation types.

The major ways to acoustically gauge phonation are spectral tilts and the relative strengths of acoustic components. Spectral tilt is the degree to which intensity drops off as frequency increases. The most commonly used measure is H1-H2 (Keating & Esposito, 2007). The relative strengths of acoustic components can be gauged by periodicity measures which quantifies the amplitude of a sound wave's harmonic component relative to its noise component, such as CPP (Hillenbrand, Cleveland, & Erickson, 1994).

1.4. The linguistic profile of Hong Kong

Hong Kong has been widely understood as a biliterate, trilingual city, but before the 1980s, Hong Kong was a diglossic but monolingual city with Hong Kong Cantonese (HKC) as its major communication means (Evans, 2016; Luke & Richards, 1982). During the first 130 years of British rule, while English was the official language of colonised Hong Kong, Standard English education is still exclusively for the wealthy and elites (Setter, Wong, & Chan, 2010). English is not a second language nor a foreign language in Hong Kong (Luke & Richards, 1982). Starting from the early 1980s, Hong Kong began to shift to a knowledge-based economy, English then acquires a prestige status in Hong Kong due to its instrumental value in the new economy (Li, 1999). Together with the implementation of nine-year compulsory education for the mass in the late 1970s, local people acquire their English through formal education. As for Putonghua, it has no official status, and very restricted social function in Hong Kong (Li, 1999, p. 70); language attitude research showed that Hongkongers are in general not very motivated to learn Putonghua (Li, 2017, p. 14). In recent years, many middle-class families have hired English-speaking foreign domestic helpers to take care of the children and these children are exposed to English, with its degree sometimes outstripping that of their native language, Cantonese. Dulay, Tong, and McBride (2017) has shown that this affects these children's native language development. On the other hand, a growing number of financially well-off parents, who are frustrated with the local education system, settle for sending their children to costly local international schools. This group of children usually develop native-like English proficiency after graduation. People believe that these children have carried their English pronunciation habits to their native tongue and form a distinctive Cantonese accent (see Dunn, 2016). This very interesting observation seems to suggest the possibility of the transfer of phonetic features from the speakers' L2 English (a non-dominant language) to their native HKC (a dominant language).

1.5. The goal of the study

The impressionistic claim of the influence of English on HKC motivates the current study on the variations and changes of the most persistent feature that characterize the accent of an individual and an entire social group in HKC. This study attempts to investigate the voice quality and its indexical information in Hong Kong Cantonese (HKC). It aims to answer the following research questions:

(1) Are there any difference in voice quality between groups of HKC speakers separated by a series of social variables including gender, age, and education background?
(2) Is there a transfer of the voice quality features of L2-English to L1-HKC, the usual and dominant language of most speakers in Hong Kong?

In the process, we also target at developing a normative profile of the voice quality of HKC containing information of different groups of HKC speakers, which we believe will benefit the further development of speech pathology, and sociophonetic studies on HKC.

2. Methodology

2.1. Participants

A total of 60 native speakers of HKC in three age groups with balanced gender and education level (5 participants × 2 genders × 3 age groups × 2 education levels) were randomly recruited. In order to build a normative profile for the HKC community, we did not particularly look for speakers attending international schools. The three age groups were young (18-25 years), middle-aged (35-45 years), and senior (55-65 years). Participants in the high education level group were university students or graduates, and those in the low education level group were all high school graduates. All 60 participants were born and raised in Hong Kong, and most of them do not speak other Chinese dialects, though some of them were minimal bidialectals or receptive bidialectals.

Among the high education level group, 14 participants have stayed abroad for more than six months (1-9 years, mean duration: 3 years). All of them stayed in English-speaking countries, except one middle-aged speaker stayed in Switzerland, and one senior speakers stayed in France and Germany. Table 2.1 summarises the percentage of participants in the high education group who have stayed abroad for longer than six months. No participants of the low education level group reported records of staying abroad for more than six months.

The label 'education level' has dual connotations: one is whether the participant is admitted to tertiary education, and the other one being English proficiency. This duality arises from the status of English as the academic lingua franca in Hong Kong (Evans, 2016). With the increased regular exposure and utility of both spoken and written English, it is expected that the higher-educated participants will attain comparative higher English proficiency than those who did not attend college. This expectation will be verified in §3.1 (see §2.3 for the assessment method).

Table 2.1. % of participants in the high education group who have stayed abroad for ≥6 months

	Male (%)	**Female (%)**	**Sub-total (%)**
Young	40	80	60
Middle-aged	60	40	50
Senior	20	40	30

2.2. Speech samples collection

Participants performed the recording tasks in a sound-proof booth of the Speech and Language Sciences Lab (SLS Lab) of the Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University. All recordings were done with a Telefunken M80 dynamic microphone connected to an M-Audio M-Track Plus sound interface. All recordings were made with mono-channel sampling at 44.1kHz.

All participants recorded two passages. The first one is a 350-character Cantonese passage printed in traditional Chinese characters. This passage is written with colloquial Cantonese lexical items, and adheres to the syntax of Cantonese. In general, the passage is descriptive in nature: it contains no dialogues nor direct speech.

After recording the Cantonese passage, the participants were asked to record a 200-word English passage. The passage is rather simple in terms of vocabulary to make sure that participants of different English proficiency can record the passage comfortably. However, the passage shows some varieties of complexity in sentence structures. The passage is descriptive in nature, and consists mainly of declarative sentences.

Participants were required to get themselves familiar with the passage before recording themselves in a natural and relaxed manner. Participants were reminded that they should keep themselves about 15cm away from the microphone as they record their own voice.

2.3. English proficiency evaluation

Two female native speakers of American English were recruited to evaluate participants' proficiency according to their performance in reading the English passage. The first evaluator is a trained linguist who holds a PhD degree, and the second evaluator is a PhD student who has never participated in language evaluation tasks as an evaluator. The evaluators were aged 25-35.

Three sentences were extracted from the recordings. The sentences were presented to the evaluators in separate blocks. For each sentence, the evaluators were required to give marks ranged from 0 (very bad performance) to 5 (native-like performance) on three aspects: fluency, intonation, and pronunciation (Iwashita, Brown, McNamara, & O'Hagan, 2008). For fluency, the evaluator considered how fluent the speaker is when s/he read a passage in her second language. Hence, speakers producing more pauses, and self-corrections received a lower score in this aspect. For intonation, the evaluator considered whether the speaker's intonation in producing compound and complex declaratives resembles that of a native speaker. For pronunciation, the evaluator considered whether the speakers pronounce the lexical items accurately at the segmental level. Placement of lexical stress, and vowel reduction, are under this category as well. Evaluators could give a half score if they find appropriate.

The evaluators were invited to the SLS Lab to evaluate the recordings by our participants. The task was implemented using the MFC Experiment function of Praat (Boersma, 2001). In essence, the recordings were presented in random order, and the evaluators could replay the recordings as many times as needed. The evaluators were given separate score sheets for the three blocks. Each participant receives a score averaged across blocks ranging from 0 to 15 from each evaluator.

2.4. Data analysis

2.4.1. Annotation

The speech outputs of the participants were segmented for acoustic measurement of voice quality. Segmentation were done manually using Praat, v.6.0.37 (Boersma, 2001) by the second author, and checked by the first author and a research personnel of the department. All long vowels were labelled except /i:/ and /u:/ (they are excluded because their low first formant (F1) values will affect the accuracy of harmonic amplitude estimation, according to Keating and Esposito 2007), and the onsets and offsets of vowels were considered as the points of onset and offset of stable second formant (F2) respectively (Thomas 2010, pp. 139-143). While studies on the relationship between voice quality and discourse suggest that creaky voice may be a cue to utterance ending (Garellek, 2015), phrase-final creak may also have sociolinguistic meaning (Callier, 2013), and more importantly, the adoption of creak as utterance-ending marker is language-specific (Kim, 2017). This study hence kept all phrase-final vowels for phonation feature correlates estimation.

2.4.2. Voice quality correlates estimation

All phonation correlates were estimated using the Matlab-based software VoiceSauce, v.1.3.1 (Shue, Keating, Vicenik, & Yu, 2011). Pitch-related correlates, including the f0 and harmonic amplitudes were estimated using the STRAIGHT algorithm incorporated in the VoiceSauce. A wide pitch dynamics (40-500Hz) for range of estimation were adopted because the recordings contain considerable number of creaky instances, and some speakers are quite dramatic in their production. Frequencies of formants and their bandwidths were estimated using the Snack toolkit incorporated in the VoiceSauce, using the default settings (pre-emphasis at 0.96). Periodicity measures were estimated using a 5-period window, and the other correlates were analysed with a 25ms window size. All correlates were estimated with a 1ms frame shift.

After the estimation process, all acoustic correlates were outputted at each msec. For f0, F1, and F2, values estimated by Praat incorporated in the VoiceSauce were also outputted, alongside the respective STRAIGHT and Snack measures for data validation. In this study, an absolute difference in estimated value between the two algorithms greater than 200 will signal poor estimation. All valid data of f0 were transformed into semitones (ST) with the formula $ST = 12 \times \log_2(f0 / f0_{ref})$, where $f0_{ref}$ was set at 1Hz to enable inter-personal comparisons.

The pitch-related correlates to be reported in this study are pitch level (mean ST), pitch floor (minimum ST), pitch ceiling (maximum ST), pitch span (the difference between pitch floor and ceiling), and pitch dynamics (standard deviation of ST). The phonation correlates to be reported are the formant-corrected narrow-, mid-, and wideband spectral tilt measures H1*-H2* (which reflects vocal fold stiffness, Zhang, Kreiman, Gerratt, & Garellek, 2013), H1*-A1* (which reflects the degree of posterior glottal opening, Hanson et al., 2001), and H1*-A3* (which the abruptness of glottal closure, Pennington, 2005), and periodicity measure CPP (Hillenbrand et al., 1994).

3. Results and Discussions

3.1. The correlation between education levels and English proficiency

§2.1 mentioned the duality of the tag 'education level.' This section verifies this claim. Table 3.1 summarises the ratings given by the two evaluators. Both inter-rater and intra-rater reliability of ratings were assessed with Intraclass Correlation Coefficient (ICC) estimates calculated via SPSS 23.0. For intra-rater reliability, ICC was calculated based on a mean-rating (k=3), absolute agreement, two-way mixed-effects model. For inter-rater reliability, ICC was calculated based on a mean-rating (k=2), consistency, two-way random-effects model. Results indicated that all measures were of good reliability (ICC>0.8 for all reliability tests).

Table 3.1. The rating profiles of the two evaluators. Highest score possible is 15.

Rater #	Min	Max	Mean	SD	Intra-rater reliability	Inter-rater reliability
1	7.83	15	11.87	1.79	0.86, 95%CI [0.75, 0.92]	0.88, 95%CI [0.79, 0.93]
2	3.67	15	9.72	2.58	0.87, 95%CI [0.80, 0.92]	

To investigate whether education level is an effective predictor of a participant's English oral proficiency (see Table 3.2 for the mean scores of different groups), a point-biserial correlation was run between education level and English oral proficiency score. There was a statistically significant correlation between education level and proficiency score, r_{pb}=0.703, p<0.0001, with higher-educated participants scoring higher in English oral proficiency (M=12.25, SD=1.29) than lower-educated participants (M=9.33, SD=1.69). Education level accounted for 49.42% of the variability in the proficiency score. The strong correlation (r_{pb}>0.5) between education level and English proficiency score confirmed the dual quality of 'education level.'

Table 3.2. Proficiency ratings of different groups of HKC speakers (maximum score: 15)

Education	Gender	Senior		Middle-aged		Young		*Age*Edu*	
high	male	11.25	(1.58)	12.32	(1.66)	13.22	(0.45)	*12.26*	*(1.50)*
	female	11.48	(0.85)	12.13	(0.57)	13.12	(1.19)	*12.24*	*(1.09)*
*Gender*High*		*11.37*	*(1.20)*	*12.22*	*(1.18)*	*13.17*	*(0.85)*	12.25	(1.29)
low	male	8.85	(1.47)	8.8	(1.60)	10.03	(0.90)	*9.23*	*(1.39)*
	female	8.9	(2.69)	10.27	(1.81)	9.15	(1.40)	*9.44*	*(1.99)*
*Gender*Low*		*8.88*	*(2.05)*	*9.53*	*(1.79)*	*9.59*	*(1.20)*	9.33	(1.69)

In the following sections, we first present our findings on the anatomically determined differences and then the socially acquired characteristics in voice quality.

3.2. Differences in voice quality caused by physiological factors

3.2.1. How the two genders differ in voice quality

Table 3.3 summarises the pitch profile of HKC speakers divided across gender and age groups. The group means were followed up with 2 (gender) × 3 (age) two-way univariate ANOVAs.

Table 3.3. Pitch profile of Hong Kong Cantonese speakers of different genders and age groups.

	Gender	senior	middle-aged	young	*Gender total*
pitch level	male	83.23 (2.45)	79.87 (2.02)	81.94 (2.45)	*81.68 (2.64)*
	female	89.51 (2.3)	89.23 (2.17)	91.4 (1.56)	*90.05 (2.2)*
	Age total	*86.37 (3.96)*	*84.55 (5.22)*	*86.67 (5.25)*	*85.86 (4.86)*
pitch span	male	13.91 (3.46)	13.74 (4.51)	13.94 (5.18)	*13.86 (4.29)*
	female	17.92 (6)	18.83 (6.98)	19.59 (3.49)	*18.78 (5.53)*
	Age total	*15.91 (5.19)*	*16.29 (6.29)*	*16.77 (5.19)*	*16.32 (5.5)*
pitch floor	male	76.56 (1.94)	74.03 (2.7)	75.33 (2.93)	*75.3 (2.68)*
	female	79.68 (6.99)	78.27 (5.5)	79.46 (4.12)	*79.14 (5.5)*
	Age total	*78.12 (5.25)*	*76.15 (4.75)*	*77.39 (4.08)*	*77.22 (4.71)*
pitch ceiling	male	90.47 (3.16)	87.77 (4.66)	89.27 (3.28)	*89.17 (3.8)*
	female	97.6 (1.92)	97.11 (3.14)	99.06 (3.33)	*97.92 (2.89)*
	Age total	*94.03 (4.46)*	*92.44 (6.16)*	*94.16 (5.97)*	*93.54 (5.54)*
pitch dynamics	male	2.79 (0.62)	2.73 (0.55)	2.75 (0.84)	*2.75 (0.66)*
	female	3.54 (1.29)	2.99 (0.77)	3.14 (0.79)	*3.22 (0.97)*
	Age total	*3.16 (1.06)*	*2.86 (0.66)*	*2.94 (0.82)*	*2.99 (0.86)*

All the pitch profile measures displayed significant gender effects. In particular, female speakers, compared to their male peers, have a higher pitch level (+8.37ST, 95%CI [7.24, 9.49], $F_{1,54}$=221.28, p<0.0001), a wider pitch span (+4.92ST, 95%CI [2.28, 7.56], $F_{1,54}$=13.96, p<0.0005), a higher pitch floor (+3.84ST, 95%CI [1.56, 6.11], $F_{1,54}$=11.43, p=0.001), a higher pitch ceiling (+8.76ST, 95%CI [7.02, 10.49], $F_{1,54}$=102.71, p<0.0001), and a slightly greater pitch dynamics (+0.50ST, 95%CI [0.031, 0.91], $F_{1,54}$=4.62, p=0.036).

While there are significant differences in pitch span, pitch ceiling, and pitch floor, the two genders differ in pitch span mainly because the female speakers adopt a disproportionately low pitch floor as compared to male. This is evident when we divide the pitch span along their pitch level. Lower and upper pitch spans are operationalised as the distance of pitch floor and ceiling from pitch level respectively. ANOVA result shows that gender effect is significant only at the lower pitch span, with female speakers having a wider lower pitch span than male speakers (+4.53ST, 95%CI [2.42, 6.65], $F_{1,54}$=18.44, p<0.0001), see Table 3.4 for details. This underlines the markedly low pitch floor of female speakers of HKC.

Table 3.4. Pitch spans divided along pitch level of different groups of HKC speakers.

Measure	Gender	senior	middle-aged	young	*Gender total*
Lower pitch span	male	6.67 (2.52)	5.84 (2.07)	6.62 (4.05)	*6.38 (2.92)*
	female	9.82 (5.45)	10.95 (5.73)	11.94 (3.24)	*10.91 (4.84)*
	Age total	*8.25 (4.44)*	*8.4 (4.95)*	*9.28 (4.5)*	*8.64 (4.58)*
Upper pitch span	male	7.24 (2.86)	7.9 (3.94)	7.32 (2)	*7.49 (2.95)*
	female	8.09 (1.06)	7.88 (2.32)	7.65 (2.06)	*7.88 (1.83)*
	Age total	*7.67 (2.14)*	*7.89 (3.15)*	*7.49 (1.98)*	*7.68 (2.44)*

Table 3.5 summarises the phonation settings of HKC speakers divided across gender and age groups. The group means were followed up with 2 (gender) × 3 (age) two-way univariate ANOVAs.

Table 3.5. The phonation setting profile of Hong Kong Cantonese speakers.

Correlate	Gender	senior	middle-aged	young	*Gender total*
H1*-H2*	male	3.72 (3.2)	0.66 (2.36)	1.81 (2.02)	*2.06 (2.8)*
	female	6.28 (2.09)	6.1 (1.99)	6.86 (1.21)	*6.41 (1.78)*
	Age total	*5 (2.94)*	*3.38 (3.51)*	*4.34 (3.06)*	*4.24 (3.2)*
H1*-A1*	male	15.56 (2.61)	14.99 (2.55)	14.61 (1.32)	*15.05 (2.2)*
	female	14.26 (1.92)	14.02 (1.42)	13.8 (0.99)	*14.03 (1.45)*
	Age total	*14.91 (2.33)*	*14.51 (2.07)*	*14.2 (1.21)*	*14.54 (1.92)*
H1*-A3*	male	11.67 (5.59)	12.92 (4.36)	13.12 (3)	*12.57 (4.34)*
	female	12.87 (3.79)	13.89 (2.63)	13.91 (3.32)	*13.55 (3.2)*
	Age total	*12.27 (4.69)*	*13.4 (3.54)*	*13.52 (3.11)*	*13.06 (3.81)*
CPP	male	22.8 (1.77)	22.45 (1.39)	23.64 (1.37)	*22.96 (1.55)*
	female	23.51 (2.24)	23.73 (1.06)	23.57 (1.42)	*23.6 (1.59)*
	Age total	*23.16 (2)*	*23.09 (1.37)*	*23.6 (1.36)*	*23.28 (1.59)*

Out of the four phonation setting correlates, only H1*-H2*, and H1*-A1* shows a significant gender effect: female speakers show a higher H1*-H2* (+4.35dB, 95%CI [3.20, 5.50], $F_{1,54}$=52.26, p<0.0001), but a subtly lower value in H1*-A1* (-1.03dB, 95%CI [-2.01, -0.04], $F_{1,54}$=4.36, p=0.042). This may suggest that female speakers have lower stiffness in their vocal folds (Zhang et al., 2013), and less posterior glottal opening compared to their male peers (Hanson et al., 2001). The comparable H1*-A3* and CPP values in the two gender groups suggest that their voice quality settings do not differ much in spectral tilt and signal periodicity.

Our gender differences result largely echoes the findings of gender differences in voice quality as reported in §1.1.1, where female speakers indeed have higher pitch level, greater pitch dynamics, and breathier voice quality, as reflected in the higher values of H1*-H2*, since vocal folds with less stiffness are more prone to air leakage and hence produce breathier voice quality. However, some of the findings are worth attention. One is the source of the difference in pitch span. While previous studies note such a difference, there are no conclusive data on its source. Our findings suggest that the wide pitch span is caused by the adoption of a rather extreme low pitch floor by the female speakers. However, given the widespread belief of frequency code, where high pitch is believed to note femininity and hence match better with the social image of a female speaker, why would female HKC speakers adopt an extreme pitch floor as compared to male? The other finding worthy of attention is the fact that female speakers of HKC do not display significantly more posterior glottal opening than male speakers as suggested in Hanson et al (2001). Considering this together with the adoption of lower pitch floor, we believe that it is a social factor affecting the voice quality of HKC speakers, and we shall discuss it in more detail in §3.3.

3.2.2. How the three age groups differ in voice quality

Given the significant differences between the two genders in pitch settings and phonation settings, it is necessary to separate the two gender groups when considering differences in age groups. ANOVA results show that among all five pitch profile measures, only pitch level has a significant interaction effect between age and gender, $F_{2,54}=3.45$, $p=0.039$. Comparing its simple effects with Bonferroni adjustments reveals a significant simple age effect on pitch level in male speakers only, $F_{2,54}=6.06$, $p=0.004$. This simple age effect is attributable to the higher pitch level of the senior male group as compared to the middle-aged males (+3.36ST, 95%CI [0.95, 5.77], $p=0.003$). There are no significant differences between other combinations of pairwise comparisons in the male group, or any pairwise comparisons between any combinations of age groups in the female group. This pitch profile finding is rather expected in the sense that differences in pitch profile caused by aging are not as straightforward as gender differences as discussed in §3.2.1. Recall that age and pitch profile has a non-linear relationship (see §1.1.1). Our findings reveal the non-linear and gender-differentiated relationship in the sense that (1) senior male speakers indeed show a higher pitch level than middle-aged speakers, and (2) this age effect is not found in the same comparison in the female group. However, the age difference in the change in pitch dynamics reported in Stathopoulos et al. (2011), which should show a similar but milder non-linear relationship with age, is not reflected in our study.

Considering the four phonation setting correlates, none of them shows a significant interaction effect, but a significant simple age effect on H1*-H2* with Bonferroni adjustments is observed in male speakers, $F_{2,54}=4.84$, $p=0.012$. This simple age effect on H1*-H2* in male speakers is attributable to a higher H1*-H2* value in senior male speakers than their middle-aged peers (+3.07dB, 95%CI [0.61, 5.53], $p=0.01$).

Recall the effect of aging on voice quality as outlined in §1.1.1, the possible acoustic consequence of such physiological change should be higher values in H1*-H2* (due to decreased stiffness), H1*-A1* (due to increased glottal gaps), and H1*-A3* (due to greater spectral tilt), and lower CPP (due to increased aspiration noise). However, our findings only revealed a higher H1*-H2* value. The very limited success in revealing the acoustic differences induced by physiological changes due to aging again seems to suggest that a social factor has a role to play in how one's voice is used when they speak their native language.

3.3. Voice quality differences caused by social factors

The discussions in the last section suggest that while the physiological differences between the two genders and different age groups affect the voice quality of different groups of HKC speakers, the picture does not match our expectation, given detailed discussions in Kreiman and Sidtis (2011). We shall consider if education background, will help differentiate the voice quality of different groups of speakers. Table 3.6 (in p. 13) and

Table 3.7 (in p. 14) summarise the pitch profile and voice quality settings profile of HKC speakers of different education levels respectively. The group means of these measures were followed up with 2 (gender) × 3 (age) × 2 (proficiency) three-way univariate ANOVA.

For the pitch profile of different groups of HKC speakers, interaction effect between gender and education is significant only in the comparison of pitch span ($F_{1,48}$=4.64, p=0.036) and pitch floor ($F_{1,48}$=9.85, p=0.003). Simple education effect is found significant only when comparing female speakers of different education levels. In particular, female speakers with a higher education level have a wider pitch span (+6.25ST, 95%CI [2.76, 9.75], $F_{1,48}$=12.93, p=0.001), and a lower pitch floor (-6.73ST, 95%CI [-9.49, -3.98], $F_{1,48}$=24.10, p<0.0001). As for phonation setting correlates, three-way interaction effect is significant for H1*-H2* ($F_{2,48}$=8.14, p=0.001), and H1*-A3* ($F_{2,48}$=4.06, p=0.024) only. For H1*-H2*, simple education effect is found significant only when comparing senior male speakers of different education levels ($F_{1,48}$=15.39, p=0.0003); in particular, senior male speakers with a higher education level has a greater H1*-H2* value compared to their peers with a lower education level (+4.84dB, 95%CI [2.36, 7.32]). For H1*-A3*, simple education effect is found significant when comparing senior male speakers of different education levels ($F_{1,48}$=4.37, p=0.042), and comparing middle-aged male speakers of different education levels ($F_{1,48}$=9.76, p=0.003). Higher-educated senior male speakers again show a greater H1*-A3* value compared to their peers (+4.66dB, 95%CI [0.18, 9.15]), whereas higher-educated middle-aged male speakers show a lower H1*-A3* value compared to their peers with a lower education level (-6.97dB, 95%CI [-11.45, -2.48]). There is no significant simple education effect in other pairwise comparisons.

In short, our findings suggest that while speakers with different education levels will have different phonation settings, the effect of education is not uniform across genders and age groups. For male speakers, education levels have effects on the spectral representations of a speaker's voice, but its effect is again not uniform in different age groups. In particular, senior male speakers who attained a higher education level is speaking with a breathier voice than their peers who attained a lower education level, since they have lower stiffness in vocal folds (as indicated by a higher value in H1*-H2*), and less abrupt glottal closure and hence a steeper spectral tilt (as indicated by a higher value in H1*-A3*). The effect of education upon senior male speakers even leads to the loss of gender difference in H1*-H2*.

Table 3.6. The pitch profile of different groups of HKC speakers (with SD)

	EDU	GENDER	senior	middle-aged	young	***GENDER*EDU***
Pitch level	high	male	83.66 (2.05)	79.4 (1.03)	81.78 (3.05)	*81.61 (2.73)*
		female	87.89 (1.37)	89.47 (1.65)	91.07 (1.33)	*89.48 (1.91)*
	*EDU*AGE*		*85.77 (2.77)*	*84.44 (5.47)*	*86.43 (5.38)*	*85.55 (4.62)*
	low	male	82.8 (2.96)	80.34 (2.74)	82.1 (2.03)	*81.75 (2.64)*
		female	91.13 (1.85)	88.98 (2.78)	91.74 (1.84)	*90.61 (2.38)*
	*EDU*AGE*		*86.96 (4.97)*	*84.66 (5.25)*	*86.92 (5.39)*	*86.18 (5.14)*
Pitch span	high	male	13.86 (4.41)	14.06 (4.86)	15.1 (4.19)	*14.34 (4.2)*
		female	20.81 (7.34)	22.67 (6.75)	22.25 (2.91)	*21.91 (5.62)*
	*EDU*AGE*		*17.33 (6.78)*	*18.36 (7.17)*	*18.68 (5.07)*	*18.12 (6.21)*
	low	male	13.96 (2.74)	13.42 (4.67)	12.77 (6.28)	*13.38 (4.46)*
		female	15.03 (2.52)	15 (5.22)	16.94 (1.11)	*15.65 (3.29)*
	*EDU*AGE*		*14.49 (2.54)*	*14.21 (4.75)*	*14.86 (4.79)*	*14.52 (4.02)*
Pitch floor	high	male	76.96 (1.99)	73.99 (1.56)	73.99 (1.92)	*74.98 (2.23)*
		female	75.75 (7.7)	74.97 (5.53)	76.6 (2.82)	*75.77 (5.33)*
	*EDU*AGE*		*76.36 (5.34)*	*74.48 (3.87)*	*75.29 (2.66)*	*75.38 (4.04)*
	low	male	76.15 (2.02)	74.07 (3.74)	76.66 (3.35)	*75.63 (3.12)*
		female	83.61 (3.49)	81.58 (3.18)	82.33 (3.13)	*82.51 (3.15)*
	*EDU*AGE*		*79.88 (4.76)*	*77.82 (5.14)*	*79.49 (4.27)*	*79.07 (4.66)*
Pitch ceiling	high	male	90.82 (3.19)	88.05 (4.45)	89.09 (3.47)	*89.32 (3.66)*
		female	96.56 (1.29)	97.64 (1.65)	98.85 (4.05)	*97.68 (2.62)*
	*EDU*AGE*		*93.69 (3.8)*	*92.84 (5.96)*	*93.97 (6.25)*	*93.5 (5.28)*
	low	male	90.11 (3.47)	87.48 (5.38)	89.44 (3.48)	*89.01 (4.06)*
		female	98.64 (1.99)	96.58 (4.33)	99.27 (2.91)	*98.16 (3.22)*
	*EDU*AGE*		*94.38 (5.22)*	*92.03 (6.65)*	*94.35 (6)*	*93.59 (5.88)*
Pitch dynamics	high	male	2.71 (0.83)	2.61 (0.22)	3.06 (0.92)	*2.8 (0.7)*
		female	4.15 (1.62)	3.29 (0.81)	3.47 (0.92)	*3.64 (1.15)*
	*EDU*AGE*		*3.43 (1.43)*	*2.95 (0.66)*	*3.26 (0.89)*	*3.22 (1.03)*
	low	male	2.86 (0.41)	2.84 (0.77)	2.43 (0.71)	*2.71 (0.64)*
		female	2.92 (0.41)	2.69 (0.67)	2.82 (0.55)	*2.81 (0.52)*
	EDU*AGE		2.89 (0.39)	2.77 (0.69)	2.62 (0.63)	*2.76 (0.57)*

Table 3.7. The phonation settings profile of HKC speakers of different education levels

	EDU	GENDER	senior	middle-aged	young	*GENDER*EDU*
H1*-H2*	high	male	6.14 (2.08)	-0.25 (1.96)	1.27 (1.27)	*2.39 (3.28)*
		female	5.07 (0.91)	6.19 (2.41)	6.61 (0.66)	*5.96 (1.57)*
	*EDU*AGE*		*5.61 (1.62)*	*2.97 (3.98)*	*3.94 (2.97)*	*4.17 (3.11)*
	low	male	1.3 (2)	1.56 (2.59)	2.35 (2.62)	*1.74 (2.29)*
		female	7.49 (2.32)	6.01 (1.76)	7.11 (1.65)	*6.87 (1.9)*
	*EDU*AGE*		*4.4 (3.85)*	*3.78 (3.14)*	*4.73 (3.25)*	*4.3 (3.33)*
H1*-A1*	high	male	17 (2.31)	13.87 (2.55)	14.53 (1.59)	*15.13 (2.46)*
		female	13.72 (1.02)	13.77 (2)	13.64 (0.66)	*13.71 (1.25)*
	*EDU*AGE*		*15.36 (2.41)*	*13.82 (2.16)*	*14.08 (1.24)*	*14.42 (2.05)*
	low	male	14.12 (2.2)	16.12 (2.23)	14.68 (1.16)	*14.97 (1.99)*
		female	14.79 (2.57)	14.28 (0.6)	13.96 (1.3)	*14.34 (1.61)*
	*EDU*AGE*		*14.45 (2.28)*	*15.2 (1.82)*	*14.32 (1.22)*	*14.66 (1.81)*
H1*-A3*	high	male	14 (3.65)	9.44 (2)	12.38 (2.34)	*11.94 (3.21)*
		female	11.96 (4.28)	12.67 (3.24)	15.43 (4.03)	*13.35 (3.91)*
	*EDU*AGE*		*12.98 (3.9)*	*11.05 (3.06)*	*13.9 (3.5)*	*12.65 (3.59)*
	low	male	9.34 (6.6)	16.4 (2.91)	13.87 (3.66)	*13.2 (5.28)*
		female	13.77 (3.45)	15.1 (1.18)	12.39 (1.67)	*13.76 (2.43)*
	*EDU*AGE*		*11.56 (5.49)*	*15.75 (2.2)*	*13.13 (2.8)*	*13.48 (4.05)*
CPP	high	male	22.06 (2.08)	23.34 (1.14)	24.41 (1.13)	*23.27 (1.72)*
		female	23.56 (2.47)	23.49 (0.33)	23.06 (1.37)	*23.37 (1.54)*
	*EDU*AGE*		*22.81 (2.29)*	*23.41 (0.79)*	*23.73 (1.38)*	*23.32 (1.6)*
	low	male	23.54 (1.16)	21.56 (1.03)	22.86 (1.21)	*22.65 (1.35)*
		female	23.46 (2.27)	23.97 (1.51)	24.08 (1.43)	*23.84 (1.67)*
	*EDU*AGE*		*23.5 (1.7)*	*22.76 (1.76)*	*23.47 (1.4)*	*23.24 (1.61)*

For female speakers, education levels have effect on pitch profile but not on the spectral representation, and the effect is not age-specific. Female speakers who attained a higher education level, is speaking with a greater pitch span than their peers with a lower education level. This difference in pitch span arises because of the adoption of a lower pitch floor, which is rather extreme because it is as low as their male peers. While education effect has no significant impact on the spectral representation of the voices of female speakers, such an extreme pitch floor suggests the adoption of vocal fry by the higher-educated females.

While the effect of education level on the way HKC speakers produce their native language may seem inconclusive at first glance, we suspect that the main drive of education effect comes from the difference in English proficiency across education levels. This is evident if we consider the variety of English that is influential in Hong Kong at different periods, together with the previous findings we reviewed in §1.1.2.

Consider the case of senior male speakers first. By the time our senior participants received education, Hong Kong was still under British rule, and that RP English, which is the standard form of British English, was considered the prestige variety in Hong Kong, since the mastery of RP English signalled the ability to legislative and administrative power in Hong Kong. Also, standard English education was still limited to the elites since mass education was not implemented until 1978, which by this time our senior participants have completed secondary education already. Hence, the higher-educated senior speakers who are more proficient in English should be exposed to RP English at large by the time they received education in the tertiary level. Recall that Henton and Bladon (1988) revealed that male speakers of RP English were considerably breathier than speakers of some northern varieties of British English, and that the effect of speaking RP English on male speakers blurred the expected gender difference in voice quality. Our findings regarding higher-educated senior male speakers actually paralleled Henton and Bladon's finding since (1) they have breathier voice quality than those who are far less proficient in English, and (2) they are as breathy as their female peers, whether or not these female speakers attained a high education level. Considering all these as a whole, our findings seem to reveal the influence of British English to senior male speakers.

However, the popular variety of English in Hong Kong has shifted to American English in recent years. According to Edwards (2016), America media is influential in Hong Kong, due to the increasing ease of access via various platforms (YouTube, Netflix, Hollywood movies etc.); this leads to the findings that there is an increasing adoption of American English features (rhoticity, and flapping in particular) by undergraduate Hongkongers, regardless of their language preference. Considering the case of our higher-educated female participants, we also find parallelisms between our findings and findings of Yuasa (2010) and Podesva (2013): our findings revealed an increased adoption of a pitch floor as low as their male peers by the higher-educated female speakers relative to their lower-educated peers, and possibly signalled a more frequent use of vocal fry; Yuasa (2010) and Podesva (2013) revealed also the prevailing trend of the adoption of a creakier speaking style by female American English speakers in California, and in Washington Metropolitan Area respectively. According to Yuasa (2010), Californian English is the major variety in the American TV series, and Hollywood films, and hence this variety has strong influence on other varieties of US English (p. 331). Our findings somewhat resemble Podesva (2013) better, in the sense that the adoption of an extra-low pitch floor is widespread among the higher-education female speakers, regardless of their age. Given this parallel trend to Podesva (2013), and to a less extent to Yuasa (2010), the difference in voice quality between female speakers of the two education levels suggests again it is the result of a transfer of American English voice quality to the HKC of this group of speakers.

4. Concluding remark

This study sets up the normative voice quality profile of HKC speakers who are essentially born and raised in Hong Kong. There are inherent differences in voice quality between males and females, and speakers of different ages that arise from physiological differences and changes. Comparing HKC speakers of different education levels, and hence different attainments of English proficiency, reveals the possibility to transfer even the most persistent phonetic feature of English voice (the speakers' L2 and non-dominant language), to HKC voice (their native and dominant language). In particular, higher-educated senior male speakers were influenced by RP English, and showed a breathier voice quality than lower-educated peers. Higher-educated female speakers, on the other hand, were influenced by American English, and speak with a very low (male-like) pitch floor, thereby showing a creaky voice quality. Our study provides new discovery regarding crosslinguistic influence: a non-dominant language, both at the individual and societal level, can influence one's dominant language, and studying crosslinguistic influence requires considerations on the variety of contact language popular in the community. In the future, we shall look into issues like language attitudes, and adopt more advanced technologies, e.g. ultrasound, MRI, to better understand the physiological, psychological and sociolinguistic drives of changing HKC speakers' voice quality.

REFERENCES

ABERCOMBIE, D. (1967). *Elements of general phonetics*. Edinburgh University Press.

BAHMANBIGLU, S. A., MOJIRI, F., & ABNAVI, F. (2017). The impact of language on voice: an LTAS study. *Journal of Voice*, *31*(2), 249. e9-249. e12.

BOERSMA, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341–345.

CALLIER, P. R. (2013). *Linguistic context and the social meaning of voice quality variation*. Georgetown University.

DULAY, K. M., TONG, X., & MCBRIDE, C. (2017). The Role of Foreign Domestic Helpers in Hong Kong Chinese Children's English and Chinese Skills: A Longitudinal Study. *Language Learning*, *67*(2), 321–347.

DUNN, P. S.-Y. (2016). The accent of international schoolers (國際學校口音). Retrieved from https://www.thestandnews.com/culture/國際學校口音/

EADY, S. J. (1982). Differences in the F0 patterns of speech: Tone language versus stress language. *Language and Speech*, *25*(1), 29–42.

EDWARDS, J. G. H. (2016). Accent preferences and the use of American English features in Hong Kong: a preliminary study. *Asian Englishes*, *18*(3), 197–215.

ESLING, J. H. (2000). Crosslinguistic aspects of voice quality. In R. D. Kent & M. J. Ball (Eds.), *Voice quality measurement* (pp. 25–35).

ESLING, J. H., & WONG, R. F. (1983). Voice Quality Settings and the Teaching of Pronunciation. *TESOL Quarterly*, *17*(1), 89–95. http://doi.org/10.2307/3586426

EVANS, S. (2016). *The English language in Hong Kong: diachronic and synchronic perspectives*. London: Palgrave Macmillan.

FUNG, R. S.-Y. (2015). Voice quality: a preliminary study on the phonetic distinctions of two Cantonese accents. In *Proceedings of 18th International Congress of Phonetic Sciences*. Glasgow.

GARELLEK, M. (2015). Perception of glottalization and phrase-final creak. *The Journal of the Acoustical Society of America*, *137*(2), 822–831.

GOBL, C., & CHASAIDE, A. N. (2010). Voice Source Variation and Its Communicative Functions. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed., pp. 378–422). Blackwell Publishing Ltd.

HANSON, H. M., STEVENS, K. N., KUO, H.-K. J., CHEN, M. Y., & SLIFKA, J. (2001). Towards models of phonation. *Journal of Phonetics*, *29*(4), 451–480.

HENTON, C., & BLADON, A. (1988). Creak as a sociophonetic marker. In L. Hyman, V. Fromkin, & C. N. Li (Eds.), *Language, speech and mind: Studies in honour of Victoria A. Fromkin* (pp. 3–29). London: Routledge.

HILLENBRAND, J., CLEVELAND, R. A., & ERICKSON, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech, Language, and Hearing Research*, *37*(4), 769–778.

IWASHITA, N., BROWN, A., MCNAMARA, T., & O'HAGAN, S. (2008). Assessed levels of L2 speaking proficiency: How distinct? *Applied Linguistics*, *29*(1), 24–49.

KEATING, P., & ESPOSITO, C. (2007). Linguistic Voice Quality. *UCLA Working Papers in Phonetics*, *105*, 85–91.

KIM, J. Y. (2017). Voice quality transfer in the production of Spanish heritage speakers and English L2 learners of Spanish. In S. Perpiñán, D. Heap, I. Moreno-Villamar, & A. Soto-Corominas (Eds.), *Romance Languages and Linguistic Theory 11: Selected papers from the 44th Linguistic Symposium on Romance Languages (LSRL), London, Ontario* (pp. 191–206). John Benjamins Publishing Company.

KLATT, D. H., & KLATT, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America*, *87*(2), 820–857.

KREIMAN, J., & SIDTIS, D. (2011). Physical Characteristics and the Voice: Can We Hear What a Speaker Looks Like? In *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception* (pp. 110–155). Walden, MA: Wiley-Blackwell Publishers.

LAVER, J. D. M. (1968). Voice quality and indexical information. *British Journal of Disorders of Communication*, *3*(1), 43–54.

LI, D. C. S. (1999). The functions and status of English in Hong Kong: A post-1997 update. *English World-Wide*, *20*(1), 67–110.

LI, D. C. S. (2017). *Multilingual Hong Kong: languages, literacies and identities*. Cham: Springer International Publishing.

LOVEDAY, L. (1981). Pitch, politeness and sexual role: An exploratory investigation into the pitch correlates of English and Japanese politeness formulae. *Language and Speech*, *24*(1), 71–89.

LUKE, K. K., & RICHARDS, J. C. (1982). English in Hong Kong: functions and status. *English World-Wide*, *3*(1), 47–64.

MAJEWSKI, W., HOLLIEN, H., & ZALEWSKI, J. (1972). Speaking fundamental frequency of Polish adult males. *Phonetica*, *25*(2), 119–125.

MENNEN, I., SCHAEFFLER, F., & DOCHERTY, G. (2007). Pitching it differently: A comparison of the pitch ranges of German and English speakers. In *16th International Congress of Phonetic Sciences* (pp. 1769–1772).

NEWMAN, M., & WU, A. (2011). "Do you sound Asian when you speak English?" Racial identification and voice in Chinese and Korean Americans' English. *American Speech*, *86*(2), 152–178.

NG, M. L., CHEN, Y., & CHAN, E. Y. K. (2012). Differences in vocal characteristics between Cantonese and English produced by proficient Cantonese-English bilingual speakers—a long-term average spectral analysis. *Journal of Voice*, *26*(4), e171–e176.

PENNINGTON, M. (2005). *The phonetics and phonology of glottal manner features*. Indiana University Bloomington.

PODESVA, R. J. (2013). Gender and the social meaning of non-modal phonation types. *Proceedings of 37th Annual Meeting of the Berkeley Linguistics Society*, 427–448.

SETTER, J., WONG, C. S. P., & CHAN, B. H. S. (2010). *Hong Kong English*. Edingurgh: Edinburgh University Press.

SHUE, Y.-L., KEATING, P., VICENIK, C., & YU, K. (2011). VoiceSauce: A program for voice analysis. *Proceedings of the ICPhS XVII*, 1846–1849.

STATHOPOULOS, E., HUBER, J., & SUSSMAN, J. (2011). Changes in acoustic characteristics of the voice across the life span: measures from individuals 4–93 years of age. *Journal of Speech, Language, and Hearing Research*, *54*(4), 1011–1021.

SZAKAY, A. (2012). Voice quality as a marker of ethnicity in New Zealand: from acoustics to perception. *Journal of Sociolinguistics*, *16*(3), 382–397.

THOMAS, E. (2010). *Sociophonetics: an introduction*. Palgrave Macmillan.

TODAKA, Y. (1993). *A cross-language study of voice quality: bilingual Japanese and American speakers*. University of California, Los Angeles.

TORGERSON, R. C. (2005). *A Comparison of Beijing and Taiwan Mandarin Tone Register: An Acoustic Analysis of Three Native Speech Styles*. Birgham Young University.

YUASA, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, *85*(3), 315–337.

ZHANG, Z., KREIMAN, J., GERRATT, B. R., & GARELLEK, M. (2013). Acoustic and perceptual effects of changes in body layer stiffness in symmetric and asymmetric vocal fold models. *The Journal of the Acoustical Society of America*, *133*(1), 453–462.