Intonation of Hong Kong English Polysyllabic Words in Young and Older Demographics



Candidate number: 1079696

Jane Xuzhen Tang

Word count: 15,227

Thesis submitted in partial fulfilment of the requirements for the Degree of

Master of Philosophy in Linguistics, Philology and Phonetics

University of Oxford

Trinity 2025

Acknowledgement

I owe my deepest gratitude to Peggy Mok and Grace Cao for doing the work of designing and conducting recording sessions to collect the corpus data used in this thesis. They trained me to process and annotate the data when I was an assistant for their projects, where I gained valuable experience in speech data processing for linguistic research. After we parted ways when my pursuit for further education took me to the UK, they very kindly and generously still allowed me access to part of their Hong Kong English data in the corpus as the basis for my MPhil thesis. Without their work and the opportunities they offered me, this thesis would not have existed in the form that it does today.

I am very grateful to my supervisors for their generous expertise and their endless patience; they tried their best to gently push me while being so understanding and accommodating of my difficult circumstances. I would also like to thank the following people for their personal support during my degree: Cath, my student rep who helped me navigate systems at the university; Xinlin, my favourite Cantonese speaker and simply one of my favourite people overall; my parents, who loved me enough to let me go, leaving them and my homeland behind in search of a better life. Finally, to A.C., who wants to read my thesis despite not being a linguist: You are part of that better life I didn't know I'd find. Thanks for everything, and I hope you find my thesis interesting.

1 Introduction

The English language has been long in use by L1 Cantonese speakers in Hong Kong, and their speech has distinctive and systematic features, which developed in its interaction with Cantonese and global varieties of English spoken in the region. It is known that people in Hong Kong speak English with roughly two types of intonation patterns, one with stronger Cantonese influence than the other. Lam (2020) thus proposed a unified approach to Hong Kong English (HKE) intonation from a variationalist perspective by treating HKE as a collection of sub-varieties on a spectrum between globally- and locally-oriented HKE. Globally-oriented HKE (HKE-gl) has similar intonation patterns to inner circle Englishes, while locally-oriented HKE (HKE-lo) borrows from Cantonese tones and assigns them as pitch accents (H(igh), M(id), and L(ow)) to each syllable. Hong Kong English is spoken in the a rapidly changing political and linguistic context, and younger and older generations in Hong Kong acquired and use English in very different contexts, and is likely reflected thus in their intonation. For this thesis, I looked at the polysyllabic intonation patterns of 20 young and old female HKE speakers based on data from reading tasks and mock interviews from a forensic linguistic corpus, and investigated to what extent the two age groups differ and where the they are located on the global-local intonation spectrum of HKE. I found that HKE intonation is complex and varied, with younger speakers having a wider distribution on the HKE-gl-HKE-lo spectrum than the older speakers, which reflects the sociolinguistic changes that took place in Hong Kong in the past decades since as early as the 1960s.

The rest of this thesis begins in Section 2 with an overview of previous literature on the sociolinguistic background of HKE, its intonation patterns and how it consists of

a spectrum of varieties with bilingual influences, after which a list of research questions are posed in Section 3; Section 4 describes the methodology of the corpus itself as well as the extraction and analysis of data from the corpus; the results section covers descriptive findings of pitch accent patterns and other intonational phenomena, as well as comparisons between age groups divided by a list of variables including intonational context, stress pattern type, and task; Finally, I discuss the phonological patterns and phonetic phenomena in this speaker group confirming most existing descriptions of HKE intonation while challenging others, compare speakers by age group while taking sociolinguistic factors into account, and argue for the benefits of a variationalist approach to HKE intonation.

2 Literature review

2.1 General background: History and identity

Hong Kong was annexed by Britain in 1842 from China after the Opium War, and it remained a colony of Britain for 155 years until its handover to the People's Republic of China in 1997. It became a partially autonomous Special Administrative Region (SAR) under the 'one country two systems' principle. Being a former British colonial administration, Hong Kong continues to have English as one of its official languages. It now has an official policy of being 'trilingual' (Cantonese, Mandarin, and English) and 'bi-literate' (English and written Chinese). HKE developed as a result of the bilingual environment in Hong Kong in this historical process.

Over 90% of Hong Kong's population self-identify as ethnically Chinese, and most of them speak Cantonese as their first language, while the percentage of English speakers in the population grew over the past three decades. According to Hong Kong

1996 Population By-census , around 3.1% of the total population reported to speak English as their ‘usual language’, and 29.1% spoke English as an additional language to their usual language (Census and Statistics Department HKSAR 1996). In 2021, 4.6% spoke English as their usual language, 54.1% spoke English as an ‘other language’, and in total, 58.7% of the population was able to speak English (Census and Statistics Department HKSAR 2021). English is taught as a subject in primary school and onward, and it is the medium of instruction in some secondary schools and in all universities. It is also a language used in government and corporate workplaces. HKE was originally doubly influenced by British English and Cantonese (a consensus in the literature, e.g. Hung 2011). Nowadays, North American English has an increasingly strong influence on younger HKE speakers due to the presence of American media and Internet cultures.

Cantonese-English bilingual speakers in Hong Kong of different age groups experienced their English language education in dramatically different ways. Since the annexation of Hong Kong in 1840, Western missionaries set up mission schools, many of which constitute the oldest and most prestigious secondary and primary schools in Hong Kong today. During the late British colonial period from the 1960s to 1997, Hong Kong saw enormous economic growth, which resulted in the growth of a ‘new middle class’ consisting of white-collar workers for the corporations that emerged rapidly during the economic boost (Leung 1996). The growth of the middle class paralleled the transformation of public education (Bolton 2000). In the 1960s, English language education was associated with the elite, who would send their children to prestigious primary schools and secondary schools, where they received an education in English, and then progress to either Hong Kong University or the Chinese University of Hong Kong, where English was used as a medium of instruction. From the 1970s, public

education reforms introduced compulsory primary and secondary schooling, providing the public with access to English language education and in some cases, English-medium education, shifting the linguistic landscape of Hong Kong from *elitist bilingualism* to *mass bilingualism* (Bratt-Paulston 1980), but also creating more variety in backgrounds in English language education. Secondary education in Hong Kong can be roughly divided into local government-funded schools and international private schools, the former providing an English language education in the classroom, sometimes in the English medium, while the latter providing an integrated English language environment with strong influences of British and North American English from the students and teaching staff alike. In the 2000s, increasing numbers of middle-class Chinese families are able to afford sending their children to international schools, which was originally meant for children of British expats, so that their children may receive a prestigious English-medium education (Forse 2010). Chinese graduates of international schools nowadays typically speak English with features of native-like North American accent; while they identify as native speakers of Cantonese via inheritance, they use English more frequently at home compared to their peers from the local government school education, and they are more likely to also identify as native speakers of English (Edwards 2018).

Since the change of sovereignty in 1997, Mandarin began to integrate into Hong Kong's language planning, which interacted with the shifting attitudes the population had towards the English language. In the wake of the handover, local attitudes toward English still preferred a 'native' exonormative model, regarding British English as the desired norm and HKE as a learner variety featuring 'errors' (Luk 1998). However, two decades on, as English increasingly served as an ethnolinguistic marker that distinguished a Hong Kong identity from the Chinese mainland in the interplay of

localisation, mainlandisation, and internationalisation (Lai 2001; Lai 2011; Lai 2012), the local endonormative model of English norms also began to gain popularity amongst speakers who identify strongly with being from Hong Kong and Cantonese. Edwards (2018) found that, while coexisting with the older exonormative attitudes, HKE increasingly gained symbolic value and has been accepted as an in-group marker of local identity, used to express solidarity with other HKE speakers.

2.2 HKE as a World English variety

Cantonese influence in HKE is reflected in its phonology, syntax, and semantics. Hung (2000; 2011) suggests that phonology is the only area that offers strong evidence for HKE being a distinctive variety of English with its own stable and systematic linguistic features, although his works only focus on segmental features. A prosodic comparison between HKE and Guangzhou Cantonese-accented English by Ran, Van De Weijer & Sloos (2020) suggests that the non-segmental features of HKE are largely similar to that of a general Cantonese-accented English. Ran, Van De Weijer & Sloos investigated the general question of whether Hong Kong English is a standardised English variety by comparing the intonation patterns between Hong Kong English and Guangzhou Cantonese-accented English. They recorded 18 L1 speakers of Guangzhou Cantonese in English in a picture description task targeting statements, wh-questions, yes-no questions, and echo questions to compare them with existing descriptions of HKE intonation patterns in the literature.

The researchers also recorded the same four types of sentences in Guangzhou Cantonese. They found that intonation patterns of Guangzhou Cantonese and Hong Kong Cantonese share the same intonation system, and that the intonation patterns of the four sentence types are largely similar between HKE and Guangzhou-accented

English. In statements, Guangzhou-accented English is only different from HKE in terms of the presence of declination in long sentences, i.e. the former shows no declination in long statements, unlike HKE. The two varieties both show a low boundary tone on the final syllable and declination in short statements. For wh-questions, the researchers could not observe any significant differences between HKE and Guangzhou-accented English. Both begin with a high boundary tone and a following fall, and both show variation between a high final boundary tone and a low final boundary tone. In yes-no questions, Guangzhou-accented English, like HKE, have a H% and a final rise starting at the penultimate syllable. The only difference observed was that Guangzhou-accented English showed less declination than HKE. Finally, the two varieties also share similar intonation patterns in echo questions, with Guangzhou-accented English only differing in the presence of declination as in statements. Ran, Van De Weijer & Sloo interpreted these results as indication of HKE as more a Cantonese-accented second language than a standardised English variety unless linguists rethink the role of linguistic structure in the classification criteria for an English variety, taking into account, for example, the role of ethnolinguistic identity and language use.

In sociolinguistics, HKE is often discussed under the Three Circles Model of World Englishes proposed by Braj Kachru in 1985. Under this model, Englishes spoken in North America, the UK, Australia, and New Zealand are considered *inner circle* Englishes. In these regions, the dominant English variety is spoken by 'native speakers' and considered to be 'standard', while the *outer circle*, which is the second circle, contains Englishes typically spoken in regions with colonial ties to Britain and where English is institutionalised. Examples of Outer Circle Englishes include Indian English and Nigerian English. English varieties from the Outer Circle are not norm-providing despite being widely used and often spoken by L1 speakers. The third circle, the

expanding circle, includes regions like China and Indonesia with rapidly increasing numbers of English speakers, where typically English is taught as a second language but is not necessarily an official language in the policy. With its relatively distinct phonological system, an identifiable accent, and unique vocabulary, HKE resembles the Englishes in the outer circle in many ways. English has an official status in Hong Kong and is institutionalised in the government and education system. Additionally, the city has a colonial history tied to Britain, the same as other outer circle regions such as India and Nigeria. Kachru noted that the expanding circle and the outer circle are not clearly distinct categories. It is still under debate whether HKE is a true variety of English of its own, and it is still unclear what influence recent political changes and language policy changes in Hong Kong may have on the development of HKE, but the attitudes towards HKE has been in flux, shifting from an exonormative model to an endonormative one due to the development of a strong local identity while attitudes toward Mandarin also became more positive as its visibility and utility increased since the handover (Edwards 2018; Lai 2012). With Cantonese remaining as a robust language in daily life but gradually existing the education domain and the increasing presence of Mandarin and English in education and in the workplace, HKE is under the influence of a different linguistic landscape from before the handover, and it may be reflected in the differences in intonation between younger and older speakers.

2.3 HKE intonation and its influences

2.3.1 Descriptions of HKE intonation and the question of tone

Previous literature on the intonation of HKE and its prosodic typology show that it has its distinct intonational features and has great internal variety. Gussenhoven (2012; 2014) proposed HKE intonation rules based on Luke (2000) and Luke (2008). Luke

observed that HKE borrows at least three lexical tones from Cantonese: High (H), Mid (M), and Low (L). Matthews & Yip (2011) described the tonal system of Hong Kong Cantonese as having 6 tones (5 being the highest level and 1 the lowest):

High level 55

High rising 25

Mid level 33

Low falling 21/11

Low rising 23

Low level 22

When pronounced in isolation, the low falling tone has a 21 contour, although the 11 contour also occurs. Traditional descriptions further distinguish the *entering* tones, which occur before unreleased stop consonants. They are phonetically equivalent to short instances of high, mid, and low level tones. Luke observed that the intonation of HKE borrow the High Level, Mid Level, and Low Level tones from Cantonese as pitch accents, assigned to every syllable in the utterance. Gussenhoven's (2012) proposed intonational rules based on this observation are as follows: (2012: 4)

1. Pitch accent: H (assigned to primary and pre-primary stressed syllables and some function words)
2. Insert L after last H in the IP
3. Place M on empty word-initial syllable
4. Copy tones right

The boundary tones in different sentence types are:

1. Declarative: Ø
2. Emphatic declarative (elicited by the exclamation mark in writing): L%

3. Interrogative: H%

Gussenhoven proposed this 'accentual analysis' of HKE intonation where lexical items are toneless and instead listed with accented syllables which are assigned pitch accents at the post-lexical stage. The accentual analysis was an alternative to Luke's (2000) account as shown below, simplified by Gussenhoven (2012: 2):

1. Place H*L on primary stressed syllable and any preceding stresses
2. Raise L to H if followed by H
3. Place M on function words and word-initial toneless syllables
4. Copy tones to right empty syllables

In the accentual analysis, the L-raising at step two in Luke's analysis will not be needed in the accentual analysis, and post-lexical assignment of H-tones to function words can take place before default M-tone assignment. Although he argued that the accentual analysis was viable, Gussenhoven continued the metaphor of 'tone assignment' to indicate the historical origin of the pitch accents. He also argued in 2014 that varieties of English with substrate tone languages such as HKE, SNE, and Ghanaian English are tone languages under Hyman's (2011: 350) definition that a tone language is a language where 'an indication of pitch enters into the lexical realization of at least some morphemes'. The examples that Gussenhoven gave are homophonic content words and function words that only differ in tone.

Gussenhoven conducted a production experiment where he asked a HKE speaker to read out loud words in declarative (no punctuation), emphatic declarative (ending in an exclamation mark '!'), and interrogative (ending in a question mark '?') conditions. For words with final main stress, he found that the final syllable had a level pitch, a pitch fall, and a pitch rise respectively across the three conditions. For words

with penultimate or ante-penultimate stress, he found that the final syllable had a low pitch in both declarative and emphatic declarative conditions, and had a low-to-high rise in the interrogative condition. Based on the elicitation results, Gussenhoven proposed that HKE has no final boundary tone in declarative IPs, a L% in emphatic declarative sentences leading to a fall (for unstressed final syllables, the fall was hypothesised to be steeper than in declarative sentences), and a H% in interrogative sentences leading to a rise. He later added a HKE L% deletion rule where the IP-final syllable has a L-tone, since a separate experiment found that an 'emphatic' fall with a steeper falling contour was not considered a significant difference and sometimes deemed unacceptable.

Stronger claims analyse HKE as a two-tone system, where H-tones are lexically encoded and assigned to stressed syllables, and the 'tones' in other syllables may phonetically surface as L or M. Contrary to the accentual analysis where 'tone-assignment' takes place post-lexically, Wee (2016) proposed that H-tones are 'lexically assigned' to certain syllables in a word, word-initial syllables are marked by M-tones, utterance-final positions are marked by a low boundary tone. H-spreading and interpolation derive the intonation of the remaining syllables, and thus not every syllable in HKE is assigned a tone/pitch accent. Cheung (2009) proposed that the span of H-tones in HKE were best represented as a single H associated with multiple syllables within the domain defined by stressed syllables. In this analysis, M tones are a surface result of the phonetic default. Yiu's (2014) proposal that HKE has a privative tone system of H and Ø with intonational boundary tones H% and L% in questions and statements respectively. The surface realisation of Ø is analysed to be sensitive to its position relative to H tones and word and utterance boundaries, thus resulting in what can be detected as a M vs. L distinction.

The privative tone system analysis can be supported by Yeung's (2023) comparison between speakers of Hong Kong English (HKE) and American English (AE), leading to preliminary evidence for tonogenesis in HKE triggered by language contact. Speakers read nine pairs of homophonic content words and function words in the carrier phrase 'say ___ again'. Yeung found that HKE speakers produced content words with significantly higher F0 than the function words, while there were no effect of syntactic category on F0 for AE speakers. This was presented as evidence for the hypothesis that HKE is a tone language with two contrastive lexical tones like in Yiu's and Wee's proposals. However, Yeung does not engage with the hypothesis of Cantonese tone borrowing; instead, since function words in non-tonal varieties of English often receive no pitch accent, and that content words are often stressed and receive a H* pitch accent, Yeung posited that the source of the tonogenesis may be the reinterpretation of F0 correlates of the non-tonal English stress system by L1 Cantonese speakers. Similar observations of the different pitches function words have from their content word homonyms have also been mentioned by Gussenhoven (2012) and Wee (2016).

2.3.2 HKE as a spectrum of varieties

Studies on HKE intonation assume a single variety of HKE that has its own features distinguished from Inner Circle Englishes such as British English and North American English. However, it has long been noted that HKE in actual use spans from minimally influenced by Cantonese and perceived to be native-like, to strongly influenced by Cantonese and having strong intonational features described in the literature summarised above. Earlier studies such as Bolton & Kwok (1990) conceptualised the HKE accent variation as a continuum: on the one end of the continuum is a variation

with ‘idealised’ HKE phonological features (‘broad Hong Kong accent’), on the other end is a variation close to what is considered native-like (‘educated Hong Kong accent). Lam (2017) renamed the two variations as HKE-lo (HKE with local orientations) and HKE-gl (HKE with global orientations) to capture the impact of globalisation on indexicality, identity, and language ideology pertaining to HKE, because neither education or proficiency can predict very well a HKE speaker’s pronunciation.

Lam conducted a study on intonational features of HKE variations and analysed how speakers style-shift and use intonation for sociolinguistic purposes. The data were all collected from young speakers at a university in Hong Kong who all attended local secondary schools. The tasks ranged from sentence and long text reading tasks, one-on-one spontaneous interactions with a non-HKE speaker, and interviews. Some of the topics of the tasks were related to Hong Kong culture. The researcher found that English proficiency was not a good predictor of the range of intonation the speakers demonstrated. Two speakers used both HKE-gl and HKE-lo intonation features during the tasks, while the other speaker only used HKE-lo features. He also found that HKE-gl intonation is similar to a ‘native variety’ of English, where only some stressed syllables receive a high pitch accent.

In addition, Lam found two violations of the HKE intonation rules proposed by Gussenhoven. Firstly, he elicited instances where L was not only inserted after the last H in the IP, but after an H in the middle of the sentence. Secondly, Lam found instances of the use of MH tone in a single syllable. Lexical items borrowed from Cantonese and sentence final particles were found to reveal tones that otherwise did not show up.

In Lam 2020, a doctoral thesis that includes the pilot study in Lam 2017 and its follow-up study, more speakers were recruited. Major characteristics of HKE-lo posited

in the pilot study was also found here. Lam found that only some syllables have pitch accent in HKE-gl, while every syllable is assigned a pitch accent in very strong HKE-lo. He also found that reading tasks related to Hong Kong culture and tasks involving compulsory sentence-final particles (SFP) were more effective at eliciting HKE-lo than the spontaneous tasks.

The results showed that in HKE-gl only some syllables have a pitch accent. They also confirmed many existing HKE-lo intonation features in the literature: 1) almost every syllable receives a pitch accent sourced from Cantonese tones H, M, and L; 2) the rules for HKE word prosody; 3) Luke's (2000) L-Raising, which seems to be an optional rule and therefore challenges the accentual analysis proposed by Gussenhoven (2012; 2014) that did not include L-Raising; and 4) common occurrence of L/L% in sentence-final intonation.

Lam found divergent and additional HKE-lo features compared to existing literature. Firstly, although almost all syllables in HKE-lo sentences were assigned a pitch accent, the percentage was not 100%. This supported the existence of a spectrum between extreme ends of local and global prosody in HKE. Secondly, occasional instances of MH and LM tones were found. These instances occurred as a result of 1) imitation of HKE-gl or native-like accent, 2) presence of Cantonese loan words, 3) English diphthongs analysed as two separate vowels, and 4) the use of contour tones sentence-finally, including in the presence of SFPs.

The model of HKE-lo intonation proposed by Lam is as follows: (2020: 160)

1. Stressed syllables are given the H tone.

2. Unstressed word-initial syllables are assigned the M tone.

3. Unstressed word-final syllables are assigned the L tone.

4. H can be optionally spread rightward to the adjacent syllable(s), replacing the word-final L tone(s).

5. Individual lexical items can optionally retain their original Cantonese pronunciations, including the use of auxiliary HKE tones, instead of following rules 1-4.

6. If an SFP is present, assign a tone based on its original Cantonese pronunciation.

7. For sentence-final intonation,

    a. Assign either the prolonged boundary tone ":" or L tone to declaratives.

    b. Assign H to interrogatives

    c. Assign either the prolonged boundary tone ":" or L tone to emphatic declaratives.

2.3.3 Cantonese and English influences

Apart from the main effect of the Cantonese tone system from which HKE speakers phonetically source their pitch accents and possibly an emerging tone system, Cantonese and English also have influences on other aspects of HKE intonation.

Gussenhoven (2012) noted that HKE has downstep if one or more M-tones occur between H-tones, although consecutive H-tones are pronounced on the same pitch. A similar effect is found in Cantonese on the High Level tone, which is said to be

the source of H tones in HKE. Gu & Lee (2009) conducted an experiment to investigate the effect of tone and emphatic focus on F0 in Cantonese. They asked 3 native speakers of Hong Kong Cantonese to read sentences that carried a disyllabic target with segmental features controlled and only varying in tones, creating 36 tone combinations. They created 3 focus conditions: without any narrow focus, focus on the first syllable, and focus on the second syllable (narrow focus conditions were created by prompt questions). They observed a number of effects of tone, including a downstep effect on High Level tones preceded by a non-H tone. The lower the preceding non-H was, the steeper the downstep. Gu & Lee found anticipatory raising to be present before low onset F0 when the preceding syllable has High Level, High Rise, and Low Rise tones, with High Rise creating the strongest effect. Gu & Lee compared the results with tonal coarticulation effects in other tonal languages (esp. Mandarin) and found them to be similar, while the emphatic focus effects were language-dependent. They argue that tonal coarticulation is realised phonetically due to universal physiological constraints and therefore are nearly identical across tone languages.

Chen & Mok (2015) investigated potential Inner Circle English and Cantonese influence in HKE sentence intonation. They recorded 19 L1 speakers of Cantonese in a sentence reading task where the speakers read out statements, wh-questions, yes-no questions, and echo questions in Cantonese and English. The researchers compared the intonation contours of HKE by sentence type with Cantonese and with English, and found mixed features of both Cantonese intonation and English intonation in HKE. In particular, HKE cues questions using final rising, which is similar to the pattern in Cantonese; however, HKE questions showed final rising starting from the penultimate syllable, whereas the final rising in Cantonese questions started from the second half of

the final syllable. Chen & Mok suggested that this was due to the influence of English using global rising in questions.

3 Research questions

The research questions of this thesis is as follows:

1. What does the data tell us about HKE intonation regarding polysyllabic lexical words?
    a) What intonational patterns emerge from the data?
    b) Do the patterns largely align with HKE-gl or HKE-lo?
2. What do the data tell us about the spectrum of global-oriented vs. local-oriented HKE intonation and its diachronic development?
    a) How do HKE-gl and HKE-lo intonation features interact?
    b) Are there differences between younger and older demographics in their HKE intonation and their positions on the HKE-gl-HKE-lo spectrum?

Note that this thesis does not intend to make claims on whether HKE is a true tone language, and 'tone assignment' and 'pitch accent assignment' are used interchangeably. The terminology of 'tone' is used to be consistent with the HKE intonation descriptions in previous literature and to reflect the historical origins of pitch accents in HKE, i.e., phonetic borrowing/transfer from Cantonese tones.

4 Methodology

This study is a corpus data analysis of the HKE data collected as part of a larger Cantonese-English-Mandarin trilingual forensic phonetics project by Grace Wenling

Cao and Peggy Pik Ki Mok (in preparation), which they kindly shared with me for the current study.

4.1 Participants

The data consists of twenty female HKE speakers, divided into two groups of ten young speakers and ten older speakers. Their data was selected from a larger cohort of mixed gender speakers in the original study. Only female speakers were chosen too eliminate the complications of an additional variable and the influence of gender differences in language change. The individuals were randomly chosen from female speakers belonging to a higher proficiency group.

The speakers all reported their L1 to be Cantonese, and that their parents' L1 was also Cantonese. The young speakers range from 18 to 22 years old, and the older speakers range from 61 to 72 years old. Nine out of the ten young speakers provided their Hong Kong Diploma of Secondary Education (HKDSE) English language levels, and they range from Level 5 to Level 5**, which is equivalent to IELTS overall band score range of 7.55-8.22 (Hong Kong Examinations and Assessment Authority 2023). One young speaker did not provide her HKDSE scores but reported her IELTS overall score to be 9.0. Only one older speaker reported her English language test results (TOEFL 590, test taken in 1976). Due to the lack of English language proficiency test data and the high heterogeneity of the contexts of acquisition and language use, Cao & Mok conducted a separate test by inviting L1 speakers of English and Hong Kong locals to listen to a 10-second recording of each participant and to rate the standardness of the participants' English accent. The scores range from 3.5 to 8.2 out of 10 for the young speakers (mean rating by L1 speakers is 5.7, mean rating by Hong Kong locals 5.95), and from 4.6 to 9.2 for the older speakers (mean rating by L1 speakers 6.8, mean rating

by Hong Kong locals 6.9). The participants' age of learning English was 0-12 years old (0-4 for the young speakers and 4-12 for the older speakers). The self-reported frequencies of language use are 90.95% Cantonese and 5.475% English on average.

4.2 Experiment

The complete procedure of the trilingual interviews is described in Cao & Mok 2023. The data relevant to this study was collected from the final tasks during the experiment where the participants finished a reading task in English and a mock interview with Cao based on the content of the reading task text. The participants read a short passage in English written like a statement from someone involved in an investment fraud case. The interviewer then role-played a British police officer interrogating them about their involvement in the case. The interviewer's English accent has clear Southern British English influence, and the participants know she speaks Cantonese as her L1 from previous tasks in Cantonese.

The experiment was recorded in a sound-proof booth. During the mock interview, which happened directly after the reading task, the participant and the interviewer sat facing each other by a desk, where a computer was placed showing the reading task passage. The participant could refer to the passage for information during the mock interview, and therefore vocabulary largely overlaps between the read speech data and interview data. An H4N recorder positioned around 30 cm away from the participant was used for recording. Each reading task lasted for around 3 minutes, and each interview lasted for about 6-10 minutes. Additionally, the participants completed a linguistic background questionnaire.

4.3 Data analysis

4.3.1 Choosing targets

Textgrid files with a word tier and a segmentation tier were generated for each recording using Montreal Forced Aligner (McAuliffe et al. 2017). A list of polysyllabic target words from the reading task and the mock interview were identified, listed below together with their HKE-lo intonation patterns predicted using the descriptions in the latest the literature, minus considerations for optional H-spreading or boundary tones. The words are sorted into stress pattern categories (written in single quotation marks) named after a word in the category. A rough survey of the data suggested that the status of secondary stresses is ambiguous, and thus the predictions include accounting for and not accounting for pre-primary secondary stresses, which (Gussenhoven 2012) assumes is treated the same way as primary stresses, but it may not be represented underlyingly for some speakers. Of particular interest are the compound words *british-made* and *chinatown*, which are of their own category, whose predictions are given for cases of both English compound stress (trochee) and phrasal stress (iamb). Another point of interest is *souvenir*, where the primary stress could fall on either the first or the final syllable depending on the English variety and individual speaker.

Targets	Possible HKE-lo intonation: with secondary stress	Possible HKE-lo intonation: without secondary stress	Position of primary stress	Stress pattern	Task
manager		HLL	1	'manager'	Reading and interview
properties		HLL	1	'manager'	Reading and interview
investment		MHL	2	'investment'	Reading and interview
celebrate		HLL	1	'manager'	Reading and interview
microsoft		HLL	1	'manager'	Reading and interview
souvenir	HHH	MMH	3	'souvenir'	Reading and interview
		HLL	1		
british-made	HHH	MMH	3	'chinatown'	Reading and interview
		HLL	1		
chinatown	HHH	MMH	3	'chinatown'	Reading and interview
		HLL	1		
company		HLL	1	'manager'	Reading and interview
consider		MHL	2	'investment'	Reading and interview
foreigners		HLL	1	'manager'	Reading and interview
consultant		MHL	2	'investment'	Interview only
photographer		MHLL	2	'photographer'	Interview only
polytechnic	HHHL or HHH	MMHL or MMH	3	'polytechnic' or 'polytech'	Interview only
university	HHHLL	MMHLL	3	'university'	Interview only

Figure 1: The predicted HKE-lo intonation of the target words as well as the position of their main stress, which stress pattern category they belong to, and which tasks they appear in.

4.3.2 Annotation for intonation

The alignment of the target words were then manually checked for accuracy and adjusted on Praat. Each token was given an impressionistic analysis of its intonation pattern facilitated with phonetic information, i.e., F0 values and contours, syllable

duration, and intensity values. Assignment of pitch accent can be typically identified by ear because I am a speaker of Hong Kong Cantonese and am sensitive to the phonetic connections between HKE pitch accents and the Cantonese tones they are sourced from. Acoustic signs of pitch accent assignment include a plateau of the F0 track for individual syllables and an abrupt jump of F0 value at the syllable boundary between two accented syllables.

Each token has also been marked with its position in relation to a large intonational boundary identified by a pause and largely coinciding with clause and sentence boundaries (labelled as initial, medial, final, and isolated; if a token is adjacent to an intonation boundary to both the right and the left, i.e. both IP-initial and IP-final, it is labelled as isolated). The list of polysyllabic words were also tagged with parts of speech and their contexts extracted from the reading task transcripts using a Python script written by Xu (2024a).

4.3.3 F0 extraction

Pitch objects were generated for every recording using the raw autocorrelation setting, and the pitch range was set to 35-600 Hz to accommodate both the high pitch range of the younger speakers and the near-consistent presence of creaky L tones in all speakers. Every token was then manually checked and adjusted if there were errors with path-finding, in particular:

a) In cases where a high pitch contour is mistakenly detected in a voiceless section, the section is manually set to be unvoiced;

b) In creaky L tone syllables where a high pitch contour is mistakenly detected, the contour is adjusted by moving it down by octaves until it reaches below 100 Hz (usually by two octaves), and then the contour is double checked by ear;

c) In cases where the algorithm failed to detect a path but candidates for a path are clearly present, the contour is manually marked and checked for accuracy by ear;

d) When the pitch contour of a part of the word is undetectable (i.e. there are no good pitch candidates) due to noise or weakness of the voice, but the whole token can still be used for impressionistic analysis, the detectable part of the pitch contour is left in and used for analysis.

F0 values at every 10 msec interval of each target token were then extracted using a Praat script modified from Arantes 2021. Time normalisation was achieved by extracting F0 values at 50 evenly divided intervals of each token using a different Praat script also modified from Arantes 2021. Tokens that are too weak to be detected are discarded. To compare different voice ranges, the F0 values in Hertz of each speaker were normalised by using the linear z-score (LZ) formula, where x_i is the F0 value from the sampling point to be normalised. The formula can also be expressed as $LZ = (x_i - M) / SD$, where M is the mean F0, and SD is the standard deviation around M. Visualisation was done using R (R Core Team 2022) and partially based on the scripts behind Xu's (2024) web-based tool originally designed for the visualisation of citation tone data. Statistical analysis was done using LibreOffice Calc and R. Smoothed average F0 contours are generated using LOESS local regression.

5 Results

5.1 Token counts

There are 545 tokens detected in total, with fixed numbers for each speaker from read tasks (11 words each speaker, each repeating once, 220 tokens in total) compared to the interviews (325 tokens in total), which has 4 additional words and some words may be repeated by the speaker. All of the tokens take place in statements. The tokens are evenly distributed between old and young groups, with 272 tokens from the older group and 273 from the younger group.

	'manager'	'investment'	'chinatown'	'university'	'souvenir'	'polytechnic'	'photographer'	'polytech'	Total
Medial	167	87	62	7	3	20	2	2	**350**
Final	50	53	18	21	20	2	17	1	**182**
Isolated	3	0	0	0	3	0	1	0	**7**
Initial	2	3	0	0	1	0	0	0	**6**
Total	**222**	**143**	**80**	**28**	**27**	**22**	**20**	**3**	**545**

Figure 2: Token counts by stress pattern and position.

Most of the tokens are of the 'manager' and 'investment' type with regards to HKE-gl stress patterns (222 and 143 respectively), followed by 80 compound tokens ('chinatown'), and then followed 'university', 'souvenir' and 'polytechnic', which are roughly found once in each speaker. This is expected, since 6 words belong to the 'manager' type (*manager*, *properties*, *celebrate*, *microsoft*, *foreigner*, and *company*), 3 words belong to the 'investment' type (*investment*, *consider*, and *consultant*), 2 to the 'chinatown' type (*british-made* and *chinatown*), while only one word each belong to the rest of the stress pattern types, which are named after the words themselves.

In terms of word position within the IP, most tokens are either IP-internal or -final, counting at 350 and 182 each. Only 6 tokens appear at the beginning of an IP,

usually after a long break in the middle of the sentence, and 7 tokens appear as there own intonational phrase, usually with long breaks before and after them or appearing as the answer to a question. Combining word position and stress pattern, the most common are IP-internal trisyllabic tokens stressed at the first syllable (167 tokens), followed by medial trisyllabic tokens stressed at the second syllable (87 tokens). In comparison, these two stress pattern types have a more even distribution for the IP-final position (50 and 53 tokens respectively). Due to the lack of tokens, the IP-internal and -initial tokens are grouped together as non-final, while isolated tokens are grouped together with the IP-final tokens. This is also so that the effect of any final boundary tones can be better observed and analysed.

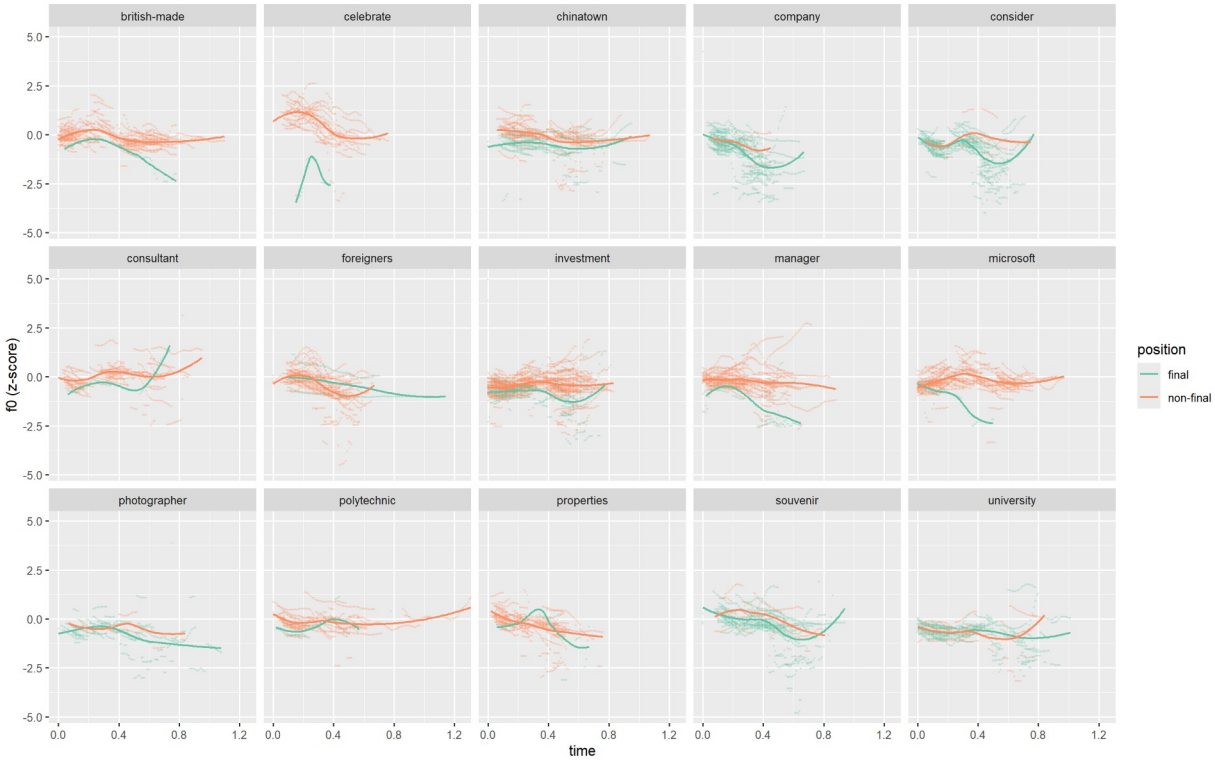


Figure 3: F0 values and smoothed contours of each word by position, in z-scores.

It should be noted that due to the nature of reading tasks and semi-scripted mock interviews, each word tends to appear in a specific sentential context and therefore a

specific position intonationally. For example, under the ‘manager’ stress type, *manager* tokens tend to appear in non-final positions, while *company* tokens tend to appear IP-finally.

5.2 Tone assignment

All speakers have at least one instance of a polysyllabic word where all syllables are assigned a pitch accent (‘fully-assigned’). 69.2% of tokens have every syllable assigned a tone (377 out of 545 tokens in total). The percentage by speaker ranges from 88.5% to 24.0%. Note that even the strongest HKE-lo speakers do not assign every syllable with a pitch accent, and even the strongest HKE-gl speakers pronounce some words in an HKE-lo intonation pattern.

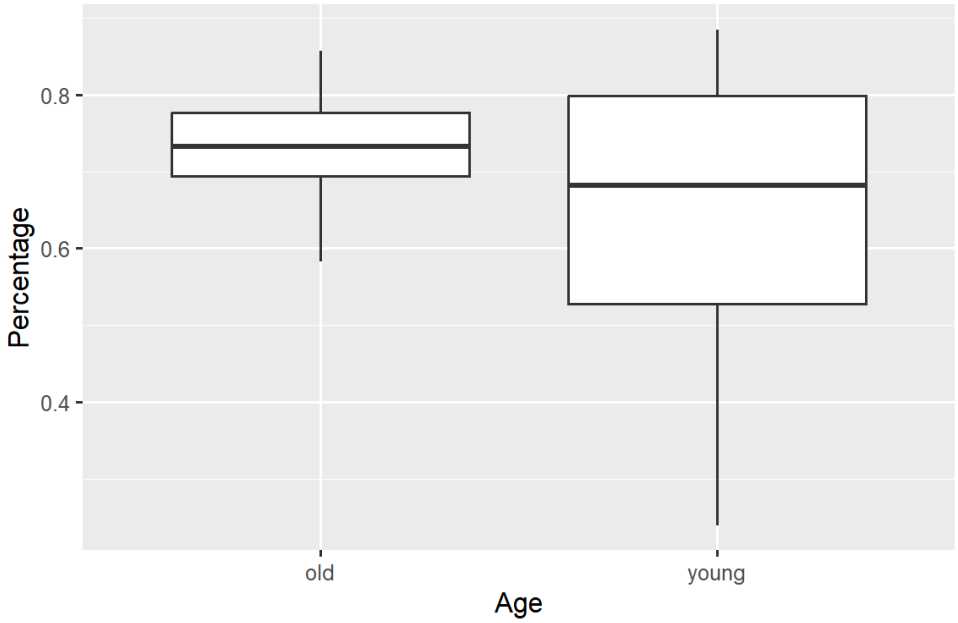


Figure 4: Percentages of 'fully assigned' tokens by speaker age.

An independent-sample t-test was conducted comparing older and younger speakers regarding the percentage of polysyllabic words they pronounced in a typical HKE-lo manner where every syllable is assigned a pitch accent. The t-test did not yield

significant results (t(12.042) = 1.24, p = .237; see also Figure 4). The mean percentage for older speakers is slightly higher than that of younger speakers (median = 73.2% vs. 64.5%), and older speakers have a narrower distribution (Q_1 and Q_3: 69.3% vs. 77.7% vs. 52.8% – 79.9%). Notably, both the highest and lowest percentages are found in the younger group. Therefore, at least in terms of tone assignment, younger speakers exhibit a slightly wider variety of their position on the local-global HKE intonation patterns, whereas older speakers are more concentrated on the more HKE-lo end of the spectrum, although the difference is not significant.

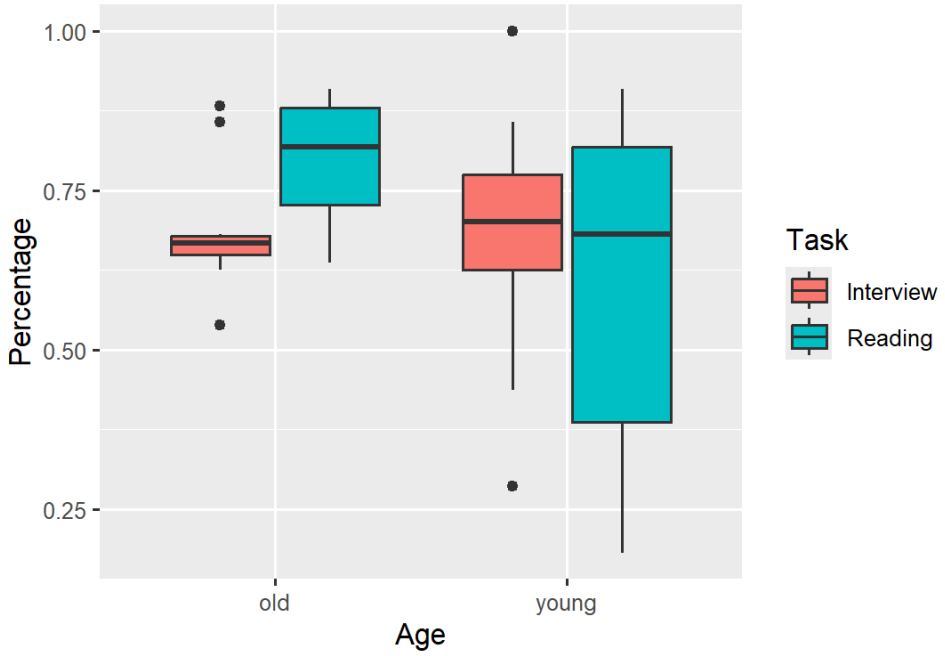


Figure 5: Percentages of 'fully assigned' words by speaker age and task.

In addition, when divided by task, wider distributions of percentages in tone assignment are found in the reading tasks (Q_1-to-Q_3 ranges for older and younger speakers respectively: 72.7% – 88.0% and 38.6% – 81.8%), compared to the more concentrated distributions in the interviews for both age groups (Q_1-to-Q_3 ranges respectively: 64.9% – 67.8% and 62.5% – 77.4%), the older group being more concentrated than the

younger one. Wider distributions in the younger speakers compared with the older speakers can be found across both tasks.

It should be noted that, amongst the 'fully assigned' tokens, some have more even distributions of intensity than others. Consider the following two HKE-lo instances of the word *consider* in the same phrasal context ('She needed more time to consider.') and both produced in the reading task by younger speakers. Both instances have a Mid-High-Low intonation pattern. The final syllable is next to a phrasal boundary, and thus it may be weakened at the end of the statement. In both instances, the final L-toned syllable is creaky, which also may affect its intensity. However, the intensity distributions between the first and the second syllables differ for these two tokens.

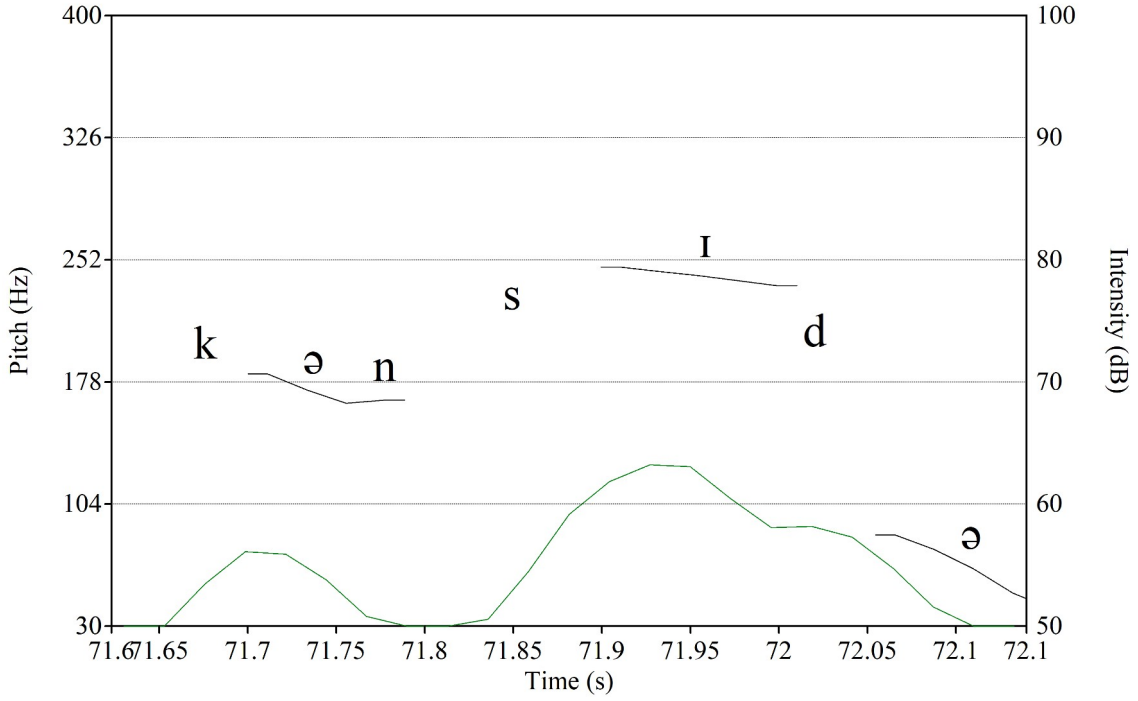


Figure 6: F0 and intensity of consider *for speaker HK14.*

For speaker HK14, it is clear that the second syllable is where the intensity peak of the whole word is located, sitting midway between 60 dB and 70 dB; whereas the initial syllable has a peak intensity roughly midway between 50 dB and 60 dB.

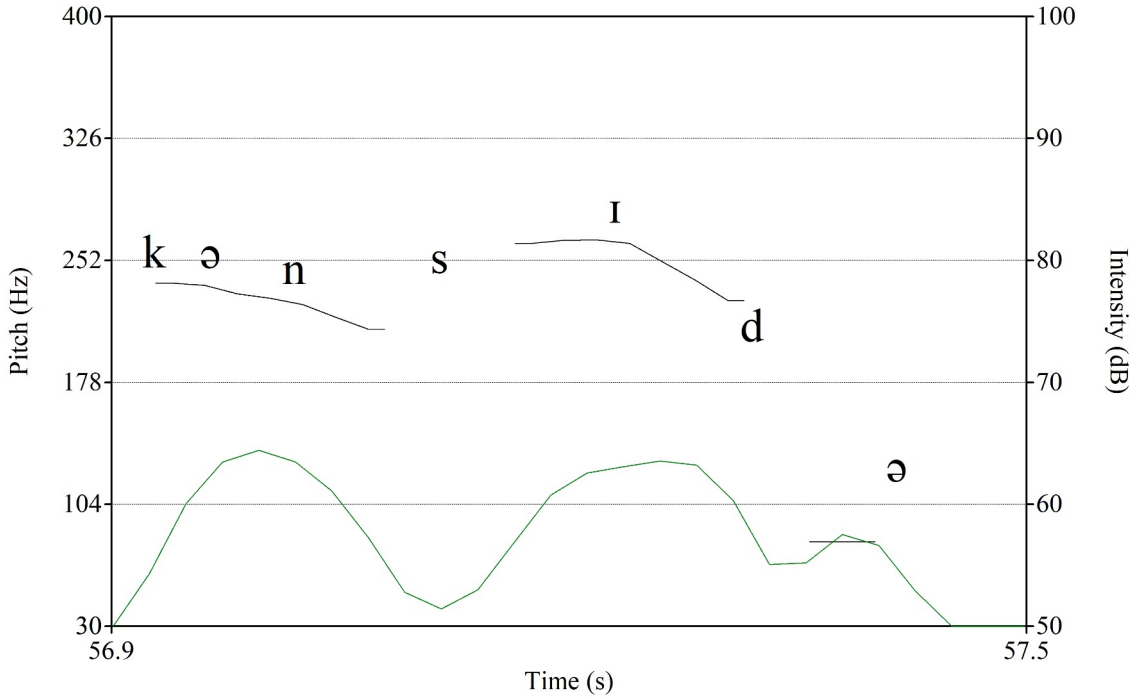


Figure 7: F0 and intensity of consider *for speaker HK49.*

In comparison, for speaker HK49, there is no obvious intensity peak across the whole word. Both the initial and the second syllables peak at around 65 dB. It should be noted that, for speaker HK49, the pitches of the first two syllables are closer to each other than speaker HK14. The higher pitch of the initial M-toned syllable may be correlated to its high intensity.

Apart from 'fully-assigned' tokens, there are some tokens that have a HKE-gl intonation pattern and a single identifiably stressed syllable. However, there are also a large number of 'in-between' tokens where more than one syllable is assigned a pitch accent, but not every syllable. Impressionistically, some of these 'in-between' tokens

have a more HKE-lo intonation, whereas others are closer to HKE-gl, depending on other factors including syllable intensity and voice quality. Tone assignment in HKE-lo and stress in HKE-gl are distinct phenomena when considering these two varieties as two ends of the spectrum. However, it is only by linking them via the 'in-between' tokens – i.e., functionally considering a HKE-gl token to be having one stressed syllable as an extremely impoverished case of tone assignment where only one or no syllable is assigned an H tone – that the HKE local-global spectrum can be quantitatively analysed in regards to tone assignment.

By the above definition of tone assignment, 1514 out of the total 1736 syllables across all tokens have been assigned a tone (87.2%). By-speaker percentages of tone-assigned syllables are higher than those of the 'fully-assigned' tokens across speakers, ranging from 64.6% to 96.3%.

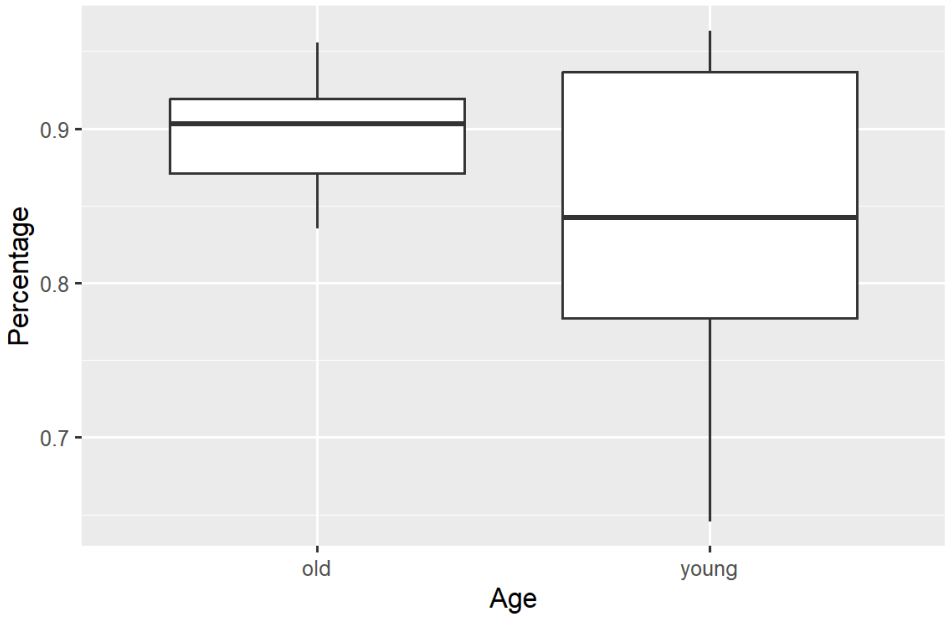


Figure 8: Percentages of syllables assigned a tone by speaker age.

An independent-sample t-test was conducted comparing older and younger speakers regarding the percentage of syllables that were assigned a tone. As the previous test on 'fully-assigned words', the t-test did not yield significant results (median = 90.3% vs. 84.4%; t(11.53) = 1.55, p = .1477; see also Figure 8). The mean percentage for older speakers is slightly higher than that of younger speakers, and, similarly to percentages of 'fully-assigned' words, older speakers have a slightly more concentrated distribution (Q_1-to-Q_3 ranges: 87.1% – 91.9% vs. 77.7% – 93.6%). Both the highest and lowest percentages come from the younger group.

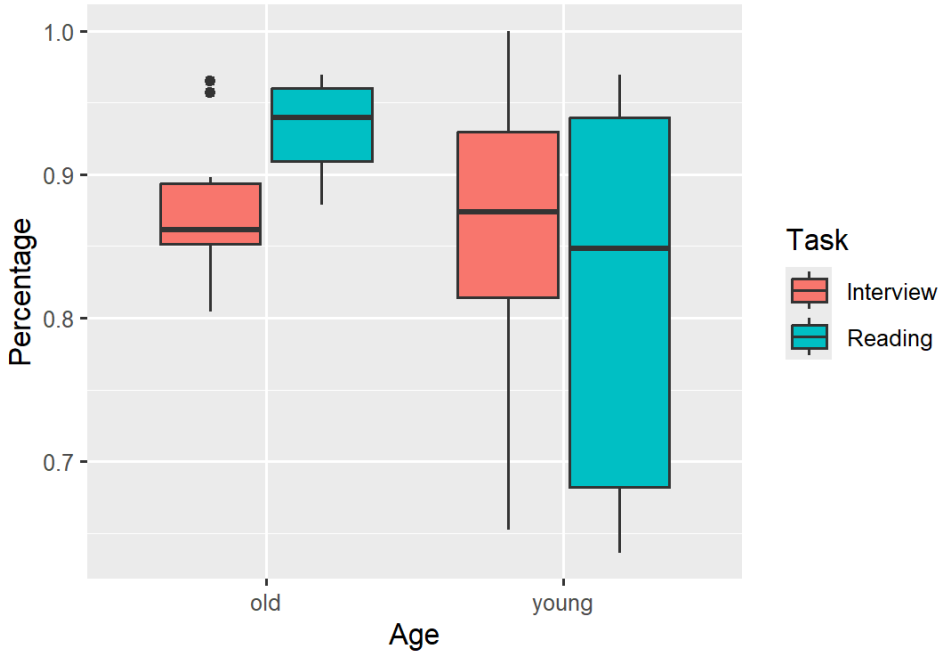


Figure 9: Percentages of syllables assigned a tone by speaker age and task.

Similarly to 'fully assigned' percentages, when divided by task, wider distributions are found in the reading tasks (Q_1-to-Q_3 ranges for old and young groups respectively: 90.9% – 96.0% and 68.2% – 93.9%) compared to the more concentrated distributions in the interviews for both age groups (Q_1-to-Q_3 ranges: 85.2% – 89.4% and 81.4% –

92.9%), with the older group being more concentrated than the younger one across both tasks.

5.3 Lexical intonation patterns

5.3.1 Overall patterns

The pitch accents assigned to the polysyllabic targets seem to be sourced from Cantonese tones and only contain the tones H(igh), M(id), and L(ow) in the target words that are examined. These tones are all level tones and tend to either present a close to plateau F0 contour within a syllable or a clear jump at the syllable boundary, neither of which can be reflected in the regression-smoothed average contours in Figure 10; however, they do reflect the overall tendencies of lexical intonation in different stress patterns. In general, IP-final contours tend to be lower than their non-final counterparts, indicative of global falling/IP-final falling pattern as a supra-lexical feature. The non-final contours also seem 'flatter', which may reflect the lack of H-tone spreading due to the presence of an intonational boundary in IP-final contours. The overall average contours of the first 3 stress pattern categories are especially worth paying attention to, which have the most tokens: 'investment' has a word-medial peak and a steeper fall towards the end of the word, indicating a MHL intonation pattern; 'manager' has a word-initial peak and a steep fall towards the end, indicating a H-tone initial syllable and a L-tone final syllable; 'chinatown' has a flat and slightly falling overall contour, especially in final positions, indicating a string of the same tone, likely H tones, superimposed with global falling or H tone downstep. The scatter plots indicate a more complicated and varied pattern, however, and close-up analyses indeed confirm the average tendencies while revealing additional patterns.



Figure 10: Time-normalised F0 values and smoothed contours (in z-scores) in IP-final and non-IP-final contexts divided by stress pattern.

		ma-	-na-	-ger
Non-final	H	99.41%	H 56.21%	H 61.54%
	M	0.59%	Ø 36.09%	L 28.40%
			L 7.69%	Ø 9.47%
				M 0.59%
Final	H	98.18%	Ø 80.00%	L 76.36%
	Ø	1.82%	L 10.91%	Ø 18.18%
			H 9.09%	H 5.45%

		in-	-vest-	-ment
	M	86.67%	H 95.56%	H 80.00%
	H	13.33%	Ø 4.44%	L 18.89%
				Ø 1.11%
	M	83.02%	H 90.57%	L 79.25%
	H	15.09%	Ø 9.43%	Ø 15.09%
	Ø	1.89%		H 5.66%

		chi-	-na-	town	
	H	100.00%	H 85.48%	H 91.94%	L% 3.23%
			Ø 14.52%	L 6.45%	
				M 1.61%	
	H	88.89%	H 66.67%	H 77.78%	L% 44.44%
	M	5.56%	Ø 33.33%	Ø 22.22%	
	Ø	5.56%			

		u-	-ni-	-ver-	-si-	-ty
Non-final	H	57.14%	H 42.86%	H 100.00%	Ø 71.43%	H 57.14%
	M	42.86%	M 42.86%		H 14.29%	L 42.86%
			Ø 14.29%		L 14.29%	
Final	M	80.95%	M 80.95%	H 95.24%	Ø 76.19%	L 80.95%
	H	14.29%	H 14.29%	Ø 4.76%	L 23.81%	H 19.05%
	Ø	4.76%	Ø 4.76%			

		sou-	-ve-	-nir	
Non-final	H	75.00%	H 75.00%	H 100.00%	L% 75.00%
	Ø	25.00%	Ø 25.00%		
Final	H	91.30%	H 91.30%	H 100.00%	L% 86.96%
	Ø	8.70%	Ø 8.70%		

		po-	-ly-	-tech-	-nic
Non-final	M	50.00%	M 45.00%	H 95.00%	H 85.00%
	H	50.00%	H 35.00%	Ø 5.00%	L 10.00%
			Ø 20.00%		Ø 5.00%
Final	M	100.00%	M 100.00%	H 100.00%	H 100.00%

		pho-	-to-	-gra-	-pher
Non-final	H	50.00%	H 100.00%	H 50.00%	H 50.00%
	M	50.00%	Ø 50.00%	L 50.00%	L 50.00%
Final	M	83.33%	H 94.44%	Ø 44.44%	L 66.67%
	H	16.67%	Ø 5.56%	L 27.78%	Ø 33.33%
				M 5.56%	

		po-	-ly-	-tech
Non-final	H	100.00%	H 100.00%	H 100.00%
Final	H	100.00%	H 100.00%	H 100.00%

Figure 11: Percentages of HKE pitch accent usage in different stress pattern types and in final and non-final positions. Individual diagrams are listed from left to right, top to bottom, in descending order by the number of tokens each has. Percentages of tones under each syllable are ranked from highest to lowest. The 'Ø' symbol represents syllables lacking a pitch accent. L% is only noted and detectable after H; as a notation it functionally marks a high falling pitch contour and does not necessarily represent a sentence-final boundary (see the discussion section).

The types of HKE pitch accents and their usage percentages for each polysyllabic stress pattern type and intonational context are listed above in Figure 11. The intonation patterns that emerge largely conform to literature predictions: 1) syllables that would be stressed in HKE-gl are assigned the H tone; 2) pre-stress syllables, if any, may be assigned the M tone; 3) post-stress syllables may be assigned the L tone; 4) when an intonational boundary is absent, H tone optionally spreads across unstressed syllables, resulting in unstressed syllables surfacing with H tones instead of M or L as they would do next to an intonational boundary. Predictions (1-3) lead to a MHL intonation pattern, while H-spreading lead to a series of adjacent H tones within a word, sometimes spanning all syllables.

As previously discussed, tokens are only divided by their adjacency to a right boundary since multisyllabic tokens adjacent to a left boundary are rare in this corpus. Nevertheless, each IP-initial token has been examined individually, and they largely conform to predictions with regards to the conditions of H-spreading, and no indications of initial boundary tones are found. There are 5 IP-initial tokens belonging to the 'manager' stress pattern category, where the stress falls on the first syllable; all of these tokens begin with H. 3 IP-initial tokens in the 'investment' category all begin with M as predicted for unstressed syllables following a boundary. Interestingly, 4 'souvenir' tokens and 1 'photography' token that follow an intonational boundary begin with H tone despite having unstressed initial syllables.

The non-IP-final tokens generally tend to feature H-spreading after the stressed syllable, whereas IP-final tokens do not. Pre-stress syllable(s) are assigned M in most cases, and often do so following an intonational boundary. In non-final contexts, the most common pattern is an all-H contour, followed by an MH contour. In IP-internal

contexts, pre-stress syllables may optionally be assigned H instead of M, and post-stress syllables may optionally be assigned H instead of L. For ‘souvenir’, which has a final stress, and for ‘chinatown’, which includes target words *chinatown* and *british-made* and are compounds consisting of a stress-initial disyllabic word followed by a monosyllabic word, the final syllable often has a high falling contour, annotated as a H tone combined with a L% boundary tone. The L% occurs more frequently in ‘souvenir’ (75% and 86.96% for non-IP-final and IP-final intonational contexts respectively) than in ‘chinatown’ (3.23% and 44.44% for either intonational contexts respectively), and more frequently in final positions than in non-final ones.

In addition, some syllables are more frequently in lack of a pitch accent than others. Stressed syllables are almost all assigned an H tone, the percentage ranging from 88.89% in IP-final ‘chinatown’ tokens to 100% in several categories. Percentages of H tone assignment of stressed syllables in categories with the largest number of tokens (‘manager’, ‘investment’, and ‘chinatown’) average above 90%. On the other hand, there is a tendency for post-stress non-final syllables to not carry a pitch accent. This is the case in most otherwise strongly HKE-lo tokens in ‘university’ (71.43% and 76.19% in either intonational contexts) and in IP-final ‘manager’ tokens (80%), for example, where the penultimate syllable is most frequently not assigned a pitch accent. Occasional non-pitch accent syllables otherwise occur in lower frequencies mostly under 20%, often as infrequent as having only one token in the category.

5.3.2 *Comparing age groups and tasks*

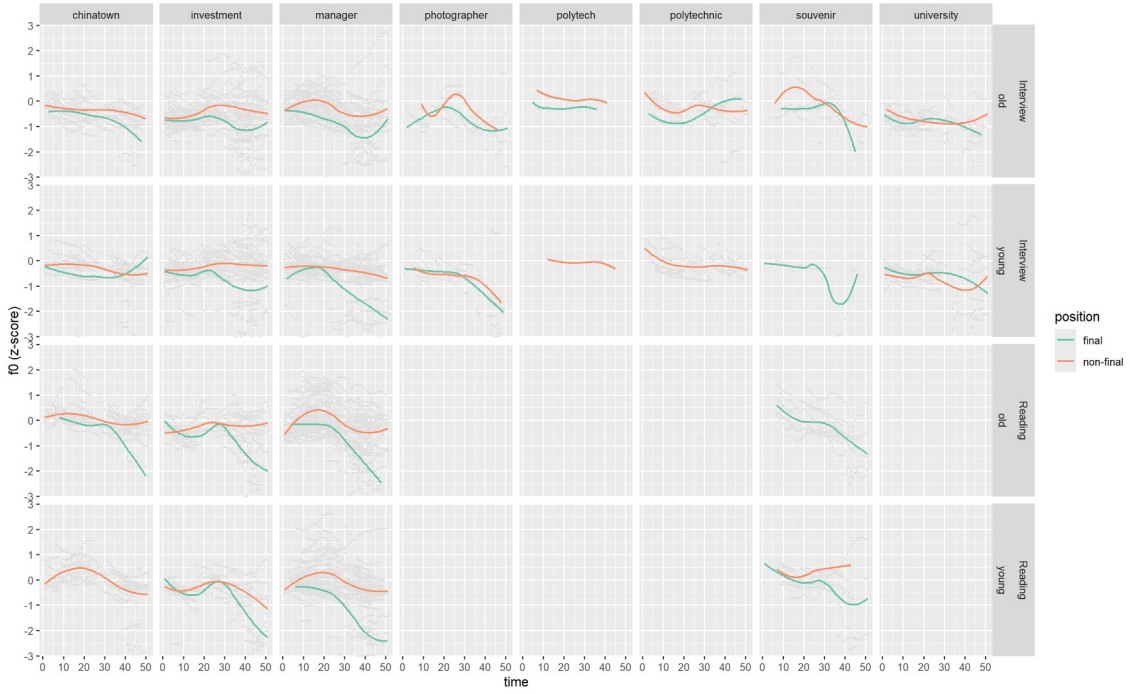


Figure 12: Time-normalised F0 values and smoothed contours (in z-scores) in IP-final and non-IP-final contexts divided by stress pattern, age group, and task.

The two age groups have largely similar overall contours that support the observations in the previous section: 1) lower overall values in IP-final positions than in non-final positions, 2) flatter contours for non-final tokens due to lack of H-spreading, 3) 'investment' tokens have a MHL average contour, 4) 'manager' tokens tend to peak word-initially and then dip towards the end of the word, and 5) 'chinatown' tokens have an overall flat contour until the end of the word. Due to small numbers of tokens divided into different categories, the contours of reflect the individual variations with regards to optional phenomena in HKE such as pitch accent assignment, H-spreading, L% boundary tone, and final rising, which will be discussed in the following sections.

Looking at ‘chinatown’, ‘investment’, and ‘manager’, which have the largest numbers of observations, there seems to be a small difference between tasks and between age groups in interviews. Contours from interviews seem to have less dramatic slopes, especially in the younger group in non-IP-final positions. This may indicate two separate phenomena: 1) H-tone spreading is slightly more common in connected semi-spontaneous speech than in read speech and in younger speakers during interviews; 2) interpolation of pitch is slightly more common as opposed to assigning pitch accents to each syllable in connected semi-spontaneous speech than in read speech and in younger speakers during interviews. Separating out the HKE-lo tokens by looking at pitch accent assignment helps tease out the difference.

5.3.3 Comparing ‘fully-assigned’ tokens

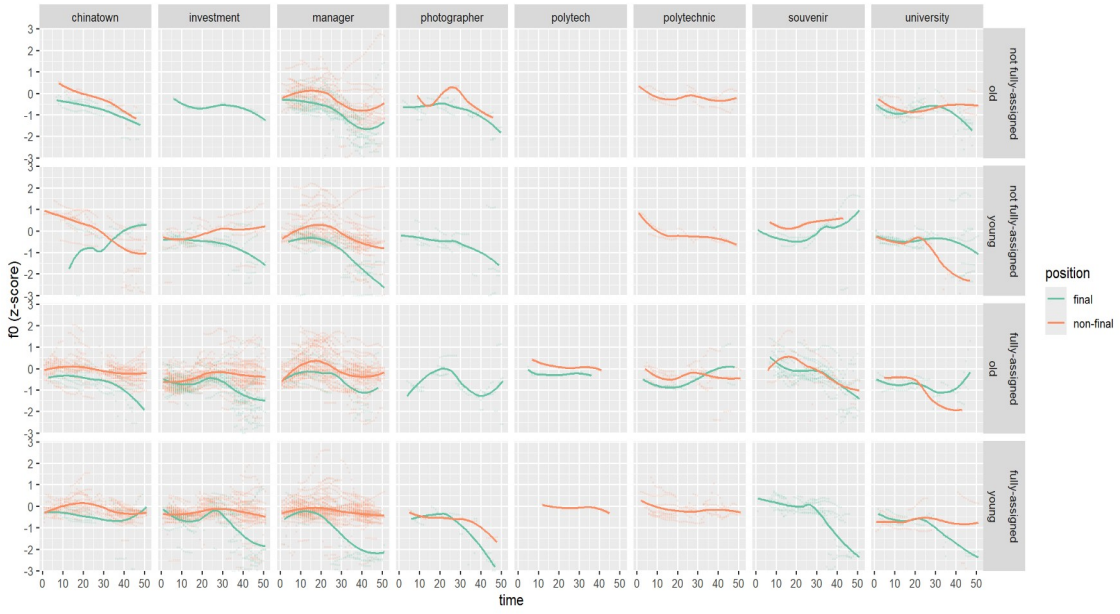


Figure 13: Time-normalised F0 values and smoothed contours (in z-scores) in IP-final and non-IP-final contexts divided by stress pattern, age, and whether the token is 'fully-assigned'.

Using pitch accent assignment as an indicator of whether a token as a stronger HKE-lo intonation pattern, it can be observed that, compared with those that do not have every

syllable assigned a pitch accent, tokens that do tend to also have F0 contours that follow predictions of HKE-lo intonation pattern. For example, ‘chinatown’ tokens have significantly flatter contours in both age demographics when ‘fully-assigned’ than when they are not; in the latter case, the contours show greater tendencies to fall towards the end of the word except for in IP-final positions in younger speakers, when the contour shows a rising tendency. It is difficult to tell the ‘manager’ contours apart, but the ‘fully-assigned’ IP-final tokens show an obvious peak word-medially, while their ‘non-fully-assigned’ counterparts do not.

The contours of non-IP-final tokens tend to have flatter shapes than tokens in IP-final positions in ‘chinatown’, ‘investment’, and ‘manager’ when they are ‘fully-assigned’, but not so much so when they are not ‘fully-assigned’, indicating that H-tone spreading may contribute to the flatter contours in non-final contexts. The effect is especially obvious in ‘investment’ and ‘manager’ in younger speakers, where an F0 peak is barely observable in non-final positions, while for older speakers a small F0 peak can still be identified.

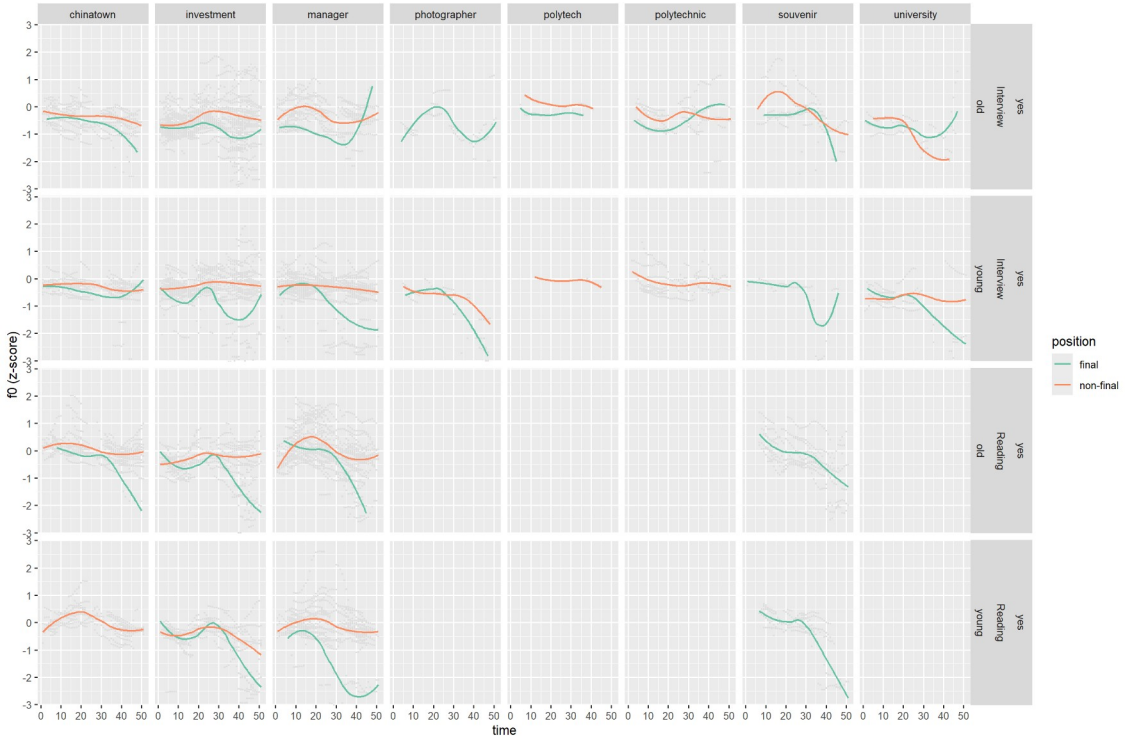


Figure 14: Time-normalised F0 values and smoothed contours (in z-scores) in IP-final and non-IP-final contexts divided by stress pattern, age, and task, filtered for only the 'fully-assigned' data.

When comparing between the tasks considering only the ‘fully-assigned’ tokens, the effect is stronger in interviews than in reading tasks, and especially stronger in interviews in younger speakers. The above figure shows that while for ‘investment’ and ‘manager’, small F0 peaks can still be identified in reading tasks for both age groups and for interviews with older speakers. The non-IP-final contours in interviews with younger speakers, however, are almost completely level; because only ‘fully-assigned’ tokens are considered, this indicates a strong presence of the HKE-lo feature H-tone spreading as opposed to HKE-gl intonational features.

5.4 Other intonational phenomena

5.4.1 L tone creak

L-toned syllables consistently feature creaky voice. Out of the 545 tokens extracted, 221 feature at least one syllable assigned with an L tone. 158 of these 221 tokens (71.49%) feature creaky voice during the lower toned syllable(s). In contrast, out of the 324 tokens not containing an L tone, only 26 (8.02%) feature creaky voice; in 8 of them, the creaky voice takes place on the M-toned syllables, and in others, creaky voice happens towards the end of the word where F0 drops. However, in most tokens not containing the L tone, no creaky voice is present despite featuring slight F0 falling.

Consider the *consider* example again, which contains a typical case of L tone creak. The entire L-toned syllable has a strong creaky quality, characterised by irregular glottal pulses and a low vibration below 100 Hz. The descend into creaky voice is abrupt, with a clear boundary between modal phonation for M- and H-toned syllables and creaky phonation for the L-toned syllable at the end.

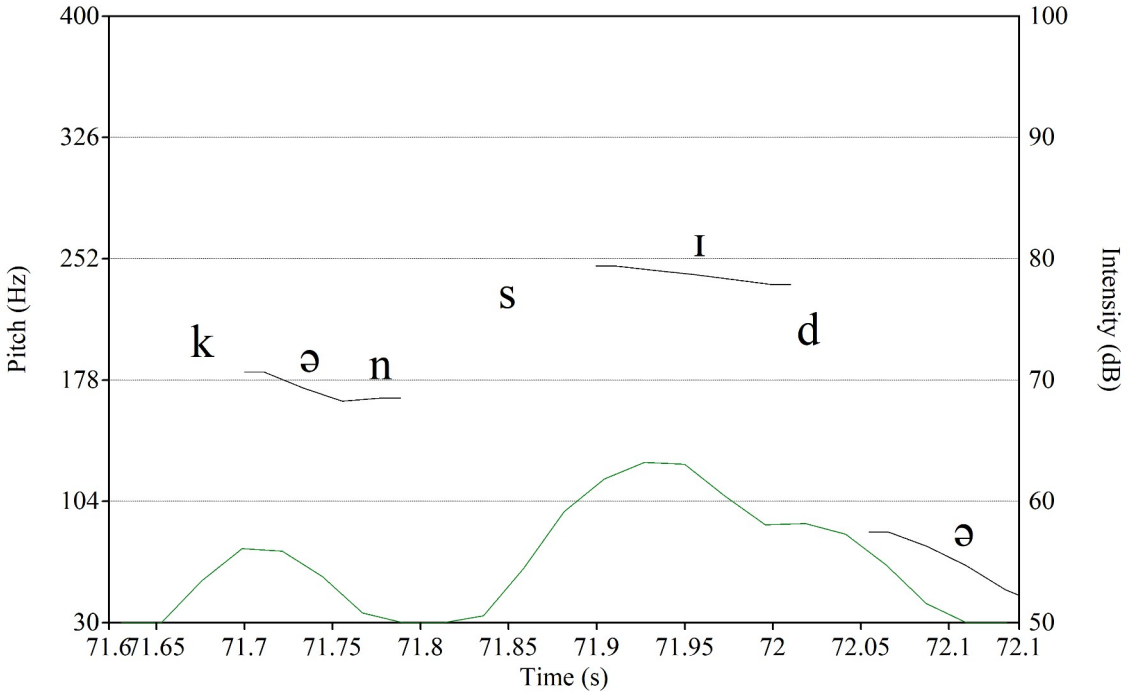


Figure 15: F0 and intensity of consider *for speaker HK14.*

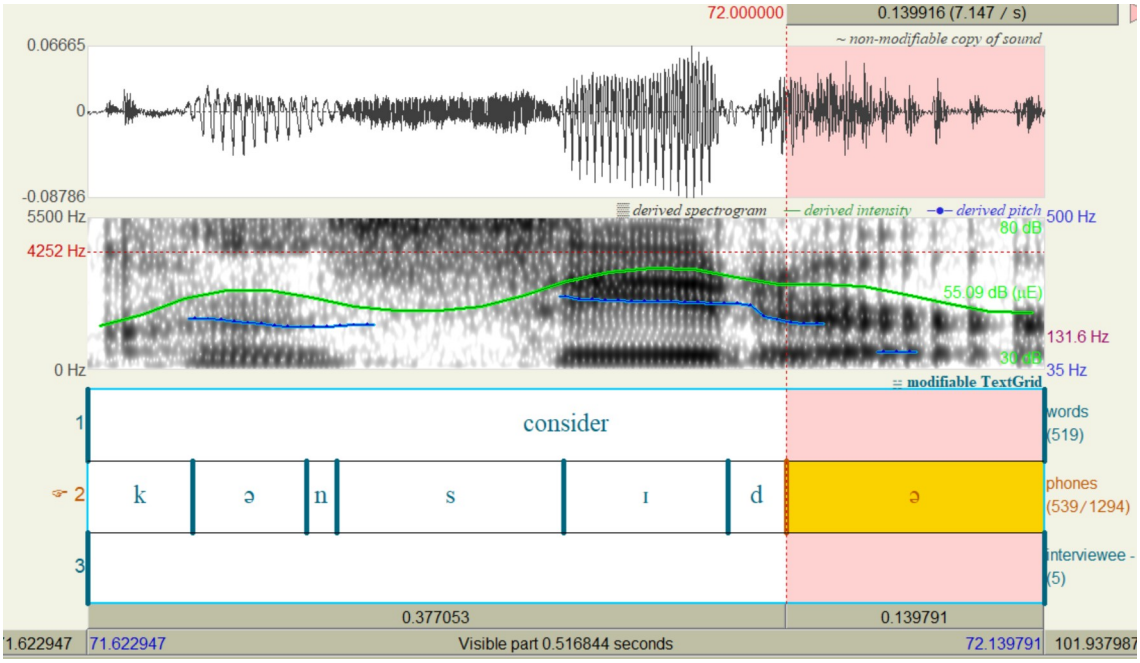


Figure 16: Wave form and spectrogram of the same consider token for speaker HK14, with the L-toned syllable highlighted.

5.4.2 H tone downstep

In some cases of H-tone spreading, the following H-toned syllables after the first one may slightly step down in terms of pitch. Out of the 545 tokens extracted, 291 feature some form of H-spreading. 77 of these tokens seem to feature downstepping between the H-toned syllables. Compare the two instances of *microsoft* by two younger speakers during a reading task. Both are IP-internal and feature a full H-tone spread from the first syllable to the other two syllables in the word, resulting in a flat contour. However, in speaker HK26's example, a slight but audible downstep happens in the final syllable compared to the previous two; the first and second syllables plateau at around 220 Hz, while the final syllable plateaus at around 180 Hz. In comparison, speaker HK55's *microsoft* does not feature H downstep. The final syllable plateaus at around 180 Hz as the initial syllable, and the three syllables in the word perceptually have the exact same tone.

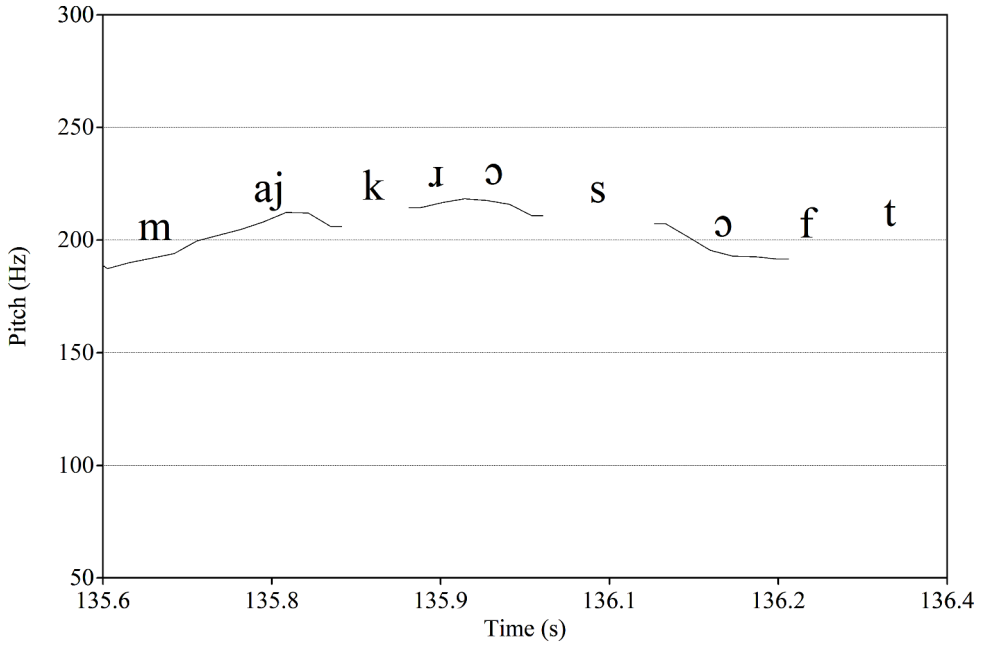


Figure 17: F0 contour of microsoft *for speaker HK26, featuring H-spreading and downstep.*

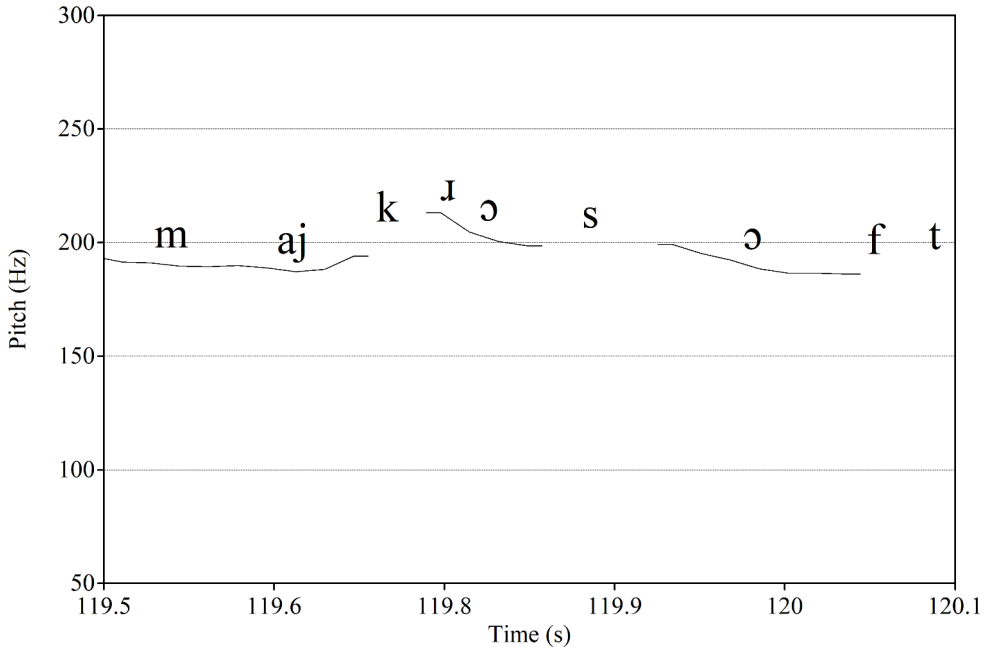


Figure 18: F0 contour of microsoft *for speaker HK55, featuring H-spreading without downstep.*

Z-score normalised average F0 contours show that there is a detectable difference between tokens that do not have this downstep and the ones that do. The two contours have largely overlapping trajectories until they begin to diverge from the 80% timing mark, with the downstepped contour falling to a slightly lower value than where it begins with, but not so low that it can be analysed as an M or L tone. The peak roughly corresponding to the second syllable in the two contours reflect a higher F0 onset on the second syllable, which can be observed in both examples of *microsoft* above. This is likely related to the aspiration in the consonant cluster /kɹ/. In several instances of IP-internal *microsoft* and *british-made* featuring H-spreading, the F0 perturbation in the second syllable becomes minimally audible, resulting in, paradoxically, the unstressed second syllable being perceptively the highest pitched syllable in the word.

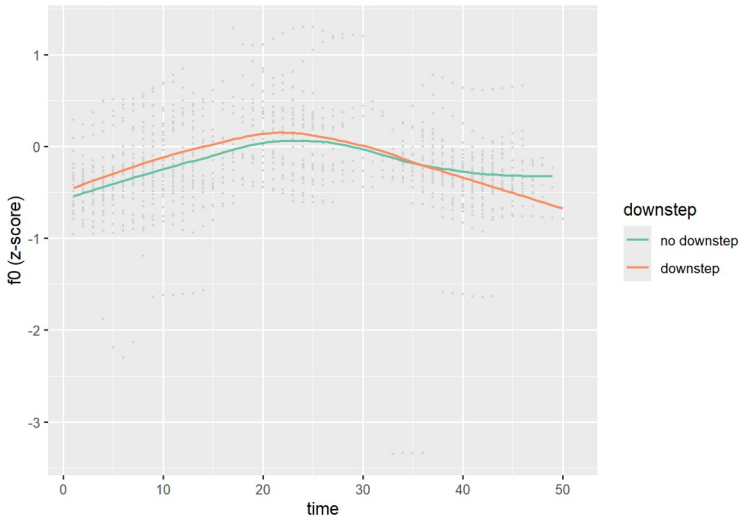


Figure 19: Time-normalised F0 values and smoothed contours (in z-score) for instances of microsoft *that feature complete H-spreading, divided by the presence of final downstep.*

5.4.3 Word-final rise

The average F0 contours in the previous sections sometimes feature a rising trend near the end of the word, and it may be more pronounced when the data is divided into smaller subgroups. This is due to a number of tokens across stress pattern types, intonational contexts, and pitch assignment conditions that have a rising contour towards the end of the word, influencing the average outcomes of the contours and more strongly so when the total number of tokens in the sub group is relatively low. However, out of the 545 tokens extracted from the corpus, only 59 (10.83%) have a rising pitch contour towards the end of the final syllable. This explains why the final rise is difficult to detect in, for example, 'manager' and 'investment' type tokens in the overall average contour in Figure 10 where there are a large number of observations being considered together.

Final rising happens across different conditions including stress pattern types, intonational contexts, pitch assignment conditions, speaker age, and task. Figure 20 below shows that there are tokens with the final rise in all stress pattern types except for 'polytech', which only contains the word *polytech*, has 3 tokens in total, and is a truncated synonym of *polytechnic*. The 'manager' category has the most instances of final rise (24), followed by 18 in 'investment', and then 7 in 'chinatown', which is consistent with the ranking of total numbers of tokens in the categories.

	'manager'	'investment'	'chinatown'	'university'	'souvenir'	'polytechnic'	'photographer'	'polytech'	Total
Final rise	24	18	7	1	4	1	4	0	59
No final rise	198	125	73	27	23	21	16	3	486
Total	222	143	80	28	27	22	20	3	545

Figure 20: Final rising count by stress pattern type.

Out of the 59 tokens with final rising, 36 (61.01%) are in IP-initial or -medial positions, and 43 (72.88%) are 'fully-assigned' tokens. Final rising happens in reading and interview tasks equally frequently. Out of the 59 final rising tokens, 35 come from interviews, and 24 come from read speech. Both older and younger speakers exhibit this feature, with 38 tokens coming from 9 speakers in the older age group and 21 tokens from 8 younger speakers out of 20 speakers in total. Interestingly, although the number of speakers that share this feature are evenly distributed between the two age groups, most individual tokens belong to two older speakers. It seems that although most speakers share this feature, two older speakers use it notably more frequently than others.

Younger speakers:

	HK14	**HK21**	**HK22**	**HK30**	**HK44**	**HK49**	**HK5**	**HK53**
Number of final rise tokens	3	2	3	4	3	1	1	4

Older speakers:

	oHK1	**oHK13**	**oHK16**	**oHK17**	**oHK18**	**oHK19**	**oHK25**	**oHK3**	**oHK4**
Number of final rise tokens	**14**	1	1	3	1	**10**	1	2	5

Figure 21: Final rising count by individual speaker.

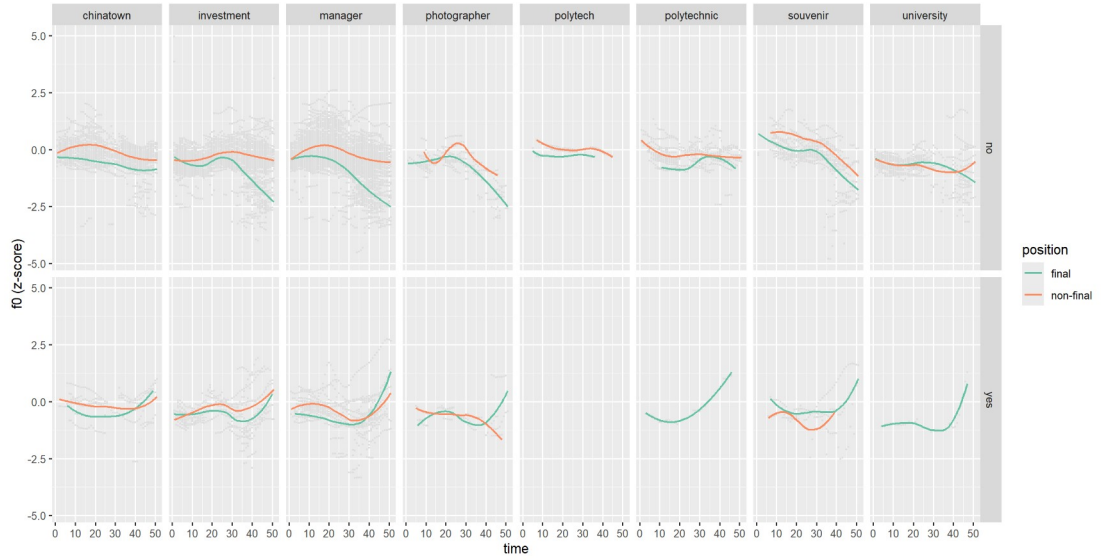


Figure 22: Time-normalised F0 values and smoothed contours (in z-scores) in IP-final and non-IP-final contexts divided by stress pattern and whether or not the token has a final rise.

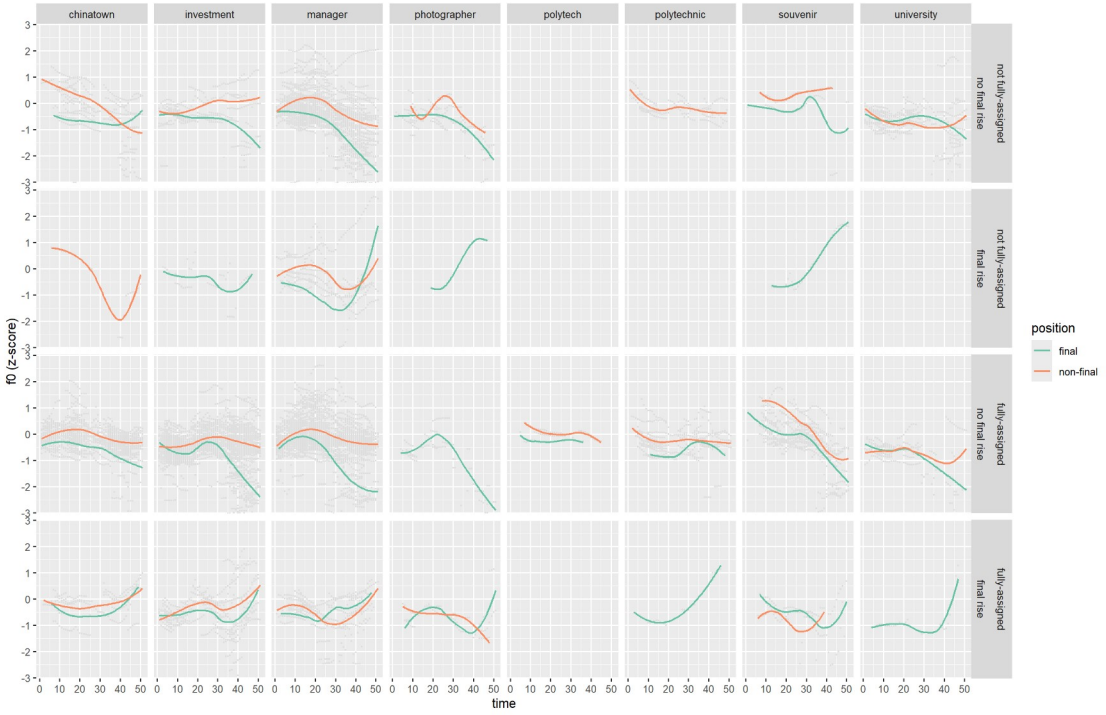


Figure 23: Time-normalised F0 values and smoothed contours (in z-scores) in IP-final and non-IP-final contexts divided by whether the token is 'fully-assigned' and whether is has a final rise.

The figures above show what the F0 contours of final rising look like overall and when divided for pitch accent assignment conditions, compared to the tokens without rising. Most instances of rising is contained in the final syllable, barring a few exceptions such as the two instances in ‘non-fully-assigned’ ‘photographer’ and ‘souvenir’ in Figure 23 that exhibit a rising trend since earlier in the word. Notably, in words with stress patterns that typically end with a L tone or otherwise falling pitch such as ‘manager’ and ‘investment’ when H-spreading does not happen, the rising often only happens after the fall. Below is an example of final rising after a L-tone in the word *investment*, produced by an older speaker in a reading task. The contour of this token otherwise shows a typical MHL intonation pattern: a mid-tone initial syllable plateauing at around 180 Hz followed by a clear jump to a higher range at the boundary before the next syllable where the contour rises to around 250 Hz, after which the contour jumps to the low end (160 Hz) and then dips towards 110 Hz. Just before it reaches clear creaky voice pitch range at the 100 Hz mark, however, the contour rises and reaches as high as the pitch of a H tone.

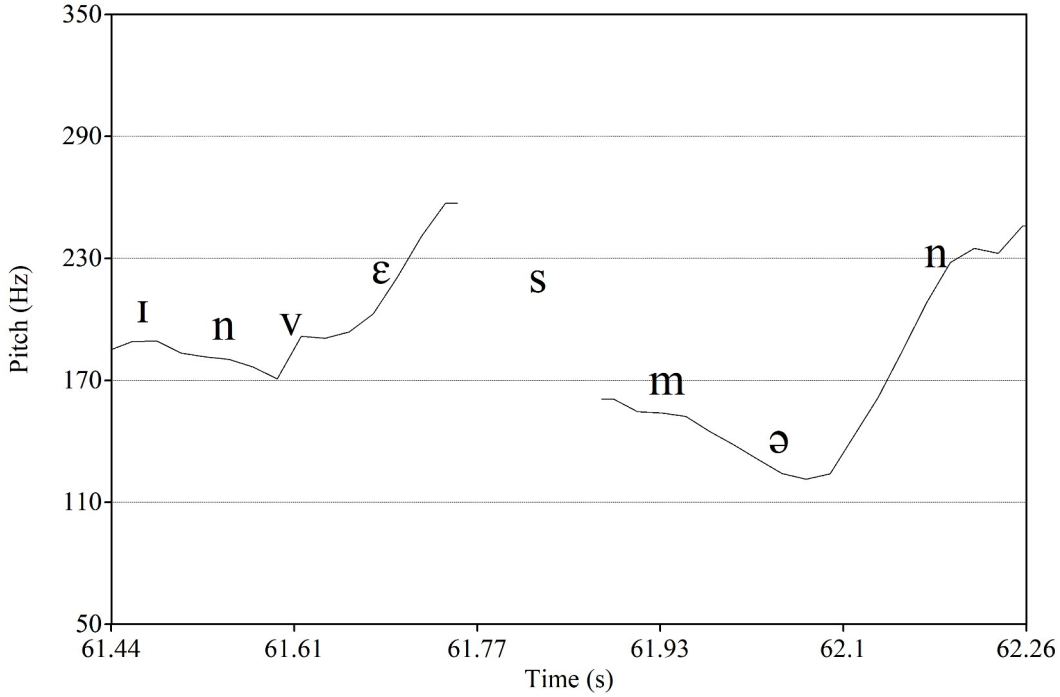


Figure 24: F0 contour of investment *for speaker oHK4, featuring word-final rising after a L tone.*

6 Discussion

6.1 HKE intonation: The expected and the unexpected

The speakers involved in this study showed reasonably strong HKE-lo intonational features. The intonation patterns largely conform to existing descriptions of HKE intonation in the literature, with some divergence from past observations, still resulting in an overall impression of HKE-lo with some HKE-gl influences, which are stronger in some speakers than in others. In this section I discuss the expected and unexpected intonational features found in the data and how they add to or challenge existing literature on HKE intonation.

6.1.1 Phonology: Pitch accents and boundary tones

The majority of tokens examined in this study have every syllable assigned a pitch accent, which confirms previous observations about HKE such as in Gussenhoven (2012) and Lam (2020). Instances of typical HKE-lo intonation patterns are found in all speakers, and the predictions in Figure 1 have largely been confirmed once H-spreading and boundary tones are accounted for. The first 4 rules in Lam's (2020) description of HKE-lo intonation plus the L% boundary rule have been confirmed in the data:

1. Stressed syllables are given the H tone.
2. Unstressed word-initial syllables are assigned the M tone.
3. Unstressed word-final syllables are assigned the L tone.
4. H can be optionally spread rightward to the adjacent syllable(s), replacing the word-final L tone(s).
5. For sentence-final intonation, assign either the prolonged boundary tone ":" or L tone to declaratives.

Lam also mentions the existence of Cantonese contour tones in HKE when Cantonese loan words or sentence-final particles (SFP) are present, but since the polysyllabic tokens analysed in this study all come from either declarative sentences in reading tasks or answers to interview questions, and none of them involve Cantonese loan words or SFPs, only the 'core HKE tones' i.e. H, M, and L are found in the HKE data.

6.1.1.1 H-spreading and secondary stress

Descriptions of HKE intonation in Luke (2000), Gussenhoven (2012) and Lam (2020) assume a rightward direction of H-tone spreading and focus on intonation

patterns of individual words, and they do not focus on the analysis of secondary stress. (Gussenhoven 2012) assumes that both primary and pre-primary secondary stresses should be equally assigned H-tones, and that is why HKE intonation for *conversation* is [HHHL]. This is a convincing analysis, because a closer examination of truly unstressed word-initial syllables in the data reveals that the few apparent H-initial instances of *investment* and *consider* can either be argued to be edge cases between H and M phonetically, or are very short in duration and follow an H-toned syllable, such as in *Central FuTai investment group*, where the apparent H tone can be analysed as interpolation instead of a pitch accent. Non-edge-case instances, on the other hand, do occur in *polytechnic*, *polytech*, and *university*, where the first syllable does have secondary stress. If we assume that rightward H-spreading only stops at the word boundary as Lam's formulation above implies, H-spreading phenomena in the 'university' tokens in the data can be accounted for by assigning the first syllable, which has secondary stress, an H tone which then spreads rightward.

However, out of the 28 instances of *university*, only 7 of them feature a span of H tones from the beginning of the word until the L tone(s) in the end, and the tokens belong to different speakers. The majority of instances in fact feature M-tones in the first two syllables, treating them as unstressed. This cannot be explained solely using the argument that the word itself is represented differently in individual speakers' mental lexicons, because out of the 7 speakers who produced H-initial tokens, 2 of them produced a different instance of *university* where the initial syllable is M-toned. This shows that pre-primary secondary stresses in HKE can be and indeed often are left out in H assignment. In addition, looking at sentential contexts reveals that when *university* begins with H, the preceding tone(s) is typically also H, as is exemplified below:

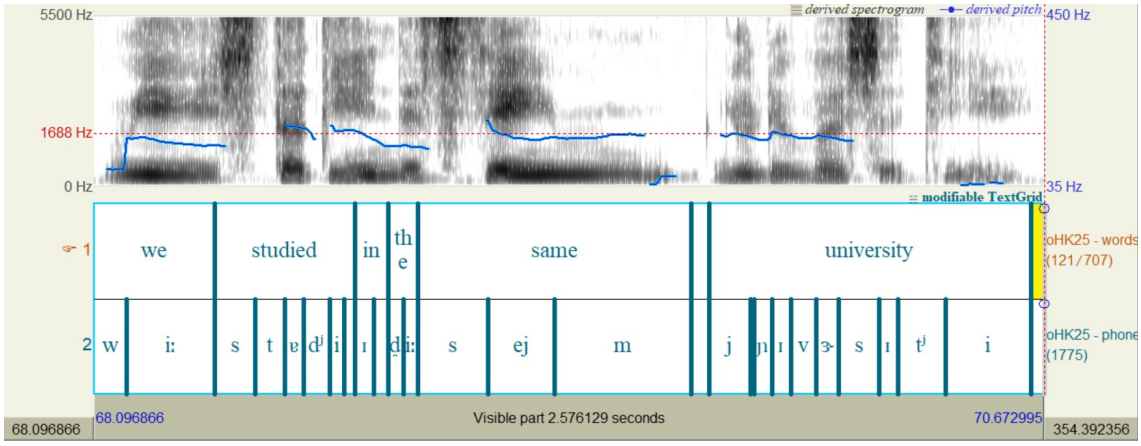


Figure 25: Spectrogram and segmental information for speaker oHK25's utterance we studied in the same university *during her interview. F0 contours are marked in blue.*

The pitch accent assignment of the sentence shown in Figure 25 is as the following:

we	stu-	-died	in	the	same	u-	-ni-	-ver-	-si-	-ty
|	|	|	|	|	|	|	|	|	|	|
M	H	H	M	M	H	H	H	H	Ø	L

A span of H-tones go from *same* to *-ver-*. The apparent low pitch at the end of the nasal coda in *same*, which interrupts the H-spanning, reflects the creaky voice quality typical of this older speaker near pauses and glottal stops and seems to be a physiological feature. The initial syllable in *university* undoubtedly has a H pitch accent, and it is thus reasonable to analyse *same univer-* as a single H-tone-spreading domain, crossing the boundary between the two words, where the H-tone in *same* spreads rightward until before *-si-*, which is either an elided L in connected speech due to hypo-articulation or not assigned a pitch accent due to influence from HKE-gl.

A similar instance of H-spreading crossing word boundaries can be found in Cheung 2009: 75:

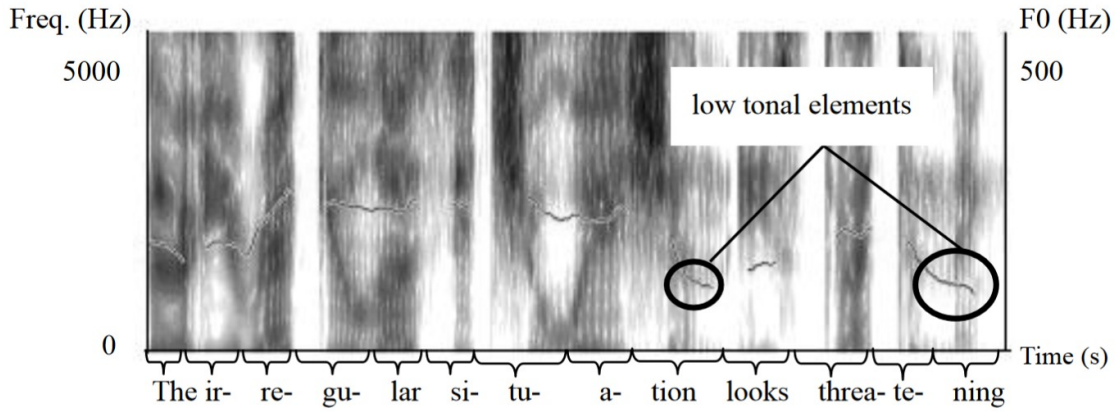


Figure 26: Spectrogram and pitch contours of the irregular situation looks threatening.

The H-tone spans from the second syllable of *irregular* to the penultimate syllable in *situation*. In Cheung's autosegmental analysis, a single H-tone would be associated with all of the syllables in its domain of H-spanning marked by two stressed syllables. There is no direction of spreading in this analysis because Cheung came from the perspective of Optimality Theory (OT). The data in this study does not provide immediate evidence whether an OT approach or an expanded derivational approach that allows rightward spreading across word boundaries, but either way, the data suggests that 1) secondary stresses in HKE are not treated the same way as primary stresses with regards to intonation, i.e., the latter is always assigned H in HKE-lo while it is not always the case for the former, and 2) H-tone spanning in HKE happens across word boundaries, and the boundary crossing seems to be more common when the immediate post-boundary tone is or can optionally be assigned H.

6.1.1.2 L% boundary tone and secondary stress

Another observation of the behaviour of secondary stresses in HKE relates to the L% boundary tone. Gussenhoven (2012), Lam (2020), and others have observed that L% in HKE has an emphatic quality and, when attached to an IP-final stressed syllable, which leads to a high falling pitch contour. In the current study, the HL contour only

occurs word-finally and is thus annotated as H+L% in Figure 11, which assumes the analysis that this contour indicates the existence of a boundary. What adds to previous literature is the fact that L% occurs not only sentence/IP-finally, but also when the word seems to be otherwise IP-internal. It is worth repeating that my methods only identify the intonational boundaries that largely coincide with clause/sentence boundaries and are separated by pauses.

Cheung analyses the H+L sequence at the end of a H-spanning domain as a marker of Phonological Phrase, which accounts for the behaviour of L% in the current study better than Lam's and Gussenhoven's descriptions. This analysis is autosegmental and does not distinguish between H+L sequences in words like *university*, where the final syllable itself is assigned a L-tone, and H+L sequences at the end of words like *souvenir*, annotated as [HHHL%] in this study, where L is not a pitch accent of its own and is analysed as sharing a single timing slot as the preceding H. What the data in the current study shows is that even in non-IP-final positions, 75% of 'souvenir' tokens and 3.23% of 'chinatown' tokens end with a H+L sequence (in IP-final positions the percentages increase to 44% and 86% respectively). It seems that compounds, which have the main stress on the first component due to English compound stress, still behave similarly to non-compounds with final main stress. The 'souvenir' category only consists of the word *souvenir*, which in the HKE speakers' variety likely has a final primary stress and an initial secondary stress. The 'chinatown' category consists of the compounds *chinatown* and *british-made*, and the former is more established as a compound than the other and thus more often take initial primary stress. However, all of the tokens with final H+L% are *chinatown*. This is likely due to the fact that all instances of *british-made* are produced in the context *a british-made bin* [M HHH HL%], where H spans from the first syllable of *british* to *bin*. Nevertheless, the

behaviour of L% in word-final secondary and primary stresses reveals that 1) L% in HKE marks more than just the end of declarative sentences, but also the boundary of a smaller phonological unit; 2) word-final secondary stresses in compounds in HKE can be assigned the H tone like primary stresses do.

6.1.1.3 The mystery of word-final rise

The phenomena of word-final rising, which I tentatively transcribe as H%, has a wide distribution across individual speakers and all levels of different variables, but is overall infrequent. Therefore, it is understandable that none of the HKE literature I reviewed for this thesis mentions such a phenomena. However, word-final rises occur enough to affect the analysis of tone assignment, and is anecdotally a noticeable marker of HKE from my personal discussion with another linguist who used to live in Hong Kong. It is mostly a locally-contained rise like L%, which only applies to one syllable (a few instances of rising in the data feature a rising contour across syllables resulting in a stronger HKE-gl impression, and may be in fact separate HKE-gl rising phenomena), but compared with L%, which only occurs noticeably after H tones and would not cause a lower fall after L tones at the end of declaratives (Gussenhoven 2012), the final rise causes a rising contour for the H-toned syllable, and for the L-toned syllable, it leads to a small rise after the pitch dips down to reach or almost reach the L target.

Since final rising happens in both IP-final positions and otherwise, it is not sensitive to the same kind of intonational boundary. In fact, it is noticeably absent where H-spreading extends into the following word and indeed would sound unnatural otherwise (e.g. *investment group* [MHH HL%; *MHHH% HL%]). It can and often does combine with H+L% sequences, resulting in a dipping contour with a small rise at the end. It may also attach directly after a H-tone span followed by a non-H (e.g. *I went to*

chinatown for a foot massage, where the underlined section has a [HHHH% M] intonation pattern). It is clear that the final rise is different from the interrogative H% boundary tone marking the end of a question that Luke, Gussenhoven, and Lam observed. I tentatively propose that the final rise, like the L% boundary tone in Cheung 2009, is an optional marker of the boundary of a smaller intonational unit. It may be an additional marker attached to the end of L% or serve as the end-of-H-span marker on its own.

The discourse function that the word-final rise serves is even less clear. One hypothesis is that it functions as a continuation rise, signalling that although the H-spanning is over, the speaker has potentially more to say. Studies based on Inner Circle Englishes have investigates different types of rises other than interrogatives (e.g. see Cruttenden 1995 on rises in dialects of British English, Australian and New Zealand English, and North American English; see Warren 2016 for an investigation into 'uptalk', and Brazil, Coulthard & Johns 1980 for a discourse analysis approach). Research on rises in HKE, however, largely limits itself to question intonation (e.g. Gussenhoven 2012 and Chen & Mok 2015). Lam (2020) did find the existence of MH boundary tones, but they are of different origin and usage from the final rises in this study. Apart from the obvious difference in pitch contour from word-final rise, which can be attached to either L or H tones, Lam's MH boundary tone is sourced directly from Cantonese, associated with sentential boundaries, and assigned to the Cantonese SFP *ge* in HKE, which expresses 'tentative or uncertain affirmation' (Matthews & Yip 2011: 391). More research therefore needs to be done on the non-interrogative rises in HKE-lo to further examine its phonological distribution and its function in the discourse.

6.1.2 Phonetics: Voice quality and coarticulation

6.1.2.1 L-tone creak: Influence from a different Cantonese tone?

Studies in the HKE literature often assume that the H, M, and L pitch accents in HKE-lo are sourced from the Cantonese tones High Level, Mid Level, and Low Level respectively. One of the first linguists who proposed and formalised this relationship was Luke (2000), who did not provide a detailed phonetic motivation. The later research that adopted this assumption, (e.g. Lam 2020, Lam 2017, Gussenhoven 2012, Gussenhoven 2014) also did not focus on the phonetic connections between the HKE pitch accents and those specific Cantonese tones. Apart from the three level tones from which HKE is said to have borrowed for its pitch accents, Hong Kong Cantonese also has three contour tones, i.e. High Rising, Low Falling, and Low Rising. Lam (2020) proposed that these tones exist in HKE as 'auxiliary tones' which occur in Cantonese loan words and SFPs but failed to record instances of Low Falling, though it is theoretically possible.

The L-tones in HKE in this study, however, happen to be phonetically closer to Low Falling than to Low Level. The first similarity is creakiness: 71.49% of all tokens containing an L-tone have creaky voice. The ones that do not feature creaky voice sometimes feature a final rise that brings the pitch upward before it can reach a low enough target to produce creakiness. The consistent and widespread presence of creaky voice in L-tones indicate that creak is a key feature of the L-tone's production and likely to be a phonation cue to its perception. In Cantonese, creaky voice serves a similar role in tone production and perception of the Low Falling tone: Yu & Lam (2014) found that creaky voice occurs consistently and systematically in the Low Falling tone in Cantonese, and that listeners use this as an important cue to the distinction between Low

Falling and Low Level, which does not feature consistent creaky voice. The second similarity is the contour: Sometimes the L-toned syllable has a level pitch track like in Figure 7, but more often they feature a falling contour like in Figure 6. Both of these contours appear in Low Falling in Cantonese (Matthews & Yip 2011), and when the pitch track has a level shape, its creakiness and low pitch become the main cues for perception.

The diachronic development of the HKE L-tone is unclear, and there is still a possibility that the Low Level was indeed its original source and then changes took place which assimilated it to a different tone, but it is unambiguous that the modern HKE L-tone is phonetically and perceptually similar to the Cantonese Low Falling, instead of the Low Level. It would not be surprising if the L-tone in HKE shifted from non-creaky to creaky over time, since Hong Kong Cantonese Low Level is less stable than Low Falling and has been undergoing mergers with the Low Falling and Mid Level simultaneously (Mok, Zuo & Wong 2013). Since the HKE M-tone is sourced from Mid Level and is only distinguished from the Low Level by F0 values, creaky voice may have become an additional cue for distinction between L and M as its function for Low Falling and Low Level in Cantonese. It would also be interesting to compare the same speakers' Cantonese tones with the L-tones in their English. In the original project that collected the data used in this study, speakers participated in reading tasks and interviews in both English and Cantonese, and it is thus a viable follow-up study for the future.

6.1.2.2 H-spanning: Internal downstep and post-aspiration 'upstep'

H-tone sequences as the result of H-spanning in HKE are assumed to be of roughly the same pitch level in previous literature. H tones may feature downstepping when

interrupted by a non-H tone, which has been noted by Gussenhoven (2012), but no downstepping within a span of H tones has been mentioned. There was also no mention of other phonetic fluctuations of the span of H-tones. The present study, however, found that in trisyllabic words assigned three H tones, regardless whether the main stress falls word-initally like in *microsoft,* or word-finally as in *british-made,* it is not uncommon to find perceptually and auditorily detectable downsteps in the final syllable and 'upsteps' after the aspirated consonant cluster in the second syllable. An example of downstepping can be found in Figure 17, while the post-aspiration uptick can be seen in Figure 18. In fact, Figure 27 below shows that these two phenomena can happen at the same time: The three syllables in *british-made* here are of roughly the same pitch level and are all part of a H-tone span, but the second syllable has a higher plateau than the others, while the final syllable jumps to a lower pitch at the syllable onset.

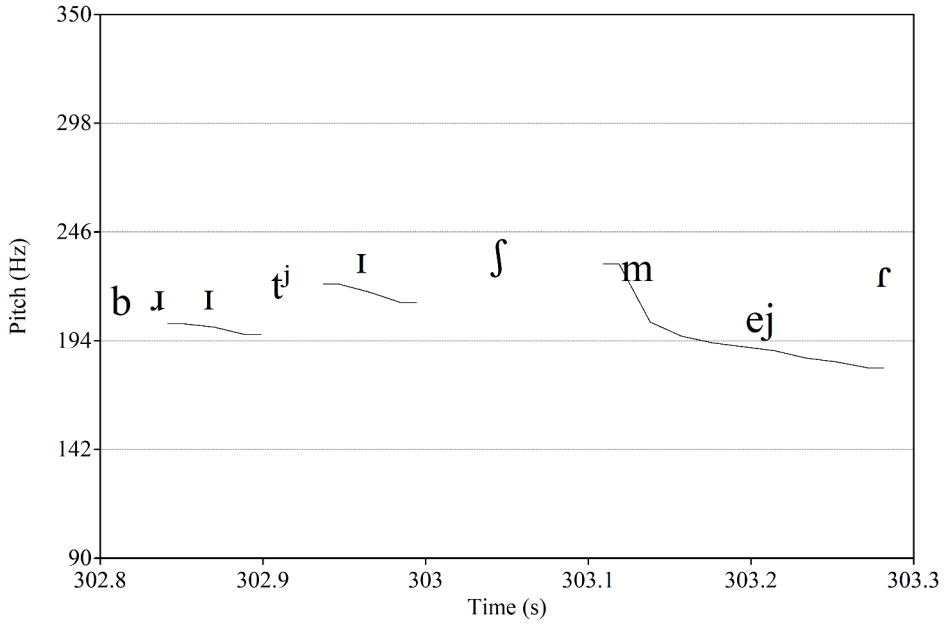


Figure 27: Pitch contour of british-made *spoken by speaker HK22 in her interview.*

The 'upstep' after an aspirated consonant is likely the result of F0 perturbation. In both English and Cantonese, vowels that follow consonants with long lag VOT tend

to have a higher pitch onset than those that follow unaspirated consonants (see for example Hombert 1978 for voicing-related F0 perturbation in English, and Francis et al. 2006 for an investigation on the same phenomena in Cantonese). F0 perturbations in tone languages have been found to last for a shorter duration than in non-tonal languages (Kirby 2018), and potentially a study on the interaction between VOT and F0 in HKE would shed light on how a largely non-tonal language with tone influence and arguable tonogenesis patterns in terms of voicing-related F0 perturbations. It will also be interesting to see if this phonetic in HKE difference becomes re-enforced in the future because not only are the pitch differences audible, but also the shape of the pitch contour in Figure 27, where both downstep and 'upstep' are happening, resembles the MHL intonation pattern in HKE enough that one could see it reanalysed it as MHL if the differences were more pronounced.

6.2 Comparing younger and older demographics

Both young and older speakers showed HKE-lo patterns that can be accounted for by descriptions of HKE intonation in previous literature, and both group also showed similar features that diverge from existing literature, including optionally H-toned secondary stress, cross-word-boundary H-spreading, non-IP-final L% boundary tone, L-tone creak, and the downsteps and 'upsteps' in spans of H-tone. Two minor differences are worth noting: 1) In general, more common H-spreading and otherwise flatter contours are seen in interviews compared with reading tasks, but the interview data from younger speakers especially has on average flatter F0 contours than older speakers, indicating that younger speakers have a slightly stronger tendency to use H-spreading and flatter contours in semi-spontaneous speech than the older speakers; 2) although both younger and older speakers have word-final rise, two older speakers in

particular exhibit this phenomenon relatively more frequently, while no younger speakers use word-final rise as much.

The most notable difference between the two age groups, however, is in the variety of levels of pitch accent saturation that individual speakers have. Both the percentages of overall pitch accents against the total number of syllables and the percentages of words saturated with pitch accents against the total number of words have a greater variety in younger speakers than in older speakers, especially in the reading tasks, where any difference in intonation patterns may become more articulated with more careful pronunciations and without the spontaneity of a semi-scripted dialogue. The mean difference is not statistically significant, but the difference in distributions is a noticeable one especially when first and third quantiles are considered (see again Figures 4, 5, 8, and 9). Assuming pitch accent assignment as a measurement for the strength of HKE-lo intonation, this difference indicates that younger speakers are distributed slightly more widely and more evenly on the HKE-gl-HKE-lo spectrum than the older speakers. In other words, both the strongest and weakest HKE-lo accents can be found in the younger group, while the older speakers, who are also fairly strong in their HKE-lo intonation, are a relatively more homogenous group.

This difference may be explained by the speakers' generational gap and what it means for the language education opportunities they were provided with, and the language ideologies and attitudes they were exposed to in their formative years. The younger and older speaker groups whose data is used in this study have an age gap of around four decades. The younger speakers were born around the 2000s and were schooled during the 2010s, living their entire life in the post-colonial period in Hong Kong; the older speakers were born around the early 1960s and grew up while the economic, social, and linguistic landscape of Hong Kong changed dramatically. The

younger speakers have had to navigate the identity of the cosmopolitan Hongkonger amidst the interplay of localism, mainland assimilation, and internationalisation, while the older speakers grew up during Hong Kong's shift from elitist bilingualism to mass bilingualism and entered the workforce as the new middle class in Hong Kong expanded, increasing the demand for English proficiency in professional settings.

Compared with the older speakers, the younger speakers have had a wider variety of options for English language education, both creating a wider working class with basic English proficiency and allowing the wealthier middle class to send their children abroad or to prestigious institutions previously exclusive to expats and local elites. In addition, the educational reforms since the 1970s that continued into the post-handover era increased the percentage of English speakers in the younger generations (Bolton 2000; Census and Statistics Department HKSAR 2021; Census and Statistics Department HKSAR 1996), thereby providing a larger speaker population as the basis for HKE variation. At the same time, parents are sending their children abroad or to international schools for their secondary education (Forse 2010), exposing the younger generation to more global varieties of Englishes. The older speakers, in comparison, had relatively less English-language resources during their younger years, and many of them learned English later in life for work, when the younger speakers acquired English earlier in life and use English in a wider variety of contexts, including with peers and at home. This may explain why the strongest HKE-gl speakers are found in the younger group.

Compared with the older speakers, the younger speakers also have grown up in a time with increasing acceptance of and positive attitudes toward the locally-oriented HKE accent. The older speakers come from a time when exonormative attitudes were common, and the past of elitist bilingualism being the dominant form of bilingualism

was within living memory. ‘Accented’ English would generally be considered to be a marker of non-standardness and lack of fluency, while Inner Circle Englishes, especially British English, was considered the desired norm (Luk 1998). Motivated by this language ideology, speakers of English may try to avoid overly strong HKE-lo intonational features in their speech. However, in more recent times after the handover, language attitudes started to shift. Increasing numbers of speakers started identifying with HKE in a positive way, as a marker of local in-group identity with symbolic values distinguishing them from both the Chinese mainland and the international English-speaking community (Edwards 2018). Indeed, Edwards found that not only do some speakers consider HKE-lo a boost to their sense of identity and belonging as Hongkongers, they also find HKE-lo features to be ‘fun to speak’ and see value and linguistic interest in the intonational features of the accent itself. These speakers may tend to accept their accent as a positive aspect of their Hong Kong identity and do not try to correct their strong HKE-lo intonation, since they may even consider it a legitimate variety of English. This may explain why the strongest HKE-lo speakers are also found in the younger group.

6.3 HKE intonation as a spectrum of varieties

Based on past literature on the varieties of HKE accents, Lam (2017; 2020) proposed a unified approach to HKE intonation from a variationalist perspective that separates it into HKE-gl and HKE-lo on two ends of a spectrum, with great possibilities of individual variation in between. This approach captures the results of the current study well.

A basic observation from this study that reflects recent literature is that not all syllables are assigned a pitch accent. Lam has already observed that pitch accent

assignment does not always saturate every syllable either in a scripted or spontaneous utterance even for speakers educated in local government funded schools who have fairly strong HKE-lo accents, and that the percentage of pitch accents decreases as speakers style shift into a more HKE-gl accent. The prototypical HKE-lo where every syllable is assigned a tone needs to be posited for a description of HKE intonation system and a model of HKE variation, but in order for it to exist, the speech not only has to be hyper-articulated to avoid effects of spontaneous connected speech, but we also have to assume the existence of the prototypical HKE speaker who is independent from the complex influences of global Englishes in contemporary Hong Kong and the still very present exonormative language ideologies that frame the HKE-lo accent as undesirable.

Since pitch accent patterns and the way boundary tones interact with them form the core features of HKE-lo intonation, the number of pitch accents indicates the material the speaker has to build HKE-lo patterns. This is why the level of pitch accent saturation can be a quantifiable indicator of where on the spectrum of HKE-gl-HKE-lo an utterance is located. In the present study, the percentage of tokens fully saturated with pitch accents of individual speakers ranges from 24.0% to 88.5%. These individuals show great variation in their frequency of using HKE-lo pitch accents, and are located on different places on the HKE-gl-HKE-lo scale.

Apart from confirming the utility of Lam's variationalist approach to HKE intonation and supporting the existence of a collection of HKE varieties on a spectrum measurable by the percentage of pitch accent assignment, this study also shows that certain kinds of pitch accents are more likely to be retained while others tend to be dropped as HKE intonation moves from local-oriented to global-oriented. In tokens with a part of their syllables assigned pitch accents, the most common position where

the pitch accent is absent is between the stressed H tone and word-final L tone (see the example of *university* [HHHØL] in Figure 25), and the syllables that most commonly retain their pitch accent are stressed H tones. Word-final L-tones and word-initial M-tones may also be retained, and either one of them can be retained without the other (e.g. *investment* [MHØ] vs. *university* [ØØHØL]. This means that if we adopt Lam's derivational account of HKE-lo intonation, as HKE moves away from the local-oriented end of the spectrum, the rules of pitch accent assignment typically stops applying before H-spreading; assuming Cheung's (2009) autosegmental approach may also lead to a good explanation, where the H-tone is associated with the boundaries of its H-tone spanning domain and optionally syllables adjacent to the stressed domain boundary, but never a syllable in the middle of a domain without 'connecting' to a stressed H-tone syllable via a span of H-tones. My proposals are preliminary since I could not control for intonational contexts precisely, but these partially saturated tokens are valuable and may shed some light on how systematic the interaction between HKE-lo and HKE-gl features is and whether there can be more definable and measurable varieties of HKE intonation that exist on the HKE-gl-HKE-lo spectrum.

7 Conclusion

Given that contemporary HKE developed against the background of economic, social, and linguistic changes in Hong Kong since the 1960s, and that HKE intonation can be viewed as a collection of varieties on a spectrum of HKE-lo and HKE-gl and has been in flux, this thesis examines polysyllabic word intonation patterns from the HKE data in a corpus originally collected for forensic linguistic projects. I investigate how HKE-lo and HKE-gl features interact, and how HKE intonation of older and younger

demographics reflect the societal changes in Hong Kong and the diachronic development of HKE over the past four decades.

Firstly, the phonology of HKE polysyllabic word intonation in this data can be described as largely local-oriented but with variations. In prototypical HKE-lo intonation, every syllable is said to be assigned a pitch accent (H, M, L) sourced from one of three level tones in Cantonese (High Level, Mid Level, Low Level), but in reality, 100% pitch accent saturation is very often not achieved, which is also the case in the present study, where some speakers do not have pitch accent saturated speech and show stronger HKE-gl pattern, while others with strong HKE-lo tendencies achieve high but not full levels of saturation. It is found that in pitch accent saturated speech, intonation patterns of multisyllabic words can be described with the following HKE-lo patterns: 1) primary stressed syllables and optionally pre-primary secondary stressed syllables are assigned H tones; 2) pre-stress syllables are assigned M, and post-final syllables are assigned L; 3) H tones optionally span from one stressed syllable to another across word boundaries, replacing any non-H syllables with H within the domain of H-spreading; 4) L% boundary tone marks the right edge of H-spreading and, when combined with a H-toned stressed syllable, surfaces as a high falling contour. In addition, I found the existence of an optional word-final rise that may combine with either H to create a rising contour, or with a L or H+L% sequence to create a dipping F0 contour which then rises again. It seems to mark the same H-spreading boundary as L% does, but it is unclear what additional functions it serves.

Secondly, there are two observations related to the phonetics of HKE intonation. Firstly, L-tones are consistently and systematically marked with creaky voice, making them phonetically similar to the Low Falling tone in Cantonese instead of Low Level,

which is said to be the source of L in HKE. Mergers in Hong Kong Cantonese involving the Low Level tone with Low Falling and with Mid Level indicate that the creaky voice quality of L in HKE may serve as an effective perceptual cue distinguishing between the L and M tones. Secondly, sequences of H-spreading sometimes feature optional internal downstepping and post-aspiration 'upstepping'. The latter is likely the result of re-enforced voicing-related F0 perturbations.

Thirdly, although the younger and older speakers largely show the same intonational patterns, the younger group has a wider variety of speakers on different positions on the HKE-gl-HKE-lo spectrum. The most HKE-lo and the most HKE-gl speakers are found in the younger group, and the younger speakers' distribution is more widespread compared to the more concentrated distribution of the older speakers. This seems to reflect the expanding public English language education in Hong Kong since the 1970s and the shift from elitist bilingualism to mass bilingualism on the one hand, and the growth of a new middle class who began to send their children abroad and to international schools on the other. Language attitudes toward HKE have also been changing since the change of sovereignty in 1997, where speakers shifted from having a dominant exonormative model treating local accents as undesired non-standard English to increasing acceptance and recognition of HKE as a positive marker of their Hong Kong identity and a legitimate variety of its own.

Finally, viewing HKE as a spectrum of global- and local-oriented varieties is a very useful approach in studying the complexity of its phonology, phonetics, and sociolinguistics. Measuring the percentage of pitch accent saturation is effective at capturing an utterance's position on the HKE-gl-HKE-lo scale. In addition, certain pitch accents are more likely to be dropped first than others as the scale slides from the HKE-

lo end to the HKE-gl end, H-spreading being the first feature to be interrupted or blocked within a polysyllabic word. More research is needed, however, to determine how systematic this pattern is, especially toward the HKE-gl end of the spectrum.

In conclusion, this thesis explores the intonation patterns of polysyllabic words in HKE in two generations of speakers through a unified variationalist approach treating HKE intonation as a spectrum defined by two varieties that interact with each other. Some findings confirm existing observations about HKE, while others diverge from them. Younger speakers show a wider distribution on the spectrum of variations compared to the older generation, reflecting the sociolinguistic changes that have taken place in Hong Kong. The linguistic environment and language ideologies are still in flux, and so does HKE intonation; thus further research is needed for a more comprehensive look at HKE intonation and its development.

References

Arantes, Pablo. 2021. parantes/time-normalized-f0: Version 2.0 of time-normalized-f0 Praat script. Zenodo. https://doi.org/10.5281/zenodo.4489748.

Bolton, Kingsley. 2000. The sociolinguistics of Hong Kong and the space for Hong Kong English. *World Englishes* 19(3). 265–285. https://doi.org/10.1111/1467-971X.00179.

Bolton, Kingsley & Helen Kwok. 1990. The dynamics of the Hong Kong accent: Social identity and sociolinguistic description. *Journal of Asian Pacific Communication* 1(1). 147–172.

Bratt-Paulston, Christina. 1980. *Bilingual education: Theories and issues*. Rowley, MA: Newbury House.

Brazil, David, Malcolm Coulthard & Catherine Johns. 1980. *Discourse intonation and language teaching*. Longman.

Cao, Grace W. & Peggy P. K. Mok. 2023. Cross-language features of filled pauses for trilingual speakers in Hong Kong. In *Proceedings of 20th International Congress of Phonetic Science*, 3814–3818. Prague. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2023/full_papers/686.pdf.

Cao, Grace W. & Peggy P. K. Mok. in preparation. Acoustic analysis of filled pauses in Cantonese-Mandarin-English trilingual speech and its implications on forensic phonetics. *Journal of Phonetics*.

Census and Statistics Department HKSAR. 1996. Hong Kong 1996 Population By-census: Main Tables. https://www.censtatd.gov.hk/en/EIndexbySubject.html?pcode=B1120084&scode=457.

Census and Statistics Department HKSAR. 2021. 2021 Population Census: Main Tables (Demographic). https://www.censtatd.gov.hk/en/EIndexbySubject.html?scode=600&pcode=D5212101.

Chen, Shuwen & Peggy P. K. Mok. 2015. Question intonation in Hong Kong English: Interaction between Cantonese and English.

Cheung, Winnie H. Y. 2009. Span of High Tones in Hong Kong English. *Annual Meeting of the Berkeley Linguistics Society* 35(1). 72–82. https://doi.org/10.3765/bls.v35i1.3599.

Cruttenden, Alan. 1995. Rises in English. In *Studies in general and English phonetics*, 155–173. Routledge.

Edwards, Jette G. H. 2018. English language schooling, linguistic realities, and the native speaker of English in Hong Kong. *Multilingua*. De Gruyter Mouton 37(3). 275–304. https://doi.org/10.1515/multi-2016-0012.

Forse, Chris. 2010. Fit for purpose? Why Chinese families choose international schools in Hong Kong. In Janette Ryan & Gordon Slethaug (eds.), *International education and the Chinese learner*, 59–72. Hong Kong University Press.

Francis, Alexander L., Valter Ciocca, Virginia K. M. Wong & Jess K. L. Chan. 2006. Is fundamental frequency a cue to aspiration in initial stops? *The Journal of the Acoustical Society of America* 120(5). 2884–2895. https://doi.org/10.1121/1.2346131.

Gussenhoven, Carlos. 2012. Tone and intonation in Cantonese English. In *Third International Symposium on Tonal Aspects of Languages*, 1–4. Nanjing, China. http://www.isca-speech.org/archive.

Gussenhoven, Carlos. 2014. On the intonation of tonal varieties of English. In Markku Filppula, Juhani Klemola & Devyani Sharma (eds.), *The Oxford handbook of World Englishes*, vol. 1. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199777716.013.29.

Hombert, Jean-Marie. 1978. Consonant types, vowel quality, and tone. In Victoria A. Fromkin (ed.), *Tone: A linguistic survey*, 77–111. New York: Academic Press.

Hong Kong Examinations and Assessment Authority. 2023. Benchmarking Studies on HKDSE: IELTS. https://www.hkeaa.edu.hk/en/recognition/hkdsebenchmarking/ielts/. (29 November, 2024).

Hung, Tony T. N. 2000. Towards a phonology of Hong Kong English. *World Englishes* 19(3). 337–356. https://doi.org/10.1111/1467-971X.00183.

Hung, Tony T. N. 2011. Hong Kong English. In Ea-Ling Low & Azirah Hashim (eds.), *English in Southeast Asia: Features, policy and language in use*, 113–133. Amsterdam: John Benjamins.

Hyman, Larry M. 2011. In defense of prosodic typology: A response to Beckman & Venditti. *UC Berkeley Phonology Lab Annual Reports* 7. https://doi.org/10.5070/P79P4074P7.

Kachru, Braj B. 1985. Standards, codification and sociolinguistic realism: The English language in the outer circle. In Randolph Quirk, Henry G. Widdowson & Yolande Cantù (eds.), *English in the world: Teaching and learning the language and literatures*, 11–36. Cambridge: Cambridge University Press.

Kirby, James P. 2018. Onset pitch perturbations and the cross-linguistic implementation of voicing: Evidence from tonal and non-tonal languages. *Journal of Phonetics* 71. 326–354. https://doi.org/10.1016/j.wocn.2018.09.009.

Lai, Mee Ling. 2001. Hong Kong students' attitudes Ttwards Cantonese, Putonghua and English after the change of sovereignty. *Journal of Multilingual and Multicultural Development*. Routledge 22(2). 112–133. https://doi.org/10.1080/01434630108666428.

Lai, Mee Ling. 2011. Cultural identity and language attitudes – into the second decade of postcolonial Hong Kong. *Journal of Multilingual and Multicultural Development*. Routledge 32(3). 249–264. https://doi.org/10.1080/01434632.2010.539692.

Lai, Mee Ling. 2012. Tracking language attitudes in postcolonial Hong Kong: An interplay of localization, mainlandization, and internationalization. *Journal of Cross-Cultural and Interlanguage Communication*. De Gruyter Mouton 31(1). 83–111. https://doi.org/10.1515/mult.2012.004.

Lam, Toni K. W. 2017. Intonational variation in Hong Kong English: A pilot study. *Asian Englishes*. Routledge 19(1). 22–43. https://doi.org/10.1080/13488678.2016.1277411.

Lam, Toni K. W. 2020. *Intonational variation in Hong Kong English*. Hong Kong: The Chinese University of Hong Kong (Hong Kong) Ph.D. https://www.proquest.com/docview/2605162084/abstract/DA104B3B16EC490CPQ/1. (8 January, 2024).

Leung, Benjamin K. P. 1996. *Perspectives on Hong Kong society*. Hong Kong: Oxford University Press.

Luk, Jasmine C. M. 1998. Hong Kong students' awareness of and reactions to accent differences. De Gruyter Mouton 17(1). 93–106. https://doi.org/10.1515/mult.1998.17.1.93.

Luke, Kapathy K. K. 2000. Phonological re-interpretation: The assignment of Cantonese tones to English words. In. Singapore. https://www.semanticscholar.org/paper/Phonological-re-interpretation%3A-the-assignment-of-Luke/bd2e6ba975c699a90a8deef599ef43d51c6beeed. (9 April, 2024).

Luke, Kapathy K. K. 2008. Stress and intonation in Hong Kong English. In *The 14th Conference of the International Association for World Englishes (IAWE)*. Hong Kong.

Matthews, Stephen & Virginia Yip. 2011. *Cantonese: A comprehensive grammar.* London and New York: Routledge.

McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner & Morgan Sonderegger. 2017. Montreal Forced Aligner: Trainable text-speech alignment using Kaldi. In *Interspeech 2017*, 498–502. ISCA. https://doi.org/10.21437/Interspeech.2017-1386.

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. https://doi.org/10.1017/S0954394513000161.

R Core Team. 2022. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Ran, Yunyun, Jeroen Van De Weijer & Marjoleine Sloos. 2020. Intonation in Hong Kong English and Guangzhou Cantonese-accented English: A phonetic comparison. *Journal of Language Teaching and Research*. London, United States: Academy Publication Co., Ltd. 11(5). 724–738. https://doi.org/10.17507/jltr.1105.07.

Warren, Paul. 2016. *Uptalk: The phenomenon of rising intonation*. First edition. Cambridge, England: Cambridge University Press.

Wee, Lian-Hee. 2016. Tone assignment in Hong Kong English. *Language* 92(2). e67–e87. https://doi.org/10.1353/lan.2016.0039.

Xu, Chenzi. 2024a. Parse.

Xu, Chenzi. 2024b. Citation tones. https://chenzixu.shinyapps.io/shinytone/. (5 February, 2025).

Yiu, Suet-yee. 2014. *Aspects of tone in Cantonese English*. *HKU Theses Online (HKUTO)*. The University of Hong Kong. https://doi.org/10.5353/th_b5481868.

Yu, Kristine M. & Hiu Wai Lam. 2014. The role of creaky voice in Cantonese tonal perception. *The Journal of the Acoustical Society of America* 136(3). 1320–1333. https://doi.org/10.1121/1.4887462.

Appendix

Material for the English reading task and mock interview created by Grace Cao and Peggy Mok:





Please read the case summary in English

1. Read the story

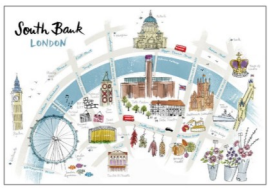
I knew Anna Sue when she worked in Hong Kong. Anna Sue was my business partner. Now she is the lead manager at the British Power Company.

I met Anna Sue twice in London. On 9th May, I flew to London and had dinner with her in the Zizzi café at South Bank, London. She was late because her daughter hurt her toe, so she came to Zizzi with her baby girl.

I ordered fish and chips, breakfast tea and she ordered sweetcorn soup. I told her that China was going to ban foreigners from buying properties in mainland China. I tried to persuade her to invest more money in the Hangzhou Fate city project. I asked for 400K pounds investment this time. Anna Sue said she needed more time to consider. She left earlier because her daughter cooed a lot and seemed sick. We then agreed to meet again on 10th May. After the brunch, I went to BigBang concert by the beach and visited a high school.







28

Please read the case summary in English

1. Read the story

On 10th May, we met in a Thai restaurant called "Sarah's Neighbour" next to the Post Office on the Labour Food corner. I told her that the shares were almost sold out, she needed to act quick. She then agreed. On the same day, Anna Sue transferred 400k pounds to my HSBC UK account.

To celebrate this, Anna and I had dinner at a restaurant called "Go and Cook". We ordered Sweet and Sour chicken with toast bread. After that, we went to a night club called "Take a sip" in South London. Before I went back to my hotel, I went to China town for a foot massage and shower massage.

I transferred 200k pounds to Sharon Li, I kept 200k pounds. With these money, I bought a house by the sea in Highland and a car. I also bought some stocks of the Microsoft searching engine Bing. I bought some souvenir. I went to Zara and bought 3 pairs of shoes and a light tan color suitcase. I also bought a British-made bin to my mum. I flew back to Hong Kong on 20th May.







29



Part 3:

You will attend an interrogation session with a British Police officer

30

Profiles



Name: Chris, So Tsz Ming
Birthplace: Hong Kong SAR
Sex: Female
Education: Hong Kong Polytechnic University, MBA
Occupation: a consultant in Central Fu Tai Investment Group; part-time photographer



Name: Anna Sue
Birthplace: London, England
Sex: Female
Relationship with you: a working partner in HK
Occupation: the lead manager of British Power company



Name: Sharon, Li Xiao Gao
Birthplace: Hangzhou China
Sex: Female
Relationship with you: Classmate in MBA courses
Education: Hong Kong Polytechnic University, MBA
Occupation: an entrepreneur

31

2. Interrogation

Location: London



1st meeting @London

Date: 9th May 2021

Location: Zizzi Café at South Bank, London

*(*Anna was late because her daughter hurt her toe. She came to the café with her baby girl)*

Order: fish & chips, Breakfast tea, sweetcorn soup

Discuss:

- You told Anna Sue: China is going to ban foreigners from buying properties soon. There is a 2nd round of investment project in the Hangzhou project (Name: The Fate city project), China.
- You asked for 400k pounds investment
- Anna Sue hesitated, she said she needed some time to consider.
- She left earlier because her baby cooed and seemed sick
- Agree to meet again on 10th May



Evening: You went to BigBang concert by the beach

2nd meeting @London

Date: 10th May2021

Location: a Thai restaurant called "Sarah's Neighbour" next to the Post Office in the Labour Food corner

Discuss:

- You told Anna: the shares are almost sold out, she needed to act fast

- Seal the deal; Anna Sue transferred 400K pounds to your HSBC UK bank account

Celebration with Anna Sue:

dinner @French restaurant called "Go and Cook"

- Order: French sweet and sour chicken with toast bread

drink @night club called "Take a sip" on High street.

- Order: cocktails

Foot massage and Shower massage @China Town (Anna left)

How did you spend Anna Sue's money?

- Transferred 200K pounds to Sharon Li, you kept 200K
- Bought a house by the sea in Highland; a car
- Bought some stocks of Microsoft searching engine Bing

Souvenir

- 3 pairs of shoes from Zara
- A light tan color Zara suitcase
- A British-made bin