Journal of Phonetics 110 (2025) 101412

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Contextual and paradigmatic effects on suspended contrast across generations: The case of Cantonese pinjam revisited

Alan C.L. Yu [a], Vivian Guo Li [b], Peggy P.K. Mok [c,*]

[a] University of Chicago Phonology Laboratory, Department of Linguistics, 1115 E. 58th Street, University of Chicago, Chicago, IL 60637, USA
[b] Department of Linguistics, Yale University, USA
[c] Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong, China

ARTICLE INFO

Article history:
Received 15 October 2023
Received in revised form 25 February 2025
Accepted 17 March 2025
Available online xxxx

Keywords:
Tone
Near merger
Incomplete neutralization
Tonal coarticulation
Phonetic analogy
Paradigmatic effect
Sound change

ABSTRACT

Suspended contrast refers to the phenomenon whereby sound change brings two phonemes into such close approximation that semantic contrast between them is suspended for native speakers of the language, without necessarily leading to complete merger or neutralization. The existence of suspended contrasts not only raises questions about the nature of the phonetics-phonology interface, but also for theories of sound change that assume sound change is biased toward selective maintenance of phonemes that contribute more to distinguishing existing lexical items in usage. Small differences supporting a suspended contrast are expected to disappear quickly given that they do not serve any apparent communicative functions. It remains a question whether a contrast can be suspended for a considerable period of time. This study revisits a case of suspended contrast in Cantonese between the lexical high rising tone and the high rising tone derived through morphological tone change (*pinjam*). We use an apparent-time approach to investigate the diachronic trajectory of this neutralization by comparing the distribution of this suspended contrast along both F0 and durational dimensions across two generations of Hong Kong Cantonese speakers. While this case of suspended tonal contrast has been in circulation for almost a century, our findings suggest that the distinction might be disappearing among the younger speakers. Only older speakers maintain a distinction between the lexical and derived rising tones, albeit in very restricted tonal contexts. The fact that this suspended tonal contrast exhibits great sensitivity to contextual and morphological influences may help explain the progression of this case of merger-in-progress.

1. Introduction

Suspended contrast refers to the phenomenon whereby sound change brings two phonemes into such close approximation that semantic contrast between them is suspended for native speakers of the language, without necessarily leading to complete merger or neutralization. A prototypical example is near merger, which refers to the existence of small and consistent articulatory differences between two lexical items that are hard to detect by native listeners. For example, certain New York speakers of English differentiate words like *source* and *sauce* in production but report no distinction between them in perception. Similar near mergers have been reported in other varieties of English, such as *fool* and *full* in Albuquerque (Di Paolo, 1988), *too* vs. *toe* and *beer* vs. *bear* in Norwich (Trudgill, 1974), *line* vs. *loin* in Essex (Labov, 1971; Nunberg, 1980), *meat* vs. *mate* in Belfast (Milroy and Harris, 1980; Harris, 1985), and *NEAR* and *SQUARE* vowel merger in New Zealand English (Warren, 2006), although cross-linguistic evidence of near merger is relatively scant despite notable exceptions (e.g., tonal near merger in Cantonese (Yu, 2007a,b), and /e/∼/ɛ/ in Shanghainese (Yao and Chang, 2016).

Similar to near merger, small but significant phonetic reflexes of underlying distinctions for phonologically or morphologically "neutralized" contrasts have also been reported under the heading of "incomplete neutralization". For example, in German, contrasts between consonants such as /t/ and /d/ (e.g. *Tier* [tiːr] 'animal' vs. *dir* [diːr] 'to you' and *leiten* [laitən] 'lead' vs. *leiden* [laidən] 'suffer') are traditionally described as neutralizing toward the voiceless consonants word-finally. Yet, various production and perceptual experimental results show that the two sets of final voiceless stops are consistently different phonetically (Port and O'Dell, 1985; Port and Crawford, 1989; Roettger et al., 2014). Similar findings of

* Corresponding author.
E-mail addresses: aclyu@uchicago.edu (A.C.L. Yu), v.li@yale.edu (V.G. Li), peggymok@cuhk.edu.hk (P.P.K. Mok).

https://doi.org/10.1016/j.wocn.2025.101412
0095-4470/

incomplete neutralization of voicing in word-final and preconsonantal positions have been reported for other languages, including Catalan (Dinnsen and Charles-Luce, 1984), Dutch (Warner et al., 2004), Lezgian (Yu, 2004), Polish (Giannini and Cinque, 1978; Tieszen, 1997), and Russian (Chen, 1970; Burton and Robblee, 1997; Dmitrieva et al., 2010; Kharlamov, 2014). Incomplete neutralization has also been reported in other types of neutralizing contexts such as flapping in English (Fox and Terbeek, 1977; Braver, 2014), coda aspiration and consonant geminiation in Eastern Andalusian Spanish (Gerfen and Hall, 2001; Gerfen, 2002; Bishop, 2007), schwa deletion in French (Fougeron and Steriade, 1997), monomoraic lengthening in Japanese (Braver and Kawahara, 2016; Braver, 2019), epenthetic stops in English (Fourakis and Port, 1986), /r/ and /l/ neutralization in post-nuclear position in Puerto Rican Spanish (Simonet, Rohena-Madrazo, & Paz, 2008), tone sandhi (Chen and Yuan, 2007; Yu, 2007b; Du and Durvasula, 2022), among others.

The existence of suspended contrasts such as incomplete neutralization and near merger raises difficult questions for theories of the phonetics–phonology interface that assume that phonological representations in the lexicon are categorical, contrastive elements, and the phonetic implementation component computes the degree and timing of articulatory gestures, which are gradient and variable (e.g., Cohn, 1990; Cohn, 1993; Keating, 1990; Pierrehumbert, 1990). Lexical information in such models cannot bypass phonological encoding to influence the phonetic implementation directly and the categorical form of a lexeme determines the phonetic outcome. Phonetic variations on the surface are considered artifacts of the context or performance-induced fluctuations. Such models are not equipped to account for systematic subphonemic differences between representations that are otherwise assumed to be identical.

Diachronic studies of suspended contrasts can offer important implications for both theories of the phonetics-phonology interface as well as theories of sound change. To begin with, if suspended contrasts can be shown to exhibit structured heterogeneity characteristic of other sound changes of progress (e.g., if they vary along sociolinguistic dimensions such as age and gender), it would suggest that the subphonemic differences observed cannot be relegated to artifacts of context or performance-induced fluctuations, but are part of the linguistic system that can undergo systematic variation and change. Diachronic studies of suspended contrasts also have important consequences for functional explanation of sound change. As Labov (1987) noted, suspended contrasts suggest that “the communicative role of phonemic contrasts can be suspended for a considerable period of time without disrupting the integrity of the word classes and the system they participate in” (p. 319). Despite the existence of near mergers (and incomplete neutralization), he maintained that phonemes function “to distinguish words”, but the historical development of the phonemic system “is not narrowly controlled by that communicative function.” While many cases of suspended contrast have been reported over the years, it remains unclear whether suspended contrasts can be maintained for a “considerable period of time”. Admittedly, what constitutes a “considerable period of time” is a matter of debate. Nonetheless, it is puzzling that no study has yet been dedicated to investigating the diachronic development of suspended contrast, be it near merger or incomplete neutralization. Understanding how suspended contrast is maintained and changed is important especially since theories of sound change that appeal to the functional load hypothesis assume that sound change is biased toward selective maintenance of contrasts that contribute more to distinguishing existing lexical items in usage (Blevins and Wedel, 2009; Wedel, 2012; Wedel et al., 2013). Such theories predict that small differences supporting incomplete neutralization and near mergers are expected to disappear quickly given that they do not serve any apparent communicative functions. On the other hand, models of sound change that permit a more articulated phonological system with phonemic and subphonemic categories (e.g., Stevens et al., 2019; Gubian et al., 2023) can accommodate the existence of suspended contrasts or intermediate subphonemic categories more readily. Empirical investigations of the diachrony of suspended contrasts can offer much needed benchmarks for evaluating model predictions.

This study revisits a case of suspended tonal contrast in Hong Kong Cantonese. We employed an apparent-time approach to investigate the diachronic trajectory of suspended contrast by comparing the distribution of phonetic differences between lexical and morphologically-derived tones across two generations of Hong Kong Cantonese speakers. Our findings not only provide much needed empirical support for Labov’s claim that suspended contrasts can maintain for a considerable period of time, but also demonstrate the maintenance of suspended contrast can be highly contextual. This study begins with an overview of tonal morphology in Hong Kong Cantonese in Section 2. A description of the methodology of our acoustic investigation appears in Section 3 and the results are reported in Section 4. Section 5 discusses the implications of our findings and concludes.

2. Cantonese *pinjam*

Cantonese is a tone language where each syllable, which often corresponds to a free morpheme, is associated with a tone. The tonal inventory of Cantonese consists of six tones: high-, mid- and low- level, high- and low- rising, and low-falling (see Table 1; Bauer and Benedict, 1997). In some varieties of Cantonese, the high level tone on open or nasal-final syllables has a high falling allotone (see, e.g., Bauer and Benedict, 1997). However, the high-falling variant is no longer used in Hong Kong Cantonese (HKC), the variety we are investigating here. For ease of reference, tones will be referred to by the tone values below (e.g., Tone 2 = T25; Tone 4 = T21).

Tone change in Cantonese, also known as 變音 *pinjam*, refers to the phenomenon where a syllable is realized with a tone other than its lexical (canonical) tone (Bauer and

Table 1
Cantonese Tones. Tone values are adapted from Bauer and Benedict (1997), where 5 represents the pitch ceiling and 1 represents the pitch floor.

Tone Type	Tone	Description	Tone Values
level tones:	Tone 1	high level	55
	Tone 3	mid level	33
	Tone 6	low level	22
rising tones:	Tone 2	high rising	25
	Tone 5	low rising	23
falling tone:	Tone 4	low falling	21

Benedict, 1997).[1] The specific type of tone changes this study is concerned with is the high-rising tone change, where a syllable with a non-high-rising tone (i.e., T33, T23, T22, T21) comes to bear a high-rising tone (i.e. T25*),[2] typically in morphological contexts associated with signaling familiarity (Liu, 2016; Yip, 2002), perfective (Yu, 2007b) or some other note-worthiness; T55 and T25 do not participate in this tone change. For example, the morpheme 年 "year" is canonically realized with a low-falling tone (e.g., *jaa33 nin21* 廿年 "twenty years"), but its tone changes to a high-rising in the phrase *gau22 nin25** 舊年 "last year". Likewise, the surname *ziu22* 趙 "Chiu" could surface with *pinjam* in the endearing vocative *lou23 ziu25** "old Chiu". While *pinjam* often targets the final syllable, it is not restricted to that position (e.g., 蛋糕 *daan21 gou55* "cake" vs. 蛋仔 *daan25* zai25* 'eggette', 黃金 *wong21 gam55* 'gold' vs. 黃 'sir' *wong25* soe4*, "Teacher/Coach Wong"). *Pinjam* is not an obligatory process (e.g., 雞蛋仔 "eggette" can be pronounced as *gai55 daan22 zai25* or *gai55 daan25* zai25*), although little is known about the nature of its optionality and frequency of usage.[3]

廿年	jaa33 nin21	"twenty-year, i.e., twenty years"
舊年	gau22 nin25*	"old-year, i.e., last year"
後趙	hau22 ziu22	"The Later Chiu Dynasty"
老趙	lou23 ziu25*	"Old Chiu"

Phonological processes that apply in the lexicon, such as tonal morphological processes like *pinjam*, are generally assumed to be structure-preserving (e.g., Kiparsky, 1982; Kiparsky, 1985). That is, they are not expected to introduce contrasts which are not otherwise motivated as underlying in the language. This prediction is generally assumed by most theoretical treatments of this phenomenon (e.g., Bao, 1999; Yip, 2002; Liu, 2016; Alderete et al., 2022) and it finds resonance in some descriptions of this tone change. Hashimoto (1972, p. 93), for example, considered the high rising tone resulting from tone change (hereafter: DR or derived rising) as similar to the lexical high rising tone (hereafter: LR) and did not mention any specific difference. Likewise, Bauer and Benedict (1997) found no evidence that LR and DR are distinct, basing their claim on acoustic evidence taken from recordings of twelve speakers from Hong Kong, Guangzhou, and Macau, whose ages ranged from 28 to about 65 years. While they did not provide direct comparison between all LR and DR pairs, for those they did, the contours, as well as the F0 values at the onset and offset of both DR and LR were qualitatively similar. However, most of the tokens for DR came from one phrase, 阿二 /aa33 ji25*/ "vocative-two; (the person who is known as) number two." The LR and DR tokens examined were also not recorded in the same context. The LR tokens were often recorded in isolation (citation form), while the DR tokens were always embedded in carrier sentences. Given the lack of a controlled tonal context and a systematic examination of the statistical properties of F0 values of LR and DR, the conclusion that LR and DR are not distinct must be taken with caution.

Despite the theoretical and empirical justifications laid out thus far, there is nonetheless a long tradition, going back to as early as the 1940s, that documented subtle differences between LR and DR, even though the details of the difference vary across reports. O'Melia (1939), cited in Whitaker (1956), observed that the contour of DR was overall very similar to that of LR, but DR had a higher offset than LR. Chao (1947) treated DR as a different tone category from LR, with LR being shorter in duration and having a slightly higher onset than DR. Whitaker (1956) described high-rising tone change as a process of attaching a high tone to the end of the original tone. In this conception, LR and DR should differ at both onset and offset: the onset of DR may be higher or lower than that of LR depending on the onset of the original tone, and the offset of DR would be higher than that of LR.[4] Yu (2007b) reported the findings of a controlled experiment where the stimuli consisted of ten LR/DR minimal pairs and systematic statistical investigations were carried out to examine potential F0 differences at the onset, inflection point, and peak of F0 contours of DR and LR among six speakers of Hong Kong Cantonese. Yu found that DR and LR were in a state of near-merger where the two were consistently distinguished in production with DR showing a higher overall F0 contour and a steeper F0 rise than LR, but the listeners in his perception study were unable to distinguish the two tones. There was considerable individual variation, however. Only one speaker showed a consistent distinction at all three measurement points. Half of the speakers showed significant differences at the inflection point and the peak but not at the onset; one speaker did not show any significant difference throughout the tone contours. Given that the participants in Yu's study were all residing in the US at the time of recording, it is also unclear if the observed individual variability might be influenced by their reduced exposure to Cantonese on account of their being in an English-dominant environment. More recently, Liu (2016) examined the durational properties of LR and DR produced by a 77 year old Cantonese speaker from Guangdong and ten Cantonese speakers between the ages of 20 to 25 studying in Taiwan. She found that DR (mean duration = 0.518 sec) is almost twice as long as LR (mean duration = 0.256 sec) for the older speaker. While DR is longer than LR among the younger speakers, the difference is much smaller by comparison (LR = 0.335 sec vs. DR = 0.395 sec). It is worth noting that Cantonese does not have a consistent quantity distinction in all environments except for the low vowels /a/ and /a:/. Given that the purported duration distinction between LR and DR is not restricted to words with low vowels, it suggests that the LR/DR might also exhibit subphonemic differences along the duration dimension. To be sure, the claim of a duration distinction between LR and DR must be taken with caution as the LR and DR stimuli in Liu (2016) were not fully comparable in terms of vowel quality and syllable structure. A more controlled examination is needed to ascertain the existence of an LR/DR duration distinction.

[1] The notion of "tone change" used here is not to be confused with non-morphologically related cases of tone mergers-in-progress. That is, while standard descriptions of Cantonese often assume a six-tone system (Cheung, 1986; Bauer and Benedict, 1997), some tone pairs have been reported to be merging: T25 vs. T23, T33 vs. T22, and T21 vs. T22 (e.g., Mok et al., 2013; Fung and Lee, 2019). These cases of tone mergers are not restricted by morphological contexts. See more discussion in Section 5.

[2] For ease of reference, all changed tones will be marked by the * symbol.

[3] Impressionistically, some forms could sound awkward without *pinjam* while others do not. For example, the first author prefers 黃"sir" 'Teacher/Coach Wong' as *wong25* soe4* and not ?*wong21 soe4*, although 盧"sir" 'Teacher/Coach Lou' sounds equally fine as *lou25* soe4* or *lou21 soe4*.

[4] Yu (2007b) found evidence of this type of tonal variation in Cantonese in the context of rising tones derived from syllable fusion (e.g., *gin33 zo25* "to meet-PERFECTIVE → *gin25*; *bong22*-PERFECTIVE → *bong25*).

The inconsistent findings of subphonemic differences between LR and DR across studies and the preliminary evidence of individual variability in Yu's study suggest that LR and DR might be in a state of flux within the speech community. Given that various sound changes in progress across generations have been documented in the language (To et al., 2015), LR and DR might very well be undergoing a change in progress where the merger between DR and LR is inching toward completion. This study aims to investigate the current status of the neutralization of the DR/LR contrast by examining the realization of DR and LR across two gender- and age-balanced cohorts of HKC speakers. If LR and DR were undergoing a merger-in-progress, the LR/DR distinction may be more apparent in the older generation than in the younger one. Given that sociolinguistic studies have consistently observed that male and female speakers participate in sound change in progress at different rates (e.g., women often lead in sound changes in progress; Gordon and Heath, 1998), the DR/LR merger might be more advanced in women than in men.

Another goal of this study is to identify what factors might be responsible for the maintenance and disappearance of this suspended contrast. Previous studies have found that linguistic context can be an important conditioning factor in mergers. Warren (2006), for instance, examined the merger between NEAR and SQUARE vowels in New Zealand English and found that the merger was complete in all contexts in young speakers' speech, but, for the middle-aged speakers, merger is complete only when the consonant preceding the vowel is coronal. Given that previous studies on tonal realization in Cantonese have consistently observed that tonal realization is heavily influenced by the tonal environment (e.g., Wong, 2006), are there tonal contexts where the LR/DR distinction are more likely to be maintained or suppressed? Morphological factors could also influence the maintenance of the LR/DR difference. To begin with, various scholars have raised the possibility of incomplete neutralization effects being due to the co-activation of paradigmatically related forms (Ernestus and Baayen, 2006; Yu, 2007a; Roettger et al., 2014). Thus, when speakers pronounce the DR form of 年 "year" *nin25**, for example, they also activate its paradigmatic counterpart *nin21*. This co-activation of the related forms with other tones could influence the speech production mechanism in subtle ways, leading to incomplete neutralization. Related to this idea are various studies that reported paradigmatic uniformity effects at the phonetic level (e.g., Steriade, 2000; Seyfarth et al., 2018; Braver, 2019). That is, morphologically complex structures exhibit phonetic properties that differ from those occurring in simplex forms. Seyfarth et al. (2018), for example, found that the stems of English inflected words (e.g., *frees*) are longer than their counterparts in homophonous mono-morphemic words (e.g., *freeze*). They attribute the extra duration to the co-activation of the longer articulatory gesture of the base stem of the inflected word (e.g., *free* in the case of *frees*), although this type of phonetic analogy effect might also be mediated by the frequency of the bare stem (e.g., Engemann and Plag, 2021). In the case of *pinjam*, Yu (2007a) found evidence that the realization of DR can be dependent on the tone of the non-derived paradigmatic neighbor. Specifically, he found that a DR that has a corresponding T22 tone (DRT22 for short) has a lower F0 profile than a DR that has a corresponding T33 tone (DRT33), but such differences are most apparent in nasal-final syllables and only marginally so in open syllables. Yu (2007a) only examined a limited number of stimuli and only focused on the difference between DR-T22 and DR-T33. This study aims to expand the scope of previous investigations by examining the effects of tonal context and paradigmatic tonal relationship across two generations of Cantonese speakers. To the extent that the LR/DR distinction is disappearing, are there tonal contexts or paradigmatic relationships that encourage the neutralization? Are these factors influenced by sociolinguistic factors such as age and gender? If the subphonemic tonal differences observed between DR and LR is influenced by sociolinguistic factors, it might complicate the co-activation explanation for incomplete neutralization as it is unclear how sociolinguistic factors could mediate effects of lexical co-activation in a regular way. To summarize then, the investigation below will focus on the following research questions:

- **Research Focus I:** Is there evidence for subphonemic differences between LR and DR? If so, are the subphonemic distinctions differently manifested across gender and age groups?
  - Predictions: If subphonemic differences between LR and DR exist and this is a case of change in progress, younger speakers might no longer maintain this difference. To the extent that females are more likely to be the leaders of sound change, we might expect LR and DR to be completely merged in females.
- **Research Focus II:** To the extent that subphonemic differences between LR and DR are observed, are they affected by
  - the preceding tonal context?
    - Predictions: If LR and DR were affected by tonal contexts differently, the subphonemic differences between LR and DR might be more robust in some contexts than others.
  - the tone of DR's paradigmatic neighbor?
    - Predictions: DRs with different paradigmatic neighbors might exhibit different neutralization patterns, depending on whether the paradigmatic effects raise or lower the F0 contour of DR toward that of the LR.

3. Methods

3.1. Participants

We recorded forty age- and gender-balanced speakers of Hong Kong Cantonese. The young speakers (N = 20) range between 21 and 27 in age (mean: 23.6, SD: 1.54), all born and raised in Hong Kong. The old speakers (N = 20) range between 64 and 85 in age (mean: 72.05; SD: 5.08); eleven of them were born and raised in Hong Kong. Nine were born in various places in Guangdong who had moved to HK at the age of twenty or younger (see Supplemental Materials for details of speaker background). Among them, there is only one old speaker whose home language was not Cantonese, but he nonetheless speaks Cantonese fluently. All forty speakers used Cantonese as their primary language, i.e., the language of daily life.

3.2. Materials

The stimuli consisted of 34 near-minimal pairs of DR and LR in CV(V)(N) syllables. An additional 16 disyllabic words were also included as fillers. The full list of lexical items used are presented in the Supplemental Materials. The target syllables were always the second syllable in a disyllabic word. Care was taken to include the full set of possible tonal combinations in the disyllabic words, i.e., the tone on the first syllable may be T55 (N = 7), T25 (N = 5), T33 (N=5), T21 (N = 6), T23 (N = 6) or T22 (N = 5), and the non-rising counterpart of the tone-change syllable may be T33 (N = 8), T21 (N = 10), T23 (N = 6) or T22 (N = 10). Syllables carrying T55 and T25 do not undergo this tone change. To the extent possible, onset consonants are matched between the target and control pairs at least in manner and voicing, if homophones cannot be found. However, given lexical restrictiveness of the tone change process, the distribution of onset types cannot be fully controlled for in the stimuli. Additional details can be found in Section 4.

3.3. Recording

All stimuli were embedded in the carrier phrase, 我讀__ 俾你聽 /ŋɔ23 tʊk22 __ pei25 nei23 t^{h}ɛŋ33/ “I read __ for you (to) hear”. Each speaker read the randomized stimuli, presented in a printed form, in colloquial Cantonese in a sound-treated room. They read the list from top to bottom in a block and repeated the block two additional times. If the speakers misread a token, s/he was prompted to reread it immediately or after a block. The recordings were made with a ZOOM H2n Handy Recorder at a sampling rate of 44100 Hz.

3.4. Measurements

Rime intervals were first manually segmented in Praat (Boersma and Weenink, 2017). Rime duration was measured from the vocalic onset to the end of the syllable. The F0 tracking was performed by taking measurements of F0 at eleven equidistant points starting at the beginning of the syllable rime using a revised version of ProsodyPro 6 beta (Xu, 2013). ProsodyPro 6 beta tracks pitch at the beginning (but not the end) of each interval, thus it could not capture the value at tone offset; we revised the script so as to include F0 at both edges.

4. Analysis and results

This section begins with a report of the results of the F0 analysis, followed by a report of the findings of the duration analysis. The dataset, the analysis scripts, and supplemental materials can be found at https://osf.io/hev8p/.

4.1. F0 analysis

Given that we are interested in comparing F0 contours with varying temporal dynamics, a statistical technique that can capture potential F0 differences occurring at different temporal span of the tonal contours is needed. To this end, we modeled the trajectory of F0 over time as a function of lexical status using Generalized Additive Mixed Models (GAMMs: Wood, 2017). GAMMs allow us to model F0 over the entire sonorous portion of the rime without being restricted to a trajectory shape, thus allowing the possibility that LR and DR differ in F0 trajectories. Unless specified otherwise, we built GAMMs using the `bam()` function from the *mgcv* package, using the `fREML` function. The smooths used thin plate regression splines as basis (Wood, 2017). The dependent variable is F0 in semitone, normalized relative to median F0 of each speaker. The following discussion is divided into three sections, each focus on a distinct model with slightly different emphases and complexities. Model 1 focuses on the main effects of lexicality, tonal coarticulatory influence from the preceding syllable, gender, and age. Model 2 examines the interaction between these four factors. Model 3 focuses on the effects of morphological tonal neighbors on the realization of the target tone in different tonal contexts.

4.1.1. Model 1: The main effect of lexicality

This section focuses on a basic model that examines the effects of lexicality, tonal coarticulatory influence from the preceding syllable, gender, and age. The model includes parametric factors of LEXICALITY (DR vs. LR), tonal specification of the preceding syllable (CONTEXT, six levels: T55, T25, T33, T21, T23, T21), and the GENDER and AGE group of the participants. All categorical variables were treatment-coded (the reference level for LEXICALITY is “DR”, CONTEXT “T21”, GENDER “female”, and AGE “old”). To model the time-course of F0 and differences between each level of the factors of interest (i.e., lexicality, tonal context, gender, age group), smooth terms were included to model (i) the non-linear pattern of F0 over time (POINT, range from 1 to 11) and (ii) F0 over time separately for each level of the parametric factors. To reduce biases on the F0 contours introduced by the onset consonant, syllable structure (i.e. open vs. closed; Yu, 2007a,b), and vowel height (i.e. high, mid, vs. low; Siddins and Harrington, 2015), random smooths of time by each of these variables were included. Variation in F0 over time according to rime DURATION (as a proxy for speaking rate) was modeled in terms of tensor product smooths. By-speaker differences were captured using factor smooths, analogous to random intercepts and slopes in linear mixed-effects models. These non-linear random effects model any non-linear differences over time with respect to the general time pattern for each speaker and for each level of LEXICALITY, allowing for the possibility that speakers differ in the effect of LEXICALITY. The model also includes by-word random intercepts to account for F0 variation across words. Model comparison was conducted by comparing the AIC of two models differing in one fixed effect using the `compareML` function in the *itsadug* package in `R`. Together with the model comparisons and model plots, the statistics provided by the model summaries were used to determine whether each predictor contributed to the variance explained by the model.

An AR1 model design was used to control for autocorrelation within F0 trajectories, by defining the starting point for each time series as the beginning of the syllable rime and the function `acf_resid()` from *itsadug* (van Rij et al., 2022) was used to estimate the autocorrelation parameter `rho`, which was used to fit a new GAMM accounting for autocorrelation of residuals. The model also uses the scaled-*t* distribution because of the residuals showing a heavy tail. The final model has an adjusted R^2 of 0.66 and explained 57.6% of the deviance, suggesting a decent fit. Model diagnostics using `gam.check()` suggested appropriate levels of smoothing (k values). The model formula used is given below. To facilitate comprehensibility of the model specification, comments are provided after each model term following the # symbol. Full model outputs with both parametric and smooth terms are provided in the Supplemental Materials.

Model 1:

```
F0 ~ LEXICALITY + CONTEXT + AGE + GENDER + #parametric factors
s(POINT, by = LEXICALITY, k = 9) +
#2 separate smooths for each of the two levels of LEXICALITY
s(POINT, by = CONTEXT, k = 9) + #6 separate smooths for each of the six levels of CONTEXTS
s(POINT, by = AGE, k = 9) + #2 separate smooths for each level of AGE
s(POINT, by = GENDER, k = 9) + #2 separate smooths for each level of GENDER
te(POINT, DURATION) + #tensor smooth for DURATION
s(POINT, ONSET, bs = "fs", k = 9, m = 1) + #factor smooth for ONSET
s(POINT, SYLLABLETYPE, bs = "fs", k = 9, m = 1) +
#factor smooth for SYLLABLETYPE
s(POINT, HEIGHT, bs = "fs", k = 9, m = 1) + #factor smooth for vowel HEIGHT
s(POINT, SPEAKER, bs = "fs", k = 9, m = 1) +#factor smooth for speaker
s(POINT, SPEAKER, by = LEXICALITY, bs = "fs", k = 9, m = 1) +
#separate factor smooths for speaker for each level of LEXICALITY
s(WORD, bs = "re") #by-word random intercept
```

Model 1 reveals a significant main effect of LEXICALITY; in general, LR is 0.172 semitone lower than DR ($p < 0.05$). The model also reveals significant difference in the smooths for LR and DR. Fig. 1 provides the estimated smooths for each level of LEXICALITY, AGE, and GENDER. The statistical significance of the difference between levels for LEXICALITY, AGE, and GENDER was evaluated by obtaining estimated marginal means using `emmeans` (Lenth et al., 2021), calculating the estimated difference between levels at each measurement point, and marginalizing over irrelevant variables (Speaker, Word, Onset, Vowel Height, Syllable Type, and rime duration), with p-values corrected using the Tukey HSD method to minimize Type I errors. The leftmost column shows the estimated smooths for each level of LEXICALITY with the pointwise 95% confidence intervals shown in shaded bands (upper panel) and a difference smooth (lower panel) between DR and LR, which allows us to see where and in what way the trajectories differ (Sóskuthy, 2021). The shaded bands represent the pointwise 95%-confidence interval of the difference smooth. Regions where the DR-LR difference are significantly different from zero, as determined by the comparison of the estimated marginal means using `emmeans` (Lenth et al., 2021) with p-values corrected using the Tukey HSD method, are highlighted in pink. The difference smooth shows that the F0 trajectory for the derived tone, DR, is significantly higher than the trajectory of the lexical tone, LR, starting from measurement point 3 (20% of the rime duration) until the end of the rime; when the difference smooth is above zero, the F0 of DR is higher than the F0 of LR. The difference is roughly between 0.1 and 0.2 semitones. The middle column shows the effects of AGE. Significance differences are observed, but in different direction at different time spans. The estimated smooths show that the

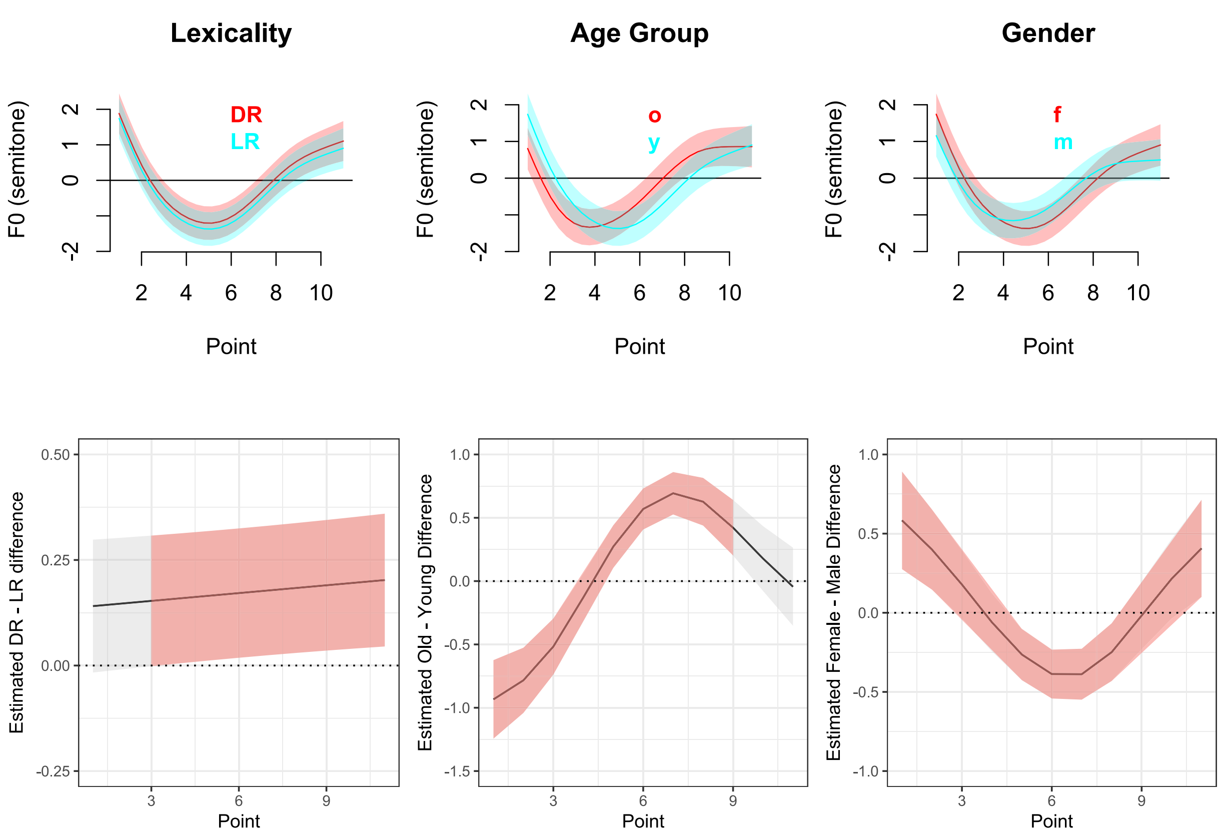

Fig. 1. Upper panels: non-linear smooths (fitted values) for LR and DR (LEXICALITY; leftmost), older and younger speakers (AGE; middle), and female and male speakers (GENDER; rightmost) in Model 1. Shaded bands represent the pointwise 95%-confidence interval. Lower panel: Differences between the two (non-linear) smooths comparing two levels of by-variable: LR (blue) vs. DR (red); older (red) vs. younger (blue) age groups; females (red) vs. males (blue). The shaded bands represent the pointwise 95%-confidence interval; the red bands represent regions where the estimated difference between DR and LR has an adjusted p-value of < 0.05 using the Tukey HSD method obtained via the `contrast()` function in `emmeans` (Lenth et al., 2021). o = older cohort; y = younger cohort; f = female; m = male.

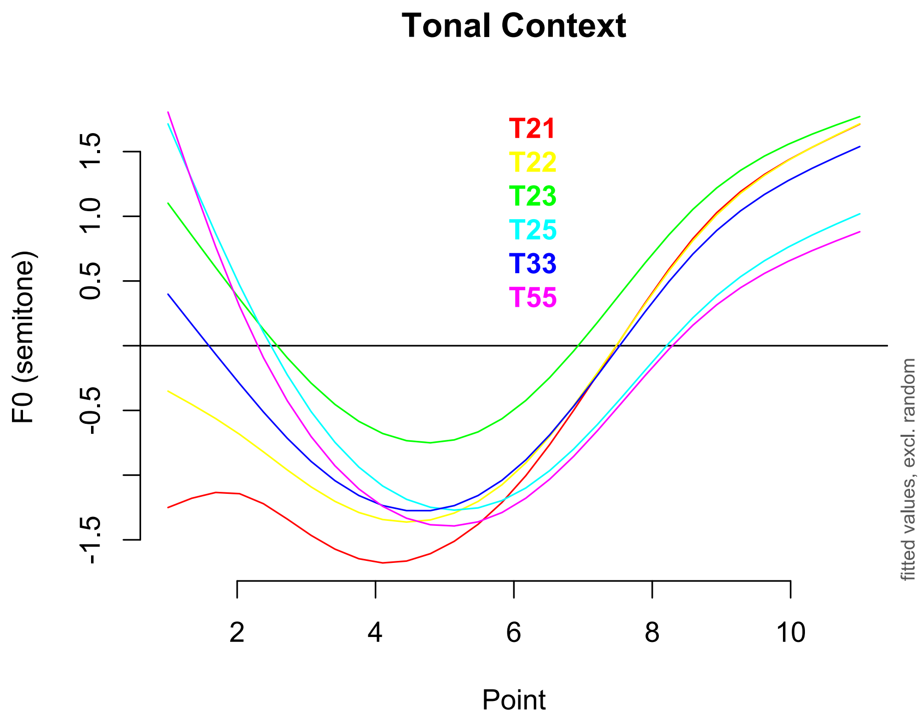

Fig. 2. Effects of the preceding tone context on the realization on the F0 of the target rising tone.

older speakers show a lower tonal onset, but a higher tonal offset beginning from measurement point 5. This pattern suggests that the older speakers have an earlier F0 inflection point for the F0 rise than the younger speakers. The rightmost column shows the estimated smooths for the male and female speakers; the difference smooth (female F0 smooth - male F0 smooth) shows that female participants have a higher F0 onset, but a lower F0 inflection point, compared to their male counter parts.

The preceding tonal context also exerts a significant influence on F0 trajectories. Fig. 2 shows the estimated smooths of the F0 contours of the rising tone in each preceding tonal context. Qualitatively, F0 onset of the target tone is highest when the preceding tone has a high tonal offset (i.e. T55 and T25) and the target tone's F0 onset is lowest when the preceding tone has a low tonal offset (i.e. T22 and T21). The F0 offset of the target rising tone is higher when following a low tone (e.g., T22 or T21) than when following a high tone (e.g., T55 or T25), which is likely the result of post-low bouncing (Gu and Lee, 2009).[5] The target rising tone also has a shallower rise when preceded by T55 and T25. The F0 inflection point is also highest when the preceding tone is T23, suggesting that the rising trajectory is shallowest in this context.

Our findings thus far point to significant differences between the tonal contours of lexical and derived rising tones. DR is higher than LR across most of the rime duration. This pattern echoes the observations reported in Yu (2007b), suggesting that this sample of Hong Kong Cantonese speakers, as a group, exhibit patterns of incomplete neutralization between LR and DR. Several observations are noteworthy, however. First, the inclusion of by-speaker random smooths for each level of LEXICALITY significantly lowers the AIC and ML scores, which suggests that there is significant individual variation in how the F0 contours of DR and LR are realized. Given that F0 trajectories differ across tonal contexts, and age and gender groups, some of the variation might originate from the interaction between these linguistic and social variables. Specifically, the current model does not take into account the possibility that the LR/DR difference might not hold across age and gender groups and across tonal contexts to the same degree, even if they remain distinct overall. The next section examines the interaction between lexicality and these other linguistic and social factors.

4.1.2. Model 2: Effects of preceding tonal context on LR and DR across gender and age groups

This section focuses on the effects of tonal context on the realization of LR and DR across gender and age groups. Specifically, a second model (Model 2) was built where F0 was modeled with a GAMM that includes a parametric term for the interaction of AGE group, GENDER group, tonal CONTEXT, and LEXICALITY (i.e. the AGEGENDERCONTEXTLEXICALITY variable) and a difference smooth term over POINT with a by-AGE/GENDER/CONTEXT/LEXICALITY interaction variable. This factor interaction (Wieling, 2018) encodes the forty-eight levels of interaction between two levels of GENDER, two levels of AGE, six levels of CONTEXT, and two levels of LEXICALITY. This factor interaction term was treatment-coded; the reference levels are as follows: GENDER = female, AGE = o, CONTEXT = T21 and LEXICALITY = DR. All other control variables and random smooth specifications are similar to Model 1. Because of the model complexity, Model 2 was built using the `fREML` method from the *mgcv* package instead of the more time-consuming `ML` method. The statistical significance of the difference between LR and DR in each tonal context for each age and gender

[5] Unlike what is observed in this study, Gu and Lee (2009) found that post-low bouncing only appeared after T21 and not T22 in their Cantonese dataset.

combination was evaluated by obtaining estimated marginal means using `emmeans` (Lenth et al., 2021), calculating the estimated difference between DR and LR at each measurement point, and marginalizing over irrelevant variables (Speaker, Word, Onset, Vowel Height, Syllable Type, and rime duration), with p-values corrected using the Tukey HSD method to minimize Type I errors. Full model outputs with both parametric and smooth terms in both models are provided in the Supplemental Materials. The model formula used for Model 2 is given below.

Model 2:

```
F0 ~ AGEGENDERCONTEXTLEXICALITY + #a parametric factor for the interaction between AGE, GENDER, tonal CONTEXT, and LEXICALITY
s(POINT, by = AGEGENDERCONTEXTLEXICALITY, k = 9) +
#separate smooths for each combination of the interaction between AGE, GENDER, tonal CONTEXT, and LEXICALITY
te(POINT, DURATION) +
s(POINT, ONSET, bs = "fs", k = 10, m = 1) +
s(POINT, SYLLABLETYPE, bs = "fs", k = 10, m = 1) +
s(POINT, HEIGHT, bs = "fs", k = 10, m = 1) +
s(POINT, SPEAKER, bs = "fs", k = 10, m = 1) +
s(POINT, SPEAKER, by = LEXICALITY, bs = "fs", k = 10, m = 1) +
s(WORD, bs = "re")
```

Fig. 3 shows the estimated smooths for LR and DR in each tonal context for each age/gender group. Qualitatively, the estimated smooths for DR are very close to the those of the LR. To the extent that there are differences at all, the estimated smooths for DR are higher than those of the LR. Fig. 4 illustrates the estimated difference in F0 (semitones) between LR and DR, over the course of the rime for each tonal context for each cohort. When the difference smooth is above zero, the F0 of DR is higher than that of LR; when the difference smooth is below zero, the F0 of DR is lower than that of LR. Regions where the DR-LR difference are significantly different, as determined by the comparison of estimated marginal means using `emmeans` (Lenth et al., 2021) with p-values corrected using the Tukey HSD method, are highlighted in red. As shown in Fig. 4, the difference smooths generally hover around zero or trend above it. DR and LR are most distinct in the T55 and T25 contexts among the older cohorts. Specifically, for both older men and women, DR and LR are distinct in the T25 context, albeit only at the initial two measurement points. In the T55 context, the older female participants show an DR/LR distinction at the final two measurement points while the older men show a consistent DR/LR difference from measurement point 4 to the end of the rime. Among the younger participants, a DR/LR difference is observed in the T23 context among the younger men at measurement points 2 and 3.

In this subsection, we examined the effects of preceding tonal context on the realization of LR and DR across age and

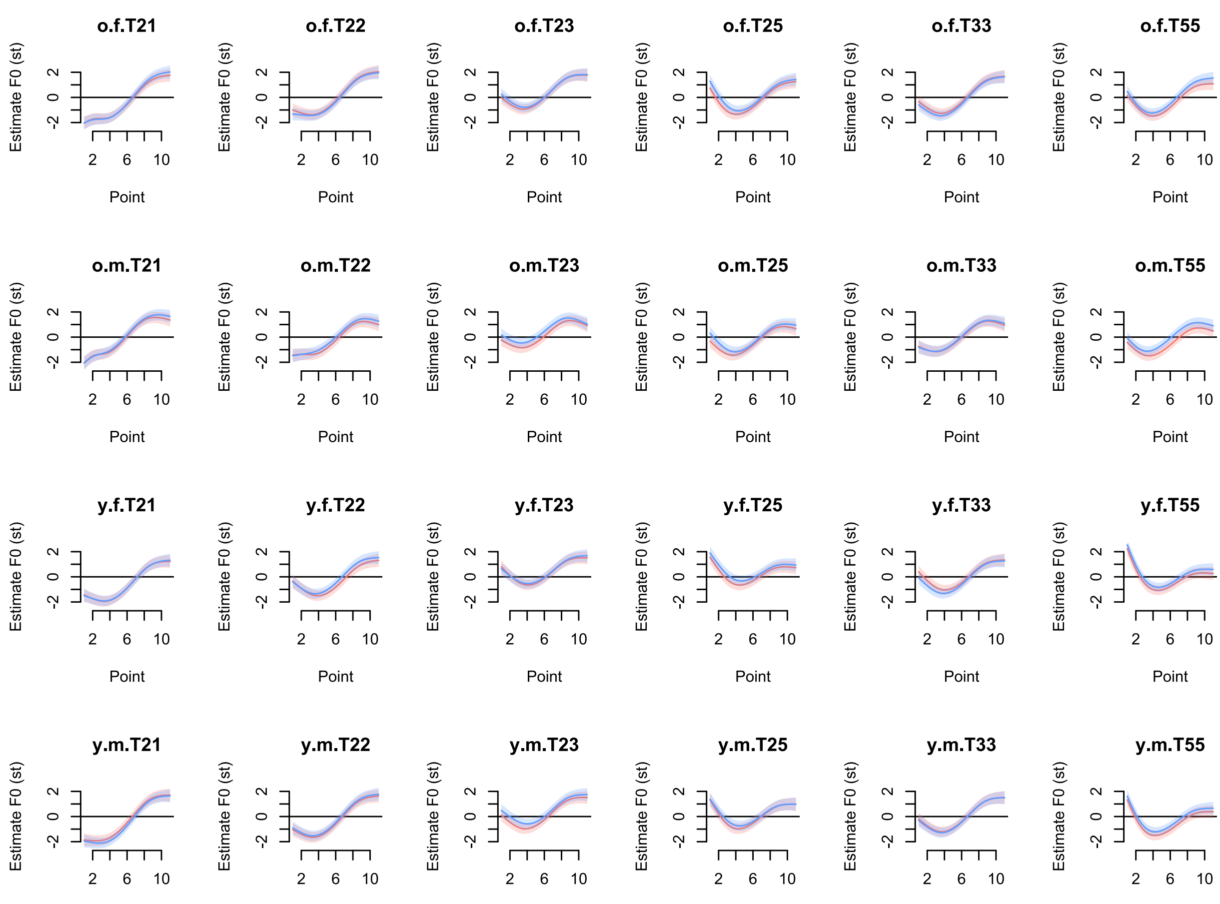

Fig. 3. Estimated smooths for LR (in red) and DR (in blue) for the older female (top row), older male (second from the top row), younger female (second from the bottom row) and younger male (bottom row) participants in Model 2. Shaded bands represent the pointwise 95%-confidence interval. o = older cohort; y = younger cohort; m = male; f = female. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

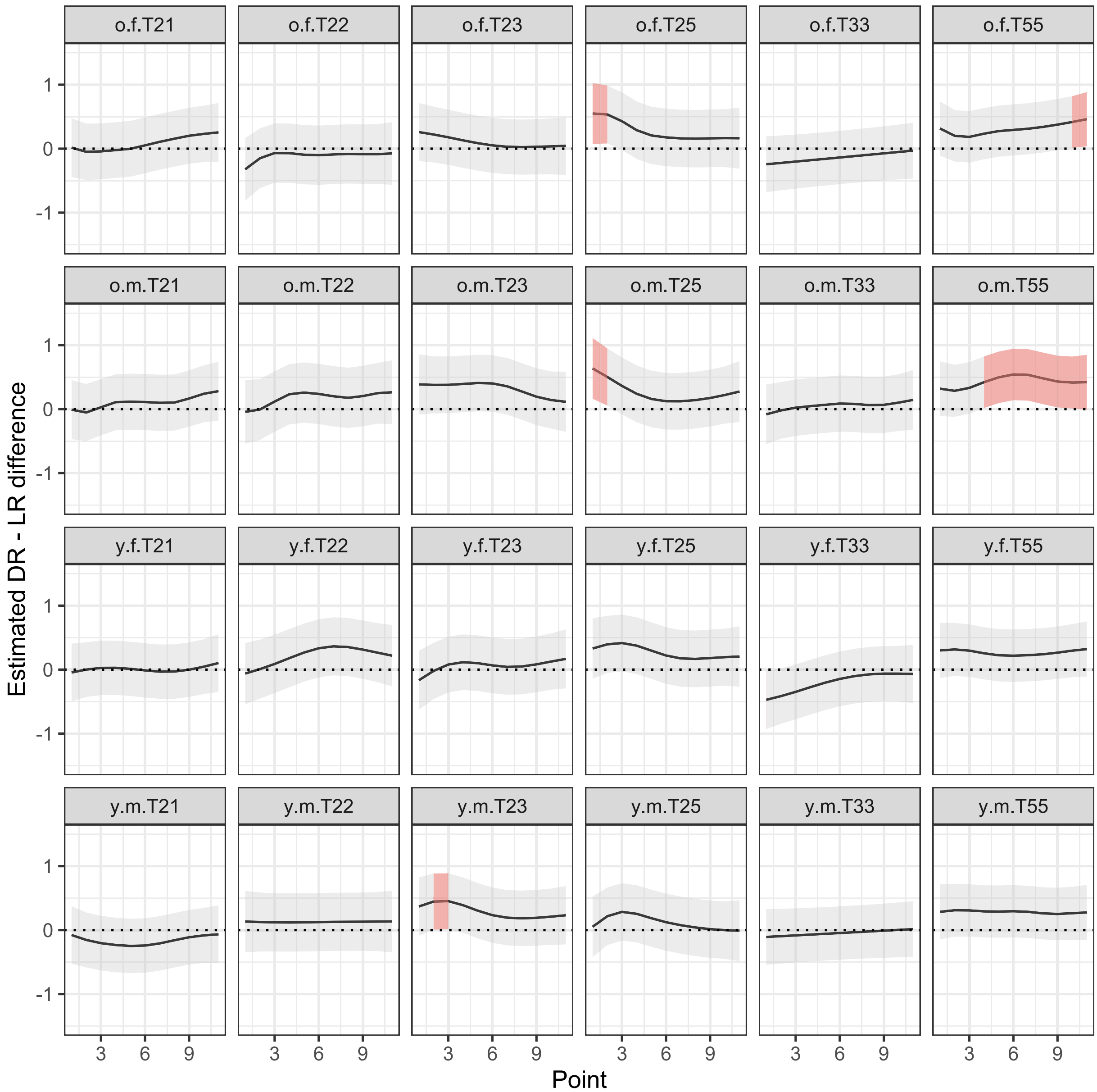

Fig. 4. Estimated differences between DR and LR in different tonal contexts by age and gender groups in Model 2. Gray bands represent the pointwise 95%-confidence interval; the red bands represent regions where the estimated difference between DR and LR has an adjusted p-value of < 0.05 using the Tukey HSD method obtained via the `contrast()` function in `emmeans` (Lenth et al., 2021). When the difference smooth is above zero, DR is higher than LR; when the difference smooth is below zero, DR is lower than LR. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

gender groups. Model 2 shows that the realization of LR and DR can be influenced differently by the preceding tonal context, resulting in different patterns of subphonemic distinctions between LR and DR. Specifically, DR has a higher F0 profile than LR most often when the preceding tone is T25 or T55. Of particular interest is the fact that the LR/DR distinction is most evident among the older participants, particularly the older men; LR and DR appeared to have largely merged among the younger participants. These findings are in keeping with the idea that the covert contrast between LR and DR is disappearing. To the extent that the apparent age effect reflects the path of diachronic change, the women appeared to be leading the change since the DR/LR distinction is less extensive among the old women compared to that of the older men and there is also no evidence of any DR/LR difference among the younger women.[6] Given that the LR/DR distinction is not consistently observed across tonal contexts, even among the older men, it seems clear that the seeds for the eventual merger between LR and DR are already present among the older generation of speakers, particularly in non-high F0 offset contexts. Another factor that might have skewed the LR and DR realizations toward complete neutralization is the influence of morphology on tonal realization. This is the focus of the next section.

4.1.3. Paradigmatic effects on DR realization

Beyond contextual influence, as noted in Section 2, previous studies also reported tonal realization differences related

[6] The observed apparent time effect is not likely to be the result of age-grading since there is no obvious reason for the DR/LR distinction to become more evident as one ages, especially in the case of the younger women where the DR/LR merger appears to be complete.

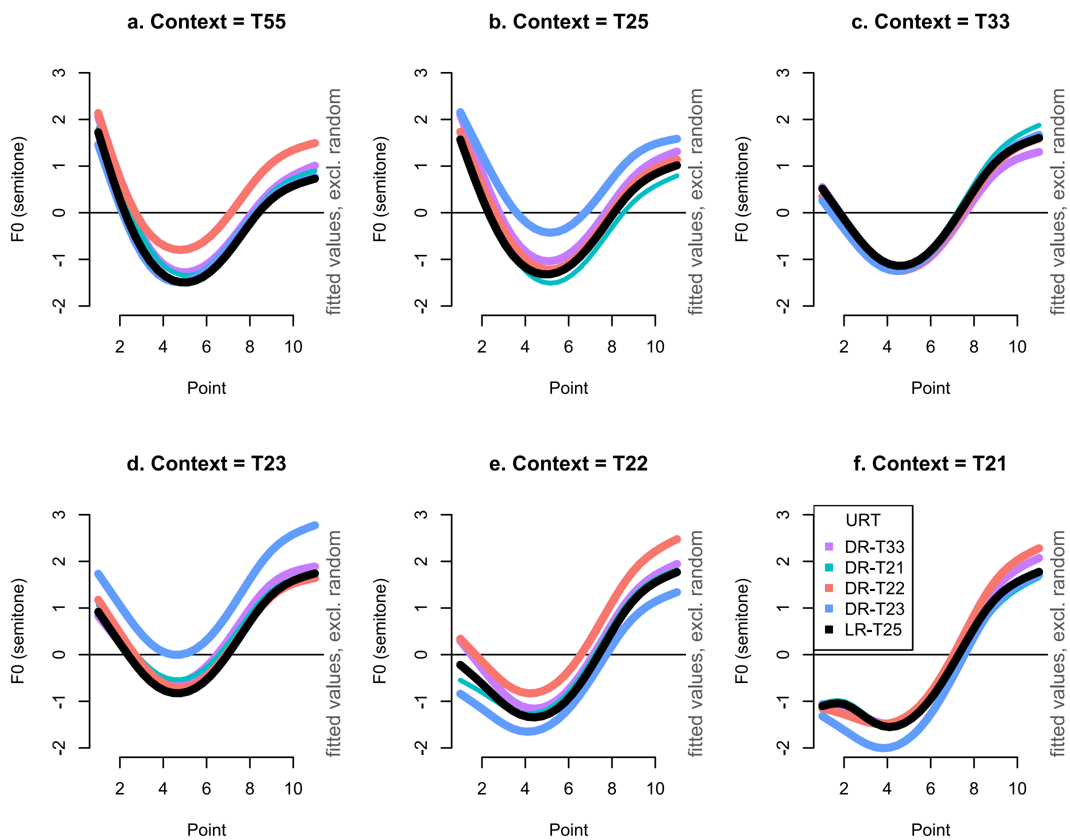

Fig. 5. Estimated smooths for each URT level across different tonal contexts in Model 3. LR is in black, DR-T21 in green, DR-T22 in red, DR-T23 in blue and DR-T33 in purple. Shaded bands represent the pointwise 95%-confidence interval. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to the tonal specification of the paradigmatic neighbor of the derived tone (UR-Tone or URT henceforth).[7] For example, Yu (2007a) found DRT33 to have a higher F0 profile than DRT22, mirroring the higher F0 profile of a lexical T33 tone relative to that of a lexical T22 tone. Thus, to the extent that paradigmatic forces can influence phonetic realization (e.g., Steriade, 2000; Yu, 2011; Braver, 2019), all else being equal, we expect the F0 profile of the DR tones to mirror the general F0 profiles of their paradigmatic neighbors. That is, the tonal contour of DRT33 is predicted to be highest among the DR tones, and the F0 profile of DRT22 and DRT21 should be the lowest. The F0 profile of DRT23 should be intermediate.

- **Phonetic analogy hypothesis I:** If the phonetic realization of a paradigmatic neighbor can influence the phonetic realization of a derived tone, then DRT33 > DRT23 > DRT22 > DRT21, where $x > y$ refers to x having a higher F0 profile than y.

To examine the paradigmatic effect on the maintenance of the LR/DR distinction, a third model (Model 3) was built where F0 was modeled with a GAMM that included a parametric term for the interaction of UR-Tones (URT with five levels: DRT21, DRT22, DRT23, DRT33, LRT25) and tonal CONTEXT (CONTEXTURT) and a difference smooth term over POINT with a by-LEXICALITY/CONTEXT interaction variable. The CONTEXTURT factor interaction term was treatment-coded; the reference level for the CONTEXTURT term is CONTEXT= T21 and URT = DRT21. By-speaker differences were captured using factor smooths for each speaker and for each level of URT. All other control variables and random smooth specifications are similar to Model 1.[8]

Model 3:

```
F0 ~ AGE + GENDER +
CONTEXTURT + #a parametric factor for the interaction
between tonal CONTEXT and URT
s(POINT, by = GENDER, k = 10) +
s(POINT, by = AGE, k = 10) +
s(POINT, by = CONTEXTURT, k = 10) +#separate smooths for
each combination of the interaction between tonal
CONTEXT and URT
te(POINT, DURATION) +
s(POINT, ONSET, bs = "fs", k = 10, m = 1) +
s(POINT, SYLLABLETYPE, bs = "fs", k = 10, m = 1) +
s(POINT, HEIGHT, bs = "fs", k = 10, m = 1) +
s(POINT, SPEAKER, bs = "fs", k = 10, m = 1) +
s(POINT, SPEAKER, by = URT, bs = "fs", k = 10, m = 1) +
s(WORD, bs = "re")
```

Fig. 5 illustrates the realization of the five UR-Tones in different tonal contexts. Qualitatively, we see that the DR F0 contours are most different from LR (in black) when the UR-Tone

[7] While we chose to use "UR" as a shorthand, it should be noted that we shall remain agnostic about whether *pinjam* and its morphological neighbor (e.g., *nin25** vs. *nin21* 'year') are morphologically derived from the same underlying representation (e.g., /nin21/) or if the two forms are merely paradigmatically related to each other. We will return to this issue in Section 5. URT is meant as a convenient way to refer to both LR and all the DR with different morphological neighbors.

[8] We did not examine the higher-order interaction between URT and CONTEXT with the social variables since we have no *a priori* hypotheses in terms of why the co-activation of neighboring tone would differ along sociolinguistic dimensions.

is T23 (in blue). The F0 contour of DRT23 is higher than that of LR when the preceding tone is rising (i.e., when the tonal context is T25 (Panel b) or T23 (Panel d)), while it is lower when the preceding tone is low (i.e. when the tonal context is T22 (Panel e) or T21 (Panel f)). The DRT22 (in red) is higher than LR when the preceding tone is T55 (Panel a) or T22 (Panel e). These observations are confirmed by the difference smooths shown in Fig. 6.

The fact that DR and LR are most distinct when the tonal neighbor is T22 or T23 suggests that the identity of the paradigmatic tonal neighbor can influence the realization of a DR tone and its contrast with LR. However, the directionality of the effect is not as straightforward as suggested by previous research. Fig. 7 shows the difference smooths for the comparisons between two UR-Tones (i.e. DRT33 - DRT23, DRT33 - DRT22, DRT33 - DRT21, DRT23 - DRT22, DRT23 - DRT21, DRT22 - DRT21) in different tonal contexts.

Consistent with the phonetic analogy hypothesis, DR-T23 is significantly higher than DR-T22 and DR-T21 in both T23 and T25 contexts. DR-T22 is also higher than DR-T21 in T55, T22, and T21 contexts. However, contrary to the phonetic analogy hypothesis, DRT33 is generally not significantly different from the other UR-Tones. The only context where DRT33 is higher than another derived tone is in comparison with DRT23 following a T22 tone. Indeed, DRT33 is found to have a LOWER F0 profile than that of DRT22 in the T55 context and DRT33 in the T23 context. Also contrary to expectation is the fact that DRT23 is LOWER than DRT22 in T22 and T55 contexts. Taken together, the phonetic analogy hypothesis is only partially supported by the results of Model 3. We return to this set of puzzling findings in the Discussion in Section 5.

4.2. Duration analysis

As noted in Section 2, some earlier descriptions suggest that LR and DR also differ along the duration dimension. Consistent with this claim is the hypothesis that *pinjam* originated from the elision of a monosyllabic morpheme that carried a high tone or a high rising tone (e.g., Whitaker, 1956; Wong, 1982; Bauer and Benedict, 1997). DR would presumably be longer than LR if speakers employ a compensatory strategy to fill the timing slot left by the elision of the neighboring syllable (e.g., Chao, 1947; Liu, 2016). Another potential source of a duration difference between LR and DR might come from

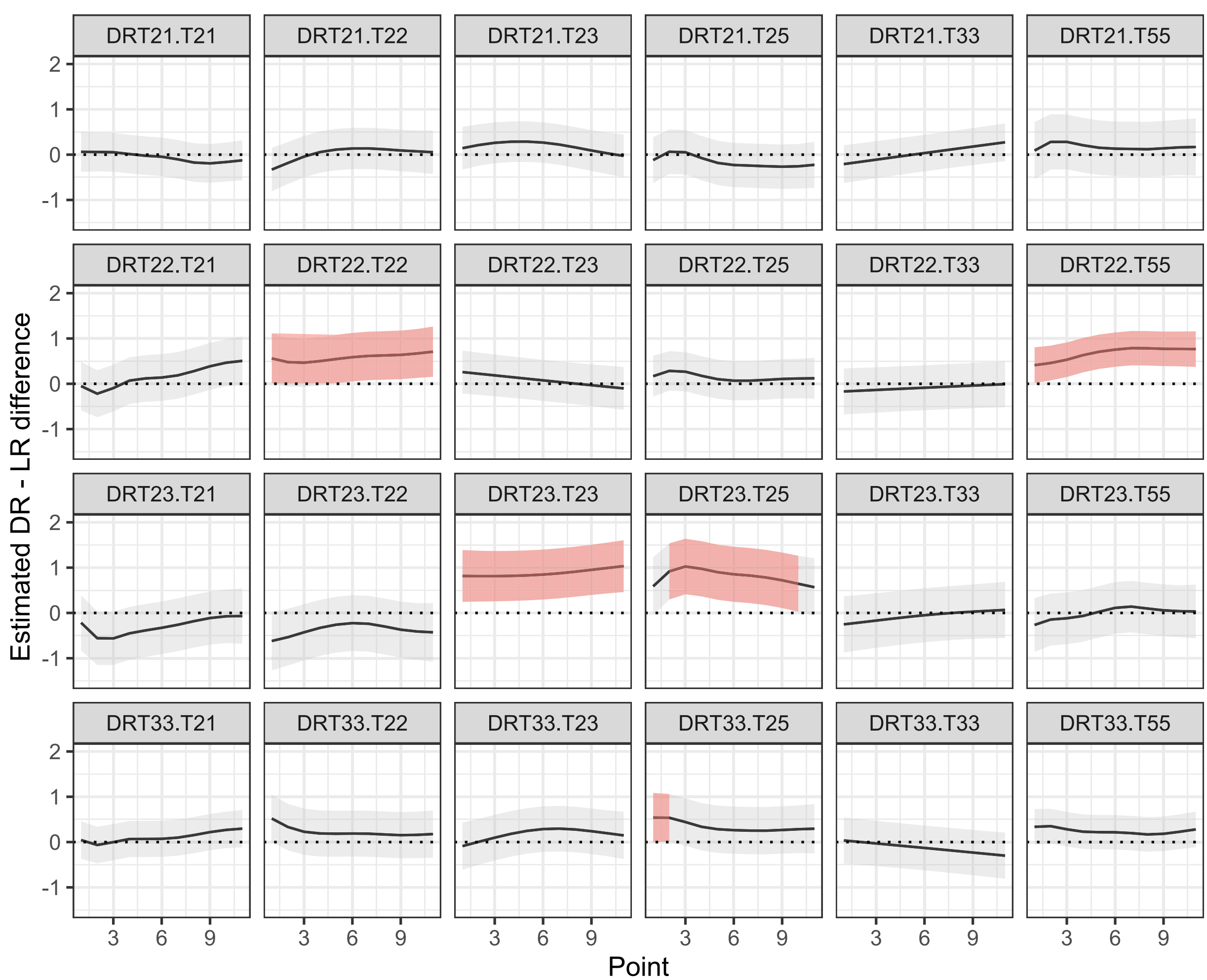

Fig. 6. Estimated differences between UR-Tones and LR (UR-Tone can be DRT21, DRT22, DRT23, or DRT33) in Model 3. The panels are arranged such that each row focuses on a different URT, while each column focuses on a particular tonal context. Thus, the panel, DRT21.T55, focuses on the difference between DRT21 with LR in the T55 context; DRT23.T21 refers to the difference between DR-T23 and LR in the T21 context. Shaded bands represent the pointwise 95%-confidence interval; the red bands represent regions where the estimated difference between DR and LR has an adjusted p-value of < 0.05 using the Tukey HSD method obtained via the `contrast()` function in `emmeans` (Lenth et al., 2021). A difference smooth above zero suggests that DR is higher than LR; when the difference smooth is below zero, DR is lower than LR.

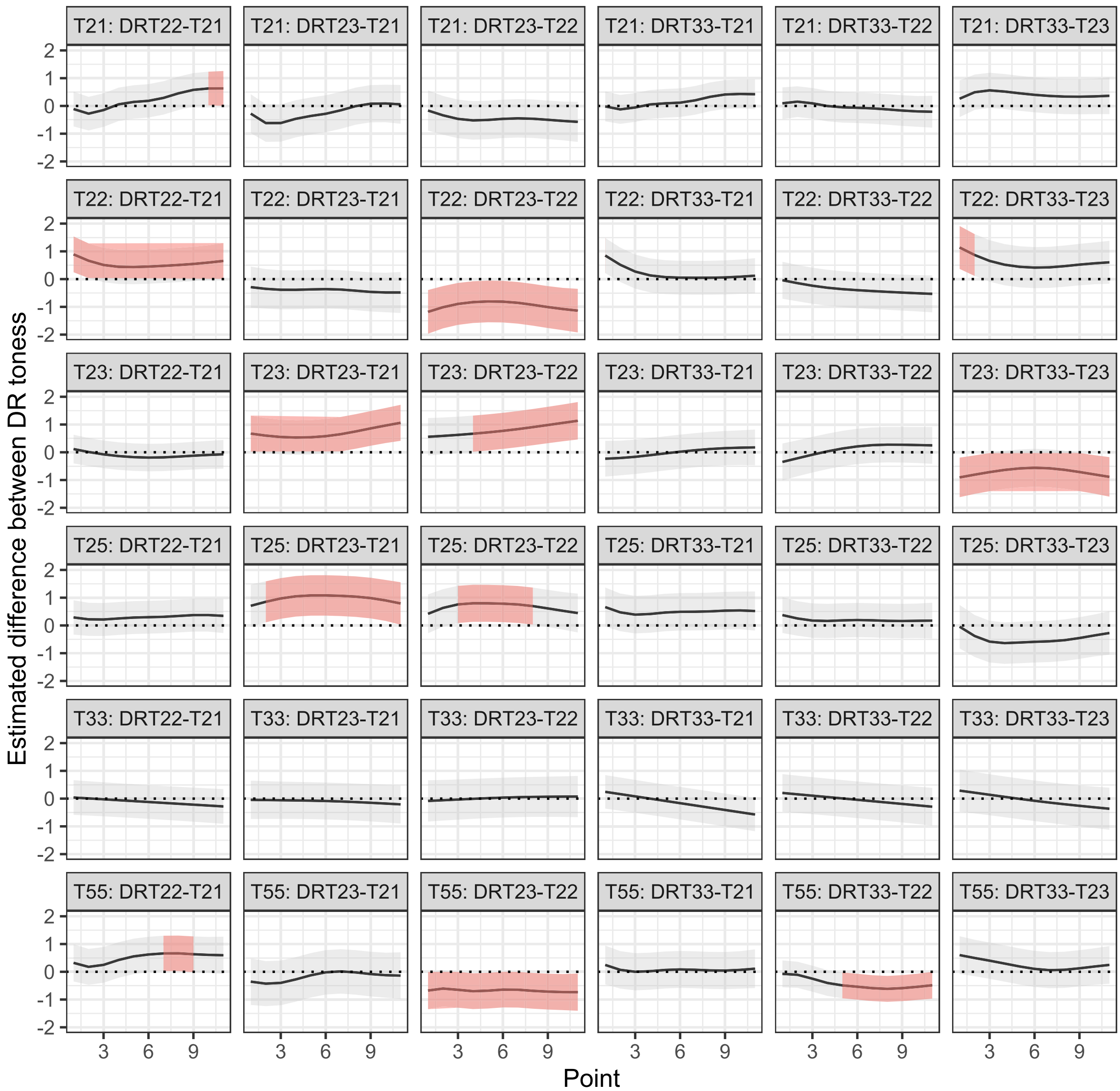

Fig. 7. Difference smooths between DRs with different paradigmatic tonal neighbors in different tonal contexts in Model 3. The panels are arranged such that each row focuses on a different tonal context, while each column focuses on a particular URT contrast. For example, the panel, T55:DRT22-T21, focuses on the difference between DRT22 and DRT21 in the T55 context; T21:DRT23-T21 focuses on the difference between DRT23 and DRT21 in the T21 context. Shaded bands represent the pointwise 95%-confidence interval; the red bands represent regions where the estimated difference between DR and LR has an adjusted p-value of < 0.05 using the Tukey HSD method obtained via the `contrast()` function in `emmeans` (Lenth et al., 2021).

the type of paradigmatic effects examined in Section 4.1.3. Not a lot is known about the durational properties of tones in Cantonese. To the extent that it has been studied, the results are variable across studies. Kong (1987) investigated the durational differences between tones in Cantonese among three native speakers of Cantonese producing the syllable /si/ in a carrier phrase and found, as a group, the following hierarchy of rime duration based on tone: T25 (272 ms) $>$ T33 (264 ms) $>$ T23 (260 ms) $>$ T22 (257 ms) $>$ T55 (236 ms) $>$ T21 (229 ms), although the exact tone-duration hierarchy differs across speakers. Recently, Yu (2023) reported that the effect of tone on duration in Cantonese is dependent on the tonal context; when the effects of tonal context are controlled for, the duration hierarchy (i.e., T22 $>$ T23 $>$T25, T21, T33 $>$ 55) differs from that reported in Kong (1987). Thus, assuming Kong's duration hierarchy, we may hypothesize LR to be longer than DR, assuming DR is influenced by the duration profiles of their nonhigh tone neighbors (recall that T55 syllables do not participate in *pinjam*). We also expect DRT33 to be longer than DR with corresponding lower UR-Tones. However, Yu's tone-duration hierarchy would predict DRT22 and DRT23 to be longer than LRT25.

- **Phonetic analogy hypothesis II**: The duration of DR should mirror the durational profile of its tonal neighbor.
  - Following Kong (1987), we expect LRT25 $>$ DRT33 $>$ DRT23 $>$ DRT22 $>$ DRT 21;
  - Following Yu (2023), we expect DRT22 $>$ DRT23 $>$ LRT25, DRT21, and DRT33.

To investigate these predictions, we investigated potential differences in rime duration across different UR-Tones by modeling rime duration using linear mixed-effects regression fitted in R, with the `lmer()` function from the LME4 package (Bates et al., 2015). UR-Tones (URT), which has five levels (LRT25, DRT21, DRT22, DRT23, and DRT33), was Helmert-coded such that each contrast compares each level of URT to the mean of the subsequent levels (i.e. LR vs. the average of all DR tones, DRT33 vs. the average of DRT23/DRT22/DRT21, DRT23 vs. the average of DRT22/DRT21, and DRT22 vs. DRT21). In addition to the URT of a rising tone, the regression model tested for the effects of vowel HEIGHT (high, mid, vs. low), SYLLABLETYPE structure (open vs. closed) and ONSET type (plain stop, aspirated stop, fricative, vs. sonorant), all variables that are known to affect syllable duration cross-linguistically (Gordon, 2006). We also tested for the effect of speaker AGE and GENDER. All categorical variables other than URT were sum-coded. The model also included by-subject random intercepts to allow for subject-specific variation, as well as by-subject random slopes for URT, HEIGHT, ONSET and SYLLABLE to allow for by-subject variability in their effects on the rime duration. The model also included by-word random intercepts to account for word-specific duration variability. In the end, neither the inclusion of URT nor GENDER significantly improve model-likelihood and therefore not included in the final analysis. The model formula in `lme4` style for rime duration is DURATION $\sim$ AGE + HEIGHT + ONSET + SYLLABLE + (1 + HEIGHT + ONSET + SYLLABLE|SPEAKER) + (1|WORD). The residuals of the initial fit were examined and found to deviate strongly from normality. As a result, residuals that were more than 2.5 standard deviations from the mean were trimmed, which amounted to no more than 2.5% of the data, and the model was refitted to the trimmed data set. The new model had a residual distribution much closer to normality, and it is the refitted models that are reported below. A summary of the duration model is shown in Table 2.

The duration model shows the expected effects of onset type (i.e. rime duration is shorter when the onset is a fricative (β =-6.903, t-value = -2,293, p = 0.022) and longest when the onset is a sonorant (β =9.005, t-value = 3.067, p = 0.002)), syllable type (i.e. rime duration is longer when it ends in a nasal rather than a vowel (β =25.232, t-value = 14.687, $p < 0.001$)), and vowel height (i.e. rime duration is shortest when the vowel is high (β =-12.974, t-value = -4.638, $p < 0.001$) and longest when the vowel is low (β =9.098, t-value = 3.605, $p < 0.001$)). The model also shows that rime duration is longer for the older speakers than the younger speakers (β =14.937, t-value = 3.310, p = 0.001). Yet, the fact that our duration analysis found no evidence that LR and DR are distinct counters the findings reported in Chao (1947) and Liu (2016). The discrepancy might be attributed to the fact that previous observations were drawn from anecdotal impressions, as in the case of Chao (1947), or phonetic comparisons that did not take into account the influence of potential confounding factors, such as syllable type, vowel height, and onset type, as well as by-speaker and by-word random effects, as in the case of Liu (2016).

Table 2
Regression model for rime duration. *** = p < 0.001; ** =p < 0.01; * = p < 0.05. The p-values were obtained using normal approximation, which assumes that the *t* distribution converges to the *z* distribution as degrees of freedom increase (see Mirman, 2014, for details).

	Coef	SE	t value	p value
(Intercept)	218.182	5.356	40.732	< 0.001***
AGE	14.937	4.512	3.310	0.001**
ONSET-fricative	−6.903	3.010	−2.293	0.022*
ONSET-sonorant	9.005	2.936	3.067	0.002**
ONSET-plain	8.263	2.495	3.312	0.001**
SYLLABLE TYPE	25.232	1.718	14.687	< 0.001***
HEIGHT-High	−12.974	2.797	−4.638	< 0.001***
HEIGHT-Low	9.098	2.524	3.605	< 0.001***

5. Discussion and conclusion

This study investigated the current state of affairs of a case of suspended tonal contrast, *pinjam*, in Hong Kong Cantonese along both F0 and duration dimensions. Our findings suggest that this subphonemic phonetic difference continues to be attested among the older speakers, albeit in limited tonal contexts. These findings lend support to Labov's idea that suspended contrast can be maintained for a "considerable period of time" (Labov, 1987, p. 319). That is, despite the fact that the subtle difference between LR and DR has been noted by early authors since as early as 1939 (O'Melia, 1939), evidence for this subphonemic difference between LR and DR remains present today. To be sure, there are also signs that the LR/DR distinction is disappearing. To begin with, as shown in Section 4.1.1, the LR/DR F0 difference is not observed among the younger speakers. While the LR/DR distinction remains among the older speakers, the distinction is preserved only in particular tonal (Section 4.1.2) and and morphological contexts (Section 4.1.3). There is also variability in terms of the locus and extent of the DR/LR distinction. Following a T25 tone, DR is only higher than LR at the tonal onset, along the line suggested by Chao (1947). However, following a T55 tone, DR is higher than LR toward the later half of the rime, more in line with the findings of Yu (2007b). Echoing earlier studies that suggest that women often lead sound changes in progress, we found no robust evidence for the LR/DR distinction among younger female participants. While the older female speakers still maintain an F0 difference between LR and DR, the extent of the difference is localized to the edges of the target tones. While the older men show similar context-sensitivity, their LR/DR distinction spans almost the entire rime in the T55 context. The age and gender distribution of the disappearance of the LR/DR contrast suggests that the merger between LR and DR is progressing toward completion. What factors might be leading to the disappearance of the LR/DR distinction? We explored two potential factors in this study, namely, the preceding tonal context and the nature of the morphological neighbor of a derived tone.

Both LR and DR are greatly affected by the nature of the preceding tone. The F0 analysis showed that the higher the preceding tonal offset, the lower the peak of the target rising tone. However, the preceding tonal context can influence LR and DR independently. Specifically, for the speakers where the LR/DR distinction is still maintained (i.e. the older cohort), high or rising tonal contexts facilitate the preservation of the LR/DR distinction. That is, when the preceding tone is T55, or T25, DR is more likely to be higher than LR. This situation is reminiscent of what Warren (2006) found in his study of the NEAR/SQUARE merger in New Zealand English. The old speakers in that study showed incomplete merger that is context-dependent. The distinction was more likely to be maintained in non-coronal contexts than in coronal ones. Warren argued that the raised tongue tip gesture of a preceding coronal consonant would likely raise the position of the SQUARE vowel, causing it to be phonetically and perceptually more similar to the NEAR vowel in coronal contexts than in non-coronal contexts. In the present case, assuming that DR was historically higher than LR and the direction of merger is toward LR, the neutralization is likely to have initiated in contexts where the LR/DR distinction is most endangered. Our analysis suggests that tonal coarticulation from a preceding tone with a low or mid tonal offset (i.e., T21, T22, T23, or T33) might have lowered DR to such an extent that it approximated the tonal profile of LR. The high and rising tonal offset contexts (i.e., T25 and T55), on the other hand, are better at preserving the higher DR F0 profile than the other tonal contexts. To be sure, this explanation assumes that DR is merging with LR and not the other way around. That is, might LR be raising to merge with DR? This question is particularly vexing as the tonal coarticulation explanation should apply to DR and LR equally *a priori*. That is, to the extent that a preceding lower tonal offset lowers DR, presumably the same coarticulatory forces should lower LR as well. What factors could lead to the stronger contextual effect on DR than LR? The answer might reside in the inherent variability of DR relative to LR. We explore this idea further below.

An important source of variability of DR comes from the fact that it is morphologically complex. Our investigation echoed earlier findings demonstrating that the realization of DR is sensitive to the tone of its morphologically-related tonal counterpart. Crucially, the morphological paradigmatic effect is mediated by tonal context. For example, we found that DRT23 is higher than LR when the preceding tone is a rising tone (T23 or T25). DRT22 is also higher than LR when the preceding tone is T55 or T22. Yu (2007a) argued that morphological effects on DR realization can be understood as the result of tonal phonetic analogy. That is, based on the concept of spreading activation, particularly from the perspective of exemplar-based models of lexical representation where morphologically related forms are expected to influence the processing and production of each other (e.g., Pierrehumbert, 2001; Pierrehumbert, 2002; Ernestus and Baayen, 2003; Ernestus and Baayen, 2006), the production of DR is expected to partially activate the exemplars of its paradigmatic neighbor. As such, the tonal target of its paradigmatic neighbor could in turn exert an influence on the production of DR. Yet, the lexical co-activation explanation of the subphonemic difference between LR and DR is complicated by the fact that the morphological neighbor effect is only observed in DRT22 and DRT23 and only in certain tonal contexts. This state of affairs might reflect a change in the understanding of the morphological relationship between DR and its paradigmatic neighbor. After all, the morphosemantics of the LR/DR distinction is complex, which prevents a transparent morphological mapping between form and meaning. This lack of morphological transparency might have endangered the LR/DR distinction as some of the DR exemplars might have been (mis)classified as LR by some speakers. Thus, for example, some speakers might treat a historically DR word such as *daan25** in *zi25 daan25** 子彈 'bullet' as unrelated to the word *daan22* in *daan22 sing33* 彈性 'elasticity, flexibility'. The *pinjam* versions of some words, like 橙 *caang25**, which was derived from *caang21* historically, have become the normative pronunciation completely. Further investigation is needed to ascertain the morphological-relatedness between DR and its paradigmatic counterparts among Cantonese speakers.

LR is also inherently more stable than DR because it is not subject to the same morphological influence as DR. LR is also more frequently observed than DR due to the morphological restrictiveness of *pinjam*. These factors make LR functionally favored to be the attractor in the case of merger. Couched within an exemplar-based model of sound change (Bybee, 2001; Pierrehumbert, 2002), the merger between LR and DR can be interpreted as resulting from a gradual overlap between their respective exemplar distributions (Pierrehumbert, 2001; Yu, 2007b). To the extent that there is any ambiguity between LR and DR, DR is more likely to be misclassified as LR not only because the preceding tonal contexts and UR-Tone can obfuscate the contrastivity between LR and DR, but also because the overwhelming frequency of LR can bias any ambiguous rising tone toward LR, rather than DR (cf. Pierrehumbert, 2001). The stability of LR might also explain why DR appears to be more susceptible to the influence of tonal context than LR. For example, the tonal trajectories of DR-T23 are higher when following a high or rising tone, but lower when following a low tone.

Several aspects of our findings remain puzzling, however. To begin with, if DR were influenced by the tonal specification of its paradigmatic neighbor as hypothesized above, F0 differences between DRs with different UR-Tones should mirror the differences between their morphologically related counterparts. As alluded earlier Yu (2007a) found that DR-T22 has a lower F0 profile than DR-T33, mirroring the fact that T22 has a lower F0 contour than that of T33. Our data show that, with the exception of DR-T23 and DR-T22, the other DR tones do not differ from LR significantly. Related to this is the finding that DR-T23, and to a lesser extent, DR-T22, particularly favor the maintenance of the LR/DR distinction. An examination of the DR-T23 words in our stimuli do not reveal any particular formal or semantic pattern (i.e. 企 kei23 'to stand', 女 neoi23 'girl', 友 jau23 'friend', 兩 loeng23 'pair/two', 母 mou23 'mother/maternal'). Further investigation is needed to resolve these unusual patterns.

Before closing, it is worth acknowledging a potential limitation of the present study. As noted above, the Cantonese tone system is undergoing some tonal mergers-in-progress. For example, in a series of discriminant analysis of tones produced by 17 Cantonese speakers who exhibit tonal mergers, Mok et al. (2013) found that that T33 and T21 were more likely to be misclassified as T22 than the other way around. Fung and Lee (2019) examined the perception and production of Cantonese tones by 120 participants with ages ranging from 20 to 58 and found that a high percentage of their participants (46.7%) were unable to produce T33-T22 distinctively, and 22.5% and 15% were unable to differentiate T25-T23 and T21-T22 respectively in production. To the extent that the UR-Tone effect on DR is due to the paradigmatic influence on phonetic realization, the on-going tonal merger between

T33 and T22 among some speakers of Cantonese today and the variability in T33 realization among such tone-merged speakers might have reduced the influence of a T33 tonal neighbor has on DR since the T33 exemplars have more varied tonal profiles. There is no clear evidence to assume that the abovementioned tonal mergers in progress have significantly impacted our findings, however. For example, if our participants showed clear signs of mergers between T33 and T22, we might expect DR-T33 to pattern with DR-T22. We found no evidence of that. To be sure, it would indeed be valuable to investigate in the future if there is a relationship between the status of tonal merger within each speaker and the realization of morphological tone.

In sum, this study revisited a case of suspended contrast in Cantonese between the lexical high rising tone and the high rising tone derived through morphological tone change (i.e., *pinjam*). Our findings suggest that the neutralization between lexical and morphological tones has been in circulation for at least almost a century as the earliest description of this suspended contrast goes back to O'Melia (1939). While there are clear signs that the distinction is disappearing, particularly among women and the younger speakers, it has not completely gone away. Our investigation also reveals significant contextual and morphological influences on the maintenance of the lexical vs. derived distinction, both in terms of their F0 contours and their durational profiles. While the longevity of this case of suspended contrast might be supported in certain tonal and morphology contexts, these same factors, perhaps in concert with the inherent variable nature of derived tone realization and the on-going tonal mergers elsewhere in the Cantonese tonal system, might be obfuscating the contrast suspension enough to tilt the balance toward complete merger. Taken together, our results highlight the complex interaction social and linguistic variables on the maintenance and disappearance of a linguistic contrast.

Acknowledgments

This study is supported by the funding provided by the deparment of the corresponding author. The authors would like to thank the anonymous reviewers for their useful comments and suggestions which have helped improve the manuscript in many ways. Any remaining errors are our own.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.wocn.2025.101412.

References

Alderete, J., Chan, Q., & Tanaka, S. (2022). The morphology of Cantonese "changed tone": Extensions and limitations. *Gengo Kenkyu, 161*, 139–169.
Bao, Z.-M. (1999). *The structure of tone*. New York: Oxford University Press.
Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.
Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese Phonology. Trends in Linguistics: Studies and Monographs 102*. Berlin and New York: Mouton de Gruyter.
Bishop, J. (2007). Incomplete neutralization in Eastern Andalusian Spanish: Perceptual consequences of durational differences involved in s-aspiration. In *In Proceedings of the International Congress of Phonetic Sciences XVI* (pp. 1765–1768). ICPhS.
Blevins, J., & Wedel, A. (2009). Inhibited sound change: An evolutionary approach to lexical competition. *Diachronica, 26*, 143–183.
Boersma, P., & Weenink, D. (2017). *Praat. Software Package*.
Braver, A. (2014). Imperceptible incomplete neutralization: Production, non-identifiability, and non-discriminability in American English flapping. *Lingua, 152*, 24–44.
Braver, A. (2019). Modelling incomplete neutralisation with weighted phonetic constraints. *Phonology, 36*(1), 1–36.
Braver, A., & Kawahara, S. (2016). Incomplete neutralization in Japanese monomoraic lengthening. In *Proceedings of the 2014 Annual Meeting on Phonology*.
Burton, M. W., & Robblee, K. E. (1997). A phonetic analysis of voicing assimilation in Russian. *Journal of Phonetics, 25*, 97–114.
Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.
Chao, Y. R. (1947). *Cantonese Primer*. Cambridge, Massachusetts: Harvard University Press.
Chen, M. Y. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica, 22*, 129–159.
Chen, Y., & Yuan, J. (2007). A corpus study of the 3rd tone sandhi in Standard Chinese. *In Interspeech, 2007*, 2749–2752.
Cheung, K.-H. (1986). *The phonology of present-day Cantonese* PhD thesis. University College London.
Cohn, A. (1990). *Phonetic and phonological rules of nasalization* PhD thesis. UCLA.
Cohn, A. (1993). Nasalization in English: Phonology or phonetics. *Phonology, 10*(1), 43–81.
Di Paolo, M. (1988). Pronunciation and categorization in sound change. In K. Ferrara, B. Brown, K. Walters, & J. Baugh (Eds.), *Linguistic change and contact: Proceedings of the 16th Annual Conference on New Ways of Analysing Variation in Language* (pp. 84–92). University of Texas.
Dinnsen, D. A., & Charles-Luce, J. (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics, 12*, 49–60.
Dmitrieva, O., Jongman, A., & Sereno, J. (2010). Phonological neutralization by native and non-native speakers: The case of Russian final devoicing. *Journal of Phonetics, 38*(3), 483–492.
Du, N., & Durvasula, K. (2022). Phonetically incomplete neutralisation can be phonologically complete: Evidence from Huai'an Mandarin. *Phonology, 39*(4), 559–595.
Engemann, M., & Plag, I. (2021). Phonetic reduction and paradigm uniformity effects in spontaneous speech. *The Mental Lexicon, 16*(1), 165–198.
Ernestus, M., & Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language, 79*(1), 5–38.
Ernestus, M., & Baayen, R. H. (2006). The functionality of incomplete neutralization in Dutch: The case of past-tense formation. In L. M. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Laboratory Phonology 8* (pp. 27–49). Berlin: de Gruyter.
Fougeron, C., & Steriade, D. (1997). Does deletion of French schwa lead to neutralization of lexical distinctions? *In Proceedings of Eurospeech, 97, volume 2*, 943–946.
Fourakis, M., & Port, R. (1986). Stop epenthesis in English. *Journal of Phonetics, 14*, 197–221.
Fox, R. A., & Terbeek, D. (1977). Dental flaps, vowel duration and rule ordering in American English. *Journal of Phonetics, 5*, 27–34.
Fung, R. S. Y., & Lee, C. K. C. (2019). Tone mergers in Hong Kong Cantonese: An asymmetry of production and perception. *The Journal of the Acoustical Society of America Express Letters, 146*, EL424.
Gerfen, C. (2002). Andalusian codas. *Probus, 14*(2), 247–277.
Gerfen, C. and Hall, K.C. (2001). Coda aspiration and incomplete neutralisation in Eastern Andalusian Spanish. Unpublished manuscript.
Giannini, A., & Cinque, U. (1978). Phonetic status and phonemic function of the final devoiced stops in Polish. In *Seminario di studi dell'Europa orientale*.
Gordon, M. (2006). *Syllable Weight: Phonetics, Phonology, Typology. Routledge, New York*.
Gordon, M., & Heath, J. (1998). Sex, sound symbolism, and sociolinguistics. *Current Anthropology, 39*(4), 421–449.
Gu, W., & Lee, T. (2009). Effects of tone and emphaticfocus on F0 contours of Cantonese speech: A comparison with Standard Chinese. *Chinese Journal of Phonetics, 2*, 133 –-147.
Gubian, M., Cronenberg, J., & Harrington, J. (2023). Phonetic and phonological sound changes in an agent-based model. *Speech Communication, 147*, 93–115.
Harris, J. (1985). *Phonological variation and change: Studies in Hiberno-English*. Cambridge: Cambridge University Press.
Hashimoto, O.-K. Y. (1972). *Phonology of Cantonese: Studies in Yue Dialects 1*. Cambridge University Press.
Keating, P. A. (1990). The window model of coarticulation: articulatory evidence. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I* (pp. 451–470). Cambridge University Press.
Kharlamov, V. (2014). Incomplete neutralization of the voicing contrast in word-final obstruents in Russian: Phonological, lexical, and methodological influences. *Journal of Phonetics, 43*, 47–56.

Kiparsky, P. (1982). From cyclic phonology to lexical phonology. In H. van der Hulst & N. Smith (Eds.). *The structure of phonological representations* (volume 1, pp. 131–175). Dordrecht: Foris.

Kiparsky, P. (1985). Some consequences of lexical phonology. *Phonology, 2*, 85–138.

Kong, Q.-M. (1987). Influence of tones upon vowel duration in Cantonese. *Language and Speech, 30*(4), 387–399.

Labov, W. (1971). Methodology. In W. O. Dingwall (Ed.), *A survey of linguistic science* (pp. 412–497). University of Maryland Linguistics Program.

Labov, W. (1987). The overestimation of functionalism. In R. Dirven & V. Fried (Eds.), *Functionalism in linguistics* (pp. 311–332). Amsterdam: John Benjamins.

Lenth, R.V., Buerkner, P., Herve, M., Love, J., Riebl, H., and Singmann, H. (2021). emmeans, 1.6.1 edition.

Liu, T.-H. (2016). On the nature of changed tones in Cantonese. In A. C. O. Chin, B. C. Kwok, & K. Y. B. Tsou (Eds.), *Commemorative Essays for Professor Y.R. Chao - Father of Modern Chinese Linguistics* (pp. 107–126). Crane Publishing.

Milroy, J., & Harris, J. (1980). When is a merger not a merger? the meat/mate problem in a present-day english vernacular. *English World-Wide, 1*, 199–210.

Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R*. Boca Raton, Florida: Chapman and Hall/ CRC.

Mok, P., Zuo, D., & Wong, P. (2013). Production and perception of a sound change in progress: tone merging in Hong Kong Cantonese. *Language Variation and Change, 25*, 341–370.

Nunberg, G. (1980). A falsely reported merger in eighteenth-century english: A study in diachronic variation. In W. Labov (Ed.), *Locating language in time and space* (pp. 221–250). New York: Academic Press.

O'Melia, T. A. (1939). *First Year Cantonese*. Hong Kong: Maryknoll House.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In Bybee, J.L. and Hopper, P., editors, Frequency and the emergence of linguistic structure, pages 137–157. Benjamins, Amsterdam. (Downloadable from http://www.ling.nwu.edu/jbp).

Pierrehumbert, J. (2002). Word specific phonetics. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology VII* (pp. 101–139). Berlin: Mouton de Gruyter.

Pierrehumbert, J. B. (1990). Phonological and phonetic representation. *Journal of Phonetics, 18*, 375–394.

Port, R., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics, 17*, 257–282.

Port, R. F., & O'Dell, M. L. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics, 13*, 455–471.

Roettger, T. B., Winter, B., Grawunder, S., Kirby, J., & Grice, M. (2014). Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics, 43*(1), 11–25.

Seyfarth, S., Garellek, M., Gillingham, G., Ackerman, F., & Malouf, R. (2018). Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience, 33*(1), 32–49.

Siddins, J., & Harrington, J. (2015). *Does vowel intrinsic f0 affect lexical tone? In International Congress of Phonetic Sciences, Glasgow, August 2015*. Glasgow, UK: University of Glasgow.

Steriade, D. (2000). Paradigm uniformity and the phonetics-phonology interface. In M. Broe & J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp. 313–335). Cambridge, UK: Cambridge University Press.

Stevens, M., Harrington, J., & Schiel, F. (2019). Associating the origin and spread of sound change using agent-based modelling applied to /s/-retraction in English. Glossa: A Journal of General. *Linguistics, 4*(1), 8. https://doi.org/10.5334/gjgl.620.

Sóskuthy, M. (2021). Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. *Journal of Phonetics, 84*, 101017.

Tieszen, B. (1997). *Final stop devoicing in Polish: an acoustic and historical account for incomplete neutralization* PhD thesis. Madison: University of Wisconsin.

To, C. K. S., McLeod, S., & Cheung, P. S. P. (2015). Phonetic variations and sound changes in Hong Kong Cantonese: Diachronic review, synchronic study and implications for speech sound assessment. *Clinical Linguistics & Phonetics, 29*(5), 333–0353.

Trudgill, P. (1974). *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Simonet, M., Rohena-Madrazo, M., and Paz, M. (2008). Preliminary evidence for incomplete neutralization of coda liquids in Puerto Rican Spanish. In Colantoni, L. and Steele, J., editors, Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology, pages 72–86. Cascadilla Proceedings Project. www.lingref.com, document 1715.

van Rij, J., Wieling, M., Baayen, R.H., and van Rijn, H. (2022). itsadug: Interpreting time teries and autocorrelated data using GAMMs, 2.4.1 edition.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics, 32*, 251–276.

Warren, P. (2006). Word recognition and sound merger. In J. Luchjenbroers (Ed.), *Cognitive linguistic investigations across languages, fields, and philosophical boundaries* (pp. 169–186). Amsterdam: John Benjamins.

Wedel, A. (2012). Lexical contrast maintenance and the organization of sublexical contrast systems. *Language and Cognition, 4*, 319–355.

Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition, 128*, 179–186.

Whitaker, K. P. K. (1956). A study on the modified tones in spoken Cantonese. *Asia Major, 6*, 184–207.

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics, 70*, 86–116.

Wong, M. K.-S. (1982). *Tone change in Cantonese. PhD thesis*. University of Illinois at Urbana-Champaign.

Wong, Y. W. (2006). Contextual tonal variations and pitch targets in Cantonese. In *In Proceedings of the 3rd International Conference on Speech Prosody* (pp. 317–320). Dresden, Germany: Speech Prosody.

Wood, S. N. (2017). Generalized additive models: An introduction with R. *Chapman and Hall/CRC*.

Xu, Y. (2013). ProsodyPro: a tool for large-scale systematic prosody analysis. In Bigi, B. and Hirst, D., editors, Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), pages 7–10. Laboratoire Parole et Langage, Aix-en-Provence, France.

Yao, Y., & Chang, C. B. (2016). On the cognitive basis of contact-induced sound change: Vowel merger reversal in Shanghainese. *Language, 92*(2), 433–467.

Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

Yu, A. C. L. (2004). Explaining final obstruent voicing in Lezgian: phonetics and history. *Language, 80*, 73–97.

Yu, A. C. L. (2007a). Tonal phonetic analogy. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the International Congress of the Phonetic Sciences XVI* (pp. 1749–1752).

Yu, A. C. L. (2007b). Understanding near mergers: The case of morphological tone in Cantonese. *Phonology, 24*(1), 187–214.

Yu, A. C. L. (2011). Contrast reduction. In J. A. Goldsmith, J. Riggle, & A. C. L. Yu (Eds.), *Handbook of Phonological Theory* (2nd Edition, pp. 292–318). Wiley-Blackwell.

Yu, A. C. L. (2023). Tonal context influences tone-duration interaction: Evidence from Cantonese. *Journal of Acoustical Society of America Express Letters, 3*, 035203.