

Article

Phonation Variation as a Function of Checked Syllables and Prosodic Boundaries

Xin Gao ^{1,2}  and Jianjing Kuang ^{1,*}

¹ Department of Linguistics, University of Pennsylvania, Philadelphia, PA 19104, USA; kauhsin@sas.upenn.edu

² Department of Chinese Language and Literature, Fudan University, Shanghai 200433, China

* Correspondence: kuangj@sas.upenn.edu

Abstract: The phonation variation in Shanghainese is influenced by both phonemic phonation contrast and global prosodic context. This study investigated the phonetic realization of checked and unchecked syllables at four different prosodic positions (sandhi-medial, sandhi-final, phrase-final, and IP-final). By analyzing both acoustic and articulatory voice measures, we achieved a better understanding of the nature of checkedness contrast and prosodic boundaries: (1) Different phonetic correlates are associated with the two laryngeal functions: The checkedness contrast is mostly distinguished by the relative degree of glottal constriction, but the prosodic boundaries are mostly associated with periodicity and noise measures. (2) The checkedness contrast is well maintained in all prosodic contexts, suggesting that the controls for the local checkedness contrast are rather independent of global prosody.

Keywords: voice quality; phonation contrast; checked syllable; Shanghainese; prosodic effects; global prosody



Citation: Gao, Xin, and Jianjing Kuang. 2022. Phonation Variation as a Function of Checked Syllables and Prosodic Boundaries. *Languages* 7: 171. <https://doi.org/10.3390/languages7030171>

Academic Editors: Christina Esposito and Sameer Khan

Received: 16 June 2021

Accepted: 13 June 2022

Published: 5 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Language under Study: Shanghainese

Shanghainese is a variety of Chinese Wu spoken in the urban area of the city of Shanghai. As laid out in Table 1, the five-tone system of this language (Chen and Gussenhoven 2015; Xu et al. 1988) can be sorted into two contrastive dimensions. On the one hand, compared to the tones in the upper register (i.e., T1, T2, and T4), the tones in the lower register (i.e., T3 and T5) are produced with a lower f_0 range and a breathier phonation (Ren and Mattingly 1989; Tian and Kuang 2019). On the other hand, there is a contrast between checked and unchecked tones, categorizing T4 and T5 against T1, T2, and T3.

Table 1. Overview of Shanghainese tonal inventory. Tones are transcribed with Chao's tone numbers (Chao 1968) and are according to Xu et al. (1988); checked tones are marked with underscores.

	Unchecked [CV]		Checked [CV]
Upper-register	T1 (high-falling): 53	T2 (high-rising): 34	T4 (high): <u>55</u>
Lower-register		T3 (low-rising): 23	T5 (low): <u>12</u>

Checked syllables are traditionally transcribed as closed syllables with a coda glottal stop in Shanghainese (Xu et al. 1988). However, the phonetic realization of checked syllables in Shanghainese has not been well understood. Based on the qualitative inspection of spectrograms from limited examples, previous studies have suggested several different phonetic realizations. For example, it is proposed that Shanghainese checked syllables at

least involve a shorter duration and a coda glottal stop (Zhu et al. 2008); similar findings have been reported for Longyou Wu, a related Wu dialect to Shanghainese, where check syllables are realized as an ‘abrupt phonatory offset and short rhymes’ (Rose 2015). However, studies also show that the checkedness of Shanghainese checked syllables is not only reflected by a coda glottal stop but can also be realized as an irregular creak throughout the vowel portion (Shen 2010). In this paper, we use [CV] to denote the check syllables in contemporary Shanghainese. The reasons for choosing this transcription are discussed in Sections 3.1.5 and 4.1. Moreover, the traditional transcription (Xu et al. 1988) has suggested that checked syllables have a lower f_0 compared to the unchecked counterpart, but no instrumental studies have been conducted to validate this claim. Therefore, in order to better understand the phonetic nature of checked syllables, it is necessary to conduct more systematic acoustic and articulatory analyses for the checkedness contrast.

Moreover, based on our observation, coda glottal stop or coda creak does not always occur in checked syllables. One potential source for this variation is prosodic contexts. However, because most of the existing studies on the voice quality of the checked syllables of Shanghainese are based on the analysis of isolated syllables, and it is unclear whether the syllable creak is driven by prosodic effects or the phonemic contrast of checkedness, little is known about how phonetic features of checked syllables are realized in connected speech (Shen 2010; Zhu et al. 2008). To tease apart the checkedness effect from the prosodic effect, it is necessary to examine the phonetic realization of checked syllables at various prosodic boundaries.

In this paper, we investigate the phonemic effect, prosodic effect, and their interaction on phonation variation by examining the checked syllables in contemporary Shanghainese in order to gain a better understanding of the nature of checkedness contrast and prosodic boundaries. In Section 1, we review the language background as well as the phonation variation associated with checked codas and different prosodic boundaries; in Section 2, we present the experimental methods and materials; in Section 3, we model the measurement results to analyze the phonemic effect and prosodic effect on various phonetic parameters; In Section 4, we discuss the phonetic nature of checked syllables in contemporary Shanghainese, the effect of prosodic effect on phonation variation, and the interaction of the phonemic effect and prosodic effect.

1.2. The Tone-Sandhi Pattern and the Prosodic Hierarchy in Shanghainese

Shanghainese has three levels of prosodic phrasing above the syllable level. The lowest level is the tone-sandhi domain level. This level is related to the ‘left-dominant’ tone-sandhi pattern of Shanghainese, which means that the tonal pattern of the entire tone-sandhi domain is determined by the leftmost syllable (Duanmu 1999; Xu et al. 1988; Yip 2002, and others). For example, as shown in (1), when the tone of the initial syllable of a disyllabic word is high-rising (34), the tonal pattern of the entire word is always high-rising (33–44), regardless of the underlying tones of the second syllables.

- (1) $s\gamma 34 + \epsilon in 53 \rightarrow s\gamma 33 \epsilon in 44$ ‘palm’
 $s\gamma 34 + \epsilon in 34 \rightarrow s\gamma 33 \epsilon in 44$ ‘souvenir’
 $s\gamma 34 + \epsilon in 23 \rightarrow s\gamma 33 \epsilon in 44$ ‘hand-shape’

Non-initial syllables in tone-sandhi domains are also subject to various phonetic reductions (Chen 2008; Kuang et al. 2018; Ling and Liang 2016; Tian and Kuang 2020). The sandhi domain is considered as an important prosodic domain in Shanghainese because it is associated with word formation and usually marks the prosodic word boundary but sometimes can be smaller than a word (Roberts 2020; Selkirk and Shen 1990; Yip 2002; Zee and Maddieson 1979). For example, in a trisyllabic personal name, it is common for the first two syllables to form a tone-sandhi domain while leaving the third syllable itself to form a monosyllabic tone-sandhi domain (see Section 2.1 for detailed discussion). An example of such a trisyllabic name is given in (2); the brackets indicate a tone-sandhi domain.

$$(2) \quad li23 + \text{ci}34 + \text{min}23 \rightarrow (li22\text{ci}44)\text{min}23$$

The tone-sandhi domain level above is the ‘Major Phrase (phrase)’ level, which can be identified by phrase-final lengthening and pitch reset but a lack of audible pause (Roberts 2020; Selkirk and Shen 1990). For example, trisyllabic personal names can often form a phrase domain; the entire trisyllabic name ‘li22ci44min23’ in example (2) forms a phrase domain.

Finally, the highest level is the ‘Intonational Phrase (IP)’ domain, which is marked by a final lengthening after the boundary of an ‘actual silence’ (Chen 2008; Roberts 2020; Selkirk and Shen 1990). Usually, the boundary of an IP domain in Shanghainese is the same as the boundary of a sentence.

In the current study, we examine how the phonetic realization of checked syllables is affected by the three levels of prosodic phrasing. For convenience, in the following discussion, from the lowest to the highest, we will refer to the three levels of prosodic phrasing as ‘sandhi domain’, ‘phrase domain’, and ‘IP domain’.

1.3. Phonation Variation Related to Checked Coda

Coda checkedness indicates the presence of a glottal stop or other glottalized stop at the end of the syllable. It has been attested in a wide range of languages, such as in Arabic (Kasim 2019), Assam Sora and Mizo (Kalita et al. 2017), Athabaskan languages (Kingston 2005), Deg Xinag (Hargus 2016), Itunyoso Trique (DiCano 2012), Maltese (Mitterer et al. 2019), Northern Vietnamese (Brunelle and Kirby 2016), Taiwan Min (Pan 2017), Western Muskogean languages (Ulrich 1993), Yucatec Maya (Frazier 2013), etc.

The phonatory attributes of the checked codas may vary from language to language. In some languages, the checked coda is realized as a full glottal stop, characterized by a complete closure of vocal folds and can be identified by a short period of silence during the glottal closure (Davidson 2020; Esling et al. 2005; Garellek et al. 2021; Ladefoged 1971; Ladefoged and Maddieson 1996, and others).

However, cross-linguistically, it has been well documented that phonemic glottal stops are rarely realized as complete glottal closure, and their occurrence is conditioned by individual differences, contextual differences, and even other contingent factors (Boroff 2007; DiCano 2012; Pan 2017; Ulrich 1993). More often, glottal stops are realized as incomplete glottal closure and can substantially influence the voice quality of the adjacent vowels and other vocalic segments (Esposito and Khan 2020; Garellek and Esposito 2021; Ladefoged and Maddieson 1996).

Due to the variation of the articulatory configuration in the larynx (and pharynx), languages vary in the extent of glottalization and types of glottalized phonation. In many languages, glottal stops are often realized as a creaky voice, or irregular glottal pulses, on the adjacent vowels. For example, in San Lucas Quiavini Zapotec, the glottal stop involved in the checked syllables often surfaces as ‘a period of strong glottalization’ in the middle of the vowel (Chávez-Peón 2008). In Itunyoso Trique, the coda glottal stop is often realized as a short portion of irregular voicing at the end of the vowel (DiCano 2012). Glottal stops in Mayan languages are also most often realized as creaky phonation on the adjacent vowels (Bennett 2016).

In addition to creaky voice, checked syllables can also be realized as other types of glottalized or laryngealized phonation such as tense voice or harsh voice. Similar to the glottalized/laryngealized voice, the tense voice is also characterized by a high glottal constriction; however, the f_0 of the tense voice is neither low nor irregular (Keating et al. 2015). For instance, the phonemic tense phonation in various Yi languages (e.g., Hani and Southern Yi) is the historical reflex of vowels in syllables that have original final stops, and the tense phonation can co-occur with a mid tone (Kuang 2013; Maddieson and Ladefoged 1985); in Daigela Wa, the tense phonation and the lax phonation differ significantly in the degree of glottal constriction, but not in f_0 (Wei 2018). Similar to creaky voice, tense voice is also produced with greater glottal constriction, but unlike creaky voice, tense voice is highly periodic and often high-pitched (Keating et al. 2015; Kuang and Keating 2014).

Apart from a tense voice, checked syllables can also be produced with a harsh voice in some languages (Garellek 2020; Rose 2015; Traill 1994). Harsh voice is a rough and noisy phonation that involves both strong laryngeal constriction and pharyngeal constriction (Garellek 2020; Gerratt and Kreiman 2001; Moisik 2012).

Since the phonetic realization of checked codas varies across languages and contexts (see Section 1.5), multidimensional phonetic features are needed to characterize the phonetic nature of checked syllables.

1.4. Phonetic Correlates of Checked Syllables

In this study, we are particularly interested in the phonation variation introduced by the checked coda. If Shanghainese checked syllables are realized as glottalized vowels, acoustically, we would expect the spectral slope, or the relative strength of the lower-frequency harmonics and higher-frequency harmonics, can reliably distinguish checked syllables from unchecked syllables. Checked vowels that are produced with greater glottal constriction should be associated with a flatter spectral slope or less prominent lower-frequency harmonics in the spectrum (Esposito 2010; Garellek and Keating 2011; Gordon and Ladefoged 2001; Keating et al. 2015; Kreiman and Gerratt 2012; Kreiman et al. 2012; Kuang and Keating 2014, among many others).

Periodicity and noise ratio are also important acoustic correlates for glottalization. Different types of laryngealized phonation have different periodicity profiles. Creaky voice is usually associated with less periodicity and lower harmonic-to-noise ratio, but tense voice is associated with higher periodicity and higher harmonic-to-noise ratio in addition to a higher pitch (Keating et al. 2015). Moreover, energy damping or energy dips are also observed in checked syllables in some languages (DiCanio 2012; Pan 2017).

Articulatory, the characteristics of glottal closure and the degree of relative glottal constriction can be non-invasively measured by an Electroglottograph (EGG). Contact Quotient (CQ), defined as the ratio of the duration of the contact phase to the period of the vibratory cycle (Rothenberg and Mahshie 1988), is the most important measure for EGG. This measure has been found to be a reliable indicator for phonation contrast in various languages, and greater CQ is correlated with greater glottal constriction (DiCanio 2009; Esposito and Khan 2012; Garellek 2020; Guion et al. 2004; Jiang et al. 2017; Kuang and Keating 2014; Li and Zhang 2020; Mazaudon and Michaud 2008; Tian and Kuang 2019, and so on).

Another important measure of EGG is Peak in Contact (PIC), also known as Derivative-EGG Closure Peak Amplitude (DECPA) (Esposito and Khan 2012; Keating et al. 2010; Kuang and Keating 2014; Michaud 2004). It is defined as the amplitude of the positive peak in the first derivative of the EGG signal. PIC is an important feature of dEGG, as the positive peak of dEGG marks the beginning of the contacting phase (Howard 1995). This measure reflects the manner of vocal fold contact, which is also an important aspect of glottal constriction. However, the specific articulatory implication of PIC is not completely clear. A popular proposal is that PIC is related to the abruptness of contact, and therefore greater PIC is related to more abrupt contact (Keating et al. 2010). Consistent with this proposal, higher PIC values are associated with the checked syllables in Northern Kam (Jiang et al. 2017). However, the opposite direction was also reported in several languages. For example, the breathier phonations (e.g., breathy or lax voice) in Gujarati, white Hmong (Esposito and Khan 2012) and Yi languages (Kuang and Keating 2014) were found to have higher PIC values instead. Although inconsistent directions were reported among the aforementioned languages, PIC has been found to be a reliable measure for distinguishing contrastive phonation in these languages (Esposito and Khan 2012; Keating et al. 2010; Kuang and Keating 2014; Michaud 2004). Therefore, in this study, we included both CQ and PIC to assess different aspects of glottal constriction.

Moreover, since f_0 affiliation is an important aspect of phonation types (e.g., vocal fry vs. tense), and phonation variation can be driven by pitch variation in tone production, as found in the case of Mandarin (Chai 2021; Kuang 2017), it is also necessary to examine f_0 values of checked and unchecked vowels and their correlations with phonation cues.

Lastly, vowel duration was measured in this study as well, as it has been reported to be a useful cue for the checkedness contrast in Shanghainese (Chen and Gussenhoven 2015; Xu et al. 1988; Zee and Maddieson 1979; Zhu et al. 2008, among many others).

1.5. Phonation Variation Related to Prosodic Boundaries

Our discussion has mainly focused on the phonation variation introduced by local phonemic contrasts in the previous sections. However, phonation also co-varies with the global laryngeal functions, such as vocal effort, prosodic boundaries, prominence, and so on (Epstein 2002; Garellek 2015; Klatt and Klatt 1990; Kuang 2018, and others). In particular, a large number of studies have shown that prosodic factors (i.e., prominence and boundary) are of significant influence on the occurrence of allophonic creak and other types of laryngealization.

On the one hand, prosodic prominence has a great impact on the relative degree of glottal constriction. Prosodically weak positions, such as unstressed syllables, are often (but not always) associated with an irregular creaky voice (Epstein 2002; Klatt and Klatt 1990; Kuang 2018). By contrast, stressed or prosodically prominent syllables are often associated with greater vocal effort or a tenser voice (Bird and Garellek 2019; Epstein 2002; Garellek 2015; Mooshammer 2010).

On the other hand, the occurrence of glottalization is significantly influenced by prosodic boundaries. It is well-known that aperiodic creak is more likely to occur at the prosodic domain-final positions: Creak is generally more likely to occur at the phrase-final position but much less likely to occur at the phrase-medial position (Davidson and Erker 2014; Dilley et al. 1996; Garellek 2013; Luthern and Clopper 2015; Seyfarth and Garellek 2020). Moreover, the likelihood of creak is strongly correlated with the strength of the prosodic boundaries: The larger the boundaries, the more frequent the creak. As such, creak is most likely to occur at the end of the utterance domain (Garellek 2013; Klatt and Klatt 1990; Pierrehumbert and Talkin 1992; Redi and Shattuck-Hufnagel 2001). The effect of phrase-final creak appears to be quite universal cross-linguistically, as it has been widely reported among tonal languages as well (Esposito 2003; Garellek 2012; Kalita et al. 2017; Kuang 2018).

Taken together, the creak of checked syllables in Shanghainese can be the consequence of the interaction of two distinct sources: the local phonemic effect from the checked coda, and the global laryngeal effect from the domain-final prosodic boundary. Therefore, to tease these two sources apart and to better understand the interaction between the two laryngeal functions, it is important to investigate the phonation variation of the checkedness contrast at different levels of prosodic boundaries.

Moreover, although the effects of large prosodic boundaries seem to be quite universal cross-linguistically, it remains largely unclear whether the smaller prosodic boundaries, especially those defined by language-specific phonology, such as the tone-sandhi domain in Shanghainese, also follow the general phrase-final effects.

1.6. Interaction between Global and Local Laryngeal Functions

This study is also more generally related to the question of the interaction between global vs. local laryngeal functions. As we discussed above, phonation variation can be driven by local functions (e.g., phonemic checkedness contrast at the syllable level), as well as by global functions, such as sentence prosody and other paralinguistic factors (e.g., vocal effort, emotion).

By far, the understanding of this topic is still extremely limited. Some studies have suggested that local laryngeal functions can be relatively independent of global laryngeal functions. For example, in Irish English, an accented syllable is consistently tenser than an unaccented syllable, regardless of the voice quality the speaker uses to pronounce the whole utterance (Yanushevskaya et al. 2016); in American English, the phonetic performance of the pitch accent of words differs from that of prosodic boundaries, with the former accompanied by a higher CQ, while the latter is accompanied by a lower CQ (Bird

and Garellek 2019). In White Hmong, a tonal language, the phonation differences of tones are not modulated by the prosodic position in the utterance (Garellek and Esposito 2021). Similarly, in Shaoxing Wu, a closely related language of Shanghaiese, it is found that the syllables of the lower register always have breathier phonation than those of the upper register, regardless of the vocal effort (Kuang et al. 2019).

In contrast, other studies have shown that the local laryngeal function is highly influenced by the global laryngeal function in some languages. Some studies have shown that the reinforcing effect of non-modal phonation, in terms of prosodic effects or other sources of vocal efforts, can sometimes weaken the phonetic differences between phonation contrasts. For example, in Santa Ana Del Valle Zapotec, the three-way phonation contrast is minimal when the target is isolated and has a focus or when the target is in the initial position (Esposito 2010). In Shaoxing Wu, while the phonation contrast is maintained in all vocal effort conditions, the contrast is less well-defined in the loud and soft conditions than in the normal condition (Kuang et al. 2019). Moreover, sometimes, it seems that the global laryngeal function can overwrite the local laryngeal function. For example, in Mandarin, larger phrasal boundaries are likely to have creak regardless of tonal categories (Kuang 2018).

In addition to the question of the distinctiveness of the phonemic phonation contrast, it is also important to understand whether the creak introduced by the global laryngeal functions has the same voice source mechanism as the creak introduced by the local laryngeal functions. By far, we still have very little knowledge about this issue. In a study of German, lexical stress and loud speech are found to be produced with similar phonetic cues, except that the extent of variation is greater for vocal effort (the global function) than for word stress (the local function) (Mooshammer 2010). Similar acoustic properties for /t/-glottalization and phrase-final in English are also observed (Garellek 2015). However, in Shaoxing Wu, the phonetic correlates of vocal effort are quite distinct from that of the phonemic register contrast (Kuang et al. 2019). Since different laryngeal functions are not always subject to the same voice source mechanism, it is necessary to examine more case studies on different languages and different types of prosodic contexts.

Furthermore, it is still unclear whether the phonetic correlates of the phonemic phonation contrast can vary according to the prosodic conditions. Limited evidence suggests that they might. For example, in San Lucas Quiavini Zapotec, stressed checked syllables are more likely to be produced with non-modal phonation (Chávez-Peón 2008); in English, more occurrence of creak is observed for word-medial /t/-glottalization at the sentence-final position (Pierrehumbert and Talkin 1992). Therefore, it is quite possible that the phonetic correlates of the checkedness contrast can contextually vary at different prosodic positions.

Overall, the interaction between global and local laryngeal functions is rather complicated and multifaceted. By looking into the phonation variation as the function of the interaction between prosodic boundaries and phonemic checkedness, this study will significantly advance our understanding of how prosodic structure manifests itself phonetically in tone languages.

1.7. Research Questions and Hypotheses

To summarize, this study aims to address several research questions:

First of all, what is the phonetic nature of checked syllables in Shanghaiese? To address this question, we examined all relevant phonetic cues for the checkedness contrast, including f_0 , duration, phonation and occurrence of creak. Due to the checked coda, it is likely that non-modal phonation is involved in the vowel portions of the checked syllables. To achieve a better understanding of the phonation involved in the checked syllables, both EGG and acoustic measures were collected and analyzed.

Secondly, how does phonation vary as a function of different levels of prosodic boundaries? In particular, does the tone-sandhi domain behave in the same way as other large prosodic domains (e.g., intermediate phrase and intonational phrase)? Furthermore, it is

likely that both prosodic boundaries and checkedness contrast involve some sort of glottalization, but are prosodic boundaries and the checkedness contrast produced with the same laryngeal mechanisms?

Lastly, how is the local checkedness contrast influenced by the global prosodic boundaries? It is possible for both checked coda and prosodic boundaries to introduce creak at the end of a syllable. By placing checked syllables in various prosodic boundaries, we are able to tease apart these two different functions and show how they interact with each other. In particular, we test whether the same set of phonation cues is involved in the checkedness contrast and the different prosodic boundaries.

2. Materials and Methods

2.1. Speech Materials

In this experiment, the participants were asked to produce the minimal pair of checked syllables [za12] vs. unchecked syllables [za23] in four different prosodic positions. The checked and unchecked rhymes to be measured occur in four different prosodic positions: (3a) sandhi-medial; (3b) sandhi-final but phrase-medial (sandhi-final); (3c) phrase-final but IP-medial (phrase-final); and (3d) IP-final. The four prosodic positions are named according to Section 1.2.

As illustrated in examples (3a–3d), the target syllables (the underscored positions in the carrier sentences) were designed as part of some pseudonyms (boldfaced syllables in the carrier sentence) and were located at the different types of prosodic boundaries. The tones in the IPA transcripts in (3a–3d) are omitted.¹ [za12] and [za23] were chosen as the target syllables because they are minimally different in other phonetic aspects, such as pitch contours, vowel quality and onsets, and they are common syllables used in names.

- (3a) tɕʰin ___ **lin** ɕiɔbaŋfiɻ tsʰaŋ tsə ku
 let.politely ___lin kid sing CLF song
 ‘Please welcome the kid ___lin to sing a song.’ (Sandhi-medial)
- (3b) tɕʰin **li** ___ **lin** ɕiɔbaŋfiɻ tsʰaŋ tsə ku
 let.politely Li ___lin kid sing CLF song
 ‘Please welcome the kid Li ___lin to sing a song.’ (Sandhi-final)
- (3c) tɕʰin **li** ___ linlin liaŋ fɯei ɕiɔbaŋfiɻ tsʰaŋ tsə ku
 let.politely Li ___ Linlin two CLF kid sing CLF song
 ‘Please welcome the kids Li ___ and Linlin to sing a song.’ (Phrase-final)
- (3d) ŋu lə zin **li** ___ .
 I PROG look.for Li ___ .
 ‘I am looking for Li ___ .’ (IP-final)
- linlin kaŋ gə tsəŋ zɕiʰi tsə fiɻ fi ɕiɔtə
 Linlin say this CLF thing only have he know
 ‘Linlin said he is the only one who knows the matter.’

Specifically, the target syllable in (3a) is the initial syllable of a disyllabic pseudonym. Because a Shanghainese disyllabic pseudonym contains one single tone-sandhi domain, the checkedness contrast codas of target syllables occur at the sandhi-medial position (Roberts 2020; Xu et al. 1988).

The target syllable in (3b) is the second syllable of a trisyllabic pseudonym, which is a phrase in Shanghainese (Roberts 2020; Selkirk and Shen 1990; Xu et al. 1988). In Shanghainese, a trisyllabic phrase can contain either one single trisyllabic tone-sandhi domain or two tone-sandhi domains (a disyllabic domain + a monosyllabic domain). However, for personal names, unless the name has special meaning or is extremely frequently used (and normally is associated with an extremely famous person), the normal personal names use the two-tone-sandhi pattern. Because the name phrases in (3b) are pseudonyms that have no special meaning, and they are not extremely common, we expect them to be realized with two tone-sandhi domains.

We found that this was exactly the case. A diagnosis to determine whether a trisyllabic phrase is realized with one or two tone-sandhi domains is to observe the third syllable's tone—if the last syllable is realized in the same way as its citation tone, the last syllable itself forms an independent tone-sandhi domain; otherwise, if the last syllable is realized as a weak low tone, the third syllable is a part of its preceding constituent, and the whole phrase forms one single trisyllabic tone-sandhi domain. We manually checked and confirmed that all the third syllables in the pseudonyms in (3b) are realized as their citation tones in every repetition of each participant; therefore, all the trisyllabic pseudonyms in (3b) are realized with two tone-sandhi domains, and the target checkedness contrast coda occurs at the phrase-medial and sandhi-final position.

In (3c) and (3d), the target syllables are the final syllable of disyllabic pseudonyms, which are phrase domains. Moreover, the disyllabic pseudonym in (3d) is also recognized as the end of an IP, which is identified by a larger boundary pause between the target syllable and its following constituent. However, there is no such perception of 'actual silence' after the target syllable in (3c). Therefore, the target checkedness contrast coda is at the phrase-final but IP-medial position for (3c), while at the IP-final position for (3d).

In each prosodic position, each carrier sentence was repeated eleven times. To avoid potential prosodic reduction in the repetitions, the onset consonants of the following syllable were changed for each repetition. The tone and vowel quality remained the same for all repetitions. Even though we have attempted to minimize phonetic reduction, the late repetition of the target syllables still potentially has some phonetic reduction due to the similarity between the carrier sentences (Ling and Liang 2017).

2.2. Data Collection

Simultaneous audio and articulatory recordings were collected in July 2019 in a double-wall sound-attenuated booth in Shanghai, using the Komplete Audio 6 Sound Card by Field-Phon (Han et al. 2013). The audio signal was recorded with an AKG C544L Headset Microphone as the first channel, and the EGG signal was recorded with Kay 6103 as the second channel. The sampling rate of the recordings was 44.1 kHz.

Sixteen native urban Shanghainese speakers (9 female and 7 male, aged 19 to 39 at the time of recording) participated in the experiment. The entire experiment took approximately 40 min per participant, with participants taking a 5-min break every 10 min, for a total of two breaks. After completing the recording of all participants and before the data analysis started, we reviewed the obtained recorded signals. Data from two speakers (one young female and one middle-aged male) were excluded from the analysis because of the poor recording quality of the EGG signals. All participants reported having normal hearing and voice, and each participant was reimbursed with 50 RMB.

2.3. Measures

Since the efficient/reliable acoustic and articulatory indicators for non-modal phonation contrasts are indeed language-specific or even contrast-specific, we opted to use a multidimensional approach to measure phonation. In the current study, we compare both acoustic and articulatory measures of voice quality in checked and unchecked syllables from audio and EGG signals.

The vowel portions of the target syllables were manually segmented in Praat (Boersma and Weenink 2021) by the first author, who is a native speaker of Shanghainese. Tokens with failed f_0 tracking were excluded from the acoustic analysis. Extensive voice measurements from both audio and EGG signals were extracted by VoiceSauce (Shue et al. 2011) and EggWorks (Tehrani 2010), using the default settings (window size = 25 ms). All measures were extracted at every millisecond. We calculated the average value of each parameter over the time interval and further performed within-speaker z-score normalization. A total of 1140 tokens were collected and annotated in the experiment; 79 vowel intervals were excluded because they are shorter than 50 ms and, therefore, could not be re-

liably measured acoustically or articulatorily; 1061 tokens were involved in the subsequent analyses.

The acoustic measures of voice quality mainly cover two aspects of the voice source: spectral measures and periodicity/noise measures. Spectral measures include the amplitudes of individual harmonics ($H1^*$, $H2^*$, $H4^*$), the amplitudes of the harmonics nearest to the first three formants ($A1^*$, $A2^*$, $A3^*$), the harmonic nearest 2000 Hz ($H2k^*$), the spectral slope or relative amplitude differences between a lower-frequency harmonic and a higher-frequency harmonic ($H1^*-H2^*$, $H2^*-H4^*$, $H1^*-A1^*$, $H1^*-A2^*$, $H1^*-A3^*$, $H4^*-H2k^*$, $H2k^*-H5k^*$). All the spectral measures were corrected for formants following the algorithm developed by Iseli et al. (2007). The periodicity and noise measures include the root mean square energy (energy), Cepstral peak prominence (CPP), the subharmonic-to-harmonic ratio (SHR), and harmonic-to-noise ratio (HNR) measures for 0–500, 0–1500, 0–2500, and 0–3500 Hz. The acoustic measures are defined after Vicenik et al. (2021).

The articulatory measures from EGG include Contact Quotient (CQ) and Peak Increase in Contact (PIC). CQ was estimated with the ‘CQ_HT’ method in EggWorks, which used the positive dEGG peak moment to define the contacting moment and takes the intersection moment of the DC contour of the EGG signal and the down sloping as the decontacting moment. The ‘CQ_HT’ method was found to be the most accurate method to distinguish the onset phonation types in Shanghaiese (Tian and Kuang 2019). PIC was measured as the amplitude of the positive peak in the first derivative of the EGG signal.

2.4. Occurrence of Creak

Most of the spectral measures rely on successful pitch tracking, which, unfortunately, is likely to fail in the segments with strong irregularity or creak. To remedy this problem, we also manually marked the occurrence of creak in the target vowels and distinguished the occurrence of different types of creak by looking at the spectrogram. The occurrence of creak in both checked and unchecked tokens was annotated by the first author.

Based on the extent and magnitude of irregular voicing, three major categories were annotated. ‘Coda glottal stop’ was coded when a full glottal stop was present at the syllable coda, and as illustrated in Figure 1A, the presence of a glottal stop is evident by the presence of a strong glottal pulse following a brief silent period. The second type, ‘coda creak’, was coded when irregular voicing was present in the last third of the vowel portion (Figure 1B). Finally, ‘broader creak’ was coded when the vowel portion began to show significant irregular voicing earlier than the last third of the vowel portion (Figure 1C).

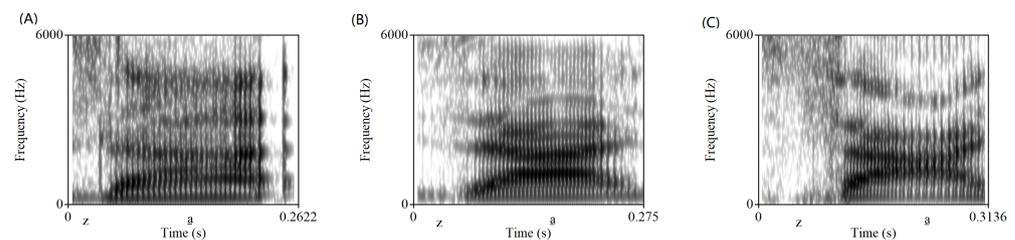


Figure 1. Three types of creakiness: (A) Coda glottal stop: short silence followed by a strong glottal pulse at the end of the syllable. (B) Coda creak: irregular voicing towards the end of the syllable. (C) Broader creak: irregular voicing occurred earlier than the last third of the vowel portion.

3. Results

3.1. Phonetic Measures

3.1.1. Acoustic Measures

Principal Component Analysis for Acoustic Measures

Since voice quality involves multidimensional acoustic cues, and many of the spectral or noise measures are correlated with each other, instead of exploring individual cues, we took a more integrative approach by fitting the high-dimensional acoustic measures into a Principal Component Analysis (PCA) model. The first principal component (PC1) and

the second principal component (PC2) together account for over 60% of the variance (score of PC1 = 40.9%, score of PC2 = 20.8%). The multidimensional acoustic space is plotted in Figure 2. The same acoustic space in Figure 2 is color-coded twice, in order to illustrate the effects of phonemic type (Figure 2a) and prosodic position (Figure 2b) separately.

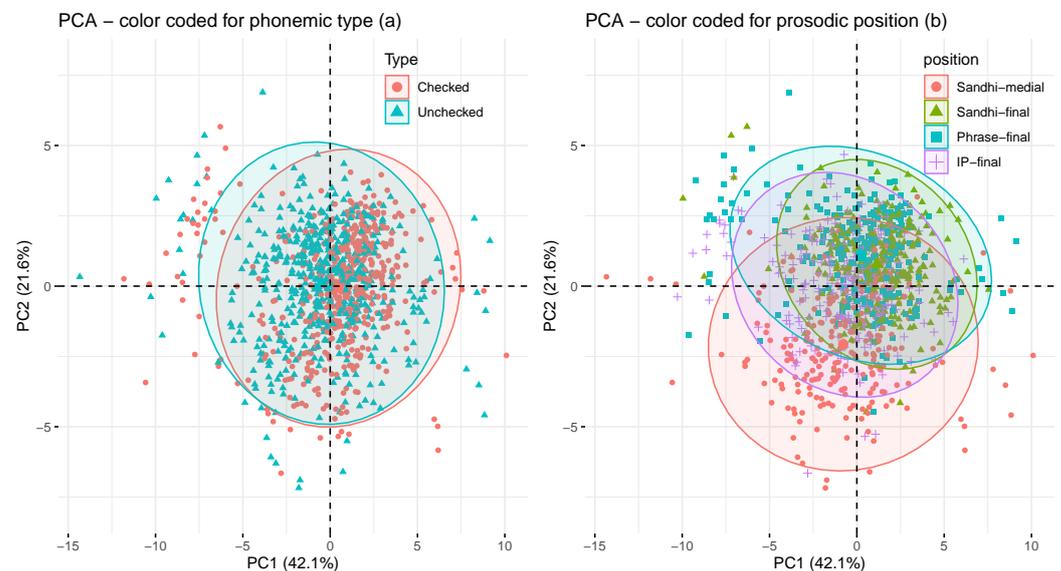


Figure 2. Principal Component Analysis of the acoustic space. (a) Color-coded for targets' phonemic type. (b) Color-coded for targets' prosodic position. Concentration ellipse level = 0.95.

As shown in Figure 2a,b, the phonemic types (checked vs. unchecked syllable) mostly contrast along PC1, while the prosodic positions (from small to large boundaries: sandhi-medial, sandhi-final but phrase-medial, phrase-final and IP-final) of the targets mostly vary along PC2. In particular, the sandhi-medial condition is generally more distinctive from all the three sandhi-final conditions (including sandhi-final but phrase-medial, phrase-final, and IP-final). This result suggests that tone sandhi is a critical prosodic condition for the phonation variation in Shanghainese.

In order to better understand the acoustic correlates of these principal components, the factor correlation loadings for PC1 and PC2 are plotted in Figure 3. A stronger correlation between an acoustic cue and a PC is when the angle between the feature vector and the PC axis is close to 0° (or 180°), and the length of the vector is relatively long. The direction of the vector indicates the direction of the correlation. As shown in Figure 3, PC1 is mostly correlated with spectral slope and the strength of individual harmonics in the higher-frequency range, and the most important cues are $A2^*$, $H1^*-A2^*$, $H1^*-A1^*$, $A3^*$, and $H1^*-A3^*$. More specifically, as the direction of the vectors suggests, PC1 is negatively correlated with spectral tilt measures ($H1^*-An^*$) and positively correlated with the strength of the individual harmonics (An^*). Therefore, greater PC1 generally means more constricted glottis. PC2 is, on the other hand, mostly correlated to periodicity/noise measures, such as HNR15, HNR25, HNR35, HNR05, and CPP; greater PC2 is correlated with greater periodicity. Therefore, PC1 and PC2 generally represent different aspects of glottal articulation. Taking Figures 2 and 3 together, it appears that the phonemic contrast of checked vs. unchecked syllables is mostly distinguished by spectral slope measures, while the prosodic boundaries mostly differ in periodicity and noise measures.

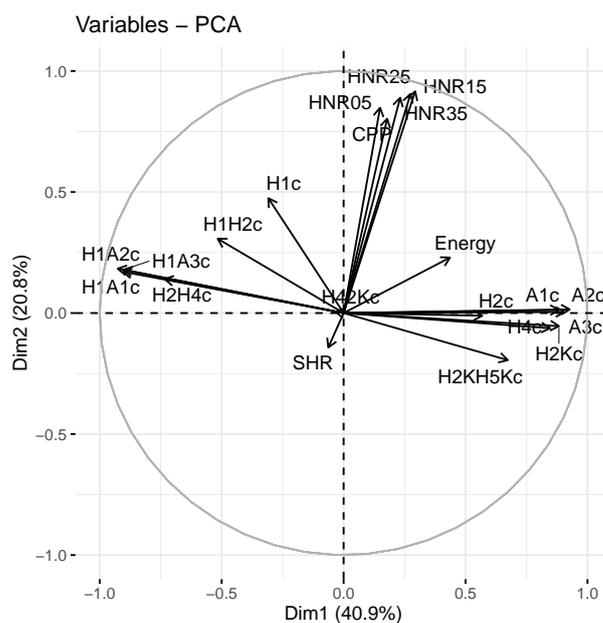


Figure 3. The loadings for PC1 and PC2 of all acoustic features. The most correlated cues for PC1 are A2*, H1*–A2*, H1*–A1*, A3*, and H1*–A3*; the most correlated cues for PC2 are HNR15, HNR25, HNR35, HNR05, and CPP.

Linear Mixed-Effect Regression Models for Acoustic Measures

To further examine the interaction between local phonemic types and global prosodic positions, linear mixed-effect regression models were fitted for each of the first two principal components in R with the ‘lmerTest’ package (Kuznetsova et al. 2017; R Core Team 2021). For each model, phonemic type and prosodic position and their interaction were set as the fixed factors; a random by-speaker intercept and a random by-following-consonant intercept were included as random effects. Categorical variables in these models were simple-coded; with such coding, the intercept of the model is the grand mean (Sonderegger 2020). If the interaction between phonemic type and prosodic position was insignificant in the model, we refitted the linear mixed-effect regression model without interaction effects and reported the output of the simpler model. To obtain pairwise comparisons among all prosodic positions, the same model was rerun multiple times for different reference levels.

PC1

As discussed in Principal Component Analysis for Acoustic Measures and shown in Figure 3, PC1 is predominantly negatively correlated with spectral slope measures (e.g., H1*–An*); therefore, greater PC1 indicates a more constricted glottis.

We fitted a linear mixed-effects regression model for PC1 with the parameters described above. Because there was no significant interaction between phonemic type and prosodic position, the model was refitted without the interaction effects.

As indicated by Table 2, both phonemic type and prosodic position have significant main effects on PC1. This result confirms our observation based on Figure 2. The pairwise comparisons also suggest that PC1 is significantly different between every two prosodic positions, except for sandhi-medial vs. IP-final. As illustrated in Figure 4, there is a general trend that larger prosodic boundaries have smaller PC1 values compared to smaller prosodic boundaries. Therefore, we see a trend of sandhi-final > phrase-final > IP-final. However, sandhi-medial is not part of this trend.

More importantly, there is no significant interaction between phonemic type and prosodic position, suggesting that checked vs. unchecked syllables are well-distinguished by PC1 (i.e., spectral slope measures, c.f. Figure 3) in all prosodic conditions. This is further confirmed by a post-hoc Games–Howell test (significant *p*-values are indicated in Figure 4).

Table 2. Outputs of linear mixed-effect regression for PC1 without interaction effects. Significant *p*-values are marked in bold.

Effect	Comparison	PC1			
		Est	SE	t	<i>p</i>
Type	Unchecked vs. Checked	−1.01	0.17	−6.00	0.00
Position	Sandhi-final vs. Sandhi-medial	1.92	0.24	7.98	0.00
	Phrase-final vs. Sandhi-medial	1.01	0.24	4.16	0.00
	IP-final vs. Sandhi-medial	0.05	0.24	0.20	0.84
	Phrase-final vs. Sandhi-final	−0.92	0.24	−3.90	0.00
	IP-final vs. Sandhi-final	−1.87	0.24	−7.95	0.00
	IP-final vs. Phrase-final	−0.96	0.24	−4.00	0.00

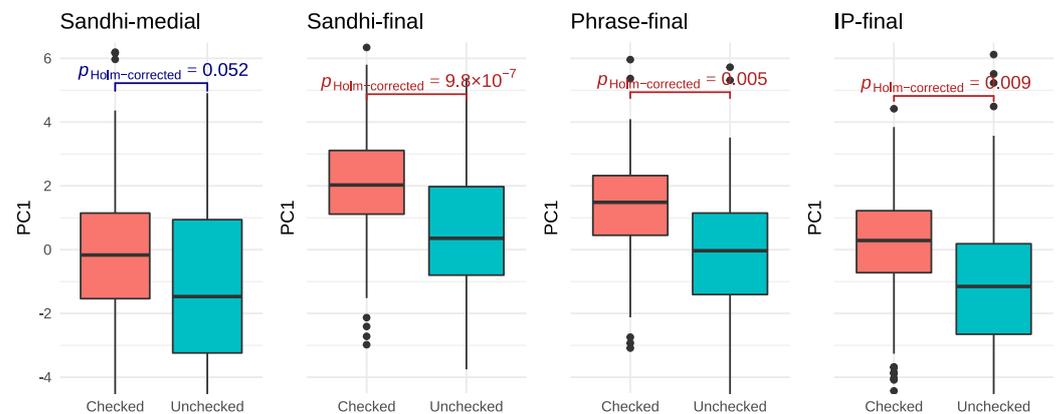


Figure 4. The variation of PC1 influenced by phonemic type and prosodic position. Greater PC1 indicates a more constricted glottis. Significant *p*-values ($p \leq 0.05$) are marked in red, which indicates that the PC1 difference between checked and unchecked syllables is significant in that prosodic position.

PC2

Similar to PC1, since there was also no significant interaction effect for PC2, the linear mixed-effect regression model was refitted without the interaction effects. The outputs of the regression models for PC2 are summarized in Table 3, and the result is visualized in Figure 5.

As can be seen here, there is a significant effect of prosodic position on PC2, but the phonemic type has no statistically significant main effects. This result further validates our observation from Figure 2—PC2 is mostly about the effects of the prosodic positions. As illustrated in Figure 5, the post-hoc Games–Howell test confirms that there is no checked vs. unchecked distinction in any of the prosodic positions.

As discussed in Principal Component Analysis for Acoustic Measures and shown in Figure 3, PC2 is mostly related to the periodicity and noise ratio. Greater PC2 is correlated with greater periodicity. IP-non-final positions (both sandhi-final and phrase-final) are generally more periodic than IP-final. Again, sandhi-medial is not part of the trend and is significantly less periodic than all domain-final positions.

Table 3. Outputs of linear mixed-effect regression for PC2 without interaction effects. Significant *p*-values ($p \leq 0.05$) are marked in bold.

Effect	Comparison	PC2			
		Est	SE	t	<i>p</i>
Type	Unchecked vs. Checked	−0.20	0.10	1.91	0.06
Position	Sandhi-final vs. Sandhi-medial	2.79	0.15	18.78	0.00
	Phrase-final vs. Sandhi-medial	3.17	0.15	21.22	0.00
	IP-final vs. Sandhi-medial	2.14	0.15	14.26	0.00
	Phrase-final vs. Sandhi-final	0.38	0.15	2.59	0.01
	IP-final vs. Sandhi-final	−0.65	0.15	−4.45	0.00
	IP-final vs. Phrase-final	−1.03	0.15	−7.05	0.00

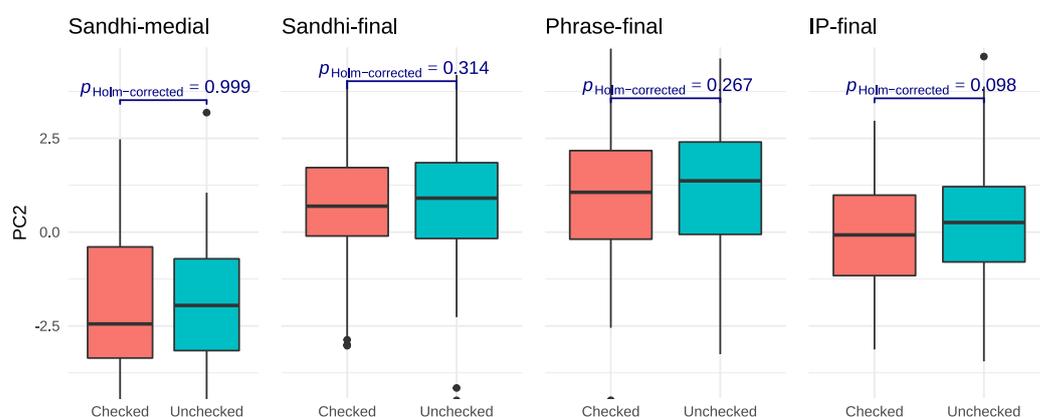


Figure 5. The variation of PC2 is influenced by phonemic type and prosodic position. Greater PC1 indicates higher periodicity during the vowel portion. The *p*-values at all prosodic positions are insignificant ($p > 0.05$, shown in blue); this indicates that the PC2 differences between checked and unchecked syllables are insignificant at all prosodic positions.

3.1.2. Articulatory Measures

Similarly, linear mixed-effect regression models were fitted for CQ and PIC, respectively. Again, phonemic type and prosodic position, as well as their interaction, were fit as the fixed factors. The by-speaker and by-following-consonant random intercepts were included as random effects. Pairwise comparisons among all prosodic positions were obtained by changing the reference levels of the model. Categorical variables were simple-coded. The modeling included all 1061 data points used in the acoustic analysis.

CQ

The outputs of the regression models for CQ are summarized in Table 4. As indicated by the table, there is a significant main effect of phonemic type. As expected, checked syllables have higher CQ values than unchecked syllables. Therefore, the phonation involved in the checked syllables has a relatively longer glottal closure duration. However, CQ is not a reliable measure of prosodic position. Essentially, only the IP-final position is significantly different from the other smaller prosodic boundaries. Moreover, there is a weak interaction between phonemic type and prosodic position. The post-hoc Games–Howell test suggests that CQ does not reliably distinguish checked vs. unchecked syllables in the sandhi-final and phrase-final positions (significant *p*-values are indicated in Figure 6).

Table 4. Outputs of linear mixed-effect regression for CQ with interaction effect. Significant *p*-values ($p \leq 0.05$) are marked in bold.

Effect	Comparison	CQ			
		Est	SE	t	<i>p</i>
Type	Unchecked vs. Checked	−0.13	0.04	−3.61	0.00
Position	Sandhi-final vs. Sandhi-medial	−0.07	0.05	−1.41	0.16
	Phrase-final vs. Sandhi-medial	−0.05	0.05	−1.03	0.30
	IP-final vs. Sandhi-medial	−0.21	0.05	−3.93	0.00
	Phrase-final vs. Sandhi-final	0.02	0.05	0.38	0.70
	IP-final vs. Sandhi-final	−0.13	0.05	−2.61	0.01
	IP-final vs. Phrase-final	−0.15	0.09	−1.64	0.13
Type:Position	Type:Sandhi-final vs. Type:Sandhi-medial	0.19	0.10	1.79	0.07
	Type:Phrase-final vs. Type:Sandhi-medial	0.13	0.10	1.24	0.21
	Type:IP-final vs. Type:Sandhi-medial	−0.05	0.11	−0.52	0.60
	Type:Phrase-final vs. Type:Sandhi-final	−0.04	0.09	−0.45	0.66
	Type:IP-final vs. Type:Sandhi-final	−0.06	0.10	−0.55	0.58
	Type:IP-final vs. Type:Phrase-final	−0.24	0.10	−2.36	0.02

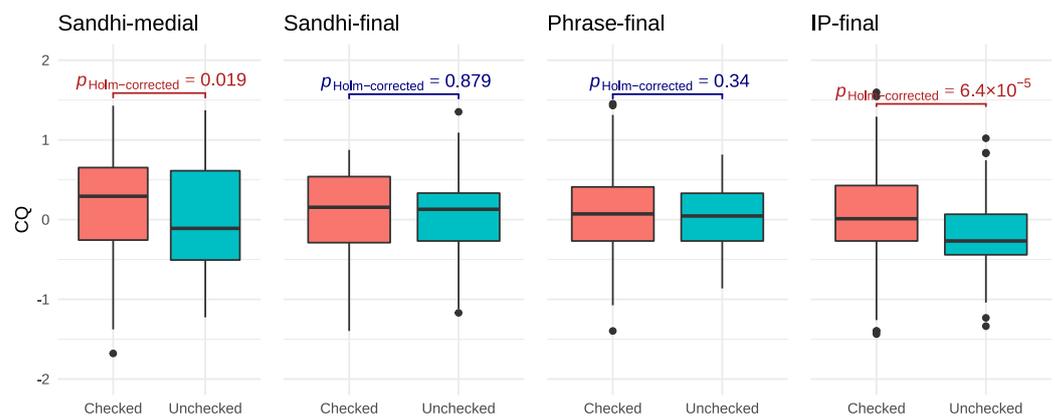


Figure 6. The variation of CQ influenced by phonemic type and prosodic position. Significant *p*-values ($p \leq 0.05$) are marked in red, which indicates that the CQ difference between checked and unchecked syllables is significant in that prosodic position.

PIC

The fixed effects on PIC are also analyzed in the linear mixed-effect regression model, and the output is in Table 5.

As indicated in the table, significant main effects were found for both phonemic type and prosodic position. In addition, PIC significantly differs between every two prosodic positions, except for sandhi-medial vs. IP-final. Moreover, there is generally no significant interaction between phonemic type and prosodic position, except that checked and unchecked syllables are more contrasting in phrase-medial position than in sandhi-final position. This result is similar to that of PC1. As illustrated in Figure 7, similar to PC1, PIC also exhibits a trend that larger prosodic boundaries have smaller PIC values than the smaller prosodic boundaries; but again, the sandhi-medial position is the exception of this trend. The post hoc test further confirms that PIC is a reliable measure for the checked vs. unchecked contrast across all prosodic positions.

Table 5. Outputs of linear mixed-effect regression for PIC with interaction effect. Significant *p*-values ($p \leq 0.05$) are marked in bold.

Effect	Comparison	PIC			
		Est	SE	t	<i>p</i>
Type	Unchecked vs. Checked	−0.48	0.05	−9.92	0.00
Position	Sandhi-final vs. Sandhi-medial	0.30	0.07	1.994.33	0.00
	Phrase-final vs. Sandhi-medial	0.14	0.07	2.06	0.04
	IP-final vs. Sandhi-medial	−0.13	0.07	−1.91	0.06
	Phrase-final vs. Sandhi-final	−0.16	0.07	−2.34	0.02
	IP-final vs. Sandhi-final	−0.43	0.07	−6.39	0.00
	IP-final vs. Phrase-final	−0.28	0.07	−4.04	0.00
Type:Position	Type:Sandhi-final vs. Type:Sandhi-medial	0.21	0.14	1.50	0.13
	Type:Phrase-final vs. Type:Sandhi-medial	0.30	0.14	2.12	0.03
	Type:IP-final vs. Type:Sandhi-medial	0.17	0.14	1.21	0.23
	Type:Phrase-final vs. Type:Sandhi-final	0.09	0.14	0.65	0.52
	Type:IP-final vs. Type:Sandhi-final	−0.04	0.14	−0.29	0.77
	Type:IP-final vs. Type:Phrase-final	−0.13	0.14	−0.93	0.35

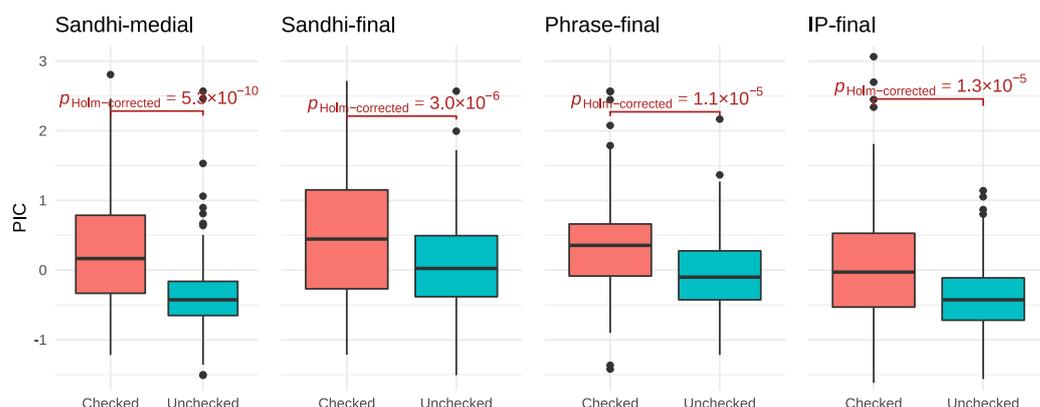


Figure 7. The variation of PIC influenced by phonemic type and prosodic position. Significant *p*-values ($p \leq 0.05$) are marked in red, which indicates that the PIC difference between checked and unchecked syllables is significant in that prosodic position.

3.1.3. F0

As described in Section 2.3, *f*₀ was extracted from both checked and unchecked vowels and was also fitted into linear mixed-effect regression models. Phonemic type, prosodic position, and their interaction were set as the fixed factors of the model; we also included a by-speaker random intercept and a by-following-consonant random intercept.

Table 6 shows that there is a significant main effect of phonemic type. Checked syllables have generally higher *f*₀ than unchecked syllables. Traditionally, checked tones were transcribed with a lower pitch than the corresponding unchecked tones (Xu et al. 1988, c.f. Table 1). However, our current study indicates the opposite way—for lower-register tones, checked tones are produced with a higher *f*₀ than unchecked tones by the younger generation of Shanghaiese.

In addition, there is a significant main effect of prosodic position on *f*₀. As can be seen in Figure 8, except for the non-final sandhi-medial position, *f*₀ of the smaller boundaries, such as sandhi-final position, is higher than that in the larger boundaries, such as phrase-final and IP-final positions, exhibiting a declination trend (i.e., Roberts 2020). The effect of prosodic position is similar to the patterns observed in PC1 and PIC. Moreover, there is a significant interaction effect between phonemic type and prosodic position for *f*₀. The difference in *f*₀ between checked and unchecked syllables is smaller in the IP-final position than in sandhi-final and phrase-final positions.

Table 6. Outputs of linear mixed-effect regression for f0 with interaction as a fixed factor. Significant *p*-values ($p \leq 0.05$) are marked in bold.

Effect	Comparison	F0			
		Est	SE	t	<i>p</i>
Type	Unchecked vs. Checked	−0.27	0.03	−7.95	0.00
Position	Sandhi-final vs. Sandhi-medial	1.43	0.05	29.82	0.00
	Phrase-final vs. Sandhi-medial	0.99	0.05	20.61	0.00
	IP-final vs. Sandhi-medial	0.10	0.05	2.10	0.04
	Phrase-final vs. Sandhi-final	−0.44	0.05	−9.35	0.00
	IP-final vs. Sandhi-final	−1.33	0.05	−28.25	0.00
	IP-final vs. Phrase-final	−0.89	0.05	−18.86	0.00
Type:Position	Type:Sandhi-final vs. Type:Sandhi-medial	0.10	0.10	1.09	0.28
	Type:Phrase-final vs. Type:Sandhi-medial	0.01	0.10	0.13	0.89
	Type:IP-final vs. Type:Sandhi-medial	0.26	0.10	2.71	0.01
	Type:Phrase-final vs. Type:Sandhi-final	−0.10	0.10	−1.09	0.28
	Type:IP-final vs. Type:Sandhi-final	−0.09	0.09	−0.98	0.33
	Type:IP-final vs. Type:Phrase-final	0.25	0.09	2.64	0.01

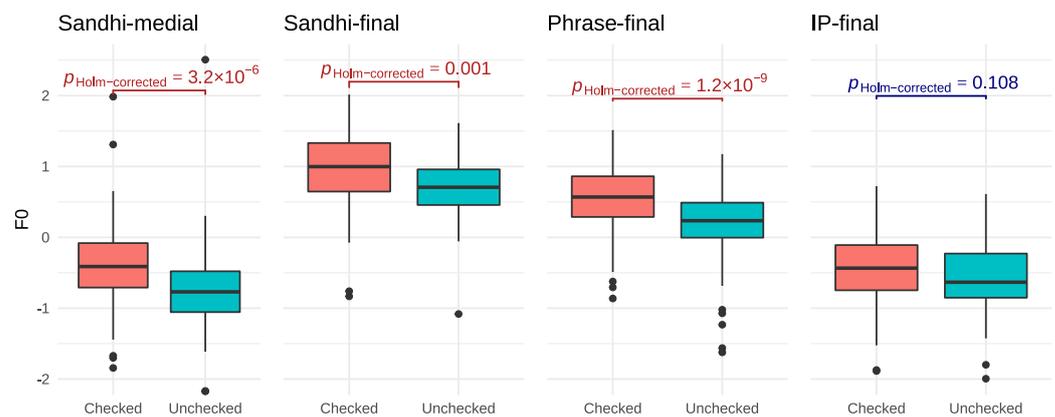


Figure 8. The variation of f0 influenced by phonemic type and prosodic position. Significant *p*-values ($p \leq 0.05$) are marked in red, which indicates that the f0 difference between checked and unchecked syllables is significant in that prosodic position.

3.1.4. Duration

The same linear mixed-effect regression model was fitted for vowel duration as well. The modeling outputs are summarized in Table 7. The significant main effects are found for phonemic type and prosodic position. As expected, checked vowels are significantly shorter than unchecked vowels. Furthermore, as demonstrated in Figure 9, vowel duration varies hierarchically with the size of the prosodic domains. The vowels located at the large prosodic boundaries (e.g., IP-final and phrase-final positions) are the longest, followed by those at the sandhi-final position, and the vowels at the sandhi-medial position are the shortest. This is consistent with the final lengthening effect widely reported among languages (Byrd and Saltzman 2003; Pan 2007; Turk and Shattuck-Hufnagel 2007, among others). Furthermore, although there is a significant interaction between phonemic type and prosodic position, significant duration differences between checked and unchecked vowels are maintained in all prosodic positions (*p*-values from the post-hoc Games–Howell tests are indicated in Figure 9).

Table 7. Outputs of linear mixed-effect regression for a duration with an interaction effect. Significant *p*-values ($p \leq 0.05$) are marked in bold.

Effect	Comparison	Duration			
		Est	SE	t	<i>p</i>
Type	Unchecked vs. Checked	1.29	0.04	32.79	0.00
Position	Sandhi-final vs. Sandhi-medial	0.14	0.06	2.58	0.01
	Phrase-final vs. Sandhi-medial	0.87	0.06	15.46	0.00
	IP-final vs. Sandhi-medial	0.97	0.06	17.15	0.00
	Phrase-final vs. Sandhi-final	0.72	0.05	13.28	0.00
	IP-final vs. Sandhi-final	0.82	0.05	15.03	0.00
	IP-final vs. Phrase-final	0.10	0.06	1.81	0.07
Type:Position	Type:Sandhi-final vs. Type:Sandhi-medial	-0.24	0.11	-2.13	0.03
	Type:Phrase-final vs. Type:Sandhi-medial	-0.54	0.11	-4.85	0.00
	Type:IP-final vs. Type:Sandhi-medial	-0.26	0.11	-2.27	0.02
	Type:Phrase-final vs. Type:Sandhi-final	0.24	0.11	2.13	0.04
	Type:IP-final vs. Type:Sandhi-final	-0.31	0.10	-2.82	0.00
	Type:IP-final vs. Type:Phrase-final	0.29	0.11	2.62	0.01

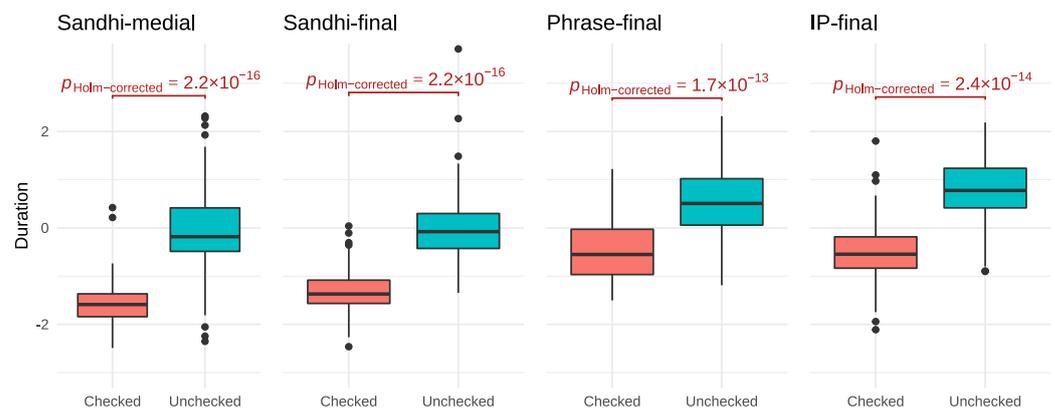


Figure 9. The variation of duration influenced by phonemic type and prosodic position. Significant *p*-values ($p \leq 0.05$) are marked in red, which indicates that the duration difference between checked and unchecked syllables is significant in that prosodic position.

3.1.5. Creak Occurrence

Finally, the occurrence of creak was manually coded for all tokens. As reviewed in Section 2.4, we specifically coded three types of creak: coda glottal stops, coda creak, and broader creak. To evaluate the effects of phonemic type and prosodic position on the likelihood of creak occurrence, a logistic mixed-effect regression model was fitted (three types combined), with phonemic type, prosodic position, and their interaction as the fixed factors; the by-speaker and by-following-consonant random intercepts were also included in the model. The fixed factors were again simple-coded. As in the previous phonetic analysis, we included 1061 tokens in the analysis. Because there is no significant interaction effect on creak occurrence, we refit the model without the interaction effect.

As indicated in Table 8, only the prosodic position has a significant main effect on creak occurrence. All prosodic positions have significantly different rates of creak from one another. This result can be better understood in Figure 10. Creak occurs more at the two larger prosodic boundaries (IP-final and phrase-final positions) than at the two smaller prosodic boundaries (sandhi-medial and sandhi-final). The breakdown of the three types of creak is also visualized in Figure 10. Overall, ‘coda creak’ appears to be the most frequent type, especially for syllables at the larger prosodic boundaries. ‘Coda glottal stops’ are also quite frequently present at the larger prosodic boundaries. However, ‘broader creak’ is more likely to occur at the sandhi-medial position.

Overall, aperiodic creak is primarily driven by prosodic boundaries. This result resonates with the PC2 variation pattern we have seen in Principal Component Analysis for Acoustic Measures. PC2, the principal component mostly correlated with noise-ratio measures, also exhibits strong effects of prosodic positions, and larger prosodic boundaries are less periodic than smaller prosodic boundaries.

Table 8. Outputs of logistic mixed-effect regression model for the occurrence of creak (three creak-types combined). Significant *p*-values ($p \leq 0.05$) are marked in bold.

Effect	Comparison	Creak Occurrence			
		Est	SE	z	<i>p</i>
Type	Unchecked vs. Checked	−0.14	0.20	−0.81	0.48
Position	Sandhi-final vs. Sandhi-medial	−1.97	0.64	−3.07	0.02
	Phrase-final vs. Sandhi-medial	1.73	0.32	5.45	0.00
	IP-final vs. Sandhi-medial	3.14	0.33	9.63	0.00
	Phrase-final vs. Sandhi-final	3.70	0.62	5.99	0.00
	IP-final vs. Sandhi-final	5.11	0.62	8.19	0.00
	IP-final vs. Phrase-final	1.41	0.23	6.08	0.00

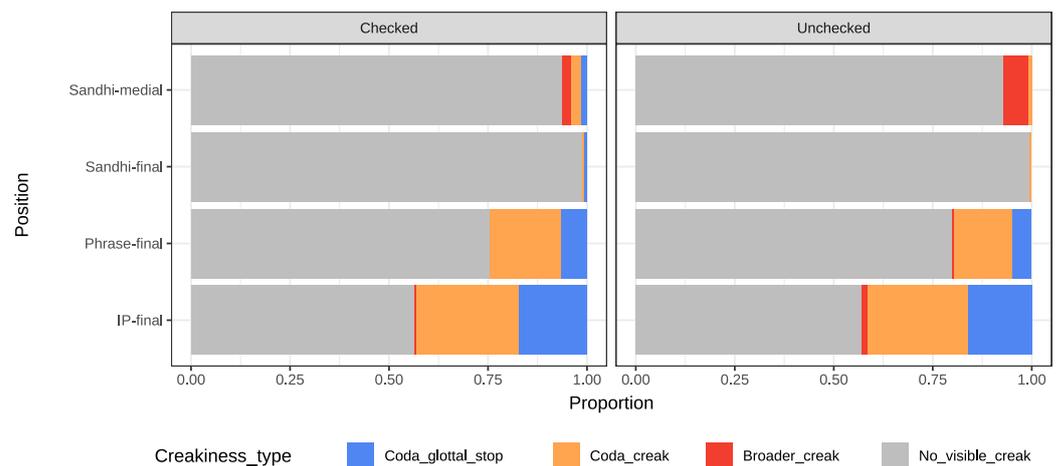


Figure 10. The distribution of tokens with three different types of creak (coded in non-gray colors) and tokens without visible creak (coded in gray) among checked and unchecked tones at various prosodic positions.

Since there is no significant difference in the occurrence of creak in contemporary Shanghainese between the checked and unchecked syllables, the coda glottal stop, based on the transcription of checked syllables in the older generation of Shanghainese (Xu et al. 1988), is no longer the most appropriate representation of checked syllables in the younger generation of Shanghainese.

3.2. Correlation between F0 and Phonation Measures

One remaining question is whether the phonation variation observed in the previous sections is pitch-driven due to the co-variation between pitch and phonation (Chai 2021; Kuang 2017). To explore this question, Pearson product-moment correlation coefficients were computed between *f*₀ and PC1, PC2, CQ, PIC, and creak occurrence (Benesty et al. 2009). Because creak occurrence is a binary variable, we performed the point biserial correlation of Pearson’s product-moment correlation when calculating the correlation between *f*₀ and creak occurrence. The correlation coefficients are summarized in Table 9.

Table 9. Correlation between f0 and the voice-quality measurements: PC1, PC2, CQ, PIC, and creak occurrence. Significant *p*-values ($p \leq 0.05$) are marked in bold.

	Correlation Coefficient with f0	<i>p</i> -Value
PC1	0.31	0.00
PC2	0.49	0.00
CQ	0.03	0.28
PIC	0.23	0.00
Creak occurrence	−0.22	0.00

As shown in Table 9, significant correlations are found between f0 and most phonation measures, including PC1 (spectral slope), PC2 (periodicity/noise ratio), PIC, as well as the creak occurrence. However, there is no significant correlation between f0 and CQ. Therefore, the phonation variation remains relatively independent of f0.

4. Discussion

This study aimed to examine the phonetic realization of the phonemic contrast between checked vs. unchecked syllables in Shanghainese and how this local coda laryngeal contrast interacts with the global prosodic contexts. To address these questions, comprehensive acoustic and articulatory voice measures were analyzed. A series of regression models were performed to evaluate the main effects and the interaction of phonemic type and prosodic position. Table 10 summarizes the results from the previous sections. Significant effects are indicated with plus signs.

Table 10. A summary of fixed factors on phonetic realization. Significant effects are indicated with plus signs.

	Factor Effect		
	Type	Position	Type:Position
PC1 (mainly spectral slopes)	+	+	
PC2 (mainly periodicity)		+	
CQ	+	+	+
PIC	+	+	+
F0	+	+	+
Duration	+	+	+
Creak occurrence		+	

4.1. Phonetic Nature of Shanghainese Checked Syllables

In general, as summarized in Table 10, checked syllables in Shanghainese are distinguished from unchecked syllables by having a non-modal phonation, a shorter duration, and a higher f0. However, the occurrence of creak does not differ significantly between checked and unchecked syllables. These results indicate that checked syllables are no longer phonetically distinguished by coda glottal stops; instead, the checkedness contrast is mostly realized as phonation, duration and f0 differences on the vowel portions.

One primary goal of our study is to better understand the phonation variation involved in the checkedness contrast. Both acoustic and articulatory measurements indicate a reliable phonation contrast between checked and unchecked syllables. There are several important aspects of the non-modal phonation involved in the checked syllables. First of all, checked syllables are associated with greater glottal constriction, as indicated by the greater CQ and flatter spectral slope (PC1). However, unlike the prototypical creaky voice (Keating et al. 2015), the non-modal phonation in the checked syllables is highly periodic, as noise-ratio and periodicity measures (e.g., PC2) do not significantly contribute to the contrast. Moreover, the checked syllables are produced with a slightly higher f0. Altogether, the checked syllables in Shanghainese are produced with a tenser phonation. Therefore, we propose to transcribe the checked syllables in contemporary Shanghainese

with [CV] to indicate the tense phonation. Moreover, contradictory to the traditional transcription (Xu et al. 1988), we found that the checked syllables have a higher f₀ than the corresponding unchecked syllables. This is another major change from the older generation of Shanghainese. The elevation in f₀ values is possibly driven by the tense phonation on the vowels (Keating et al. 2015; Kuang 2013; Maddieson and Ladefoged 1985).

Furthermore, there are some unique articulatory details reflected in the EGG measurements. In addition to CQ, PIC also reliably distinguishes the types of syllables in our study. As reviewed earlier in this paper, CQ and PIC reflect different aspects of glottal constriction. Greater CQ indicates a relatively longer contacting phase in each vibratory cycle. PIC is generally related to the status of the vocal folds at the moment of initiating contact. In our case, checked syllables with tense phonation are associated with larger PIC and positively correlated with PC1 ($r = 0.16, p < 0.05$). This association is consistent with the notion that abrupt glottal closure boosts the high-frequency energy in the spectrum (Hanson et al. 2001; Stevens 1977). It is useful to note that such significant association was not reported for some other phonation contrasts, such as the tense vs. lax contrast in Yi consonants (Kuang and Keating 2014), the breathy-nasalized vs. non-breathy-nasalized contrast in Yi vowels (Garellek et al. 2016), and the breathy vs. model syllable-onset contrast in Shanghainese (Tian and Kuang 2019). Therefore, the actual articulatory implication of PIC could be rather language-specific and contrast-specific.

4.2. Prosodic Effects on the Phonation Variation

In general, the laryngeal functions involved in the prosodic boundaries are quite different from the laryngeal functions involved in the phonemic type. As shown in Figure 2 and summarized in Table 10, unlike the phonemic contrast, the prosodic boundaries are primarily differentiated by the measures associated with periodicity (e.g., PC2 and the occurrence of creak), although the measures related to glottal constriction (e.g., PC1 and CQ) are also relevant.

Importantly, the prosodic positions investigated in our study do not simply co-vary uniformly with the phonetic features; instead, each prosodic position exhibits some unique phonetic properties. Our finding suggests that multiple mechanisms are contributing to the prosodic variation.

On the one hand, there is a rather gradient effect of the phrasal boundaries—larger prosodic boundaries are associated with less vocal constriction (e.g., PC1), less periodicity (e.g., PC2), and more creak occurrence. The IP-final position is the extreme end of this trend. This effect is consistent with the ‘non-constricted creak’ at the larger prosodic boundaries attested in many languages (Keating et al. 2015; Slifka 2006).

However, on the other hand, tone-sandhi is a distinct prosodic domain from the other levels of prosodic phrases. As shown in Figure 2b, the sandhi-medial position is generally distinguished from all the three sandhi-final positions. Furthermore, as demonstrated by PC1 and PC2, the sandhi-medial position does not participate in the general trend of the gradient creak effect of phrasal boundaries. In addition, the sandhi-final (but phrase-medial) position differs from the phrase-final position in that it has a much lower rate of creak. Overall, these findings suggest that tone-sandhi domains and prosodic phrases influence the phonation variation differently. This may be because the tone-sandhi domain in Shanghainese is specified by morphology and phonology, which is planned differently from the higher levels of prosodic domains (Roberts 2020; Selkirk and Shen 1990).

4.3. Interaction between the Global vs. Local Laryngeal Functions

Based on the discussions from the previous sections, it is clear that the global prosodic effects and the local phonemic contrast are subject to different voice mechanisms, as they involve distinct phonation cues. In particular, periodicity is only related to the global prosodic boundaries but not the local checkedness contrast.

Moreover, as summarized in Table 10, there is no interaction between the global vs. local functions for acoustic features. Interactions were only found for individual EGG

measures. For the sandhi-medial and IP-final positions, the phonemic contrast is well distinguished by both CQ and PIC, but for the sandhi-final and phrase-final positions, the phonemic contrast is only distinguished by PIC. Taking all the phonetic measures into consideration, the phonemic contrast between checked and unchecked syllables is essentially well-maintained in all prosodic contexts. Therefore, speakers of Shanghainese manage to control the global vs. local levels of phonation variation rather independently.

5. Conclusions

This study investigated the phonetic realization of the checked vs. unchecked contrast at various prosodic positions in Shanghainese. By extensively analyzing both acoustic and articulatory measures, we found that the Shanghainese checked codas are realized with a tenser phonation, a higher f_0 , and a shorter duration in the vowel segment. This study also provides us with a better understanding of the prosodic hierarchies of Shanghainese. The tone-sandhi domain, which is specified by phonology and morphology, is not subject to the general mechanism of prosodic phrases. Moreover, this study achieved a comprehensive understanding of the range of phonation variation as the function of the interaction between the global laryngeal function (the prosodic boundaries) and local phonemic contrast (the checked vs. unchecked contrast) in Shanghainese. Overall, the laryngeal functions of the global prosodic contexts can be rather independent of that of the local phonemic contrast, and the local phonemic contrast can be well maintained in all prosodic contexts.

Author Contributions: Conceptualization, X.G and J.K.; methodology, X.G.; software, X.G.; validation, X.G. and J.K.; formal analysis, X.G.; investigation, X.G. and J.K.; resources, X.G.; data curation, X.G.; writing—original draft preparation, X.G. and J.K.; writing—review and editing, X.G. and J.K.; visualization, X.G.; supervision, J.K.; project administration, X.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study, as it was not related to bio-medical and health research.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw data is available with the first author and can be made available for non-profit, academic and research purposes.

Acknowledgments: We thank the Phonetics Laboratory at Fudan University for providing the space and facilities for the experiments. We thank the Penn Phonetics Laboratory for providing a venue for presentation and discussion while this research was in its early stages. We are especially grateful to Huan Tao, Mark Liberman, Zihao Wei, May Chan, Meredith Tamminga, Zhuocheng Zhao, and Yifei Zheng for their constructive feedback. We sincerely thank Christina Esposito, Sameer Khan, Marc Garellek, and three anonymous reviewers for their thorough and insightful comments during the review process. We thank all experiment participants for sharing their language knowledge with us.

Conflicts of Interest: The authors declare no conflict of interest.

Note

¹ CLF = classifier; PROG = progressive. The IPA transcription is based on Xu et al. (1988).

References

- Benesty, Jacob, Jingdong Chen, Yiteng Huang, and Israel Cohen. 2009. Pearson correlation coefficient. In *Noise Reduction in Speech Processing*. Berlin: Springer, pp. 37–40.
- Bennett, Ryan. 2016. Mayan phonology. *Language and Linguistics Compass* 10: 469–514. [CrossRef]
- Bird, Elizabeth, and Marc Garellek. 2019. Dynamics of voice quality over the course of the English utterance. Paper presented at 19th International Congress of Phonetic Sciences, Melbourne, Australia, August 5–9. pp. 2406–410.
- Boersma, Paul, and David Weenink. 2021. Praat: Doing Phonetics by Computer [Computer Program]. Version 6.1.42. Available online: <https://www.fon.hum.uva.nl/praat/> (accessed on 15 April 2021).

- Borroff, Marianne. 2007. A Landmark Underspecification Account of the Patterning of Glottal Stop. Ph.D. thesis, Stony Brook University, Stony Brook, NY, USA.
- Brunelle, Marc, and James Kirby. 2016. Tone and phonation in Southeast Asian languages. *Language and Linguistics Compass* 10: 191–207. [\[CrossRef\]](#)
- Byrd, Dani, and Elliot Saltzman. 2003. The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31: 149–80. [\[CrossRef\]](#)
- Chai, Yuan. 2021. The source of creak in Mandarin utterances. In Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia, August 5–9.
- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chávez-Peón, Mario E. 2008. Phonetic cues to stress in a tonal language: Prosodic prominence in San Lucas Quiavini Zapotec. In Proceedings of the 2008 Annual Conference of the Canadian Linguistic Association, Vancouver, BC, Canada, May 31–June 2.
- Chen, Yiya. 2008. Revisiting the phonetics and phonology of Shanghai tone sandhi. In Proceedings of the Fourth Conference on Speech Prosody, Campinas, Brazil, May 6–8. pp. 253–56.
- Chen, Yiya, and Carlos Gussenhoven. 2015. Shanghai Chinese. *Journal of the International Phonetic Association* 45: 321–37. [\[CrossRef\]](#)
- Davidson, Lisa. 2020. The versatility of creaky phonation: Segmental, prosodic, and sociolinguistic uses in the world's languages. *Wiley Interdisciplinary Reviews: Cognitive Science* 12: e1547. [\[CrossRef\]](#)
- Davidson, Lisa, and Daniel Erker. 2014. Hiatus resolution in American English: The case against glide insertion. *Language* 90: 482–514. [\[CrossRef\]](#)
- DiCanio, Christian T. 2009. The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association* 39: 162–88. [\[CrossRef\]](#)
- DiCanio, Christian T. 2012. Coarticulation between tone and glottal consonants in Itunyoso Trique. *Journal of Phonetics* 40: 162–76. [\[CrossRef\]](#)
- Dilley, Laura, Stefanie Shattuck-Hufnagel, and Mari Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24: 423–44. [\[CrossRef\]](#)
- Duanmu, San. 1999. Metrical structure and tone: Evidence from Mandarin and Shanghai. *Journal of East Asian Linguistics* 8: 1–38. [\[CrossRef\]](#)
- Epstein, Melissa Ann. 2002. Voice Quality and Prosody in English. Ph.D. thesis, University of California, Los Angeles, CA, USA.
- Esling, John H., Katherine E. Fraser, and Jimmy G. Harris. 2005. Glottal stop, glottalized resonants, and pharyngeals: A reinterpretation with evidence from a laryngoscopic study of Nuuchahnulth (Nootka). *Journal of Phonetics* 33: 383–410. [\[CrossRef\]](#)
- Esposito, Christina Marie. 2003. Santa Ana del Valle Zapotec Phonation. Ph.D. thesis, University of California, Los Angeles, CA, USA.
- Esposito, Christina Marie. 2010. Variation in contrastive phonation in Santa Ana del Valle Zapotec. *Journal of the International Phonetic Association* 40: 181–98. [\[CrossRef\]](#)
- Esposito, Christina Marie, and Sameer ud Dowla Khan. 2012. Contrastive breathiness across consonants and vowels: A comparative study of Gujarati and White Hmong. *Journal of the International Phonetic Association* 42: 123–43. [\[CrossRef\]](#)
- Esposito, Christina Marie, and Sameer ud Dowla Khan. 2020. The cross-linguistic patterns of phonation types. *Language and Linguistics Compass* 14: e12392. [\[CrossRef\]](#)
- Frazier, Melissa. 2013. The phonetics of Yucatec Maya and the typology of laryngeal complexity. *Language Typology and Universals* 66: 7–21. [\[CrossRef\]](#)
- Garellek, Marc. 2012. The timing and sequencing of coarticulated non-modal phonation in English and White Hmong. *Journal of Phonetics* 40: 152–61. [\[CrossRef\]](#)
- Garellek, Marc. 2013. Production and Perception of Glottal Stops. Ph.D. thesis, University of California, Los Angeles, CA, USA.
- Garellek, Marc. 2015. Perception of glottalization and phrase-final creak. *The Journal of the Acoustical Society of America* 137: 822–31. [\[CrossRef\]](#)
- Garellek, Marc. 2020. Acoustic discriminability of the complex phonation system in! Xóõ. *Phonetica* 77: 131–60. [\[CrossRef\]](#)
- Garellek, Marc, Yuan Chai, Yaqian Huang, and Maxine Van Doren. 2021. Voicing of glottal consonants and non-modal vowels. *Journal of the International Phonetic Association* 2021: 1–28. [\[CrossRef\]](#)
- Garellek, Marc, and Christina M. Esposito. 2021. Phonetics of White Hmong vowel and tonal contrasts. *Journal of the International Phonetic Association* 2021: 1–20. [\[CrossRef\]](#)
- Garellek, Marc, and Patricia Keating. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association* 41: 185–205. [\[CrossRef\]](#)
- Garellek, Marc, Amanda Ritchart, and Jianjing Kuang. 2016. Breathily voice during nasality: A cross-linguistic study. *Journal of Phonetics* 59: 110–21. [\[CrossRef\]](#)
- Gerratt, Bruce R., and Jody Kreiman. 2001. Toward a taxonomy of nonmodal phonation. *Journal of Phonetics* 29: 365–81. [\[CrossRef\]](#)
- Gordon, Matthew, and Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of phonetics* 29: 383–406. [\[CrossRef\]](#)
- Guion, Susan G., Mark W. Post, and Doris L. Payne. 2004. Phonetic correlates of tongue root vowel contrasts in Maa. *Journal of Phonetics* 32: 517–42. [\[CrossRef\]](#)

- Han, Xia, Long Li, and Wuyun Pan. 2013. Computer based field investigation and processing system for languages. *Journal of Tsinghua University (Science and Technology)* 53: 888–92.
- Hanson, Helen M., Kenneth N. Stevens, Hong-Kwang Jeff Kuo, Marilyn Y. Chen, and Janet Slifka. 2001. Towards models of phonation. *Journal of Phonetics* 29: 451–80. [\[CrossRef\]](#)
- Hargus, Sharon. 2016. Deg Xinag word-final glottalized consonants and voice quality. In *The Phonetics and Phonology of Laryngeal Features in Native American Languages*. Leiden: Brill, pp. 71–128.
- Howard, David M. 1995. Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *Journal of Voice* 9: 163–72. [\[CrossRef\]](#)
- Iseli, Markus, Yen-Liang Shue, and Abeer Alwan. 2007. Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America* 121: 2283–95. [\[CrossRef\]](#)
- Jiang, Ying, Yize Tang, Wenda Lu, Zhongfeng Wang, Zepeng Wang, and Luming Zhang. 2017. Intelligent acoustic data fusion technique for information security analysis. *Journal of Physics: Conference Series* 887: 012090. [\[CrossRef\]](#)
- Kalita, Sishir, Wendy Lalhminghlu, Luke Horo, Priyankoo Sarmah, S. R. Mahadeva Prasanna, and Samarendra Dandapat. 2017. Acoustic Characterization of Word-Final Glottal Stops in Mizo and Assam Sora. Paper presented at Interspeech, August 20, pp. 1039–43.
- Kasim, Ziyad Rakan. 2019. An acoustic investigation of the glottal stop in Arabic. Paper presented at 19th International Congress of Phonetic Sciences, Melbourne, Australia, August 5–9.
- Keating, Patricia, Christina M. Esposito, Marc Garellek, and Jianjing Kuang. 2010. WPP, No. 108: Phonation Contrasts Across Languages. In *UCLA Working Papers in Phonetics*, Los Angeles: Department of Linguistics, UCLA, vol. 108, pp. 188–202.
- Keating, Patricia A., Marc Garellek, and Jody Kreiman. 2015. Acoustic properties of different kinds of creaky voice. Paper presented at 18th International Congress of Phonetic Sciences, Glasgow, Scotland, August 10–14. vol. 2015, pp. 2–7.
- Kingston, John. 2005. The phonetics of Athabaskan tonogenesis. *Amsterdam Studies in the Theory and History of Linguistic Science Series* 4 269: 137.
- Klatt, Dennis H., and Laura C. Klatt. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America* 87: 820–57. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kreiman, Jody, and Bruce R. Gerratt. 2012. Perceptual interaction of the harmonic source and noise in voice. *The Journal of the Acoustical Society of America* 131: 492–500. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kreiman, Jody, Yen-Liang Shue, Gang Chen, Markus Iseli, Bruce R. Gerratt, Juergen Neubauer, and Abeer Alwan. 2012. Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. *The Journal of the Acoustical Society of America* 132: 2625–32. [\[CrossRef\]](#)
- Kuang, Jianjing. 2013. The tonal space of contrastive five level tones. *Phonetica* 70: 1–23. [\[CrossRef\]](#)
- Kuang, Jianjing. 2017. Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice. *The Journal of the Acoustical Society of America* 142: 1693–706. [\[CrossRef\]](#)
- Kuang, Jianjing. 2018. The influence of tonal categories and prosodic boundaries on the creakiness in Mandarin. *The Journal of the Acoustical Society of America* 143: EL509–EL515. [\[CrossRef\]](#)
- Kuang, Jianjing, and Patricia Keating. 2014. Vocal fold vibratory patterns in tense versus lax phonation contrasts. *The Journal of the Acoustical Society of America* 136: 2784–97. [\[CrossRef\]](#)
- Kuang, Jianjing, Jia Tian, and Yipei Zhou. 2018. The common word prosody in Northern Wu. Paper presented at 6th International Symposium on Tonal Aspects of Language, Berlin, Germany, June 18–20. pp. 7–11.
- Kuang, Jianjing, Jia Tian, and Bing'er Jiang. 2019. The effect of vocal effort on contrastive voice quality in Shaoxing Wu. *The Journal of the Acoustical Society of America* 146: EL272–EL278. [\[CrossRef\]](#)
- Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82: 1–26. [\[CrossRef\]](#)
- Ladefoged, Peter. 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.
- Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell, vol. 1012.
- Li, Yinghao, and Jinghua Zhang. 2020. Electroglottographic-Phonetic Study on Korean Phonation Induced by Tripartite Plosives in Yanbian Korean. Paper presented at Interspeech, Shanghai, China, October 25–29. pp. 666–70. [\[CrossRef\]](#)
- Ling, Bijun, and Jie Liang. 2016. Organizing Syllables into Sandhi Domains-Evidence from F0 and Duration Patterns in Shanghai Chinese. Paper presented at Interspeech, San Francisco, CA, USA, September 8–12, pp. 72–76. [\[CrossRef\]](#)
- Ling, Bijun, and Jie Liang. 2017. Focus encoding and prosodic structure in Shanghai Chinese. *The Journal of the Acoustical Society of America* 141: EL610–EL616. [\[CrossRef\]](#) [\[PubMed\]](#)
- Luthern, Erin, and Cynthia G. Clopper. 2015. Variation in glottalization at prosodic boundaries in clear and plain lab speech. Paper presented at 18th International Congress of Phonetic Sciences, Glasgow, Scotland, August 10–14.
- Maddieson, Ian, and Peter Ladefoged. 1985. “Tense” and “lax” in four minority languages of China. *Journal of Phonetics* 13: 433–54. [\[CrossRef\]](#)
- Mazaudon, Martine, and Alexis Michaud. 2008. Tonal contrasts and initial consonants: A case study of Tamang, a ‘missing link’ in tonogenesis. *Phonetica* 65: 231–56. [\[CrossRef\]](#) [\[PubMed\]](#)
- Michaud, Alexis. 2004. A measurement from electroglottography: DECPA, and its application in prosody. Paper presented at Speech Prosody, Nara, Japan, March 23–26. pp. 633–36.

- Mitterer, Holger, Sahyang Kim, and Taehong Cho. 2019. The glottal stop between segmental and suprasegmental processing: The case of Maltese. *Journal of Memory and Language* 108: 104034. [CrossRef]
- Moisik, Scott Reid. 2012. Harsh voice quality and its association with blackness in popular American media. *Phonetica* 69: 193–215. [CrossRef]
- Mooshammer, Christine. 2010. Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in German. *The Journal of the Acoustical Society of America* 127: 1047–58. [CrossRef] [PubMed]
- Pan, Ho-hsien. 2007. The effects of prosodic boundaries on nasality in Taiwan Min. *The Journal of the Acoustical Society of America* 121: 3755–69. [CrossRef]
- Pan, Ho-hsien. 2017. Glottalization of Taiwan Min checked tones. *Journal of the International Phonetic Association* 47: 37–63. [CrossRef]
- Pierrehumbert, Janet, and David Talkin. 1992. Lenition of /h/ and glottal stop. In *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press, pp. 90–117.
- R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Redi, Laura, and Stefanie Shattuck-Hufnagel. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29: 407–29. [CrossRef]
- Ren, Nianqi, and Ignatius G. Mattingly. 1989. Spectral slope as a cue for the perception of breathy and non-breathy stops in Shanghainese. *The Journal of the Acoustical Society of America* 86: S102. [CrossRef]
- Roberts, Brice David. 2020. An Autosegmental-Metrical Model of Shanghainese Tone and Intonation. Ph.D. thesis, University of California, Los Angeles, CA, USA.
- Rose, Philip. 2015. Tonation in three Chinese Wu dialects. Paper presented at 18th International Congress of Phonetic Sciences, Glasgow, Scotland, August 10–14.
- Rothenberg, Martin, and James Mahshie. 1988. Monitoring vocal fold abduction through vocal fold contact area. *Journal of Speech and Hearing Research* 31: 338–51. [CrossRef] [PubMed]
- Selkirk, Elisabeth, and Tong Shen. 1990. Prosodic domains in Shanghai Chinese. *The Phonology-Syntax Connection* 313: 337.
- Seyfarth, Scott, and Marc Garellek. 2020. Physical and phonological causes of coda/t/glottalization in the mainstream American English of central Ohio. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 11: 24. [CrossRef]
- Shen, Xiangrong. 2010. The Acoustic Performances of Glottal Stop. *Studies in Language and Linguistics* 30: 35–39.
- Shue, Yen-Liang, Patricia Keating, Chad Vicenik, and Kristine Yu. 2011. VoiceSauce: A program for voice analysis. Paper presented at 18th International Congress of Phonetic Sciences, Glasgow, Scotland, August 10–14. pp. 1846–49.
- Slifka, Janet. 2006. Some physiological correlates to regular and irregular phonation at the end of an utterance. *Journal of Voice* 20: 171–86. [CrossRef]
- Sonderegger, Morgan. 2020. *Regression Modeling for Linguistic Data*. Cambridge: The MIT Press.
- Stevens, Kenneth N. 1977. Physics of laryngeal behavior and larynx modes. *Phonetica* 34: 264–79. [CrossRef]
- Tehrani, Henry. 2010. EGGWorks, 2010. Available online: <http://phonetics.linguistics.ucla.edu/facilities/physiology/EGG.htm> (accessed on 15 December 2020).
- Tian, Jia, and Jianjing Kuang. 2019. The phonetic properties of the non-modal phonation in Shanghainese. *Journal of the International Phonetic Association* 51: 202–28. [CrossRef]
- Tian, Jia, and Jianjing Kuang. 2020. The phonetic realization of contrastive focus in Shanghainese. Paper presented at 10th International Conference on Speech Prosody 2020, Tokyo, Japan, May 25–28. pp. 265–69.
- Traill, Anthony. 1994. The perception of clicks in !Xóõ. *Journal of African Languages and Linguistics* 15: 161–74. [CrossRef]
- Turk, Alice E., and Stefanie Shattuck-Hufnagel. 2007. Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics* 35: 445–72. [CrossRef]
- Ulrich, Charles H. 1993. The glottal stop in Western Muskogean. *International Journal of American Linguistics* 59: 430–41. [CrossRef]
- Vicenik, Chad, Spencer Lin, Patricia Keating, and Yen-Liang Shue. 2021. Online Documentation for VoiceSauce. Available online: <http://www.phonetics.ucla.edu/voicesauce/documentation/index.html> (accessed on 15 December 2020).
- Wei, Jiuqiao. 2018. A Study of the Tense and Lax Contrast in Daigela Wa. Master's Theses, National University of Singapore, Singapore.
- Xu, Baohua, Zhenzhu Tang, Rujie You, Nairong Qian, Ru-jie Shi, and Ya-ming Shen. 1988. *Shanghai Shiqü Fangyan Zhi [Urban Shanghai Dialects]*. Shanghai: Shanghai Educational Publishing House.
- Yanushevskaya, Irena, Ailbhe Ní Chasaide, and Christer Gobl. 2016. The interaction of long-term voice quality with the realisation of focus. Paper presented at 8th International Conference on Speech Prosody 2016, Boston, May 31–Jun 3, pp. 931–935.
- Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.
- Zee, Eric, and Ian Maddieson. 1979. Tones and tone sandhi in Shanghai: Phonetic evidence and phonological analysis. In *UCLA Working Papers in Phonetics*. Los Angeles: UCLA, vol. 45, pp. 93–129.
- Zhu, Xiaonong, Lei Jiao, Zhicheng Yan, and Ying Hong. 2008. Three ways of Rusheng (入声) sound change. *Studies of the Chinese Language*. 4: 324–38.