



Universiteit
Leiden
The Netherlands

Phantom stress in tone sandhi: evidence from Shaoxing Wu Chinese Wen, X.

Citation

Wen, X. (2026, September 17). *Phantom stress in tone sandhi: evidence from Shaoxing Wu Chinese*. LOT dissertation series. LOT, Amsterdam. Retrieved from <https://hdl.handle.net/1887/4309711>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4309711>

Note: To cite this publication please use the final published version (if applicable).

CHAPTER 1

Introduction

1.1 General background

Tone and stress are two prosodic systems that may co-occur in the same language and interact with each other (de Lacy, 2002; Goldsmith, 1988; Hyman, 2006, 2009).

One widely observed interaction is stress-dependent tonal distribution or realization: tones are usually attracted to stressed positions or preferentially preserved on them.¹ In Lamba (Bantu), for instance, high tones surface on the heads of trochaic feet regardless of their underlying position (Bickmore, 1995; de Lacy, 2002). Similarly, in Zulu and Xhosa (both Bantu), high tones are attracted to the stressed position in the final foot (Goldsmith, 1988). Notably, such tone-to-stress attraction is commonly observed in languages where tones are not lexically specified on every syllable (Brunelle, 2017). In languages

¹ Although the interaction between tone and stress can also be in the opposite direction (i.e., when stress placement is tone-dependent), here we focus primarily on stress-dependent tonal patterns, as they are more commonly employed to account for tonal phenomena in Sinitic varieties, the central concern of this dissertation.

2 *Phantom stress in tone sandhi*

where each syllable carries a lexical tone with multiple levels of tonal contrasts, stressed syllables tend to preserve their underlying tonal contrasts, while unstressed syllables undergo tonal modification or neutralization. For instance, in Thai, stressed syllables preserve their underlying tones, with unstressed syllables undergoing tonal reduction (Potisuk et al., 1994, 1996). Burmese exhibits complete neutralization of tonal contrasts in unstressed syllables (Gruber, 2011).

In Sinitic varieties,² stress has also been considered to condition tonal behaviors, particularly in the processes of tone sandhi (i.e., when a tone undergoes changes due to the effect of adjacent tones). Despite the diversity in the triggers and locations of sandhi application, tone sandhi patterns across Sinitic varieties generally fall into two categories, namely left-dominant and right-dominant ones (M. Y. Chen, 2000; Yue-Hashimoto, 1987; J. Zhang, 2007). In right-dominant sandhi, mainly reported in Mandarin, Min, and Southern Wu, the final syllable preserves its underlying tonal contrast while non-final syllables undergo tonal changes. An example from Xiamen Min is shown in (1): the underlying high level tone of the non-final syllable *p'ang* is changed to a low level tone, while the final syllable preserves its underlying tone (M. Y. Chen, 1987, p. 112). Conversely, in left-dominant sandhi, the initial syllable preserves its underlying tone and determines the sandhi contour, with non-initial syllables undergoing tonal neutralization. This pattern is commonly observed in Northern Wu. In Shanghai Wu, for instance, the non-initial tones are deleted, while the initial tone is preserved and spread towards the following syllables, as exemplified in (2) (B. Xu et al., 1981, p. 152):

(1) Right-dominant (Xiamen Min)

/p'ang ⁴⁴ /	/tsui ⁵³ /	→	[p'ang ²²	tsui ⁵³]
fragrant	water		'perfume'	

²Following Chappell (2001), in this dissertation, we use “Sinitic varieties” as a general term when discussing patterns across the Sinitic branch or when addressing topics of broader scope, but use “dialect” when referring to specific varieties (e.g., Shaoxing Wu dialect) or traditionally recognized groupings (e.g., Wu dialects).

(2) Left-dominant (Shanghai Wu)

- a. /sɿ³⁴/ /ɕin⁵³/ → [sɿ³³ ɕin⁴⁴]
 hand center ‘palm’
- b. /ɕi³⁴/ /tɕiɿ³⁴/ → [ɕi³³ tɕiɿ⁴⁴]
 happy wine ‘wedding ceremony’

The systematic variation in tonal preservation and neutralization aligns well with the tonal behaviors observed in languages with stress-dependent tonal realization. This has prompted researchers to account for the tone sandhi patterns across Sinitic varieties with stress (M. Y. Chen, 2000; Duanmu, 1992, 1993, 1995; Yue-Hashimoto, 1987, among others). For example, Shanghai Wu is regarded as having left-headed word stress; the unstressed (i.e., non-initial) syllables undergo tonal deletion, while the underlying tone of the stressed syllable is preserved and reassociated one-to-one with the unstressed syllables, resulting in the typical left-dominant sandhi (e.g., M. Y. Chen, 2000; Duanmu, 1993, 1995). In contrast, dialects with right-dominant sandhi are proposed to have right-headed word stress (e.g., R. L. Cheng, 1973; Duanmu, 1995; Wright, 1983 on Xiamen Min), with stressed syllables preserving citation tone and unstressed ones undergoing tonal modification. Building on cross-dialectal evidence, Duanmu (1995, 1999, 2000) demonstrated that stress provides a unified account of tone sandhi patterns, particularly in the formation of sandhi domains. On that basis, Duanmu (2005) further advanced a general theory of stress, claiming that all languages have stress and share a common stress-based foot formation mechanism, which plays a pivotal role in tone sandhi processes.

The stress-based account has been influential. Many subsequent studies have adopted it as a working assumption, treating the existence of stress and its interaction with tone as established in their descriptions or analyses of tone sandhi across Sinitic varieties (Hou, 2011; Jiang, 2005; Jiang & Xie, 2011; Xia, 2018, among others). Although first proposed decades ago, the stress-based account has experienced renewed interest in recent years, with an increased number of acoustic investigations of tone sandhi patterns and their relationship to stress (e.g., Guo & Chen, 2022; Hu & Feng, 2024; M. Zhang et al., 2023;

Zhu, 2023). This continued scholarly attention suggests the importance of tone-stress interaction as a topic of research.

1.2 Motivation

Despite the prevalence of stress-based accounts for tone sandhi in Sinitic varieties, two fundamental problems deserve further consideration and motivate this dissertation.

First, existing research on tone-stress interaction mainly focuses on dialects exhibiting typical left- or right-dominant sandhi patterns, which is typologically limited. Over the past decades, a growing body of studies have documented dialects with features of both left- and right-dominant sandhi patterns (e.g., Kuang et al., 2018; Rose & Yang, 2024; Shen & Rose, 2016; Ying, 2025). Although many studies have shown that both sandhi dominant patterns can co-exist within the same dialect, they typically occur in different morphosyntactic structures (e.g., Duanmu, 1995; B. Xu et al., 1981, on Shanghai Wu), which has been proposed to reflect stress of different levels (see further discussion in this chapter and in Chapter 4). Unlike such structure-sensitive cases, in these dialects, left- and right-dominant sandhi can occur within words. The relationship of these complex sandhi patterns to stress remains underexplored, leaving a critical gap given the widespread appeal of stress-based accounts.

Recent acoustic studies on previously under-studied dialects, particularly on Wu dialects, provide converging evidence that the dichotomy between left- and right-dominant sandhi is less clear-cut than traditionally assumed. In Southern Wu, which are conventionally considered as undergoing right-dominant sandhi, the final syllables do not always preserve their underlying tones. For instance, in Maodian Wu, the tonal realizations of the final syllables are conditioned by the initial tones in disyllabic words ending with mid level tones (Shen & Rose, 2016). In Northern Wu, despite the general left-dominant tendency, non-initial tones have been repeatedly observed to condition sandhi contours as well. For example, in Xinzhuang and Zhuji Wu, the sandhi contours of disyll-

labic words starting with high falling tones are conditioned by the non-initial tones (Kuang et al., 2018). These findings collectively suggest that left- and right-dominant sandhi patterns can coexist within the same dialect.

Notably, such dialects present theoretically important test cases for stress-based accounts: if stress truly underlies tone sandhi patterns across Sinitic varieties, these complex patterns where sandhi dominance varies with tonal context should also be compatible with stress-based analysis. However, these dialects have not received due attention in the discussion on tone-stress interaction. As reviewed earlier, the majority of existing studies focus on dialects with typical left- or right-dominant sandhi patterns. While some researchers have attempted to extend stress-based accounts to complex sandhi patterns (e.g., Rose & Yang, 2024), such analyses encounter the second fundamental problem.

The second problem concerns the lack of independent evidence for identifying stress in Sinitic varieties. Unlike other tonal languages with stress, the purported stress in many Sinitic varieties lacks supporting evidence beyond tonal patterns themselves. This leads to a circular reasoning where tonal preservation and neutralization are used to identify stressed and unstressed syllables, while stress is at the same time invoked to explain these same tonal patterns.

Stress is generally realized through various acoustic parameters: stressed syllables tend to exhibit longer duration (e.g., Arvaniti, 2000; Fry, 1955; Sluijter & Van Heuven, 1996), higher mean f_0 or greater f_0 change (e.g., Dogil & Briony, 1999; Plag et al., 2011; Vogel et al., 2016), smaller intensity reduction towards higher frequencies (e.g., Sluijter & Van Heuven, 1996), or more peripheral vowel space (e.g., De Jong & Zawaydeh, 1999; Prieto & Ortega-Llebaria, 2006; Remijsen & Van Heuven, 2005) than their unstressed counterparts (see Gordon & Roettger, 2017, for a cross-linguistic survey). In tonal languages with well-attested stress systems, although f_0 also encodes lexical tone, stress can usually be identified through acoustic correlates independent of f_0 . For example, in Thai, duration serves as the primary acoustic correlate of stress (Potisuk et al., 1996); in Pirahã, stress is realized predominantly through duration (Everett, 1998). Even when stress lacks clear acoustic correlates, its

existence can still be justified through systematic phonological patterns. For instance, in Lamba, high tones consistently surface on odd-numbered syllables within a string of level-1 prefixes (i.e., object and certain tense/aspect markers preceding verbal roots), suggesting a rhythmic effect (Bickmore, 1995; de Lacy, 2002). The systematic alternating pattern can justify positing stress even without robust phonetic cues, as metrical analysis provides coherent accounts of otherwise random phonological processes.

However, such independent evidence is largely absent in many Sinitic varieties.³ Previous studies claiming stress in particular Sinitic varieties often rely on auditory impressions from fieldworkers or native speakers, typically based on small sample sizes (e.g., Zeng & Niu, 2006; Zhong, 2003). The reliability of these perceptual judgments concerning stress thus remains questionable. In addition, as dynamic pitch movement is considered important to stress perception (e.g., Rietveld & Gussenhoven, 1985; 't Hart et al., 1990), there may be a perceptual bias toward interpreting syllables with pitch movement, typically those preserving underlying tones, as stressed. This raises the possibility that perceived stress is not the cause of tonal preservation but rather its consequence.

While recent years have witnessed increased acoustic investigation of stress in Sinitic varieties, many studies have not sought **independent** evidence for stress, but presuppose tone-stress interaction and explore acoustic correlates of stress on that basis instead (e.g., Guo & Chen, 2022; Hu & Feng, 2024; M. Zhang et al., 2023). For instance, Guo and Chen (2022) examined two dialects with left-dominant sandhi (i.e., Chengdu Mandarin and Changsha Xiang) and two with right-dominant sandhi (i.e., Fuzhou Min and Xiamen Min) as representatives of dialects with different metrical structures. They investigated f0 re-

³One might point to Mandarin as a case providing such evidence. In Mandarin, some syllables surface with a reduced pitch contour whose realization depends on the preceding tone, a phenomenon known as neutral tone (e.g., Cao, 1986; Chao, 1968; Lin & Yan, 1980; see also Y. Chen, 2015, for a review). In addition to tonal reduction, neutral tone syllables often exhibit shorter duration and vowel reduction compared to their full tone counterparts (e.g., Chao, 1968; C.-C. Cheng, 1973), which has led many researchers to regard them as unstressed. This case, however, is not comparable to languages with a well-attested stress system, as neutral tone occurs in only a small proportion of words in Mandarin (e.g., W. Li, 1981). Other than in these words, stress is not readily observable in Mandarin.

alization and duration of each syllable in disyllabic words, and found that f_0 is a more consistent correlate of stress than duration across these dialects: duration patterns varied across dialects, whereas stressed syllables consistently surfaced with their underlying pitch forms while unstressed syllables surface with levelish contours. However, since sandhi dominance was used to diagnose stress in the first place, this finding does not resolve the circular argumentation but rather reinforces it.

It is worth noting that the stress discussed above is lexical stress, which should be distinguished from focus-induced prominence (i.e., marking new or contrastive information at the utterance level). In Sinitic varieties, focus is commonly signaled through f_0 (typically via on-focus expansion and/or post-focus compression), longer duration, and/or higher intensity, though the detailed implementation varies across dialects (e.g., Y. Chen, 2006; Jin, 1996; Shih, 1988; Y. Xu, 1999, on Standard Chinese; Y. Chen, 2009; B. Ling & Liang, 2017, on Shanghai Wu; Gu & Lee, 2007; Wu & Xu, 2010, on Cantonese Yue). In addition, the distinctiveness of lexical tones may also be enhanced under focus (e.g., Y. Chen, 2009; Y. Chen & Gussenhoven, 2008). While a full discussion of the prosodic encoding of focus and its interaction with tonal realization is beyond the scope of this dissertation, readers may refer to Y. Chen et al. (2014) for some relevant discussion.

Admittedly, stress can still be justified through its systematic effects on tonal patterns across different domains (i.e., from words, compounds, to phrases), despite the lack of clear-cut phonetic evidence. Consider Shanghai Wu as an example. As reviewed in Section 1.1, in Shanghai Wu, the underlying tones of word-initial syllables are usually preserved and spread rightward with non-initial underlying tones deleted, which are claimed to be stress-conditioned. Specifically, tone sandhi domain is regarded as a left-headed stress foot, within which only stressed syllables can preserve their underlying tones (e.g., M. Y. Chen, 2000; Duanmu, 1993, 1995). The stress-based account is further supported by asymmetric sandhi domain formation in compounds with different structures: 1+2 compounds (e.g., [red [wood bed]], ‘red wooden-

bed') consistently form one sandhi domain, while 2+1 compounds (e.g., [[red wood] bed], 'red-wood bed') form either one or two sandhi domains (Duanmu, 1993), which can be explained with metrical processes such as stress assignment, stress clash resolution, and stress reduction (see M. Y. Chen, 2000, for more detailed analysis). Within each sandhi domain, left-dominant sandhi applies consistently (M. Y. Chen, 2000). More importantly, tonal realizations at the phrasal level provide additional evidence. In contrast to the left-dominant sandhi observed in words and compounds, disyllabic phrases undergo right-dominant sandhi: the final syllable preserves its underlying tone while the initial syllable surfaces with a level contour (e.g., B. Xu et al., 1981). Building on cross-linguistic evidence that phrasal stress assignment differs from that of words and compounds (e.g., Cinque, 1993), Duanmu (1995) proposed right-headed phrasal stress in Shanghai Wu, which well explains the right-dominant sandhi pattern observed at the phrasal level. Taken together, stress of different levels can provide consistent accounts of tonal realizations across words, compounds, and phrases. This justifies the postulation of stress even in the absence of robust acoustic correlates.

Nonetheless, such comprehensive investigations are limited to dialects with typical left- or right-dominant patterns. For those with complex tone sandhi patterns, independent evidence for stress remains lacking, even when these patterns are analyzed in terms of stress-induced tonal neutralization and preservation. For example, Jinshan Wu is a Northern Wu reported to exhibit both left- and right-dominance depending on tonal context, and Rose and Yang (2024) posited underlying left- and right-headed stress for the respective cases to account for the varying sandhi dominance. Yet, as Rose and Yang (2024) reports only f_0 data, acoustic evidence for stress beyond tonal realization is absent. Furthermore, there has been no systematic investigation across various domains comparable to those conducted for Shanghai Wu. Without such evidence, it is hard to determine whether the variable sandhi patterns genuinely reflect context-dependent stress assignment, or whether they arise from processes unrelated to metrical structure. The issue of circular argumentation thus remains

unresolved for these theoretically crucial cases.

To advance our understanding of tone-stress interaction in Sinitic varieties, it is essential to address these two fundamental problems. More attention should be paid to dialects exhibiting complex sandhi patterns, with particular focus on seeking acoustic evidence for stress beyond tonal realizations and, more importantly, investigating whether stress can provide consistent accounts of sandhi patterns across different domains.

1.3 The current studies

This dissertation explores the potential interaction between tone and stress in Shaoxing Wu Chinese (绍兴话), a dialect exhibiting complex tone sandhi patterns, with controlled experiments investigating tonal realizations across different domains. Before delving into the details of the current studies, a brief overview of this dialect and its tonal system is provided.

Shaoxing Wu Chinese is a Wu dialect (吴语, ISO 639-3; code: wuu) spoken in Yuecheng district (越城区),⁴ which belongs to the prefectural-level municipality of Shaoxing city (绍兴市) in Zhejiang province (浙江省), the People's Republic of China (see Figure 1.1 for the location of Shaoxing city). This dialect belongs to the Suhujia cluster (苏沪嘉小片) within the Tai Lake subgroup (太湖片) of the Northern Wu dialect group (北部吴语), a Sinitic branch within the Sino-Tibetan family (Wurm et al., 1987).

There are eight lexical tones in Shaoxing Wu (hereafter labeled T1 to T8), with four different contours in both high and low pitch registers (e.g., Wang, 1959, 2008; J. Zhang, 2006). Figure 1.2 illustrates the f₀ trajectories of the eight lexical tones produced in isolation by a female speaker born in the 1960s. T1 is a high-register tone falling from the higher to the lower end of the speaker's pitch range. T2 is a low-register falling tone, starting in a low register with a rise towards the middle, and falling to the lower end. T3 is a high-register rising tone with a slightly falling contour at the start. T4 is a low-register rising

⁴According to the census in 2020, there are around 1,020,000 residents in Yuecheng district.



Figure 1.1: Map showing the location of Shaoxing. The locations of Beijing and Shanghai are also indicated as reference.

tone. T5 is a high-register level tone with a slightly falling contour. T6 is a low-register level tone. T7 and T8 are rising tones in high and low registers with relatively short durations. High-register tones (T1, T3, T5, T7) co-occur exclusively with voiceless onsets, while low-register tones (T2, T4, T6, T8) only with voiced onsets. T1 to T6 only occur in open syllables or syllables with a nasal coda, while T7 and T8 only occur in syllables closed by a glottal stop.

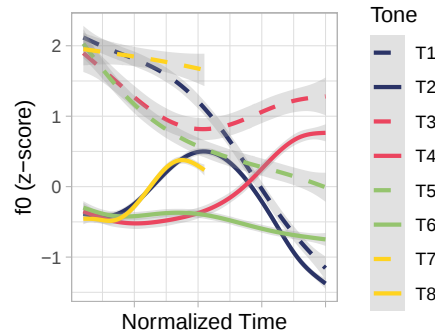


Figure 1.2: Lexical tones in Shaoxing Wu.

Despite being classified as a Northern Wu dialect, Shaoxing Wu does not exhibit the typical left-dominant sandhi pattern characteristic of this dialect group. According to impressionistic studies, sandhi patterns of disyllabic words can be conditioned by the underlying tones of both syllables depending on tonal

context (Wang, 1959, 2015, see Chapter 2 for more details). These patterns, like those recently documented in other Wu dialects, exhibit features of both left- and right-dominant sandhi, and their relationship with stress deserves further investigation.

What makes Shaoxing Wu more intriguing is its tone sandhi patterns in trisyllabic compounds. Impressionistic studies suggest that they are solely determined by the initial underlying tone (Wang, 2015; J. Zhang, 2006). This contrasts with the patterns observed in disyllabic words, raising the question of whether stress can offer consistent accounts of tone sandhi across different domains. If tonal preservation and neutralization in Shaoxing Wu are stress-driven, disyllabic sandhi patterns seem to suggest that stress location varies depending on tonal context, whereas trisyllabic sandhi patterns suggest consistent initial stress. It is worth investigating whether these seemingly contradictory patterns are accurately described in impressionistic studies, and if so, what underlies such inconsistency.

Furthermore, tonal realizations at the phrasal level in Shaoxing Wu show more intricacy than those typically observed in Northern Wu. While phrases in Shaoxing Wu generally display right-dominant sandhi, impressionistic descriptions reveal that disyllabic phrases with verb-object (VO) and subject-predicate (SP) structures differ in their tonal realizations of initial syllables (Wang, 2015). Such VO-SP distinction is rarely reported in other dialects, and its underlying mechanism remains unclear. Given that right-dominant phrasal sandhi is often cited as evidence for tone-stress interaction (e.g., M. Y. Chen, 2000; Duanmu, 1995), the added complexity in Shaoxing Wu presents an important opportunity to evaluate whether stress-based accounts can accommodate this distinction, or whether alternative analyses are required.

Taken together, these complex patterns suggest Shaoxing Wu as an interesting test case for investigating the relationship between tone and stress. It should be noted, however, that existing observations are based primarily on impressionistic descriptions for documentation purposes. Therefore, these patterns require further verification through acoustic analysis, and systematic ex-

perimental investigation is needed to understand whether stress can offer consistent accounts of sandhi patterns across different domains.

Against this background, the current studies addressed the following research questions with controlled experiments. To establish an empirical foundation, we first examined whether left- and right-dominant sandhi patterns co-exist at the word level (RQ1). As an initial attempt of seeking independent evidence of stress, we also investigated whether words with different sandhi dominance differ in their durational patterns (RQ2). Building on these findings, we then examined whether stress can consistently account for sandhi patterns across different domains. At the compound level, we investigated whether trisyllabic sandhi patterns align with stress-based predictions derived from disyllabic patterns. Since asymmetric sandhi domain formation between 1+2 and 2+1 compounds has been cited as supporting evidence for stress in Shanghai Wu, yet remains underexplored in Shaoxing Wu, we examined whether and how these compound structures differ in sandhi domain formation (RQ3). We further investigated tonal realizations in trisyllabic compounds, specifically whether non-initial tones are preserved in some tonal contexts as stress-based accounts would predict, or neutralized as impressionistic descriptions have suggested (RQ4). At the phrasal level, we focused on the reported VO-SP distinction. We examined whether this distinction can be confirmed acoustically and discussed whether it can be accommodated within stress-based accounts (RQ5). We further investigated whether the distinction extends to phrases with disyllabic initial constituents, which could shed light on its underlying nature (RQ6).

Through acoustic analysis of tone sandhi across words, compounds, and phrases in Shaoxing Wu, and systematic investigation of whether stress of different levels can consistently account for these patterns, this dissertation seeks independent evidence for stress in a dialect with complex sandhi, thereby addressing both the typological and methodological limitations of previous research. The findings will advance our understanding of the generalizability of stress-based accounts and offer broader insights into the mechanisms underlying complex tone sandhi patterns.

1.4 Outline

This dissertation investigates whether stress-based accounts can adequately explain the complex sandhi patterns in Shaoxing Wu Chinese. Specifically, it progressively evaluates whether stress can account for tone sandhi in words, compounds, and phrases, and if not, consider alternative analyses.

Chapter 2 lays the empirical foundation by investigating disyllabic tone sandhi at the word level. It first examines whether left- and right-dominant sandhi patterns coexist (RQ1), then explores whether different dominance types show systematic durational differences that could serve as independent phonetic correlates of stress (RQ2). Building on these results, Chapter 3 turns to trisyllabic compounds. It investigates whether sandhi domain formation in 2+1 vs. 1+2 structures differs in ways predicted by stress-based analyses (RQ3), and whether tonal realizations, particularly the preservation or neutralization of non-initial tones, conform to stress-based expectations (RQ4). Chapter 4 shifts the focus to the phrasal level, examining the reported VO-SP distinction. It first tests acoustically whether this distinction is robust (RQ5), and then evaluates whether it extends to phrases with disyllabic initial constituents (RQ6), thereby probing its theoretical implications. Finally, Chapter 5 synthesizes the findings across different domains, discusses their theoretical implications, and outlines directions for future research.