Tone 3 Sandhi in Huai'an Mandarin Applies Based on Cyclicity and Binarity: An Experimental Study

Naiyan Du and Karthik Durvasula
Baiko Gakuin University and Michigan State University

1. Introduction

We propose that Tone 3 Sandhi (Low Tone Sandhi) in Mandarin languages (Huai'an Mandarin in this case; Huai'an hereafter) applies cyclically based on syntactic structures (Cyclicity), and the prosodic unit has to be 2-syllable long (Binarity). The predicted patterns are observed in production experiment that we conducted. The Tone 3 Sandhi process is shown in (1), which states that a Tone 3 undergoes tone sandhi to become a Tone 2 (rising tone) with reference to another Tone 3 to the immediate right. For conciseness, in examples, we will use '3' to refer to Tone 3 and '2' to refer to Tone 2 hereafter.

(1) Tone 3 Sandhi (Low Tone Sandhi) in Mandarin Languages

Tone 3 (Low Tone) + Tone 3 (Low Tone) → Tone 2 (Rising Tone) + Tone 3 (Low Tone)

As one of the few productive phonological processes across Mandarin Languages, Tone 3 sandhi has enjoyed great spotlight not only because of its complicated patterns, but also because of its potential to testify important theories in phonology and syntax. For example, Tone 3 Sandhi patterns in Standard Mandarin have been used to show that prosodic constituents should match syntactic constituents in every detail if possible (Shih, 1986, 1997; Chen, 2000), which is later formalized by Selkirk (2011) as Match Theory. Within another theoretical framework, Duanmu (2007) proposes a novel theory that syntactic Nonhead must bear stress, and Tone 3 Sandhi application domains in Standard Mandarin have been shown to form according to these Nonhead Stresses.

The current study focuses on Huai'an, which belongs to the Jianghuai Guanhua Group (Lower Yangtze Mandarin) of the Mandarin language family. Native speakers are mainly from Huai'an city, which is located in the northern part of Jiangsu Province (Li, 1989). Through the current experimental study, we confirmed that in Huai'an, Tone 3 Sandhi can be explained by Cyclicity (Chomsky et al., 1956), which means Tone 3 Sandhi applies from the smallest syntactic units and move to larger units step by step. It is worth noting that the patterns in Huai'an can also be perfectly explained by Match Theory, which is developed under Optimality Theory framework (Prince & Smolensky, 1993). For simplicity, the current proposal is only made within a rule-based framework. Moreover, we confirmed that Binarity does play an important role in Mandarin Languages, as formalized with a Binarity constraint within an Optimality Theory framework by Chen's (2000).

2. Issue

One big issue with previous studies on Mandarin Tone 3 Sandhi is that theories are usually based on much debated basic data. To give one example from Standard Mandarin, for the right-branching

utterance in 0, Chen (2000) assumes the only possible surface representation to be '3 2 3' while Duanmu (2007) allows both '3 2 3' and '2 2 3'. Another disputed case in Standard Mandarin is whether a sequence of consecutive Tone 3s is allowed on the surface. Duanmu (2007) suggests such pattern as in SR2 in 0 is possible, while Chen (2000) and Wang and Lin (2011) reject the possibility of adjacent Tone 3s in the surface under any circumstances.

(2) a. Left-branching
[[mai xau] tɕiou]
buy good wine
'finished buying wine'

UR 3 3 3
SR1 2 2 3
SR2 2 3 3?

b. Right-branching
[[mai xau] tɕiou]
buy good wine
'to buy good wine'

UR 3 3 3
SR1 3 2 3
SR2 2 2 3?

One possible reason for the discrepant empirical observations is simply language variation. The so-called 'Standard Mandarin', although officially defined by Chinese Government (1958) as the language used in Luanping County of Hebei Province, should have many variations considering the span of Chinese territory and the large amount of native speakers (Ramsey, 1989). Also, the participants of previous studies may be bilingual, namely they are native speakers of both Standard Mandarin and another local Mandarin variation, and these two languages may affect each other.

A second possible reason for the discrepant empirical observations is that some of the observed tonal patterns (in this case a sequence of Tone 3s) are actually due to non-linguistic factors and production errors. Many previous researchers have noticed that Tone 3 Sandhi domain may be influenced by speech rate (Shih, 1986, 1997; Chen, 2000; Cheng, 1973; inter alia). According to these researchers, for right-branching 3-syllable utterance in (2b), '2 2 3' becomes normal under accelerated speech rate. So, it is possible that at slower speech rates, there is a prosodic boundary between the first two underlying Tone 3s, thereby allowing the existence of '3 2 3' in the surface.

To solve these issues and to move this field forward in a productive fashion, the scope of the current study is limited to Huai'an. The participants come from a small geographical area, which makes sure that the speakers actually share the same language. A controlled production experiment is employed to eliminate the interference of non-linguistic factors as much as possible.

3. The Production Experiment

3.1 Participants

We recruited 19 native speakers of Huai'an Mandarin through personal relationships in Huai'an City. The age range was from 27 to 69 years. Among them, nine self-identified as female, and ten as male. All the participants were born and raised in Huai'an City. None of them had participated in any linguistic studies before or heard about the concepts of Tone 3 Sandhi domain or Syntax-Phonology Interface.

3.2 Stimuli

To further simplify the picture and get controlled data for theoretical analysis, test items consist of only right-branching utterances. The length ranges from 3-syllable long to 6-syllable long. The underlying

syllables are always Tone 3. Moreover, each word forms a separate phrase by itself, therefore Tone 3 Sandhi only applies at the post-lexical level.

Each participant produced two repetitions of 32 test and 61 filler sentences at a natural speech rate. Therefore each participant read a total of 186 sentences. Also, each participant was assigned with a different randomized stimuli.

3.3 Procedure

The entire experiment was conducted in Huai'an city. Each participant was recorded by the first author using Audacity (Audacity Team, 2024) and a Popu Line BK USB microphone on an Apple laptop in a quiet room that was located in the participant's home or workplace. The participants were not informed of the real research question. Instead, they were told that the purpose of the study was to collect some general information on Huai'an. In post-experiment interviews, none of the participants reported noticing utterances with all tones being Tone 3, or that tones were the real focus of the study. The participants were instructed to use their everyday voice and read at normal speech rate. The participants were also encouraged to read through the stimulus list to be familiar with the reading materials before producing them. All stimuli were presented in Chinese characters.

3.4 Measurement

Using Praat (Boersma & Weenink 2021), the recordings were manually annotated by the first author, who is a native speaker of Huai'an. The penultimate syllable is marked for the purpose of a different study, an example is shown in Figure **1**. The underlying representation is marked in the third tier and the surface representation is indicated in the fourth tier. The first tier uses an IPA-like transcription to record the vowel of the penultimate syllable, the second tier shows the segments in Pinyin (the phonetic alphabet for Mandarin Languages), and the quality of the recording is documented in the last tier. Tokens with speech errors or recording issues are not marked as 'good' and are excluded from the analysis. 67 out of 1216 tokens are excluded, which accounts for 5.51% of the total.

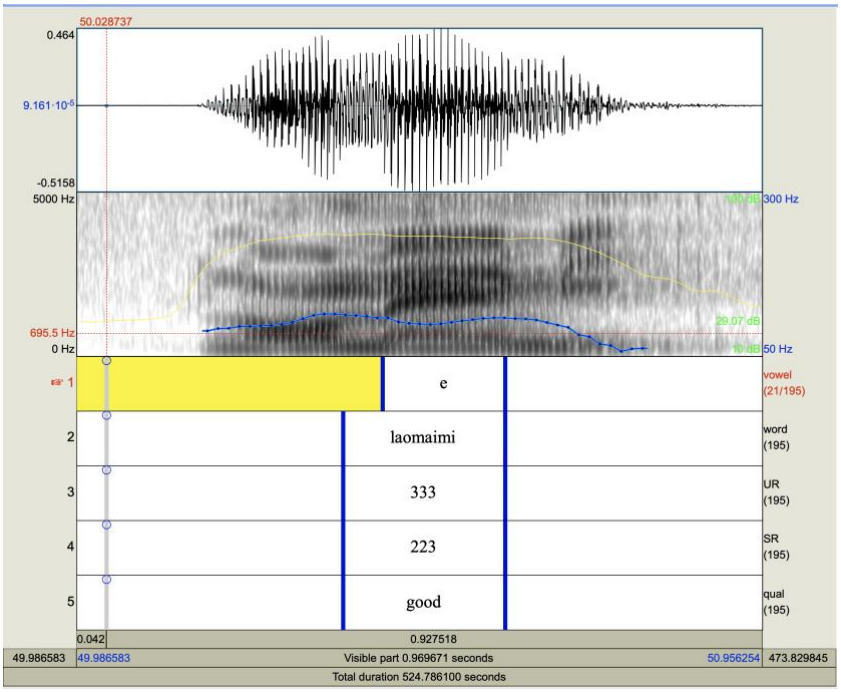

Figure 1: Annotation Scheme of Experiment

3.5 Predictions

Our proposal is stated again as follows: Tone 3 Sandhi applies cyclically based on syntactic structures (Cyclicity), and the prosodic unit has to be 2-syllable long (Binarity). Based on these, the predictions should be straightforward. The derivations for utterances from 3-syllable long to 6-syllable long are shown in (3).

(3) a. [tsuŋ [ta tɕin]]
always dig well
'(Someone) always digs a well.'

UR 3 3 3
3 (**2** 3) Cycle 1
(3) (2 3) Cycle 2
SR 3 2 3

b. [o [tsuŋ [ta tɕin]]]
1sg always dig well
'I always dig wells.'

UR 3 3 3 3
3 3 (**2** 3) Cycle 1
3 (3) (2 3) Cycle 2
(**2** 3) (2 3) Cycle 3
SR 2 3 2 3

c. [o [tsuŋ ɕiæ̃ [ta tɕin]]]
1sg always want dig well
'I always want to dig wells.'

UR 3 3 3 3 3
3 3 3 (**2** 3) Cycle 1
3 3 (3) (2 3) Cycle 2
3 (**2** 3) (2 3) Cycle 3
(3) (2 3) (2 3) Cycle 4
SR 3 2 3 2 3

d. [o [tsuŋ ɕiæ̃ li [ta tɕin]]]
1sg always want 2sg dig well
'I always want you to dig wells.'

UR	3 3 3 3 3 3	
	3 3 3 3 (**2** 3)	Cycle 1
	3 3 3 (3) (2 3)	Cycle 2
	3 3 (**2** 3) (2 3)	Cycle 3
	3 (3) (2 3) (2 3)	Cycle 4
	(**2** 3) (2 3) (2 3)	Cycle 5
SR	2 3 2 3 2 3	

The Underlying Representations and corresponding predicted Surface Representations are summarized in Table **1**.

Underlying Representation	Predicted Surface Representation
3 3 3	3 2 3
3 3 3 3	2 3 2 3
3 3 3 3 3	3 2 3 2 3
3 3 3 3 3 3	2 3 2 3 2 3

Table 1: Underlying Representations of Corresponding Predicted Surface Representations

3.6 Results

All data analyses in this article were performed in R (R Core Team, 2025) using the tidyverse suite of packages (Wickham et al., 2019). The data analysis was done using the lme4 package (Bates et al., 2021). The results for utterances of different lengths will be discussed separately in this section.

The results for 3-syllable utterances are summarized in Figure **2**. The x-axis lists all the surface representations, and the y-axis represents the mean of proportions, which is calculated as the average proportion of a surface representation across all speakers. The 'Sub' on top of each bar represents the number of speakers who actually produced this surface representation in the experiment, and the 'C' represents the count (number of tokens). The syntactic structures are shown on the top. All the following figures follow the same format. Two syntactic structures are tested for 3-syllable utterances, both are pro-drop. Based on Figure **2**, it is quite clear that '3 2 3' is the dominant surface representation for 3-syllable right-branching utterance, which verifies the prediction. The few exceptions may simply be assigned to speech error, potentially related to the introduction of higher prosodic boundaries due to pausing --- a possibility that needs further experimental testing.

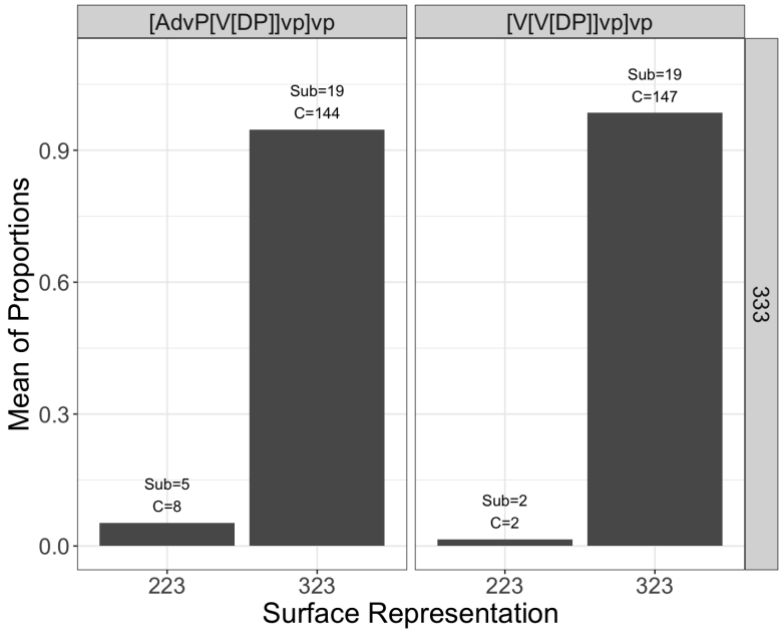

Figure 2: Surface Representations and Their Means of Proportions for 3-syllable Stimuli

The results for 4-syllable utterances are summarized in Figure **3**. Again, two syntactic structures are tested. For the syntactic structure shown in the left part of Figure **3**, the last syllable itself functions as the syntactic object, and '2 3 2 3' is the only dominant surface representation as predicted. Speech error may again account for other observed surface forms. However, for the other syntactic structure where the last 2 syllables function as the object, multiple surface presentations get enough tokens in production as shown in the right part of Figure **3**. To explain the existence of '3 2 2 3', we propose that there is a tendency for verb and its complement to be in the same prosodic unit (in this case, phonological phrase). In this case, such tendency causes the last 3 syllables to be grouped together. There are many morpho-syntactic constituents in Mandarin Chinese that can be analyzed as verb plus complement. However, many native speakers of Mandarin Chinese tend to ignore the inner structures of these constituents and treat them as single words (Huang & Liao, 2012). For example, the word 'to eat' is usually translated as /ʈʂɨ fan/ in Mandarin. The /ʈʂɨ/ part means 'to eat' and the /fan/ part means 'meal'. Although it may be more plausible to be analyze /ʈʂɨ fan/ as a verb phrase, many native speakers report that /ʈʂɨ fan/ is a single word. Such a phenomenon usually happens when both the verb and its complement are monosyllabic and have a high chance of co-occurrence. However, such a phenomenon may indicate a general tendency, and as a result, when forming Tone 3 Sandhi domains, a verb and its complement tend to be grouped together. Inside any prosodic units, Tone 3 Sandhi simply applies from left to right whenever possible. In this case, the last 3 syllables form a prosodic unit and the first syllable forms a prosodic unit by itself, which generates '3 2 2 3'. Moreover, the effect of minimal word may also play a role (McCarthy & Prince 1986). In the case of Mandarin Chinese, prosodic unit has to be 2-syllable long, which means monosyllabic prosodic unit tends to be avoided (Chen, 2000; Zhou, 1964). Therefore, when the last 3 syllables form a prosodic unit, the first syllable tends to join. When this happens, all 4 syllables form a prosodic unit and Tone 3 Sandhi simply applies from left to right to generate '2 2 2 3'.

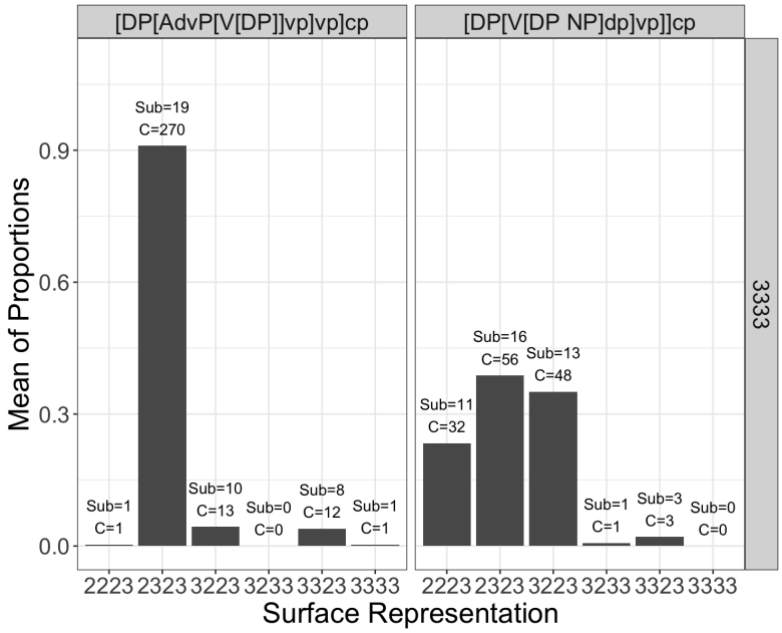

Figure 3: Surface Representations and Their Means of Proportions for 4-syllable Stimuli

The results for 5-syllable utterances are summarized in Figure **4**. Only one syntactic structure is tested, and two surface representations are possible, namely '2 2 3 2 3' and '3 2 3 2 3'. Among them, only '3 2 3 2 3' is predicted by our proposal. Again, minimal word effect may participate in and cause the first syllable to join the adjacent prosodic domain. As a result, the first 3 syllables are grouped together and Tone 3 Sandhi applies from left to right inside prosodic domains to generate '2 2 3 2 3'.

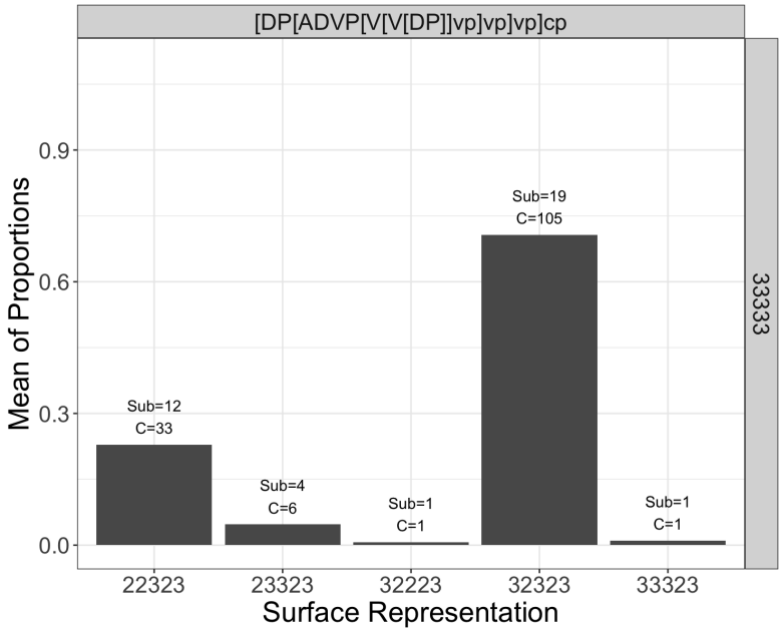

Figure 4: Surface Representations and Their Means of Proportions for 5-syllable Stimuli

The results for 6-syllable utterances are summarized in Figure **5**. Only one syntactic structure is tested but the picture is rather complicated. 4 surface representations, including '2 2 2 3 2 3', '2 3 2 3 2 3',

'3 2 2 3 2 3' and '3 2 3 3 2 3', need to be explained. Among these, only '2 3 2 3 2 3' is predicted by our proposal. For other observed patterns, the explanation may lie in speakers failing to process the syntactic structure of the utterances. Both utterances with only Tone 3 syllables and utterances with each syllable forming a separate phrase by itself are rare in Mandarin Chinese. When such utterances extend to 6-syllable long, speakers may fail to process the sentences in the experiment and apply Tone 3 Sandhi whenever possible. This may explain the pattern in the first 4 syllables of '2 2 2 3 2 3'. Also, boundaries may be placed unexpectedly due to pause in processing. With this, patterns like '3 2 2 3 2 3' and '3 2 3 3 2 3' can be explained.

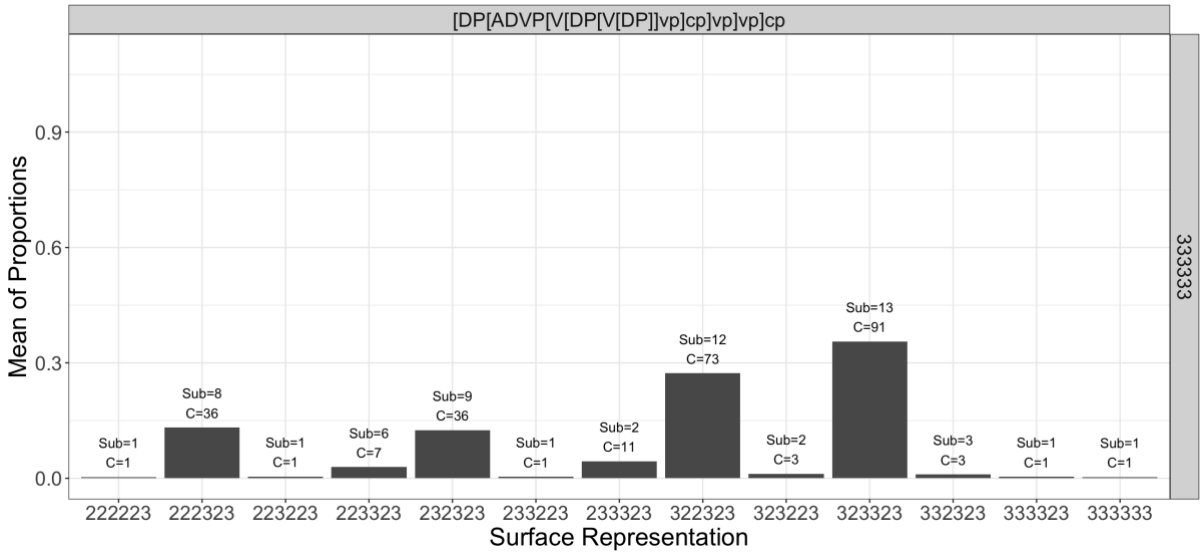

Figure 5: Surface Representations and Their Means of Proportions for 6-syllable Stimuli

To summarize, through the production experiment, all the predicted surface representations are testified. Moreover, other surface representations are also observed in actual production. Herein lies a potential explanation for why researchers have disagreed on the data. One is able to observe complex tone patterns in more complex sentences. However, such patterns are almost impossible to observe in simpler sentences. This suggests to us that the complex tonal patterns appear in more complex sentences due to non-linguistic or performance factors related to the effect of Minimal Word, the tendency to group verb and complement together, and the tendency for speakers to be unsure of the syntactic structures in experiments involving more complex sentences.

4. Limitation of the Current Study

It is worth noting that in order to avoid the issue of the variation of 'Standard Mandarin', we focused on Huai'an for the current study. However, we believe our claims based on Cyclicity and Binarity can (and likely should) be generalized to Tone 3 Sandhi patterns in all Mandarin languages. Future studies on other Mandarin languages are needed to support this generalization.

Also, although all the observed patterns are accounted for, we would like to suggest that our results are preliminary and further additional evidence from other sources is necessary to firmly establish the claim, as claims based on any single source of data are always subject to (unknown) confounding factors. We leave this for future work.

5. Conclusion

To eliminate interfering factors as much as possible, a careful production experiment was conducted on right-branching Huai'an utterances varying in syllable counts from 3 to 6 syllables. Overall, a succinct account of the observed patterns is provided based on Cyclicity and Binarity.

Finally, we would like to reiterate our claim that the discrepant claims about Tone 3 sandhi patterns in the extant literature might be due to the fact that researchers are depending on different sorts of structures to convince themselves of the possibility of a certain potential surface tonal pattern. In general, we would like to recommend the use of very short sequences to establish the domain of application of tone sandhi in order to avoid non-linguistic (processing) factors from affecting speaker productions/intuitions. When we focused on 3-syllable utterances, underlying Tone 3 sequences almost exclusively surfaced as '3 2 3' sequences, suggesting that '2 2 3' is not really an option, and therefore Cyclicity and Binarity are the relevant phonological restrictions on Tone 3 Sandhi.

References

Audacity Team (2024). *Audacity. Version 3.4.2*, retrieved 2 August 2024 from https://www.audacityteam.org/.

Bates, Douglas, Martin Mächler, Ben Bolker & Steven Walker (2021). *Lme4: linear mixed effects models using 'Eigen' and S4*. https://CRAN.Rproject.org/package=lme4.

Chen, Matthew. Y. (2000). *Tone sandhi: Patterns across Chinese dialects*. Cambridge University Press.

Cheng, Chin-Chuan. (1973). *A synchronic phonology of Mandarin Chinese*. The Hague: London.

Chinese Government (1958). *Hànyǔ pīnyīn fāngàn* [The Scheme of Chinese Phonetic Alphabet].

Chomsky, Noam. Morris Halle & Fred Luckoff. (1956). On accent and Juncture in English. In Morris Halle, Horace Lunt, Hugh MacLean & Cornelis van Schooneveral (eds.) *For Roman Jakobson*. The Hague: London. 65–80.

Duanmu, San. (2007). *The phonology of standard Chinese*. Oxford University Press.

Huang, Borong & Xudong Liao (2012). *Xiàndài hànyǔ* [Contemporary Chinese]. Higher Education Press.

Li, Rong (1989). Hànyǔ fāngyán de fēnqū [The geographic division of Chinese dialects]. *Fāngyán* [Dialect] *4*. 241–259.

McCarthy, John & Alan Prince. (1986). *Prosodic Morphology*. Manuscript. University of Massachusetts and Brandeis University.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. New Brunswick: Rutgers University.

R Core Team (2025). R: a language and environment for statistical computing. Version 1.4.1106, retrieved 16 April 2025 from https://www.posit.co.

Ramsey, S. Robert (1989). *The languages of China*. Princeton, NJ: Princeton University Press.

Selkirk, Elisabeth. (2011). The syntax-phonology interface. In John A. Goldsmith, Jason Riggle & Alan C. L. Yu (eds.) *The Handbook of Phonological Theory. 2nd edition*. Blackwell Publishing. 435–485.

Shih, Chi-Lin. (1986). *The Prosodic Domain of Tone Sandhi in Chinese (Phrasal Phonology, Tonal Typology, Mandarin, Syntax-Phonology Interface)*. PhD diss. University of California, San Diego.

Shih, Chi-lin. (1997). Mandarin third tone sandhi and prosodic structure. *Linguistic Models*, *20*, 81-124.

Wang, Chiung-Yao & Yen-Hwei Lin. (2011). Variation in Tone 3 Sandhi: The case of prepositions and pronouns. *Increased Empiricism: New Advances in Chinese Linguistics*, edited by Zhuo Jing-Schmidt, 271-291. Amsterdam, Netherlands: John Benjamins.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo & Hiroaki Yutani (2019). Welcome to the Tidyverse. *Journal of Open Source Software 4*. 1686.

Zhou, Youguang. (1964). *Hànzì gǎigé gàilùn* [On Chinese Orthographic Reform]. Beijing: Wenzi Gaige Chubanshe Press.