Durational Reflexes of Syllable Structure in Shanghai Chinese

Mingqiong Luo

Shanghai International Studies University (SISU), Shanghai, China
joanlmq@shisu.edu.cn

Abstract

In citation form, syllables in Shanghai Chinese (SH) contrast in rhyme duration: open syllables and those closed with a nasal coda (in general, syllables with 'smooth' rhymes, henceforth SR syllables) are much longer in rhyme duration than those closed with a glottal coda (i.e. syllables with 'checked' rhymes, henceforth CR syllables). The former pattern with contour tones and the latter level tones. In connected speech, however, all citation tones are substituted with short level ones due to the tone sandhi process, and the duration of the SR syllables is significantly reduced. This study reports on a phonetic experiment to see whether the CR syllables are also shortened in connected speech, and whether the durational contrast between SR syllables and CR ones holds in word-initial and word-final positions. The results show that CR syllables are also significantly shortened in connected speech and that the durational contrast holds across contexts.

Key words: syllable structure; Shanghai Chinese; rhyme duration; tone sandhi

1. Introduction

SH is a member of the Wu dialect group in China, which is mainly spoken in East China's Shanghai municipality. Although SH has many varieties, both in terms of geographical location and history, these varieties share two important phonological properties: (i) the durational contrast between SR syllables and CR ones, which, in citation form, pattern with long contour tones and short level ones, respectively; (ii) the tone sandhi process in which all tones in a tonal domain are deleted except for the initial one, whose contour decides that of the whole tonal domain (e.g. [18][19][20][21]). This research takes New SH as the data source, because it is representative of the current mainstream variety.

Based on the presence or absence of the onset, the medial glide and the nature of the coda, SH syllable structure can be divided into twelve types, illustrated in Table 1 below:

Table 1: Syllable structure and tonal inventory in SH (excluding nasal-only syllables):

SH Syllables	With smooth rhymes	open	CV	V	Smooth tones	Tone 1 (52)
			CGV	GV		Tone 2 (34)
		closed	CVN	VN		Tone 3 (13)
			CGVN	GVN		
	With checked rhymes	closed	CV?	V?	checked tones	Tone 4 (5)
			CGV?	GV?		Tone 5 (12)

In citation form, SR syllables are perceptually much longer in duration than CR syllables. However, in connected speech, SR syllables are significantly reduced in duration compared with those in citation form, whereas CR syllables remain perceptually unchanged in duration [2][3]. Moreover, the contour tones that the SR syllables carry in citation form are reduced to level ones in connected speech. The result is that all syllables carry a simple level tone in connected speech except for those at the end of a sentence, where an intonational falling tone is added optionally. Examples in (1) below are monosyllabic, disyllabic, tri-syllabic and quadri-syllabic words/compounds in SH with the same initial syllable and tone (i.e. /sɑ̃ŋ/ with T1). The citation tone and sandhi tone are annotated in the upper right corner of each syllable using Chao digits:

(1) a. $sɑ̃ŋ^{52}$ 'birth'
b. $sɑ̃ŋ^{52\text{-}5}$. $ji^{52\text{-}21}$ 'business'
c. $sɑ̃ŋ^{52\text{-}5}$. $ji^{52\text{-}3}$. $nin^{13\text{-}21}$ 'businessman'
d. $sɑ̃ŋ^{52\text{-}5}$. $ji^{52\text{-}3}$. $dɤ^{13\text{-}3}$. $nɔ^{13\text{-}21}$ 'business awareness'

The tone sandhi patterns in SH monosyllabic and polysyllabic compounds are summarized as follows (adapted from [18]):

Table 2: Tone sandhi patterns in SH words and compounds:

	1σ	2σ	3σ	4σ	5σ
T1	52	5+21	5+3+21	5+3+3+21	5+3+3+3+21
T2	34	3+4	3+4+21	3+4+3+21	3+4+3+3+21
T3	13	1+4	1+4+21	1+4+3+21	1+4+3+3+21
T4	5	3+4	3+4+21	3+4+3+21	3+4+3+3+21
T5	12	1+23	1+1+13	1+4+3+21	1+4+3+3+21

The distinctive tone sandhi patterns in SH have given rise to a heated debate on the underlying and surface form of SH syllable structure, and their relation with syllabic weight and tone.

2. Literature review

Duanmu [9] presents a detailed formal analysis on SH syllable structure, syllabic weight and tone, and proposes that: (i) all syllables in SH are CV syllables underlyingly. The nasal coda and the glottal coda are just diacritics on the nuclear vowel. That is, CVN syllables and CV? syllables are in fact CV^{N} and $CV^{?}$ syllables underlyingly; (ii) all syllables are light in SH underlyingly; (iii) the position where the syllable is in a prosodic word determines its weight on the surface form. Specifically, syllables in the word-initial position are stressed and heavy because of their ability to keep their base tones, whereas those in the non-initial position(s) are unstressed and light because of their inability to keep their base tones. Thus, there is no weight distinction between SR and CR

syllables in the same word position (also refer to [7][8][11][12]).

Duanmu [10] did a phonetic experiment to support the proposal that all syllables in Mandarin-like languages are heavy, whereas those in SH-like languages are light. The study took syllable duration to be the phonetic correlate for syllabic weight. However, the experiment did not distinguish between SR syllables and CR ones, nor did it distinguish between syllables in the word-initial position and those in the non-initial position. Therefore, the results failed to provide supportive evidence for the proposals in Duanmu [9].

Zhu's [22][24] phonetic experiment on SH citation tones considered all the possible 25 combinations between all the five citation tones in SH disyllabic words. The mean duration of all the five citation tones in word-initial and word-final positions of di-syllabic words, respectively, are summarized in Table 3 below:

Table 3: Mean duration of the citation tones in SH disyllabic words, adapted from [23]:

(in ms)	T1 + X	T2 + X	T3 + X	T4 + X	T5 + X	Ave X
X=T1/T2[1]	158 + 73	174 + 121	172 + 122	55 + 139	57 + 169	125
X=T3	187 + 81	191 + 135	202 + 129	77 + 161	65 + 193	140
X=T4	162 + 39	174 + 51	177 + 55	59 + 54	55 + 63	52
X=T5	180 + 46	202 + 62	211 + 60	78 + 66	68 + 85	64
Ave T	172	185	190	67	61	---
Citation T	212	240	238	71	92	---

T1=53, T2=35, T3=13, T4=5, T5=12(short)

In Table 3, the bottom row shows the mean duration of the five citation tones; the pre-final row shows duration in word-initial position, and the right-most column shows mean duration in word-final position. The table shows that the checked tones (i.e. T4 and T5) are significantly shorter in duration than the smooth tones (i.e. T1, T2 and T3) both in citation form and in word-initial and word-final positions. Specifically, the ratio of average duration between smooth and checked tones in the citation form is around 2.5:1; and those in the word-initial position and word-final position are around 3:1 and 2:1, respectively. Since tone duration parallels rhyme duration in Chinese languages, Zhu's results actually show that the durational contrast between SR syllables and CR ones in SH exists in both word-initial and word-final positions.

Different from both Duanmu and Zhu, Zhang&Meng [21] found that in left-dominant tone sandhi in SH disyllabic words, the rhyme duration of the initial syllable is similar to that of the final syllable. Their experiment took the average duration among syllables carrying all the five citation tones in each of the word positions and did not distinguish between SR and CR syllables. Therefore, the results failed to provide sufficient evidence for the durational difference between SR and CR syllables in SH.

The contradictory findings in Duanmu [9][10][11][12], Zhu [22][23][24] and Zhang&Meng [21] motivate two research questions: do CR syllables also reduce in rhyme duration in connected speech? Does the durational contrast between SR and CR syllables exist in different word positions? We approach these two questions through a phonetic experiment described in Section 3.

3. The phonetic experiment

We constructed a dataset with six major syllable types in SH, i.e. CV, CGV, CVN, CGVN, CVq, and CGVq. We distinguished between the glides /j/ and /w/ to see whether they influence rhyme duration in different ways. This results in six SR syllable types, namely, CV, CjV, CwV, CVN, CjVN and CwVN, and three CR syllable types, namely, CVq, CjVq and CwVq, in the corpus. Each SR syllable type patterns with three citation tones, and each CR syllable type patterns with two citation tones (cf. Table 1), so altogether there are 6×3+3×2=24 target monosyllables in the corpus[1].

In order to test the effect of word position on syllable duration, we constructed two disyllabic words for each monosyllable, with the target monosyllable in word-initial position and word-final positions, respectively, resulting in 24×2=48 disyllabic words in the corpus.

All the monosyllables and disyllabic words were set in the carrier phrase: [ŋu kɒŋ gəʔ tsəʔ _____ pəʔ nuŋ t^{h}iŋ] (I say this _____ to you.). In addition, the two disyllabic words were further embedded in two non-carrier sentences, which were constructed by the researcher with the help of two experiment subjects. Thus, each target monosyllable was produced in five sentences, resulting in 24×5=120 sentences in the corpus. The nuclear vowel was held constant as /a/ because this vowel is subject to fewer phonotactic and tonal constraints than other vowels in SH.

Eight native speakers of SH, four male and four female, participated in the experiment. They were asked to read the sentences at a normal speed and as naturally as possible. The recordings were made with a DR-05 Tascam linear PCM recorder in a quiet university classroom.

Altogether 960 tokens were collected and analyzed acoustically in Praat (cf. [2]). Segmentation of the rhyme was done straightforwardly by visual inspection of the waveform and spectrogram, aided by listening. The first three formants were used as further indicators of the segmental boundary.

For CR syllables, only the voiced part of the rhyme was measured, because previous researches, e.g. [2][3][4], suggest that there is no regular and constant glottal pulsing at the end of CR syllables in connected speech. Even in the citation form, the glottal pulsing is only occasionally detected. These findings were matched with ours in this experiment. Therefore, the offset of CR syllables was determined at the last regular vocal pulsing of the nuclear vowel.

4. The experimental results

4.1 General comparison between SR and CR syllables across contexts

Table 4 below presents the average rhyme duration (in ms) of nine syllable types in five contexts, based on the raw data elicited from the eight experiment subjects.

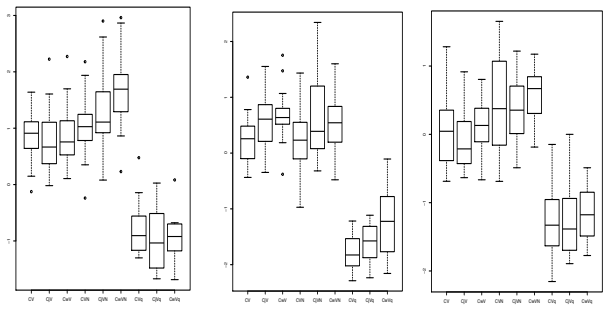
Table 4: Mean rhyme duration (in ms) of nine syllable types in five contexts:

	citation	phini	phfin	spini	spfin
CV	175	149	141	139	157
CjV	172	163	135	143	138
CwV	176	167	145	159	156
CVN	184	148	159	141	141
CjVN	193	166	155	160	133
CwVN	211	163	164	138	145
CVq	106	62	84	59	77
CjVq	99	71	86	64	91
CwVq	100	85	89	79	93

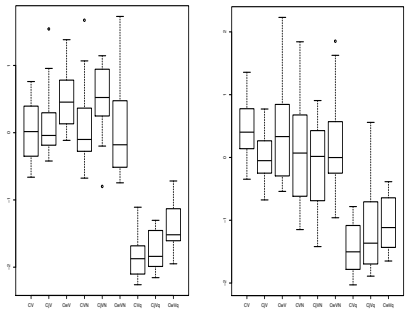
Note: phini=word-initial position in carrier phrase; phfin=word-final position in carrier phrase; spini=word-initial position in non-carrier sentences; spfin=word-final position in non-carrier sentences.

In order to assess the statistical significance of the differences between SH syllable/rhyme types in different contexts, a mixed-effects linear regression analysis was run in R using the *lme4* package [1]. Duration is the dependent variable and normalized into Z-score. Syllable structure and tone[2] were set as the fixed effects. Subject was set as the random effect. The data is divided into five subsets according to the context. Figure 1 below is a boxplot presentation of the comparison in rhyme duration between the nine syllable types in SH across the five positions surveyed in this experiment.

Figure 1: comparison in rhyme duration between SH syllable types (CV, CjV, CwV, CVN, CjVN, CwVN, CVq, CjVq and CwVq in order) in citation form (i), word-initial (ii) and word-final (iii) positions in carrier phrase, and word-initial (iv) and word-final (v) positions in non-carrier sentences.



(i) (ii) (iii)



(iv) (v)

Table 4 shows that in citation form, SR syllables that are structurally more complex are longer in rhyme duration than structurally less complex ones. Specifically, CjV, CwV and CVN syllables are longer in rhyme duration than CV syllables, and CjVN and CwVN syllables are the longest. This result indicates a possible effect of the medial glide and the nasal coda on rhyme duration.

However, in word-initial and word-final positions in both carrier phrase and non-carrier sentences, no such structural effects on rhyme duration are observed. Moreover, it is worth noting that in non-carrier sentences, syllables that are structurally more complex have a longer rhyme duration than CV syllables word-initially, but are shorter word-finally. This observation might seem to support Duanmu's [9][11][12] proposal that syllables in SH are stressed word-initially and unstressed word-finally. However, a closer look at Table 4 reveals that CV syllables are much longer word-initially than they are word-finally in non-carrier sentences. Moreover, all CR syllables are shorter word-initially than they are word-finally, both in carrier phrase and non-carrier sentences. These results provide negative evidence against Duanmu's [9][11][12] proposal.

Figure 1 shows that all CR syllables are significantly shorter in rhyme duration than the SR syllables. The difference is at the highest zero-level significance ($p < 0.001$). It can thus be concluded that duration contrast between SR and CR syllables holds across contexts.

On the other hand, in each context, durational differences between SR syllables (i.e. CV, CjV, CwV, CVN, CjVN, and CwVN syllables) are not as significant as those between SR and CR syllables, which is also true with the differences between CR syllables (i.e. CVq, CjVq and CwVq syllables).

4.2 Contextual effects on rhyme duration for each syllable type

Our second research question concerns the effect of contextual effects on syllable duration in SH. In order to tackle this question, we did a multiple comparison for each syllable type in rhyme duration (normalized in Z-score) between the five contexts examined in this experiment, using the *lsmeans* package in R (cf. [17]). Duration is the dependent variable. Context and tone were set as the fixed effects, and subject was set as the random effect. Relevant results are summarized in Table 5 below:

Table 5: Multiple comparison between citation form, forms in word-initial and word-final positions in carrier phrase and in non-carrier sentences, respectively, for each syllable type:

Baseline~target	CV	CjV	CwV	CVN	CjVN	CwVN	CVq	CjVq	CwVq
cit~phini	***(−)	(−)	(−)	***(−)	***(−)	***(−)	***(−)	***(−)	**(−)
cit~phfin	***(−)	***(−)	***(−)	***(−)	***(−)	***(−)	***(−)	*(−)	*(−)
cit~spini	***(−)	***(−)	*(−)	***(−)	***(−)	***(−)	***(−)	***(−)	***(−)
cit~spfin	**(−)	***(−)	**(−)	***(−)	***(−)	***(−)	***(−)	(−)	(−)
phini~phfin	(−)	***(−)	***(−)	(+)	(−)	(+)	***(+)	**(+)	(+)
spini~spfin	***(+)	(−)	(−)	(−)	***(−)	(+)	**(+)	***(+)	**(+)

Note: cit = citation. (+) indicates longer in duration; (−) indicates shorter in duration. '***' for $p < 0.001$, '**' for $0.001<p<0.01$, '*' for $0.01<p<0.05$.

The part of Table 5 above the horizontal dash line shows the difference in rhyme duration of each syllable structure between citation form and forms in word-initial and word-final positions in both carrier phrase and non-carrier sentences. The table shows that: firstly, all SR syllables, except for the CjV and CwV syllables in

word-initial position in carrier phrase, are significantly longer in citation form than in either of the two word positions in both carrier phrase and non-carrier sentences.

Secondly, among all CR syllables, CVq syllables are most significantly shorter in rhyme duration both word-initially and word-finally than those in citation form, but CjVq and CwVq syllables are as significantly shorter only in word-initial position. In word-final position in both carrier phrase and non-carrier sentences, rhyme reduction for CjVq and CwVq syllables is not as significant as that in word-initial position. Moreover, the shortening of rhyme duration for CjV and CwV syllables word-initially in carrier phrase and that for CjVq and CwVq syllables word-finally in non-carrier sentences are not significant at all, indicating a possible effect of medial glide on rhyme duration.

The part of Table 5 below the horizontal dash line is a summary of the comparison in rhyme duration for each syllable type between word-initial and word-final positions in carrier phrase and non-carrier sentences, respectively. It shows that SR syllables are more likely to have a shorter rhyme duration word-finally than word-initially in both carrier phrase and non-carrier sentences. In contrast, CR syllables tend to be significantly longer word-finally than word-initially in both carrier phrase and non-carrier sentences. This result confirms our findings in Section 4.1.

The interim summary for this section is that rhyme duration in CR syllables are reduced in connected speech as compared to that in citation form, but CjVq and CwVq syllables are not as significantly reduced word-finally as CVq syllables.

4.3 Effects of pre-nuclear glide and nasal coda on rhyme duration

Both Section 4.1 and Section 4.2 showed that the medial glide and the nasal coda may affect rhyme duration in SH. In order to better understand the effects of the medial glide and the nasal coda on rhyme duration, we ran the mixed-effects linear regression analysis again using the *lme4* package in R on syllables differing only in medial glide (Table 6 below), those differing only in the nasal coda (Table 7 below), and those differing both in medial glide and the nasal coda (Table 8 below), respectively:

Table 6: Multiple comparisons in rhyme duration between syllables differing only in medial glide:

Baseline~Target	citation	phini	phfin	spini	spfin
CV~CjV	(-)	*(+)	(-)	(+)	***(-)
CVN~CjVN	(+)	**(+)	(-)	***(+)	(-)
CVq~CjVq	(-)	(+)	(+)	(+)	(+)
CV~CwV	(+)	**(+)	(+)	***(+)	(-)
CVN~CwVN	***(+)	**(+)	(+)	(-)	(+)
CVq~CwVq	(-)	***(+)	(+)	**(+)	(+)

Table 7: Multiple comparisons in rhyme duration between SR syllables differing only in nasal coda:

Baseline~Target	citation	phini	phfin	spini	spfin
CV~CVN	(+)	(-)	*(+)	(+)	*(-)
CjV~CjVN	***(+)	(+)	***(+)	***(+)	(-)
CwV~CwVN	***(+)	(-)	***(+)	***(-)	(-)

Table 8: Multiple comparisons in rhyme duration between syllables differing in both medial glide and nasal coda:

Baseline~Target	citation	phini	phfin	spini	spfin
CV~CjVN	**(+)	**(+)	**(+)	***(+)	**(-)
CV~CwVN	***(+)	*(+)	***(+)	(+)	(-)

The results summarized in Table 6 show that in word-initial position, syllables with medial glide generally have a longer rhyme duration than those without, whereas in word-final position, no consistent difference is observed. That is to say, the presence of the medial glide does not necessarily increase the rhyme duration. This result is expected if the medial glide is part of the rhyme.

In Table 7, syllables with the nasal coda are significantly longer in citation form than those without. In word-initial position in both carrier phrase and non-carrier sentence, no consistent difference is observed. However, it is interesting to note that in word-final position in carrier phrase, syllables with the nasal coda are significantly longer than those without, but in the same position in non-carrier sentence, the former are consistently shorter. These results indicate that: (i) the nasal coda is flexible in phonetic realization in connected speech; (ii) factors that affect the phonetic realization of segments in non-carrier sentence can be different from those in carrier phrase.

Table 8 shows that the joint effect of medial glide and nasal coda on rhyme duration is still inconsistent: they tend to significantly increase the rhyme duration in all contexts except word-finally in non-carrier sentences, where both CjVN and CwVN syllables are shorter than CV syllables in rhyme duration.

The interim summary for this section is that neither the medial glide nor the nasal coda has a consistent effect on rhyme duration. Their presence does not make further qualitative distinctions in structure between SR syllables in SH, which is also true with the CR syllables.

5. Conclusion

Based on the experiment results in Section 4, the following conclusions about SH syllables are warranted: (i) CR syllables are significantly reduced in rhyme duration in connected speech as well; (ii) the durational contrast between SR and CR syllables holds both word-initially and word-finally in disyllabic words. These findings support those in Zhu [22][23][24] on tone duration in SH.

Since rhyme duration is one of the most important phonetic correlates for syllabic weight (e.g. [6][10][13][14][15][16]), we can also conclude that in SH, SR syllables are heavy underlyingly and CR ones are light, and that this weight contrast holds both word-initially and word-finally in disyllables.

Acknowledgements: Thanks to Michael Kenstowicz, Marjoleine Sloos and Jeroen van de Weijer for their insightful comments on this paper. This research is supported by a project fund from Shanghai International Studies University (SISU) in China. The project name is "*Acquisition of English Sounds by Native Speakers of A Non-Mandarin Dialect in China---A Third Language Perspective*" (Project NO.: 20161140011).

6. REFERENCES

[1] Bates D, Mächler M, Bolker B, Walker S. 2015. "Fitting Linear Mixed-Effects Models Using lme4." Journal of Statistical Software, 67(1), 1–48. doi:10.18637/jss.v067.i01.

[2] Boersma, P., Weenink, D., 1992—2011. Praat: Doing Phonetics by Computer. Version 5.2.21, http://www.praat.org

[3] Chao, Y. 1928. Studies in the Modern Wu-Dialects. Beijing: Qinghua College Research Institute.

[4] Chao, Y. 1967. Contrasting aspect of the Wu Dialects. In *Linguistic Essays by Yuenren Chao*. Pp:802-821. Beijing: the Commercial Press.

[5] Chen, Y. 2008. The acoustic realization of vowels of Shanghai Chinese. Journal of Phonetics (36): 629-648.

[6] Cohn, A. 2003. Phonological Structure and Phonetic Duration: the Role of Mora. Working Paper of the Cornell Phonetics Laboratory: 69-100.

[7] Duanmu, S. 1990. A formal study of syllable, tone, stress and domain in Chinese languages. PhD dissertation. MIT.

[8] Duanmu, S. 1991. Stress and syntax-phonology mismatches: tonal domains in Danyang and Shanghai. Proceedings of the Tenth West Coast Conference on Formal Linguistics. 127-137.

[9] Duanmu, S. 1993. Rime length, stress, and association domains. Journal of East Asian Linguistics (2): 1-44.

[10] Duanmu, S. 1994. Syllable weight and syllable duration [J]. Phonology (1): 1-24.

[11] Duanmu, S. 1999. Metrical Structure and Tone: Evidence from Mandarin and Shanghai. Journal of East Asian Linguistics 8: 1-38.

[12] Duanmu, S. 2008. Syllable Structure: the limits of variation. New York: Oxford University Press.

[13] Gordon, M. 2002. A Phonetically Driven Account of Syllable Weight. *Language* (78): 51-80.

[14] Gordon, M. 2004. Syllable weight. In Hayes, B., Kirchner, R.&Steriade, D. (eds.) *Phonetically based phonology*: 277-312.

[15] Gordon M. 2006. Syllable weight: phonetics, phonology, typology[M]. Routledge.

[16] Hyman, L. 1984. On the weightlessness of syllable onsets. Annual Meeting of the Berkeley Linguistics Society (10). 1-14.

[17] Lenth, R. 2016. Least-Squares Means: The R Package lsmeans. Journal of Statistical Software (69): 1-33. URL: https://www.jstatsoft.org/v069/i01.

[18] Qian, N., Xu, B., Tang, Z. 2007. A Grand Dictionary on Shanghai Chinese [Shanghaihua Dacidian]. Shanghai: Shanghai Lexicography Publishing House.

[19] Tang, Z. 2000. A Practical Coursebook in Shanghainese. Shanghai: Shanghai Education Press.

[20] You, R. (ed.) 2013. An Investigation and Research on the Dialects in Shanghai Municipality [Shanghai Diqu Fangyan Diaocha Yanjiu]. Shanghai: Fudan University Press.

[21] Zhang, J., Meng, Y. 2016. Structure Dependent Tone Sandhi in Real and Nonce Disyllables in Shanghai Wu. Journal of Phonetics (54): 169-201.

[22] Zhu, X. 1995. Shanghai Tonetics. PhD dissertation. Australian National University, Canberra.

[23] Zhu, X. 1996. Two stress patterns of Shanghai compounds. In P. McCormack and A. Russell (ed.) *Proceedings of the sixth Australian international conference in speech science and technology*. Canberra, Australia: Australian Speech Science and Technology Association. 581-586.

[24] Zhu, X. 1999. Shanghai Tonetics. München: LINCOM EUROPA.

[1] We categorized the rhyme duration according to citation tones, normalize the data into Z-score and ran a mixed-effects linear regression analysis using the lme4 package in R. Rhyme duration is the dependent variable. Citation tone was set as the fixed effects. Subject was set as the random effect. The statistical results and boxplot presentation show that although there are statistically significant differences in rhyme duration between syllables carrying the three smooth tones (i.e. T1, T2 and T3) respectively, the differences are far less significant than that between smooth tones and checked tones, which is also true with syllables carrying the two checked tones (i.e. T4 and T5) respectively. The results provide further evidence for the hypothesis that rhyme duration in Chinese languages is closely related with tone duration.

[2] The ANOVA test shows that the model with syllable structure and tone as the fixed effects has the smallest AIC value, compared to a model with no effect for tone, one with tone as the random effect, and one with the interaction of syllable structure and tone as the fixed effects.