

# PHONETICS AND PHONOLOGY OF THE -ER SUFFIX IN THE HANGZHOU WU CHINESE DIALECT

Yang Yue<sup>1</sup>, Fang Hu<sup>2,1</sup>

<sup>1</sup>Department of Linguistics, Graduate School, University of Chinese Academy of Social Sciences

<sup>2</sup>Institute of Linguistics, Chinese Academy of Social Sciences  
kk24325535@qq.com, hufang@cass.org.cn

## ABSTRACT

This paper describes the phonetics and phonology of the -er suffix in the Hangzhou Wu Chinese dialect. 12 speakers, 6 male and 6 female, provided speech data. The phonetic characteristics of the -er suffix was discussed on the basis of articulatory and acoustic data. And the impact of the -er suffix on the preceding syllables was examined. Results suggest that the -er suffix is a somehow retroflexed lateral, or a lateral rhotic. Different to Beijing Mandarin, there is no resyllabification process regarding the -er suffixation in Hangzhou. Rather, multisyllabic words with -er suffixation are subject to general rules of tone sandhi. However, the 2 falling diphthongs [ei ou] are monophthongized into [e o] respectively during the process of -er suffixation.

**Keywords:** -er suffixation, syllabic consonant, lateral, rhotic, the Hangzhou Wu dialect.

## 1. INTRODUCTION

Hangzhou was the capital of China during the Southern Song dynasty (A.D. 1127-1279) and is the capital city of the Zhejiang province. The Hangzhou dialect belongs to the Wu dialect family but has certain characteristics of Mandarin ([1], [2], [3], [4]).

The -er suffix occurs as a typical case where Wu meets Mandarin. The -er suffixation is a morphophonological process expressing a diminutive meaning. The phonetic forms of -er suffix vary from fricative, approximant, rhotic, rhoticized vowel, plain vowel, apical vowel, to nasal among Chinese dialects ([5]). A basic typology is that -er in Mandarin dialects usually takes a form of fricative or rhotic consonant, whereas -er in Wu and other southern dialects is associated with nasal consonant. This is the reason why Karlgren reconstructed \*ŋz for the historical consonant in Middle Chinese ([6]), because he failed to give an explanation how a nasal could change into a fricative or vice versa ([5], [7]). But -er in Hangzhou is neither a rhotic nor a nasal; it is basically a lateral. This is believed to be the influence of the Mandarin speaking immigrants in the Southern Song dynasty ([8], [9]). That is, the rhotic consonant for -er, which was not available in Hangzhou phonology, was substituted by a lateral.

The morphological -er suffixation is associated with phonological resyllabification in Beijing Mandarin, and consequently affects the phonetics of the preceding host syllable. In Hangzhou, there is no resyllabification for -er suffixation. The -er suffix stands as an individual syllable, and the disyllabic or multisyllabic words with -er suffixation is subject to general rules of tone sandhi. However, -er suffixation could still have effect on the vowel in the preceding syllable. Hangzhou has a rich vowel inventory. There are 8 monophthongs [ɿ ʏ a i u y ɛ ɔ], 2 falling diphthongs [ei ou] and 10 rising diphthongs [ia io iɛ ua uɛ uo yo ɥa ɥɛ ɥo], and a triphthong [uei] in open CV syllables. A few vowels and rising diphthongs [i y a ə o ia io ua uə ɥa ɥə] occur in syllables with a nasal coda -ŋ; and fewer [a o iɛ io ua yɛ ɥa] also occur in syllables checked by a glottal stop -ʔ. All the vowels can be followed by the -er suffix to form a word expressing a diminutive meaning.

This paper describes the phonetics and phonology of the -er suffix in Hangzhou. As mentioned earlier, the -er suffix is basically a lateral in Hangzhou and is often transcribed as syllabic lateral [l̥] or [əl] in dialectological works ([10]). But there is controversy that Simmons argued that the -er suffix in Hangzhou is not a lateral, but a rhoticized vowel [əɾ] ([11]). Xu further pointed out that the pronunciation of the -er suffix might be affected by sociolinguistic factors and prosodic conditions, namely younger speakers tend to use more retroflex articulations for the -er suffix in unstressed syllables ([12]). However, previous studies were basically dialectological impressionistic descriptions without phonetic details. This paper aims to query into this issue with more phonetic details. And also, this paper examines how -er suffixation affects the vowel production on the preceding syllable.

## 2. METHODOLOGY

12 native speakers, 6 male and 6 female, provided speech data. All of them were born and raised up in Hangzhou, and had no reported history of speech or hearing disorders.

Meaningful -er suffix words with a preferable mid-tone monosyllabic noun root were used as test

words in the acoustic study. Audio sounds were recorded into a laptop PC with a DMX 6 Fire USB sound card through a SHURE SM86 microphone. Speakers were instructed to speak the test words in a natural way with a normal tempo. The sample rate is 22,050 Hz. Five repetitions were recorded.

Acoustic data were annotated and analyzed by PRAAT 6.0.19 ([13]) and VoiceSauce ([14]). The annotation first divide the target -er suffixation into two parts: the monosyllabic root and -er. The lowest three formants were extracted at the midpoint of the -er syllable and of the vowel or diphthong element in the root, respectively. Due to the space limit, only the monophthong data were included in this paper. In order to check if there is a lingual gesture of rhoticization during -er suffixation, the third formant (F3) were measured at the 70-100% point of the entire duration on the root vowel.

The articulatory study is based on one male speaker who is not included in the acoustic study. Monosyllabic CV words were used as control, where C includes all coronal consonants and V is the low vowel [a] to facilitate the measures of linguopalatal contacts. The test words are onset-less V syllables followed by the -er suffix. And the monosyllabic word er ‘son’ was also included.

### 3. RESULTS

#### 3.1. Properties of the -er suffix

##### 3.1.1. Articulation

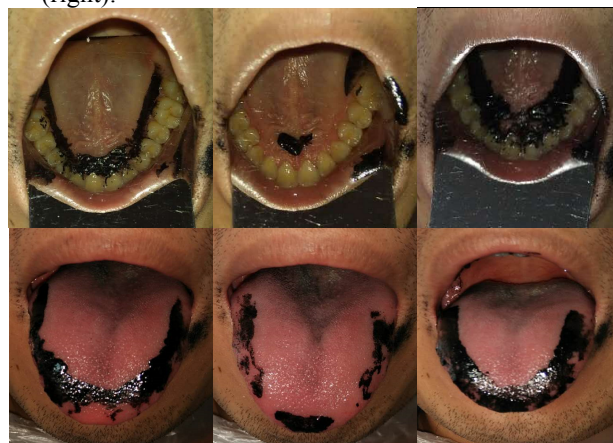
Figure 1 shows palatograms (upper) and linguograms (lower) for the monosyllabic word er [ɪ<sup>23</sup>] ‘son’ (mid), as compared to the alveolar stop [tʰa<sup>53</sup>] ‘attack’ (right), and affricate [tʰsa<sup>33</sup>] ‘squeeze’ (left) in citation form. The black areas in the figure indicate linguopalatal contacts. Palatograms show the contacts on upper teeth and the palate, and linguograms show the contacts on the tongue. The palatogram and linguogram for the target syllable were recorded separately. And Figure 2 below is arranged in the same way.

It can be observed from Figure 1 that there are three differences between er [ɪ<sup>23</sup>] and its coronal counterparts [tʰsa<sup>33</sup>] and [tʰa<sup>53</sup>]. First of all, they differ in manner of articulation. That is, er [ɪ<sup>23</sup>] is lateral, as there is linguopalatal contact only along the midsagittal plane. This suggests that the airflow has a median blockage but lateral passages. By contrast, there is full linguopalatal contact during the production of [tʰsa<sup>33</sup>] and [tʰa<sup>53</sup>], suggesting a complete blockage of airflow for stop and affricate. Second, they differ in passive place of articulation. The lateral er is postalveolar, while [tʰsa<sup>33</sup>] and [tʰa<sup>53</sup>] are basically alveolar. Third, they differ in terms of

active articulator: namely er is apical, while [tʰsa<sup>33</sup>] and [tʰa<sup>53</sup>] are more laminal.

Figure 2 compares palatograms (upper) and linguograms (lower) for [ɪ<sup>23</sup>] ‘son’ (mid), [ɸa<sup>2</sup>ɪ<sup>23</sup>] ‘box’ (left) and [u<sup>23</sup>ɪ<sup>23</sup>] ‘kettle’ (right) in citation form. It can be seen from the figure that there is no essential difference between the monosyllabic er and the suffix -er.

**Figure 1:** Palatograms (upper) and linguograms (lower) for [tʰsa<sup>33</sup>] (left), [ɪ<sup>23</sup>] (mid) and [tʰa<sup>53</sup>] (right).



**Figure 2:** The palatogram (upper) and linguograms (lower) for [a<sup>23</sup>əl<sup>23</sup>] (left), [ɪ<sup>23</sup>] (mid) and [u<sup>23</sup>əl<sup>23</sup>] (right).



##### 3.1.2. Acoustics

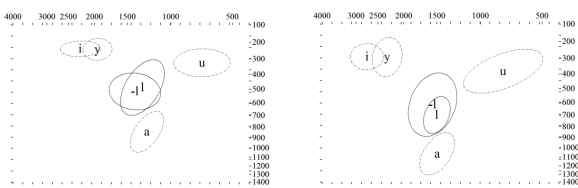
Both the monosyllabic er ‘son’ and -er suffix have a schwa-like formant structure. Figures 3 shows 2-sigma ellipses for the monosyllabic er (l) ‘son’ and -er (-l) suffix in the acoustic F1/F2 vowel plane. The vowel plane is determined by using the first formant (F1) as ordinate and second formant (F2) as abscissa with the origin of the axes to the top right. The axes are Bark-scaled ([15]), while the values along the coordinates are still labelled in Hertz. The ellipse for the monosyllabic er ‘son’ is based on 30 data points (5 repetitions × 6 speakers); and the ellipse for the -

er suffix is based on 300 data points (10 vowels  $\times$  5 repetitions  $\times$  6 speakers). And the ellipses for the monophthongs [i y a u], each of which is based on 60 data points (2 test words  $\times$  5 repetitions  $\times$  6 speakers), are also superimposed in the figure for reference.

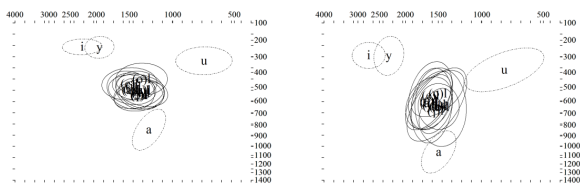
It can be observed from Figure 3 that the ellipses for the monosyllabic er ‘son’ and -er suffix occupy a central vowel position and extensively overlap with each other. As mentioned earlier, -er was usually transcribed as [ɨ] or [əɭ] in the dialectological literature. In fact, the schwa [ə] and the lateral [ɭ] have similar formant structure, as the vocal tract is basically in a neutral situation during the production of both sounds. The lateral [ɭ] differs with the schwa [ə] in that in addition to formants, the lateral [ɭ] has antiformants. -er is transcribed as [ɨ] in this study, as the data suggest that -er is composed of a single segment, rather than a sequence of a schwa and a lateral [əɭ].

Figure 4 examines the ellipses for the -er suffix in the context of different preceding vowels. Each ellipse is based on 30 data points. The preceding vowels include the 8 monophthongs [ɿ ʏ a i u y ɛ ɔ] and 2 falling diphthongs [ei ou] that are realized as monophthongs [e o] respectively after -er suffixation. It can be seen from the figure that although there is certain contextual effect by the preceding vowel, the -er suffixes occupy a central position and all ellipses extensively overlap with each other in the acoustic F1/F2 vowel plane.

**Figure 3:** 2-sigma ellipses for the monosyllabic er (l) ‘son’ and -er (-l) suffix in male (left) and female speakers (right).



**Figure 4:** 2-sigma ellipses for the -er suffixes in the context of different preceding vowels in male (left) and female speakers (right).

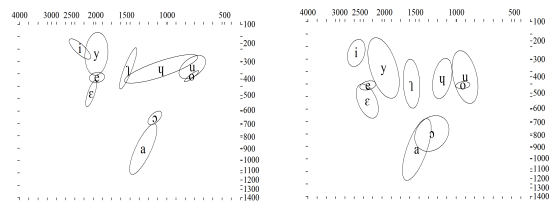


### 3.2. The impact of -er suffixation on the preceding vowel

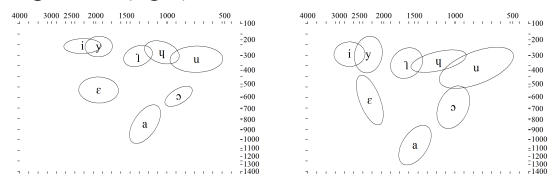
Unlike in Beijing Mandarin where the process of -er suffixation triggers resyllabification ([16]), it does not in Hangzhou. The -er suffix [ɨ] stands as an individual syllable, and has a duration comparable to the preceding syllable in general. Moreover, -er suffixation obeys general rules of tone sandhi in Hangzhou. The -er suffix has a citation tone that is transcribed as [23] according to Chao’s 5-digit tone letters ([17]). For instance, disyllabic words with -er suffix undergo tone sandhi according to the rule applied to the disyllabic words whose second syllable is the tonal category of [23]. In summary, unlike in Beijing Mandarin where -er suffix is an unstressed syllable with a neutral tone, the -er suffix in Hangzhou is a stressed syllable with a citation tone.

However, -er suffixation still has impact on the vowel on the preceding syllable. The most salient effect is that the two falling diphthongs [ei] and [ou] are monophthongized under the process of -er suffixation. Figures 5 and 6 compare the distribution of Hangzhou vowels with and without -er suffixation in the acoustic F1/F2 vowel plane. Each ellipse in Figure 5 is based on 30 data points, and each ellipse in Figure 6 is based on 60 data points.

**Figure 5:** 2-sigma ellipses for Hangzhou vowels with -er suffixation in male (left) and female speakers (right).



**Figure 6:** 2-sigma ellipses for Hangzhou vowels without -er suffixation in male (left) and female speakers (right).



As can be seen from the figures, the eight Hangzhou vowels [ɿ ʏ a i u y ɛ ɔ] have a three-way distinction of height in open CV syllables: high [i y ɿ ʏ u], mid [ɛ ɔ], and low [a]. But under the environment of -er suffixation, there is a four-way distinction of height. That is, the mid-high vowels [e o] emerge and are contrastive to their mid-low counterparts [ɛ ɔ]. has its relative independent

distribution in acoustic plane, while [o] is overlapped with [u]. But it can also be seen from Figure 5 that ellipses for [o] and [u] extensively overlap with each other, especially in female speakers. That is probably why mid-high vowels become dynamic in open CV syllables in Hangzhou, since being dynamic helps to distinguish the mid-high vowels from the neighboring vowels ([18]).

Another salient impact on the preceding vowel concerns the third formant (F3). It is observed that the ending part of the preceding vowel with -er suffixation, especially the non-high vowel, has a declined F3, as compared to the corresponding plain vowel without -er suffixation. Table 1 summarizes mean F3 values on the 70-100% point of the entire duration for every vowel with and without -er suffixation. And the results of one-way ANOVAs yielded significant differences in most cases. The exceptions are mainly from high or apical vowels. The F3 lowering of the preceding vowel indicates the presence of rhoticization and signifies a retroflex gesture of the -er suffix. That is, the lateral [l̥] is a lateral rhotic. This is why Simmons, a native American English speaker, described it as a sound similar to the r-sound, rather than an l-sound in American English ([11]).

**Table 2:** One-way ANOVAs for F3 in the paired vowels with and without -er suffixation.

	Male speakers		Female speakers	
	Mean	P-value	Mean	P value
il	2864	0.157	3202	<0.0001
i	2958		3478	
ul	2732	0.0007	3051	0.1
u	3049		3175	
yl	2195	0.01	2195	0.11
y	2367		2759	
ɟl	2673	0.007	3227	0.326
ɟ	2909		3181	
ɥl	2777	0.007	3056	0.03
ɥ	3022		3202	
al	2626	0.008	2975	<0.0001
a	2798		3235	
ɛl	2486	0.015	2605	<0.0001
ɛ	2605		2894	
el	2404	0.0007	2868	<0.0001
e	2585		3089	
ol	2674	0.004	3095	<0.0001
o	2950		3433	

#### 4. CONCLUSION

This paper describes the phonetics and phonology of the -er suffix and -er suffixation in the Hangzhou Wu Chinese dialect.

The palatograms and linguograms confirmed that the Hangzhou -er suffix is basically a lateral [l̥]. Based on the sampled speaker, it is not an alveolar lateral, but an apical postalveolar lateral. The lateral has a schwa-like formant structure. And the acoustic data suggest that the lateral [l̥] is also a rhotic simultaneously, as it triggers vowel rhoticization on the preceding syllable.

Different to that in Beijing Mandarin, the morphological process of -er suffixation in Hangzhou does not result in phonological resyllabification. And the -er suffix stands as an individual syllable in the words with -er suffixation. The -er suffix is neither a weakened syllable, as the multisyllabic words with -er suffixation obey general rules of tone sandhi in Hangzhou. However, the -er suffixation still has impact on the preceding host syllable. In addition to the rhoticization of the vowel on the preceding syllable, it is observed that the two falling diphthongs are monophthongized, [ei ou] > [e o] respectively, in the environment of -er suffixation.

#### 5. ACKNOWLEDGEMENTS

This work is supported by CASS Innovation Program and Chinese National Social Science Foundation (Project number: 15BY073; P.I.: Hu Fang).

## 5. REFERENCES

- [1] Simmons, R. V. 1992. *The Hangzhou Dialect*. Ph.D. dissertation, The University of Washington, Seattle.
- [2] Li, R. and Bao, S. 1998. *The Hangzhou dialect Dictionary* [in Chinese]. Nanjing: Jiangsu Educational Publishing House.
- [3] Qian, N. 1992. *Studies in the contemporary Wu dialects* [in Chinese]. Shanghai: Shanghai Educational Publishing House.
- [4] Wang, Q. 1999. The Hangzhou phonology [in Chinese]. *Journal of Tsinghua University (Philosophy and Social Sciences)* (1), 57-67.
- [5] Anonymous. Post-oralization of nasal initials and the emergence of rhotic and vowel rhoticization in Mandarin Chinese. Submitted.
- [6] Karlgren, B. 1915-1926. *Etudes sur la Phonologie Chinoise*. Leyde: E.-J. Brill.
- [7] Hu, F. 2007. Post-oralized nasal consonants in Chinese dialects – Aerodynamic and acoustic data. In *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, Germany, 1405-1408.
- [8] Zhengzhang, S. 1979. The -er suffix in Wenzhou [in Chinese]. *Fangyan (Dialect)* (3), 207-230.
- [9] Bao, S. 1988. The Hangzhou dialect and the Mandarin dialects [In Chinese]. In *Studies on Wu dialects*, Shanghai: Shanghai Educational Publishing House, 282-288.
- [10] Bao, S. 1998. *A dictionary of the Hangzhou dialect* [in Chinese]. Nanjing: Jiangsu Educational Publishing House.
- [11] Simmons, R. V. 1989. A note on the articulation of the -er suffix in the Hangzhou dialect [in Chinese]. *Fangyan (Dialect)* (3), 180-181.
- [12] Xu, Y. 2002. On the -er suffix words in the Hangzhou dialect [in Chinese]. *Journal of Hangzhou Teachers College (Philosophy and Social Sciences)* (2), 93-97.
- [13] Boersma, P. & Weenink, D. 2016. Praat: doing phonetics by computer [Computer program]. Version 6.0.19.
- [14] Shue, Y.-L. 2010. *The voice source in speech production: Data, analysis and models*. Ph.D. Dissertation, UCLA.
- [15] Smith, J. O. and Abel, J. S. 1999. Bark and ERB Bilinear Transforms. *IEEE Trans. Speech and Audio Proc.*, 7(6), 697-708.
- [16] Chao, Y.-R. 1968. *A grammar of spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- [17] Chao, Y.-R. 1930. A system of "tone-letters". *Le Maître Phonétique* 45, 24-27.
- [18] Hu, F. 2017. Toward a dynamic theory of vowel production. *The Journal of the Acoustical Society of America*, 142 (4), Pt.2, 2579.