Acoustic Analysis of Nasalization in Mandarin Prenasal Vowels Produced by Wenzhou and Rugao Speakers

Xinya Zhang, Yanyang Chen, Jiazheng Wang and Ying Chen*
Nanjing University of Science and Technology, Nanjing, Jiangsu, China
E-mail: zhang_xinya@njust.edu.cn Tel: +86-18312922601
E-mail: ChenYanyang@njust.edu.cn Tel: +86-18921688672
E-mail: WangJiazheng@njust.edu.cn Tel: +86-18362116101
*E-mail: ychen@njust.edu.cn Tel: +86-17714341101

***Abstract*— This study investigates the effects of first dialect, age, vowel type, nasal type and vowel timing on the nasalization of prenasal vowels in second dialect standard Mandarin produced by respectively Wenzhou and Rugao speakers. The nasalization was quantified by the values of the acoustic correlates *A*1-*P*1 (dB) and *A*1-*P*0 (dB) in the amplitudes of spectral frequencies of the vowels. The statistical results of mixed-effects linear regression model reveal different patterns of nasalization of prenasal vowels in Mandarin produced by young and middle-aged speakers of Wenzhou and Rugao dialects. The degree of nasalization was found varying across vowel type and time course of the vowel but not nasal type. The appropriate acoustic parameter for testing nasalization of the mid vowel /ə/ remains uncertain because *A*1-*P*1 is supposed to be used for measuring nasalization of nonlow vowels while *A*1-*P*0 for nonhigh vowels [11, 12].**

I. Introduction

The production of (C)VN syllables with a vowel preceding a coda nasal involves a process of coarticulation. The result of nasalization affects the acoustic structure of the vowel, and in turn, the perception of the vowel quality [1]. Lowering the velum before articulation of nasal consonants necessarily involves some amount of overlap with the articulation of the preceding vowel [1, 2, 3, 4]. Coarticulatory cues from the preceding vowels have also been found to strongly affect the perception of place of articulation of postvocalic nasals [5, 6, 7, 8, 9, 10]. Furthermore, there is a tendency to nasalize low vowels more than high ones, which is because when tongue position is low, velum position tends to be relatively lower [2, 3]; and the velum lowers earlier during a low vowel than during a high one [2]. Thus low vowels showed longer nasalization than high ones [11]. Therefore, the effect of coda nasal on the quality of the preceding vowel can be detected by examining nasalization in the vowel [11].

Acoustics analysis of nasalized vowels in frequency domain indicates the presence of extra peaks: one between the first two formants with amplitude *P*1 and one at lower frequencies, often below the first formant, with amplitude *P*0. The values for the acoustic correlates *A*1-*P*1 and *A*1-*P*0 can be used as measures of nasalization quantification. Here *A*1 refers to the first-formant amplitude. *A*1-*P*1 correlates with the nasalization of nonlow vowels and *A*1-*P*0 with nonhigh vowels, with the value (dB) of both *A*1-*P*1 and *A*1-*P*0 negatively related to the degree of nasalization [12]. Previous studies have found that the measurements of *A*1-*P*1 and *A*1-*P*0 showed lower values for nasal vowels than for oral ones in most cases in English and higher values for the least-nasalized portions than for the most-nasalized portions in French, whose nasalized vowels do not have corresponding oral vowel counterparts [12]. In CVN syllables, more cues from the nasal coda were present in the vowel region in Mandarin than in English [6]. Vowels /a, ə, i/ in Mandarin were found to have greater degree and faster rate of nasalization in nasal coda contexts than in non-nasal contexts [11].

Based on the above reviewed findings in the literature, the current study aims to quantify the nasalization of vowels /a, ə, i/ preceding coda nasals /n, ŋ/ in Mandarin Chinese produced by native speakers of Wenzhou dialect and Rugao dialect.

Wenzhou is a city located in the southeast of Zhejiang Province and connected to the north border of Fujian Province, China. Wenzhou dialect belongs to Wu, a Chinese language spoken in the area from the south bank of Yangtze River to Zhejiang Province. Wenzhou dialect is the southernmost Wu dialect, which is mostly in contact with Southern Min, a Chinese language mostly spoken in Fujian and Taiwan. From the perspective of language contact, it has been found that there is a historical relationship between Wu dialect and Southern Min with a large number of similarities in phonetic features [13]. However, Southern Min preserves three coda nasals /m, n, ŋ/ from old Chinese whereas Wenzhou dialect has only one coda nasal phoneme /ŋ/. In Wenzhou dialect, only /a, ə, o/ can be followed by the coda nasal /ŋ/; /i/ can never precede a coda nasal but work as a monophthong [13].

Rugao is a city located in the southeast of Jiangsu Province and on the north bank of Yangtze River. Rugao dialect belongs to Jianghuai Mandarin, a Mandarin variety spoken in the area between Huaihe River and Yangtze River. Rugao dialect mostly contacts with northern Wu dialect. To a great extent, it shows transitional characteristics between Jianghuai Mandarin and Wu dialect [14]. Unlike standard Mandarin, which has two coda nasals /n, ŋ/ and no nasal vowel, Rugao dialect has both nasal vowels and coda nasal phoneme /ŋ/. In Rugao dialect, the VN syllable and nasal vowels are /ẽ, õ, ĩ, ɔŋ, əŋ/, plus a low vowel with nasal is either transcribed as /aŋ/ [14] or /ã/ [15].

Diaglossia happens in both Wenzhou and Rugao. All residents in Wenzhou and Rugao are societal bilinguals. Wenzhou speakers learn Wu first and standard Mandarin follows. Rugao speakers also learn standard Mandarin as a second dialect but Jianghuan Mandarin as the first dialect. However, with the national popularization of standard Mandarin in school education and mass media, younger bilinguals learned standard Mandarin earlier and speak more Mandarin in daily life than the old generation. Therefore, the present study was designed to explore the following research questions:

(1) With different nasal rimes in their first dialect, do Wenzhou speakers and Rugao speakers produce prenasal vowels differently in their second dialect Mandarin?

(2) With the same first dialect, i.e. Wenzhou dialect or Rugao dialect, do young generation produce prenasal vowels differently from the middle-aged?

(3) Are there any other factors that impact the production of vowel nasalization in second dialect Mandarin, e.g. vowel type, nasal type, time course of the vowel?

II. Methods

A. Participants

Twelve native speakers of Wenzhou dialect and Rugao dialect were respectively recruited for two age groups: young (mean age 21.3 with SD 4.46 for Wenzhou and 19.8 with 0.90 for Rugao) and middle-aged (mean age 42.8 with SD 5.24 for Wenzhou and 42.3 with 5.28 for Rugao). Each group had three males and three females. They are all bilingual speakers of L1 dialect (either Wenzhou or Rugao) and L2 standard Mandarin. To avoid the diversity within the dialect as much as possible, all Wenzhou participants were recruited in Longwan District, Wenzhou City, Zhejiang Province, and all Rugao participants were recruited in Xiayuan Town, Rugao City, Jiangsu Province.

B. Stimuli

Six real monosyllabic words in Mandarin were selected. All the selected words are high-frequency words and with Tone 1—high-level tone in standard Mandarin: 攀[pʰan], 乓[pʰaŋ], 喷[pʰən], 烹[pʰəŋ], 拼[pʰin], 乒[pʰiŋ] [8]. The target words were embedded in a carrier sentence: /__, tʂɤ kɤ ʂi __ tsi/ (“__, this is __ character”) for recording.

C. Procedure

All the talkers were requested to read the six stimulus sentences presented in a random order in standard Mandarin with five repetitions. The second, third and fourth repetitions of the first presentation of target words in the carrier sentence in each speaker’s production were extracted for acoustic analysis, which yielded 432 tokens (6 words * 3 times * 24 participants) in second dialect Mandarin. The first and the fifth were abandoned for potential uncertainty and exhaustion respectively.

All recordings were conducted in quiet rooms, using a Marantz PMD661 professional recorder and a Shure SM10A-CN head-worn dynamic microphone in mono channel with a 44,100Hz sampling rate. The sound files were digitized into a SD card to save on a personal computer.

D. Analysis

The acoustic correlates of vowel nasalization were calculated on the value of the amplitudes in dB of the first formant ($A1$) subtracted by that of the extra peak below the first formant ($P0$) or the extra peak between the first two formants ($P1$) in the spectrum of the vowel. For /a/, $A1$-$P0$ is a better measure than $A1$-$P1$ since the $P1$ peak sometimes is not distinctly observable due to the high F1 and low F2; For /i/, $A1$-$P1$ is a better measure than $A1$-$P0$ since F1 at a low frequency affects and is affected by the P0 peak; For /ə/, both $A1$-$P1$ and $A1$-$P0$ can be used to quantify nasalization [11, 12]. Since $A1$-$P1$ is appropriate acoustic parameter for nonlow vowel nasalization and $A1$-$P0$ for nonhigh [11, 12], we analyse $A1$-$P1$ for /i, ə/ and $A1$-$P0$ for /a, ə/. For all the extracted words, $A1$-$P1$ and $A1$-$P0$ were examined at the start, middle, and end points of the prenasal vowels.

R (version 3.5.3) was used for statistical analysis. The lme4 and lmerTest packages [16] were employed to perform linear mixed-effects model. Two sets of mixed linear model were separately conducted for the two dependent variables $A1$-$P1$ and $A1$-$P0$. Subject and repetition were set as random factors and dialect (2 levels: Wenzhou, Rugao), age (2 levels: young, middle-aged), nasal (2 levels: /n, ŋ/), vowel (2 levels: /ə, i/ for $A1$-$P1$, /ə, a/ for $A1$-$P0$), point of vowel timing (3 levels: start, mid, end) as fixed factors. The first level of each fixed factor was set as the reference level to predict the vowel nasalization. The lsmeans package [17] was used for post hoc comparisons and ggplot2 package [18] for graphics.

III. Results

A. A1-P1 of Vowels /i, ə/

The statistical results show both vowel type and vowel timing significantly predicted the acoustic parameter $A1$-$P1$ of prenasal vowels /i/ and /ə/, suggesting main effects on the nasalization of these two vowels (see Table 1).

TABLE 1.
MIXED-EFFECTS LINEAR REGRESSION RESULTS OF $A1$-$P1$ (dB) IN VOWELS /i, ə/

Reference	Predictor	Est.	SE	*df*	*t*	*p*
	(Intercept)	8.020	1.530	66.481	5.241	0.000
/ə/	/i/	9.034	1.231	821.001	7.341	0.000
Start	Mid	3.994	1.376	821.001	2.903	0.004
	End	3.420	1.376	821.001	2.486	0.013
Interaction	Rugao:End	5.673	1.231	821.001	4.610	0.000
	Middle-aged: /i/	2.923	1.005	821.001	2.909	0.004
	End:/i/	5.163	1.231	821.001	4.196	0.000

Table 1 indicates that /ə/ overall had a lower $A1$-$P1$ than /i/, start point overall lower than midpoint and endpoint. There were significant interactions between dialect and vowel timing, between age and vowel type, and between vowel type and vowel timing. Post hoc comparisons were conducted based on the interactions and reported in Table 2.

TABLE 2.
POST HOC COMPARISON RESULTS OF *A*1-*P*1 (dB)

Inter-action	Level	Contrast	Est.	SE	*df*	t.ratio	*p*
Dialect: Timing	Start	Wenzhou-Rugao	3.63	1.32	39.4	2.742	0.009
	Mid	Wenzhou-Rugao	3.42	1.32	39.4	2.585	0.014
Age: Vowel	/i/	Middle-aged-Young	3.98	1.22	29.9	3.251	0.003
Vowel: Timing	Start	/i/-/ə/	9.51	0.87	821	10.931	0.000
	Mid	/i/-/ə/	7.45	0.87	821	8.564	0.000
	End	/i/-/ə/	4.35	0.87	821	4.998	0.000

Table 2 indicates that overall Rugao speakers produced lower *A*1-*P*1 than Wenzhou speakers at both start and mid points of vowels /i, ə/. Young speakers produced lower *A*1-*P*1 than middle-aged speakers for /i/. Vowel /ə/ had a lower *A*1-*P*1 than /i/ at all of the three points of vowel timing. The detailed data of *A*1-*P*1 were plotted in Figures 1-4 by dialect first, then vowel type, nasal, age, and vowel timing.

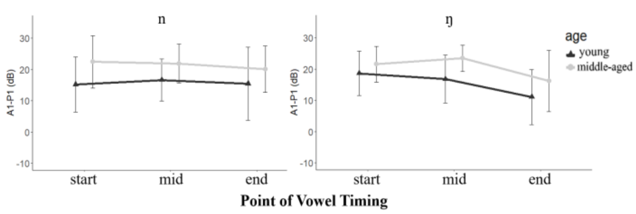

Figure 1. *A*1-*P*1 in Wenzhou speakers' vowel /i/ by nasal, age and timing

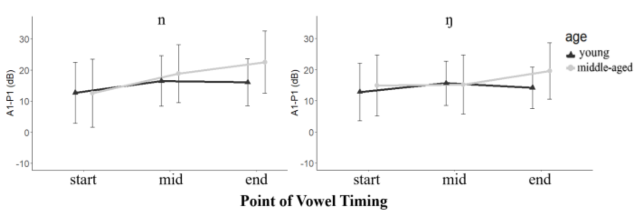

Figure 2. *A*1-*P*1 in Rugao speakers' vowel /i/ by nasal, age and timing

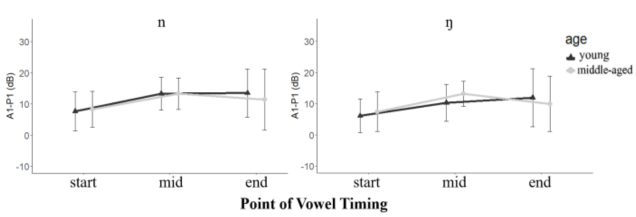

Figure 3. *A*1-*P*1 in Wenzhou speakers' vowel /ə/ by nasal, age and timing

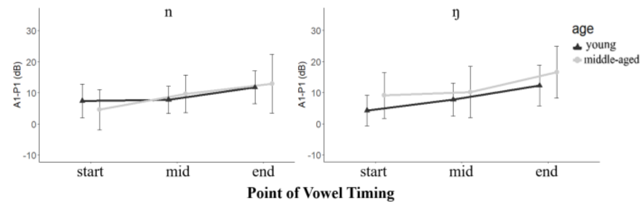

Figure 4. *A*1-*P*1 in Rugao speakers' vowel /ə/ by nasal, age and timing

Figures 1 and 3 illustrate that Wenzhou speakers produced lower *A*1-*P*1, suggesting more nasalization, especially in the middle-aged speakers' production and preceding the velar nasal /ŋ/, at the endpoint of vowel /i/ than start and mid points, which is the boundary between the nasalized vowel and the nasal murmur in the syllable [1]. This expected result did not occur in Rugao speaker's production (see Figures 2 and 4). In contrast, Rugao speakers, especially the middle-aged, produced higher *A*1-*P*1 for both /i/ and /ə/ at the endpoint than the start and mid points, which means less nasalization of vowel where it is connected to the nasal murmur.

B . A1-P0 of Vowels /a, ə/

The statistical results show first dialect, vowel type and vowel timing significantly predicted the acoustic parameter *A*1-*P0* of prenasal vowels /a/ and /ə/, suggesting main effects on the nasalization of these two vowels (see Table3).

TABLE 3.
MIXED-EFFECTS LINEAR REGRESSION RESULTS OF *A*1-*P*0 (dB) IN VOWELS /a, ə/

Reference	Predictor	Est.	SE	*df*	*t*	*p*
	(Intercept)	1.348	1.508	31.115	0.893	0.379
/ə/	/a/	4.055	0.799	823.001	5.073	0.000
Start	Mid	5.157	0.894	823.001	5.771	0.000
Wenzhou	Rugao	-4.572	2.018	24.936	-2.266	0.032
Interaction	Rugao:/a/	2.471	0.653	823.001	3.787	0.000
	Rugao:End	3.710	0.799	823.001	4.641	0.000
	Middle-aged:Mid	2.228	0.799	823.001	2.788	0.005
	Middle-aged:End	-1.685	0.799	823.001	-2.109	0.035
	/a/:End	-3.346	0.799	823.001	-4.186	0.000

Table 3 indicates that /ə/ overall had a lower *A*1-*P*0 than /a/, start point overall lower than midpoint, and also overall a lower *A*1-*P*0 produced by Rugao speakers than by Wenzhou speakers. There were significant interactions between dialect and vowel type, between dialect and vowel timing, between age and vowel timing, and between vowel type and vowel timing. Post hoc comparisons were conducted based on the interactions and reported in Table 4.

TABLE 4.
POST HOC COMPARISON RESULTS OF *A*1-*P*0 (dB)

Inter-action	Level	Contrast	Est.	SE	*df*	t.ratio	*p*
Dialect: Timing	Start	Wenzhou-Rugao	3.391	1.5	24.1	2.261	0.033
	Mid	Wenzhou-Rugao	4.037	1.5	24.1	2.692	0.013
Dialect: Vowel	/ə/	Wenzhou-Rugao	3.61	1.46	21.7	2.463	0.022
Vowel: Timing	Start	/a/-/ə/	5.71	0.57	821	10.102	0.000
	Mid	/a/-/ə/	6.03	0.57	821	10.674	0.000
	End	/a/-/ə/	2.36	0.57	821	4.181	0.000

Table 4 indicates that Rugao speakers produced lower *A*1-*P*0 than Wenzhou speakers for /ə/ and at both start and mid points of the two vowels /a, ə/. Vowel /ə/ had a lower *A*1-*P*0 than /a/ at all of the three points of vowel duration. The detailed data of *A*1-*P*0 were plotted in Figures 5-8 by dialect first, then vowel type, nasal, age, and vowel timing.

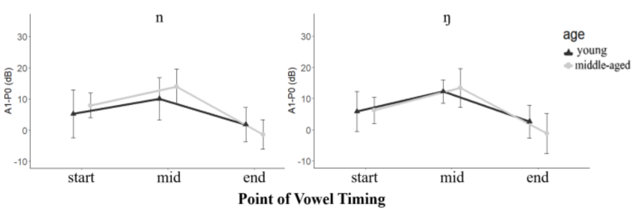

Figure 5. *A*1-*P*0 in Wenzhou speakers' vowel /a/ by nasal, age and timing

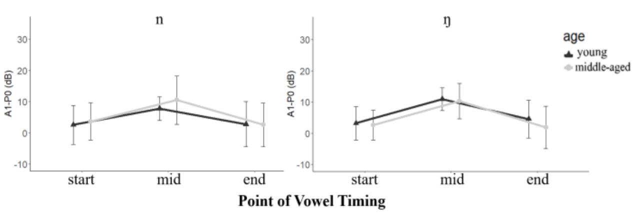

Figure 6. *A*1-*P*0 in Rugao speakers' vowel /a/ by nasal, age and timing

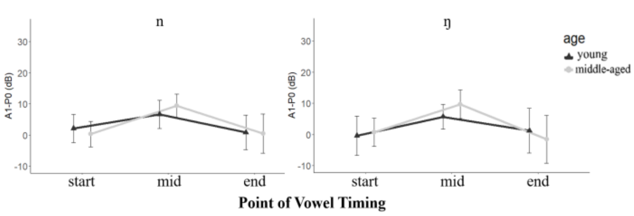

Figure 7. *A*1-*P*0 in Wenzhou speakers' vowel /ə/ by nasal, age and timing

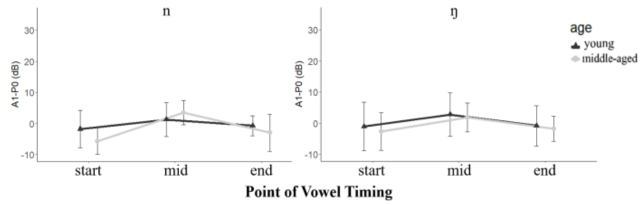

Figure 8. *A*1-*P*0 in Rugao speakers' vowel /ə/ by nasal, age and timing

Figures 5-8 illustrate that all the speakers across dialect and age produced lower *A*1-*P*0 at the endpoint than the midpoint of both vowels /a, ə/ preceding both nasals /n, ŋ/ as predicted. However, the *A*1-*P*0 was also surprisingly lower at the start point than the midpoint for both vowels.

IV. DISCUSSION

The above reported statistical results answer the first research question that Wenzhou speakers did produce prenasal vowels in Mandarin with a different pattern as in the Rugao speakers' production. Though there was only main effect of the speakers' first dialect on the production of *A*1-*P*0, dialect significantly interacted with point of vowel timing on *A*1-*P*1. Wenzhou speakers demonstrated more nasalization when the vowel approaching the nasal murmur as predicted for vowel nasalization in CVN syllables. However, in Rugao speakers' standard Mandarin production, *A*1-*P*1 suggested more nasalization at the start point of /i/ and /ə/ than mid- and endpoints. This difference between Wenzhou and Rugao speakers' production can be attributed to the difference of the phonological system in the two dialects [13, 15]. Rugao dialect has nasal vowels while Wenzhou dialect does not. Specifically, the vowel /i/ in Rugao dialect only occurs as a nasal vowel and never be followed by a nasal in a syllable. This may result in the fact that Rugao speakers habitually lower the velum too early for a /CiN/ syllable.

The statistical results also answer the second research question that the two generations of Wenzhou and Rugao speakers produced prenasal vowels in Mandarin somewhat differently. Though there was no main effect of speakers' age, age significantly interacted with vowel type and vowel timing. Figure 1 reveals that young Wenzhou speakers produced more nasalization for /i/ than the middle-aged speakers. Similar result happened to Rugao speakers' production, especially at the endpoint of /i/ (see Figure 2). Younger Rugao speakers also demonstrated lower *A*1-*P*1, suggesting more nasalization, in /ə/ than the middle-aged speakers did (see Figure 4). Because we did not include a control group's data, i.e. the production of native speakers of standard Mandarin, in the present paper, we do not know whether younger speakers' production was more nativelike or the middle-aged speakers' production. However, from the sociolinguistic and SLA (second language acquisition) point of view and based on the language background of the two generations [19], younger speakers have more experience than the middle-aged speakers in second dialect standard Mandarin and thus their production could be more nativelike.

As for the third research question, first, we do not see main effect of nasal type. That means nasal place of articulation did not impact with the degree of nasalization of prenasal vowels in standard Mandarin produced by Wenzhou and Rugao speakers. Second, we can see the main effect of vowel type in the results of both $A1$-$P1$ and $A1$-$P0$. We found that overall /ə/ was more nasalized than both /i/ and /a/, respectively a high vowel and a low vowel. Although this is not completely consistent with the expected correlation between vowel height and degree of nasalization, i.e. lower vowels being more nasalized than higher ones, the result of the present study that /a/ was less nasalized than /ə/ is still reasonable for Clumeck's finding that the correlation aforementioned occurred more often in cases of high vs nonhigh vowels than low vs nonlow vowels [2]. Theoretically, $A1$-$P1$ is the appropriate acoustic parameter for nasalization for nonlow vowels while $A1$-$P0$ for nonhigh vowels [11]. The schwa /ə/ is supposed to be mid vowel and thus both parameters are applicable to the measurement of its nasalization [12]. However, we can see the difference in Figures 3 and 4 vs. Figures 7 and 8. $A1$-$P1$ tended to increase gradually from the start point to the endpoint while $A1$-$P0$ had the highest value at the midpoint in the entire time course. Which parameter is more proper for measuring the nasalization of /ə/ remains unknown. Third, point of vowel timing showed main effect on both $A1$-$P1$ and $A1$-$P0$. The unexpected result of lower value at the start point than the mid- and/or endpoints is also difficult to explain. There are two plausible reasons. One could be the coarticulation of voiceless bilabial stop /p/ with the vowel, whose aperiodic waveform of the aspirated burst right before the vowel might have distorted the regular formant structure of the vowel in the /pVN/ syllable. The other could be the complexity and confounded factors in the non-native speech of standard Mandarin produced by Wenzhou and Rugao speakers, whose first dialect closely contacts with other Chinese languages and/or dialects.

V. Conclusions

The present study has found different patterns of nasalization of prenasal vowels in standard Mandarin produced by young and middle-aged speakers of Wenzhou and Rugao dialects. The degree of nasalization also varied across vowel type and time course of the vowel. No effect of nasal place of articulation was found on the degree of nasalization of prenasal vowels. The appropriate acoustic parameter for testing nasalization of the mid vowel /ə/ remains uncertain.

Future work will involve the acoustic analysis of nasal rimes in Wenzhou and Rugao dialects produced by the same speaker groups as in the current study and also their perception of nasal rimes in standard Mandarin produced by native speakers for insight into the correlations between first dialect production and second dialect production and between the perception and production of second dialect.

Acknowledgment

This work is supported by Humanity and Social Science Foundation of Ministry of Education of China, approval number 18YJA74007; Also, this paper is funded by the JiangSu Undergraduate Training Program for Innovation and Entrepreneurship. Item number: 201910288153Y.

References

[1] Goodin-Mayeda, C. E. (2016). *Nasals and nasalization in Spanish and Portuguese: Perception, phonetics and phonology* (Vol. 9). John Benjamins Publishing Company.

[2] Clumeck, H. (1976). Patterns of soft palate movements in six languages. *Journal of Phonetics*, *4* (4), 337–351.

[3] Cohn, A. C. (1990). Phonetic and phonological rules of nasalization. *UCLA Working Papers in Phonetics, 76*.

[4] Hayes, B. (2011). *Introductory phonology* (Vol. 32). John Wiley & Sons.

[5] Zee, E. (1981). Effect of vowel quality on perception of post-vocalic nasal consonants in noise. *Journal of Phonetics*, *9*(1), 35–48.

[6] Mou, X. (2006). *Nasal codas in Standard Chinese: a study in the framework of the distinctive feature theory* (Doctoral dissertation, Massachusetts Institute of Technology).

[7] Yang, J. H. (2003). Auditory accounts of the word-final nasal merger in Mandarin: Perception experiments and limitations. Paper presented at *NWAV-36* Philadelphia.

[8] Chen, Y., & Guion-Anderson, S. (2011). Perceptual confusability of word-final nasals in Southern Min and Mandarin: Implications for coda nasal mergers in Chinese. *Proceedings of the 17th International Congress of Phonetic Sciences*, 464–467.

[9] Chen, Y., Kapatsinski, V., & Guion, S. (2012). Acoustic cues of vowel quality to coda nasal perception in Southern Min. *Proceedings of the 13th Annual Conference of the International Speech Communication Association*, 1412–1415.

[10] Chen, Y., Chen, Y. P., Wang, L., & Cui, J. (2019). Vowel effects on the perception of nasal place of articulation in Wuxi Wu and Fuqing Min. *Proceedings of the 19th International Congress of Phonetic Sciences*, ISBN 978-0-646-80069-1, 344–348.

[11] Chen, M. Y. (2000). Acoustic analysis of simple vowels preceding a nasal in Standard Chinese. *Journal of Phonetics*, *28*(1), 43–67.

[12] Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *The Journal of the Acoustical Society of America*, *102*(4), 2360–2370.

[13] Zheng Zhang, S. F. (2008). *Dialect record of Wenzhou city (Wenzhou fangyanzhi)*, Beijing: Zhonghua Book Company.

[14] Li, L. Y. (2015). *The study of Rugao dialect's phonetic system*. M.A. Thesis, Nanjing University.

[15] Wu, F. S. (2006). *Studies of Rugao dialect (Rugao fangyan yanjiu)*, Beijing: China Federation of Literary and Art Circles Publishing House.

[16] Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

[17] Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33.

[18] Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.

[19] Chen, Y., Xu, Y., & Guion-Anderson, S. (2014). Prosodic realization of focus in bilingual production of Southern Min and Mandarin. *Phonetica*, *71*(4), 249–270.