Acoustic Analysis of Nasal Vowels in Monguor Language

Hanbin Zhang

School of Foreign Languages, Northwest Minzu University, Lanzhou730030, China

E-mail:zhanghb06@163.com

Abstract. The purpose of the study is to analyze the spectrum characteristics and acoustic features for the nasal vowels [ɑ̃] and [ɔ̃] in Monguor language. On the base of acoustic parameter database of the Monguor speech, the study finds out that there are five main zero-pole pairs appearing for the nasal vowel [ɑ̃] and two zero-pole pairs appear for the nasal vowel [ɔ̃]. The results of regression analysis demonstrate that the duration of the nasal vowel [ɑ̃] or the nasal vowel [ɔ̃] can be predicted by its F1, F2 and F3 respectively.

1. Introduction

When a vowel is pronounced, if the soft palate lowers down and oral and nasal passages open at the same time, the vowel will take nasal feature. Under the circumstance, the vowel is called nasal vowel [1]. A nasal vowel is shown with an additional symbol "˜" over a vowel in the international phonetic alphabet. When a nasal vowel is pronounced, the breath flows from the mouth and nose at the same time. Here are some examples of various kinds of language. In mandarin, [ə̃] in "DengEr" [tə̃˞55] is a nasal vowel, but [ɤ] in "Deng" [tɤŋ55] is not a nasal vowel. Polish ¥ę¥ is the nasalized sound of [e]. French letter combinations "an", "am", "en" and "em" are nasal vowels in the general case. Portuguese letter combinations "am" and "in" are also nasal vowels.

Studies on nasal vowels based on acoustic analysis are mainly concerning French, Portuguese, Hindi, Japanese, and some Chinese dialects. For example, the study on lingual differences in Brazilian Portuguese oral and nasal vowels found that the lingual configurations are in line with the effects of vowel nasalization on formant frequencies [2]. The investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of French showed that lingual and pharyngeal articulatory configurations enhance formant-frequency-related acoustic effects associated with nasalization [3]. A research on the oral articulations of French nasal vowels suggests that many of the oral articulatory configurations of Northern Metropolitan French nasal vowels enhance the acoustic effect of velo-pharyngeal coupling on F1 and F2 frequencies [4].

In Monguor language, when "a" and "o" appear before velar nasal sound [ŋ], "a" and "o" will become nasal vowels [ɑ̃] and [ɔ̃] respectively [5]. For example, while Monguor word "hairghan" (cute) is pronounced as [xaɪʐGan], "rghang" (power) is pronounced as [ʐGɑ̃ŋ]; while "manqog" (cockscomb) as [mʌntɕʰɔk], "manglii" (forehead) as [mɑ̃ŋli:]; while "jog" (aspect) as [tɕɔk], "jong" (hundred) as [tɕɔ̃ŋ]; while "rzogla" (overturn) as [ʐtsɔklʌ], "zongna" (honey) as [tsɔ̃ŋna:]. In Monguor language, the appearance of nasal vowels [ɑ̃] and [ɔ̃] has their particular conditions and rules, but nasal vowels [ɑ̃] and [ɔ̃] have no the function of distinguishing the meaning of words.

This paper analyzes the nasal vowels [ɑ̃] and [ɔ̃] in Monguor language from the respect of acoustic features, such as duration and formants, especially the first formant (F1), the second formant (F2) and the third formant (F3), in which F1 corresponds to the height of the tongue, F2 the front or back of the tongue, and F3 the lip rounding [1].

2. Method

2.1 Source of Corpus

An acoustic parameter database of the Monguor speech is built for this study. The Monguor words recorded for this study are just the words containing the nasal vowels [ɑ̃] and [ɔ̃]. There are 34 Monguor words with the nasal vowel [ɑ̃] and 37 Monguor words with the nasal vowel [ɔ̃]. These words are composed of monosyllable, disyllable, trisyllable, and quadrisyllable words.

2.2 Speech Signal Collection

A Dell Notebook, a Behringer recording microphone, and a sound card of YAMAHA Steinberg are used in this study as the recording equipment. The recording sampling rate is 44.1 kHz with 16 bits resolution ratio. The sounds are saved with *.wav files. The sounds are recorded in a recording studio. Each word is read three times.

2.3 Speakers

One speaker, male, aged 45 reads all the selected words. The speaker serves as a teacher in a local primary school. His speech belongs to Tianzhu dialect, similar to the Huzhu dialect.

2.4 Research Questions

The questions addressed in this study are:

(1) What are the spectrum characteristics for the nasal vowels [ɑ̃] and [ɔ̃]?

(2) What are the acoustic features for the nasal vowels [ɑ̃] and [ɔ̃]?

3. Results And Discussion

3.1 Spectrum Characteristics

When a sound comes from a branching acoustic tube, it has not only zero point but also pole point on its spectrogram [1]. When a nasal vowel is pronounced, the nasal cavity coupled with the oral cavity comes into being a branching acoustic tube. Figure 1 shows the spectrums of /a/ for [ɑ] before nasalization in Fig. 1 (A), and for [ɑ̃] after nasalization in Fig. 1 (B).

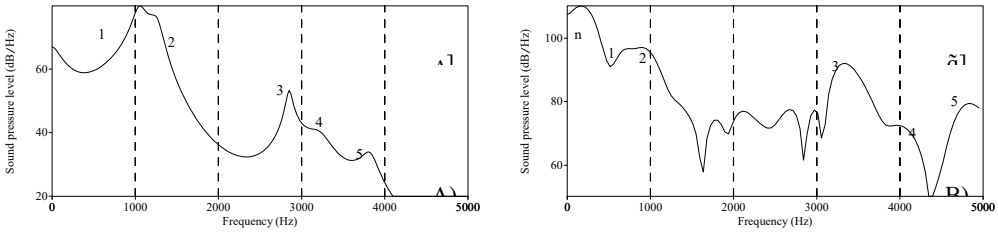

FIGURE 1. Spectrums of [ɑ] and [ɑ̃]

In Fig. 1 (A), F1, F2, F3, F4 and F5 of the oral vowel [ɑ] descend successively in the sound pressure level. However, in Fig. 1 (B), a zero point appears on the left of F1 of the nasal vowel [ɑ̃]. At the point of around 200Hz, a new formant comes into being, which is called nasal formant (Fn). Between F2 and F3, a group of zero points appear, and between F4 and F5, there also exists a zero point. Figure 1 (B) shows there are five main zero-pole pairs appearing for the nasal vowel [ɑ̃]. For the first zero-pole pair, the zero point is about 530Hz and the pole point is about 160Hz. The zero point is about 1630Hz and the pole point is about 1450Hz for the second zero-pole pair. The zero point is about 2840Hz and the pole point is about 2670Hz for the third zero-pole pair. The zero point is about 3060Hz and the pole point is about 2970Hz for the fourth zero-pole pair. The zero point is about 4380Hz and the pole point is about 4070Hz for the fifth zero-pole pair.

In Fig. 2 (A), F1, F2, F3 and F4 of the oral vowel [ɔ] are almost the same in the sound pressure level. However, in Fig. 2 (B), F1, F2, F3 and F4 of the nasal vowel [ɔ̃] vary greater. Two zero-pole pairs appear

on the left of F1 of the nasal vowel [ɔ̃]. At the point of around 400Hz, a nasal formant (Fn) comes into being. The zero point is about 230Hz and the pole point is about 40Hz For the first zero-pole pair. The zero point is about 950Hz and the pole point is about 400Hz for the second zero-pole pair.

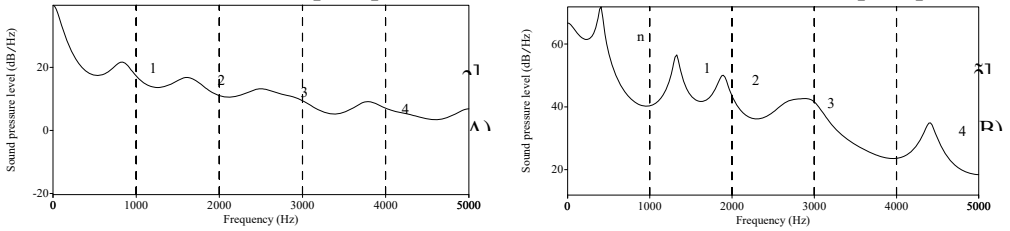

FIGURE 2. Spectrums of [ɔ] and [ɔ̃]

3.2 Acoustic Features

Acoustic features mainly include formants and duration in this study. Figure 3 shows the formants for [ɑ̃] and [ɔ̃]. F1 and F2 stay close for both [ɑ̃] and [ɔ̃], while F3, F4 and F5 stay close for [ɑ̃], and F3 and F4 stay close for [ɔ̃].

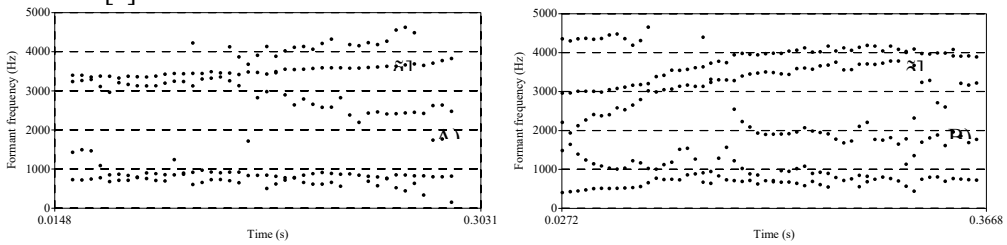

FIGURE 3. Formants for [ɑ̃] and [ɔ̃]

Table 1 shows the descriptive statistics of formants and duration for [ɑ̃]. F1, F2 and F3 of [ɑ̃] are 712Hz, 1442Hz, and 2769Hz respectively. The mean duration of [ɑ̃] is 276 milliseconds (ms).

TABLE 1. Descriptive statistics for [ɑ̃]

	Mean	**Std. Deviation**	**Std. Error**	**Range**
F1	712	126	22	552
F2	1442	211	36	717
F3	2769	205	35	734
Duration	276	88	15	357

The standard deviation for F1, F2 and F3 of [ɑ̃] is 126Hz, 211Hz, and 205Hz respectively. The standard deviation for duration is 88ms. The standard error for F1, F2 and F3 of [ɑ̃] is 22Hz, 36Hz, and 35Hz respectively. The standard error for duration is 15ms. The range for F1, F2 and F3 of [ɑ̃] is 552Hz, 717Hz, and 734Hz respectively. The range for duration is 357ms.

The result of ANOVA shows that there exists a regression relationship among F1, F2, F3 and duration of the nasal vowel [ɑ̃] ($F = 135$, $p = 0.000 < 0.05$). The result of regression analysis demonstrates that the relationship among F1, F2, F3 and duration of the nasal vowel [ɑ̃] can be the equation:

$$D = -0.212F1 + 0.153F2 + 0.1F3 \quad (1)$$

The equation (1) shows that the duration of the nasal vowel [ɑ̃] can be predicted by F1, F2 and F3. The coefficient of F1 is -0.212, the coefficient of F2 0.153, and the coefficient of F3 0.1. F1, F2 and F3 are the independent variables, and duration is the dependent variable. In Fig. 4 (A), the scatterplot verifies the prediction.

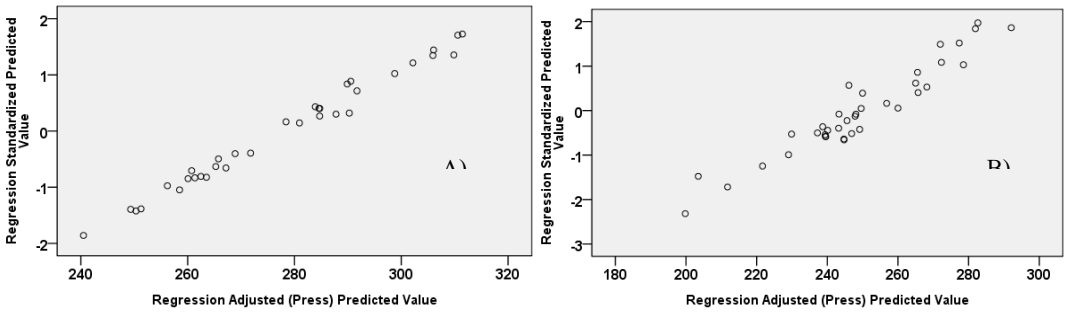

FIGURE 4. Scatterplots of formants and duration for [ã] and [õ]

Table 2 shows the descriptive statistics of formants and duration for [õ]. F1, F2 and F3 of [õ] are 720Hz, 1558Hz, and 2907Hz respectively. The mean duration of [õ] is 249 milliseconds (ms). The standard deviation for F1, F2 and F3 of [õ] is 142Hz, 276Hz, and 206Hz respectively. The standard deviation for duration is 75ms. The standard error for F1, F2 and F3 of [õ] is 23Hz, 45Hz, and 34Hz respectively. The standard error for duration is 12ms. The range for F1, F2 and F3 of [õ] is 513Hz, 1023Hz, and 818Hz respectively. The range for duration is 316ms.

TABLE 2. Descriptive statistics for [õ]

	Mean	**Std. Deviation**	**Std. Error**	**Range**
F1	720	142	23	513
F2	1588	276	45	1023
F3	2907	206	34	818
Duration	249	75	12	316

The result of ANOVA shows that there also exists a regression relationship among F1, F2, F3 and duration of the nasal vowel [õ] ($F = 151$, $p = 0.000 < 0.05$). The result of regression analysis demonstrates that the relationship among F1, F2, F3 and duration of the nasal vowel [õ] can be the equation:

$$D = -0.072F1 + 0.017F2 + 0.096F3 \quad (2)$$

The equation (2) shows that the duration of the nasal vowel [õ] can also be predicted by F1, F2 and F3. The coefficient of F1 is -0.072, the coefficient of F2 0.017, and the coefficient of F3 0.095. F1, F2 and F3 are the independent variables, and duration is the dependent variable. In Fig. 4 (B), the scatterplot also verifies the prediction.

4. Conclusion

On the base of acoustic parameter database of the Monguor speech, the study discusses the spectrum characteristics and the acoustic features for the nasal vowels [ã] and [õ]. The study has found that there are five main zero-pole pairs appearing for the nasal vowel [ã]. A zero point appears on the left of F1. A new formant, a nasal formant, comes into being at the point of around 200Hz. Between F2 and F3, three obvious zero-pole pairs appear, and the fifth zero-pole pair emerges between F4 and F5. Two zero-pole pairs appear on the left of F1 of the nasal vowel [õ]. A new nasal formant comes into being at the point of around 400Hz. F1, F2 and F3 of [ã] are 712Hz, 1442Hz, and 2769Hz respectively. The mean duration of [ã] is 276ms. F1, F2 and F3 of [õ] are 720Hz, 1558Hz, and 2907Hz respectively. The mean duration of [õ] is 249ms.

The results of regression analysis demonstrate that the duration of the nasal vowel [ã] or the nasal vowel [õ] can be predicted by F1, F2 and F3. The relationship among F1, F2, F3 and duration for the nasal vowels [ã] and [õ] can be the equations: [ã]: $D = -0.212F1 + 0.153F2 + 0.1F3$； [õ]: $D = -0.072F1 + 0.017F2 + 0.096F3$.

Monguor language is disappearing nowadays. As professor Wang Yuanxin [6] says, a certain minority language is not only the historical creation of the nation, but also the excellent culture of the nation. It is also the common cultural wealth of all mankind. The main significance of the study lies in

language application and protection. The disappearance of a large number of languages and a large number of unique cultures, without a doubt, is a great loss of human civilization. Therefore, the purpose of the study is for the preservation of the precious minority language. The results of the study can be applied in speech recognition and speech synthesis.

Acknowledgments
This research was supported by Fund Yxm2015014 to Northwest Minzu University.

Refrences
[1] H. Q. Bao and M. C. Lin, *Essentials of Experimental Phonetics* (Peking University Press, Beijing, 2014).
[2] M. Barlaz, M. J. Fu, Z. P. Liang, S. Ryan, S. Brad, "Lingual Differences in Brazilian Portuguese Oral and Nasal Vowels: An MRI Study," In Proceedings of the 18th International Congress of Phonetic Sciences (Glasgow, Scotland, UK, 2015).
[3] C. Christopher, K. S. Ryan, M. J. Fu, Z. P. Liang, P. S. Bradley, "A real-time MRI investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of French," Journal of Phonetics 50, 34–51(2015).
[4] C. Christopher, "An acoustic and articulatory examination of the "oral" in "nasal": The oral articulations of French nasal vowels are not arbitrary," Journal of Phonetics 46, 23–33(2014).
[5] Chinggeltei. *The Monguor language and Mongolian,* (Inner Mongolia people's publishing house, Hohhot, 1991).
[6] Y. X. Wang, "The investigation and protection of the vulnerable and endangered language of the less populous ethnic groups," Journal of Xinjiang Normal University, (2008).