http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 13 (11): 1793-1799, 2014
ISSN 1812-5638 / DOI: 10.3923/itj.2014.1793.1799

The Acoustics Properties of the Nasals and Nasalization in Standard Chinese

Jian Li and Le Cheng
School of Foreign Languages, China University of Political Science and Law, 100088, China

Abstract: The acoustic effect of nasal coupling is well understood, but quantification of the extent of acoustic nasalization is difficult. This study attempts to develop an objective approach to investigate the acoustic properties of the nasals and vowel nasalization in Standard Chinese (SC). The amount of acoustic nasal coupling is quantified based on the A1-H1 (amplitude difference between the first formant and the first harmonics) and A1-P1 (amplitude difference between the first vowel formant and the nasal formant around 1000 Hz) measures. Results show that these two indices successfully distinguish between the vowels in oral and different nasal contexts. The range, rate and period of their changes provide significant information for the evaluation of nasalization.

Key words: Acoustic, nasal, nasalization, standard Chinese

INTRODUCTION

The acoustic study of nasality has a long history and the results are complicated and diversified. There is a general agreement that the principal and most consistently observed modifications in the acoustic spectrum of nasalized vowels are at the low frequency, in the region of the first formant (F1). When a vowel is nasalized, the F1 intensity is often reduced with its bandwidth increased and a pole-zero pair is introduced in the vicinity of F1 (Fant, 1960; Ohala, 1975; Hawkins and Stevens, 1985).

Acoustic indices of nasality have been developed for the evaluation of nasalization. For example, Huffman (1990) found that A1-H1 (the amplitude difference between F1 and the first harmonics) were correlated with listeners' perception of nasal vs. oral vowels. On the other hand, Chen (1995, 1997, 2000) found A1-P1 and A1-P0 (the amplitude differences between F1 and the two nasal formants at about 450 and 1000 Hz, respectively) to be significant acoustic cues for vowel nasalization. Chen (2000) applied A1-P1 and A1-P0 to investigate six of the simple vowels in the nasal rhymes in SC. However, such a method is found not to be effective in our pilot study which includes all the eight simple vowels allowed in the nasal rhymes in SC. Based on the past studies and our pilot study, the present experiment investigates the acoustics of the nasals and vowel nasalization in SC based on the A1-H1 and A1-P1 measures.

MATERIALS AND METHODS

Corpus: The corpus of the present study consists of A monosyllabic words with (a) vowels followed by a nasal coda, i.e. nasal rhymes (VN, V = [a ə i u y], N = [n, ŋ]), (b) vowels preceded by a nasal initial (NV, N = [n, m], V = [a i u]) and (c) the above mentioned vowels in oral context (V); B bisyllabic words with the first syllable containing a nasal rhyme and the second syllable begins with a stop (VN.S), fricative (VN.F), nasal (VN.N) or a vowel (VN.V); and C bisyllabic words with the first syllable ending with an oral vowel, followed by a second syllable beginning with a nasal initial (V.N, eg. [ta51 nau213] "brain").

The selection of such a corpus enables comparisons between (1) vowels in oral and nasal contexts, (2) pre-nasal (NV) and post-nasal vowels (VN), (3) nasal rhymes followed by different types of segments, namely, stop, nasal, fricative, or vowel and (4) tautosyllabic (VN, where a vowel is followed by a nasal coda) and heterodyllabic (V.N, where a vowel is followed by a nasal initial across syllable boundary) nasalization.

Recording and analysis: Two male and two female speakers were required to read the word lists at a normal rate of speech. The wordlist B is also recorded at a fast speech rate. For each speech rate, three repetitions were obtained. The recorded signal was digitalized at 11 kHz. The speech data were read into Praat and the onset and the end of the vowels and the nasals are labeled in an annotation file. A V:N boundary is defined by visual inspection of the waveform and the spectrographic display with a reference to the author's perceptual judgment.

Following the procedure described by Chen (2000), the position of V:N boundary is decided in a more objective way. Figure 1a shows the narrow band spectrogram of [tan] in [tan.ta].

Corresponding Author: Le Cheng, School of Foreign Langnages, China University of Political Science and Law, 100088, China

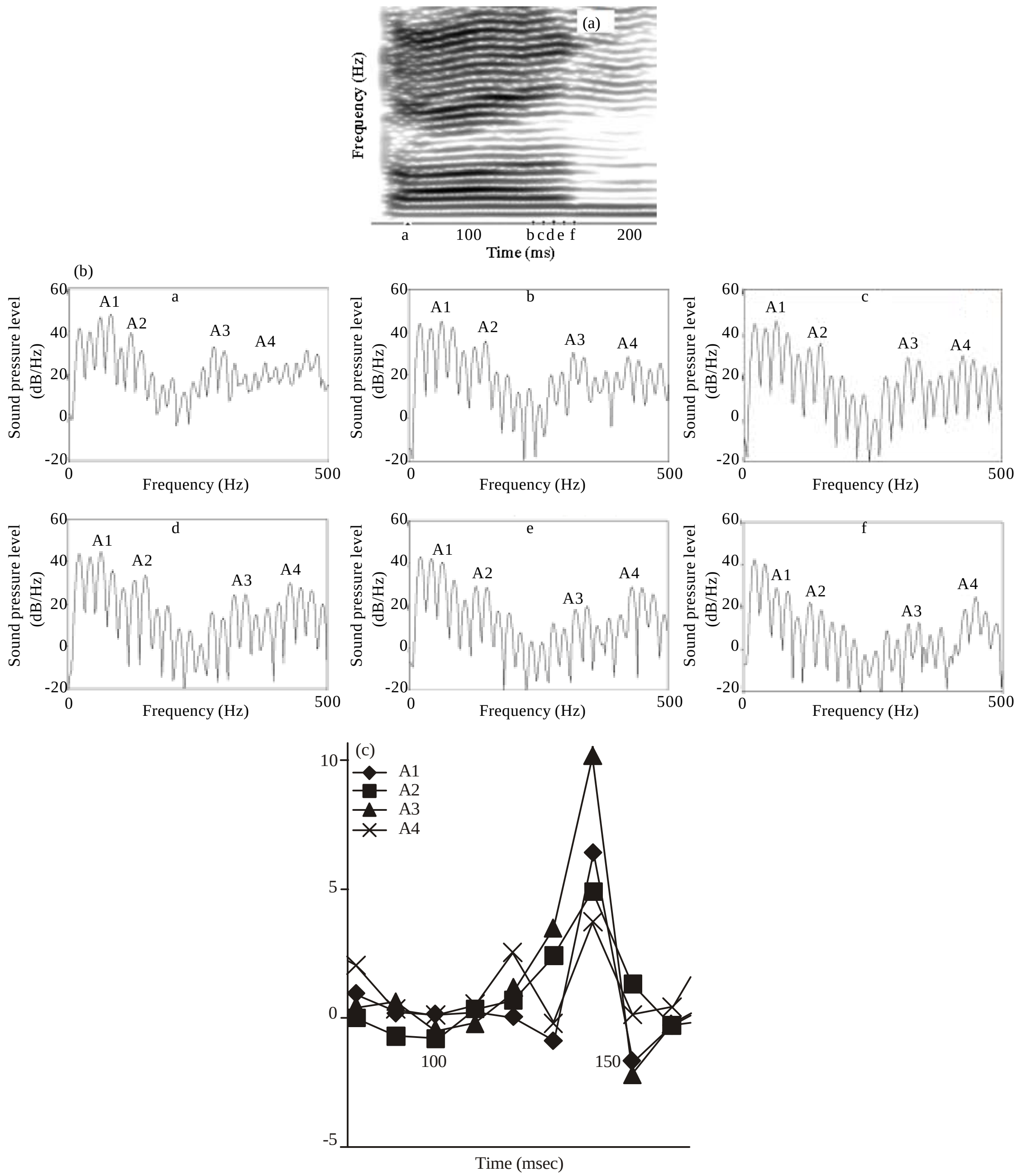

Fig. 1(a-c): (a) The narrow-band spectrogram of [tan], (b) FFT spectra at the six points labeled a-f in the spectrogram (A1, A2, A3 and A4 indicate the amplitude of the first four formants) and (c) Time contours of the first-difference amplitude of the first four formants

FFT spectra analyses were performed at every 10 msec of the whole VN sequence using Praat, with the analyzing window set to 25 msec and the prediction order set to 12 msec. The amplitude of the first harmonics (H1), the nasal formant at around 1000 Hz (P1) and the four vowel formants (A1, A2, ...A4) are obtained through a script composed with Praat. For the sake of illustration, six of the spectra are shown in Fig. 1b with the amplitudes of the first four formants labeled as A1, A2, A3 and A4. After obtaining the formant amplitude values, the first-difference amplitude is calculated as the formant amplitude at a given time minus the corresponding amplitude sampled 10 msec later. Then the time contour of "the first difference amplitude" was plotted (Fig. 1c) and the maximum first-difference for each formant was found. The maximum first-difference amplitudes of the four formants in Fig. 1 are at about 150 msec.

Threshold: Measurement of first-difference amplitude was first made in VN.S words, where a distinct V:N boundary can always be found. Then the maximum first-difference amplitude in VN.S is compared with that of the oral vowel (V), where no V:N boundary is expected. If the maximum first-difference amplitude in VN.S and V show a significant difference for a given vowel across repetitions by a given speakers, the maximum first-difference amplitude averaged across repetitions was used as a threshold for VN boundary detection for the given vowel by the given speaker. For each vowel, nasal coda and speaker, there would be four thresholds for the four formants. If the maximum first-differences of any of the four formants of a test token are above the relevant thresholds, the point where the maximum first differences are determined is defined as the V:N boundary. Only when the maximum first-differences of all the four formants are below the relevant thresholds, the V:N boundary is not determined and it is assumed that there is no oral closure for the nasal ending.

RESULTS

Results about the V:N boundary: The maximum first differences of the four formants for the eight VN combinations in VN.S context (Table 1) are compared with the corresponding vowels in pure oral context. The significant comparison results based on independent-sample t-tests in the Table 1 ($p<0.05$). It can be seen that the maximum first difference amplitude of any one of the four formant amplitudes, except for the A1 of the high vowels, can be used to distinguish an oral vowel from a vowel in nasal context.

The duration of the vowel and the nasal coda of the nasal rhymes are measured based on the V:N detection technique. Figure 2 shows the averaged relative duration of the vowels and the nasals in the nasal rhymes in normal speech rate.

Table 1: Maximum first difference of the first four formant amplitudes (DA1, DA2, ...DA4) of the nasal rhymes in /VN.S/ averaged across three repetitions by four speakers (M = mean S = standard deviation)

Nasal rhymes		DA1	DA2	DA3	DA4
an	M	9.6*	9.6*	12.4*	10.7*
	S	3.8	2.5	2.5	3.4
aŋ	M	7.5*	9.9*	10.5*	9.9*
	S	2.8	2.9	2.3	3.1
ən	M	7.5*	8.3*	11.3*	10.8*
	S	1.7	1.6	3.5	2.2
əŋ	M	5.6*	8.2*	9.3*	10.7*
	S	1.2	3.4	3.1	3.5
in	M	3.5	10.9*	11.8*	10.2*
	S	3.2	3.2	3.7	3.6
iŋ	M	4.2	11.9*	12.2*	12.2*
	S	3.9	3.2	4.6	3.1
uŋ	M	4.6	8.4*	7.1*	8.6*
	S	1.9	4.3	2.3	2.1
yn	M	3.3	7.4*	10.1*	13.8*
	S	2.6	2.0	3.1	4.5

*Significant at $p<0.05$ using t-test

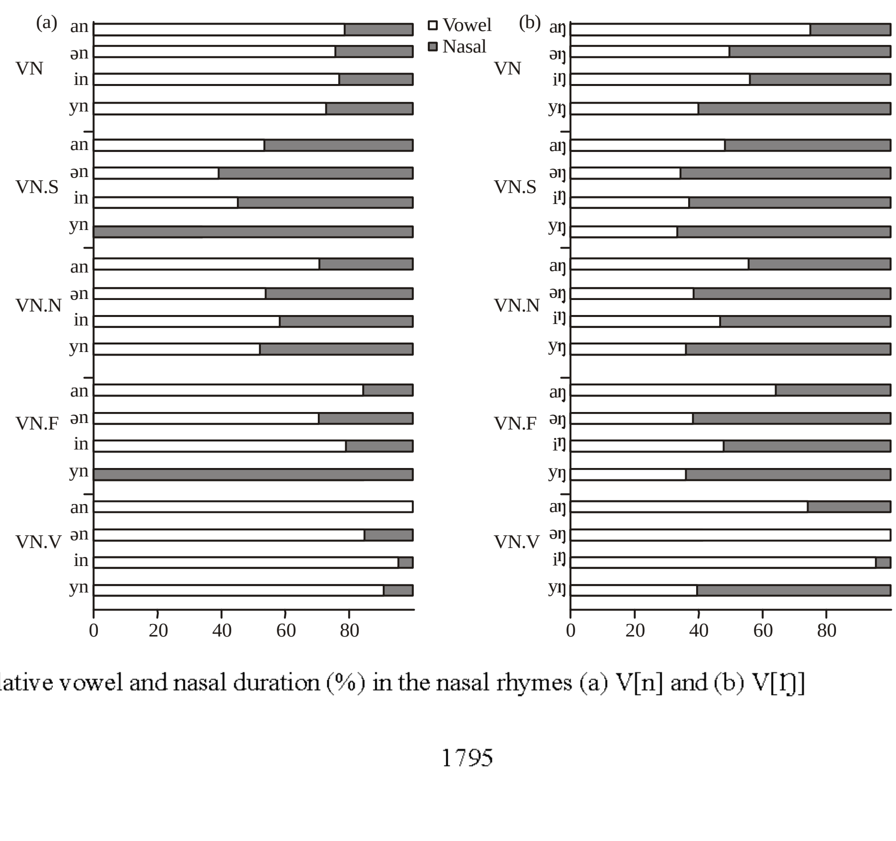

Fig. 2(a-b): Relative vowel and nasal duration (%) in the nasal rhymes (a) V[n] and (b) V[ŋ]

In VN in monosyllabic words, a V:N boundary is usually detected. The velar nasal is longer than the alveolar nasal which has been documented in the literature (Chao, 1968). In addition, the velar nasal following a low vowel [a] has a shorter duration than when it follows the other vowels.

In VN.S and VN.N, a V:N boundary is always detected which imply that the nasal coda in the nasal rhymes is always produced with an oral closure when they are followed by a stop or a nasal consonant. In VN.F, a V:N boundary is detected for [an, aŋ, en, in, yn] in some cases, but is absent in the rest. A V:N boundary is always detected in [iŋ, əŋ, uŋ]. In VN.V, the nasal coda in [an, aŋ, ən, in, yn] are seriously shortened or deleted.

In sum, the nasal coda is always produced with an oral closure when followed by a stop or a nasal consonant. The nasal coda is subject to shortening or even deletion (as indicated by the absence of V:N boundary) when followed by a vowel or a fricative. The velar nasal tends to have longer duration and is less likely to be shortened or deleted than the alveolar nasal. Low vowel is another factor to trigger nasal shortening or deletion.

Acoustic evaluation of nasalization: Figure 3 shows the range of nasal coupling during vowel production, as indicated by the range of A1-H1. It can be seen that A1-P1 or A1-H1 makes a good distinction between the vowels in oral and nasal contexts. In the oral context, the A1-H1 or A1-P1 values are relatively high and the ranges of A1-H1 or A1-P1 change are relatively small. In the nasal context, the A1-H1 or A1-P1 values are relatively low and

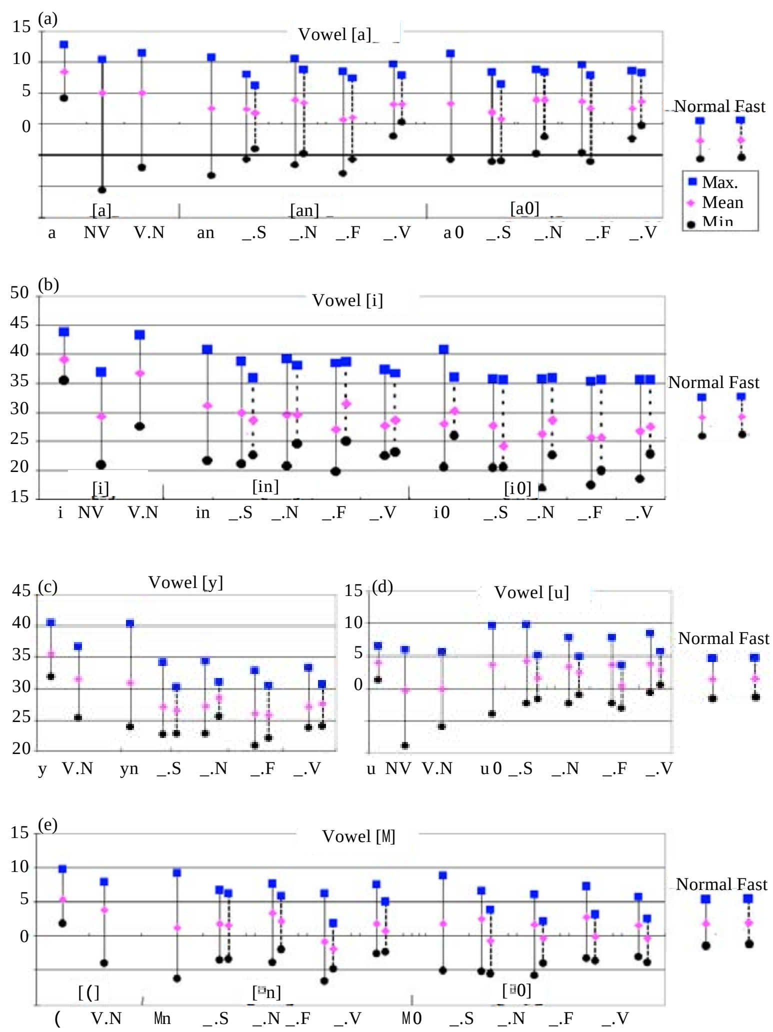

Fig. 3(a-e): The range of A1-H1 (reflected by the maximum, mean and minimum values) averaged across repetitions for the vowels (a) a, (b) i, (c) y, (d) u and (e) ə in different contexts (Note: _.S, _.N, _.F and _.V refer to VN.S, VN.N, VN.F and VN.V contexts, respectively)

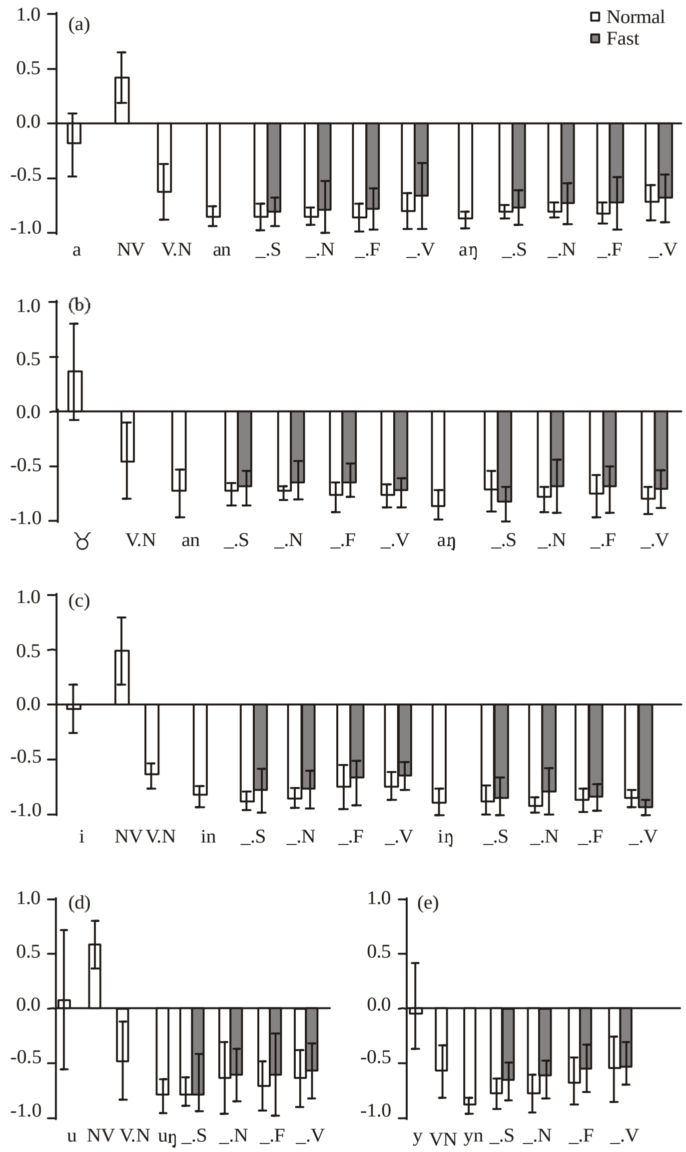

Fig. 4(a-e): Slope of the A1-P1 or A1-H1 time contour (Note: _.S, _.N, _.F and _.V refer to VN.S, VN.N, VN.F and VN.V contexts, respectively) (a) a, (b) e, (c) i, (d) y and (e) u

the A1-H1 or A1-P1 range is large. The A1-P1 range is generally comparable for the vowels in different nasal context (pre-nasal NV or post-nasal VN, tautosyllabic V.N or heterosyllabic VN). When speech rate increases, the range of A1-H1 or A1-P1 change become smaller.

Figure 4 shows the results of the slope of A1-H1 or A1-P1 change as an indicator of the rate of nasalization, or in articulatory terms, the rate of velopharyngeal opening/closing gesture.

It can be seen from Fig. 4 that the slope of A1-P1 or A1-H1 change can successfully distinguish among oral vowels, pre-nasal vowels and post-nasal vowels. The slope for the vowels in oral context can be either positive or negative, with the average value relatively close to zero. The slope in NV is always positive which means that the initial parts of the vowels are nasalized and the degree of nasalization decreases into the later parts of the vowels. The slope for vowel+nasal is always negative, indicating increasing nasalization into the vowels due to anticipation of the nasal coda.

The slope is less steep (with smaller absolute value) in heterosyllabic context than it is in taotosyllabic context which imply that the rate of nasal coupling is faster when a vowel is followed by a nasal coda than when if it is

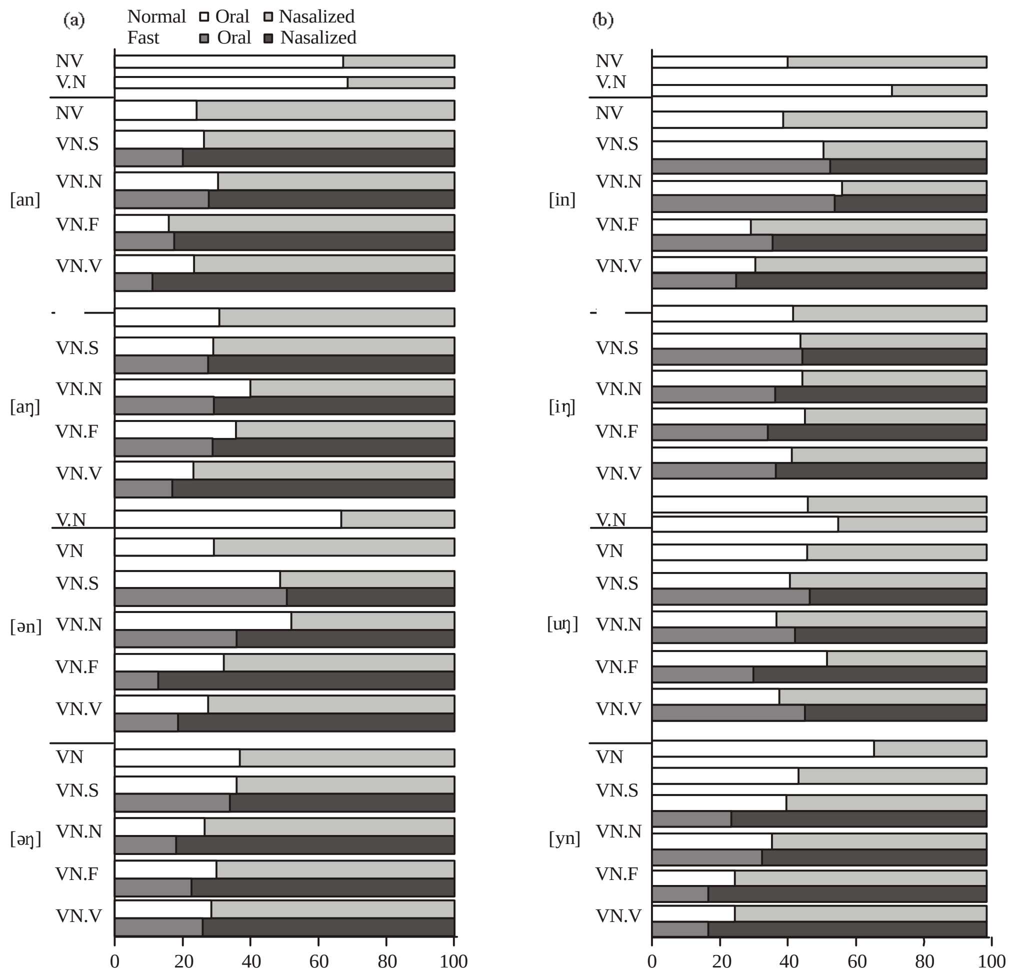

Fig. 5(a-b): Relative duration of vowel nasalization (in percentage) in different VN combinations

followed by a nasal initial across syllable boundary. In addition, the slope is steeper, hence the rate of vowel nasalization is higher, in VN.S and VN.N (where the nasal coda is realized as nasal murmur) than in VN.V and VN.F (where the nasal murmur of the underlying nasal coda is not always realized) contexts, or in normal speech rate than in fast speech rate.

Figure 5 plots the percentage of vowel nasalization (PVN = duration of nasalized vowel portion/overall vowel duration)for the vowels in nasal contexts.

The onset of vowel nasalization is defined as the point where A1-H1 or A1-P1 exceeds the average A1-H1 or A1-P1 value of the corresponding vowel in oral context. Figure 5 shows that PVN is the highest for the low vowel [a], followed by the schwa and the vowel [y], then followed by the vowel [i] and [u]. One-way ANOVA show that the effect of vowel identity on the percentage of vowel nasalization is significant ($p<0.05$). *Post-hoc* t tests show that the percentage of vowel nasalization of the vowel [i] and [u] are significantly lower than the vowels [a, y] and the schwa. The differences among the vowels [a, y] and the schwa are not are not significant. PVN is higher when a vowel is followed by a nasal coda than when it is followed by a nasal initial across syllable boundary. PVN for post-nasal vowels is characterized by extensive variation both within and across speakers. The average percentage of nasalization is between VN and V.N contexts, but is not significantly different from either of them.

CONCLUSION

This study uses a quantitative technique suggested by Chen (2000) to determine the V:N boundary. The V:N boundary is determined by using the maximum first-difference amplitude obtained from the first four formant amplitude time contours in the vowels and the adjacent nasal consonants. The boundary determined according to this quantitative method is highly correlated with the boundary measured in the traditional subjective way, if the presence of a V:N boundary is known a priori. The significance of such a quantitative method is reflected in analyzing the nasal rhymes where the nasal murmur may be absent and thus a V:N boundary is absent. This method offers an objective way to determine whether there is a V:N boundary and where the boundary is. The results show that in the context of VN.F and VN.V, a V:N boundary is often absent, especially when the nasal coda is alveolar or when the velar nasal is preceded by a low vowel [a]. For such cases, other methods are needed to reflect the presence of the underlying nasal coda. A combination of A1-H1 and A1-P1 is adopted to quantify the extent of coarticulatory nasalization of the vowels. The results indicate that all the five vowels in the nasal contexts show greater extent and faster rate of vowel nasalization than their counterparts in oral context. The present results also show that anticipatory vowel nasalization is greater when a vowel is followed by a nasal coda than when it is followed by a nasal initial across syllable boundary, as indicated by faster rate of nasalization and a longer period of nasalization. Such a result is in agreement with the result by Chen (2000). In addition, a comparison between pre-nasal and post-nasal vowels show that the extent of nasal coupling is similar for these two groups, but the rate of nasalization is faster and the duration of nasalization is longer for the former which indicate that the extent of anticipatory nasalization is greater than progressive nasalization.

REFERENCES

Chao, Y.R., 1968. A Grammar of Spoken Chinese. University of California Press, Berkeley, Los Angeles.

Chen, M.Y., 1995. Acoustic parameters of nasalized vowels in hearing impaired and normal-hearing speakers. J. Acoust. Soc. Am., 98: 2443-2453.

Chen, M.Y., 1997. Acoustic correlates of English and French nasalized vowels. J. Acoust. Soc. Am., 102: 2360-2370.

Chen, M.Y., 2000. Acoustic analysis of simple vowels preceding a nasal in standard Chinese. J. Phonetics, 28: 43-67.

Fant, G., 1960. Acoustic Theory of Speech Production. Mouton de Gruyter, Bordeaux, France.

Hawkins, S. and K.N. Stevens, 1985. Acoustic and perceptual correlates of the non-nasal distinction for vowels. J. Acoust. Soc. Am., 77: 1560-1575.

Huffman, M.K., 1990. The role of F1 amplitude in producing nasal percept. J. Acoust. Soc. Am., 88: S54-S54.

Ohala, J.J., 1975. Phonetic Explanations for Nasal Sound Patterns. In: Nasalfest: Papers from a Symposium on Nasals and Nasalization, Ferguson, C., L. Hyman and J.J. Ohala (Eds.). Stanford University, Stanford, California, pp: 289-316.