BIDIRECTIONAL INTERACTION BETWEEN TONE AND SYLLABLE CODA: ACOUSTIC EVIDENCE FROM CHINESE.

Phil Rose

Department of Linguistics (Faculties), Australian National University.

ABSTRACT: The acoustic phonetic interaction between tone and syllable-final velar nasal coda is examined for a dialect of Chinese. A bidirectional effect is demonstrated in that the nasal coda causes significant differences in the tonal duration and F0 height of its syllable, whilst at the same time having its duration conditioned by the tone.

INTRODUCTION

With few exceptions (e.g. Zee 1980), work on the phonetic interaction between tone and segmentals (e.g. Hombert 1978) has generally concentrated on the effect of segments on tone, rather than vice-versa; and specifically the effect on tone of features on syllable-*onset* segments. This paper examines the possibility of the converse of these effects, namely effect of tone on segment; and effect of syllable-*coda* on tone. It was prompted by the observation, during my initial transcription of a Chinese dialect, that pitches on some tones on syllables with final nasal consonants were higher than on syllables without such a nasal, and that the duration of the nasal sounded longer in some tones than in others. This paper reports a small part of the results of a detailed acoustic investigation into the interaction between segments, tone and register in a Chinese dialect. It examines specifically the mutual interaction between the phonological categories of tone and syllable-final velar nasal coda, as reflected in their acoustic properties of fundamental frequency and duration.

The dialect under examination, Zhenhai, has four tones which allow a Rhyme contrast between a long monophthong, and a short "lax" vowel plus velar nasal coda (written 'ng' below), e.g. [ti: 324] 'resist' vs. [tɪng 324] 'top'. These will be referred to below as 'oral' and 'nasal' syllables respectively. The citation pitch values of the four tones are: high falling (Tone 1); mid concave (Tone 2); low convex (Tone 3); and low concave (Tone 4). In tone sandhi they can have a wide variety of additional values, including high, mid, and low level; high and low rising; high and low falling (Rose 1990:6).

EFFECT OF NASAL CODA ON TONE

In order to investigate the effect of the nasal coda on tone, the acoustic properties of fundamental frequency (F0) and duration of tones on syllables with the nasal coda, e.g. [tɕɪng 243] 'town', were compared with those of tones with monophthongal Rhymes, e.g. [tɕ i : 243] 'flag'. Monosyllabic citation forms from a young male native speaker (NYJ) were used. The effect was examined for his four tones on syllables with voiceless unaspirated obstruent Initials. For the low convex tone 3, the effect was also examined on syllables with sonorant Initials e.g. [mɪng 243] 'bright' vs. [mi : 243] 'rice'. The duration of the tonal Rhyme was first determined from wide band spectrograms. This duration was taken to be from the onset of modal phonation to phonation offset, the latter defined as the point of discontinuity in increase of fundamental period, and/or discontinuity in peak-to-peak amplitude. F0 was then sampled at percentage points of the duration of the Rhyme from temporally aligned narrow-band spectrograms.

RESULTS

Fig.1 plots mean F0 values as a function of absolute duration for NYJ's four citation tones on syllables with and without the velar nasal. The point of onset of the velar nasal, the determination of which is described below, is also shown. (The measurements on which fig. 1 are based are in Rose (1982:12-16)). Fig. 1 shows that the velar nasal has an effect, differential with respect to tone, on both F0 and Rhyme duration. It can be seen that in the high falling and low convex tones 1 & 3 the Rhyme is shorter in nasal syllables. The magnitude of the difference is from 3 to 4 centiseconds, and *t* tests show this to be significant at at least the 0.05 level. In contrast to this, in the high and low concave

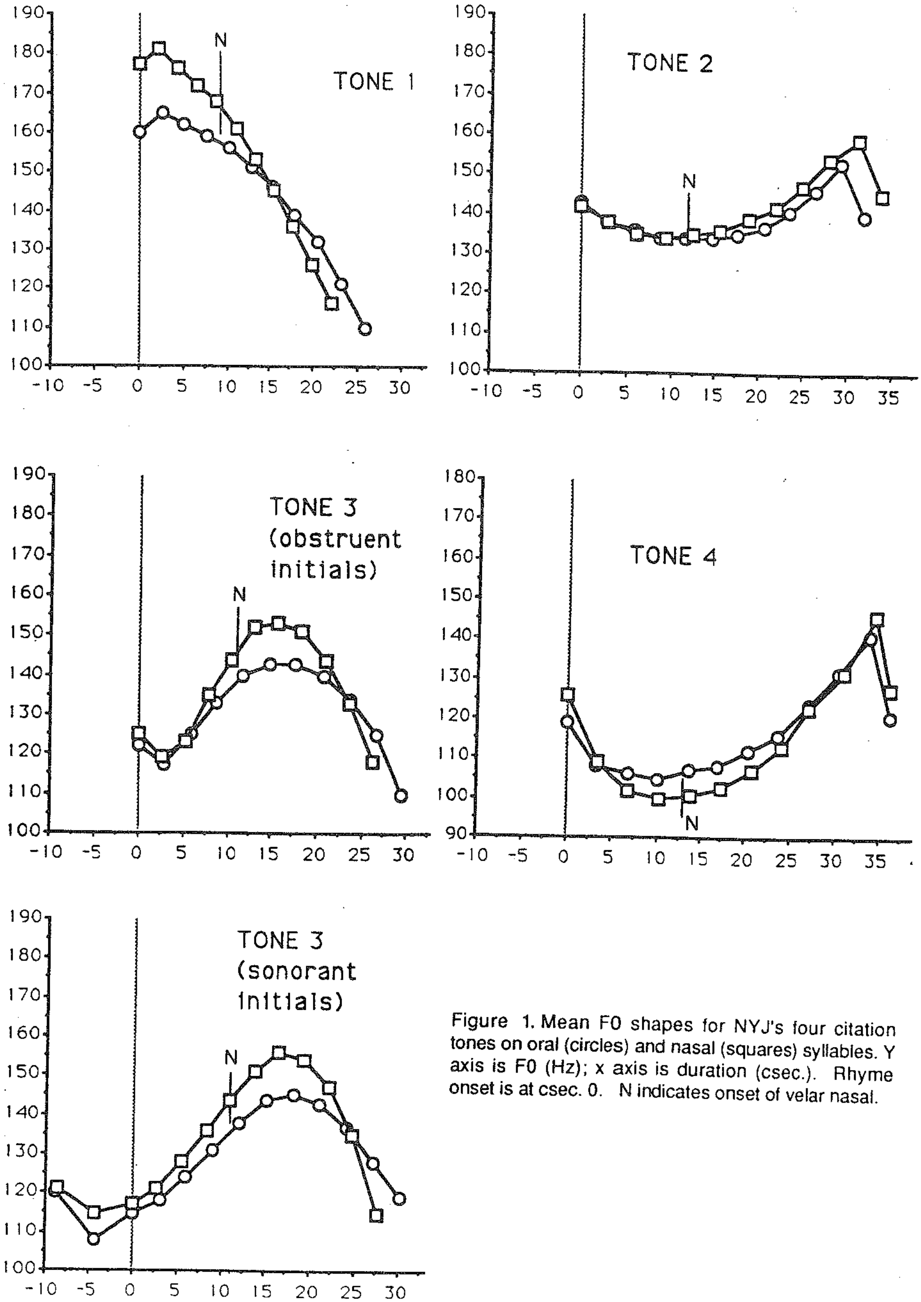

Figure 1. Mean F0 shapes for NYJ's four citation tones on oral (circles) and nasal (squares) syllables. Y axis is F0 (Hz); x axis is duration (csec.). Rhyme onset is at csec. 0. N indicates onset of velar nasal.

tones 2 & 4, durational differences between oral and nasal syllables are small, non-systematic, and non-significant.

As far as F0 is concerned, no clear differences between oral and nasal syllables can be demonstrated in the concave tones 2 & 4. F0 is slightly higher at peak in nasal syllables in both these tones, but non-significantly so. Moreover, F0 in nasal syllables is generally slightly lower over most of tone 4, but slightly higher for the latter part of tone 2. Again, none of these differences is significant. In contrast to this, tones 1 & 3 show consistently and significantly higher peak F0 values for nasal syllables. Since oral and nasal syllables differ as to duration in tones 1 & 3, the precise relationship between the nasal and oral F0 curves depends, of course, on the nature of the duration basis according to which they are being compared. If F0 curves are being compared as a function of equalised duration, it can be said that both types of syllables have the same contour, but nasal syllables have significantly higher F0 throughout almost all the duration of the Rhyme, and that the difference decreases with overall F0 height. If comparison is on an absolute basis, F0 can be said to be significantly higher in nasal syllables at peak, and nasal syllables have a more abruptly falling post-peak contour. In both cases, differences between oral and nasal syllables at peak (about 16 Hz in tone 1 and 10 Hz in tone 3) are significant at at least the 0.03 level. The syllable coda velar nasal is thus associated with significant differences in F0 height and contour, and duration, in tones 1 & 3, but not in tones 2 & 4. Tones 1 & 3 with final nasal have a higher F0 peak, more abrupt post-peak decay, and a shorter Rhyme duration. This therefore constitutes evidence for a phonetic effect of syllable-coda on tone. It is worth noting that it has been assumed (Gsell 1985:391) that the presence of a postvocalic nasal will not affect tonal F0. This study has shown that, on the contrary, in the investigation of intrinsic segmental effects on tonal acoustics it is advisable to control for a nasal coda.

On the basis of tone sandhi rules in Zhenhai phonology (Rose 1990), tones 1 & 3 constitute a natural class, as do also tones 2 & 4. The phonetic basis of this class, however, is not totally clear, since the falling and convex tones 1 & 3 appear to share only the pitch feature [+fall] (or the tonal sequence HL). The present results are interesting because they show that the four tones divide into the same two natural classes on the basis of phonetic effects, namely their sensitivity to the nasal coda influence on F0 and duration. It is thus possible that, despite their apparent difference in F0 shape, the sensitivity to the nasal effect in tones 1 & 3 reflects some shared productional feature.

EFFECT OF TONE ON NASAL CODA

It was mentioned above that the nasal coda is audibly longer in some tones. The precise relationship between tone and nasal coda duration was investigated by comparing the duration of the nasal and the whole Rhyme for the four citation tones of NYJ. The corpus for this part of the investigation consisted of 154 syllables with velar nasal coda, divided into the four tonal groups. Each tonal group was further subdivided into three groups according to the vowel phoneme (/i/, /u/, /a/) preceding the /ng/, viz: / i / [ɪ] e.g. [tɪng] 'nail'; / u / [ɷ] (tones 1 & 2) ~ [ɒ] (tones 3 & 4) e.g. [t U ng 441] 'east', [t ɒ ng 243] 'copper'; / a / [ɐ] e.g. [tɐ ng 441] 'stool'.

In order to investigate the relationship between tone and nasal coda, the onset of the nasal coda had to be identified. Laver (1980:77) notes that the velum takes 13 centiseconds to move from closure to an open state. Therefore the onset of the nasal cannot be equated with the lowering of the soft palate, *pace* phonetic or phonological representations, since this would have to occur right at the beginning of the preceding short vowel. Rather, the appropriate point of nasal onset is best specified in terms of the active articulator, namely the onset of velar occlusion as effected by the tongue dorsum. This point is clearly visible acoustically as the point at which an abrupt drop occurs in the overall wide band spectral amplitude over about 2 KHz. (The discontinuity in amplitude occurs when the oral resonator is decoupled from the nasal and pharyngeal resonators.) The nasal onset point is also usually clearly reflected in small perturbations in the F0 visible in the upper harmonics of narrow band spectrograms. In nearly all cases, the nasal onset point could be ascertained with a precision of one glottal pulse, which for NYJ is somewhat less than a centisecond, given his mean citation tone F0 of 131 Hz. The nasal onset point is also shown for the tones in fig. 1. It was pointed out above that the F0 of tones 2 & 4, with final rising pitch, shows an abrupt drop after peak in the few centiseconds before phonation offset which is associated with a characteristic syllable-final glottal stop. Because of this, it seemed advisable to distinguish two possible Rhyme duration values for the concave tones: duration to F0 peak (DP) and duration to F0 offset (DF).

RESULTS

Table 1 gives the mean and standard deviation values for Rhyme duration and point of onset of nasal coda (NASON) in NYJ's four tones. The results for each tone are given separately for the three different vowel phonemes preceding the coda, and for tones 2 & 4, separate values are given for duration to F0 peak (DP) and duration to F0 offset (DF). The duration of the velar nasal in each case can be found by subtracting the NASON value from the appropriate Rhyme duration. Thus in tone 2, the mean point of onset of velar occlusion in /ing/ syllables occured at csec.11.8, and the mean duration of the Rhyme in /ing/ syllables was 33.6 csec. The mean duration of the velar nasal for tone 2 in /ing/ syllables is therefore (33.6 - 11.8 =) 21.8 csec. If the portion of the Rhyme after the F0 peak is ignored, the duration of the velar nasal is of course somewhat shorter: (30.6 - 11.8 =)18.8 csec.

Table 1. Means and standard deviations (csec.) and number of items per sample (**X**, sd, *n*) for Rhyme duration and nasal onset in NYJ's four tones.

		/ing/	/u ng/	/ang/
Tone 1	Rhyme	22.6, 4.3, *30*	23.0, 3.8, *17*	21.6, 3.0, *10*
	NASON	9.3, 2.2	8.2, 1.5	8.9, 1.1
Tone 2	Rhyme	(DP) 30.6, 4.6, *18*	29.1, -, *1*	28.5, 4.0, *6*
		(DF) 33.6, 4.8	31.7, -	31.8, 4.0
	NASON	11.8, 2.5	11.0, -	12.7, 2.4
Tone 3	Rhyme	27.1, 2.5, *19*	27.3, 3.6, *9*	26.3, 3.2, *10*
	NASON	11.5, 1.6	11.2, 1.2	10.8, 1.5
Tone 4	Rhyme	(DP) 34.1, 5.0, *19*	35.5, 4.3, *10*	32.0, 2.4, *4*
		(DF) 37.0, 5.1	38.4, 4.6	35.1, 2.4
	NASON	13.9, 3.5	14.7, 3.0	14.7, 1.1

Table 1 shows that for tones 1 to 3 the nasal onset point occurs later (though usually not significantly so) in /ing/ and /ang/ syllables than in /ung/ syllables. This possibly reflects the relative proximity of the articulators involved, since under conditions of constant velocity it would take the tongue body longer to move from the vowel target in /i/ and /a/ to /ng/ than from /u/. The fact that nasal onset in tone 4 is not earlier for /ung/ syllables correlates with the lower target [ɒ] for /u/ in this tone. This explanation is, however, incompatible with the tone 3 NASON values, which also show earlier onset despite lower vowel allophones for /u/.

Table 1 also shows that there are large differences in Rhyme duration depending on the tone. Tone 1, for example, is about 14 csec. -- or three-fifths -- shorter in Rhyme duration than tone 4. One-way ANOVA on the mean Rhyme duration values for the four tones shows that each tone has significantly different Rhyme durations at at least the 0.05 level. From the shortest to the longest they are: high falling *Tone 1* **22.4** csec. < low convex *Tone 3* **26.9** csec. < high dipping *Tone 2* **29.4** csec. (DP) / **32.4** (DF) < low dipping *Tone 4* **33.9** csec. (DP) / **36.8** (DF). These durational relationships are also evident in fig.1, which reflects a smaller corpus non-differentiated with respect to vowel quality.

It is also clear that the duration of the nasal coda shows large differences depending on the tone: the nasal in tone 1, for example, is about 8 to 10 csec. shorter than in tone 4. One-way ANOVA on the mean duration of the nasals followed by Tukey's test showed that the duration of the nasal does not differ significantly between the longer tones 2 & 4, or between tone 2 and tone 3; other differences are significant at at least the 0.01 level. This then is evidence for the phonetic effect of tone on segmentals.

What model could account for the different durations for the nasal coda, or, equivalently, the timing of the velar occlusion, in the different tones? It is clear that the differences in the duration of the nasal coda are not the result of an interaction between different Rhyme durations (intrinsically determined by ease of production of tonal F0 shapes) and a velar occlusion timed to occur at some value independent of the Rhyme duration. This is because, except for tones 2 and 3, the NASON values show significant differences between tones at at least the 0.01 level (ANOVA followed by Tukey"s test). One obvious approach is to see whether the timing of the velar occlusion relates to the Rhyme duration. To investigate this, NASON was regressed linearly on the durations of the Rhyme for all syllablus in the corpus (DP measurements were used for tones 2 & 4). This gave an r of 0.832 (df = 152), indicating a clearly significant linear relationship between the two variables, and a slope of 0.415, showing the mean point of nasal onset to occur at around 40% into the Rhyme duration.

The percentage of the Rhyme duration taken up by the nasal coda - the 'nasal duration ratio' - was then examined in detail for the individual tones. This was done for each syllable by subtracting the NASON value from the Rhyme (to find the duration of the nasal) and dividing this by the Rhyme duration. Mean and standard deviation nasal duration ratios are given in table 2. So for example the value of 58.9% for /ing/ in tone 1 indicates that after /i/ the velar nasal takes up 58.9% of the duration of the Rhyme; or equivalently the nasal onset is timed at 41.1% of the Rhyme. Note that in table 2 tonal means are also given. These represent pooled tokens for all three vowel conditions. Since each sample had different numbers of tokens, there is a potential distortion involved. However, comparison between these tonal means and those calculated from the means of the three vowel conditions shows the error to be minimal (a maximum of 0.6% in tone 4 for example).

Table 2. Means, standard deviations (%) and number of items per sample (X,sd,*n*) for nasal duration ratio in NYJ's four tones.

		/ing/	/ung/	/ang/	tonal mean
Tone 1		58.9, 5.2, *30*	64.5, 6.8, *17*	59.5, 6.1, *10*	60.3, 6.2, *57*
Tone 2	(DP)	61.1, 6.5, *18*	62.0, -, *1*	55.7, 2.7, *6*	59.8, 6.1, *25*
	(DF)	64.7, 5.9	65.0, -	60.3, 2.7	63.6, 5.5
Tone 3		57.6, 4.8, *19*	58.8, 5.8, *9*	58.7, 5.0, *10*	58.1, 5.0, *38*
Tone 4	(DP)	59.4, 6.8, *19*	60.7, 3.8, *10*	54.3, 4.5, 4	58.7, 6.1, *33*
	(DF)	62.7, 6.2	63.7, 3.7	58.0, 2.9	62.0, 5.5

Overall Mean (ratio of /ng/ : DF in tones 1 & 3; ratio of /ng/ : DP in tones 2 & 4): 59.4, 5.8, *154*

As can be seen from table 2, the nasal duration ratio is an effective normalising transform. The nasal duration ratios are very similar for all tones, with only about 4% difference between the highest and lowest values. T tests show that the values of 60.3% and 58.1% for the ratio expressed in terms of DF (i.e. duration of whole Rhyme) for tones 1 & 3 differ very significantly from the DF values of 63.6% and 62.0% for tones 2 & 4. However, the values of 60.3% and 58.1% for the ratio expressed in terms of DF in tones 1 & 3 do not differ from the values of 59.8% and 58.7% for the ratio expressed in terms of DP (i.e. duration to F0 peak) in tones 2 & 4. It appears, therefore, that NYJ times the onset of the velar nasal, or equivalently the duration of the short vowel, with respect to the Rhyme duration in which it occurs. Specifically, he times it with reference to the duration of the whole Rhyme (DF) in tones 1 & 3, and the duration of the Rhyme to peak F0 (DP) in tones 2 & 4, such that nasal onset is effected at two-fifths of DF tones 1 & 3, and two-fifths of DP in tones 2 & 4. (The overall mean nasal duration ratio is 59.4%, which is very near 60%, or three-fifths.)

These data provide evidence that the speech production mechanism must have good knowledge of intrinsic relationships in order to function. In order to time nasal onset successfully, NYJ must take into account at least two apparently intrinsic relationships in speech production: the variation of Rhyme duration with tonal F0 shape; and the amount of time it takes for his cords to stop vibrating in tones 2 & 4 after the neural command to produce the glottal stop has arrived at the laryngeal muscles. An obvious next question in this area of speech timing -- one that requires information on between-speaker variation in the nasal duration ratio -- is whether the nasal duration ratio constitutes a linguistic phonetic parameter. Do (tone) languages and dialects differ with respect to the timing of onset of final nasals?

REFERENCES

Gsell, R. (1985) 'Hauters Spécifiques, Types Consonantiques et Tons Statiques en Thai', in S. Ratatnakul et al. eds. *Southeast Asian Linguistic Studies presented to André-G. Haudricourt*, (Mahidol University: Bangkok), 389-427.

Hombert, J-M. (1978) 'Consonant Types, Vowel Quality, and Tone' in V. Fromkin ed. *Tone: A Linguistic Survey* (Academic Press: New York), 77-111.

Laver, J. (1980) *The phonetic description of voice quality* (Cambridge University Press: Cambridge).

Rose, P. (1982) 'Acoustic characteristics of the Shanghai-Zhenhai syllable types', in D. Bradley ed. *Tonation* Pacific Linguistics Papers in South-East Asian Linguistics 8,1-53.

Rose, P. (1990) 'Acoustics and Phonology of Complex Tone Sandhi', *Phonetica* 47, 1-35.

Zee, E. (1980) 'Peak intraoral pressure in [p] as a function of F0.' *UCLA Working Papers in Phonetics* , 80-88.