6th International Conference on Spoken Language Processing (ICSLP 2000)
Beijing, China
October 16-20, 2000

A NEW PROPOSAL OF LARYNGEAL FEATURES FOR THE TONAL SYSTEM OF VIETNAMESE

Masaaki Shimizu and Masatake Dantsuji

Center for Information and Multimedia Studies, Kyoto University,
Yoshida-Nihonmatsu-cho, Sakyo-ku, Kyoto, 606-8501, Japan

ABSTRACT

We have examined a tonal system of modern Hanoi dialect of Vietnamese. This has been described for a long time as a pitch contour prominent system with secondary phonatory variations. The present study aims to propose certain laryngeal features, [tense] and [glottal constriction], in addition to the former prosodic feature, [high], to cover the overall distinction within the system. Some acoustical and laryngographical analyses were made in order to present phonetic evidence for our proposal.

I. INTRODUCTION

Modern Hanoi capital dialect of Vietnamese possesses 6 contrastive tones which are neutralized into 2 in the syllables of stop coda, and each of them are labeled in English (Vietnamese) as LEVEL (ngang) tone, FALLING (huyen) tone, RISING (sac) tone, CURVE (hoi) tone, BROKEN (nga) tone, and DROP (nang) tone. There have been several investigations into them from an acoustic phonetic point of view, such as M. S. Han (1969), Vu Thanh Phuong (1982), Nguyen Van Loi and J. A. Edmondson (1998), etc. Most of the previous works regarded the system as pitch contour prominent, and Vu Thanh Phuong, for instance, applied the 5-level notation to each tone in order to show the phonetic property: LEVEL [33], FALLING [21], RISING [35], CURVE [212], BROKEN [325] (laryngealised), DROP [21] (laryngealised). However, in fact, in the modern Hanoi dialect there exits one most significant variant of CURVE tone which has [21] value exactly as that of FALLING tone. What guarantees the distinction between them is a certain kind of phonation type. Nguyen Van Loi and Jerold A. Edmondson (1998) analyzed the whole system from this point of view, making use of the inverse filtering method. In the present study, we attempt to go one step further than them to propose a set of prosodic plus laryngeal features to cover the whole system.

II. METHOD

Pitch contour (F0) value, speed quotient (SQ) value of larynx waveform, and spectrogram of each tone are the main objects of our consideration. F0 and SQ values are analyzed by utilizing EGG program (ver.3.01), and spectrogram by CSL program (ver.5.05). We observe F0 values to distinguish between high and non-high tones. SQ values are useful to judge if one is tense or non-tense. And we check the appearance of vertical striations in spectrogram to make sure the creakiness of syllables in question. The second procedure concerning SQ values has not been common so far, but the result coincides with our auditory impression very well, that is, the higher the SQ value is, the less tense the syllable is.

III. DATA ACQUISITION

A set of syllables with all 6 tones were pronounced by 5 informants of Vietnamese. All of them are from Hanoi. The set consists of syllables of 6 variations of initials, /t/, /th/, /d/, /n/, /s/, /z/, with 1 vowel, /a/, and 6 variations of tones. Each syllable was repeated three times for each recording. Acoustic data for F0 measurement and spectrogram reading were captured by way of microphone, and laryngographic data were captured through electrode collar set on the right and left outsides of larynx.

IV. FO MEASUREMENT

As stated above, the pitch contour of each tone has been examined exhaustively by the former authors, chiefly for the purpose of describing the pitch value and contour pattern of each tone. It is quite obvious that some main contour patters, such as [rising] or [falling], are actually audible, and our F0 measurements also support the existence of them. But our concern here is the 2-way distinction between high and non-high. Therefore, we measure the mean FO value of each tone for each informant, regardless of the initial consonant variation. The result is shown in Table 1. The mean value of one informant is different from that of the others, and the boundary between high and non-high of each informant is varied. Thus, a specific boundary for any individual informant divides 6 tones into 2 categories, giving a conclusion that LEVEL, RISING, and BROKEN tones are high tones and

	LEVEL	FALLING	RISING	CURVE	BROKEN	DROP
Informant1	144.9	115.5	169.1	113.9	176.4	137.1
Informant2	178.2	136.2	182.4	141.9	204.3	165.8
Informant3	132.0	109.2	144.3	104.8	146.1	113.9
Informant4	154.7	124.0	161.8	120.6	172.6	134.6
Informant5	153.9	119.6	152.0	117.5	152.9	135.3

Table 1. Mean F0 values. (Hz)

FALLING, CURVE, and DROP tones are non-high tones. The boundary between high and non-high is around 120Hz for Inf. 3, 140Hz for Inf. 1, 4, 5, and 170Hz for Inf. 2, respectively.

V. SPEED QUOTIENT MEASUREMENT

Speed quotient is the ratio of signal rise time to fall time of larynx waveform. The relationship between SQ and F0 values in certain phonation types was examined by Esling, Dickson, and Snell (1992). According to their result, the breathy voice and modal voice are distinguished clearly by their SQ values, regardless of their F0 values: the SQ values of modal voice are always lower than those of breathy voice. In a tonal system of dynamic pitch contour pattern like Vietnamese, this point is very helpful for the judgement of phonation type concerning each tone. Based on this point of view, we make a hypothesis: as the SQ value decreases, the tenseness increases, which coincides well with our auditory impressions. The problem here is the treatment of creaky voice, which will be discussed in later section. As a technique of SQ measurements, because voiced stop /d/ and fricative /z/ initials tend to lower the SQ values, those portions must be excluded from the subject.

At first, we present the mean SQ values of all the syllables with different tones spoken by each informant in Table 2. Mean values here are measured for the whole portion of syllables, that is, from the beginning point of the F0 contour to the end of it. Only a glance of this table would not give us any good idea about the tenseness of each tone, except for RISING tones which always have the lower SQ values than the others. The portion within one syllable which characterizes the phonatory property of that syllable is quite limited in duration, therefore it is meaningful to measure only that portion of syllable to make the distinction more obvious. In fact, the pitch rising part of RISING tone syllable and falling part of CURVE tone syllable show extra-low SQ values, and we have measured those portions. The results of measurement are listed in Table 3. Now it is proved that RISING and CURVE tones show the more tenseness than the other tones. The problems arising here are the slight high SQ value of CURVE tone of informant 2, which makes the mean values of LEVEL tones and CURVE tones almost the same, and the extra lowness of BROKEN SQ values of informant 3. As for the second problem, BROKEN and DROP tones are characterized by the creaky voice audible in those syllables, and the SQ measurements only for the creaky portion results in the high SQ value (the mean is 345.6%), which will solve this problem immediately.

Based on the above hypothesis, we may set the boundary between tense and non-tense for each informant as follows: around 230% for Inf.1, 250% for Inf.2, 240% for Inf.3, 270% for Inf.4, and 350% for inf.5.

VI. SPECTROGRAM READING

The third procedure of our analysis is the spectrogram reading to check the characteristic of creaky voice. Ladefoged, Maddieson, and Jackson (1988) pointed out that, for the spectrogram of Jalapa Mazatec, an Otomanguean language spoken in Mexico, the vertical striations (i.e., glottal pulses) occur at irregularly spaced intervals in the middle of vowels that are marked by creaky voice. As shown in Figure 1. and Figure 2., such kind of vertical striations actually appear in the middle of BROKEN tone syllables, and in the end of DROP tone syllables, which is a strong evidence for the creakiness of BROKEN and DROP tones.

VII. DISCUSSION

As we have mentioned in the first section, CURVE tone has two variants of pitch contour patterns: one is a falling followed by a slight rising in the end of syllables, and the other is just a falling and no rising. This kind of variation occurs context-freely, but the occurrence depends on the speakers. In fact, 3 of our 5 informants pronounce the CURVE tone with the second variant, in which case the contour pattern of FALLING and CURVE tones are almost the same. (See Figure 3.) In those and other cases, the mean SQ values of FALLING tones are always higher than those of CURVE tones, and the difference between them is

	LEVEL	FALLING	RISING	CURVE	BROKEN	DROP
Informant1	280.8	244.5	230.0	234.3	241.5	315.7
Informant2	254.9	294.6	237.5	275.3	287.2	309.1
Informant3	243.9	263.1	223.9	256.6	229.7	316.3
Informant4	323.1	325.4	295.1	311.1	300.7	377.4
Informant5	347.4	364.7	351.6	354.6	352.7	463.4

Table 2. Mean SQ values. (%)

	LEVEL	FALLING	RISING*	CURVE*	BROKEN	DROP
Informant1	280.8	244.5	189.5	214.6	241.5	315.7
Informant2	254.9	294.6	216.4	256.7	287.2	309.1
Informant3	243.9	263.1	174.9	233.5	229.7	316.3
Informant4	323.1	325.4	251.0	259.5	300.7	377.4
Informant5	347.4	364.7	299.5	318.1	352.7	463.4

Table 3. Mean SQ values. (%)
(*) Measurements for the characteristic portion only.

around 10% for the whole syllable measurements and 30~60% for the partial measurements. Therefore it is quite sure that what guarantees the distinction between these two tones is the difference of the degree of tenseness.

Figure 1. contains the spectrograms of RISING and BROKEN tones. In the middle of BROKEN tone, we can observe the obvious characteristics of the creaky voice, while RISING tone does not show the same kind of feature. According to Vu Thanh Phuong (1982), the pitch pattern of RISING tone is [35] and BROKEN [325] (with glottalizations). Our data also show the similar pitch contour patterns. and there is an abrupt drop of pitch contour in the middle of BROKEN tone. Actually that part coincides with the creaky voice part. Therefore, the abrupt dropping of F0 will be interpreted as the physiological effect of creaky voice. Consequently, if we follow the 5-level notation of Vu Thanh Phuong (1982), we can give [35] pitch pattern for both tones and analyze that what distinguishes them is the appearance of creaky voice in the middle of BROKEN tone.

Both of FALLING and DROP tones can be labeled as [+falling] tone. But their difference in phonetic property is quite obvious and diverse. First, DROP tone is characterized by the existence of creaky voice in the end of syllables, while FALLING tone is not. (See Figure 2.) Secondly, the relative F0 value of DROP tone is higher than that of FALLING tone. And finally, the duration of FALLING tone is much longer than that of DROP tone. Here we regard the first point as basic.

The relationship between tenseness (SQ value) and creakiness remains problematic. According to our data, the syllables with creaky voice show high SQ values, which means that creaky voice is categorized as non-tense. This result agrees with the consideration by Halle and Stevens (1971), who categorized creaky voice vowels as [-stiff, +slack]. But their approach has been criticized by a number of phoneticians, and we would like to leave this problem open for future considerations.

VIII. CONCLUSION

Now we have enough basis for attributing the prosodic and laryngeal features to each tone, which covers the overall distinction within the system. As for high and non-high F0, we propose a feature [+/-high]. Secondly, we propose [+tense] for low SQ value and [-tense] for high SQ value. And finally, [+/- glottal constriction] is for creaky and non-creaky voice. The consequent feature attribution is as follows:

	[high]	[tense]	[glottal constriction]
LEVEL	+	-	-
FALLING	-	-	-
RISING	+	+	-
CURVE	-	+	-
BROKEN	+	(-)	+
DROP	-	(-)	+

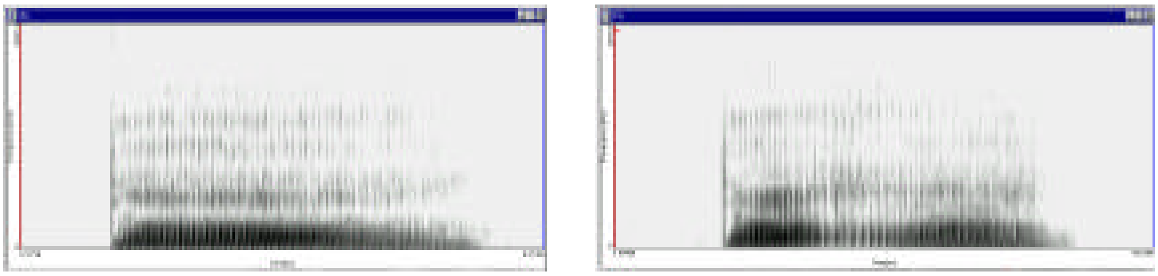

Figure 1. Spectrograms of RISING (left) and BROKEN (right) tones (/ta/ of Informant1)

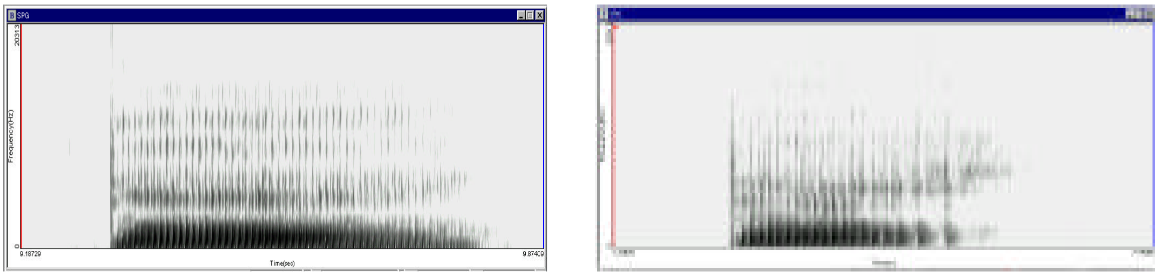

Figure 2. Spectrograms of FALLING (left) and DROP (right) tones (/ta/ of Informant1)

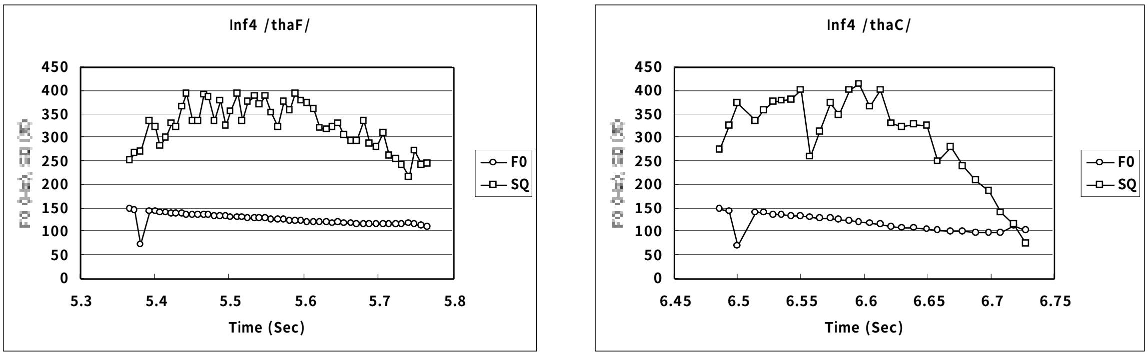

Figure 3. FO and SQ contour of FALLING (left) and CURVE (right) tones (/tha/ of Informant4)

In this way, we can distinguish each member in the whole system using one prosodic feature [high] and tow laryngeal features [tense] and [glottal constriction].

ACKNOWLEDGEMENT

We are very grateful to our Vietnamese informants for all their kind assistance. This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research on Priority Areas, No.12040106, 2000.

REFERENCE

[1] M. S. Han. *Studies in the Phonology of Asian Languages VIII: Vietnamese tones.* Los Angeles: University of Southern California. 1969.

[2] Vu Thanh Phuong. "Phonetic properties of Vietnamese tones across dialects." In D. Bradley (ed.), *Papers in South-east Asian linguistics. Tonation.* No.8, pp. 55-76. Canberra: Pacific Linguistics A62. 1982.

[3] Nguyen Van Loi and J. A. Edmondson. "Tones and voice quality in modern northern Vietnamese: Instrumental case studies." *Mon-khmer Studies.* vol.28, pp.1-18. 1998.

[4] J. H. Esling, B. C. Dickson, and R. C. Snell. "Analysis of Phonation Type Using Laryngographic Techniques." *ICSLP* '92-2, Banff. 1992.

[5] P. Ladefoged, I. Maddieson, and M. Jackson. "Investigating Phonation Types in Different Languages." In O. Fujimura (ed.) *Vocal Physiology: Voice Production, Mechanisms and Functions.* pp.297-317. New York: Raven Press, Ltd. 1988.

[6] M. Halle, and K. N. Stevens. "A note on laryngeal features," *MIT Quarterly Progress Report* vol.101, pp.198-213. 1971.

[7] M. Dantsuji. "An Acoustic Study on Glottalized Vowels in the Yi (Lolo) Language - A Preliminary Report -." *Studia Phonologica*. vol.16, pp.1-25. 1982.