TONATION IN THREE CHINESE WU DIALECTS

Phil Rose

philjohn.rose@gmail.com

ABSTRACT

Auditory, acoustic and limited physiological data are presented, together with audio, to describe different types of interaction between citation tone pitch and breathy, whispery, growly and creaky phonation type in three varieties of Wu 吳 Chinese: Lóngyóu 龍游, Yǒngjiāng 甬江 and Jìnyún 縉雲. Interaction in both lexical and verb-object tone sandhi is briefly described for the last two varieties.

Keywords: Tone, phonation type, Wu dialects.

1. INTRODUCTION

Acoustic energy generated by modulation of airflow at the larynx is recruited for many functions associated with the implementation of communicative intent [7]. Its periodic time-domain modulation underwrites the suprasegmental trinity of tone, intonation and stress; and its mode – for laryngeal structures can not only vibrate but vibrate in different ways – underlies different phonation types. Since they share a laryngeal generator, tonal pitch and phonation type are likely to interact in tone languages where extrinsic phonation type is an integral part of the realisation of the linguistic category of tone – so-called *tonatory* languages [2, 12, 26]. This paper describes different types of extrinsic phonation and their interaction with tonal pitch in three Chinese Wu dialects spoken in Zhèjiāng province. The description is supplemented with audio from field-work recordings made over the last 40 years.

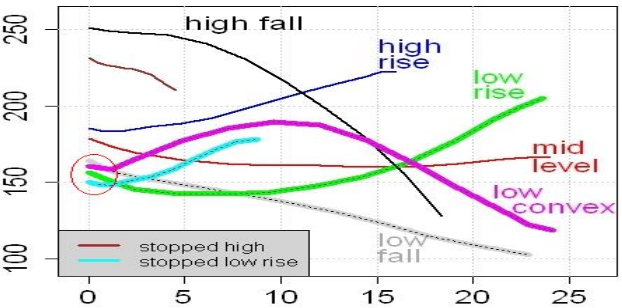
2. TONATION AND NATURAL CLASSES

Table 1: The Wu tonation template: Eight citation tones from a speaker of Lóngyóu. Click blue pitch contour descriptors to listen.

	-trunc			*+trunc*
+modal phonation	**mid level (-rise)**	**high rise**	**high fall**	**stopped high**
-modal phonation	**low fall**	**low rise**	**low convex**	**stopped low rise**

The primary phonological function of tonation in Wu is to characterise natural classes of importance for tone sandhi and phonotactics. The dialect of Lóngyóu, belonging to the Chǔ subgroup of Wu [4], provides a straightforward example. Table 1 shows its eight tones, a typical number for Wu. The tones of a 33 y.o. male speaker, recorded in 1988, can be heard by clicking on the table's tonal pitch descriptors (*high rise* etc.). Figure 1 shows his mean tonal acoustics – F0 plotted as function of absolute duration (number of samples per tone = 5, including those exemplified in table 1).

Figure 1: Mean tonal acoustics of a male Lóngyóu speaker. X-axis = duration (csec.), y-axis = F0 (Hz). Non-modal tones are plotted with thicker lines. Red ellipse denotes their shared low onset.



The data in table 1 and figure 1 show that the Lóngyóu tones do not simply constitute an unrelated octuplet, but organise themselves according to two orthogonal dimensions, intentionally reflected in the layout of table 1. The first, horizontal, dimension, which can be called *phonatory register*, divides the eight tones into two natural classes of four on the basis of phonation type and pitch/F0. Thus tones in the bottom row have *non-modal phonation* (NMP) characterised mostly by a noisy d.c. component. This is the most commonly mentioned phonation type for Wu, especially Shanghai, variously described as breathy, murmured, lax and whispery [18, 23]. The magnitude of the phonatory effect correlates with the pitch contour: the low falling tone sounds more breathy; the other tones, with rising pitch components, more whispery. There are also three tokens (low convex [d̥ɪa], stopped low rise [z̥ɐʔ b̥ɐʔ]) audibly tending towards harsh. Given current knowledge relating laryngeal gesture to phonation type [8, 10] it is likely that the Lóngyóu NMP involves a relatively open phonating glottis [19], with laryngeal sphinctering in the whispery rising pitch tones (this would also go to explaining the odd harsh tokens). The breathiness of the low falling tone would be expected if it had been produced with the more patent, unconstricted airway associated with lowered larynx [13]. It would be interesting to see how easily the NMP phonation can be quantified with interharmonic noise or conventional spectral tilt measures, as in [3]. As far as F0 is concerned, figure 1 shows the F0 values of the NMP tones lying largely in the lower half of the speaker's F0 range, but still overlapping with the F0 of the modal tones. The F0 feature that most clearly defines the NMP tones is their low onset, itself probably a function of the non-modal phonation type [22]. In contrast to the NMP tones, tones in the upper row of table 1 have modal voice, non-low onset F0, and lie largely in the upper F0 range.

The two natural classes thus defined constitute probably the most basic phonological dichotomy in Wu, governing tone sandhi rules, the distribution of syllable Onsets, and realisation of vowels. For example, in modal register, fortis coincident and lag VOT stops contrast; in non-modal register, voiceless lenis and voiced stops alternate morpho-phonemically, conditioned by word position.

The second, vertical, dimension of table 1, which can be called *truncation* and is also phonatory in nature, partitions the tones into two natural classes on the basis of phonation offset and Rhyme duration. One class, with the two stopped tones, has abrupt phonatory offset ([ʔ]) and short Rhymes; the other class, with the remaining tones, has gradual offset and long Rhymes. This truncation dimension defines natural classes of importance for tone sandhi and Rhyme. For example, [+truncated] tones often behave differently in tone sandhi from [-truncated] tones and also have a different, smaller set of Rhymes [21].

Within this matrix defined by the two primary phonatory dimensions of Register and Truncation there is a further subdivision according to pitch contour, whereby, for example, the high rising tone contrasts with the high falling tone. The pitch contours of three of the NMP tones are the same as those in the corresponding upper register tones but with a depressed onset [22], and this pairing-by-depression of pitch contours is often found. The contour pairing is obviously not totally systematic: the pitch of the modally phonated mid level(-rising) tone cannot be related by depression to the pitch of the low falling tone.

3. INDEXICAL TONATION: YǑNGJIĀNG GROWL

The non-modal phonation is found exaggerated as a socio-phonetic marker in varieties within the Yǒngjiāng sub-subgroup of Wu. These varieties were often locally stereotyped as sounding *stone-bone hard* 石骨硬, a well-known Chinese saying from the Wu area declaring it *better to argue with someone from Suzhou than speak with someone from Ningbo* (Ningbo is Yǒngjiāng's main city and Suzhou a Wu-speaking city outside the Yǒngjiāng area). In his 1928 survey of the Wu dialects [5] Chao Yuenren observed that the low onset of two Ningbo tones contributed to *an especially indistinct and heavy Ningbo resonance* 特別濁重的"寧波腔".

Table 2: Yǒngjiāng Tonation. S = speaker. Click blue pitch descriptor to listen.

	modal	non-modal
S1	**high fall**	**low convex**
	mid dipping	**low rise**
	stopped high	**stopped low rise**
S2	**high fall**	**low convex**
	mid dipping	**low rise**
	stopped high	**stopped low rise**
S3	**high fall**	**low rise(-fall)**

From inferences based on more recent investigation of Yǒngjiāng speakers [20], it appears these stereotypes relate to a harsh phonation type often accompanied with epilaryngeal trilling (when it is termed *growl*), which shows many features of a socio-linguistic marker indexing the speaker as someone from the Yǒngjiāng area. Audio from three Yǒngjiāng males is presented in Table 2. The first speaker was a 69 y.o. from Zhènhǎi town 鎮海 recorded in 1977. He has six tones on isolated monosyllables, exemplified with recordings of carefully read out single Chinese characters. As with Lóngyóu, there is a natural class of NMP tones but in this case they are mostly growled. Here is an additional growled example, with clear trilling, in a disyllabic word: [d̥ɛ̃ˤⁿ tsz̩ 11.31] 彈子 *marbles* (click to hear); it makes a nice minimal pair with the modally voiced [tɛ̃ⁿ tsz̩ 33.42] 單子 *list*. (Growl is transcribed [ˤ], after [9, 15], to imply epilaryngeal, as opposed to specifically aryepiglottal vibration.) The same truncated natural class as Lóngyóu is also present (the "stopped" tones), truncation obviously combining non-problematically with growl.

The second speaker exemplified in table 2 is from Hòushī 后施 village in the Zhènhǎi countryside. He has the same six citation tone system as the first. He was also recorded in 1977, when he was 30 (28 years later he is still growling), but his isolation tone examples are edited from a much less formal elicitation session where he was given a list of characters and asked to simply chat about them. Thus his utterances in table 2 are mostly of the type *this (character) is pronounced x*, the morphemes with his six isolation tones occurring either at the end of each utterance, or penultimately before the word for *character* (realised as [z̩]). His growl is audibly more tense than the first speaker.

The last speaker, recorded in 2007, is a 30 y.o. male from Pŭtuóshān 普陀山. Only two of his non-truncated tones are exemplified: one modal, with high falling pitch, and one non-modal, with free-varying low rise and low convex pitch. His non-modal phonation sounds more harsh than the other two, with no clearly trilled examples.

Figure 2: Laryngoscopic view of author's epilarynx. left: growling; right: nil phonation.
ae = *aryepiglottic fold*, c = *cuneiform tubercle*,
k = *cornicluate cartilage*, v = *vocal fold*,
f = *ventricular fold*, et = *epiglottic tubercle*.

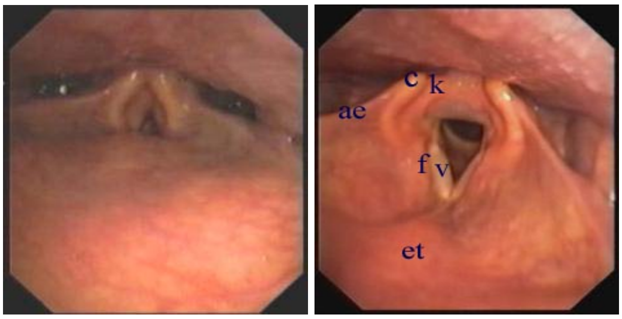
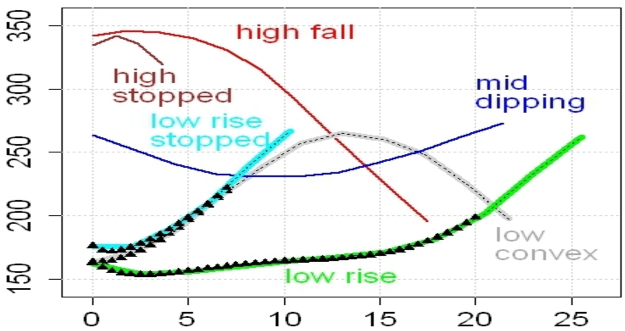


Figure 2 shows a still of the laryngeal configuration in a growled [ɛ] (click to hear) taken from a fibrescopic examination of the author (who has spoken conservative Zhènhǎi dialect non-natively, but quite fluently, for over 35 years). A labelled nil phonation configuration is also provided for orientation. Interpreted within the *Valves Model* [8, 16] an engagement of the laryngeal sphincter mechanism of valve III can be seen, involving anteroposterior approximation of the epiglottis and aryepiglottic folds. The degree of epilaryngeal constriction, together with the arytenoidal adduction for phonation, allows but a glimpse of a small portion of the right cord with a possibly slightly open glottis. Otherwise the glottal and ventricular activity of valves I and II is totally obscured. Oscillation is clear in the laryngoscopic video from which the still was taken, but not between the borders of the aryepiglottic folds and epiglottis as in the epiglottal trills demonstrated in [11, 17] (it is possible that such trilling is inhibited by my relatively massive cuneiforms). Rather, source modulation appears to come predominantly from the arytenoidal cartilage complex, especially the cuneiform tubercles, with some sympathetic vibration around the inner mucosa of the aryepiglottic folds [14]. This is presumably an example of the between-speaker variation in execution of epilaryngeal vibration noted in [15].

Figure 3 reproduces from [20] the mean tonal acoustics of speaker 2 in table 2. Not surprisingly, automatic F0 extraction does not work well with shimmery growl, and F0 has to be measured direct from the wave form. As with Lóngyóu, the NMP tones can be seen to have a common low onset, and overall lower F0 than the modal. Growl obtains on low F0 values and dies out when they rise out of the lower 1/3 of his F0 range. That it does not return with the falling F0 values of the low convex tone shows that it is associated with morpheme onset. The same pairing-by-depression of pitch contours can be seen as in Lóngyóu, except that the high fall / low convex pairing is a reflex of different historical tonal categories: *Ia/Ib in Zhènhǎi; *IIIa/IIIb in Lóngyóu (different historical categories realised by the same tones is a feature of Wu [25]).

Figure 3: Mean citation tonal acoustics for Zhènhǎi speaker 2. X-axis = duration (csec.), y-axis = F0 (Hz). Non-modal tones are plotted with thicker lines. Triangles indicate growled portions.



Tone and phonation type also interact as part of the complex tone sandhi for which Wu is famous [6]. Typical Wu tonal sensitivity to morpho-syntactic structure - most varieties, for example, use sandhi to distinguish between phonological word and syntactic phrase – can be illustrated in the realisation of the Zhènhǎi morpheme {rice飯} in speaker 2's conversational utterance (click to hear) *"This character is pronounced 'ɣ̥ɛˤ'; (as for example in) breakfast, lunch, dinner"*. The {rice} morpheme's first occurrence, as part of a verb-object *phrase*, is a nominal object governed by the verb /doʔ/ *to be read as*, where, as an independent word, it preserves its low rising pitch and growl: [kǐ z̩ dǒ **ɣ̥ɛˤ**] (click to hear). In the *compound words* that follow – *breakfast* [tʰiɲɪã**vɛ** 44.44.44] 天亮飯, *lunch* [tɕɪukɔʊ**vɛ** 44.44.44] 晝過飯 and *dinner* [jaˤ**vɛ** 11.34] 夜飯 – the {rice} morpheme's pitch has undergone regular lexical tone sandhi changes to high level and high rising pitches (transcribed *[44]*, *[34]*), and the growl has been lost, restoring modal phonation [21]. This is typical of so-called *left-dominant* Wu tone sandhi systems, where tone and phonation type tend to be preserved on the left morpheme of a word, with changes to the right [1, 28]. Note that growl is preserved, as a regular part of left-dominant lexical tone sandhi, on the word-initial syllable of [**ja**ˤvɛ] *dinner*, as also in [**d̥ɛ**ˤn tsz̩] *marbles* above.

Growl's currency at the time of writing is not clear. Speaker 3 in table 2 volubly attests to the presence of harsh Yǒngjiāng speakers in 2007, but a thesis on phonation types in Ningbo written at the same time [27] was only able to find speakers with breathy or whispery NMP.

4.CONTRASTING NON-MODAL PHONATION: JÌNYÚN

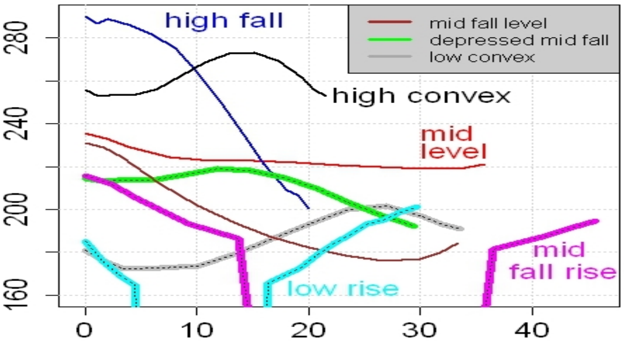
Table 3: Eight citation tones from a speaker of Jìnyún. Click blue pitch descriptors to listen.

+modal phonation	**mid level**	**high fall**	**high convex**	**mid fall level**
-modal phonation	**low convex**	**depressed mid fall**	**mid fall rise**	**low rise**

An additional tonational complexity may be found in the eight tones of the Wǔyúnzhèn 五雲鎮 variety spoken in Jìnyún county. Like Lóngyóu and Zhènhǎi, the same phonatory-register partitioning of tones is found; but unlike them the NMP differs starkly depending on the tone. Two tones – with low convex and depressed mid falling pitch – have breathy phonation; two tones – with mid falling-rising and low rising pitch – are creaky. All eight Jìnyún tones are exemplified in table 3. Recordings made in 1997 of a 20 y.o. female can be heard by clicking on the pitch descriptor. Figure 4 shows her mean tonal acoustics. Jìnyún also differs from Zhènhǎi and Lóngyóu in lacking a separate tonational truncation dimension: a glottal-stop coda can indeed be heard on some of the (mid fall level / low rise) examples of historical category IV, but it is not obligatory; and neither are the mid fall level and low rise tones shorter than the other tones. This variety also uses more complex tonal pitch contours, even though its tonology is already complicated by phonation type differences. There is also not a convincing separation of modal and non-modal tones by F0 onset and pairing-by-truncation; and a case can be made for contrastive phonation type: modal vs. creaky between mid fall level and mid fall rise tones; and breathy vs. creaky between low convex and low rise tones.

Like Zhènhǎi, Jìnyún tone sandhi is sensitive to morphosyntactic structure. Unlike Zhènhǎi, and in common with most varieties in the southern half of Zhejiang, it has a *right-dominant* tone sandhi system: tone and phonation type are preserved on word-final morphemes, but can change on morphemes towards the beginning of a phonological word [1, 28]. For example, both free morphemes in the word [d̥ə̤ʊ lṵ] 大路 *main road* (click to hear) have the mid fall rise creaky tone. The tone and phonation type are preserved on the word-final syllable (allowing for pitch changes due to influence from the first syllable), but change by lexical sandhi to the depressed mid falling breathy tone on the first. In a *phrase*, however, like *sell cattle* [ma̰ ṇi̤ʊ̃] 賣牛, (click to hear) the creak of the fall rise tone is preserved on the first syllable (*sell*) although it loses its rising pitch. (The second syllable (*cattle*) carries a low rising allotone of the low convex breathy tone). A nice contrast then obtains with *buy cattle* 買牛, the first syllable of which (*buy*) carries the mid falling breathy tone: [ma̤ ṇi̤ʊ̃] (click to hear).

Figure 4: Mean citation tonal acoustics of female Jìnyún speaker in table 3. X-axis = duration (csec.), y-axis = F0 (Hz). Non-modal tones are plotted with thicker lines.



5. SUMMARY

This paper has documented three different types of interaction between tonal pitch and phonation type found in Wu. Although not exhaustive – another type being the non-modal phonation extending to the upper register tones in Táizhou Wu varieties [29], and in Lóngquán [24] – the examples described contribute to the wider typology of interaction between tone and phonation type.

6. ACKNOWLEDGEMENTS

I am exceedingly grateful to the many people who made this paper possible. Profs. Wang Wei and Chen Jie and their staff at the *Shanghai Ear Nose and Throat Hospital* scoped my growl, and Dr. Zhang Xixi provided timely assistance with the audio recording. Prof. Zhu Xiaonong arranged the laryngoscopy and supplied the Jìnyún recordings. The Lóngyóu recording was made by Prof. William Ballard. Prof. John Esling, and particularly Dr. Scott Moisik, provided extensive, instructive and invaluable comments on growl and related phonatory phenomena.

7. REFERENCES

[1] Ballard, W. 1984. Wu, Min and a Little Hakka - Tone Sandhi: Right and Left. *Cahiers de Linguistique Asie Orientale* 13, 3–34.

[2] Bradley, D. (ed) 1982. *Tonation*. Papers in South East Asian Linguistics No. 8. *Pacific Linguistics.*

[3] Cao, J., Maddieson, I. 1992. An exploration of phonation types in Wu dialects of Chinese. *JP* 20, 77–92.

[4] Cao, Z. 曹志耘 2002. 南部吴语语音研究 *Nanbu Wuyu Yuyin Yanjiu* [*Phonetic Studies on Southern Wu dialects*], Peking: Commercial Press.

[5] Chao, Y. 1928. 現代吳語的研究 *Studies in the Modern Wu Dialects.* Tsinghua College Research Institute monograph no. 4. Peking: Tsinghua College Research Institute.

[6] Chao, Y. 1967. Contrastive Aspects of the Wu Dialects. *Language* 43/1, 92–101.

[7] Christer, G., Ní Chaside, A. 2013. Voice Source Variation and its Communicative Functions. In: Hardcastle, W., Laver, J., Gibbon, F.E. (eds), *The Handbook of Phonetic Sciences*. 2nd Ed. Chichester: Wiley-Blackwell, 378–423.

[8] Edmondson, J. A., Esling, J. H. 2006. The valves of the throat and their functioning in tone, vocal register and stress: laryngoscopic case studies. *Phonology* 23, 157–191.

[9] Esling, J.H. 1996. Pharyngeal consonants and the aryeppiglottic sphincter. *JIPA* 26/2, 65–88.

[10] Esling, J.H., Harris, J.G. 2005. States of the Glottis: An Articulatory Phonetic Model Based on Laryngoscopic Observations. In: Hardcastle, W.J., Beck, J.M. (eds), *A figure of speech: A Festschrift for John Laver*. Mahwah, New Jersey: Erlbaum, 347–383.

[11] Hassan, Z. M., Esling, J.H., Moisik, S.R., Crevier-Buchman, L. 2011. Aryeppiglottic trilled variants of /ʕ ħ/ in Iraqi Arabic. *Proc. 17th ICPhS*, 831–834.

[12] Henderson, E.J.A. 1967. Grammar and Tone in S.E. Asian Languages. *Wissenschaftliche Zeitschrift der Karl-Marx-Universität Leipzig, Gesell- u. sprachwissenschaftliche Reihe* 1/2, 171–178.

[13] Laver, J. 1980. *The Phonetic Description of Voice Quality*. Cambridge: CUP.

[14] Moisik, S.R., 2015. Pers.comm.

[15] Moisik, S.R. 2013. *The Epilarynx in Speech*. Unpublished Ph.D. thesis, University of Victoria.

[16] Moisik, S.R., Esling, J. H. 2011. The 'Whole Larynx' approach to Laryngeal Features. *Proc. 15th ICPhS*, 1406–1409.

[17] Moisik, S.R., Esling, J. H., Crevier-Buchman, L. 2010. A high-speed laryngoscopic investigation of aryepiglottic trilling. *JASA* 127/3, 1548–1558.

[18] Norman, J. 1988. *Chinese*. Cambridge: CUP.

[19] Ren N. 1992. *Phonation types and stop consonant distinctions: Shanghai Chinese*. Unpublished Ph.D. thesis, University of Connecticut.

[20] Rose, P. 1989. Phonetics and Phonology of the Yang Tone Phonation types in Zhenhai. *Cahiers de Linguistique Asie Orientale* 18/2, 229–245.

[21] Rose, P. 1990. Acoustics and Phonology of Complex Tone Sandhi. *Phonetica* 47, 1–35.

[22] Rose, P. 2002. Independent Depressor and Register Effects in Wu dialect tonology: evidence from Wenzhou Tone Sandhi. *Journal of Chinese Linguistics* 30/1, 39–81.

[23] Sherard, M. 1972. *Shanghai Phonology*. Cornell: Cornell University Press.

[24] Steed, W. 2006. Phonation Type and Tone Sandhi as Evidence of Stress in Longquan Wu. In: Warren, P., Watson, C. (eds), *Proc. 11th Australasian Intl. Conf. on Speech Science and Technology*, 82–87.

[25] Steed, W., Rose, P. 2009. Same Tone, Different Category: Linguistic-Tonetic Variation in the Areal Tone Acoustics of Chuqu Wu. *Proc. Interspeech*, 2295–2298.

[26] Thurgood, G. 2002. Vietnamese and tonogenesis: Revising the model and the analysis, *Diachronica* 19, 333–363.

[27] Whitehead, D.F. 2007. *Phonation types and speaker variation in Ningbo Chinese*. Unpublished M.Phil. Thesis, Hong Kong University of Science & Technology.

[28] Zhang, J. 2007. A directional asymmetry in Chinese tone sandhi systems. *Journal of East Asian Linguistics* 16, 259–302.

[29] Zhu, X. 2006. Creaky voice and the dialectal boundary between Taizhou and Wuzhou Wu. *Journal of Chinese Linguistics* 34/1, 121-133.