WENZHOU TONAL ACOUSTICS - DEPRESSOR AND REGISTER EFFECTS IN CHINESE TONOLOGY

Phil Rose

Department of Linguistics (Arts), Australian National University

ABSTRACT Mean F0 and duration data from one male speaker of Wenzhou dialect are presented for the eight citation tones, and tones in selected disyllabic lexical sandhi environments. Morphotonemic sandhi alternations are adduced to argue for the inclusion of a Depressor component in Wu tonological representation in addition to Register.

INTRODUCTION

This paper describes some tonological phenomena in the Southern Wu Oujiang dialect of Wenzhou. Wenzhou (Wz) has highly complex tone sandhi of the so-called right dominant type found in the southern Wu area. In this type of sandhi the tones on syllables towards the end of a word determine the sandhi shape and tend to show values similar to citation forms, and tone contrasts on syllables towards the beginning of the word tend to be neutralised (Ballard 1988:43).

The Wz citation tones are described first, followed by some tone sandhi changes that, it is argued, point to the existence of a Depressor as well as a Register component in Wz tonological representation. The auditory descriptions and acoustic analyses reported here are from my own continuing research into Wz tonetics and tone sandhi. They are based on my analysis of extensive recordings, very generously furnished by Lew Ballard, of one male native Wz speaker (ZGQ) made by Ballard in China in 1988. Two sets of F0 and duration measurements of the Wz data, made with two different extraction devices, were reported in Ballard (1989) and (1994). Ballard sampled F0 only at inflection points, which is inadequate for this kind of descriptive and analytical work: it does not give a sufficiently detailed picture of the F0 time course, and, without auditory transcription, does not permit differentiation between intrinsic and extrinsic effects on tonal F0. The data presented in this paper are the results of a complete remeasurement, therefore.

CITATION TONES

The auditory characteristics of the speaker's eight tones are described below with respect to pitch and length. The tones have been named according to their Middle Chinese tonal categories ("Ia", "IIIb" etc.) and also according to a simplified pitch description, ("mid-level" etc.). Their traditional Chinese phonological name is given in parentheses. **Tone Ia**: *mid-level* (yin ping). This tone has level pitch in the speaker's mid pitch range, and average length. Examples are [sei 33] 'west'; [ka 33] 'to close'. **Tone Ib**: *mid-falling* (yang ping). The pitch of this tone starts at about the same pitch as in tone Ia, has a short initial level component and then falls; length is average: [mai 331] 'plum'; [ni 331] 'year'. **Tone IIa**: *high-rising* (yin shang). This tone rises abruptly within the upper pitch range. Length is notably short, but the phonation offset is gradual, and not truncated by a glottal stop. Examples are [ɕou 34] 'arm', [tshə 34] 'grass'. **Tone IIb**: *low-rising* (yang shang). This tone has a short initial low level component, and then rises abruptly into the upper pitch range. It has average length, and phonation offset is as for tone IIa: [pai 114] 'blanket'; [tœy 114] 'stomach'. **Tone IIIa**: *high-falling* (yin qu). This tone falls abruptly from high in the upper pitch range into the lower pitch range. Length is somewhat shorter than average: [tɕi 51] 'paste', [thiɛ 51] 'jump'. **Tone IIIb** *low-level* (yang qu). This tone has a level pitch in the mid-low pitch range. Length is notably longer than average. Examples are [mi 222] 'face'; [tei 222] 'ground'. **Tone IVa** *mid-dipping* (yin ru). The pitch of this tone falls after a short initial level component in the mid pitch range, then rises slightly. Length is much longer than average: [pai 3312] 'north'; [tso 3312] 'make'. **Tone IVb** *low-dipping* (yang ru). This tone has similar prosodic characteristics to tone IVa, but its pitch onset lies slightly lower: [pa 2212] 'white'; [wu 2212] 'learn'. The eight Wenzhou tones thus comprise upper ("a") and lower ("b") values of the same four pitch shapes: *level* (tones Ia and IIIb); *rising* (tones IIa and IIb); *falling* (tones IIIa and Ib); and *dipping* (tones IVa and IVb). Length also appears to be an important auditory dimension for some tones. Thus the high-rising tone IIa is notable for its shortness, and the low-level and dipping tones IIIb, IVa and IVb sound long.

The upper and lower Wenzhou citation tones are distributed in typical Wu fashion with respect to several segmental and suprasegmental features of the syllable, in particular the manner of articulation of syllable-initial obstruents. Like other Wu dialects, Wenzhou has three morphophonemically separate sets of syllable-initial stops (voiceless aspirated; voiceless unaspirated; and voiced), and two sets of syllable-initial fricatives (voiceless and voiced). The first two sets of stops occur on syllables with the upper ("a") tones and are realised by voiceless aspirated and voicless unaspirated allophones. The third set of stops, which co-occurs with the lower ("b") tones, has different realisation depending on position in the word. Word-internally the realisation is modally voiced. Word-initially, the third series of stops is realised in this corpus predominantly by voiceless, coincident VOT articulations, but there is a small amount of free variation with modal, VOT lead tokens, although not in the citation monosyllables. The same applies *mutatis mutandis* to the two sets of fricatives: voiceless in upper, "a" tones; voiced word-internally, and voiceless in free variation with voiced word-initially in lower, "b" tones. In this corpus, however, some monosyllabic citation forms do have a fully voiced realisation. In this paper, the morphophonemically voiced series are transcribed as voiceless word-initially, and voiced word-internally.

Cao and Maddieson (1992) provide aerodynamic and acoustic evidence for a different, more breathy, mode of phonation in the Wz low-level tone IIIb than in the high falling tone IIIa. They claim that it is a phonological property of the initial stop. However, they crucially failed to examine syllables without an initial consonant, and so logically cannot claim a causal relationship between stop and phonation type. This kind of phonatory difference is audible in all type "b" tones irrespective of what, if anything, fills the onset slot, and, as was emphasised by Rose (1989:243), it is more likely a realisation of a syllable prosody which is phonologically related to the tone, not the consonant. Most Wu dialects, e.g. Shanghai, show this tonatory behaviour.

Acoustical values for the eight Wz citation tones (fundamental frequency (F0), duration, and radiated amplitude (Ar)) are given in table 1. Measurements are derived from between four to six tokens of differing segmental structure. In the mid level tone 1a, for example, it can be seen from table 1 that the mean F0 at the 40% sampling point, which occurred at (32.9 csec. x 40% =) 13.2 csec. into the Rhyme, was 142 Hz, with a standard deviation of 6 Hz. Figure 1 shows the F0 values of the Wz citation tones plotted as a fuction of absolute duration. The F0 shapes of the speaker's citation tones lie within a range of about 130 Hz, from 70 Hz to 200 Hz. There is a greater concentration of F0 shapes in the lower half of the F0 range. The F0 shape of the high falling tone IIIa falls fairly abruptly, traversing all but the last 10 Hz of the F0 range, and bisecting the F0 shapes of the level tones Ia and IIIb. The F0 shapes of these two tones lie in the middle of the F0 range separated by about 30 - 40 Hz. The F0 shapes of the two rising tones IIa and IIb both rise concentrically to the same value in the upper F0 range, but from different onsets separated by about 40 Hz. The F0 of the high rising tone IIa has a slightly lower onset than the mid-level tone Ia, and rises directly to its

Table 1. Means and standard deviations (x,sd) for F0, Duration (D) and Radiated Amplitude (Ar) in Wenzhou citation tones. Units are in Hz, csec., and dB respectively. SP = sampling point; DP = duration to F0 peak; D off = duration to phonation offset.

	mid level Ia	n = 6		**mid falling Ib**	n = 6
SP	**F0**	**Ar**	SP	**F0**	**Ar**
0%	148,4	26.9,3.5	0%	126,4	30.7,2.6
5%	-	31.7,3.5	5%	-	33.9,1.3
10%	147,4	33.0,2.8	10%	132,4	34.3,1.4
20%	144,4	32.4,2.3	20%	135,3	34.0,1.0
40%	142,6	31.1,2.2	40%	131,4	33.0,1.3
60%	141,6	30.8,1.9	60%	109,3	30.4,1.2
80%	138,6	29.1,2.3	80%	85,6	24.6,1.3
90%	-	24.1,2.5	90%	-	21.5,1.0
95%	-	,2.6	95%	-	21.1,0.8
100%	133,13	,2.8	100%	71,7	20.4,0.7
D	32.9,2.6		D	36.3,3.1	

Table 1. (con't).

high rising IIa		n = 5
SP	F0	Ar
0%	135,6	20.5,2.2
5%	-	26.3,1.3
10%	136,4	29.5,1.6
20%	137,5	31.7,1.5
40%	140,11	30.0,3.2
60%	149,12	30.6,2.2
70%	-	31.2,2.9
80%	161,13	30.7,3.0
90%	-	29.6,2.9
100% (F0 peak)	176,14	28.4,2.9
F0 off	174,19	23.4,3.0
DP	16.6,2.7	
D off	19.3,2.5	

low rising IIb		n = 5
SP	F0	Ar
0%	94,11	20.0,1.9
5%	-	22.7,3.8
10%	96,11	26.1,4.7
20%	96,10	25.8,4.1
40%	99,5	27.5,3.0
60%	112,4	30.8,1.7
70%	-	30.8,0.6
80%	139,3	31.1,1.5
90%	-	30.5,0.8
100% (F0 peak)	168,7	27.4,1.8
F0 off	142,17	21.4,2.1
Ar offset (at csec. 35.9)		15.5,2.9
DP	26.7,1.3	
D off	32.0,1.8	

high falling IIIa		n = 4
SP	F0	Ar
0%	201,10	21.6,2.2
5%	-	25.8,2.7
10%	202,11	27.9,2.5
20%	197,13	29.6,2.0
40%	170,16	29.0,2.2
60%	133,10	28.5,2.4
80%	103,6	26.4,2.3
90%	-	22.8,1.8
100%	81,2	19.8,0.4
D	21.3,3.1	

low-mid level IIIb		n = 5
SP	F0	Ar
0%	113,16	27.3,4.5
5%	-	30.8,3.9
10%	117,12	31.9,3.3
20%	121,10	32.6,3.3
40%	123,7	32.6,2.4
60%	123,4	32.3,2.7
80%	121,4	31.3,2.0
90%	-	29.2,2.0
95%	119,3	26.3,1.9
100%	109,12	22.7,1.8
D	44.1,4.3	

mid dipping IVa		n = 5
SP	F0	Ar
0%	137,6	24.4,2.0
5%	-	32.2,2.2
10%	132,8	33.8,1.7
20%	122,7	32.3,1.7
40%	105,5	30.4,2.1
60%	97,5	27.3,0.3
80%	99,6	26.3,1.0
90%	104,8	25.8,1.3
95%	101,8	23.9,1.9
100%	93,7	20.9,1.7
D	50.0,4.4	

low dipping IVb		n = 5
SP	F0	Ar
0%	109,8	27.7,5.5
5%	-	30.6,3.1
10%	110,8	30.4,1.9
20%	106,7	30.4,1.3
40%	96,7	28.5,1.6
60%	89,6	25.1,1.3
80%	91,7	24.9,1.9
90%	97,9	24.2,2.5
95%	100,8	23.2,2.7
100%	91,7	19.6,2.6
D	50.2,2.6	

target. The F0 shape of the low rising tone IIb has the lowest onset of all the tones, and shows a short quasi-level component before rising. The F0 of the low-falling tone Ib traverses the lower half of the F0 range. It onsets between the values for the two level tones Ia and IIIb, and falls, after a mild convexity, to the lowest offset point of all the tones. The F0 shapes of the dipping tones IVa and IVb converge slightly after onset and then exhibit almost parallel concave F0 shapes in the lower half of the F0 range. The falling component lasts longer, and shows a greater amount of fall, than the rising component. The F0 of the higher of the two tones shares the same onset with the high rising tone IIa. The low dipping tone shares the same onset as the low-level tone IIIb. Short falling offset perturbations are observable on tones with final rising pitch, and the low level tone IIIb. These are not audible as pitch contour. Rising onset perturbations are visible for up to 10 csec. on the "b" tones. These may be related to the more breathy phonation noted above to occur on the "b" tones. One conspicuous feature of the citation F0 configuration, which also agrees well with the auditory impression, is the large between-tone differences in duration. For example the high rising tone IIa has less than half the duration of the dipping tones IVa and IVb.

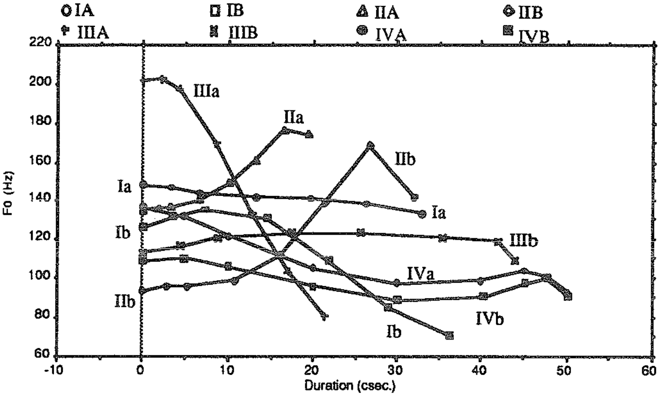


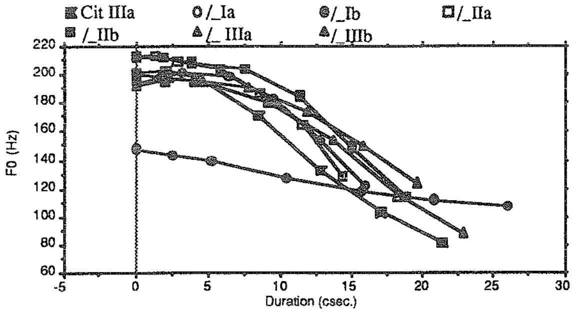
Figure 1. Mean F0 shapes and duration for Wenzhou citation tones.

Tonological representation in Asian languages using a combination of register and tone pioneered by Yip (1980) is now widely accepted. The Wz data presented above, with their upper and lower versions of the same four pitch/F0 shapes, seem a particularly clear instance of the realisation of this framework. Thus the high falling and mid falling tones IIIa and Ib would be represented as upper [+U(pper)] and lower [-U] register versions of the same HL tonal sequence; the high rising and low rising tones IIa and IIb would be upper and lower register versions of a LH sequence; the dipping tones IVa and IVb would be [+U, HLH] and [-U, HLH], and the mid and low-mid level tones Ia and IIIb would be [+U, L] and [-U, H] respectively. However, evidence from morphotonemic alternations in Wz tone sandhi points to another interpretation for the difference between the high and low versions. These alternations are described below.

TONE SANDHI

The relevant Wz sandhi combinations are found in disyllabic words with the following four input tones on the first syllable: high rising and low rising (IIa and IIb respectively), high falling (IIIa), and low level (IIIb), and all tones except the two dipping tones IVa and IVb on the second syllable. These will be dealt with in turn.

In combinations with the input high falling tone IIIa on the first syllable, this syllable has a high falling pitch, e.g. [pə 51] 'to report' + [tsei 34] 'paper' -> [52 23] 'newspaper'. There is one exception to the high falling realisation, which will not be considered further in this paper: when the second syllable input tone is low falling Ib. In this case the pitch on the first syllable still falls, but is audibly lower, and the second syllable is atonic, e.g. [tshe 51] 'vegetable' + [y 331] 'garden' -> [32 1] 'vegetable garden'. Figure 2 shows mean F0 shapes corresponding to these [52] and [32] pitches, together with the mean F0 shape for high falling citation IIIa. Seven shapes are shown - one for the [32] allotone before input tone Ib, five for the [52] allotones before input tones Ia, IIa, IIb, IIIa, and IIIb, and one for the IIIa citation tone. All second syllable b tones have a modally voiced syllable-initial consonant (see above). Figure 2 shows that the [52] allotones and the high falling citation tone F0 form one group clearly separate from the lower F0 shape of the [32] allotone. The former group displays the same level-falling F0 contour which onsets at about the same position in the speaker's F0 range, but which is truncated at different durations, yielding slightly different F0 offset values. The obvious tonological interpretation is one of a citation high falling tone target, with allotones conditioned by voicing on the following consonant, and atonicity on the following syllable. 'Citation target' is one of the categories of relation between citation tones and tones in sandhi, and refers to cases where a tone in sandhi 'can be identified as one of the citation tones, once allowance is made for intrincsic influence of various conditioning factors...' (Rose and Toda 1994:271). Tonologically, then, no change to the high falling input tone IIIa is involved in these combinations.



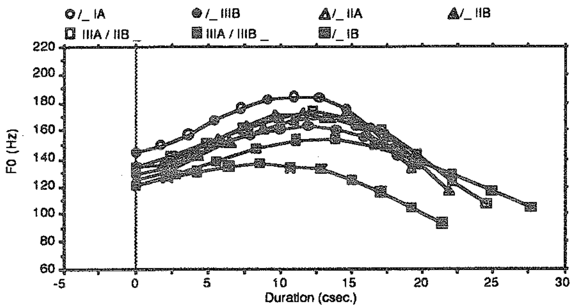


Figure 2. (above) Mean F0 for citation tone IIIa, and rhyme in IIIa + _ combinations.

Figure 3. (below) Pooled mean F0 shapes for IIb and IIIb tones before input tones Ia IIIb IIa IIb and Ib.

The same pitch shapes as just described are observed in combinations with the input high rising tone IIa on the first syllable, e.g. [kə 34] 'draft' + [tsei 34] 'paper' -> [52 23] 'manuscript'. In this case, however, a lower falling pitch on the first syllable does not occur when the second syllable carries an input Ib tone, although the second syllable still loses its tone, as with IIIa + combinations, e.g. [tɔ 34] 'party' + [y 331] 'member' -> [52 1] 'party member'.

The mean F0 shapes of these input IIa tones do not differ in any systematic way from that of the IIIa input tones in figure 2, except that there is no separate low F0 corresponding to the F0 of input IIIa before input tone Ib. The near identity in acoustic values between the input IIa and input IIIa tones confirms the auditory impression that the input high rising tone IIa changes to high falling tone IIIa before all except the dipping tones IVa and IVb. Structurally, this results in a neutralisation of the high rising and high falling tones IIa and IIIa in favour of the high falling tone IIIa before tones other than IVa and IVb. Given the tonological representation outlined above, this change is in the tonal component, with the register constant, thus: [+U, LH] -> HL / _ ~HLH.

When considered in conjunction with the change just described, the behaviour of the low register non-falling input tones IIb and IIIb on the first syllable indicates that the tonological representation with +/- U(pper register) is best augmented in Wz with a component indicating a depressor effect: "+/- D(epressed)". This manifests as a lowered onset to an otherwise unchanged F0/pitch contour. The changes are as follows. When the first syllable has low rising or low level input tones IIb and IIIb, there is a change to a low convex pitch tone. For example, [tœy 114] 'stomach' + [pei 331] 'skin' -> [tœy bei 232 1] 'stomach'; [vu 114] 'rain' + [sz 34] 'water' -> [232 23] 'rain'; [pɐŋ 222] 'ill' + [nɐŋ 331] 'person' -> [232 1] 'patient'; [tei 222] 'ground' + [xwɔ 33] 'direction' -> [232 33] 'place'. One possible exception to this is the combination IIIb + IIIa, which was given a slightly different, lower [221 1] pitch on the first syllable than the others. I suspect this is not a tonal change, but due to the intrinsic effect of the vowel, which in all three tokens was a low central-back [a]. Figure 3 shows mean F0 shapes for pooled IIb and IIIb tones before input tones Ia, IIIb, IIa, IIb and Ib. This figure shows a set of convex F0 shapes the offset and duration of which are conditioned in the same way as in the previous examples.

It has been shown above that input tones IIa and IIb both undergo changes in comparable environments (i.e. before a non-dipping tone), and constitute a well-defined natural class (LH, or rising). Under these circumstances it might be expected that the already demonstrated change of

LH -> HL for tone IIa would also apply to IIb. However, this is manifestly not the case if IIb is represented as [-U, LH], with a lower ([-U]) register, since the result of the change - HL - in conjunction with an unchanged low register would give a low *falling*, and not the observed low *convex* pitch. Rather, it appears that the LH -> HL change is taking place on input tone IIb in conjunction with a constant *depressor* effect, viz [+D, LH] -> HL / _ ~HLH. That is tone IIb, together with the other 'b' tones, is represented with +D, and its change from LH to HL results in [+D, HL], which is a plausible tonological interpretation of the convex pitch, or the depressed equivalent of the high falling pitched tone IIIa. Since there is no difference between the F0/pitch reflexes of input IIb and IIIb, it must be concluded that what is happening in the small part of Wz tone sandhi described in this paper is a neutralisation of the tonal and register components in tones IIa, IIb, IIIa and IIIb in favour of an upper register HL tone, with unchanged depression specification. The Wz data further show that the depressor representation must be in addition to the register specification +/-U. Otherwise the convex reflexes of IIb and IIIb ("[+D, HL]") would merge with mid falling citation tone Ib, which is also [+D, HL]. That these are in fact kept distinct reflects their different register specifcation: Ib is [-U, +D, HL]; the concave reflexes of IIb and IIIb are [+U, +D, HL].

The changes associated with tones IIb and IIIb are different from those which occur with their upper counterparts IIa and IIIa, in that the resulting low convex tone is not a citation target. However, the coherence of the sandhi behaviour lies in the fact that the convex tone and the high falling tone IIIa differ in the +/- depressor effect, which is revealed in the emergence of a new, non-citation tone. It is interesting to note that no mention is made of this convex pitch on input IIb and IIIb in Zhengzhang's (1964) description of Wz tone sandhi, where generally no differentiation between input b tones and input a tones on first syllables is shown (p108).

One final point to emerge from the Wz tone sandhi changes concerns the significance of the depressor effect for the by now widely accepted decomposed representation of contour tones. The depressor effect on a *contour* is conceptually coherent. A depressed onset to a fall, for example, results in a convex shape. It is not so easy to see how the effect can be conceived on a single, internally unstructured unit in a tonal sequence (for example, an H in a decomposed HL sequence). This therefore argues for the integrity of contour tones, and against their decomposition.

REFERENCES

Ballard, L. (1988) "Distribution and Historical Significance of Varying Tone Sandhi Types in South China", Computational Analysis of Asian & African Languages 30, 43-55.

Ballard, L. (1989) Visipitch Data on Tones and Sandhi in some Southern Zhejiang Wu Dialects. M.s.

Ballard, L. (1994) "Phonetic Data on Tone Sandhi in Two Oujiang Wu Dialects via CECIL, WORKS, and Harvard Graphics", in H. Kitamura et al. eds., *Current Issues in Sino-Tibetan Linguistics*, pp. 86-391, (Organising Committee, 26th Intl. Conf. on Sino-Tibetan Languages and Linguistics: Osaka).

Cao J, and Maddieson, I. (1992) "An exploration of phonation types in Wu dialects of Chinese", Journal of Phonetics 20, 77-92.

Rose, Phil. (1989) "Phonetics and phonology of the yang tone phonation types in Zhenhai",Cahiers de Linguistique Asie Orientale 18/2, 229-245.

Rose, Phil. & Toda, Takako. (1994) "A Typology of Tone Sandhi Rules in Northern Wu", in H. Kitamura et al. eds., *Current Issues in Sino-Tibetan Linguistics*, pp. 267-273, (Organising Committee, 26th Intl. Conf. on Sino-Tibetan Languages and Linguistics: Osaka).

Yip, M. (1980) *The Tonal Phonology of Chinese*, (Indiana University Lingusitics Club:Bloomington).

Zhengzhang S. (1964) Wenzhou Fangyande Liandu Bian Diao [Tone Sandhi in Wenzhou Dialect], Zhongguo Yuwen 2, 106-143.