ERRATUM - the name Wencheng is written incorrectly in Chinese.
The second character should be wen2 'accomplish', not wen2 'city'. PJR

The long and the short of Wencheng tones – acoustic and auditory description of tonologically challenging phenomena in an Oujiang Wu dialect of Chinese

Phil Rose

Linguistics, College of Arts & Social Sciences, The Australian National University

philip.rose@anu.edu.au

Abstract

Auditory and acoustic data are presented to document some tonologically challenging aspects of the seven tones of Wénchéng, a south-west Oujiang Wu dialect of Chinese. Tones are described both in isolation form and in selected lexical tone sandhi combinations. It is shown how the complexity of the tonal system results in clashes between tonological representation and phonetic reality.

Index Terms: Tone sandhi, Chinese, Wu dialects

1. Isolation tones

Language sometimes requires speakers to do strange things with their vocal cords in the name of tone realisation, and some of the strangest things are to be found in the Wu 吳 dialects of East Central China. These dialects, especially the coastal varieties in Zhejiang province, are notorious for tonal complexity. This is encountered primarily in their tone sandhi [4] where the relationship between sandhied and isolation tones is often opaque and involves phonetically indefinable natural classes [6]. Their tonal contrasts, too, are complex, involving a large number of tones differentiated with respect to complex contours, duration, and often phonation type [10, 3, 8]. Apart from its inherent interest there is also theoretical merit in such complexity. For there is likely to be less underdetermination by such complex observational data of tonological theories put forward to explain them. This paper documents a little of this complexity in the tones of the Oujiang 甌江 Wu dialect of Wénchéng 文城, pointing out several areas where the complexity presents problems for conventional models of tonological representation.

1.1. Wu, Oujiang, Wencheng

The Wu dialects are considered a well defined first-order sub-group of Sinitic typologically, on the basis of their systematic three-way VOT contrast for syllable-initial stops and affricates [1], and in this respect Wencheng is typically Wu. Subminimal triplets showing the three-way contrast intervocalically on an alveolar stop & alveo-palatal affricate are :

wa **t**ɛɪ 2.34 襪底	*sole of socks*	pɪ **tɕ**ɪŋ 2.33 別針	*safetypin*
ɲɥɪ **d**ɛ 2.33 月台	*platform*	tɕɪ **dʑ**aŋ 2.33 接近	*adjacent*
tʰɪ **tʰ**a 2.34 鐵塔	*Eiffel Tower*	ɲɪ tɕʰi 2.22 熱氣	*heat*

(underlining indicates the pitch shape is of short duration). In many Wu dialects, including Wencheng, the voiced series of initial stops is normally realised with coincident VOT word-initially, and thus the contrast between voiced and voiceless unaspirated appears neutralized word-initially.

The Oujiang dialects are a well defined first-order sub-group of Wu spoken in SE Zhejiang province, and Wencheng is spoken in the SW of the Oujiang area [7]. All Oujiang varieties show an unusual innovation from proto Wu involving the loss of a syllable-final stop and compensatory over-lengthening and elaboration of tone contour [7]. Wencheng shows this typical Oujiang *Ersatzdehnung*. For example the word for *white* 白, which in most other Wu dialects has a tone with a short low (often rising) pitch truncated with a glottal-stop ([paʔ 23]), as in Wencheng a tone with a long unstopped dipping pitch: [pa 213]. Oujiang dialects are typically tonally homogenous and have very similar tonal reflexes of Middle Chinese tones. Wencheng, however, does not partake of this homogeneity, and shows rather different tone shapes from the other Oujiang dialects [7].

This paper first describes the isolation tones of Wencheng, and then the lexical tone sandhi in a tonally well-defined set of disyllabic words. The term *tone* is used here in the sense analogous to *phone*, that is, as a constellation of audibly different properties, with pitch predominating but also including length and phonation type, that constitutes observation data for tonological analysis. An *isolation tone* is the tone given to a morpheme when it occurs on its own. This may be when the morpheme is free and occurs as a monosyllabic word in normal speech, or when its Chinese character is read out. In the latter case, where the morpheme can be either free or bound, the term *citation tone* is often used, the implication being that this is the tone with which the particular character is read. Both auditory and acoustic descriptions are given. The data are from a fine recording of one male speaker made by Professor W. Ballard in the 80's and his contribution is gratefully acknowledged here. The recordings were first segmentally and tonally transcribed, and then the tones analysed acoustically on the basis of the auditory transcription.

2. Isolation tones

The variety of Wencheng described here has seven isolation tones which I have named after their pitch. Their auditory properties, with some examples, are as follows.

The **Mid Level** tone has a level pitch contour just below the middle of the speaker's pitch range. The nearest representation in Chao's [2] procrustean but descriptively useful five-point 'tone letters' is [33]. Examples are: [fɛɪ *fly* 飛; tʊŋ *east* 東; kɔɐ *close* 關; fʊŋ *wind* 風].

The **Depressed Mid Level** tone has the same mid level pitch as the Mid Level tone, but its onset is depressed so that pitch over the initial third to half of duration is rising: [233], e.g. [pɛɪ *blanket* 被; z̥oʊ *sit* 坐; mɔ *horse* 馬; təʉ *belly* 肚; ʑ̊ʲɛ *elephant* 象]. Depression – lowering the pitch onset of a tone so that, for example, a depressed high falling tonal pitch is realised as convex – is a common feature of Wu tones, and

ISBN 978-0-9581946-3-1 © 2010 ASSTA
Accepted after peer review of full paper

is associated with word-initial position [6]. Figure 1 shows a typical F0 time course for the depressed mid level tone, on [pɛɪ] *blanket,* plotted against its wideband spectrogram. The initial F0 rise from about 110 Hz to 125 Hz over the first 10 csec. of the Rhyme is clear. Note that this lower F0 onset is clearly not a perturbatory function of any voicing on the syllable-initial stop, because there is none, the VOT being coincident.

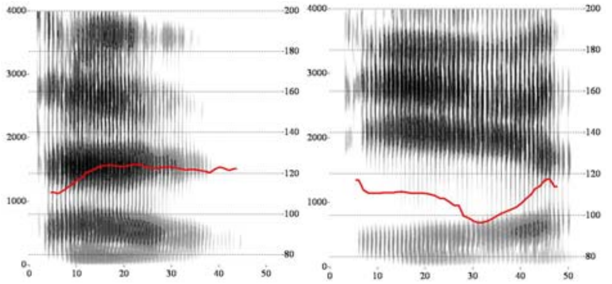

Figure 1: *left = F0 of depressed mid level tone in* [pɛɪ] *'blanket'; right = F0 of mid dipping tone, in* [tɪɛ] *'electricity'. F0 scale (Hz) at right; spectrogram scale (Hz) at left; duration (csec.) at bottom.*

The **Low Level** tone has level pitch a little lower than the mid level tone: [22], e.g. [ɲɪɛ *year* 年; tɪɛ *field* 田; maŋ *door* 門; tɕɪɛ *long* 長, kʰɑ *quick* 快; sz̩ *four* 四; saŋ *letter* 信].

The **High Rising** tone has a rising pitch from the middle into the upper pitch range, and is short. Phonation is optionally truncated with a glottal stop: [34(ʔ)], e.g. [tɕʊ *wine* 酒; ɕʊ *hand* 手; kʷoə *vast* 廣; fu *fire* 火; hæ *good* 好].

The **Low Rising** tone has pitch which rises from low in the speaker's pitch range to mid: [23 ~ 24], e.g. [pe *north* 北; ɕyə *snow* 雪; kʉə *bone* 骨; fʷɔ *put forth* 出; tɕʲoʊ *bamboo* 竹].

The **Low Dipping** tone has a dipping contour, first falling within the speaker's low pitch range to the their lowest pitch, and then rising into the mid pitch range: [213]. As might be expected from this complex contour, length is above average: [ɲyɛ *moon* 月; ɔ *study* 學; z̩eɪ *stone* 石; pa *white* 白; z̩a *ten* 十].

The **Mid Dipping** tone also has a dipping pitch, but its onset is higher, in the mid pitch range. It often has a prolonged initial level component; and it does not fall so low: [3(3)23]. Its length sounds well above average. E.g. [joɐ *use* 用; wɑɐ *rice* 飯; tɪɛ *electricity* 電; mɪɛ *face* 面; pɛŋ *ill* 病]. Figure 1 shows a typical F0 time course for this complex tone, on [tɪɛ] *electricity,* plotted against its wideband spectrogram. Over about the first third of the Rhyme, the F0 can be seen to maintain a level contour at about 110 Hz. It then falls to just below 100 Hz in the second third, and finally rises to just under 120 Hz in the last third.

It can be seen that the seven Wencheng isolation tones can be neatly characterised as the intersection of three pitch contours: *level*, *rising*, and *dipping*, each with a relatively higher and lower pitch. (The depressed mid level tone can be taken as an example of a level pitch contour.) The system is highly unusual – at least in global typological terms: Wu dialects actually prefer complexity in their isolation tones – for the following reasons. It contains at least two complex tones (tones with more than two pitch targets [9]; no falling tone, and two level tones only minimally separated.

2.1. Isolation tone acoustics

Figure 2 shows the mean F0 of the seven Wencheng tones plotted as a function of mean absolute duration (means are over five tokens). The two parallel level F0 shapes for the mid and low level tones can be seen in the middle of the F0 range, separated by a little more than 10 Hz. The F0 of the depressed mid level tone can be seen first rising, then flattening-out at just above the F0 of the mid level tone. The conspicuously short high rising F0 of the high rising tone can be seen at the top of the plot. Confined mostly to the lower half of the F0 range are no less than three tones with rising F0 components.

The low rising tone and the low dipping tone have near identical onset values at about the same level as the low level tone, and also have very similar F0 derivatives over the first few centiseconds of their Rhyme. After about 5 centiseconds into the Rhyme, the low rising tone F0 diverges from that of the low dipping tone and rises to the F0 onset height of the high rise tone. The F0 shape of the low dipping tone continues to fall to the lowest value in the speaker's F0 range at just above 90 Hz, and then recovers to a value in the middle of the F0 range. Finally, the mid dipping tone F0 onsets at a value around the same height as the low level tone, falls at first gently and then more abruptly to a trough at about 100 Hz, and then rises to a peak in the mid F0 range. There appear to be three F0 onset loci: the highest at about 135 Hz for the high rising and mid level tones; one between about 115 Hz and 120 Hz for the mid dipping, low level and depressed mid level tones; and the lowest at about 110 Hz for the low dipping and low rising tones.

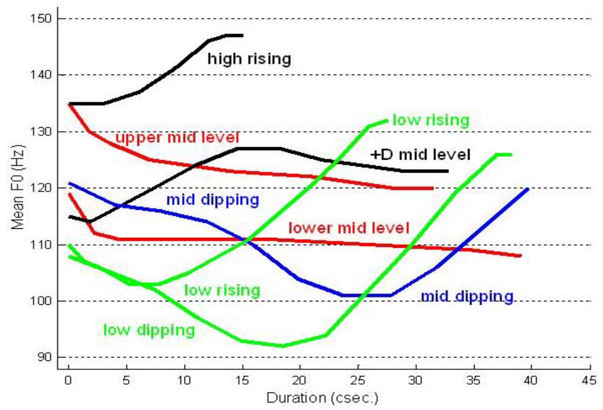

Figure 2: *Mean F0 of the seven Wencheng isolation tones plotted against mean absolute duration.*

2.1. Tonological representation of isolation tones

The isolation tone configuration just described presents some problems for tonological representation. Currently the favoured model for Asian tone languages [9] represents tones in terms of two components. *Register* divides the pitch range into an upper and lower half to define upper [+U] and lower [-U] registers. *Contour* is specified using (sequences of) high (H) and low (L) autosegmental tones. Thus for example a tone with a high rising pitch is represented as [+U, LH], i.e. a sequence of LH tones in the upper pitch range. This model will get us some way in assigning representations to Wencheng tones. The mid and low level tones for example live in the upper and lower halves of the F0 range, and can be non-controversially represented as differing in register and contour thus: [+U, L] (the mid level tone), and [-U, H] (the low level tone). Likewise the two rising tones must be represented with a LH contour, differing in register, thus: [+U, LH] (high rising tone), and [-U, LH] (low rising tone).

Table 1: *Examples of Wencheng lexical tone sandhi in combinations with Ru (low rising/ low dipping) tones on the first syllable*

low rising isolation tone on S1 followed by ..		low dipping isolation tone on S1 followed by ..	
... mid level isolation tone on S2:			
tɕe kwɔ 2.33	浙江 *Zhejiang province*	ɲɥə kwɔ 2.33	月光 *moon*
... low level isolation tone on S2:			
kɯə dʑi 2.22	國旗 *national flag*	ɲɥə de 2.22	月台 *platform*
tɕă tɕʰi 2.22	腳氣 *beriberi*	pa pu 2.22	白布 *calico*
... high rising isolation tone on S2:			
tɕʰɥə kʰaʊ 2.34	出口 *exit*	z̥a paŋ 2.34	日本 *Japan*
... depressed mid level isolation tone on S2:			
tɕɪ dʑaŋ 2.33	接近 *adjacent*	wu dʊŋ 2.33	活動 *activity*
... mid dipping isolation tone on S2:			
ɪ dɛŋ 3.331	一定 *definitely*	ɲɪ ʋɑ 3.331	熱飯 *hot rice*
... low rising isolation tone on S2:			
tɕʰɪ tɕa 3.23	赤腳 *barefoot*	pa tʰɪe 3.23	白鐵 *galvanized iron*
... low dipping isolation tone on S2:			
tɕʰɪ leɪ 2.21	吃力 *tired*	pa ja 2.21	白藥 *medicinal powder*

The remaining tones are more problematic. The depressed mid tone appears a problem, firstly because modeling its rising-level pitch/F0 as LHH violates the so-called Obligatory Contour Principle. This can be obviated by incorporating mora in the tonal representation, and interpreting the tone as LH associated with two moras, the H being associated with both. However, there then remains the problem that its pitch/F0 also straddles the boundary between upper and lower registers, such that even were it modeled as LH, it would fit in neither upper nor lower registers. These problems can be solved by recalling that depression in Wu is a word-initial phenomenon. The depressed onset can then be modeled as a L tone which is attached to the word, not the syllable. This allows the depressed mid level tone to be represented plausibly as <L> [+U, L] (thus showing the relationship with the mid level tone).

Far more problematic are the two dipping tones. As these have the same HLH contour, the model requires them to be represented with different registers, thus: [+U, HLH] (mid dipping), [-U, HLH] (low dipping). But this clearly is at odds with the phonetic reality, as both lie almost totally in the low pitch/F0 half of the speaker's range. Another clash with reality comes with a comparison of the low rising and low dipping tones. As pointed out, both these low register tones clearly have the same onset pitch/F0, but it is represented with opposite contour specifications (L vs H): [-U, LH] in the low rising tone, and [-U, HLH] in the low dipping. This clash between phonetic reality and phonological representation also raises its head in modeling aspects of the Wencheng lexical tone sandhi, which is demonstrated in the following section.

3. Sandhi tones

Given the complexity of the Wencheng isolation tones, it is of interest to see what happens to them in tone sandhi. In Oujiang, as in many Zhejiang Wu dialects, the simplest sandhi environment – at least for disyllabic words – is word-finally after tones of the historical *Ru* category on the first syllable [5, 4]. In Oujiang, disyllabic words with first syllable *Ru* input tones usually have tones on their final syllable which are very similar to the isolation tones, and very short tones on the first syllable. The two *Ru* tones in Wencheng are the low rising tone and the low dipping tone, and Wencheng disyllabic words with these tones on the first syllable do indeed have the characteristics of short first syllable tone followed by isolation type tone. Table 1 gives some examples, where it can be seen, for example, that in a disyllabic word with an input mid dipping tone on the second syllable and a low rising tone on the first (ɪ dɛŋ 一定) the resulting pitch shape is short mid [3] on S1 followed by mid fall [331] on S2.

What is happening in these sandhi combinations is transparent and easily expressed in words. The pitch of the second syllable tone closely resembles that of its isolation counterpart. For example a mid level [33] tone is realised as [33] on the second syllable; a high rising [34] is realised as [34]. Of interest are the two isolation dipping tones. On second syllables they lose their final rise to become tones with mid falling [331] and low falling [21] pitch. This change, which can be nicely represented as a loss of the final H in the tones' HLH contour sequence, is probably an instance of the well-known phonological constraint against tone crowding.

The depressed mid level tone loses its depression to become [33]. Since depression is associated with word-initial position, its non-occurrence is to be expected non-word-initially.

The pitch on the first syllable (recall that this has input low rising or low dipping values) is short and located in the middle of the pitch range [3], or just below [2]. The default realisation seems to be [2], with the two higher [3] values occurring before the low rising pitch tone [23] and the mid falling pitch tone [331] on the second syllable. The shortness of the first syllable tones is clearly the result of truncation of the corresponding isolation tone: both low rising [23] and low dipping [213] tones have lost all but their first components to give [2]. One would probably motivate this as the simplification of complex tones on metrically weak first syllables. The slightly higher value of [3] appears to result from conditioning by the pitch height of the following tone: a dissimilation before the [23] pitch of the low rising tone; and an assimilation before the [331] pitch of the truncated mid dipping tone.

3.1. Sandhi tone acoustics

Figure 3 shows the mean acoustics corresponding to the different sandhi combinations described in table 1. In each panel, two sets of F0 curves are shown corresponding to the two different first-syllable isolation tones: (red) triangles indicate F0 of combinations with isolation low rising tone on the first syllable; (black) circles indicate F0 of combinations with isolation low dipping tone on first syllable. F0 on voiced intervocalic consonants is shown with thinner lines.

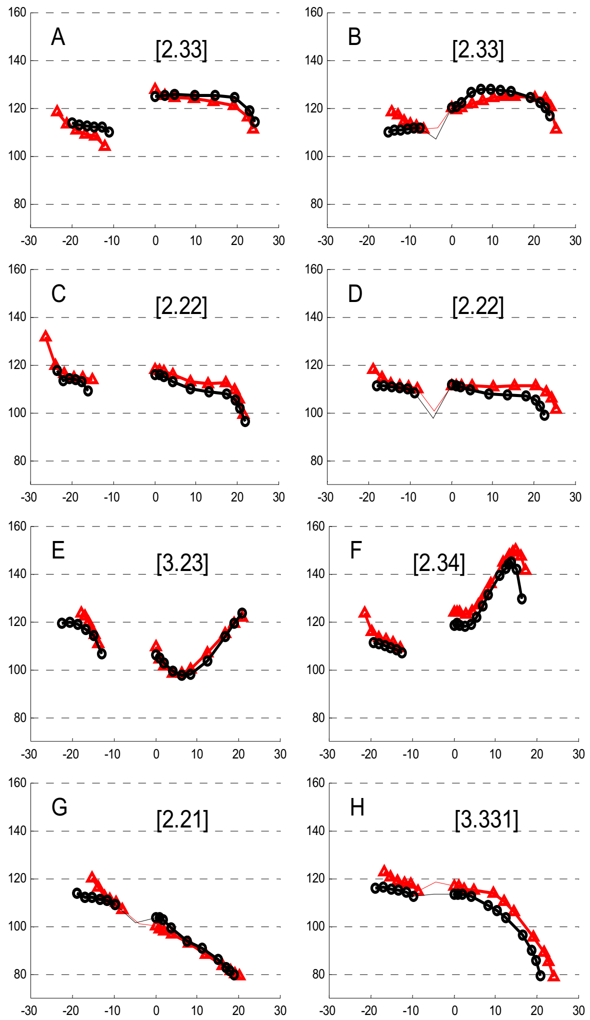

Figure 3: *Mean F0 shapes for disyllabic Wencheng words with low rising and low dipping ("Ru") isolation tones on the first syllable. Second syllable tones are: mid level, depressed mid level (panels A, B); low level(panels C, D); low rise (panel E); high rise (panel F); low dipping (panel G); mid dipping (panel H).*

The acoustics generally reflect the auditory impression well. The short nature of the first syllable tone is clear, with its Rhyme duration of about 10 csec. - about half that of the second syllable Rhymes. Pitch height dissimilation of the first syllable tone is clearly seen in panels E and F; assimilation in panels G and H. Of considerable interest is the consistent difference in F0 corresponding to the truncated low rising and low dipping tones on the first syllable. This is too small to reflect a tone difference and is likely to relate to phonatory differences associated with depression. Finally, note that the acoustic and auditory evidence of truncation on the first syllable confirms and reinforces the problems of tonological representation discussed for the low rising and dipping isolation tones. The remaining portion of these tones on the first syllable is clearly the same, and we would therefore expect them to have the same tonological representation. This is not possible if input low rising and low dipping tones are represented with their first mora carrying different, L and H, tones.

4. Discussion

The less than satisfactory performance of conventional tone feature geometry to account for the Wencheng morphotonemic alternations described above is at least partly due to historical tonal changes. In Wu, the shapes of isolation tones have changed relatively more than the sandhi shapes of corresponding tones in phonological words [11]. For example, the Wencheng low rising [24] and low dipping [213] tones are cognate with rather different tones in Wenzhou 60 kilometers to its NE [5]. In Wenzhou, these tones are either +/- depressed mid falling-rising tones [312] and [212], or +/- depressed mid falling tones [31] and [21]. Yet the first-syllable sandhi forms for these tones are very similar. For example, *Zhejiang province* and *moon* are [tse kwɔ, ɲy kwɔ 1.33] in Wenzhou; and *exit* and *Japan* are [tɕ^{h}y k^{h}aʊ, z̥aɪ paŋ 1.34] (the Wencheng values are in table 1). When complex sandhi and isolation tones have parted company in this way it is not surprising that phonetically-based features fail. However, a more important part of the solution might be to jettison the unique underlier condition on the phonological representation of morphemes: to not insist that the tone on a morpheme in a polysyllabic word be derived from the same tonological representation as the morpheme's isolation tone.

5. Acknowledgements

Thank you to my reviewers for their spot-on comments.

6. References

[1] Chao Y. R., 現代吳語的研究 *Studies in the Modern Wu Dialects,* Tsing Hua College Research Institute Monograph 4, 1928.

[2] Chao Y. R., A system of tone letters, Le Maître Phonétique 45:24-27,1930.

[3] Rose, P.J., "Phonetics and phonology of the yang tone phonation types in Zhenhai", Cahiers de Linguistique Asie Orientale 18(2):229-245, 1989.

[4] Rose, P.J., "Acoustics and Phonology of Complex Tone Sandhi", Phonetica 47: 1-35, 1990.

[5] Rose, P.J. "Wenzhou Dialect Disyllabic Lexical Tone Sandhi with First Syllable Entering Tones", in M. Barlow [Ed], Proc. 8th Australian Intl. Speech Science and Technology Conf., 230-235, Australian Speech Science and Technology Association, 2000.

[6] Rose, P.J., "Independent depressor and register effects in Wu dialect tonology: Evidence from Wenzhou tone sandhi". Journal of Chinese Linguistics, 30(1):39-81, 2002.

[7] Rose, P.J., "Oujiang Wu tones and acoustic reconstruction", in C. Bowern, B. Evans & L. Miceli [Eds], Morphology and Language History, 235-250, John Benjamins, 2008.

[8] Steed, W., "Phonation Type and Tone Sandhi as Evidence of Chinese Stress in Longquan", in C. Watson & P. Warren [Eds], Proc. 11th Australian Intl. Conf. on Speech Science and Technology, 82-87, The Australian Speech Science and Technology Association, 2006.

[9] Yip, M., Tone, Cambridge University Press, 2002.

[10] Zhu Sean, and Rose, P.J., "Tonal Complexity as a Dialectal Feature: 25 Different Citation Tones from Four Zhejiang Wu Dialects", in R. Mannell & J. Robert-Ribes [Eds], Proc. 5th Intl. Conf. on Spoken Language Processing, 3:919-922, The Australian Speech Science and Technology Association, 1998.

[11] Rose, P.J., "Reverse-engineering tones and tone sandhi in Wenzhou dialect: between-speaker differences and historical development逆序制造温州方言聲調和連讀變調兩個人差異與歷史發展", Bulletin of Chinese Linguistics, (in print).