Draft for Rose (2008) in Bowern et al. (eds.) Morphology and Language History, John Benjamins.

OUJIANG WU TONES AND ACOUSTIC RECONSTRUCTION

PHIL ROSE
Linguistics (Arts), Australian National University

1. Introduction

We do not expect tones, as phonological phenomena, to vary without limit, but to show similarities, or so-called tonological universals, across varieties (Maddieson 1978). Nevertheless, when one samples tones from different Asian tone languages, or even different dialects, they often show a great variety of pitch shapes. Zhu and Rose (1998) for example demonstrated no less than 25 different tones in four varieties from different sub-groups within the Wu dialects of east central China. (I use the term *tone* in this paper in a 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 phonological – usually tonemic – analysis.)

In exploring the nature of phonological variation in Language, we naturally concentrate on differences. This paper, on the other hand, makes use of how similar tones can be. It looks at variation in tones at seven localities distributed over a fairly small geographical area where the varieties, which belong to the same dialectal sub-group, have been described as homogeneous. Although this study reveals some interesting synchronic linguistic-tonetic variation, the main reason for undertaking it is not typological, but, as befits a festschrift for Harold Koch, historical. As mentioned, the Wu dialects of Chinese are well-known for their tonal complexity. Not only do they exhibit probably the most complex tone sandhi in the world (Rose 1990), but they also show a bewildering variety in their citation tones. Most Wu varieties have eight citation tones, but these can differ considerably in their shape. It is a major challenge to find out how it got that way. In order to apply the comparative method, it is normal to start from the smallest well-defined groupings. That is what this paper does for a first order sub-group of Wu called Ōujiāng 甌江. It brings together acoustic descriptions of tones of speakers from several different Oujiang sites, in order to reconstruct two of the tones of proto-Oujiang.

Attempts to reconstruct actual tonal values from modern source material are rare. Possibly because of the bewildering synchronic variety of tonal shapes, and few clear indications of common ways tones can change, it is normal to reconstruct abstract tonal categories with no phonological content (e.g. Brown 1985). One notable, and relevant, exception is Ballard (1969: 70, 92-98) who reconstructed pitch values for proto-Wu tones from cognate sets, and tone sandhi behaviour, in 16 Wu sources. This paper is novel in its attempt to go one further and reconstruct acoustic values for proto tones. It has the following structure. I first briefly characterise the Wu dialects and their Oujiang sub-group. Next, acoustic data of Oujiang citation tones are described from seven sites. Then, as an important precursor to acoustic reconstruction, a method of quantifying homogeneity in tonal acoustics is presented. Reconstructed acoustics and measures of proto variation are then derived by means of normalisation of two demonstrably homogeneous tones.

2. Wu and Oujiang

The Wu dialects, of which Shanghai is a well-known variety, are spoken in the two eastern Chinese coastal provinces of Zhejiang and (southern) Jiangsu (Rose 2001). They are considered a first-order sub-group of Middle Chinese, along with five other major dialect groups of Chinese: Mandarin, Hakka, Yue, Xiang and Gan. The sixth major dialect group of Sinitic, Min, is assumed to have diverged before Middle Chinese (Norman 1988).

Oujiang is one of the sub-groups currently recognised for Wu (Zhengzhang 1987). Sources differ a little on the number of sub-groups. Zhengzhang (1987) lists six: *Taihu, Taizhou, Oujiang Wuzhou Chuqu, Xuanzhou*. The Oujiang sub-group, eponymous with the river bisecting it in a N.W. to S.E. direction (jiāng = *river*), is located in the S.E. corner of Zhejiang province, and is approximately congruent with the Wenzhou administrative area. About five million people speak Oujiang varieties, of which Wenzhou is probably the best known.

The Wu dialects have traditionally been considered a well-defined group either typologically, on the basis of their systematic three-way VOT contrast for syllable-initial stops and affricates (Chao 1967: 92, 94), or diachronically, since this tripartite contrast is taken to be a retention of the Middle Chinese three-way division (Chao 1928: 1). A subminimal triplet exemplifying this three-way contrast on palatal affricates, from the Wu dialect of Zhenhai (Taihu sub-group), is [mɪŋ tɕjɔ 11.41] 明朝 *tomorrow*, [mɪŋ dʑjɔ 11.41] 明朝 *Ming Dynasty*, and [lɪŋ tɕʰjɔ 11.41] 靈巧 *skillful*. A subminimal triplet from the Oujiang dialect of Wenzhou showing the three-way contrast intervocalically on an alveolar stop is: [tə te 23 331] 道德 *morality*, [tseɪ da 23.331] 直達 *to go straight to*, and [tʰi tʰa 4.331] 鐵塔 *Eiffel tower* (underlining indicates the pitch shape is of short duration).

As Harold is fond of pointing out, it is an important methodological principle in historical linguistics that sub-groups can only be established on the basis of shared, unusual innovations. The sub-group status of the Wu dialects, resting as it does on a retention from Middle Chinese, is therefore suspect from an orthodox point of view. However, from a Bayesian point of view, the traditional position articulated by Chao is perfectly reasonable. Since the Wu dialects are very nearly the only varieties of Sinitic to have retained such a structural property (the vast majority of the other dialects has devoiced the Middle Chinese voiced series of obstruents leaving only a two-way contrast), the existence of a systematic three-way contrast in an idiolect is strongly diagnostic of Wu-hood.

Most of the Wu sub-groups, too, have not been established on innovations, but are typically characterised as showing various constellations of phonological, lexical and morpho-syntactic features. Oujiang, however, is different. Oujiang can be considered a *bona-fide* sub-group because of its unusual development of two proto-Wu tones. Proto-Wu is reconstructed with eight tones, two of which - tones *IVa and *IVb - occurred on syllables with short Rhymes and final stops. Elsewhere in the Wu area, reflexes of these tones are usually short, with a word-final glottal-stop, e.g. Zhenhai [pǎʔ 5̲] (< proto-Wu *pak 45) 百 *hundred*; & [bǎʔ 24] (< proto-Wu *bak 23) 白 *white* (reconstructions from Ballard 1969: 70). Oujiang dialects, however, show an unusual compensatory lengthening, whereby the proto-Wu final stop in *tones IVa/b has been lost, and the original short tone has developed a long, complex pitch. The cognates of the Zhenhai IVa & b tone morphemes *hundred* & *white* in the Oujiang dialect of Wenzhou, for example, are [pa:: 312] & [ba:: 212].

The published auditory descriptions of Oujiang tones show them to be generally homogeneous (Fu et al. 1985: 20). With the exception of Wencheng, Oujiang dialects have very similar pitch shapes for reflexes of Middle Chinese (and by extension proto-Wu) tones. This makes the reconstruction of most proto Oujiang tones from the auditory descriptions non-controversial. Such a reconstruction is shown in table 1. The tones show a neat pattern with three pitch contours - level, rising and falling - each with a relatively higher and lower pitch resulting from the depressor effect. (Depression is a phonological effect associated with word-initial position in many Wu dialects, whereby the pitch onset of a tone is lowered, so for example a depressed *falling* tone will have a *convex* pitch shape (Rose 2002). One interesting fact is the asymmetry in the level and falling tones: the upper and lower paired contours are not reflexes: the level tones come from *Ia & *IIIb; the rising tones from *IIa & *IIb; and the falling tones are from *IIIa & *Ib. In each case, the *a* reflex is higher than the *b*. The two remaining tones – the elongated reflexes of *IVa & *b – also typically form a higher and lower pair. Although the acoustics will uncover some interesting variation in this pattern, it helps to keep it in mind as a useful synchronic template when the individual speakers' tones are discussed.

	I	II	III	IV
a	high level	high rising	high falling	mid dipping
b	mid falling	low rising	low level	low dipping

Table 1. *Proto Oujiang tones reconstructed from auditory data.*

3. Acoustic data

Acoustic descriptions of tonal fundamental frequency (F0) and duration are available for seven Oujiang sites: Wēnzhōu 溫州, Xiàngyáng 象陽, Yǒngjiā 永嘉, Yuèqīng 樂清, Qīngjiāng 清江, Píngyáng 平陽 and Wénchéng 文成. Figure

1 shows their location in the Oujiang dialect area as delineated by Wurm et al. (1987). This selection of sites gives a fairly evenly distributed sampling of all the Oujiang area except for the N.W, where the Oujiang dialects border on the Chuqu sub-group of Wu. The data are from my measurements and analysis of some very fine recordings made by W.L. Ballard in the 1980's and by Zhu Xiaonong in the 1990's. The procedure used to extract the acoustic tonal features can be found in Rose (1990, 2002).

***Figure 1:* Map of the Oujiang dialect area, after Wurm et al. (1987), showing location of sites mentioned in this paper. Dashed line indicates borders with other Wu sub-groups. Solid line indicates border with Min dialects.**

Figure 2 shows the mean citation tonal acoustics for the seven Oujiang sites. Each panel has the tones of a single speaker (means of between five and ten tokens per tone). F0 is plotted as a function of absolute duration: this is because there are some characteristically big differences in the duration of different tones in Oujiang which would otherwise be obscured. The top three rows show Wenzhou, Xiangyang and Yongjia, each row with two speakers shown side by side. The remaining four sites are represented by one speaker each, and are shown in the bottom two rows. The tones are identified by their Middle Chinese tonal categories: Ia/b, IIa/b, IIIa/b, and IVa/b. This makes it easier to see which tones are cognates, an important consideration because there are some potentially confusing differences between speakers and varieties. The tones are also labeled with simple pitch descriptors, like 'mid level'; '+D' indicates depression.

3.1 *Wenzhou citation tone acoustics*

The two Wenzhou speakers' tones are shown in the first row of Figure 2: a middle-aged male recorded by Ballard, and a young female I recorded in 2006. The corpus and elicitation is described in Rose (2002), and was exactly the same for both speakers. Both show the three level, rising and falling F0 contours already described: tones Ia & IIIb have level pitch; tones IIa & IIb rising; and tones IIIa & Ib falling pitch.

Although both speakers conform largely to the typical Oujiang (OJ) configuration, they differ from the auditory descriptions, and each other, in several ways. (1) The female does not distinguish an upper and lower rising tone: both her IIa and IIb tones are realised with the same low rising pitch/F0 (F0 differences at onset and peak are not statistically significant). Compare this with the male's separate reflexes of IIa and IIb, which present a typical Oujiang configuration of higher and lower rising tones. Thus the female can be said to only have seven citation tones. (2) Whereas the male has typical long complex falling-rising realizations for his IVa and IVb tones (the falling part being the more salient), the female lacks the rising component for both these tones. They consequently have higher and lower mid falling pitch. This realisation for IVa/b, without a rise, does not occur in any published OJ transcription. (3) The female's mid falling tonal F0 (Ib), seems to fall from a relatively higher value than the male's, and this is presumably related to the fact that she already has a mid falling tone: IVb. It can be seen that the F0 of her IVb tone has a very similar level-falling shape to that of the male's mid falling Ib tone. The female's Ib tone also has a clearer rising onset perturbation from the depressor effect than the male. In her speech, the depressor was conditioned by the [+/- sonorant] feature of the onset consonant: it was audible on syllables with an obstruent onset, but not on those with sonorants. (4) The speakers differ in the pitch height of their Ia tone. The female's Ia tone has a upper mid level pitch and corresponds to the auditory transcription in table 1 of [44]. The pitch of the male's Ia tone, on the other hand, sounds lower, in the middle of his pitch range, and would be best transcribed as [33]. The difference in F0 height corresponding to this pitch difference is easy to see in figure 2, with the male's Ia F0 lying considerably lower in his F0 range than the female's. It is of interest to note that the historically paired Ia and Ib tones, although they have different contours, appear to be linked in the sense that they both involve similar initial targets. This relationship can also be seen in the other speakers.

3.2 *Xiangyang citation tone acoustics*

The two Xiangyang speakers, a brother and a sister, were recorded by Ballard with the same corpus and elicitation as the Wenzhou speakers above. No published auditory descriptions exist, therefore I have given mine for the two

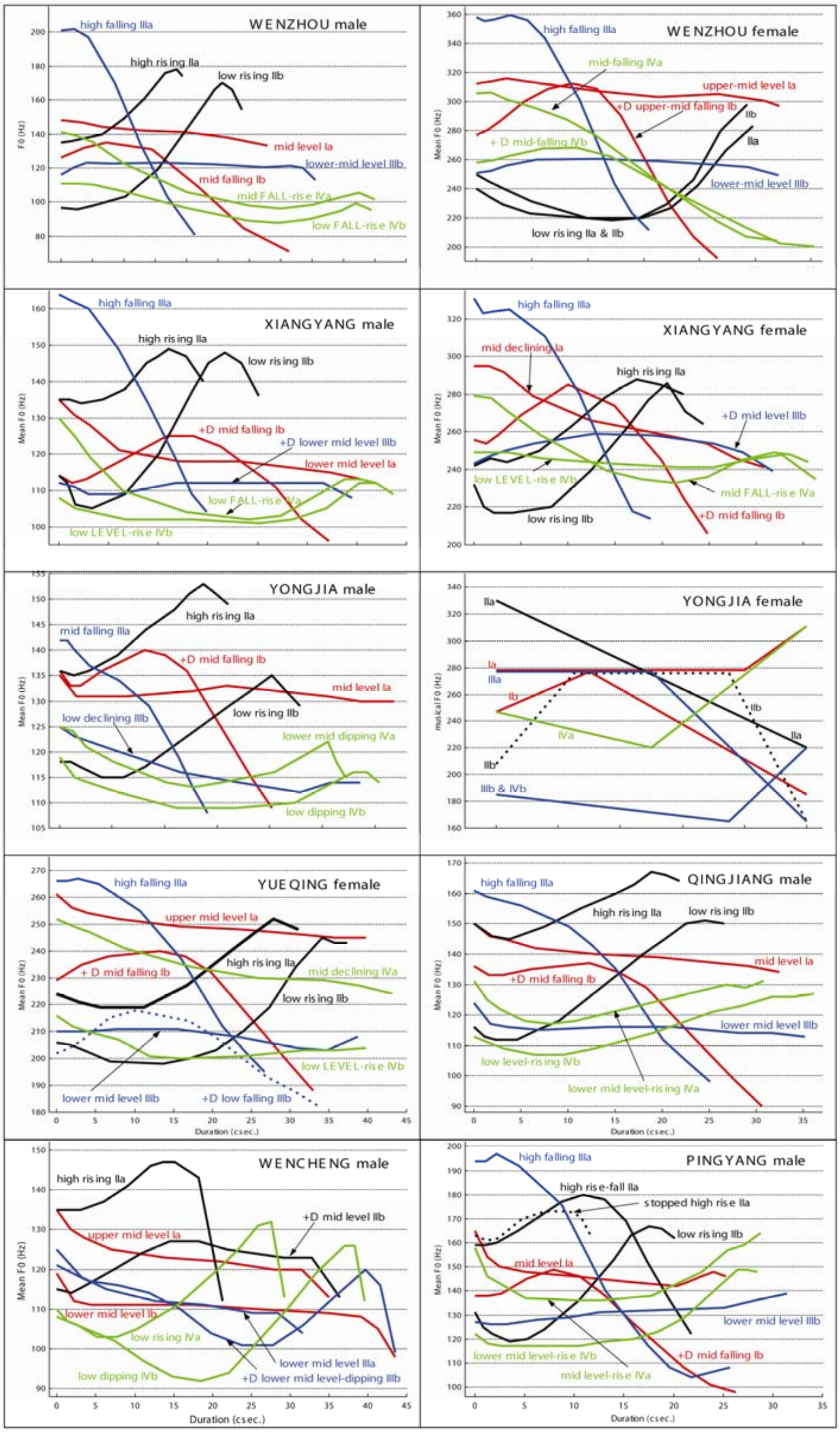

Figure 2: *Citation tone acoustics of speakers from seven Oujiang varieties.*

speakers in table 2. It can be seen that both have eight tones, configured in the typical OJ way, with the paired level (or quasi level), rising, falling and complex contours seen in the Wenzhou male. There is not a lot of difference between the speakers in the tones' pitch features. Their acoustics are shown in the second row of Figure 2. As with Wenzhou, tone IIIa and tone Ib constitute a high and mid-falling pair, the latter also showing a clear depressed onset for both speakers. It can be seen in figure 2 that both speakers have very similar F0 shapes for their IIIa tone, with an F0 which falls through most of their F0 range. They also have similar rising-falling F0 contours for their Ib tone, (the rising portion being a reflex of the depressor) but in slightly differing locations: the F0 peak of the female's Ib lies slightly higher than that of the male. This is also reflected in their pitch transcription ([231 vs 341]). The Xiangyang speakers are also similar to the majority of the Wenzhou data in showing a pair of rising-pitched tones for tone IIa & tone IIb. There is little between-speaker difference in the pitch of these tones, with the female having just a slightly lower pitch onset to her upper rising tone IIa. Figure 2 shows the F0 of these tones rising into the upper half of the speakers' F0 range from low (IIb) and either a mid or low-mid location.

Tones Ia and IIIb constitute a further contour pairing, as in Wenzhou, but the details are a little more complicated. Both speakers have a mid or low level pitch with depressed onset for IIIb; their Ia tone has very similar pitch level as their IIIb, but for the female it falls slightly [43], whereas for the male it is level [22]. These relationships can be seen in the F0 data. The male's Ia and IIIb F0 shapes lie close together in the lower third of his F0 range, with the IIIb shape lying slightly lower and showing a depressed onset. The female's Ia & IIIb lie slightly higher in her F0 range, with the Ia shape falling steadily to an offset which is similar to her IIIb tone. Her IIIb shape also shows a depressed onset. The fourth pairing obtains, as with Wenzhou, between tone IVa and tone IVb. In both speakers, IVa has a falling-rising pitch, with the falling component more salient, whereas IVb has a low level-rising pitch (thus this – low level rising - is another pitch shape for IVb). The speakers differ in the location of these complex contours, with the female's tones located slightly higher in her pitch range. The tonal F0 for tones IVa & IVb reflects their pitch details nicely. Figure 2 shows that both speakers show remarkably similar F0 contours for tones IVa & IVb, but with the male's tones located lower.

	Ia	Ib	IIa	IIb	IIIa	IIIb	IVa	IVb
female	[43]	[341]	[24]	[14]	[51]	[233]	[423]	[223]
male	[22]	[231]	[34]	[14]	[51]	[122]	[312]	[112]

Table 2: *Pitch of Xiangyang citation tones.*

3.3 *Yueqing citation tone acoustics*

The Yueqing speaker's tones are plotted in the fourth row of Figure 2 and are from a recording made by Zhu Xiaonong of a female speaker. She has eight tones with the usual OJ configuration of paired level, rising and falling contours. Her tonal pitch characteristics correspond largely with the published descriptions, which show the common pairing of level, rising, falling and dipping contours already documented for other sites. Figure 2 shows the two falling F0 shapes for her tones Ib and IIIa; the two rising F0 shapes for tones IIa and IIb, and a high level F0 shape for tone Ia. Her tones agree too, by and large, with the previous speakers. There are two main differences. (1) Zhu's speaker has a new shape for IVa: a slightly falling pitch [32] in her mid pitch range. The corresponding gradually falling F0 shape for her IVa tone can be easily picked out in figure 2. The published descriptions indicate an expected mid-dipping [323] for this tone, so perhaps this female lacks the final rise. (2) She also has two allotones for the reflex of IIIb. One is the low level pitch version common to most other speakers; the other, plotted with dots in Figure 2, is a depressed low falling tone [121]. There is no obvious segmental conditioning to the split which would predict which morphemes are said with level pitch and which with falling, and it may be that her depressed low falling IIIb morphemes represent a merger in progress with Ib which also has a depressed falling pitch/F0. Against this hypothesis is the fact that, as can be seen from Figure 2, the falling IIIb morphemes have much lower F0, and sound much lower in pitch than the Ib morphemes.

3.4 Qingjiang citation tone acoustics

The Qingjiang speaker's citation tone acoustics are on the right of the fourth row in Figure 2. They are from recordings made by Zhu of a young male speaker. There are no published auditory descriptions. From table 3, which contains my pitch transcription of Zhu's speaker, it can be seen that these data are very similar to those already presented for the other speakers, comprising four contours, each with an upper and lower version. Three of the contours are the same as Wenzhou, Xiangyang and Yueqing, i.e. level, rising and falling. The F0 corresponding to these can be seen in Figure 2. Of note is that the lower level tone IIIb lacks the depressor found in the other speakers, and the high falling tone IIIa has a slight onset shoulder. Of greatest interest is the contour for the fourth pair of tones – tone IVa & tone IVb – which is delayed low rising, and therefore differs from that found in the other speakers who have falling-rising, or falling pitch contours. Comparison with the other speakers' tone IV F0 shapes in Figure 2 shows that the Qingjiang shapes are clearly different. The concentric F0 for this pair of tones starts to rise about a third of the way into the Rhyme (it is likely that their initially falling F0 is a consonantally perturbed onset perturbation and not perceived as a fall in pitch). Although this contour is not found in the other speakers, something obviously like it – [23] and [12] – is noted for IVa and IVb in two of the Wenzhou sources (ZH 1962, CH 1964).

Ia	Ib	IIa	IIb	IIIa	IIIb	IVa	IVb
[33]	[331]	[34]	[14]	[441]	[22]	[223]	[113]

Table 3: *Pitch of Qingjiang citation tones.*

3.5 *Yongjia citation tonal acoustics*

The Yongjia tonal acoustics in the third row of Figure 2 are from two sources: a male speaker recorded by Ballard, and the female speaker described by Chao way back in 1928. In his pioneering description of Wu dialect phonetics, Chao transcribed tonal pitch musically, using a sliding pitch pipe. Chao's musical description of Wu tones remains unsurpassed, and his notation permits easy transformation into F0 values on the simplifying assumption that tonal pitch is a primary function of F0 – see Rose (1993: 215, 216) for similarly reconstructed Shanghai tonal F0. Unfortunately, the duration of the tones cannot be reconstructed, and so the tonal F0 is plotted as a function of equalized duration in Figure 2.

Yongjia tones Ia and Ib show the same values as in the previous sites: upper-mid, or mid, level, and (depressed) mid-falling respectively. The acoustics of these tones in Figure 2 show nothing new. There is a lot more variation between the different sources' descriptions for the other Yongjia tones, however, and they show somewhat less agreement between the four Yongjia sources than is found for Wenzhou and Xiangyang.

Firstly, some variation is evident for tone IIIa. Two sites (Fu et al. 1995, Cao 2002) show the same [42] high-falling tone as for the other sites' tone IIIa. However, Chao's and Ballard's speakers are a little different. The male speaker Ballard recorded had a fall with a much lower onset: [31], and interestingly Chao also noted a *mid* fall (with an initial level component) for his speaker in 1928. This lower onset is shown clearly in Figure 2, where the F0 shape for tone IIIa can be seen to fall from a lower position in the F0 range for both speakers.

There is considerable between-source variation in the Yongjia reflexes of tones IIa and IIb. Ballard's Yongjia speaker, as well as one of the sources (Fu et al. 1985), has the same typical OJ pattern for IIa/b as already observed in all except the female Wenzhou speaker, namely realised as a pair of upper and lower rising pitch tones. The concentric F0 shapes of these two rising tones are easily identified in Figure 2. However, as also can be seen, Chao's 1928 Yongjia speaker had a very different pattern. His tone IIa had high falling pitch (musically transcribed as $\underline{53}^{b}$), and his IIb tone had a mid convex pitch ($\underline{241}$) which looks interpretable as its depressed counterpart (IIb is shown in Figure 2 with a dotted line for clarity). Finally, in the last source (Cao 2005) both IIa and IIb are shown as convex pitch tones. Thus Cao's notation seems to partake of features of the other three. The difference between rising and falling pitched tones, as demonstrated for Yongjia IIa and IIb, is very big. An extrinsic falling pitch/F0 for these two tones would of course present a problem for reconstruction, given that the other speakers and other sites have rising pitches for IIa and IIb. Zhengzhang (1995: 359), who is a native Wenzhou speaker, has commented on this variation. He notes the falling pitch for these tones in the earliest description of Wenzhou tones, by Montgomery in 1893, as well in Chao's 1928 Yongjia description. He says that a falling pitch for these tones occurs if the speaker tries to prolong their intrinsically short duration, as is the case when Wenzhou scholars declaim citation tones. (The falling part would then be construed nowadays as an intonational L# boundary tone.) Thus, he says, Chao's 1928 transcription was in error in interpretation of the pitch fall in IIa/b. It is difficult to dismiss the inside information from a very experienced native-speaker linguist like Zhengzhang, even if it impugns the field-work of a titan like Chao. And it is worth noting that the Pingyang speaker below shows free variation of a very similar kind, at least for IIa, between a stopped high rising pitch and an unstopped rise-fall pitch.

There are also between-source differences in the description of tone IIIb. For two sources (Cao 2002, Fu et al. 1995), IIIb is the same as for the other sites: a low level [22] tone, pairing tonetically with Ia. In contrast, Chao's 1928 speaker is shown as having a IIIb tone with a long low falling pitch component, followed by a little rise (2^b13^b) (its reconstructed F0 can be seen clearly in Figure 2). A very similar F0 contour for Ballard's speaker's IIIb tone, although it lies a little higher in the F0 range, is unlikely to be a coincidence, and so it seems that both Chao's and Ballard's speaker share a different reflex for IIIb than the low level reflex, with depressor, found in the other descriptions.

Tones IVa and IVb are a little less complicated. Firstly, although three descriptions show the usual pairing between high and low reflexes, Chao says his speaker lacks a reflex for tone IVb, and that it has merged with IIIb. This is why there are only seven tonal shapes shown for his speaker in Figure 2. Apart from this, there is little discrepancy between the descriptions of Yongjia IVa pitch.

3.6 Pingyang citation tone acoustics

The male Pingyang speaker's tones in the last row of Figure 2 are from a recording made by Zhu. The speaker shows the same basic configuration as the others, and the published sources, with two exceptions. (1) His IIa morphemes showed free variation between a short stopped high rising pitch [34], and an unstopped high rise-fall pitch [342]. The F0 corresponding to these two variants can be easily seen in Figure 2, where the Pingyang stopped IIa tone F0 is plotted with dots. It is of interest to note, too, that one of the modern sources gives a slightly falling [54] pitch for this tone, while the other cites the typical OJ high rising [45] pitch. We may see here, then, an actual example of the problem mentioned above for Yongjia, whereby the high rising tone is intrinsically short, but in citation form may optionally acquire a falling component. This could be modeled by having a LH(L) melodic component with the first two tones associated with a single mora, and the last, optional, L tone associated with a second mora. One can note further that the source that describes a falling pitch for IIa also describes a falling pitch component [243] for IIb, whereas Zhu's Pingyang speaker does not have this falling component for his IIb. This is perhaps an additional indicator of the optionality of the falling tail for these tones. In any case, the Pingyang example is important historically because it shows a possible source of a falling from a high rising tonal pitch. (2) The Pingyang speaker's IIIb tone appears to have an overall rising F0 contour, but this is due to two factors: a long depressed onset, and a clear final glottal stop, which raised its F0, and pitch, a little at the end. (3) The published sources show mostly non-complex rises for tones IVa/b. Zhu's Pingyang speaker, however, had rather long IV tones, with a substantial level pitch before the final rise. The F0 corresponding to this long level then rising pitch contour can be easily seen in Figure 2.

3.7 *Wencheng citation tone acoustics*

Ia	Ib	IIa	IIb	IIIa	IIIb	IVa	IVb
[33]	[22]	[34□]	[233]	[22]	[3323]	[23]	[213]

Table 4: *Pitch of Wencheng citation tones.*

Wencheng is dramatically different. The acoustic data on the last row of Figure 2, from a recording of a male speaker by Ballard, show far more differences from the other dialects than similarities. There are, to be sure, paired higher and lower level tones, and higher and lower rising tones, but they are not the same reflexes as for the other OJ varieties. The lower rising tone, for example, is a reflex of IVa, not IIb, and the lower level tone is a reflex of Ib, not IIIb. There are not even any falling tones. The only thing that Wencheng clearly shares with the other varieties is a mid level tone for Ia, and a high rising tone for IIa (its final falling F0 is an intrinsic effect from a [ʔ]). The other tones do not appear elsewhere in the OJ group examined. For example, there is no depressed upper-mid level tone (IIb), neither is there the marvelous super-complex reflex for IIIb, which seems to consist of a lower-mid *level* target, most likely a depressor, followed by a *dipping* contour. Even the low dipping contour for IVb rises much higher than in the other varieties, and the apparent merger of Ib and IIIa is a very rare phenomenon in Chinese.

There are also some discrepancies between the published descriptions of Wencheng tones and Ballard's Wencheng speaker's acoustics. My pitch transcriptions are therefore given in table 4. There is, by and large, agreement for tones Ia, Ib, IIa IIIb IVa and IVb. For example, corresponding to Ballard's speaker's [33] and [22] for tones Ia and Ib, both published sources describe a bigger contrast in pitch height, between high and low level pitches for Ia and Ib, with one source having a slight final rise: [44(5)] and [11(3)]. For tone IIIb, which has a [3323] pitch, no depressor is transcribed, but otherwise the low dipping contour [312, 313] is the same. The biggest difference is for tone IIIa, which has merged with tone Ib in Ballard's speaker (the clear difference in duration between IIIa and IIb in figure 2 is due to an intrinsic truncation effect from the initial aspirated consonant in the IIIa morphemes), but which has a separate high rising [334] or dipping [434] pitch in the sources. There is also clearly a difference in the IIb realisations. Ballard's speaker has a depressed mid-level pitch [233], whereas one of the sources shows a mid level pitch without the depressor, and the other shows a mid dipping [324] pitch.

Despite these differences, Wencheng still shows elongated reflexes of proto-Wu *IVa/b tones, and thus must count as a synchronic Oujiang variety. However, whether it is diachronically an Oujiang variety that has undergone considerable subsequent changes independently of the other Oujiang dialects; or whether it is originally not Oujiang but has changed under the latter varieties' influence, remains to be determined.

4. Quantifying tonal homogeneity in Oujiang

The sections above have described the tonal acoustics of speakers from seven different sites in the Oujiang sub-group. It is clear that some tones show very little variation from site to site. The reflex of Ia has mid or upper-mid level pitch in most sites, for example, and the reflex of IIb has a low rising pitch. How can one quantify this variation, however? One possibility is to measure the degree of homogeneity of the different speakers' tones by how well they cluster after normalisation. Normalisation is a mathematical procedure for getting rid of the individual characteristics in the speech acoustics to leave a linguistic-phonetic residue - what is common to the variety (Rose 1987, 2000). As well as speech-dependent information, a particular tonal F0 value contains speaker-dependent information. It is determined not only by the particular tonal target a speaker is aiming for, but also by the length and mass of their vocal cords.

A very simple normalisation method is to first find a speaker's mean and standard deviation F0 over all their tones, and then transform their individual F0 values by subtracting them from the mean and dividing by the standard deviation. This is called a z-score normalisation, after the equivalent transform in statistics. This method maximally eliminates variation in F0 due to between-speaker differences in cord mass and length. Speakers' normalised F0 values can then be averaged to derive a representation of the variety, as well as a quantification of the amount of variation in the variety, which can be modeled as a function of the standard deviation around the mean normalised curve.

Now, if it can be shown that the amount of variance around the mean normalised F0 shape for a given tone across the Oujiang varieties is the same as that found for a comparable tone *in a single dialect*, it can be claimed that the tones from the different Oujiang sites are as homogenous as in a single dialect. This idea can be demonstrated using two tones from the Wu dialect of Shanghai, a dialect for which there is data on 18 speakers' citation tones (Rose 1993, Zhu 1999), and for which there is consequently a reasonable estimate of the within-dialect between-speaker variance in normalised tonal F0. Shanghai (SH) has five tones, one of which has a high falling pitch (Tone 1) and another a low rising pitch (Tone 3). These sound very similar to the high falling IIIa and low rising IIb tones of Oujiang, and it was these two tones - high falling in SH and OJ, and low rising in SH and OJ - that were compared by z-score normalisation.

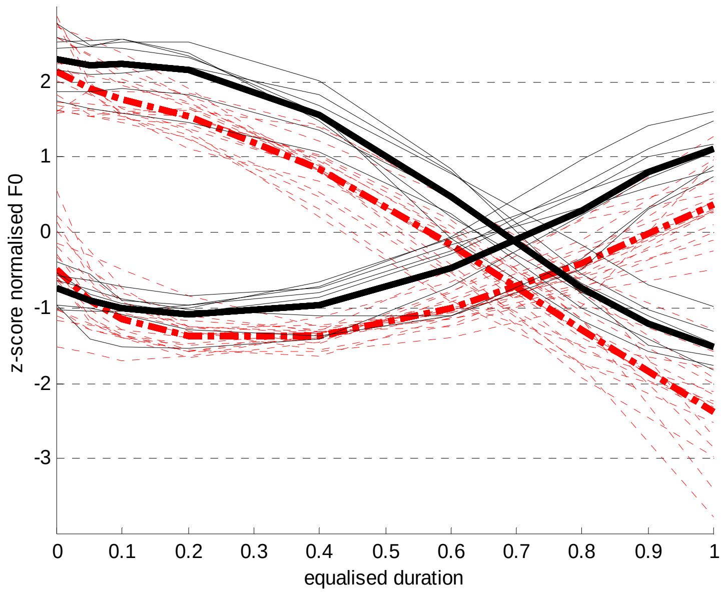

Figure 3: *Comparison by z-score normalisation of Shanghai and Oujiang high falling and low rising tones*

Figure 3 shows the z-score normalised values for the four tones plotted against equalized duration. The thin dotted lines represent the normalised tones for the 18 Shanghai speakers and the thick interrupted line their mean normalised value. The thin solid lines represent the normalised OJ tones and the thick solid line their mean normalised value. It can be seen that the OJ falling and rising tones lie less than a standard deviation above the corresponding SH tones, but that otherwise the two varieties' normalised contours are remarkably similar. (The vertical displacement is probably because OJ has more tones distributed in the lower F0 range than SH, and consequently a lower mean F0, leading to higher normalised values.) The standard deviations at each of the sampling points are given in Rose (2004: table 2), which showed generally very little difference between OJ and SH in these tones (the differences are of the order of one tenth of a standard deviation, which is exiguous).

Thus, in spite of the fact that the Oujiang descriptions span some twenty years, it is possible to claim considerable homogeneity for these two tones in the Oujiang dialects examined. The same could probably be demonstrated for all the other OJ tones except reflexes of IVa/b, which, as the results of the acoustic analyses above have shown, do vary over the Oujiang area.

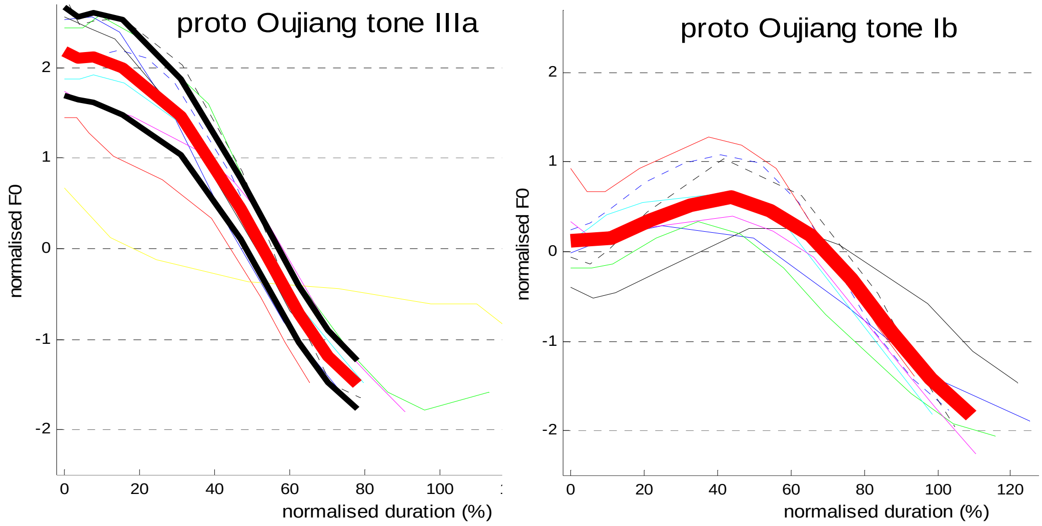

Figure 4: *Reconstruction of proto-Oujiang tones IIIa & Ib from acoustics of modern varieties.*

5. Acoustic reconstruction of proto-Oujiang tones

An important point, and one that does not seem to have been realised before, emerges from the demonstrated homogeneity of the normalised tonal acoustics in the Oujiang area. This is that they can be interpreted in a completely new light, namely as *historical* constructs. The normalised tonal F0 contours in Figure 3, together with a quantification of their variation, constitute the best estimate of the proto-Oujiang tones from which the modern forms derive. As an example, Figure 4 shows the acoustic reconstruction of two proto-Oujiang tones. On the left is the high falling proto-Oujiang tone IIIa, based on F0- and duration-normalised values from all varieties except Yongjia (which, it will be recalled, has a mid falling IIIa) and Wencheng (which has a mid level IIIa). The thick line shows the z-score normalised mean F0, which falls from about two standard deviations above the mean to about 1.5 sds below it. This is the best estimate of the F0 of proto-Oujiang IIIa and, together with the normalised duration quantifies what is meant by the reconstructed "high falling" for IIIa of table 1. The slightly thinner lines mark one standard deviation corridors above and below the reconstructed mean, and give a crude idea of the variation around it: something that is not really possible with auditory based reconstructions. A more sophisticated model of the variation must take into account the fact that values at the different sampling points are not independent, and give the probability of a particular set of values conditional upon a starting value for the tone, and the mean contour. A further sophistication would be to model the reconstructed F0 time-course with discrete cosine transforms or polynomials.

On the right of Figure 4 is the reconstructed mid falling proto-Oujiang tone Ib, based on F0- and duration-normalised values from all varieties except Wencheng, which had a lower mid level reflex for Ib (a standard deviation corridor is not shown). The reconstructed F0 shows that the tone is not straightforwardly mid falling, but has an initial, probably depressed, shoulder. (This is notated only in a few of the published sources, most of which describe the OJ Ib reflex as [31]). It also falls from higher than the middle of the range than implied by the synchronic acoustic descriptors.

Of course, one cannot just mechanically reconstruct across the board from mean acoustic values: this approach would not work with the OJ reflexes of IVa and IVb, which have been shown to differ considerably over the area (see Figure 2). There needs to be demonstrated homogeneity, and an indication that a more abstract reconstruction is not indicated, based on what we know about the way tones change (Brown 1985).

6. Summary and Envoi

This paper has described the citation tones in the small Wu Chinese sub-group of Oujiang in order to show how proto tones can be reconstructed from their modern acoustics. The key to this is a novel interpretation of the results of normalisation – the mean normalised values and the variation around the mean – as historical constructs.

The proto-Oujiang tones thus estimated may then be compared with similarly reconstructed cognates from other Wu dialect sub-groups for which there are modern acoustic data. It can be appreciated that the quantified aspect of this acoustic reconstruction approach, where each level of abstraction is explicit, and one operates with continuous variables, opens up the possibility of a whole new level of precision in historical tonological reconstruction and comparison.

Harold was the one to meet me at Canberra Airport in 1979 when I arrived to start my linguistics career in Australia. Then, the time-consuming process of extracting, measuring and processing tonal acoustics, (bent over the now antiquated analog spectrograph and spectrogram with their ever-present smell) meant the kind of work demonstrated in this paper, where acoustic data from 35 speakers have been used, was not possible. It is now. However, the ease with which we can now process data, and the new insights we win therefrom, should not blind us to the fact that computers will not supplant a prior, orthodox, linguistic analysis of the kind that Harold has always championed.

References

Ballard, W.L. (1969) *Phonological History of Wu*. Ph.D. thesis, UC Berkeley.

Brown, Marvin. (1985) *From Ancient Thai to Modern Dialects*. Bangkok: White Lotus Co.

CH (1964) CH 1964 汉语方言词汇 *Hanyu Fangyan Cihui* [Dictionary of Vocabulary in Chinese Dialects]. Peking: Wenzi Gaige Chubanshe.

Cao Z 曹志耘 (2005) 南部吴语语音研究 [Studies in the Phonetics of the Southern Wu dialects]. Peking: Shangwu Press.

Chao Yuen Ren (1928) 現代吳語的研究 / *Studies in the Modern Wu Dialects* Monograph 4. Peking: Tsing Hua College Research Institute.

Fu Guotong 傅国通, Fang Songxi, 方松熹 Cai Yongfei 蔡勇飞, Bao Shijie 鲍士杰, Fu Zuozhi 傅佐之(1985) 浙江吴语分区[Sub-groups of the Wu dialects of Zhejiang]. Zhoushan: Zhejiang Linguistics Society.

Maddieson, Ian (1978) 'Universals of Tone'. In J.H. Greenberg (ed.) *Universals of Human Language* vol. 2: Phonology, pp. 335-66. Stanford: Stanford University Press.

Norman, J. (1988) *Chinese*. Cambridge: CUP.

Rose, Phil (1987) 'Considerations in the normalisation of the fundamental frequency of linguistic tone'. *Speech Communication* 6/4: 343-352.

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

Rose, Phil (1993) 'A Linguistic Phonetic Acoustic Analysis of Shanghai Tones'. *Australian Journal of Linguistics* 13: 185-219.

Rose, Phil (2000) 'Hong Kong Cantonese Citation Tone Acoustics: A Linguistic-Tonetic Study'. In Michael Barlow (ed.) *Proc. 8th Australian International Speech Science and Technology Conference*, pp. 198-203 Canberra: Australian Speech Science and Technology Association.

Rose, Phil (2001) 'Wu', In J.Garry & C.Rubino (eds.) *Facts about the World's Languages*, pp. 158-161. New York and Dublin: New England Publishing Associates.

Rose, Phil (2002) 'Independent depressor and register effects in Wu dialect tonology: Evidence from Wenzhou tone sandhi'. *Journal of Chinese Linguistics* 30/1: 39-81.

Rose, Phil (2004) 'Zooming-in on Oujiang Wu: tonal homogeneity and acoustic reconstruction in a small sub-group of Chinese dialects'. In C. Watson and P. Warren (eds.) *Proc. 11th Australasian Intl. Conf. on Speech Science and Technology*, pp. 58-63. Auckland: Australasian Speech Science and Technology Association.

Wurm, S.A. T'sou, B. Bradley, D. Li R, Xiong Z. & Zhang Z. (eds.) (1987) *Language Atlas of China* Canberra: Longmans/Australian National University Research School of Pacific Studies.

ZH (1989) 汉语方音字汇 *Hanyu Fangyin Zihui* [A dictionary of pronunciations of characters in Chinese dialects]. 2nd ed. Peking: Wenzi Gaige Chubanshe.

Zhengzhang S. (1987) 'Wu Group'. In Wurm et al. (eds.) 1987 (no page nos.)

Zhu S. Xiaonong (1999) *Shanghai Tonetics*. Munich: Lincom.

Zhu S. & Rose, Phil (1998) 'Tonal complexity as a dialectal feature: 25 different citation tones from four Zhejiang Wu dialects'. In R. Mannell & J. Robert-Ribes (eds.) *Proc. 5th ICSLP,* Vol 3 pp. 919-922. Sydney: Australian Speech Science and Technology Association.