Citation	Phil ROSE. 2016. *Complexities of Tonal Realisation in a Right-Dominant Chinese Wu Dialect – Disyllabic Tone Sandhi in a Speaker from Wencheng*. *Journal of the Southeast Asian Linguistics Society* 9:48-80
URL	http://hdl.handle.net/1885/99985
Reviewed	Received 02/05/2015, revised text accepted 01/12/2015, published March 2016
Editors	Editor-In-Chief Dr Mark Alves | Managing Eds. Dr Sigrid Lew, Dr Paul Sidwell
Web	http://jseals.org
ISSN	1836-6821

www.jseals.org | Volume 9 | 2016 | Asia-Pacific Linguistics, ANU

COMPLEXITIES OF TONAL REALISATION IN A RIGHT-DOMINANT CHINESE WU DIALECT – DISYLLABIC TONE SANDHI IN A SPEAKER FROM WENCHENG

Phil Rose
Emeritus Faculty, Australian National University
<philjohn.rose@gmail.com>

Abstract
An acoustically-based description is given of the isolation tones and right-dominant tone sandhi in disyllabic words of a male speaker of the Chinese Oūjiāng 甌江 Wú 吳 dialect of Wénchéng 文成. His seven isolation tones show typical Wu complexity, comprising two mid-level, two rising, two falling-rising and one depressed level pitch shapes. Typical too is his three-way voicing contrast in syllable-Onset stops. However, the typical Wu relationship between tonal register and phonemic Onset voicing is shown to be disrupted, Onset voicing no longer correlating with tonal pitch height. The word-final tones in sandhi are shown to be straightforwardly related, phonologically and phonetically, to the isolation tones, with biuniqueness preserved. The realization of the word-initial tones in sandhi, on the other hand, involves complex mergers conditioned by largely non-phonetic factors related to historical tone categories, resulting in five extra sandhi tones that do not occur in isolation. It is suggested that they not be related phonologically to the isolation tones. The historical implications of the patterns are also briefly explored.
Keywords: Right-dominant tone sandhi, Wu dialects
ISO 639-3 codes: wuu

1. Introduction

1.1 Wu tone sandhi typology and aims of paper
Through the coastal provinces of S.E. China (Southern Jiangsu, Zhejiang, Fujian) runs a ‘sandhi zone’ (Norman 1988: 202). Varieties in this area, which belong mostly to the Wu and Min dialects of Chinese, exhibit highly complex morphotonemics (Chen 2000). One of the intriguing aspects of the morphotonemics found in the Wu dialects of the sandhi zone, in Southern Jiangsu and Zhejiang, is the typological distinction between so-called *left*- and *right-dominant* (or *-focused*) lexical tone sandhi first pointed out in Ballard (1984) and analysed in depth in Zhang (2007).

In left-dominant sandhi, it is largely the input tone on the leftmost syllable morpheme of a word which determines its sandhi shape. All tones in a polysyllabic word except that associated with the initial syllable tend to be lost, and the initial syllable tone is then realised, or spread, over the whole word, resulting in wholesale neutralisation of morphotonemic contrasts. Paradigm examples are the much-discussed Shanghai (Sherard 1980, Zee & Maddieson 1979, Zhu 2004) and the dialect for which the phenomenon was first demonstrated: Tángxī 塘溪 (Kennedy 1953).

In right-dominant varieties, it is the tones on the morphemes on the rightmost syllables of a word which determine the sandhi shape. The tone on the word-final syllable is said to be ‘preserved’, ‘unchanged’ or ‘in agreement with' the citation tone, and tonal contrasts on the preceding syllables tend to be neutralised, although not by spreading: the neutralisation groupings are often complicated.

The typological distinction between left- and right-dominant systems, which is primarily conceived as applying to words, but has also been extended to syntactic constructions (Zhang 2007), appears to be geographically distributed (Pan 1991: 287): dialects in the north of the Wu area tend to be left-dominant; southern varieties (including Min dialects) right. There may also be some intermediate varieties. Zhènhǎi 鎮海, for example (Rose 1990), appears in retrospect to partake of both types, with two tones involving left dominance, two right, and two both.

Given the acknowledged importance for phonological theory of empirical quantitative studies on the realisation of tones in sandhi (Chen 2000: 27-28, Zhang 2007: 284, 288, 296; 2009: 11-12), it is surprising that very little acoustic description and analysis exists on putative right-dominant varieties. I know of three, all from the Chuqu subgroup of Wu: Lóngquán 龙泉 Lìshuǐ 丽 and Qīngtián 青田 Steed (2006, 2010, 2012 respectively). The main aim of this paper, therefore, is to add to the documentation of the right-dominant types by describing, in sufficient detail for phonetically grounded phonological analysis – that is, by trying to account for all systematic acoustic observations – disyllabic lexical sandhi in a speaker of a variety from another Wu subgroup that appears to be right-dominant: Wénchéng 文成.

From the south-western part of the south-eastern Oujiang 甌江 subgroup of Wu, Wencheng is additionally interesting from the complexity of its isolation tone system (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, which may represent either a free or bound morpheme, is read out.) Five of the speaker's seven isolation tones have contour pitch, three of which are complex, and four of which involve rising pitch (Rose 2010). It is of obvious interest to show how such complex tones interact in sandhi and the description will therefore aim to specifically investigate the nature of the relationship between isolation tones and tones in disyllabic words. I will focus on two aspects of the relationship described as criterial for right-dominant sandhi. Firstly, the extent to which the isolation tone values are preserved on word-final syllables, and how one might demonstrate this; and secondly the nature of the apparent changes to tones on word-initial syllables. For the reader who wants the bare minimum without the phonetic argumentation, this may be distilled from table 1 and figure 3 (isolation tones); table 5 and figure 10 (sandhi on disyllabic words), and table 6 and figure 12 (first-syllable sandhi changes).

Although the primary aim of the paper is description, I have also made some phonological observations where I think they are warranted. I have been deliberately derivational, with a default assumption, common in Chinese Linguistics, that the isolation form is basic; and I have pointed out where this approach is found wanting. It would be nice to see how well the data yield to a constraint-based approach. One role of complex data, after all, is to better test models.

1.1 Data

The data are from a recording which was part of a wider survey of tones and tone sandhi in ca. 30 southern Wu varieties conducted by Professor W. Ballard in 1988 (Ballard 1992: 41-42), and his generosity in making the data available is gratefully acknowledged here. The recording consists of three replicates each of ca. 40 monosyllabic and 190 disyllabic utterances elicited from a then 21 year old male speaker. The informant, selected for Ballard by his Chinese Wu dialectologist colleagues, was born in Wencheng city (大嶨鎮) and lived there until the age of 19. I have also made use of the several pages of Ballard's supplementary phonetic transcriptions of body-part names from the same informant. The digitised recordings can be listened to on the author's web-page (http://philjohnrose.net), where their individual and mean acoustics are also plotted. A digitised version of the complete cassette recording is available for further analysis (vowel acoustics, perhaps?) upon request from the author.

1.2 General procedure

As expected from a Wu variety, the speaker's morphotonemics are very complex, involving considerable variation in tonal properties of pitch and duration, as well as related variation in segmental properties. For example, the free morpheme {*skin, cover* 皮} has at least four different rule-governed pitch realisations. In isolation [**b̥ɛɪ 22**] and in the word *stomach* 肚皮[d̥œy**bɛɪ** 342.**22**], for example, the morpheme is realised with a lower-mid level pitch; in *pigskin* 豬皮[tɕi**bɛɪ** 55.**44**] it is realised with an upper-mid level pitch; in *tape measure* 皮尺[**b̥ɛɪ**tɕʰi **342**.23] its pitch is depressed high falling, and in *football* 皮球 [**b̥ɛɪ**dʑaʊ **11**.22] it is low level. The free morpheme {white 白} provides examples of morphotonemic variation in both duration and pitch. On its own it has a long low falling-rising pitch: [**b̥a**˥ **214**] *white*; in *understand* [mɛŋ**ba** 342.**21**]明白 it has a low falling pitch of average length, and in *medicinal powder* 白藥 [**b̥ǎ**ja 2.21] its pitch is lower-mid and too short to reliably specify a contour. As can be seen, the Onset bilabial stop also varies in these three examples, between voiceless lenis word-initially and voiced word-internally.

It would be difficult to make sense of much of this surface tonal variation without morphological information indicating which phonic strings were realisations of the same morpheme. With very few exceptions, Chinese characters identify morphemes, and in addition one generally has a very good idea, from both diachronic and synchronic linguistic information, what tonal category a particular morpheme is supposed to belong to. I made judicious use of this information wherever possible, the general procedure being to first segmentally and tonally transcribe the recordings – with respect to the primary auditory tonal properties of pitch, length and phonation type – and then quantify the tones acoustically on the basis of the transcription and the morphological information encoded in the Chinese characters with which the data were elicited.

1.3 Terminology

Because the description is both impressionistic and acoustic, and provides plenty of examples where fundamental frequency does not reflect tonal pitch, it is best not to follow current practice, where *pitch* is often used as a synonym for *F0*. Instead, the referents of both terms are clearly distinguished: *pitch* as the main perceptual correlate of the linguistic category of tone, and *F0* as its main acoustic correlate (cf. Rose 1989a). The acoustic quantification is described in appropriate detail in the sections dealing with acoustic properties.

Another terminological problem occurs with the word *tone*, which can be found in the literature in many different senses, for example *toneme*, *allotone*, *autosegmental tone*, *morphotoneme*, *tonal allomorph*, *pitch*, *F0*, *linguistic/phonological/phonetic category*. The complex nature of the relationship between the Wencheng speaker's isolation and tone-sandhi forms means that sometimes a statement like *a high rising tone becomes a falling tone before an upper level tone* is capable of rather different construal depending on whether underlying phonological constructs (morphotonemes, tonemes) are meant by *tone*, or their realisations. I have been specific, using the term *morphotoneme*, or the expression *tone on a given morpheme*, where there seemed danger of confusion.

When describing tonal allomorphy in disyllabic words, it is likely that *word-final* is the correct environmental specification for tones on their second syllables, and this is therefore the term I have used. However, it is not clear what the correct environmental specification is for the first-syllable forms: in most cases it is likely to be *penultimate-*, or *pre-final syllable*. This is because the pitches that occur towards the onset of words longer than two syllables in Ballard's supplementary transcriptions are often simpler than those for the disyllabic words described here. However, it is also possible that in some cases, depending on the tones involved, *word-initial* may be the correct structural description. Because of this indeterminacy, I have used the term *first syllable* as the environmental specification.

1.4 Structure of Paper

As stated above, this paper deals with the relationship between isolation tones and tones in disyllabic words. For describing synchronic morphotonemics of this kind, a structural *item and arrangement* model (Matthews 1970) is probably best. Thus I will describe the allomorphic tonal realisations in three different environments. The isolation tones are handled first, in section 2. As well as a description of the individual tones, this section contains a demonstration, with the Wencheng tones, of the limitations of the Chao 5 point system of tonal description. Because of the tight relationship between isolation tones and syllable Onset voicing, itself reflecting the way phonological information is packaged into the Wencheng morpheme, it is also necessary to discuss these topics in this section.

Tones in the disyllabic word are described in sections 3 through 5. Section 3 gives the essential overview necessary to this complex topic, section 4 describes word-final tones, and section 5 describes first-syllable tones.

2. Isolation Tones

2.1 Overview

The speaker has seven isolation tones, exemplified in table 1. The **upper-mid level tone** has a level pitch contour just above the middle of the speaker's pitch range. The **depressed upper-mid level tone** has the same level pitch as the upper-mid level tone, but its onset is depressed so that pitch over the initial third to half of duration is rising. As this is quite a rare tonal pitch contour, it warrants acoustic illustration. Figure 1 shows the typical F0 time course for the depressed upper-mid level tone on three syllables with Onsets differing primarily in VOT and manner: [b̥ɛɪ] *blanket* 被, [mo] *horse* 馬 and [z̥o] *to sit* 坐. The F0 is superimposed on a wideband spectrogram to show the relationship between the segments and the F0. The initial F0 rise of between 20 and 30 Hz over about the first 15 csec. of the Rhyme is clear, after which the F0 remains reasonably stable until the falling offset perturbation over the last five centiseconds or so. Note that this lower F0 onset and rise is clearly not a perturbatory function of any voicing on the Onset consonant, since there is none in [b̥] and [z̥], both being voiceless with coincident VOT.

Table 1: *Examples of isolation tones.*
**MC = Middle Chinese tonal category*

tone name	*MC					
upper-mid level	Ia	Yīnpíng	[kɔɐ	*close* 關	fɛɪ	*fly* 飛
depressed upper-mid level	IIb	Yángshǎng	z̥o	*sit* 坐	b̥ɛɪ	*blanket* 被
lower-mid level	Ib IIIa	Yángpíng Yīnqù	t^{h}ɔɐ	*jump* 跳	b̥ɛŋ	*flat* 平
high rise	IIa	Yīnshǎng	koɐ	*vast* 廣	hɛ	*good* 好
low rise	IVa	Yīnrù	foɐ	*put forth* 發	pe	*north* 北
mid fall-rise	IIIb	Yángqù	joɐ	*use* 用	b̥ɛŋ	*ill* 病
low fall-rise	IVb	Yángrù	o	*study* 學	z̥ɛɪ]	*stone* 石

The **lower-mid level** tone has level pitch a little lower than the upper-mid level tone. 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. The **low rising** tone has pitch which rises from low in the speaker's pitch range to mid. The **low fall-rise** tone has a dipping contour, first falling within the speaker's low pitch range to their lowest pitch, and then rising into the mid pitch range. As might be expected from this complex contour, length is above average. The **mid fall-rise** 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 as low as the low fall-rise tone. Its length sounds well above average. This tone has quite a complex pitch contour and its typical F0 time course, superimposed on a wideband spectrogram, is shown in Figure 2 for two syllables differing in Onset: [d̥ɪɛ] *electricity* 電, and [mɪɛ] *cheek* 面. The two portions of the Rhyme with first (level-) falling and then rising F0 are clear, with the rising portion being a little shorter than the falling, and being followed by the same type of abrupt falling offset perturbation seen in the depressed upper-mid level tone examples in figure 1. The difference between the two syllables in the F0 time course over ca. the first 20 csec. of the Rhyme is free variation and not a function of the difference in Onset consonants.

The speaker's seven isolation tones thus comprise two level-pitched tones: one upper-mid and one lower-mid. Of its five contour tones, two are simple rises: a short high and a low; and three are complex: a depressed upper-mid, a mid fall-rise and a low fall-rise[1]. As is shown in table 1, these seven different

[1] These isolation values differ somewhat from three previous impressionistic descriptions (Fu et al. (1985: 110), Cao (2002: facing page 100) and Zhu (1996: 243), which also differ among themselves. They notate, respectively, Ia = [44] [445] [55]; Ib = [11] [113] [13]; IIa = [45] [454] [54]; IIb = [33] [324] [33]; IIIa = [434] [334] [435]; IIIb = [312] [313] [312]; IVa = [23] [24] [35]; and IVb = [12] [213] [213]. Cao's (2002) description seems to have incorporated some rising and falling offset pitch perturbations in some tones. The main difference from the data in the present paper is that none of these sources show a merger between reflexes of *IIIa and *Ib, and have rather different, non-depressed reflexes for *IIb. Unfortunately these sources provide no acoustic, background or acquisition data so it is not possible to interpret the between-description variation, as for example age- or location-related.

tonal pitch shapes are the reflexes of eight historical Middle Chinese tonal categories, two of which – *Ib and *IIIa – have merged to the lower-mid level tone. This diachronic fact needs to be mentioned in the middle of an essentially synchronic description for the interesting reason that the two merged historical tones (*Ib *IIIa) are clearly distinguished in the synchronic tone sandhi because, as will be shown below, they evince different tonal behaviour on the preceding syllable. This synchronicity often necessitates indicating, clumsily I am afraid, the two different origins of the lower-mid level tone with “<IIIa” or “<Ib”.

Figure 1. *F0 of* ***depressed upper-mid level*** *tone on three syllables with different Onsets: from left to right, [b̥ɛɪ] [mo] [z̥o]. Vertical axis = frequency (Hz) with F0 values at right, and spectrogram values at left. Horizontal axis = duration (csec.).*

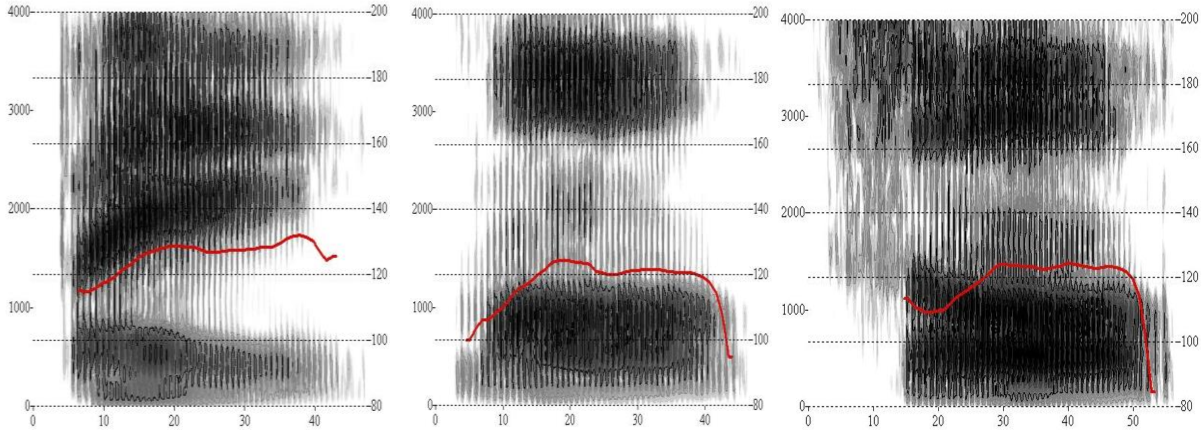
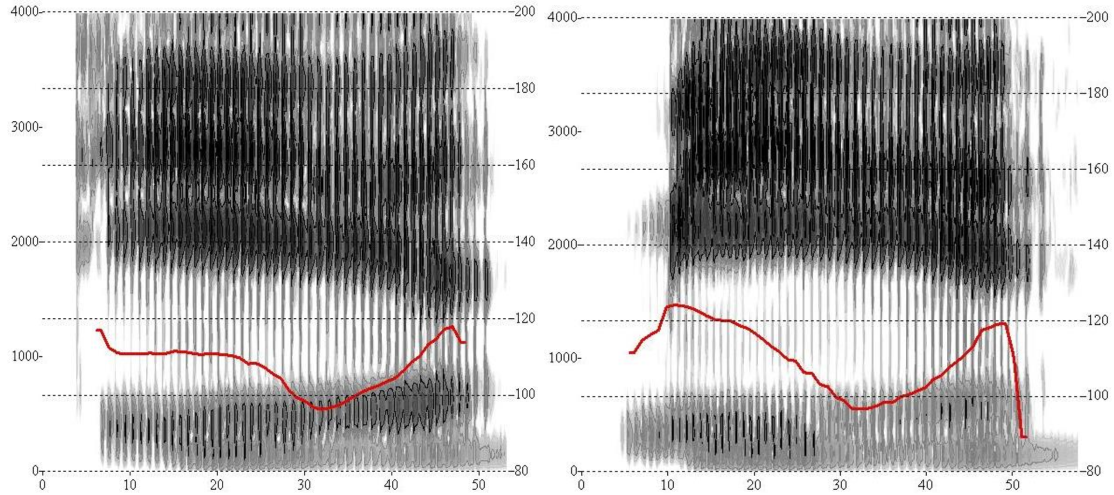


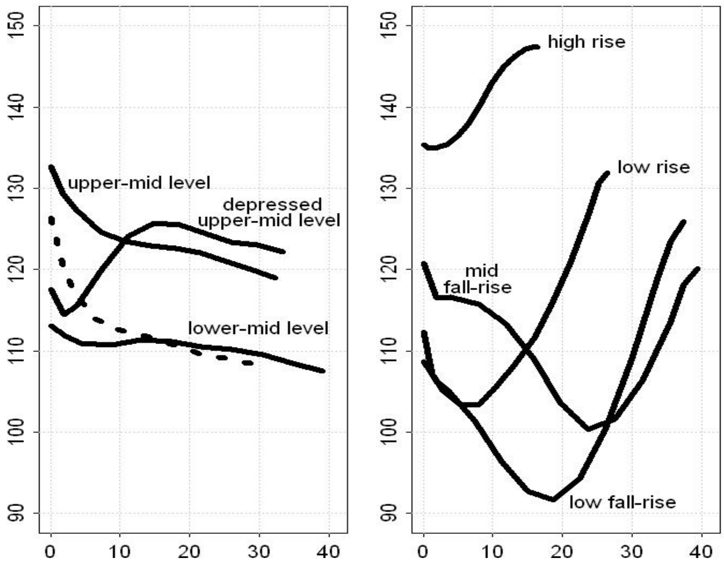
Figure 2. F0 of **mid fall-rise** tone on syllables with different Onsets: [d̥ɪɛ] (left), [mɪɛ] (right). Vertical axis = frequency (Hz) with F0 values at right, and spectrogram values at left. Horizontal axis = duration (csec.).



2.2 Mean tonal acoustics

Mean tonal acoustic values (F0 plotted as a function of absolute duration) for the seven isolation tones are shown in figure 3. Means were calculated over five replicates per tone, except in the case of the lower-mid level tone where five replicates were measured separately for reflexes of both its historical categories, *Ib and *IIIa. In order to measure each token, a wideband spectrogram was generated in *Praat*, together with its wave-form and superimposed F0. The token’s tonally relevant F0 was then identified, extracted and modelled in *R* by an eighth order polynomial. This enabled F0 values to be sampled from the polynomial F0 curve with a sufficiently high sampling frequency (usually at 10% points of the curve as well as 5% and 95%) to capture the details of its time-course. (When the tonal pitch contours are as complex as this, it is necessary to have a relatively high sampling rate.) Following work on the phonetic domain of tone (Howie 1974, Rose 1982b, 1998), the onset of the tonally relevant F0 was taken to be the onset of the Rhyme (adjudged as its first strong glottal pulse). Thus F0 on syllable-initial sonorants, like those shown in figures 1 and 2, was excluded. The offset of the tonally relevant F0 was taken to be at the start of the offset perturbation. In tones with rising pitch, such as those shown in figure 2, the start of the offset perturbation is easily determined with reference to the peak F0 value. However for level pitched tones, such as those illustrated in figure 1, the end of the tonally-relevant F0 is not clearly marked with an acoustic discontinuity such as a change in the sign of the F0 derivative: instead the F0 derivative is gradual, and a different strategy had to be used. In these tones, then, F0 was first extracted to the point at which discontinuities in wave-form period or peak-to-peak amplitude, or both, occur. As can be appreciated from figure 1, this will include the falling offset perturbation. The tonally relevant F0 portion was then taken to end at the 90% point of the extracted polynomial. It can be seen from figure 3 that this removes the offset perturbation.

Figure 3. *Mean F0 (Hz) of the seven isolation tones plotted against mean duration (csec.). (Dotted line = reflex of historical *IIIa.)*



The isolation tonal F0 shapes in figure 3 have been plotted in two panels, as their complex configuration would have made it difficult to identify their shapes otherwise. Upper- and lower-mid level and depressed upper-mid level tones are shown in the left panel, where two F0 shapes are shown for the lower-mid level tone corresponding to the two historical reflexes that have merged. The remaining contour tones are in the right panel. From figure 3 it can be seen that the F0 shapes correspond well to their impressionistic descriptions. The speaker's isolation tonal F0 range, from the high rising tone peak to the low fall-rising tone trough, is about 55 Hz. The upper and lower-mid level tones, and the depressed onset tone are located in the middle third of the speaker's F0 range, above and below the mid-range point.

Although it is possible to get rid of some non-tonally relevant acoustics, for example the F0 offset perturbations, and the F0 on syllable-initial sonorants, it is not the case that the residue will be free from non-tonal effects: tones can never be acoustically observed independent of their segmental realisation, due to the well-known assumed intrinsic effects from concomitant segmental articulation, like intrinsic vowel and consonantal F0. The isolation tone corpus was actually well controlled for intrinsic vowel F0, but consonantal effects can still be observed in the tonal F0 in figure 3. For example, the abruptly falling F0 over about the first five centiseconds of the upper-mid level tone and the *IIIa reflex of the lower-mid level tone probably reflect [+spread] vocal cord Onset consonants like voiceless fricatives and aspirated stops; and shorter perturbatory effects can be seen in some other tones. Because of co-occurrence relationships between tone and syllable-initial consonants, to be described below, it is not possible to control completely for such effects, and therefore it is best to understand figure 3 as showing the tonal acoustics plus the effect of the syllable-initial consonant on the Rhyme.

2.3 Auditory phonetic representation

An appropriately detailed impressionistic description of tones – one that describes the forms accurately enough to make clear their linguistic conditioning – obviously needs a way of representing tonal pitch sufficiently fine-grained to enable a so-called systematic allophonic transcription (Abercrombie 1967: 127-

129, Laver 1994: 549-551). The IPA sanctions two methods of pitch notation (IPA Guide 1999: 14) one of which, based on Chao's (1930) five-point iconic 'tone letters', is also the preferred method for impressionistic-phonetic transcription of the tones of Chinese varieties. It has been demonstrated, however, that the five-point scale does not model the pitch percept equally well for all tones. Figure 4, after Rose (2014), shows a transformation of the mean tonal F0 of figure 3 taking into account two factors relevant to the perception of linguistic pitch: the appropriate scale, and F0 declination. Since it has been shown that linguistic pitch perception involves taking into account F0 declination (Lieberman 1967: 53:61, Pierrehumbert 1979), the mild declination observable in the F0 of the tones with level pitch components (upper-mid and lower-mid level, and depressed upper-mid level) was modelled with linear regression and the F0 of all tones adjusted by the mean slope (-0.198 Hz/csec.). Secondly, since the semitone scale has been shown to give the best approximation for perception of intonational pitch (Nolan 2003), the resulting F0 values were then transformed into semitones relative to the lowest tonal F0 value (the trough in the low fall-rise tone).

It can be seen that the transformed F0 range covers about eight semitones, which can be nicely mapped onto the five points of the Chao scale (shown on the right of each panel in brackets). It can be also seen that the level tones now have effectively level pitch, and the slope of the contour tones is also not as steep as with their F0. It can be appreciated from figure 4 that most of the contour tones in the right panel can be rather well represented by Chao tone letters. Thus the high and low rising tones may be non-controversially represented as [45] and [24] respectively, and the low fall-rise tone is [214]. The mid fall-rise tone, however, although its onset and turning points are clearly [3] and [2], has an offset intermediate between [4] and [3], and representing it as either [4] or [3] will therefore be procrustean.

As far as the level tones in the left panel of figure 4 are concerned, it can be seen that they do indeed fall either side of the mid pitch range point, but that their values are neither high enough nor low enough to be represented as [44] or [22]. In fact their pitch values look about equally probable, given a representation of either [44] or [33] (for the upper-mid level tone), or [33] or [22] for the lower-mid (Rose (2014) used a simple Gaussian Mixture Model to show that the likelihoods of their tonal pitch targets given either representation are equally minimal). It is also clear that the height of the upper-mid level tone is the same as the offset of the mid fall-rise tone, and that this, at least for this speaker, represents a separate tonal pitch target in their range. The same applies to the pitch of the lower-mid tone and the onset of the depressed upper-mid tone.

Although the five-point scale is demonstrably procrustean for some tones, it obviously fits rather well for others, and I will make use of it in subsequent description, with the caveat that [4] and [2] in the upper- and lower-mid level tones [44] and [22], the depressed upper-mid level tone [244], and the offset of mid fall-rise tone [324] may represent upper and lower-mid pitches which are not as high or low as their integers imply. Alternatively, it would also be possible to make use of the upstep and downstep notation sanctioned by the IPA (IPA Guide 1999: 15), notating the upper-mid level tone as [↑33] and the lower-mid level tone as [↓33]. But this would be contrary to the definition of "modifications (raising or lowering) of the pitch indicated by ordinary tone symbols", and in any case it will be necessary to make proper use of the step notation when transcribing sandhi in disyllabic words.

Figure 4. *Perceptual transform of isolation tone F0. Vertical scale at left of panels = semitones, (vertical scale at right of panels = Chao tone letters, horizontal scale = duration (csec.).)*

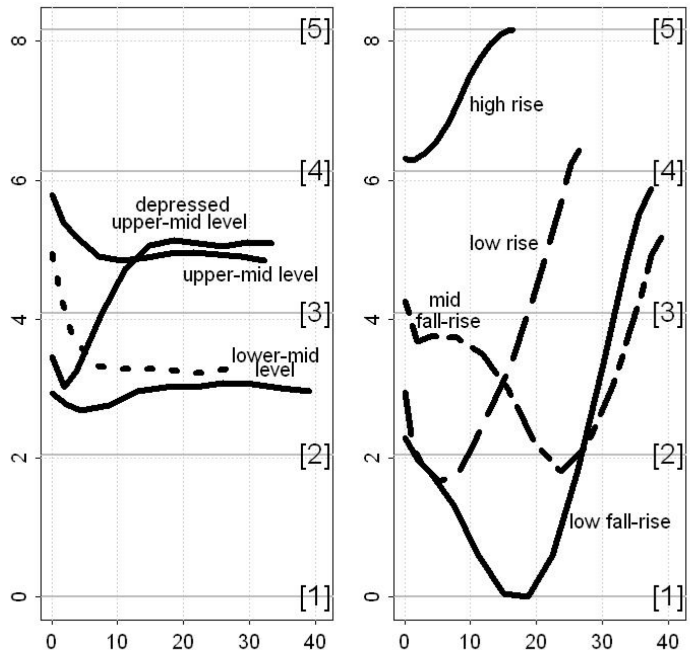


Table 3 exemplifies the obstruent types word-internally on alveolars and alveolo-palatals. Figure 5 shows the acoustics of the alveolar stop examples. It shows the three word-internal alveolar stops differing, unremarkably, in VOT, hold phase duration, and length of preceding vowel. Both [t] and [tʰ] have lag VOT values – a short 1.3 csec for [t], and a much longer 6.8 csec. for [tʰ] – whereas [d] VOT is lead with the phonation, whence the intrinsic drop in F0 in the hold phase. The hold duration for [t] (9.0 csec.) is similar to [tʰ] (7.9 csec.) and both are greater than in [d] (4.9 csec.).

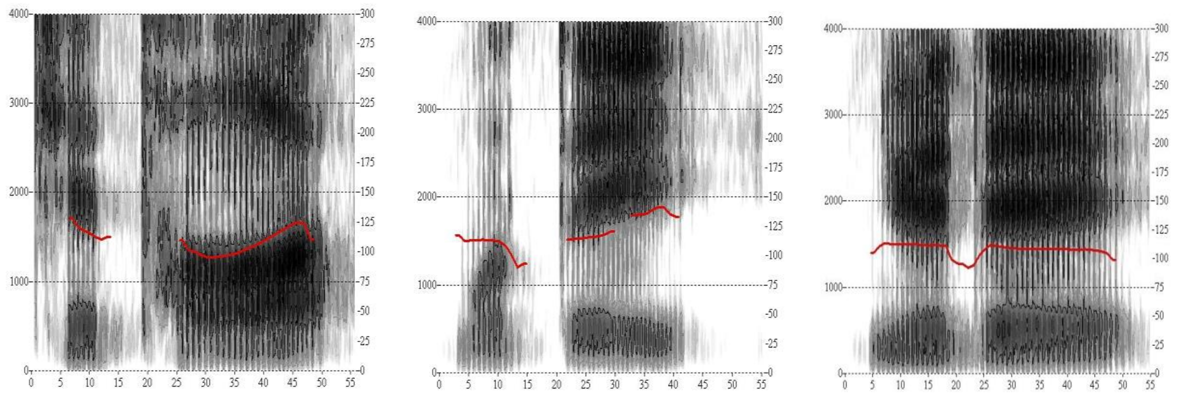
Table 2*. Co-occurrence relationships between tone and Onset obstruents.*

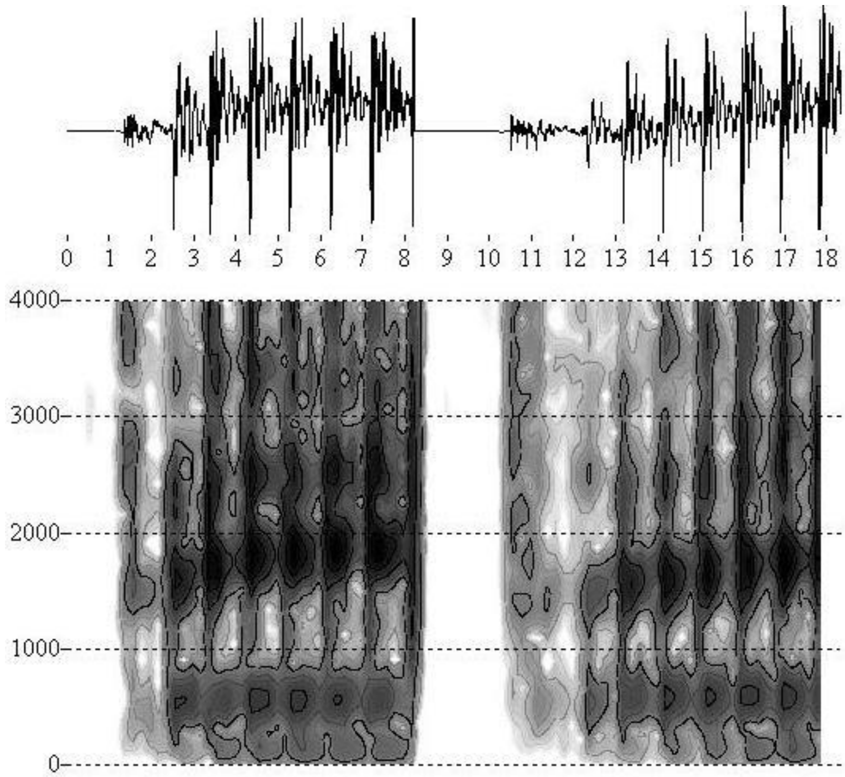
obstruent	voiceless aspirated & unaspirated stops & affricates; voiceless fricatives	voiced word-internally, voiceless lenis word-initially
tone	upper-mid level [44]	depressed upper-mid level [244]
	high rising [45]	mid fall-rise [324]
	low rising [24]	low fall-rise [214]
	lower-mid level [22]	

Table 3. *Examples of contrasting VOT types in obstruents.*
Underlining indicates the pitch shape, and associated vowel, is of short duration.

[mo **t**ɛɪ 2.35	*sole of socks* 袜底	b̥ɪ **tɕ**ɪŋ 2.44	*safety pin* 别针
ɲɥɪ **d**e 2.22	*platform* 月台	tɕɪ **dʑ**ʲaŋ 2.44	*adjacent* 接近
tʰɪ **tʰ**ɑ 3.24	*pylon* 铁塔	ɲɪ **tɕʰ**i 3.22	*heat* 热气
kʊŋ **s**ɛɪ 52.12	*formula* 公式	sʊŋ **ɕ**ʲɛ 55.↑44	*rosin* 松香
ləʉ **z**ɛɪ 342.44	*guesthouse* 旅舍	b̥a **ɲ**ʲɛ 2.344]	*white elephant* 白象

Figure 5. *Spectrograms, with superimposed F0, illustrating the three-way intervocalic VOT contrast in stops. Left = [tʰǐ tʰa], middle = [mǒ tɛɪ], right = [ɲɥǐ dɛ]. Horizontal axis = duration (csec.), vertical axis = frequency (Hz).*



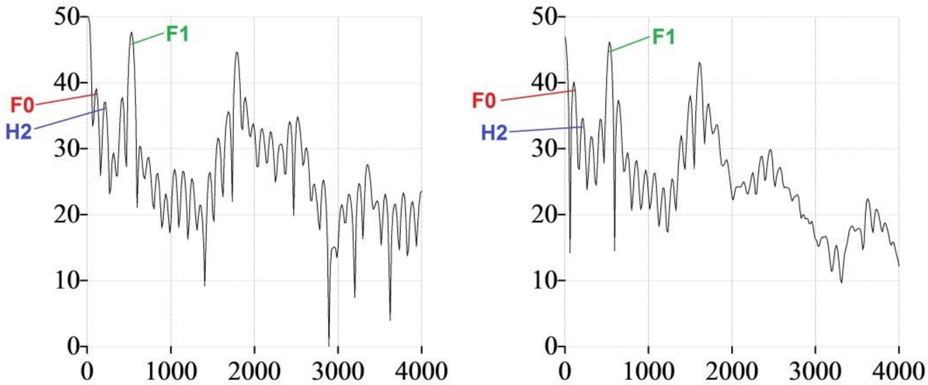


The word-internal voiced obstruents alternate with voiceless lenis realisations word-initially. The lenis nature may be related acoustically to a slightly *longer* VOT in lenis stops (i.e. [b̥ d̥] etc.) and a *shorter* duration in lenis fricatives (i.e. [ɣ̥ z̥] etc.), both differences being statistically significant. To illustrate this, figure 6 shows an example of the first few centiseconds of two morphemes - [pe 24] *north* 北 and [b̥ɛŋ 324] *ill* 病, which have been chosen to maximally control effect of vowel height and initial tonal pitch on VOT. Both have voiceless, coincident VOT syllable-initial stops: the former is an audibly fortis [p], the latter an audibly lenis [b̥]. It can be seen that the two stops differ in VOT, the [p] having a VOT of just over a centisecond, and the [b̥] a slightly greater VOT of just under 2 centiseconds[2]. The peak-to-peak amplitude of

[2] The difference between fortis and lenis stop VOT was tested with a two-way ANOVA with bilabial and alveolar stops in the data ([b̥] *vs* [p]; [d̥] *vs* [t]). This showed the fortis stops to have a little less than a centisecond shorter VOT than the lenis stops. Mean fortis and lenis VOT values were: 1.5 csec. & 2.2 csec. (bilabials); and 1.9 csec. & 2.8 csec. (alveolars). The fortis-lenis difference in VOT is very highly significant (F = 14.6, DF =1, *p*=0.0006). An analogously small but significant VOT difference was also demonstrated for fortis *vs* lenis stops in the Northern Wu dialect of Zhenhai (Rose 1982a: 262, 263 *et pass*.).

the first glottal pulse is also noticeably greater for [p]. The corresponding spectra taken over the first 5 centiseconds from the first glottal pulse show a small difference in F0-H2 and F0-F1, with the difference after the lenis stop about 2 - 3 dB greater. I hear no regular difference in phonation type on the following vowel, however.

Figure 6. *Illustrative acoustic characteristics of word-initial [p] and [b̥]. Top two rows show aligned wave-form and wide-band spectrogram of first few centiseconds of [pe 24] north and [b̥ɛŋ 324] ill. Horizontal axis = duration (csec.). Bottom row shows corresponding spectra over first 5 csec. after first glottal pulse. F0, H2 = fundamental and second harmonic, F1 = harmonic with greatest amplitude under first formant transfer function. Axes = spectral amplitude (dB) and frequency (Hz).*



The corresponding difference for fricatives is shown in figure 7, which illustrates typical acoustic differences between the word-initial alveolar fricative consonants [s] and [z̥] occurring in the monosyllabic words [saŋ 22] *letter* 信 and [z̥a 214] *ten* 十. This pair was also chosen to control for following vowel quality, and tonal pitch onset value (the F0 at onset is about 14 Hz higher for the token with [s]). The first five glottal pulses of the following vowel are also shown. It can be seen that both [s] and [z̥] are voiceless, but [s] has about twice the duration of the [z̥][3] and also a greater peak-to-peak amplitude. The presence of the 4 kHz pole in the [s] does not correlate with the fortis/lenis difference. As in the stop examples, the first glottal pulse is weaker for the lenis, and the spectra for the first 5 csec. after the first glottal pulse also show it to have a stronger fundamental relative to both H2 and the harmonic under the F1. Again, I hear no difference in phonation type on the following vowel.

2.5 Atypical relationship between tone and phonemic obstruent voicing

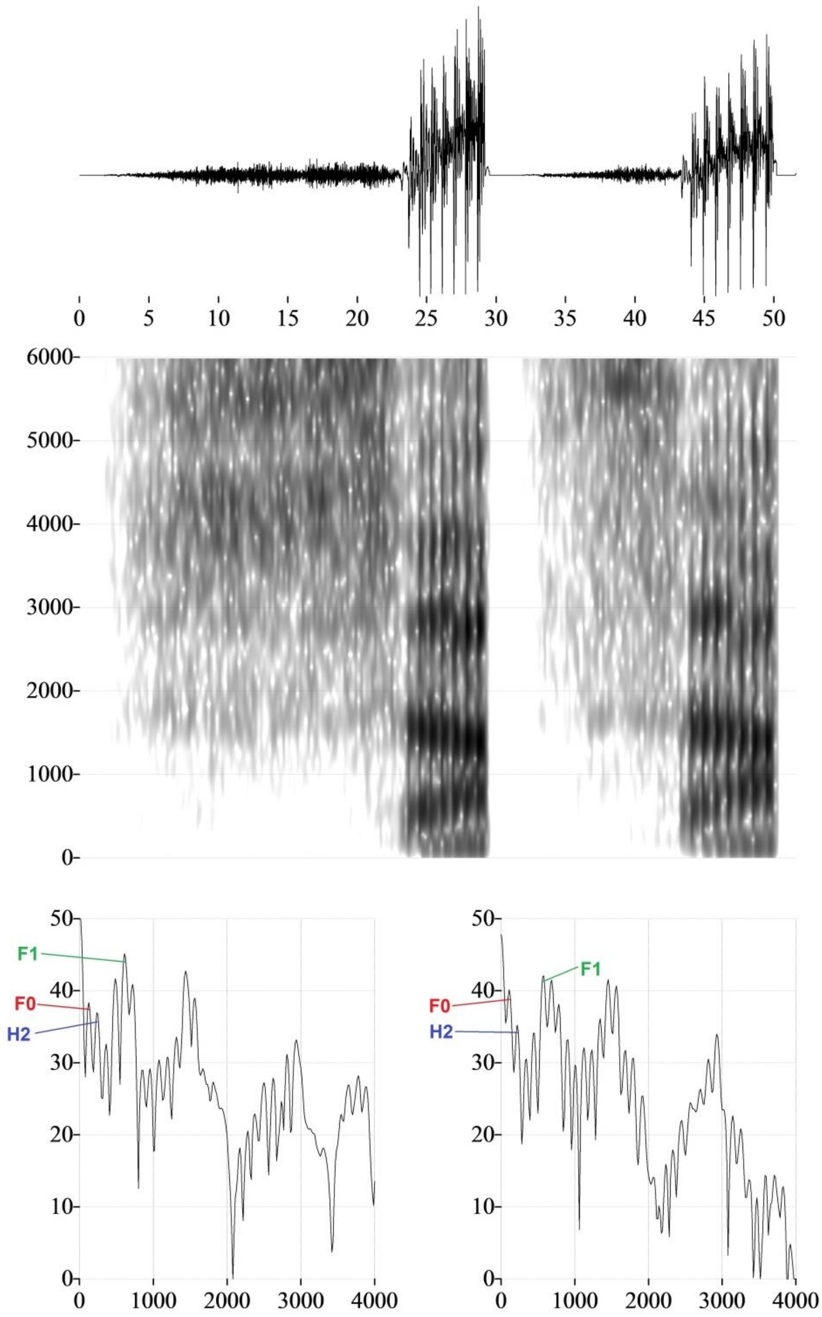
The phonemic interpretation of the variation in stop and fricative realisation is clear from the contrastive and complementary distributions involved. A three-way VOT-based contrast can be established for stops - /aspirated/, /voiceless unaspirated/ and /voiced/ - the first two realised by aspirated and voiceless unaspirated allophones respectively, and the third by voiced allophones word-internally and voiceless lenis allophones word-initially. The same applies *mutatis mutandis* for fricatives: a two-way contrast between /voiceless/ and /voiced/, with the latter realised by voiced and voiceless lenis allophones.

A systematic three-way contrast in syllable-initial stops and affricates is actually the defining typological diagnostic for Wu dialects as a first-order sub-group of Sinitic (Chao 1967: 92, 94), and in this respect the data is therefore typically Wu. It deviates markedly from the canonical Wu pattern, however, in the relationship between these obstruents and the actual tonal pitch values. In many Wu dialects, including Wencheng's Oujiang subgroup neighbours, phonemically voiced obstruents co-occur with low-pitch isolation tones, and phonemically voiceless obstruents occur on isolation tones with high pitch onset and high register. It can be seen from table 2 that this intimate and commonly found connection between tonal pitch and syllable-initial obstruent voicing has been weakened, where tonal pitch cannot be predicted from

[3] The difference between fortis and lenis fricative duration was t-tested with alveolar fricatives in the data ([z̥] *vs* [s]). The mean duration of the data's 33 word-initial [s] tokens was 13.3 csec. (sd = 2.5), and its 39 word-initial [z̥] tokens were on average ca. 4 csec shorter at 9.4 csec. (sd = 1.2). These values mean that a t-test will show the difference between the means to have a *p* value considerably less that 0.0001. An analogous durational difference between fortis and lenis fricatives was also demonstrated in the Northern Wu dialect of Zhenhai (Rose 1982a: 262, 263 *et pass.*).

phonemic obstruent voicing. Only two out of four of the tones that co-occur with voiceless obstruent phonemes can be considered high register: [44] and [45]; and only two of the tones co-occurring with voiced obstruent phonemes – [22] and [213] – can be considered low.

Figure 7. *Illustrative acoustic characteristics of word-initial [s] and [z̥]. Top two rows show aligned wave-form and wide-band spectrogram of first few centiseconds of /saŋ/ letter and /za/ ten. Horizontal axis = duration (csec.). Bottom row shows corresponding spectra over first 5 csec. after first glottal pulse. F0, H2 = fundamental and second harmonic, F1 = harmonic with greatest amplitude under first formant transfer function. Axes = spectral amplitude (dB) and frequency (Hz).*



2.6 Phonological classification of morphemes

The discussion above has shown that phonemic Onset obstruent voicing partitions the tones into two explicit groups: a /voiceless obstruent/ group ([44] [45] [24]) and a /voiced obstruent/ group ([244], [324], [214]), with the lower-mid level [22] tone belonging to both classes. This arrangement was implicit in table 2. These two groups of tones constitute unnatural classes by virtue of the fact that they cannot be defined by any obvious pitch features, but, as will be shown below, the grouping is also strongly reinforced by tone sandhi behaviour. The lower-mid level [22] tone, for example, behaves very differently in sandhi depending on its Onset obstruent properties. The useful coinage *unnatural class* referring to a group of tones that behaves as a natural class (i.e. as input, or conditioning to a rule), but with no phonetic basis – a commonplace in Wu – is from Steed (2010).

Phonatory activity also partitions the tones into the same two groups in the case of syllables beginning with a sonorant (i.e. sonorant consonant, glide, vowel). In these syllables, phonation onsets word-initially with a weak glottal-stop in the case of isolation tones [44], [45] and [24] and their sandhi variants. For example [ˀø 44] *peace* 安, [ˀɥyə 45] *wrist* 腕. In the case of isolation tones [244], [324], [214], phonation onset is non-abrupt, e.g. [maŋ 22] *door* 門. Sometimes, weak anticipatory noise excitation of following formant patterns indicates that phonation starts from an open glottis word-initially, for example [ʰɛɪ 214] *together* 合, [ʰɥyə 244] *far* 遠. Word-internally, phonation is usually continuous and so the distinction between two different types of Onsets in sonorant-initial syllables is not usually maintained.

Table 4. *Phonological classification of Wencheng morphemes.*

Morpheme class →	*Yin*	*Yang*
Tone	upper-mid level [44] high rising [45] low rising [24] lower-mid level [22]	depressed upper-mid level [244] mid fall-rise [324] low fall-rise [214] lower-mid level [22]
Onset phonemic obstruent voicing	/voiceless unaspirated stops, affricates/ e.g. /t/ → [t], /ts/ → [ts] /voiceless aspirated stops, affricates/ e.g. /tʰ/ → [tʰ], /tsʰ/ → [tsʰ] /voiceless fricatives/ e.g. /s/ → [s]	/voiced stops, affricates/ e.g. /d/ → [d̥] /#_ , [d] / elsewhere e.g. /dʑ/ → [d̥ʑ̊] /#_ , [dʑ] / elsewhere /voiced fricatives/ e.g. /z/ → [z̥] /#_ , [z] / elsewhere
Sonorant onset phonation	word-initial weak glottal-stop e.g. /ye/ → [ˀɥyə]	Non-abrupt phonatory onset e.g. /ei/ → [ʰɛɪ], /m/ → [m]

The parallel in tone-associated phonatory onset behaviour between the obstruents and sonorants is not, of course, fortuitous, but a manifestation of a more basic *partition of morphemes* into two classes according to both phonation onset behaviour and tone. Morphemes in one class – I will borrow the term *Yin* from traditional Chinese phonology to name it – have isolation tones [44], [45], [24], [22] and their sandhi variants; voiceless obstruent phonemes; and glottalised onset to syllable-initial sonorants. Morphemes in the other, *Yang*, class have isolation tones [244], [324], [214], [22] and their sandhi variants, voiced obstruent phonemes and non-glottalised onset to syllable-initial sonorants[4]. Table 4 summarises this.

Although morphemic identity is normally redundantly underwritten by both tone and Onset features, sometimes one is more important than the other. An isolation [22] tone, for example, can realise both types of morpheme depending on its phonatory onset behaviour, but the morpheme's identity will be signalled by its Onset's phonatory behaviour. I suspect, too, that it is tonal pitch rather than phonatory onset behaviour which is most salient for morphemes on syllables beginning with sonorants word-initially.

This pattern of over-engineering morpheme identity with laryngeal features is canonical Wu, and found in all northern varieties, where it is often reinforced by another laryngeal feature of phonation type. Conservative varieties in the Yongjiang 甬江 subgroup for example are well-known for their non-modal

---

[4] The terms *Yin* and *Yang* are understood in traditional Chinese phonology to refer to classificatory features of Middle Chinese *tones*, not *morphemes*. It is interesting to note that, whilst adhering to this terminology, Ballard chose to start his book-sized (1980) paper on *tones* by pointing out the small number of well-defined phonological shapes for Middle Chinese *morphemes*. It might be worth considering that *Yin* and *Yang* refer to types of morpheme rather than types of tone. This would go to explaining the problem of the separate listing of so-called *Ru* tones, which, ending in Coda stops, would have been in complementary distribution with the other Middle Chinese tones and not separate from them. If *Ru* is a morphological term, this is no longer a problem.

phonation on *Yang* morphemes (Rose 1989b), and Rose (2015), in which audio is embedded, gives examples from two other Wu sites in addition to Yongjiang. Such morpheme reinforcement by phonation type is absent for this speaker, however.

One of the theoretically interesting consequences of the typically Wu tight co-occurrence of laryngeal features just described is that it is very difficult unarbitrarily to distinguish contrastive from conditioned (Sherard 1980: 20 *et pass*.). So one shouldn't: I will assume below that there are both seven contrasting tones *and* phonemic obstruent voicing, and that these combine to characterise two different kinds of morpheme (*Yin/Yang*) in a phonetically non-transparent way.

3. Disyllabic Tone Sandhi

The following sections describe the tones on morphemes within disyllabic phonological words. For example, the word [fʊze 2̲.21] *complicated* 復雜 consists of two syllabic morphemes: {fu 24 *complex* 復} and {ze 214 *miscellaneous* 雜}. Most examples of words elicited were compounds with morphological structure of the quasi-syntactically defined *co-ordinate* type, as in the previous example, or *attribute-head* type, as for example the word [futɕʰɔ 52.44] *train* 火車 from {fu 45 *fire* 火} and {tɕʰɔ 44 *vehicle* 車} (terms are from the description of Modern Standard Chinese morphological constructions in Kratochvil (1968: 73-78) . The sandhi applies irrespective of major word class, and word-classes of noun, verb, adjective and adverb were all represented in the corpus. Since compounding is the most common word-formation process in Chinese varieties, the tone alternations involved in lexical tone sandhi of the kind described in this paper are ubiquitous, and absolutely central to the phonology.

3.1 Overview

Table 5 shows the details of the disyllabic tone sandhi in example words, their segments transcribed phonemically. It is arranged firstly in eight panels according to the isolation tone of the morpheme on the word-final syllable (S2). This is shown along the top. Then within each panel the arrangement is according to the isolation tone of the first-syllable morpheme (S1), shown in the leftmost column. Thus it can be seen, for example, in the top left panel A of the table that when the word-final morpheme has an upper-mid level [44] isolation tone, and the isolation tone on the first-syllable morpheme is high rising [45], the resulting tonal pitch on the disyllabic word is [52.44], i.e. high falling on the word-initial syllable followed by upper-mid level on the word-final syllable. Although there are only seven tones, 64 disyllabic tonal combinations are shown. This is because it is necessary to show two sets of shapes (in panels B and C) for words with the lower-mid level [22] tone on the word-final syllable: as mentioned above, this tone has different sandhi behaviour correlating with its different historical provenance.

The same caveat applies to the representation of the sandhi tone pitch as to the isolation tones: the five-point tone letters indicate approximate pitch shape only and primarily show identity, i.e. which words sound to have the same tonal pitch, and which different. In particular, [4] and [2] represent, as before, upper- and lower-mid pitch values. The transcription of tonal pitch on stretches of speech longer than monosyllables, as here, introduces additional problems. The pitch of speech is highly discriminable, and this creates problems for low resolution transcription methods like the five-point tone letters. It is easy to hear, for example, a slight declinational drop in pitch occurring between two adjacent syllables carrying the same high tone, and if different pitches have to be represented by different transcriptions, as in a transcription at the level of discriminable phones (Laver 1994: 556), then, assuming that the pitch on the first high tone is transcribed [55], a notational system that only allows for five different pitch heights will be exhausted after the fifth syllable. An analogous, but real, example occurs in the data, where after a falling tone, all word-final tones are realised audibly lower. After a falling tone, morphemes with the low fall-rise [214] tone have a low falling pitch word-finally (table 5 panel H). But the pitch of this tone is audibly lower than in the falling part of the isolation tone, so it cannot just be transcribed [21], which would have to imply the same pitch. There are no integers left to represent any lower pitch, however. The solution to this is to make use of the downstep (↓) notation, which is intended to represent "… modifications … of the pitch indicated by ordinary tone symbols" (IPA Guide 1999: 15). The low falling pitch after a falling tone can thus be transcribed [↓21], indicating that the pitch of word-final morphemes carrying the isolation low fall-rise [214] tone is a low fall sounding lower than the falling part of the isolation fall-rise tone.)

Table 5. *Disyllabic lexical tone sandhi examples.*
S1, 2 = word-initial, word-final syllable. Underlining indicates a pitch of short duration.

S1 tone	A S2 tone = upper-mid level [44] Ia			B S2 tone = lower-mid level [22]		
upper level [44] Ia	55.↑44	soŋ ɕie	*rosin* 松香	55.44	tʰie dʑye	*suspension bridge*
high rise [45] IIa	52.44	fu tɕʰo	*train* 火車	52.22	tɕu beŋ	*wine bottle* 酒瓶
lower level [22] IIIa		pai seŋ	*waistcoat* 背心		tʰɑ beŋ	*peace* 太平
lower level [22] Ib	11.44	die tɕi	*frog* 田雞	11.22	ɲaŋ ɑ	*bank* 銀行
dpr. level [244] IIb	342.44	mai po	*tail* 尾巴	342.22	dʑie dʑi	*chess* 象棋
mid fall-rise [324] IIIb		zøy kye	*tree root* 樹根		mie mɛæ	*whiskers* 面毛
low rise [24] IVa	<u>2.</u>44	pe foŋ	*north wind* 北風	<u>2.</u>22	kye diu	*bone* 骨頭
low fall-rise [214] IVb		za seŋ	*honest* 實心		ɲye de	*platform* 月台

S1 tone	C S2 tone = lower-mid level <IIIa [22]			D S2 tone = mid fall-rise [324] IIIb		
upper level [44] Ia	55.22	tɕʰɔ kʰu	*storehouse* 倉庫	55.31	sɑ doŋ	*cave* 山洞
high rise [45] IIa		ta sø	*intend* 打算		tʰei mie	*dignity* 體面
lower level [22] IIIa	44.11	ku i	*intentionally* 故意	44.31	tʰe døy	*attitude* 態度
lower level [22] Ib	11.22	dɕie taŋ	*bench* 長凳	22.31	doŋ z	*workmate* 同事
dpr. level [244] IIb		doŋ tʰe	*developments* 動態		zei ʋai	*society* 社會
mid fall-rise [324] IIIb		nai tɕie	*civil war* 內戰		meŋ leŋ	*order* 命令
low rise [24] IVa	<u>3.</u>22	tɕa tɕʰi	*beri-beri* 腳氣	<u>3.</u>331	i deŋ	*definitely* 一定
low fall-rise [214] IVb		o fei	*tuition* 學費		za o	*truth* 實話

S1 tone	E S2 tone = high rise [45] IIa			F S2 tone = depressed level [244] IIb		
upper level [44] Ia	52.23	kɛæ kʰɛæ	*matric. exam* 高考	52.44	soŋ dei	*brothers* 兄弟
high rise [45] IIa		ɕu tɕie	*palm* 手掌		tɕʰy lei	*manage* 處理
lower level [22] IIIa		jau ɕu	*left hand* 右手		pɛæ zei	*newspaper office* 報社
lower level [22] Ib	342.23	ɑ tei	*sole of shoe* 鞋底	342.44	weŋ dʑie	*documents* 文件
dpr. level [244] IIb		lɛæ tɕu	*booze* 老酒		løy zei	*hostel* 旅社
mid fall-rise [324] IIIb		du tʰai	*thigh* 大腿		dɑ mei	*rice* 大米
low rise [24] IVa	<u>2.</u>35	tɕo pʰeŋ	*goods* 作品	<u>2.</u>344	kyə ny	*national language* 國語
low fall-rise [214] IVb		ɲiu tʰei	*flesh* 肉體		ba ɲie	*white elephant* 白象

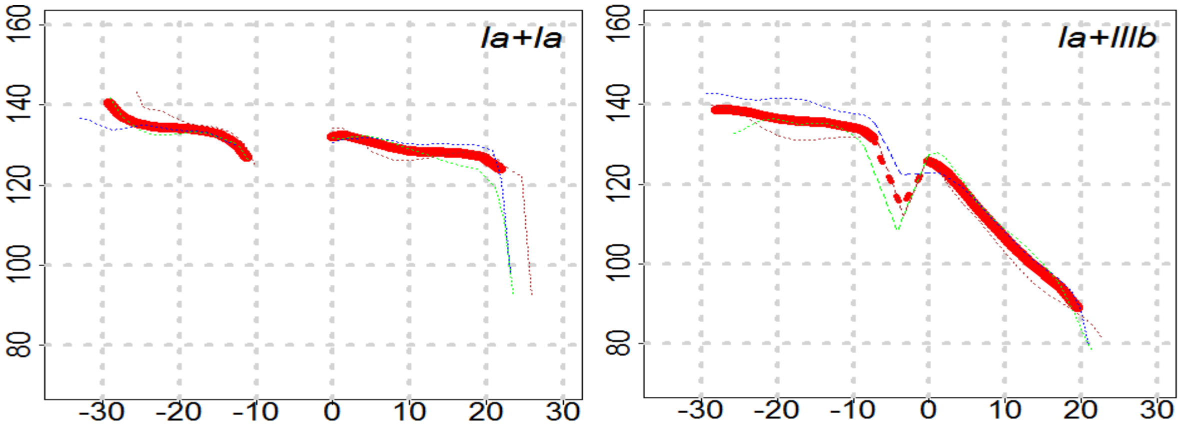
S1 tone	G S2 tone = low rise [24] IVa			H S2 tone = low fall-rise [214] IVb		
upper level [44] Ia	52.12	fuo fuo	*method* 方法	52.21	tɑ dou	*alone* 單獨
high rise [45] IIa		ɕu kye	*arm* 手骨		kʰau ɲie	*tongue* 口舌
lower level [22] IIIa		kʰɑ sou	*high-speed* 快速		sei lei	*power* 勢力
lower level [22] Ib	342.12	mɛæ pe	*writing brush* 毛筆	342.21	meŋ ba	*understand* 明白
dpr. level [244] IIb		dɛæ te	*moral* 道德		dø zø	*cut off* 斷絕
mid fall-rise [324] IIIb		maŋ tø	*questions & answers* 問答		dɑ o	*university* 大學
low rise [24] IVa	3.24	tɕʰi tɕa	*barefoot* 赤腳	2.21	kye ɲu	*flesh and blood* 骨肉
low fall-rise [214]		e tɕɔ	*cooperate* 合作		o za	*study* 學習

3.2 Disyllabic tone sandhi acoustics - general procedure

The tonal acoustics of the disyllabic words, like the isolation tones, were measured by first generating their wideband spectrogram in *Praat*, together with its wave-form and superimposed F0 contour. The disyllabic words were then segmented with reference to the spectrogram into first syllable Rhyme, intervocalic consonant, and second syllable Rhyme. Rhyme onset was adjudged at its first strong glottal pulse. Word offset was adjudged to be at F0 peak in tones with rising pitch, and in tones with level or falling pitch it was adjudged to occur at the start of offset perturbations identified at the point in the wave-form at which discontinuities in either period or peak-to-peak amplitude, or both, occurred. The intervocalic consonant was defined to extend from the offset of the first Rhyme, defined by the observed onset of the consonantal stricture (fricatives) or hold phase (stops), to (for voiced consonants) the release of the occlusion or structure, or (for voiceless consonants) the onset of periodicity after release. Thus the intervocalic consonantal duration comprised both hold and release. The F0 over each Rhyme was then extracted with *Praat*, and modelled with an eighth-order polynomial in *R*. The resulting F0 was then sampled with a sufficiently high frequency to capture the details of its F0 time-course, usually at 10% points of the Rhyme as well as 5% and 95%. F0 was also sampled in mid-duration of a voiced intervocalic consonant.

Three different words were measured for each tonal combination, and their mean values calculated and plotted. Figure 8 shows the results for two tonal combinations, both with upper-mid level morphotonemes on the first syllable. The combination in the left panel has a word-final upper-mid level morphotoneme (the examples were /fei tɕi/ *aeroplane* 飛機, /soŋ ɕie/ *rosin* 松香 and /sei ko/ *watermelon* 西瓜). As can be read from panel A of table 5, these words have [55.↑44] pitch. The combination in the right panel has a word-final mid fall-rise morphotoneme (the examples were /sɑ doŋ/ *cave* 山洞, /fuo bie/ *convenient* 方便, and /koŋ z/ *public business* 公事). These combinations had [55.31] pitch. The F0 for the individual words is plotted with thin dotted lines and their mean F0, with word-offset perturbation removed, with thick solid lines. F0 is plotted as a function of absolute duration aligned at onset of second syllable Rhyme (csec.0). The intervocalic consonants in the examples in the right panel are voiced, because the mid fall-rise morphotoneme co-occurs with obstruents which have fully voiced allophones word-internally – see table 2, section 2.4, and their F0 can be seen plotted with a dashed line showing the characteristic dip associated with the decrease in the transglottal pressure drop caused by supraglottal constriction. Inter-vocalic consonantal F0 is lacking in the left panel of figure 8, as the upper-mid level morhphotoneme co-occurs with voiceless Onset allophones. It can be seen that, as expected, the duration of the voiceless consonants is greater than that of the voiced (mean values for the whole corpus were 12.6 csec. (voiceless) vs. 8.0 csec. (voiced)). Concomitant with this is the expected slightly longer duration (ca. 1 csec.) of the first syllable Rhymes before voiced consonants. The fairly tight clustering of the individual words' F0 values is typical. Figures of all 64 such tonal combinations are available on the web at http://philjohnrose.net/Wu_tones/Wencheng_tones_webpage/Wencheng_2sTS_table.html, where the actual recordings can also be heard from which the words were measured. Their acoustic values constitute the observation data for the description and analysis to follow.

***Figure 8**. Derivation of disyllabic mean acoustics from three individual token words. Both panels show three Wencheng words with upper-mid level morphotonemes on the first syllable. Word-final morphotoneme is upper-mid level (left panel) and mid fall-rise (right panel). Thick solid line = mean F0, dotted line = individual tokens' F0. X axis = duration (csec.) aligned at onset of second syllable Rhyme; y axis = F0 (Hz).*



4. Word-Final Tones in Disyllabic Words

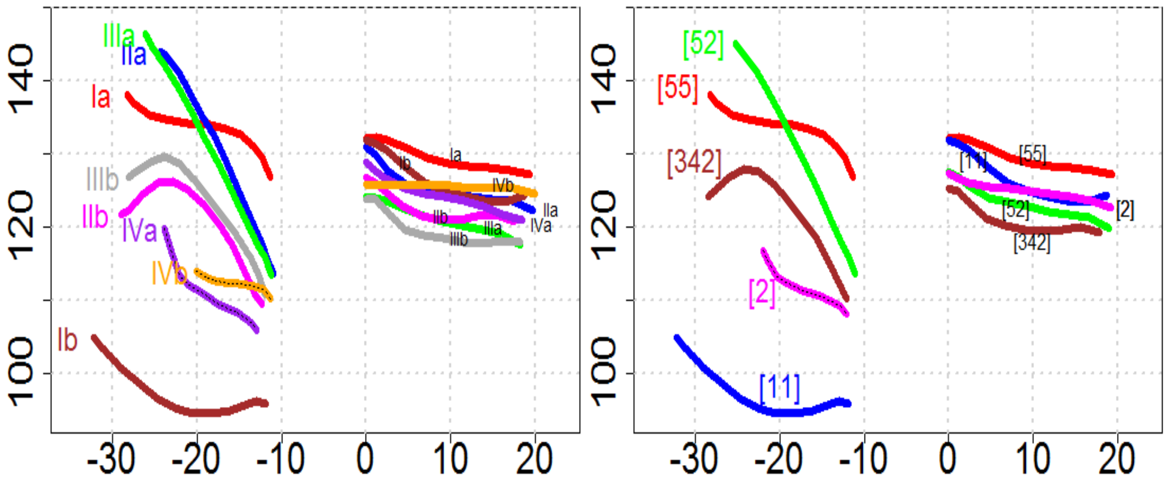
4.1 Acoustic pre-processing

As already pointed out, one of the assumed characteristics of right-dominant tone sandhi is so-called preservation of tones word-finally, whereby the term *tone* is usually understood as isolation tone. Although this can hardly be a necessary and sufficient criterion – one might even expect it perhaps to be a typological default for tone languages – it is nevertheless interesting to examine what *word-final preservation of tone* might mean for these data, since it is unlikely of course to mean the preservation of the exact isolation tone pitch values. This section thus describes the realisation of word-final tones in disyllabic words. It will be shown that the notion of word-final tone preservation can be a little more complex than it initially appears.

Eyeballing the impressionistic descriptions in table 5 – scanning down its columns and comparing the word-final syllable tone in each panel with the corresponding isolation tone at the top – it can be seen that tones on word-final syllables do generally resemble their isolation tones. There are instances of word-final tone and isolation tone represented with the same pitch, but generally there are audible differences, most of which clearly result from the expected perseverative assimilation to the preceding tone. For example, all except one instance of the lower-mid level morphotoneme in panel B of table 5 are shown with the same [22] pitch word-finally as in isolation. After a high level [55] tone on the first syllable, however, the lower-mid level tone is realised with a higher [44] pitch, the height assimilation also enabled by the voiced intervocalic consonant.

Even if there are no assimilatory pitch changes, however, the acoustic properties of tones will be perturbed by occurrence word-finally after another syllable (Xu 1997). In order to examine any preservatory relationship between isolation and word-final tones, therefore, it is necessary first to determine how word-final tone is conditioned by the preceding tone. Then comparison can be made between the word-final conditioned variants and the isolation tone. The section below thus uses acoustic analysis to get a more precise picture of how the word-final tone varies as a function of the preceding tone, and how its acoustic allotones relate to the isolation tone.

***Figure 9**. Steps in simplifying disyllabic mean tonal acoustics of disyllabic Wencheng words with an upper-mid level morphotoneme on the final syllable. Left panel shows all eight combinations; right panel shows combinations pooled with respect to same first-syllable tones. X axis = duration (csec.) aligned at onset of second syllable Rhyme; y axis = F0 (Hz).*



The left panel of figure 9 plots the raw mean acoustic data for all eight disyllabic combinations with the upper-mid level morphotoneme [44] on the word-final syllable (these combinations are shown in panel A of table 5)[5]. To avoid clutter, combinations are identified according to the Middle Chinese category (IIa IVb etc.) of their *first-syllable* morphotoneme. Panel A of table 5 shows that, before a word-final [44] morphotoneme, both first-syllable high rise [45] and lower-mid level [22] (*IIIa) morphotonemes are realised as a high falling [52] tone; depressed upper-mid level [244] and mid fall-rise [324] are both realised as depressed high falling [342]; and low rise [24] and low fall-rise [214] are both realised with low short [2̲]. These mergers are confirmed by the corresponding very similar first-syllable F0 shapes, e.g. for IIa and IIIa, in the left panel of figure 9.

The eight F0 shapes on the word-final syllable in the left panel of figure 9 lie in a narrow corridor of about 10 Hz between about 120 Hz and 130 Hz. Since these eight shapes all realise the same upper-mid level [44] morphotoneme, and have similar pitch, it can be assumed that their variation is an intrinsic assimilatory function of features of the different first-syllable tones preceding them, together with an error term probably mostly associated with other intrinsic factors like vowel and consonantal effects on F0.

The first step towards specifying intrinsic acoustic allotones of the word-final tones involves simplification of the F0 shapes in the left panel by pooling those F0 shapes for combinations with the same first-syllable tone. This is shown in the right panel of figure 9, where F0 values for combinations with IIa and IIIa morphotonemes on the first syllable have been pooled, as well as IIb and IIIb, and IVa and IVb. The resulting five combinations are now labelled according to the pitch of their first-syllable tone.

The variation in the word-final tonal acoustics in the right panel of figure 9 is the result of conditioning by tonal features of the preceding syllable, but which features? It is clear from both phonetic considerations on the nature of progressive assimilation, and statistical considerations of significant differences between word-final F0 shapes, that, for example, the F0 shapes of the word-final tones after the high [55] tone, after the short [2̲] and low [11] tones, and after the depressed fall [342] tones represent three different allotones conditioned by different features on the preceding syllable tone. But consider the two lowest word-final F0 shapes occurring after the high fall [52] and the depressed fall [342]. Are these the result of a single [+Fall]

---

[5] The actual words measured were as follows.
Ia+Ia [55.↑44]: /fei tɕi/ *aeroplane* 飛機, /soŋ ɕie/ *rosin* 松香, /sei ko/ *watermelon* 西瓜
IIa+Ia [52.44]: /tie seŋ/ *dimsum* 點心, /fu tɕʰo/ *train* 火車, /pʰu tʰoŋ/ *common* 普通
IIIa+Ia [52.44]: /kʰua tɕʰo/ *express train* 快車, /si fuo/ *on all sides* 四方, /pai seŋ/ *waistcoat* 背心
Ib+Ia [11.44]: /die tɕi/ *frog* 田雞, /noŋ tɕʰø/農村 *village*, /dʑie sɑ/ 長衫 *long gown*
IIb+Ia [342.44]: /bei kʰu/ *sleeping quilt* 被窩, /maɪ po/ *tail* 尾巴, /zo tɕʰo/ *to ride* (verb-object) 坐車
IIIb+Ia [342.44]: /zøy kye/ *tree root* 樹根, /wɑ tʰeŋ/ *canteen* 飯廳, /dei fuo/ *place* 地方
IVa+Ia: [2̲.44] /tɕie kuo/ *Zhejiang* 浙江, /kye kʰu/ *orthopedics* 骨科, /pe foŋ/ *north wind* 北風
IVb+Ia: [2̲.44] /ɲye kuo/ *moon* 月光, /bi tɕeŋ/ *pin* 別針, /za seŋ/ *honest* 實心

feature on the preceding tone – in which case they should be pooled as a single acoustic allotone? Or are there two separate conditioning features involved, one for the high fall and one for the depressed fall, with the depressed fall lowering the following tone more? In that case, the word-final F0 shapes after [52] and [342] are two separate acoustic allotones and should be kept apart.

In order to resolve this question, competing hypotheses of conditioning by the preceding syllable tone were compared by seeing which hypothesis makes best sense in terms of the observed variation in the word-final F0 shapes (i.e. asking under which hypothesis the word-final F0 shapes are the most probable). Since the separate allotonic status of F0 shapes after high [55], short/low [2]/[11], and depressed fall [342] tones is clear, three hypotheses were established, differing in the integration of the F0 shape after the high fall [52] tone:

- Hypothesis 1: the high fall [52] tone exerts the same lowering effect on the following tone as the depressed falling [342] tone.
- Hypothesis 2: the high fall [52] tone exerts a lowering effect on the following tone different from, and intermediate between, depressed falling [342] and short/low [2]/[11] tones.
- Hypothesis 3: the high fall [52] tone exerts the same lowering effect on the following tone as the short/low [2]/[11] tones.

These three hypotheses were evaluated with a standard linear model of word-final F0 shape as a function of first-syllable F0 shape, with different fixed factors corresponding to the three hypotheses, thus:

- For H_1 the first-syllable F0 shape was treated as a factor with three levels: *High*, *Short+low*, and *Fall*. That is, the word-final F0 values after high fall [52] and depressed fall [342] tones were grouped together under *Fall*; word-final values after the short [2] and low [11] tones were grouped together as *Short+low*; and word-final F0 values after the high [55] were indexed as *High*.
- For H_2, first-syllable F0 shape was treated as a factor with four levels: *High, Short+low, Depressed-fall* and *High-fall*.
- For H_3, first-syllable F0 shape was treated as a factor with three levels of *High, Short+low+high-fall, and Depressed-fall*.

The word-final F0 shape was represented by its F0 value in mid-Rhyme as dependent variable. To take into account the effect of first-syllable tonal shape on all six relevant word-final morphotonemes, data was used, not just from combinations with the word-final upper-mid level morphotoneme under examination here, but from all combinations with first-syllable high falling, depressed falling, short, low and high tones (i.e. all combinations except word-final lower-mid level [22] (<III) and mid fall-rise, which lack high and depressed falling first-syllable tones). This variation in word-final morphotoneme was included as a second fixed effect in the linear model.

Evaluation with Bayes' Information Criterion (BIC) and ANOVA post-hoc *p* values selected Hypothesis 1, thus explicitly recognising the conditioning effect of a shared [+Fall] feature in the preceding high fall and depressed falling tones[6]. Word-final F0 shapes after high falling [52] and depressed-falling [342] first-syllable tones were therefore pooled. The final result of this pooling of word-final tonal shapes for the upper-mid level [44] morphotoneme is represented in panel A of figure 10. The F0 for the upper-mid level isolation tone is also plotted, with a thicker line, to show how the word-final acoustic allotones thus derived relate to their corresponding isolation tone.

---

[6] The Bayes Information Criterion (BIC) is a commonly used metric for comparing models representing different hypotheses. It is given by BIC = -2 * ln(L) + ln(*n*) * *npar*, where L and *npar* denote respectively the likelihood and degrees of freedom of the model, and *n* is number of observations. BIC is thus essentially a function of the likelihood of the model, given the data, i.e *P*(data|model), weighted by a term penalising the model's complexity, itself weighted by a term taking the number of observations into account. Given two competing models equally probable before the data are adduced (which is the case here), lower BIC values select the preferred model. In this case, the evidence favouring H_1 (BIC = 774) was only marginally greater than H_2 (BIC = 777), with both considerably stronger than H_3 (BIC = 821). Since both H_1 and H_2 were associated with very similar log-likelihoods, the BIC superiority of H_1 over H_2 was due to its model parsimony. H_1 is also superior in having the means of each of its three word-final groups (after high; after high falling/depressed falling; and after short/low tones) highly significantly different from each other; whereas with H_2, the means of word-final tones after high fall and depressed fall, being almost exactly the same, do not differ significantly.

Panel A of figure 10 thus shows the upper-mid level morphotoneme to have three word-final acoustic allotones. The lowest, plotted with a dotted line, represents mean values lowered by occurrence after first-syllable falling and depressed falling tones (also plotted with dotted lines). The highest word-final shape, plotted with a solid line, represents values raised after the high level tone. Values after low and short tones, plotted with a dashed line, form the third, intermediate, allotone. The similarity between these intermediate values and the isolation tone F0 can be interpreted as indicating that preceding low and short tones do not exert any assimilatory influence on the realisation of the word-final tones: they are the same as the isolation tone, allowing for the differences in duration associated with occurrence in disyllabic words (the Rhyme duration of word-final syllables is fairly uniform at about 20 csec., and, except for the high rising tone, some 10 – 20 csec. shorter than the corresponding isolation tones).

As a whole, figure 10 thus makes explicit the conditioning of word-final tonal acoustics as a function of the tone on the preceding syllable, and also shows the relationship between the word-final allotones and the isolation tones. It is divided into eight panels, labeled A through H, corresponding to the panels in table 5, with each panel showing the mean tonal acoustics for words with a different word-final morphotoneme. The specific morphotoneme is named in the one-line legend for the isolation tone. Panel A, for example, shows acoustics for words with the upper-mid level morphotoneme on the word-final syllable. The figure has been arranged so that its right-hand column contains the tonal acoustics of disyllabic words with word-final *Yang* morphemes – they have isolation tones that co-occur with the phonemically voiced syllable-initial consonants that have voiced allophones word-internally (see section 2.6). Words with *Yin* morpheme final syllables are in the left-hand column morphemes – they have isolation tones that co-occur with the phonemically voiceless Onset obstruents. Typical intrinsic consonantal F0 perturbations associated with the voiced and voiceless syllable-initial Onsets can also be seen at the onset of most second-syllable Rhymes. As with the isolation tones, word-final offset F0 perturbations have been identified and omitted.

It can be seen that some panels have more first-syllable tones than others. Panel A, for example, shows F0 shapes for five first-syllable tones, compared to panel E, with only three. This is because before word-final low rising morphotonemes (panel E) there are only three possible tones on the first syllable, whereas five first syllable tones can occur before the upper-mid level morphotoneme (this information can be seen in the corresponding panels of table 5).

4.2 Relationship between word-final and isolation tones

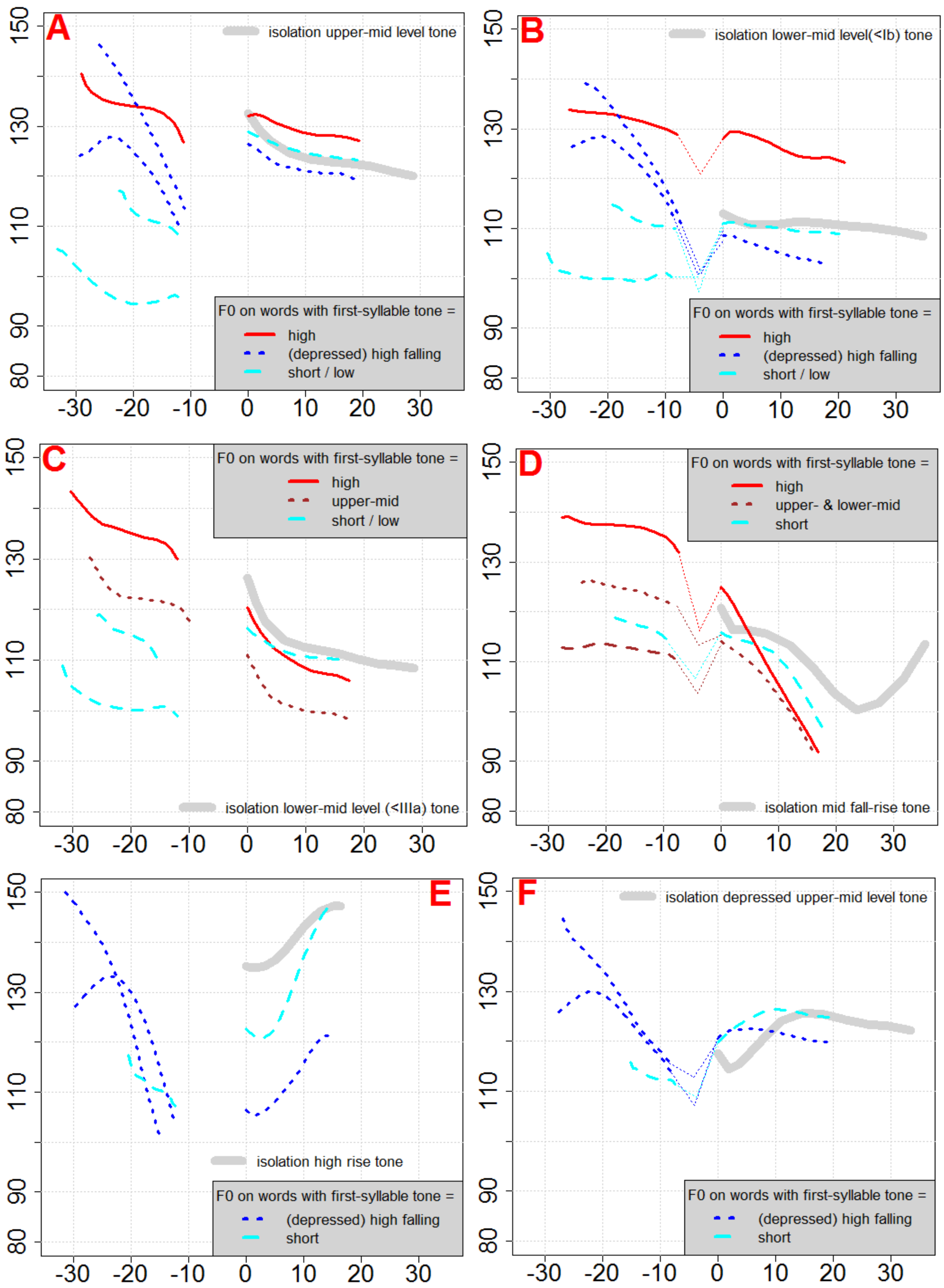
The acoustic data in figure 10 indicate little difference in F0 between isolation and word-final level tones – i.e. when the tone on the word-final morpheme is upper-mid level (panel A), depressed upper-mid level (panel F), or lower-mid level (panels B, C) – preceded by either a short tone (transcribed as [<u>2</u>] or [<u>3</u>]) or a low (transcribed as [11]). In all other cases, the word-final realisation differs from the isolation tone. The differences, and what conditions them, are listed below. It will be seen that they clearly result from phonetically plausible processes: co-articulation, dissimilation, contour loss.

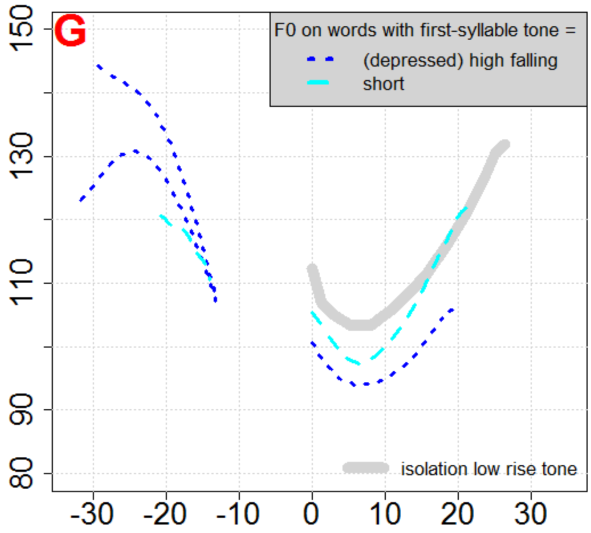
4.2.1 Loss of final rise in fall-rise tones

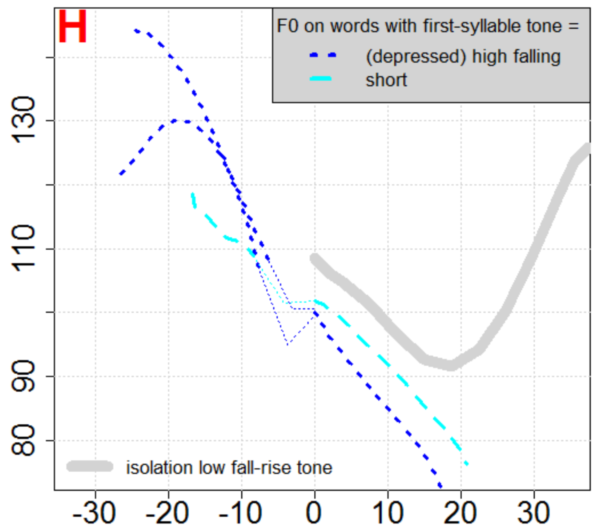
Probably the most salient auditory difference between the isolation and word-final tones occurs with the two fall-rise isolation tones [324, 214], which, as table 5 shows, lack a final rise word-finally and consequently have a mid falling and low falling pitch respectively. Panels D and H of figure 10 show the acoustics, where it can be seen that the F0 of all acoustic allotones continues to fall to offset. The mean offset of the low fall-rise allotone after a high fall, at about 80 Hz, is the lowest value in the speaker's F0 range.

This may be an instance of a no-crowding constraint, easily formalisable with the deletion of the last tone of a three autosegmental HLH tone sequence for the isolation tones. Deletion of a second H tone-bearing-mora would also model the process. Since three autosegmental tones crowd together quite happily in the mid and low fall-rise isolation tones, however, occurrence as part of a longer unit, i.e. a word, must also figure in an explanation for the deletion. Figure 10 shows that all final tones in disyllabic words have a Rhyme duration of about 20 csec., so a simple truncation of the isolation fall-rise tonal Rhyme to the 20 csec. required for disyllabic words would indeed leave just the falling portion of the F0, thus motivating the no-crowding construal phonetically

Figure 10*. Acoustics of disyllabic lexical tone sandhi showing intrinsic acoustic variation on word-final tone. Corresponding isolation tone is plotted in thick grey. X axis = duration (csec.) aligned at onset of word-final Rhyme; y axis = F0 (Hz).*







4.2.2 Lowering and depression-absorption after preceding falling tone

Figure 10 shows that all tones have acoustic allotones with lower F0 after both high falling [52] and depressed-falling [342] tones. This lowering after a falling tone is presumably an inertial effect from the contraction of strap muscles involved in the latter part of the fall.

The effect of the preceding fall is greatest on the high rising tone (panel E of figure 10), converting it from a high rising [45] isolation tone into what sounds clearly like a low register low rising [23] word-final tone, with an F0 lying about 25 – 30 Hz lower than in isolation. For the remaining tones the magnitude of the effect associated with the preceding fall appears to decrease with height. The upper-mid level and depressed level tones (panels A, F) show only a slight, ca. 4 - 8 Hz, lowering, and the low rising and low fall-rising tones (panels G, H) show a mean drop of about 10 Hz relative to isolation. In several cases the magnitude of the effect is big enough to be reflected in pitch, in which case I have used the appropriate downstep (↓) notation in the impressionistic transcription in table 5. The overall shifting down of the word-final tones relative to their isolation forms gives the impression that the lower boundary of the tonal space has been stretched downwards.

An additional effect of occurrence after a fall is the audible loss of the depression in the depressed upper-mid [244] tone (the dotted line in figure 10, panel F) which, as well as lowering slightly, loses its initial rise and becomes upper-mid level [44]. Since one might expect the low onset of the depressed upper-mid level tone to be preserved precisely in the environment of a preceding low offset, its absorption must be due to a factor other than assimilation in cord tension.

4.2.3 Lowering after preceding non-falling low/mid tone

As noted above, a short first-syllable tone has no effect on a following level tone. On contour tones, however, figure 10 (for example panels G, H, dashed line) shows that the low short tone exerts a lowering effect, but of a magnitude less than the high fall. Again the greatest difference from the isolation form occurs with the high rising [45] tone. The acoustics are in panel E of figure 10, where both isolation and word-final high rise allotones can be seen to share the same peak F0 target (as well as the same duration), but the mean onset of the word-final allotone after [2] is some 15 Hz lower than in isolation. The fact that the high rising isolation peak F0 target is achieved word-finally, even with a lower onset, seems to indicate its importance. This is not the case with the low [24] rising tone, which does not reach the same F0 peak after a short tone as in isolation (panel G). Apart from the high rising tone, the F0 values of the contour tones lie about 5 Hz lower than in isolation and I cannot hear regular differences in pitch from the isolation tones for these. The lowering effect on contour tones is presumably an assimilation to the preceding tone.

The overall shifting down of word-final tones relative to their isolation forms is not so great after a short or low tone as after a fall, described in 4.2.2. The fact that lowering is greater after the falling tones than after the non-falling low [11] and short tones is interesting, given that figure 10 shows they have very similar offset F0 values. To achieve explanatory adequacy this would presumably have to be represented as conditioned by a contour feature [+Fall] on the preceding tone rather than the use of an autosegmental sequence of [HL], the L of which would have the same value for both short and fall tones and would therefore not be able to plausibly condition differential lowering.

4.2.4 Raising after preceding high tone

Panel B of figure 10 shows a clear assimilation after the high level [55] tone on the preceding syllable, where a following lower-mid level [22] tone is raised to the same F0 and [44] pitch height as the upper-mid level tone, presumably enabled by the continuous phonation of the word-internal voiced Onset consonant. Occurrence after a preceding [55] also raises the pitch and F0 of the upper-mid level tone a little (panel A), and the F0 at onset of the mid falling-rising tone (panel D) is also raised.

4.2.5 Lowering after preceding upper-mid level tone

Panel C shows the lower-mid level [22] <IIIa tone is lowered after a preceding upper-mid [44] tone (dotted line). As can be seen in row 3 of panel C in table 5, this creates a [11] word-final tonal pitch which does not occur in isolation. Given the persevatory raising effect after [55] just described in 4.2.4, this lowering is unusual: one would have expected the [22] also to have raised, not lowered, after [44]. The lowering looks therefore to be a dissimilation in height, which is unusual for following tones (it is the preceding tone which usually dissimilates). Panel D (dotted line) shows a similar effect on the following mid fall-rise tone.

4.3 Word-final tonal neutralisation

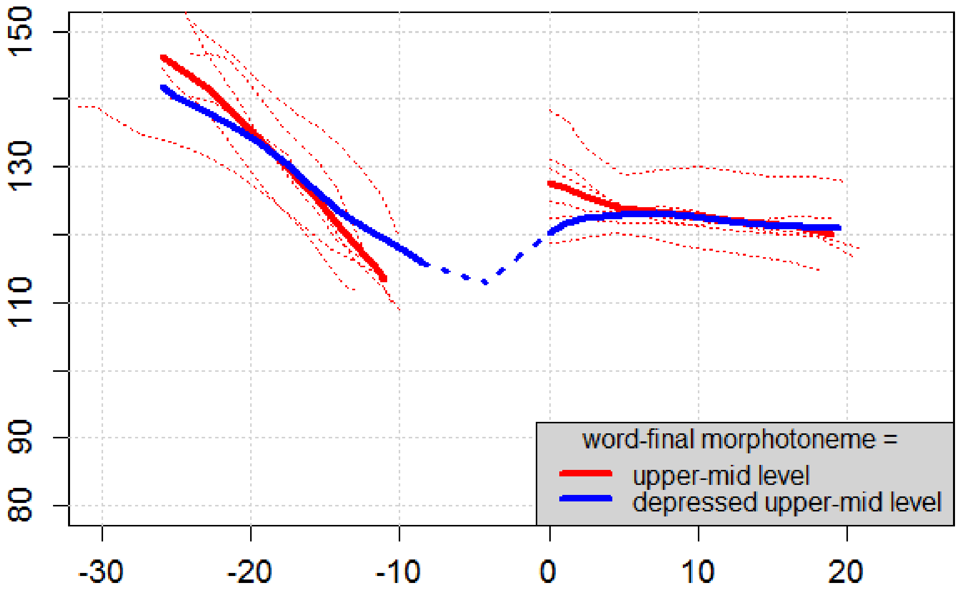
The conditioning just described complicates the idea of tonal preservation, by giving rise to three instances of apparent tonal neutralisation on the word-final syllable, where morphemes with different isolation tones have the same pitch word-finally. Consider for example the tonal realisation of morphemes with the depressed upper-mid level [244] isolation tone after a high falling or depressed-falling tone. As indicated in the previous section 4.2.2 the depression is lost, giving rise to a tone with [44] pitch, the same as for the word-final upper-mid level tone. Examples are [sʊŋdɛɪ 52.44] *brothers*, which has a word-final morpheme with a depressed upper-mid level tone: [d̥ɛɪ 244], and [fu tɕʰɔ 52.44] *train* with a word-final morpheme with an upper-mid level tone: [tɕʰɔ 44].

The transcriptional identity for the two word-final tonal realisations implies, of course, that their pitch is the same. This is confirmed in figure 11, which shows that the mean acoustic realisation of the [244] morphotoneme after a first-syllable falling tone is almost exactly the same as that of the [44] morphotoneme in the same environment, except for the initial perturbation caused by the difference in Onset voicing associated with the two morphotonemes. (The [244] tone belongs to *Yang* morphemes, and so its Onset stop will be voiced; the [44] tone belongs to Yin morphemes, so its Onset fricative will be voiceless.) The same degree of similarity is also observed after depressed-falling first-syllable tones.

Obviously, if tone contrasts are neutralised, tones cannot be said to be preserved. Or can they? For it to be meaningful, neutralisation must have perceptual implications. For these data, due to the tight morpheme-internal relationship between tones and Onsets, this is not the case: the identity of the morphemes involved, and hence their tone, is still guaranteed by the voicing of the syllable-initial consonant. For example, given the word-final [44] pitch in the words [fu tɕʰɔ 52.44] *train* and [sʊŋdɛɪ 52.44] *brothers*, the respective provenance of the word-final tone from upper-mid level and depressed upper-mid level morphemes is signalled by their voiceless and voiced syllable-Onset consonants /tɕʰ/ and /d/.

This phenomenon, whereby word-final morphemic identity is guaranteed by information external to its tonal pitch, occurs quite widely in the data. Mostly it is the syllable-Onset voicing which is important – another case is the raising of the lower-mid level (Ib) tone to the same [44] pitch as the upper-mid level after a preceding high tone (panel B in table 5 & figure 10). But there is also a case where the sandhi-shape of the preceding tone carries the information. For example word-final morphemes with the lower-mid isolation tone [22] (<Ib) and with the depressed upper-mid level isolation tone [244] are both found realised with a [44] pitch word-finally, even though their syllable-initial consonants are voiced (panels B & F in table 5 & figure 10). In the first instance however (for example in the word [œβɑ 55.44] *arrange* < [44] + [22]), the relationship of the [44] to the lower-mid level tonal morpheme can be recovered from the [44] pitch on the first syllable, which is different from the [52] first syllable pitch before morphemes carrying the depressed level isolation tone (in for example [sʊŋdɛɪ 52.44] *brothers* < [44] + [244]). It is interesting to note that in fact in every case where there is apparent word-final tonal neutralisation, the morpheme identity, and hence its tone, is signalled by some other phonic information in the word, and biuniqueness is preserved.

Figure 11. *Mean acoustics of disyllabic words with word-final upper-mid level (red) and depressed upper-mid level (blue) morphotonemes after a falling tone on the first syllable. Dotted lines = individual tokens with word-final upper-mid level morphotoneme. X axis = duration (csec.) aligned at onset of second syllable Rhyme; y axis = F0 (Hz).*



4.4 Summary – word-final tones

The description and analysis presented in the sections above have examined how well the idea of word-final preservation of isolation tones applies as a characteristic of right-dominant sandhi. Whether one analyses the variation phonemically or morpho-phonemically, the low level phonetic differences demonstrated are not enough to result in loss of biuniqeness. The set of word-final tones thus relates straightforwardly to the set of isolation tones, in the sense that both can be considered realisations of the same phonemic or morpho-phonemic form, the word-final realisations having been conditioned by phonetically plausible factors associated with occurrence in disyllabic words. It is in this sense that tone is preserved word-finally.

The demonstrable similarity between word-final and isolation tones, once plausible intrinsic effects are discounted, allows us to say that the same tone is involved both word-finally and in isolation. However, the situation has also been shown to be a little more complex in two senses. Firstly, tones are also preserved word-finally in the sense that, thanks mostly to tight correlation with segmental properties, they are recoverable from preserved morphemic identity even though there may be phonetic changes that neutralise different tonal pitches. Secondly, the more accurate picture afforded by the acoustics shows that the word-final tone is actually conditioned by the phonetic shape of the preceding tone, especially when the latter is falling. Thus even in right-dominant sandhi the left tone still exerts an influence. This presents certain problems of metrical interpretation. Preservation of word-final tone surely counts as strong evidence for the metrical strength of the word-final position; yet the conditioning of the word-final tone by the preceding tone suggests rather the latter as the metrically strong element.

5. Word-Initial Tones in Disyllabic Words

The second characteristic of right-dominant sandhi is apparent neutralisation of often complicated subsets of tonal contrasts on non-word-final syllables, and the disyllabic data provide some nice examples of this. Table 6, with word-final morphotoneme shown along the top, and the tone on the preceding syllable morpheme down the left hand side, shows the details of tonal realisation on the first syllable. It has been arranged to optimally show the neutralisation patterns.

It can be seen, firstly, that the realisation of the first-syllable tone is partly a function of the tone on the following, word-final syllable, although in most cases the conditioning is not phonetic. Thus for example morphemes with the upper-mid level [44] tone are realised with a high level [55] pitch on the first syllable if the word-final syllable carries either upper-mid level [44], lower-mid level [22] or mid fall-rise [324] tones; but otherwise first syllable upper-mid level morphotonemes are realised with a high falling [52] pitch. It is difficult to see any phonetic conditioning in this. Secondly it is clear that considerable loss of tonal contrast is involved. For example, the depressed upper-mid level [244] tone and the mid fall-rise [324] tone are never distinguished.

Table 6. *Realisation of first-syllable tone.*
T_1, T_{wf} = tones on first and word-final morphemes.

<table>
<tr><th></th><th colspan="8">T_{wf}</th></tr>
<tr><th>T_1 ↓</th><th>Ia [44]</th><th>Ib [22]</th><th>IIa [45]</th><th>IIb [244]</th><th>IVa [24]</th><th>IVb [214]</th><th>IIIa [22]</th><th>IIIb [324]</th></tr>
<tr><td>Ia [44]</td><td colspan="2">[55]</td><td colspan="4" rowspan="3">[52]</td><td colspan="2" rowspan="2">[55]</td></tr>
<tr><td>IIa [45]</td><td colspan="2" rowspan="2"></td></tr>
<tr><td>IIIa [22]</td><td colspan="2">[44]</td></tr>
<tr><td>Ib [22]</td><td colspan="2">[11]</td><td colspan="4" rowspan="3">[342]</td><td rowspan="3">[11]</td><td rowspan="3">[22]</td></tr>
<tr><td>IIb [244]</td><td colspan="2" rowspan="2"></td></tr>
<tr><td>IIIb [324]</td></tr>
<tr><td>IVa [24]</td><td colspan="4" rowspan="2">[2]</td><td rowspan="2">[3]</td><td rowspan="2">[2]</td><td colspan="2" rowspan="2">[3]</td></tr>
<tr><td>IVb [214]</td></tr>
</table>

5.1 Relationship between first-syllable and isolation tones

In relating the isolation tones to the first-syllable tonal shapes in table 6 the most general statements that can be made are as follows. Note that, from the point of view of pitch shapes, the natural classes involved nearly always constitute *unnatural* tonal classes.

- The low rise [24] and low fall-rise [214] isolation tones are always realized with a short [2] or [3] pitch on the first syllable;
- otherwise, in the majority of cases, isolation tones are realised with a high falling pitched tone. This is depressed [342] when the isolation tone belongs to the set occurring with the *Yang* class of morphemes i.e. lower-mid level (<Ib), mid fall-rise, or depressed upper-mid level. The first syllable realisation is non-depressed [52] when the isolation tone belongs to the *Yin* class of morphemes, i.e. upper-mid level, high rise, or lower-mid level (<IIIa). These two phonological classes of morphemes were discussed in section 2.6.
- There are two exceptions to this high falling tone realisation. One occurs when the word-final tone is lower-mid level (<IIIa) or mid fall-rise, where all first-syllable tones are realized with various level pitches [55] [44] [22] and [11] (see table 5 panels C, D). The second exception occurs when first-syllable upper-mid level [44] or lower-mid level [22] (<Ib) tones occur before themselves on the second syllable. The realisations are then [55] and [11] respectively.

A further complexity in the realisation of the first syllable tones is their differential merging. Consider for example the first syllable upper-mid level, high rise and lower-mid level (<IIIa) tones. Before lower-mid level (<IIIa) and mid fall-rise tones, table 6 shows upper-mid level and high rise isolation tones are merged as [55], separate from the lower-mid level (<IIIa) isolation tone, which is realized as [44]. This grouping is reversed before upper-mid level and lower-mid level (<Ib) tones, where it is the high rise and lower-mid

level (<IIIa) isolation tones that are merged (to a fall), separate from the [55] realisation of the upper-mid level isolation tone.

As far as the realisation is concerned, table 6 shows eight distinct pitch shapes on the first syllable of disyllabic words. Unlike the phonetically transparent set of word-final tonal shapes, which relate clearly to the isolation tones, the set of first syllable tones resulting from the sandhi is rather different from both word-final and/or isolation tones. There are six level pitches: [55 44 22 11 2 3], with the only contour pitch shapes being high falling [52] and depressed falling [342]. Thus rising pitch tones found in isolation and word-final position are missing; and the majority high falling realisation for first-syllable tones does not occur in isolation or word-finally. This means only three pitch shapes are shared between isolation and first-syllable occurrence: [44, 22, 11].

Although there is nearly the same number of first-syllable pitch shapes as with isolation tones, the merger patterns mean that the relationship between the two sets is rather complicated and largely not bidirectionally unique: it is not possible to recover an isolation tone from the shape of the first syllable tone alone; neither is it possible to predict the first syllable tone shape from the isolation tone alone. For example, first syllable morphemes with a word-initial [55] pitch correspond to morphemes with both upper mid-level [44] and high rising [45] isolation tones; and the upper mid-level isolation [44] tone is realised as both high level [55] and high falling [52].

These examples serve to highlight the nature of the complexity of apparent neutralisations in non-word-final position: they are mostly not phonetically conditioned; they group different sub-sets of isolation tones; and the result of the apparent neutralisation is also often not one of the isolation tones.

5.2 Intrinsic variation in first-syllable tone acoustics

The acoustics of the first-syllable tones have already been shown in figure 10, where it is easy to pick out the F0 shapes corresponding to the eight pitches shown in table 6. For example, in panel A of figure 10 can be seen the high fall [52] and depressed high fall [342]; the high level [55], the low level [11] and the short [2], or in panel D can be seen the high level [55], the upper-mid level [44], the lower-mid level [22] and the short [3].

Like the word-final tones, the first-syllable tones are also expected to have intrinsic acoustic allotones conditioned by the word-medial consonant and the following tone. A clear example of co-articulation is in the offset of the falling and depressed falling tones, which is higher before the high [44] allotone of the depressed upper-mid level tone than before the low allotone of the [22] lower-mid level tone (figure 10, panels F & B). Since the high falling and depressed falling tones, not occurring in isolation, are clearly the result of sandhi changes, we see here co-articulation applying to sandhi tones. Moreover, since the upper-mid level tone has lost its depressed onset by co-articulation with the preceding fall (sections 4.2.2, 4.3), these examples show co-articulation working in both directions in addition to tone sandhi.

In addition to the expected slightly longer duration of the preceding Rhyme, a voiced intervocalic consonant also conditions slightly higher offset F0 values on the preceding Rhyme than a voiceless. Interestingly, figure 10 also shows that the F0 *onset* to the high falling [52] tone is higher before a voiced intervocalic consonant (compare the F0 onset values of the high falling tone in panels A and B, and E and F, for example)[7]. The only example of a possible expected dissimilation in height from a following tone occurs with the low [11] tone before the upper-mid level [44] tone. A comparison of the F0 shapes corresponding to the first-syllable low [11] tone in panels A and B shows that F0 falls lower before the [44] than before the [22] tone. I cannot hear a consistent pitch difference for these acoustic differences, however, and therefore have given them the same [11] pitch transcription in table 5.

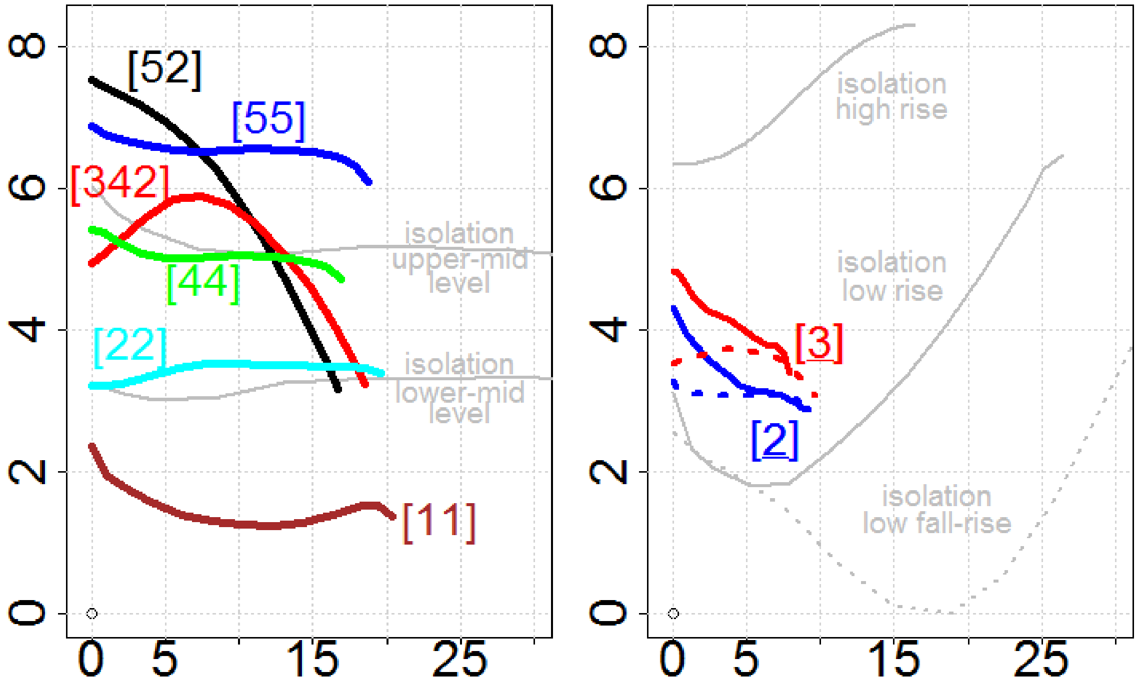
5.3 Long first-syllable tones

Figure 12 shows the mean characteristics of the first-syllable tones pooled across occurrences before all word-final tones. F0 values have been perceptually transformed analogously to figure 4 in section 2.3. In its left panel are shown, plotted with thicker lines, the mean values of the long tones corresponding to the Chao tone letters in table 6, namely [52], [342], [55], [44], [22], and [11]. The transformed F0 for the upper and

[7] As this was not a function of word-Onset consonantal features associated with [+/- spread glottis], the difference was checked with a two-way ANOVA on the F0 onset as response and the following tone and consonantal voicing as factors. This showed the difference at F0 onset to have a significant *p* value of 0.08.

lower-mid isolation tones is also plotted for reference. The mean values for these six long tones were calculated from the groups identified with the same Chao pitch letters in table 6. Thus the mean shape for the [52] tone was calculated from the six separate means shown in panels A B E F G and H of figure 10. The mean values are thus calculated over all the intrinsic allotones and therefore give an approximation of the first-syllable tone target independent of the factors affecting the individual tone.

Figure 12. *Perceptual transform of pooled mean acoustics of first-syllable tones. Relevant isolation tones are plotted with thinner lines. Dotted lines indicate values associated with the low fall-rising tone. X axis = duration (csec.) aligned from onset of Rhyme, y axis = F0 (semitones).*

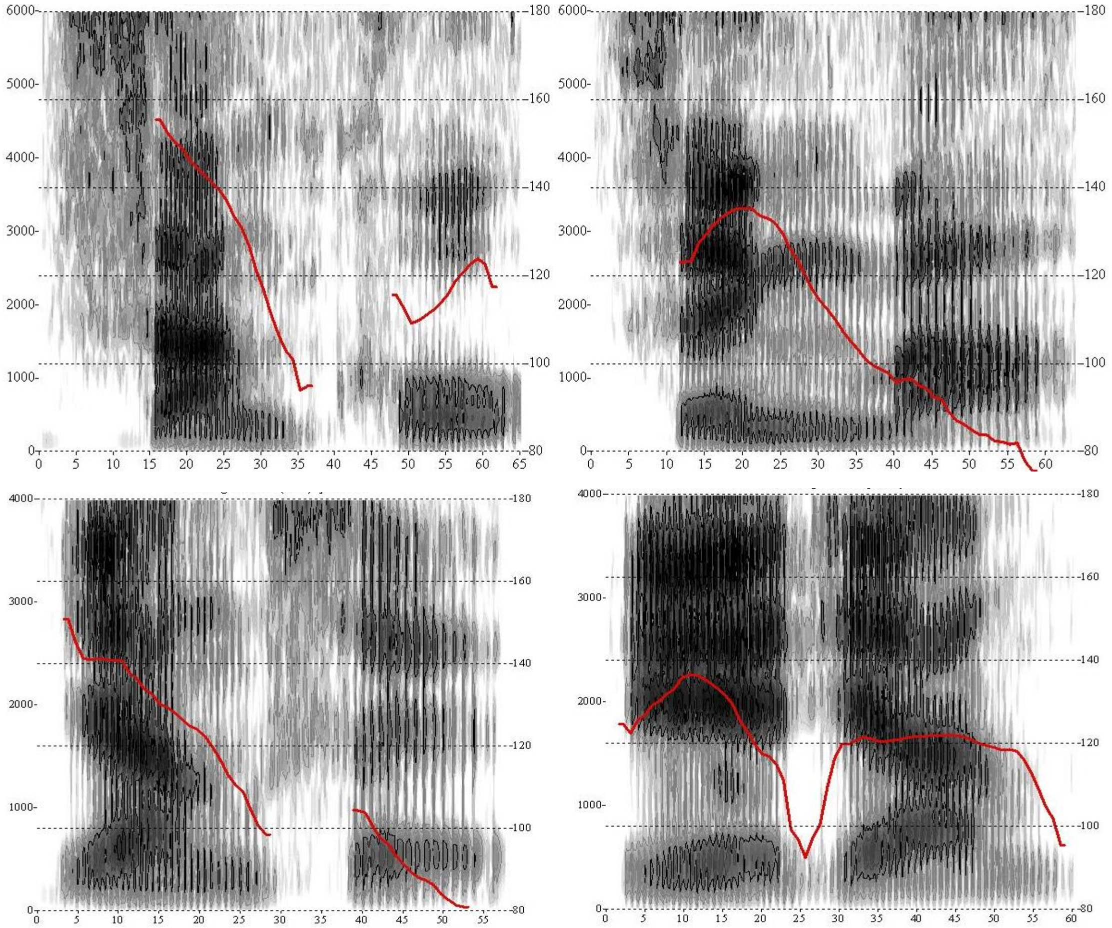


The configuration in the left-hand panel of figure 12 is simple. The transformed F0 shapes corresponding to the four level pitches are clear and fairly evenly spaced, and it can be seen that the F0 on the upper-mid level and lower-mid level isolation tones corresponds closely to the F0 on the first syllable tones denoted [44] and [22] (if these were plotted as a function of equalised duration the agreement would be considerable). The similarity is deceptive, however: the sandhi upper- and lower-mid level tones do not necessarily realise the same morphotoneme as the isolation upper- and lower-mid level tones: the upper-mid sandhi tone reflects the lower-mid level tone < IIIa*, not the upper-mid isolation tone (see table 5, panels A and C).

The transformed F0 shapes corresponding to the high falling and depressed falling shapes are also clear. The first-syllable data is simple in another sense, however. The first-syllable sandhi results in a much more phonetically transparent relationship between tonal pitch and the two *Yin* and *Yang* morpheme classes introduced in section 2.6 than obtains for the isolation and word-final shapes shown in table 4. Thus the high fall [52], high level [55] and upper-mid level [44] first-syllable tones participate in the realisation of *Yin* morphemes – i.e. those with invariantly voiceless Onset obstruents and upper-mid level, high rise, low rise and lower-mid level isolation tones. *Yang* morphemes – i.e. those with Onset obstruents with alternating voiced and lenis voiceless realisations and lower-mid level, depressed upper-mid level, mid and low fall-rise isolation tones – are realised by the depressed high fall [342], the lower-mid level [22] and the low level [11] first-syllable tones.

It is important to emphasise again that the depressed or lower pitch realisations on word-initial *Yang* morphemes are not causally related to phonetic obstruent voicing, since their obstruents lack phonetically voiced allophones word-initially. Moreover, as already pointed out, even morphemes with sonorant Onsets show the same tonal differentiation with respect to sandhi depending on whether they are *Yin* or *Yang*. Word-initially, therefore, one can find for example the contrast between high falling and depressed high falling tones on syllables regardless of whether the Onset is phonetically voiced or voiceless. The acoustics in figure 13 (wide-band spectrogram with superimposed F0) help to make this clear. The top row shows the disyllabic words /saŋkʰu/ [52.23] *tough* 辛苦 and /zɛŋma/ [342. ↓21**]** *vein* 靜脈, both beginning with alveolar fricatives. The first morpheme in /saŋkʰu/ is *Yin*, with an upper-mid level [44] isolation tone which, before the high rising morphotoneme of the second syllable, corresponds to the observed high falling [52] sandhi tone. The first-syllable morpheme in /zɛŋma/ is *Yang*. It has a depressed upper-mid level [344] tone in isolation which corresponds to the observed depressed high falling sandhi tone. Its Onset fricative /z/ is realised, word-initially, as [z̥]. It can be seen that both word-initial fricatives in these words are voiceless (and display the typical realisations already demonstrated above in figure 7), and so the depressed falling tone on the first syllable of /zɛŋma/ cannot be a result of phonetic voicing on the obstruent.

Figure 13. *Spectrograms, with superimposed F0, illustrating independence of depression in word-initial high falling tones from phonetic voicing on syllable Onset segments. Top row = words with Onset obstruents (left = /saŋ kʰu/ [52.34]; right = /zɛŋ ma/ [342.21]. Bottom row = words with Onset glides (left = /jaŋsø/ [52.21]; right = /ɥødʑaŋ/ [342.44]). Vertical axes = frequency (Hz), with F0 (right) and spectrographic (left); horizontal axis = duration (csec.).*



The bottom row of figure 11 shows the disyllabic words /jaŋsø/ [52.↓21] *print* 印刷 and /ɥødʑaŋ/ [342.44] *distance* 遠近, which both have glide Onsets. The first word /jaŋsø/ has a high falling pitched sandhi tone on the first syllable which reflects its *Yin* morpheme's [22] < *IIIa isolation tone. The second glide-initial word /ɥødʑaŋ/ has a depressed high falling pitch on its first syllable, which reflects its *Yang* morpheme's isolation depressed upper-mid level tone. The difference in tonal realisation between high falling and depressed falling obtains despite the agreement in glide Onset voicing.

5.4 Short first-syllable tones

It remains to discuss the short [2̲] and [3̲] first-syllable tones. These can be seen in the right-hand panel in figure 12. It will be recalled from table 6 that short first-syllable tones occur on morphemes which have isolation low rise or low fall-rising tones. Figure 10 shows that they have about half the Rhyme duration of the long tones. The higher of the two short tones, transcribed [3̲], occurs when either the low rise or low fall-rise occurs before morphemes with low rise, mid fall-rise or low level (<IIIa) tones. The lower of the two

short tones, transcribed [2], occurs before the remaining morphotonemes. Thus it appears that the low rising and low fall-rising tones are neutralised to a short tone on the first syllable, its pitch height – [2] or [3] – being conditioned by the following tone. The right-hand panel of figure 12 shows mean transformed F0 for these two first-syllable short tones, together with the low rising and low fall-rising isolation tones to which they are morphophonemically related. The isolation high rising tone F0 is also shown to indicate the upper limit of the isolation tonal F0 range.

It will be noted that in the right-hand panel of figure 12 *four* short F0 shapes have been plotted corresponding to the *two* short tones [2] and [3]. Linestyles have been chosen to highlight the relation with the isolation forms: solid lines represent the short tone F0 from morphemes with low rising isolation tones, thus the higher solid line represents the mean F0 of *Yin* morphemes with the low rising tone occurring before low rising, mid fall-rise and low level (<IIIa); and the lower solid line represents the mean F0 of the low rising tone before the other isolation tones. Dotted lines represent the short tone F0 from *Yang* morphemes with the low fall-rise isolation tone: the higher dotted line represents the mean F0 of morphemes with the low fall-rise tone before low rising, mid fall-rise and low level (<IIIa); the lower dotted line represents the low fall-rise tone before other isolation tones.

The configurations for both the [3] and [2] tones are very similar, and separated by a small (but statistically significant) amount of about 6 Hz. In both cases, over the ca. 10 csec. of their Rhyme duration, which is quite short relative to the isolation tones to which they are related, the mean F0 of the forms corresponding to the isolation low rising tone (solid lines) falls to converge with those of the corresponding isolation low fall-rise tone (dotted lines) from onsets differing by ca. 8 Hz.

The plotting of two F0 shapes for both tones – one dotted, one solid – was done to demonstrate the first-syllable neutralisation of the low rise and low fall-rising tones. It can be seen from figure 10 that the magnitude of the difference between the F0 shapes corresponding to the low rise and low fall-rise tones is too small, and the duration involved too short, to be tonal. Furthermore, these short F0 shapes also look very much like perturbations induced by the Onset consonants involved. Recall that the low rising tone belongs to *Yin* morphemes and therefore has a different set of Onsets from the *Yang* morpheme low fall-rising tone, and indeed the same kind of F0 patterning observed in the short tones can be seen at the onset of their corresponding isolation tones in figure 12, although it is of a shorter extent. I think, therefore, that the best interpretation of these acoustics is to assume that the tonal difference has been neutralised, but the difference between the morpheme classes has been preserved by the different F0 perturbations associated with the morpheme classes' Onsets. It is worth contrasting this situation with that on the long first-syllable tones. Here the differences in F0 shapes – for example between the high falling [52] and depressed high falling [342] tones – represent the *same* difference between *Yin* and *Yang* morphemes with their typical Onsets, but in this case they *are* big enough to constitute tonal differences.

The acoustics in the right-hand panel of figure 12 appear to give some support for an analysis of the short first syllable tone as a truncated version of the isolation tone. Which part of these low rising and low fall-rising tones is truncated (mora? autosegmental tone?) depends on how the tones are formally represented. It is also worth mentioning that a synchronic truncation interpretation would be exactly the opposite of the historical facts, since the low rising and low fall-rising isolation tones were actually historically short and ended in a glottal stop – it is the *Ersatzdehnung* of the short tones following the loss of the glottal-stop which constitutes the unusual innovation guaranteeing the Oujiang dialects its subgroup status within Wu.

5.5 Summary – first-syllable tones

The impressionistic and acoustic descriptions in the preceding sections on the first-syllable tones have shown the typically complicated apparent merger pattern of isolation tones predicted of right-dominant systems. It has made clear that the complication resides mainly in the phonetic opacity of the subsets of tones that appear neutralised; the subsets of second-syllable tones that condition the neutralization; and the result of the neutralization, which is usually not one of the isolation tones. What may pass for phonetic tonal behavior is only sporadically glimpsed, and there is little biuniqueness. It is probably best, *pace* Lass (1984: 46, 47), not to characterize this messy relationship between isolation and first-syllable tone contrast as neutralization, which was after all originally restricted to bilateral oppositions (Trubetzkoy 1939).

One wonders in fact if any relationship at all should be posited between first-syllable and isolation or word-final tones, especially given the simplicity of the first-syllable tone system if considered independent of such a relationship. For with one exception, all long first-syllable tones – [55 44 52 22 11 342] – can be analysed, using conventional critera of phonetic similarity and complementary distribution, as realisations of a tonemic system consisting of just /level/ and /fall/, conditioned by the following tone and the *Yin/Yang* morpheme class[8]. An even neater analysis would be possible in the prosodies and phonematic units of an old-style Firthian terms (Palmer 1970, Lass 1984: Ch. 10). *Yin*, *Yang* and *Short* would then be prosodies governing the co-occurrence of Onset phonation, Rhyme duration and pitch register/depression, leaving just two phonematic tonal units: /level/ and /fall/.

But perhaps we focus too much on sychnronic system and it is more fruitful to ask how things got like that. This first-syllable tonal configuration, unlike the isolation or word-final tones, was shown, for the long tones, to clearly reflect the relationship between tonal register/depression and the *Yin*/*Yang* morpheme type. Upper register and non-depressed first-syllable tones [55 44 52] belong to *Yin* morphemes; lower register and depressed first-syllable tones [22 11 342] are *Yang*. Given that such a relationship is assumed to be the historical antecedent, it is likely that the first-syllable tones reflect an earlier state of affairs than the isolation and word-final tones, which comparative evidence (e.g. Rose 2008, 2011) shows have probably changed more in the interim. Reflect history, but do not preserve it. The distribution of first-syllable tones in table 6, and some of their values, is reminiscent of a once more differentiated array that has simplified by the spread of a few tones. There is no high falling tone isolation tone in these data, but other Oujiang sites have a high fall for reflexes of IIIa. In the first-syllable high fall/depressed fall Wencheng 'sandhi tone' one may be glimpsing a bit of the past tonal history of Oujiang: it would be worth testing the validity of a reconstruction along these lines.

6. Closing Remarks: Caveat, Summary, Discussion

6.1 Caveat

This paper has been a targeted analysis in the incipient era of Big Data, so a caveat is probably in order before its summary and discussion. The clearly restricted nature of the data – one speaker, one occasion – means of course that the issue of generalizability arises. The results of over a decade of forensic voice comparison testing have shown that, in reality, all speakers are regularly idiosyncratic to the extent that same-speaker speech samples can be discriminated rather well from different-speaker speech samples on the basis of between- and within-speaker variation in all sorts of features (e.g. Rose 2005, 2013). So, of course some of the speaker's behaviour described above, despite its consistency, may be idiosyncratic, and thus possibly atypical, either of the particular socio-economic group of Wencheng to which he belonged[9], or of Wu right-dominant tone sandhi. And he will be guaranteed to vary in this behaviour on different occasions[10].

Fortunately, comparable data exists from other Oujiang sites that reveal at least some of the speaker's sandhi typicality. Ballard, for example, recorded the same corpus from speakers from four other counties in the Oujiang area, mostly quite far from Wencheng, which I have processed similarly: Pingyang 平陽,

[8] For example, first-syllable [11] would be an allotone of /level/ on *Yang* morphemes before a word-final [22] or [44]; [22] would be another first-syllable allotone of /level/ on *Yang* morphemes before a word-final [324]. First-syllable [52] and [342] would be allotones of /Fall/ on *Yin* and *Yang* morphemes respectively. The one, messy, exception is provided by the contrast between [55] and [44] before [31] (panel D, table 5 & figure 10), which cannot both be realisations of /level/ in *Yin* morphemes before a high tone.

[9] A possible candidate for this is the characteristic prenasalisation noted for Wencheng *Yin* morpheme /p/ and /t/ stops in Zhu (1996: 243). Since no more information is given, its clear absence in this speaker cannot be elucidated.

[10] He did, in fact. One example of within-speaker variation occurred with combinations with word-final morphemes etymologically carrying the low rising [24] tone. These would be expected to have a word-final [12] realization (table 5, panel G), but six occurrences were said with the same low falling [21] pitch as the low fall-rise morphotoneme in panel H. Examples are /bɑ fuo/ 342.21 辦法 *method* < /bɑ 22/ + /fuo 24/; and /pɛæ tʰɑ/ 52.21 寶塔 *pagoda* < /pɛæ 45/ + /tʰɑ 24/. This variation may be heard on the examples on my webpage; there is some evidence that it is free. The speaker recorded the word *hair* 頭髮, for example, as diu fuo 342.**21** with a low falling tone on the second syllable, but the word is transcribed by Ballard in a separate session eliciting body part names with a low rising tone. The putative free-varying morphemes involved all have [+ spread glottis] onset consonants, but since there is only a small number of occurrences, it is not clear whether this observation constitutes the correct generalization over which morphemes show the free variation.

Yongjia 永嘉, Xiangyang 象陽 and Wenzhou 溫, the latter already described in Rose (2000, 2002a, 2002b, 2004). I am also lucky in having access to extensive recordings of disyllabic sandhi in several other Oujiang sites made by Prof. Zhu Xiaonong in the mid-nineties. Many important aspects of the tone sandhi behaviour demonstrated for the Wencheng speaker above can be found in speakers from these other sites. I mention just three:

- the first-syllable realisation of reflexes of Middle Chinese IVa/b tones as a single short tone, its pitch height conditioned by the following tone;
- the word-final allotones most resembling the isolation tones occur after the short tones;
- the massive neutralisation of first-syllable tones to high falling/depressed falling tones depending on the *Yin/Yang* class of morpheme involved.

One aspect of the speaker's sandhi that does not occur in the Oujiang varieties I have examined so far is the separate pitch shapes for words with reflexes of the Middle Chinese IIIa (Yinqu) tone before IIIa/b (table 5 & figure 10 panels C, D, solid and dotted lines). In the other varieties these have the same shapes as words with IIa on the first syllable.

6.2 Summary and Discussion

This paper set out firstly to document right-dominant tone sandhi in a previously undescribed variety of Wu. Secondly, it aimed to explore with (acoustic-) phonetic analysis theoretically important typological aspects of the sandhi, like preservation and neutralisation of tone, by examining in detail the relationship between isolation tones and tones in disyllabic words. It has shown a generally clear, unproblematic, relationship between the word-final tones and their corresponding isolation forms, with differences between them being explainable by phonetically plausible processes like co-articulation with the first-syllable tones. On the other hand, the relationship between isolation and first-syllable tones is byzantine. Complicated by an almost total lack of phonetic motivation, it is also generally not the case that a particular first-syllable tone corresponds to a specific isolation tone, since some complicated mergers obtain, non-phonetically conditioned by the following tone. Moreover, some first-syllable tones do not have a corresponding isolation tone – the lack of a falling isolation tone, for example.

Phonological models since SPE have been biased towards production, focussing on how an underlying form might surface. But the complexities of the right-dominant sandhi apply to perception too. In a word with [11.44] tonal pitch, the [11] before the [44] uniquely identifies the first-syllable morpheme as having a lower-mid level <Ib tone (table 5 panel A). In a word with the same first-syllable [11] pitch before [22], however, the lexical search is only narrowed down to a word-initial morpheme having one of three isolation tones – either a lower-mid level <Ib, depressed upper-mid level, or mid fall-rise (table 5 panel C).

The description of tones in disyllabic words presented in the paper reveals a problematic dichotomy. The straightforward relationship between isolation and word-final tones means that the shapes of the former can be plausibly accounted for by synchronic rule- or constraint-based operations on the isolation forms: this is a normal, non-problematic phonological model. But the complicated relationship between isolation and first-syllable tones makes it sensible, I think, to question the extent to which the first-syllable tones should be modelled as synchronic rule- or constraint-governed behaviour operating on their corresponding isolation forms. As suggested, a more realistic analysis of the first-syllable tones might be simply to jettison the *unique underlier condition* and make their underlying forms as abstract as a tonemic (or phonematic), as opposed to a morphotonemic, representation.

As with many Wu dialects, it is only possible to see a systematic relationship between isolation and and first syllable tones by invoking historical tone categories. Representing the tones by their Middle Chinese (MC) historical tonal categories gives a much neater picture, with the categories acting as natural classes. For example the merger and realization patterns for first-syllable tones described above correlate with the MC natural tonal classes of which the word-final tones are reflexes: the word-final tones in table 5 sections A & B are reflexes of MC *ping/I tones; in sections C & D reflexes of MC *Qu/III; and in sections E – H reflexes of MC *shang & *ru/ II & IV. This is why the lower-mid level tone, coming as it does from two different MC categories, appears to evince the differential first-syllable behaviour observable in section B, where it is a reflex of *Ib; and in section C, where it is a reflex of *IIIa.

Of course, this explains nothing synchronically. It has been assumed, however, that the path from historical coherence to synchronic phonetic opacity in Wu is due to the isolation tones historically changing more than the sandhied tones in polysyllabic words (Rose 2011: 42). But there must be more to it than that. Diachronic changes do not generally result in a synchronic mayhem imposing such a cognitive burden in production and perception for no reason. The explanandum is why there has been a shift to a more complex position in phonological hyperspace. Perhaps the demarcative function of tone sandhi in Wu, which clearly signals a phonological word, has provided a perceptual advantage which has come to outweigh the lexical function of tone. Additional perceptual cues to parsing a phonetic input are of great use in L1 acquisition, for example. This might provide a competing OT-theoretic constraint to help motivate both diachronic developments and synchronic patterns.

7. References

Abercrombie, D. 1967. *Elements of General Phonetics*. Edinburgh University Press, Edinburgh.

Ballard, W.L. 1980. On some aspects of Wu tone sandhi. *Journal of Asian and African Studies* 19: 83-164.

Ballard, W.L. 1984. Wu, Min and a little Hakka – tone sandhi: right and left. *Cahiers de Lingistique Asie Orientale* 13: 3-34.

Ballard, W.L. 1992. Oujiang Wu tone sandhi: Visi-Pitch Results. *Chinese Language and Linguistics 1: Chinese Dialects*. *Symposium Series of the Institute of History and Philology Academia Sinica* 2, Taipei: 41-66.

Cao Z. 曹志耘 2002. 南部吴语语音研究. [Studies on the phonetics of Southern Wu dialects]. Shangwu Yinshuguan, Peking.

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

Chao Y. R. 1930. A system of tone letters. *Le Maître Phonétique* 45: 24-27.

Chao Y. R. 1967. Contrastive aspects of the Wu dialects. *Language* 43/1: 92-101.

Chen, M. 2000. *Tone Sandhi - Patterns Across Chinese Dialects*. Cambridge University Press, Cambridge.

Fu G. 傅国通, Fang S. 方松熹, Cai Y. 蔡勇飞, Bao S. 鲍士杰, Fu Z. 傅佐之 1985. 浙江吴语分区 [Wu dialect subgroups of Zhejiang]. Zhejiang Linguistics Society, Hangzhou.

Howie, J. 1974. On the domain of tone in Mandarin. *Phonetica* 30: 129-148.

IPA Guide. 1999. *Handbook of the International Phonetic Association.* Cambridge University Press, Cambridge.

Kennedy, G.A. 1953. Two tone patterns in Tangsic. *Language* 29/3: 367-373.

Kratochvil, P. 1968. *The Chinese Language Today*. Hutchinson, London.

Lieberman , P. 1967. *Intonation, Perception, and Language*. M.I.T. Research Monograph 38. M.I.T. Press, Cambridge Mass..

Lass, R. 1984. *Phonology*. Cambridge University Press, Cambridge.

Laver, J. 1994. *Principles of Phonetics*. Cambridge University Press, Cambridge.

Matthews, P.H. 1970. Recent developments in morphology. In J. Lyons (ed.) *New Horizons in Linguistics*. Penguin, London: 96-114.

Norman, J. 1988. *Chinese*. Cambridge University Press, Cambridge.

Nolan, F. 2003. Intonational equivalence: an experimental evaluation of pitch scales. *Proc. 15th Int'l. Congr. Phonetic Sciences*: 771-774.

Palmer, F.R. 1970. *Prosodic Analysis*. Oxford University Press, Oxford.

Pan W. 1991. An introduction to the Wu DIALECTS. In Wang S. (ed.) *Languages and Dialects of China*. Journal of Chinese Linguistics Monograph 3: 237-293.

Pierrehumbert, J. 1979. The perception of fundamental frequency declination. *JASA* 66/2: 363-369.

Rose, P. 1982a. An Acoustically-based phonetic description of the syllable in the Zhenhai dialect. Unpublished Ph.D. thesis, University of Cambridge.

Rose, P. 1982b. Acoustic characteristics of the Shanghai-Zhenhai syllable-types. In D. Bradley (ed.) *Tonation.* Pacific Linguistics Papers in South East Asian Linguistics 8: 1-53.

Rose, P. 1989a. On the non-equivalence of fundamental frequency and pitch in tonal description. In D. Bradley, E. Henderson and M. Mazaudon (eds.) *Prosodic Analysis and Asian Linguistics: to Honour R.K.Sprigg.* Pacific Linguistics: 55-82.

Rose, P. 1989b. Phonetics and phonology of the yang tone phonation types in Zhenhai**.** *Cahiers de Linguistique Asie Orientale* 18/2: 229-245.

Rose, P. 1990. Acoustics and phonology of complex tone sandhi. *Phonetica* 47: 1-35.

Rose, P. 1998. The Differential status of semivowels in the acoustic phonetic realization of tone. *Proc. 5th Intl. Conf. on Spoken Language Processing* 2: 61-64.

Rose, P. 2000. Wenzhou dialect disyllabic lexical tone sandhi with first syllable entering tones. *Proc. 8th Australian Intl. Speech Science and Technology Conf.*: 230-235.

Rose, P. 2001. Independent Depressor and Register Effects in Wu Dialect Tonology: Evidence from Wenzhou Tone Sandhi. *Journal of Chinese Linguistics* 30/1: 39-81.

Rose, P. 2002. Tonal complexity as conditioning factor: more depressing Wenzhou dialect disyllabic tone sandhi. *Proc. 9th Australian Int'l. Speech Science and Technology Conf.*: 64-69.

Rose, P. 2004. "Defying explanation?" – Accounting for tones in Wenzhou dialect disyllabic lexical tone sandhi. *Proc. 10th Australian Int'l Conf. on Speech Science & Technology*: 237-242.

Rose, P. 2005. Technical forensic speaker recognition: evaluation, types and testing of evidence. *Computer Speech and Language* 20/2-3: 159-191.

Rose, P. 2008. Oujiang tones and acoustic reconstruction. In Bowern, Evans, Miceli (eds.) *Morphology and Language History*. Current Issues in Linguistic Theory 298: 235-250.

Rose, P. 2010. The long and the short of Wencheng tones – acoustic and auditory description of tonologically challenging phenomena in an Oujiang Wu dialect of Chinese. *Proc. 10th Australasian Int'l. Speech Science & Technology Conf.:* 54-57.

Rose, P. 2011. Reverse-engineering tones and tone sandhi in Wenzhou dialect: between-speaker differences and historical development. *Int'l. Bull. Chin. Linguistics* 5/2: 29-48.

Rose, P. 2013. Where the science ends and the law begins: likelihood ratio-based forensic voice comparison in a $150 million dollar telephone fraud. *Int'l J of Speech Language and the Law* 20/2: 277-324.

Rose, P. 2014. Transcribing tone – A likelihood-based quantitative evaluation of Chao's tone letters. *Proc. 2014 Interspeech*: 101-105.

Rose, P. 2015. Tonation in Three Chinese Wu Dialects. *Proc. 2015 Int'l Cong. Phonetic Sciences* (no page numbers).

Sherard, M. 1980. *A Synchronic Phonology of Modern Colloquial Shanghai.* Computational Analyses of Asian & African Languages Monograph 5.

Steed, W. 2006. Phonation type and tone sandhi as evidence of Chinese stress in Longquan Wu. In C. Warren and P. Watson (eds.) *Proc. 11th Australian Int'l. Speech Science & Technology Conf.:* 82-87.

Steed W. 2010. Lishui Wu Tone and Tone Sandhi: An Acoustically-Based Description. Unpublished Australian National University Ph.D. thesis.

Steed, W. 2012. Qingtian Wu lexical tone sandhi: Voiceless depressors and allotones. In C. Donohue, S. Ishihara and W. Steed (eds.) *Quantitative approaches to problems in linguistics*: 21-32. Lincom Studies in Phonetics 8. Lincom, Munich.

Tatham, M.A.A. 1971. Classifying allophones. *Language and Speech* 14: 140-145.

Trubetzkoy, N.S. 1939. *Grundzűge der Phonologie*. Travaux du Cercle Linguistique de Prague 7, Prague.

Xu Y. 1997. Contextual tonal variation in Mandarin. *Journal of Phonetics* 25: 61-83.

Zhang J. 2007. A directional asymmetry in Chinese tone sandhi systems. *J. East Asian Linguistics* 16: 259-302.

Zhang J. 2009. Contour tone distribution is not an artifact of tonal melody mapping. *Studies in the Linguistic Sciences* 33 /1, 2: 73-132.

Zee, E. & Maddieson, I. 1979. Tones and tone sandhi in Shanghai: phonetic evidence and phonological analysis. *Glossa* 14: 45-88.

Zhu L. 朱礼 1996. 文成县志 [Wencheng County Annals]. Zhonghua Shuju 中华书局.

Zhu X. 朱晓农 2004. 上海声调实验陆. *An Experimental Study in Shanghai Tones*. Shanghai Educational Publishing House, Shanghai. [Published in English in 1999 as *Shanghai Tonetics*, Lincom Studies in Asian Linguistics 32. Lincom, Munich].