



# Yuhuan Wu tone sandhi and tone contrast maintenance

Carlos Gussenhoven<sup>1</sup>, Lu Wang<sup>2</sup>, Louis Hendrix<sup>3</sup>

<sup>1</sup> Department of Linguistics, Radboud University Nijmegen, Netherlands

<sup>2</sup> School of Foreign Languages, East China University of Political Science and Law, Shanghai

<sup>3</sup> School of Interdisciplinary Area Studies, University of Oxford, UK

c.gussenhoven@let.ru.nl 1.1.wang@outlook.com ljchendrix@gmail.com

## Abstract

Yuhuan Wu Chinese deletes non-final lexical tones within a small post-lexical tone domain, except HL if its syllable begins with a sonorant consonant. The explanation of this phonological specification lies in the number of tone contrasts in the location concerned (H, L, HL, ML and LH), which is higher than in syllables with a voiceless obstruent or zero onset (H, HL, LH) or in syllables beginning with a voiced obstruent (L, ML, LH). The interpretation as contrast maintenance is supported by features of two phonetic implementation rules which enhance tone contrasts. One enhances the incompleteness of the neutralization of tones in pre-final position, while the other enhances the ML-HL contrast in phrase final position.

**Index Terms:** Chinese, Wu, Yuhuan, tone sandhi, speaker control, incomplete neutralization, phonetic implementation

## 1. Introduction

What may at first sight look like arbitrary exceptions to phonological rules often turn out to have explanations in functional factors governing the process of vocal communication. Contrast maintenance is a major such factor. Phonetic implementation typically aims to produce signals that are distinct from signals that compete with them acoustically [1]. This factor is pitted against ergonomic considerations affecting articulation, which promote shortcuts and deletions [2,3]. This contribution singles out an exception to a tone deletion rule whose explanation will be argued to lie in a mitigation of the effects of ergonomically motivated tone deletions on discriminability of lexical items. The case concerns generalization (1) about Yuhuan Wu tone sandhi.

- (1) Pre-final lexical tones are deleted, except HL when it appears in a syllable with a sonorant onset.

Support for the explanation of (1) in terms of ‘damage limitation’ will be presented with the help of data that reveal a similar concern for the loss of lexical distinctions in the phonetic implementation. Yuhuan Wu is spoken by about 300,000 speakers in Yuhuan (玉环, Yu<sup>o</sup>huan<sup>H</sup>, 28°08'09" N;121°13'35" E), Taizhou Prefecture, Zhejiang Province, China.

## 2. The Yuhuan Wu syllable

The segmental structure of the syllable is given first, after which we state the distribution of the tones as a function of the segmental structure.

The syllable structure is (C)(G)V(C). The obligatory element V can be filled by any of the vowels in Table I if there is no

coda. In addition, syllabic /z/ occurs after onset /s, z/, while syllabic /ŋ/ occurs more generally.

Table I. Yuhuan Wu vowels in open syllables

	front unround	front round	back unround	back round
high	i	y		u
mid	ɛ	ø		o
diphthong	ei			əu
low			a    ǎ	ɔ    ǒ

Some of these vowels may appear before a coda, along with /ə/, which exclusively appears before codas. Of the consonants that are listed in Table II, only /ʔ/ and /ŋ/ can occur in the coda. The vowels /i o ə/ may appear before them, and additionally /ɛ, ø/ occur before /ʔ/. All consonants occur as onsets, except /ʔ/.

Table II. Yuhuan Wu consonants

	labial	coronal	velar	glottal
plosive	b p p <sup>h</sup>	d t t <sup>h</sup>	g k k <sup>h</sup>	ʔ
affricate		dz ts ts <sup>h</sup>		
nasal	m	n	ŋ	
fricative	v f	z s	x	ɦ
lateral		l		

The way onsets and rimes are segmentally filled determines the distribution of the lexical tones. There are two such factors. One concerns the segmental composition of the rime, and the other that of the onset.

*The rime factor.* Fully sonorant rimes, with or without prevocalic glide, maximally have a five-way tone contrast. Glottal rimes, again with or without prevocalic glide, maximally have a two-way contrast. The lexical tones appearing in sonorant rimes are H, L, HL, ML and LH, while glottal rimes have H or L, henceforth notated as H<sub>q</sub> and L<sub>q</sub> to indicate their occurrence in glottal rimes. These contrasts are shown in Table III. The explanation for the reduced number of contrasts in glottal rimes is the short duration of the voiced portion of glottal rimes [4]. Sonorant portions of glottal rimes average 122 ms, against 438 ms for sonorant rimes, calculated over phrase-final five tokens with H-tone.

### The onset factor

The second factor affecting the number of lexical tone contrasts is the nature of the onset. Obstruent onsets, which include the plosives, affricates and fricatives shown in Table II, occur in syllables with a three-way tone contrast in sonorant rimes and no contrast in glottal rimes. Sonorant onsets occur in syllables with the full set of five tones in sonorant rimes and the full set of two tones in glottal rimes. Table IV shows the contrasts in sonorant rimes with obstruent onsets. Voiceless obstruents, either aspirated or unaspirated, co-occur with H,

HL and LH, while voiced obstruents, located in leftmost position in the cells of Table II, with L, ML and LH.

Table III. Tone contrasts in sonorant and glottal rimes

miɔ	H	<i>glance</i>
miɔ	L	<i>temple</i>
miɔ	HL	<i>second (unit of time)</i>
miɔ	ML	<i>depict</i>
miɔ	LH	<i>seedling</i>
loʔ	Hq	<i>sway</i>
loʔ	Lq	<i>fall down</i>

Table IV. Tone contrasts in sonorant and glottal rimes after obstruent onsets

p <sup>hi</sup>	H	<i>fart</i>
p <sup>hi</sup>	HL	<i>drape over shoulder</i>
p <sup>hi</sup>	LH	<i>semi-finished product</i>
pɔ	H	<i>leopard</i>
pɔ	HL	<i>precious things</i>
pɔ	LH	<i>bag</i>
bɔ	L	<i>carpenter's plane</i>
bɔ	ML	<i>hug</i>
bɔ	LH	<i>robe</i>
pəʔ	Hq	<i>give</i>
bəʔ	Lq	<i>white</i>

The explanation of the greater tolerance of tone contrasts after sonorants is their low degree of interference with the glottal airflow. They have a constriction in the vocal tract that allows air pressure levels on either side to be relatively equal. Despite the somewhat higher air pressure behind the oral constriction, the generous aperture left by the constriction prevents air pressure from building up behind it and thus the concomitant impedance of the airflow through the glottis. Voiced obstruents, by contrast, have a constriction behind which the air pressure builds up, and vocal fold vibration during the constriction will be counteracted by the impedance created by the high oral air pressure above the glottis. During voiceless obstruents, the vocal folds are abducted to prevent voicing. When voicing sets in after them, the tenser vocal folds will tend to vibrate faster than after the voiced sonorants (cf. [1]). Tones after voiced obstruent onsets are therefore predicted to be lower than after voiceless onsets. This tendency is reflected in the tone distribution in Table IV, where voiced obstruents appear before L, Lq and ML, but voiced ones appear before H, Hq and HL. Only LH occurs after both obstruent types.

Syllable-initial phonetic glides [j, w], indicated by (G) above, are interpreted as prevocalic /i u/ occurring in the nucleus. Prevocalic /i/ appears before /u ε ø ɔ α ã ɔŋ/ and may occur after all consonants except /f v x ŋ/. Prevocalic /u/ appears before /ε ei ð α ã ŋ oʔ/ and may occur after /g k k<sup>h</sup> ŋ x fi/. Prevocalic [j w] either start from weakly released glottal closures, [ʔ<sup>j</sup> ʔ<sup>w</sup>], or have weak breathy voice without initial glottal closure. Following [5], we interpret the contrast as onsetless vs. /fi/, as shown in Table V for prevocalic /u/. The distribution of the tones provides the motivation for analyzing [j w] as /i u/ inside the nucleus of the syllable, as opposed to

taking the glides to be onset consonants /j w/, after which the five-way contrast would have been expected. Voiced /fi/ shares the distribution with voiced obstruents, while onsetless syllables behave like syllables with voiceless onsets.

Table V. Tone contrasts in onsetless and /fi/-initial rimes

uei	[ <sup>j</sup> wei]	H	<i>fed up</i>
uei	[ <sup>j</sup> wei]	HL	<i>bowl</i>
uei	[ <sup>j</sup> wei]	LH	<i>hello</i>
fiuei	[wei]	L	<i>meeting</i>
fiuei	[wei]	ML	<i>refuse</i>
fiuei	[wei]	LH	<i>proper name</i>

### 3. Tones in phrase-final sonorant rimes

We recorded isolated syllables with all five tones and onset and coda conditions as illustrated in Tables III and IV from three speakers. Two are the father (FW) and mother (MW) of the third, female, speaker (DW). Average time-normalized f0 tracks of the five tones in sonorant rimes spoken by FW, MW and DW are given in Fig. 1, pooled over onset conditions.

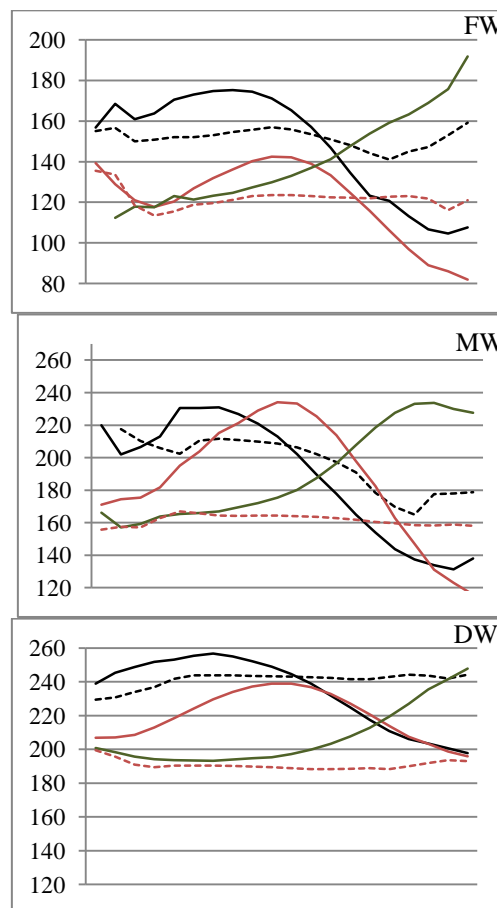


Figure 1: F0 tracks on normalized times scales for H, L, HL, ML and LH for FW, MW and DW. N varies from 2 to 56. — = HL, - - - = H, — — — = ML, - - - - = L, — — — = LH.

While the five contours are similar across the speakers, speaker MW aligns the peak of ML considerably later than the peak of HL. Speaker FW has a lower and earlier realization. The timing difference produced by speaker MW is the sole distinguishing feature between HL and ML. Interestingly, the realization of the HL-ML contrast by daughter DW looks like

a compromise between those by the parents: the alignment difference is greater and the peak higher for ML than in the case of speaker FW, but both features fall short of how speaker MW pronounces the contrast.

#### 4. Damage control of tone deletion

Data by DW reported in [5] showed three features in sequences of two tones in noun-noun and adjective noun combinations. One is that the final tone is preserved and the first deleted. Second, the pitch in the pre-final syllable, which was low to mid, was 15 to 20 Hz higher when the deleted tone was H or LH than when it was L or ML. Third, whenever a pre-final syllable whose onset was /m n ŋ l/ has a lexical HL, its tone was preserved. The data by FW and MW confirmed all three features. First, Fig. 2 shows mean  $f_0$  tracks of pre-final H, L, ML and LH separately before LH (speaker FW), before ML (speaker MW) and before L and Lq (speaker DW). All three speakers have higher pitch in the case of deleted H and LH than in the case of deleted L and ML. Strikingly, the enhancement carries over into the final syllable for all three speakers. The data for DW additionally show that deletion appears to be indistinguishable from the occurrence of L-tone and shows that tones are deleted in glottal rimes (or neutralized to L, in an alternative formulation). In [5], conclusions regarding the DW data were presented with some caution, since their systematic nature was based on a single speaker. The additional data allow an interpretation of this effect of the underlying contrast as a controlled property of the phonetic implementation, whose function is to preserve some of the functional load of the tones. This strongly supports the position taken by [1] that phonetic implementation is non-automatic and under speaker control. Their argumentation concerned enhancement of surface phonological contrasts with the help of ancillary phonetic cues in order to counteract an erosion of the distinctiveness of the primary feature involved in the expression of the contrast. What is different about Yuhuan Wu pitch raising for deleted H and LH is that phonological representations after the deletions are identical to those after deletion of L and ML. This suggests that cases of incomplete neutralization [6] should be conceptually merged with the notion of ‘phonetic knowledge’ introduced by [1]. Incomplete neutralization is a morphologically induced effect. Our data suggest that this effect may be enhanced in situations where the phonological deletion of a contrast leads to excessive loss of functionality, i.e., to some critical degree of reduced lexical discrimination.

As for the third feature in DW’s data, both FW and MW displayed the same exception to tone deletion. Whenever a sonorant rime with HL has a sonorant onset, its tone is unambiguously preserved in pre-final position, a systematic phonological exception to the tone sandhi rule. The motivation for the exception must again be sought in the preservation of lexical distinctions. Table VI lists the number of tone contrasts on pre-final sonorant rimes before and after deletion as a function of the major class features of the onset. After obstruents, three lexical tones are lost due to pre-final deletion. After sonorants, four are lost, but since these tones are enhanced two by two, the perceptual damage is to some extent mitigated. In the last column, this mitigation is estimated at 25%, such that the perceptual damage done by pre-final tone deletion is equalized across all words with sonorant rimes.

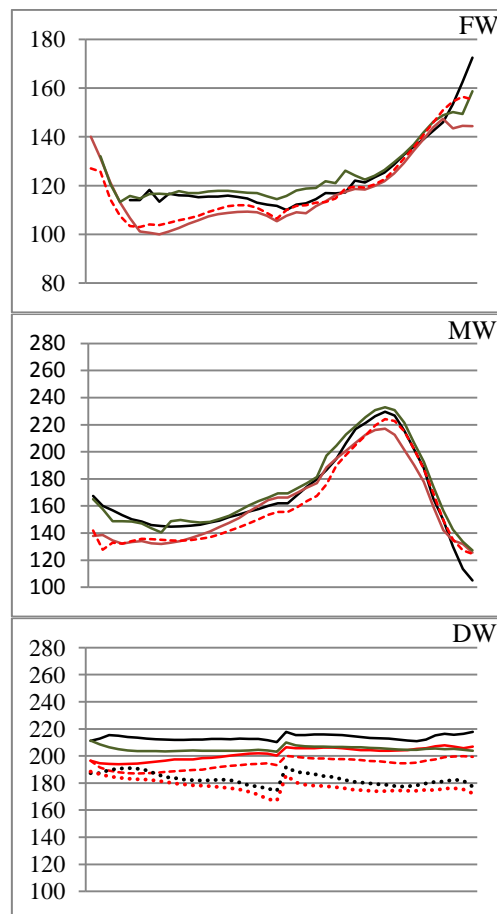


Figure 2:  $F_0$  tracks on normalized times scales for deleted H, L, ML and LH before other tones by speaker FW (before LH), speaker MW (before ML) and of H and Hq by speaker DW (before L and Lq).  $N$  varies from 4 to 8. — = H, — = LH, — = ML, - - - = L, ..... = Hq, ..... = Lq.

Table VI. Number of lexical tones in sonorant rimes as a function of the value for [sonorant] of the onset before and after deletion due to pre-final tone deletion, the number of deleted tones, any enhanced tones and estimates of the perceptual damage in terms of the number of lost lexical contrasts.

Onset	Lexical	Remaining	Deleted	Enhanced	Damage
+son	5	1	4	H, LH	≈3
-son	3	0	3	-	3

#### 5. Phrase-final phonetic enhancement

In this section, we point to a second case of phonetic enhancement, one that is in evidence in phrase-final syllables. Recall that sonorant rimes with sonorant onsets are the only syllables that contrast for the two falling tones in Yuhuan Wu, ML and HL, since voiceless or absent onsets only allow HL and voiced obstruent onsets only allow ML. As shown by [5] for speaker DW, pitch perturbations after these three classes of onset run counter to what may be expected on the basis of speech ergonomics. Sonorant onsets should have no effect on the  $f_0$  of the initial part of the rime, since their articulation barely creates any impedance to the air flow from the glottis, so that speakers can maintain the flow much as during vowels. After voiced obstruents, the  $f_0$  of the first part of the rime should be expected to be depressed due to the impedance

offered by the oral constriction and the lowered f0 that will result. After voiceless obstruents, the initial f0 after HL is typically raised relative to the situation after sonorant consonants, due to the tensing of the vocal folds when they are abducted to facilitate voicelessness. Counter to these expectations, speaker DW pronounces HL after sonorants at the same f0 as after voiceless obstruents, while the f0 of the peak for ML is lower after sonorants than after voiced obstruents, as shown in panel DW of Fig. 3. In [5], this finding was interpreted as enhancement of the HL-ML contrast. The support for that conclusion was somewhat precarious, because of the single speaker involved and the fact that the raising of HL after sonorant onsets did not exceed the f0 of HL after voiceless onsets. In panels FW and MW of Fig. 3, we display the results for the two new speakers. FW in fact produces higher f0 after sonorant onsets than after voiceless onsets, while producing lower f0 after sonorant onsets than after voiced onsets, thus showing the behavior described in [5] for DW in its fullest form. The results for speaker MW confirm the DW and FW data for HL: the f0 after sonorants is very much higher than after voiceless onsets, again showing the enhancement. The difference is created as much by f0 raising after sonorants as by f0 lowering after voiceless obstruents. At first sight, it is therefore odd to see that the f0 peak of ML is higher after sonorant onsets than after voiced obstruent onsets. This is what would be expected if the speaker pronounced ML after sonorants naturally, without the kind of speaker control required for the enhancement of the HL-ML contrast. Recall from section 3, however, that MW realizes the distinction between HL and ML in terms of alignment, not in terms of a difference in peak f0. ML is later than HL. Exactly this feature is enhanced by MW when she pronounces ML after sonorants: it is even later than after voiced obstruents. This confirms that alignment is what MW employs in order to distinguish between HL and ML, while also confirming enhancement of the contrast after sonorants, in a manner as outlined by [1].

## 6. Conclusion

The tone grammar of Yuhuan Wu reveals features in its phonology and phonetics which are to be interpreted as the limitation of damage which would be inflicted on lexical distinctiveness by a rule of tone sandhi if those feature were not present. The rule amounts to the deletion of non-final lexical tones in short phrasal domains, the mirror image of the well-known Shanghai version, where phonetic enhancement has been shown to target the three-way obstruent contrast in syllables with neutralized tone ([7]). In light of the fact that there are five lexical tones after sonorants, against three after obstruent or empty onsets, it is revealing that whenever the onset of a prefinal syllable is sonorant, one of the five tones, HL, is preserved. Second, the deleterious effect of the deletion of the other four tones, H, L, ML and LH, is mitigated in the phonetic implementation by raising f0 in syllables from which H and LH are deleted relative to syllables in which ML and L are deleted. This was characterized as enhancement of incomplete neutralization, a strategy which has not been diagnosed before, as far as we are aware. Outside the tone deletion context, we additionally reported enhancement of the phonological HL-ML contrast in phrase-final syllables, a strategy of the type first reported in [1]. These effects amply illustrate that speech production is a goal-directed activity under the control of the speaker, as claimed in [1], but there is no implication whatsoever that any of these aspects reach speaker awareness.

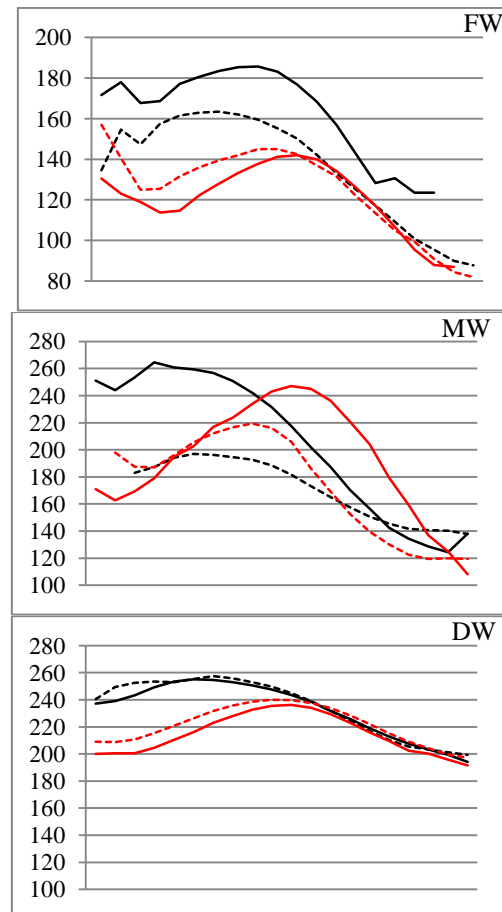


Figure 3. F0 tracks on normalized times scales for HL in sonorant and voiced obstruent nuclei and ML in sonorant and voiceless obstruent nuclei for speaker FW, speaker MW and speaker DW. *N* varies from 4 to 8. — = HL after son, - - - = HL after voiceless obs, — = ML after son, - - - = ML after voiced obs.

## 7. References

- [1] Kingston, J. and Diehl, R.L., "Phonetic knowledge", *Language* 70: 419-454. 1994.
- [2] Flemming, E., "Auditory Representations in Phonology". New York: Routledge. 2002. [PhD dissertation MIT 1995.]
- [3] Boersma, P., "Functional Phonology: Formalizing the Interactions between Articulatory and Perceptual Drives". The Hague: Holland Academic Graphics. 1998.
- [4] Zhang, J. "The effects of duration and sonority on contour tone distribution: Typological survey and formal analysis". Doctoral dissertation UCLA. 2001.
- [5] Gussenhoven, C. and Wang, L., "Yuhuan Wu tone and the role of sonorant onsets". In N. Campbell, D. Gibbon and D. Hirst (eds.), *Social and linguistic speech prosody. Proceedings of the 7<sup>th</sup> International Conference on Speech Prosody*. 619-622. 2014.
- [6] Port, R. and O'Dell, M., "Neutralization of syllable-final voicing in German". *Journal of Phonetics* 13: 455-471. 1986.
- [7] Chen, Y., "How does phonology guide phonetics in segment-*f*0 interaction?" *Journal of Phonetics* 39: 612-625. 2011.