Aug.15.2015 Note: the version of this paper presented at ICPhS 2015 has the following changes: 1. duration data added to show that the sandhi tone duration is neutralised despite the underlying tones having different durational profile (T242 is longer) and to justify time-normalising the F0 contour. 2. SS ANOVA was used to calculate the 95% confidence interval for the F0 contour. I also calculated a Bayes Factor, showing substantial favouring of the null hypothesis for the neutralisation experiment. 3. More emphasis was placed on the anticipatory dissimilation, or in the words of Dr. Albert Lee and Prof. Yi Xu, "pre-low raising". I also borrowed their very appealing account of this phenomenon, involving the enhancement of crycothyroid muscle activity in anticipation of falling pitch.

TONE SANDHI AND TONAL COARTICULATION IN FUZHOU MIN

Yang Li
Phonetics Laboratory, Department of Theoretical and Applied Linguistics, University of Cambridge
yl492@cam.ac.uk
yangli-linguistics.com

ABSTRACT

This study examines the completeness of Fuzhou tonal neutralisation and its interaction with tonal coarticulation. In Fuzhou, Tone 44, 242 and 53 are allegedly neutralised into Tone 44 preceding Tone 53. Acoustic data, analysed with Linear Mixed Modelling, show no significant difference in pitch height between neutralised tones, suggesting complete neutralisation in production. A forced-choice identification test reveals that Fuzhou speakers are unable to distinguish these neutralised tones perceptually. Further acoustic data show how the output of categorical tone sandhi may be modulated by tonal coarticulation, with dissimilatory effects superimposed on the sandhi tone.

Keywords: Fuzhou, tone sandhi, tonal coarticulation, neutralization.

1. INTRODUCTION

Fuzhou is a Min Chinese dialect spoken around Fuzhou city, Fujian. Fuzhou has seven citation tones, of which five are full tones and two are checked tones (carried by syllables that end in glottal stops). Fuzhou tone sandhi has been the subject of major phonological analyses [1,2,15,19], although there were few acoustic studies (exceptions are [4,5,8]).

Fuzhou has a paradigmatic, right-dominant tone sandhi system. In a disyllabic sandhi domain, the final syllable retains its citation tone, while the tonal identity of the pre-final syllable depends on both its underlying tone and the final tone. Table 1 illustrates this, with the top row showing the context (final) tone, leftmost column showing pre-final (target) tone, and cells in the middle showing the realisations of target tones before each context tone.

Table 1: Fuzhou tone sandhi table (full tones only).

Context tone →	**44**	**53**	31	**213**	**242**
Target tone ↓	Sandhi tone ↓				
44	**44**		53		
213					
242					
53	**44(31)**		31	**21**	
31	21		24	**44**	

This paper will focus on processes (1) and (2) (corresponding cells in Table 1 are in bold).

(1) 44, 242, 53 → 44 / ___ 53, 44
(2) 31 → 44 / ___213, 242

In (1), three distinct target tones are neutralised in unified tonal contexts. Fuzhou listeners should therefore not be able to distinguish among underlying 44-53, 242-53 and 53-53 pairs, all realised as 44-53. This forms the hypothesis to be tested in the perception experiment. Analogously, Mandarin third-tone sandhi (T21 → T24 / ___ T21) has been shown [8,10] to be perceptually neutralising, while evidence has been mixed regarding neutralisation in production.

1.1. Incomplete neutralisation

Incomplete neutralisation refers to small, sub-phonemic differences between pairs of words which contain putatively neutralised segments or tones [9,12]. In the tone sandhi realm, studies on Mandarin [6,18,20,21] showed that surface T24 resulting from tone sandhi is lower in pitch than lexical T24, although the detected effect size varied from 20 Hz [20] to 3.2 Hz [13]. Taiwanese, by contrast, seems to exhibit complete tonal neutralisation. In the most recent, well-controlled and statistically informed study, [10], no significant differences were found between the tonal contours of sandhi T55 and T24, which both resemble that of a lexical T33.

If Fuzhou tone sandhi involves categorical substitution of tonal units, no significant differences should be found in pitch height and contour on neutralised tones. Based on process (1), the first production experiment will test the completeness of tonal neutralisation by comparing the F0 profile of sandhi tones from differing underlying tones.

Linear Mixed Modelling (LMM) was selected in favour of repeated-measure ANOVA for comparing F0 between putatively neutralised tones, following, among others, [10]. LMM with subject as a random effect is more sensitive to nanovariation, and can maximally guard against "pseudoreplication" and its attendant, artificially high Type-I error rate [17]. This is important, as F0 differences between sandhi tones, if exist at all, are expected to be small.

1.2. Tonal coarticulation

Tone sandhi and tonal coarticulation have a blurred boundary, with [3] contending “there is no essential difference”. However, Fuzhou processes (1) and (2) may evidence their independence and interaction. Here we focus on four underlying tonal sequences: 53-53, 53-44, 31-242 and 31-213, whose initial tones are all putatively realised as T44. Do they indeed do so? If so, would the resulting high-level sandhi tones be modified differently by following tones of varying shapes? A theorist with a classical, modular view of phonology-phonetics interface might say that it was tone sandhi which turns T53 and T31 into T44 in a pre-articulatory, phonological process, described in (1) and (2); the resulting T44 will then be modified through the kind of tonal coarticulatory effects observed in tone languages without pervasive sandhi (e.g. Mandarin [13,17]). The second production experiment will test this view by comparing the F0 profile of underlying 53-53, 53-44, 31-242 and 31-213 pairs, with the expectation that the realisations of initial tones will be similar, because of sandhi, but distinct, due to coarticulation.

2. PRODUCTION EXPERIMENTS

2.1. Subjects

15 native Fuzhou speakers (M=8, F=7) were recruited, of whom 10 (M=5, F=5) passed a character-reading test and participated in both the production and perception experiments. During screening, subjects were shown 15 common words written in simplified Chinese on a computer screen, drawn from [7]. To pass, subjects had to correctly pronounce the Fuzhou readings of at least 13 of the 15 words, with 5 seconds of reaction time. All subjects come from the city of Fuzhou, and speak, read and write Mandarin to native proficiency. They had a mean age of 37 (SD=4), and reported no speech or hearing problem.

2.2. Materials and Procedure

The materials for the production experiments are disyllabic words (see Table 2) embedded in the carrier phrase [ŋuai^{31} puoʔ^{44} t^{h}øy^{44} ___ k^{h}øy^{44} ny^{31} t^{h}iaŋ^{55}], “I want to read ____ for you to listen”.

Table 2: Stimulus words for production experiments.

Tone pair	Syllable	Meaning
Neutralisation experiment		
44-53	/suoŋ.nøyŋ/	“businessman”
53-53	/suoŋ.nøyŋ/	“normal guy”
242-53	/suoŋ.nøyŋ/	“master”
44-53	/touŋ.nøyŋ/	“treat as man”
53-53	/touŋ.nøyŋ/	“tall man”
242-53	/touŋ.nøyŋ/	“father-in-law”
44-53	/gao.t^{h}ao/	“converse”
53-53	/gao.t^{h}ao/	“monkey head”
242-53	/gao.t^{h}ao/	“thick head”
Coarticulation experiment		
53-53	/suoŋ.nau/	“keep flowing”
53-44	/suoŋ.nau/	“keep scooping ”
31-242	/suoŋ.nau/	“want to release”
31-213	/suoŋ.nau/	“want to get old”

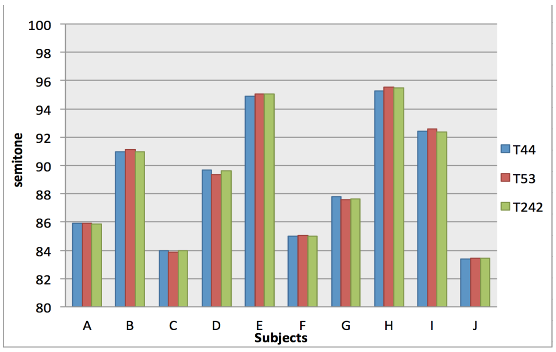
For the neutralisation experiment, there were 9 syllable-tone combinations: the three-way sandhi tone contrast was tested on three syllable types. Each combination was then repeated 6 times, generating 54 tokens from each subject. Before the session began, subjects were given 5 minutes to become familiar with the stimulus words, then listened to three example sentences. For the coarticulation experiment, there were 4 different underlying tonal pairs, carried by the same syllable. Each subject repeated each of these 6 times, generating 24 tokens.

All experiments were conducted in individual subject’s home. Subjects read the stimuli (presented in random order) from self-running slides with 10 seconds of interstimulus interval. A digital audio recorder was placed 20 cm away from their mouths, and the recording sampled at 44.1 kHz. A *Praat* script extracted F0 (semitone re 1 Hz) from five equally spaced points for statistical analysis.

2.3. Results: neutralisation experiment

Figure 1 shows the mean semitone values of sandhi tones originating from three underlying tones by 10 speakers. F0 differences between the sandhi forms of T53 and T242 and the realisation of underlying T44 are negligible across the board. The mean F0 (semitones) are 88.926 for sandhi T44, 88.946 for sandhi T53 and 88.951 for sandhi T242, which are remarkably close (the largest difference translates into a mere 0.3 Hz).

Figure 1: Mean F0 (semitone re 1 Hz) of sandhi forms of T44, T53 and T242 by 10 subjects.



LMM analysis was conducted in *R* [13] with the package *lme4* and function *lmer*. A full model was constructed, which included semitone as the dependant variable; gender, underlying tone, measurement point and syllable type as fixed effects; and speaker and repetition as random effects. Normality and homogeneity of variance were checked by visually inspecting the quantile plot of the pooled model residuals. A reduced model without the factor under study was then compared against the full model using a likelihood ratio test. The *p*-values thus obtained for some factors are summarised in Table 3. Based on the large *p*-value for the underlying tone factor and the aforementioned, negligible effect size, we cannot reject the null hypothesis nor conclude a finding of incomplete tonal neutralisation in our data.

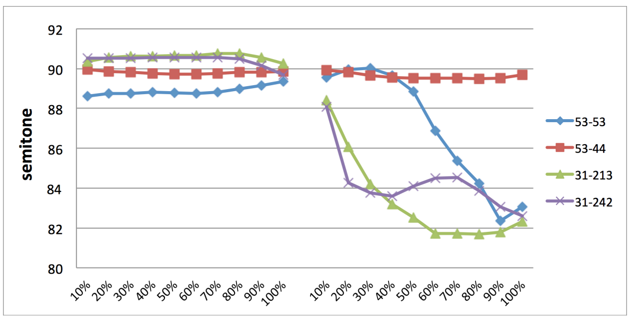
Table 3: Summary of *p*-value of various factors in the neutralisation experiment.

Factor	*p*
Underlying tone	= 0.585
Gender	< 0.001
Sample point	< 0.001
Syllable type	< 0.001

2.4. Results: coarticulation experiment

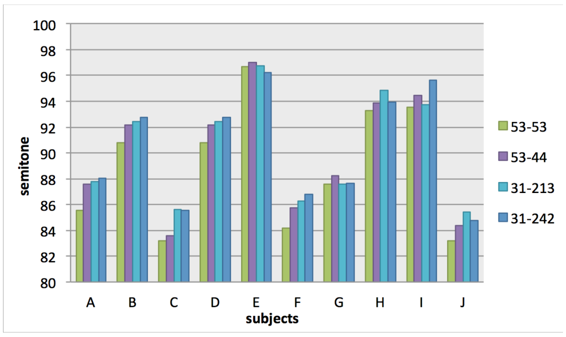
Figure 2 shows the mean surface F0 trajectory for the four underlying tone pairs. Each point is the average of 60 tokens (10 speakers * 6 repetitions). Notably, the initial tones in all four tone pairs are realised on a largely level pitch and reside at or near the top of subjects' pitch range. This presumably is what sandhi rules (1) and (2) are capturing. There are apparent dissimilatory effects from the context tones that are superimposed on the sandhied initial tones: they have the highest F0 preceding T213 and T242, somewhat lower F0 preceding T44, and the lowest F0 preceding T53.

Figure 2: Mean F0 (semitone) for underlying 53-53, 53-44, 31-213, 31-242 pair, with each tone sampled at 10 equally spaced points.



Individual variation of sandhi tone realisation can be observed in Figure 3. All subjects had the lowest mean F0 for the sandhi tone in the 53-53 pair. There is less agreement for the relative height of the other sandhi tones, with only subject C producing clearly higher sandhi tones for both 31-213 and 31-242.

Figure 3: Mean F0 (semitones) of the sandhi tone in 53-53, 53-44, 31-213, 31-242 by 10 subjects.



LMM analysis was conducted as in section 2.3, except that the models here did not include syllable type. A significant effect of underlying tone was found ($p<0.001$). Results of post-hoc Tukey Contrast are shown in Table 4, and they confirm that the sandhi tone in the 53-53 pair is significantly lower than those in other pairs, and by larger margin when contrasted with those in 31-213 and 31-242 pairs.

Table 4: Summary statistics of Tukey Contrasts. The "Contrast" column shows the two tone pairs whose F0 on initial tones are being compared.

Contrast	Effect size	SE	*p*
53.44 – 53.53	1.0471	0.1788	<0.001
31.213 – 53.53	1.7007	0.4413	<0.001
31.242 – 53.53	1.5349	0.3370	<0.001
31.213 – 44.53	0.6536	0.4510	=0.432
31.242 – 44.53	0.4879	0.2573	=0.204
31.242 – 31.213	-0.1657	0.4681	=0.982

3. PERCEPTION EXPERIMENT

3.1. Stimuli

A 2AFC identification experiment tested if Fuzhou listeners could detect the three-way underlying tonal contrast (T44 v. T242 v. T53) after tone sandhi. The experimental stimuli were recordings of the 9 tone-syllable combinations (words) used in the neutralisation experiment (see Table 2). Within each syllable type, there would be three oppositions by underlying tone: 44-53 v. 53-53, 44-53 v. 242-53, 53-53 v. 242-53.

As the above oppositions were expected to sound similar, a set of control stimuli was introduced to avoid back-to-back similar-sounding stimuli. The underlying 31-53 pair, with its sandhi form 21-53 [5,8], should be clearly distinguishable from the experimental tone pairs. We therefore pitted 31-53 combinations (listed in Table 5) with the three existing ones, creating three further oppositions: 31-53 v. 53-53, 31-53 v. 242-53, 31-53 v. 242-53.

Table 5: Control stimuli for perception experiment.

Tone pair	Syllable	Meaning
31-53	/suoŋ.nøyŋ/	“miss someone”
31-53	/touŋ.nøyŋ/	“block people”
31-53	/gao.t^{h}ao/	“nice heads”

The 18 stimulus words (9 experimental + 9 control) were recorded by an extra female speaker, who passed the character-reading test with 100% accuracy. The words were uttered in isolation and the perception test was conducted shortly after the production experiments. Such arrangements gave the most favourable condition for listener identification, as our hypothesis was that listeners would not perceive the difference among neutralised tones.

3.2. Experimental procedure

Each trial tested one tonal opposition on one syllable. To test the distinctness of 44-53 and 53-53 on syllable /suoŋ.nøyŋ/, for example, subjects would see the characters for /suoŋ^{44}.nøyŋ^{53}/ on the response sheet and hear the audio of /suoŋ^{44}.nøyŋ^{53}/ and /suoŋ^{53}.nøyŋ^{53}/. Subjects then indicated whether the written words were heard first or second. Two sources of bias were dealt with by counterbalancing: which of the two words, /suoŋ^{44}.nøyŋ^{53}/ or /suoŋ^{53}.nøyŋ^{53}/, was presented in written form, and which of them was played to subjects first (see Table 6). It thus took four trials to test one tonal opposition; with 3 syllable types, 6 oppositions and 10 subjects, we obtained 720 identification responses.

Table 6: Counterbalancing by which word was presented in writing and which audio stimulus appeared first, with four possibilities.

Opposition tested	44-53 v. 53-53	
Written word	Audio stimulus order	
44-53	44-53, 53-53	53-53, 44-53
53-53	53-53, 44-53	44-53, 53-53

3.3. Results

The overall percentages of accurate identification within each opposition were presented in Figure 4. There is a clear divide in identification rates between oppositions that are supposedly homophonous (experimental) and those that are distinct (control). Z-tests were conducted to see whether the differences between the assumed proportion (50%, if responses were random) and the observed proportion were significant. Results were tabulated in Table 7, showing that subjects were not identifying the correct sandhi tone significantly above chance level in any of the experimental oppositions. By contrast, they could identify the control oppositions significantly better than chance.

Figure 4: Percentages of accurate identification within each tonal opposition, grouped by whether the initial tones are supposedly neutralised.

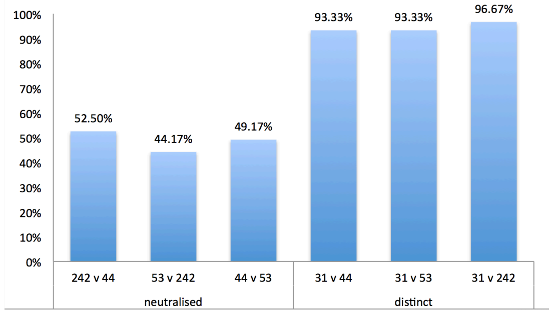


Table 7: Summary statistics for Z-tests.

Group	Accuracy	N	*p*
Overall “control”	94.34%	360	<0.001
Overall “experiment”	48.67%	360	=0.6138
242.53 v. 44.53	52.50%	120	=0.5839
242.53 v. 53.53	44.17%	120	=0.2015
53.53 v. 44.53	49.17%	120	=0.8266

4. CONCLUDING REMARKS

This study was motivated by two questions surrounding Fuzhou tone sandhi – whether it completely neutralises various underlying tones in a unified context, and how the neutralised sandhi tone may be further modified by non-categorical tonal coarticulation processes, when context tones differ.

Based on acoustic and perception data, we tentatively conclude that Fuzhou tone sandhi is neutralising, consistent with [10]’s conjecture that languages in the Min Chinese family have more categorical sandhi systems than Mandarin. Care must be taken, however, in interpreting the large *p* value found in the experiment, as absence of evidence does not equate evidence of absence.

The experiment on coarticulation revealed regressive dissimilatory effects, where the global F0 of the initial tone seemed to deviate from the mean in a direction opposite to the direction of the following pitch target. These effects can be seen as a sort of add-on to tone sandhi, which does the “heavier” work of turning tones as disparate in shape as T53, T31 and T242 into T44. Seen in this light, the present findings would support a modular, feed-forward model of phonology-phonetics interface, where the output of categorical phonology is fed as input to gradient, phonetic processes. Further research is needed to argue for this view, and a follow-up study, which manipulates speech rate and attempts at controlling lexical frequency, is currently underway.

5. ACKNOWLEDGEMENT

Thanks are due to Prof. Francis Nolan, Dr. Adrian Leemann, Professor Yi Xu, Dr. Albert Lee, Derek Johnson and Cherry Lam for valuable discussion. This work is supported by the External Research Studentship from Trinity College, Cambridge.

6. REFERENCES

[1] Chan, L.-L. 1998. *Fuzhou Tone Sandhi*. University of California, San Diego, doctoral dissertation.

[2] Chan, M. 1985. *Fuzhou Phonology*: *A Non-Linear Analysis of Tone and Stress.* University of Washington doctoral dissertation.

[3] Chen, M.Y. 2000. *Tone sandhi: Patterns across Chinese dialects*. Cambridge: CUP.

[4] Donohue, C. 1992. The phonetics of register in the Fuzhou dialect of Chinese. *Proc. Fourth Aust. Int. Conf. Speech Sci. Technol.*, 602–607.

[5] Donohue, C. 1992. *The Phonetics and Phonology of Fuzhou Tones*. ANU Honours thesis.

[6] Kratochvil, P. 1987. The case of the third tone. In: Chinese Language Society of Hong Kong (eds), *Wang Li Meml. Vol. Engl. Vol.* Hong Kong: Chinese University Press, 253–277.

[7] Li, R., Liang, Y., Zou, G., Chen, Z. 1994. *Fuzhou fangyan cidian [A dictionary of Fuzhou Dialect]*, Fuzhou: Fujian Renmin Chubanshe.

[8] Li, Y. 2014. An experimental study of Fuzhou tone sandhi. University of Cambridge MPhil thesis.

[9] Manaster, R. 1996. A letter from an incompletely neutral phonologist. *J. Phon.* 24, 477–489.

[10] Myers, J., Tsay, J. 2008. Neutralization in Taiwan Southern Min tone sandhi. In: Hsiao Y., Hsu H., Wee L-H, Ho D. (eds), *Interfaces in Chinese Phonology.* Taipei: Academia Sinica, 47–78.

[11] Peng, S. 1996. *Phonetic Implementation and Perception of Place Coarticulation and Tone Sandhi*. OSU doctoral dissertation.

[12] Port, R., Crawford, P. 1989. Incomplete neutralization and pragmatics in German. *J. Phon.* 17, 257–282.

[13] R Development Core Team, 2014, *R: A Language and Environment for Statistical Computing*, Vienna, Austria.

[14] Shen, X.S. 1990. Tonal coarticulation in Mandarin. *J. Phon.* 18, 281–95.

[15] Shih, C.-L. 1985. From Tonal to Accentual: Fuzhou Tone Sandhi Revisited. *Proc. Annu. Meet. Berkeley Linguist. Soc.* 11.

[16] Wang, W.S., Li, K. 1967. Tone 3 in Pekinese. *J. Speech Hear. Res.* 10, 629–636.

[17] Winter, B. 2011. Pseudoreplication in phonetic research. *Proc. of ICPhS,* Hong Kong, 2137–2140.

[18] Xu, Y. 1993. *Contextual Tonal Variation in Mandarin Chinese*. University of Connecticut doctoral dissertation.

[19] Yip, M. 1980. *The Tonal Phonology of Chinese*. MIT doctoral dissertation.

[20] Zee, E. 1980. A Spectrographic Investigation of Mandarin Tone Sandhi. *Univ. Calif. Work. Pap. Phon.* 49, 98–116.

[21] Zhang, J., Liu, J. 2011. Tone sandhi and tonal coarticulation in Tianjin Chinese. *Phonetica* 68, 161–191.