# GRADIENCE IN CONTEXTUAL TONAL REALIZATION PROCESSES: AN INSTRUMENTAL STUDY OF NANJING CHINESE

Xiaole Sun Tsan Huang
Department of Linguistics, SUNY at Buffalo
xiaolesu@buffalo.edu thuang3@buffalo.edu

## ABSTRACT

This acoustic study has investigated contextual tonal variations in 125 disyllabic words produced by six native speakers of Nanjing Chinese, with five phonemic tones. Three types of tonal variation were observed in the data: neutralizing sandhi, non-neutralizing sandhi and tonal coarticulation. Comparison between the three types of tonal variation suggests that tone sandhi may be rooted in phonetic coarticulation and influenced by the perception-production interaction. Non-neutralizing sandhi processes demonstrate a transition stage from tonal coarticulation to category-changing and neutralizing tone sandhi. The contextual tonal realization in Nanjing Chinese exhibits a gradient effect in terms of tonal category change, indicating a blurry boundary between phonetic and phonological processes.

**Keywords**: speech production, tonal coarticulation, tone sandhi, Nanjing Chinese, gradience

## 1. INTRODUCTION

Nanjing Chinese has five phonemic tones described in previous reports as high falling, rising, low falling, level, and checked high [1,2]. However, there was a lack of comprehensive analysis of tonal realization in Nanjing Chinese in these studies. When tones are produced in citation forms, the f0 contours seem well defined and stable [3]. When produced in connected speech, it has been found cross-linguistically that tonal variations take place due to the effect of neighboring tones [3,4,5]. These types of tonal interaction and adjustment are commonly referred to as either tone sandhi or tonal coarticulation [4,5,6,7]. The goal of this study is to provide sufficient acoustic description of lexical tones and contextual tonal realization processes in Nanjing Chinese and to further address the relationship between tone sandhi and tonal coarticulation.

## 2. EXPERIMENT

### 2.1. Procedures

Six female native speakers of Nanjing Chinese participated in the production experiment and completed a reading task. The age of all the speakers at the time of recording was 25 years. The selected test items were words in everyday use found in [1]. These disyllabic words included all possible combinations of the five lexical tones resulting in a total of 25 (5*5) tonal combinations. Each tonal combination was represented in five different test words, resulting in a total of 125 (5*5*5) disyllabic words. In addition, 250 (125*2) monosyllabic words originating from the 125 disyllabic words were recorded. The monosyllabics allowed for an acoustic description of the tonal inventory. These citation tones also provided us with a departing point for investigation into contextual tonal variation. Carrier sentences were used for natural pronunciation. The speakers were instructed to read each sentence twice, with the semi-randomized test words embedded in them. Each participant's utterances were recorded at the sampling rate of 44100Hz with Praat [8]. Independent equal variance two-tailed t-tests were conducted on five points (onset, midpoint, offset, minimum and maximum) of the pitch contours on monosyllabic words and the corresponding pitch contours in disyllabic conditions to examine pitch change in adjacent tones. T-tests were also used to determine whether the contrast between a surface tone and another phonemic tone is neutralized in a certain context.

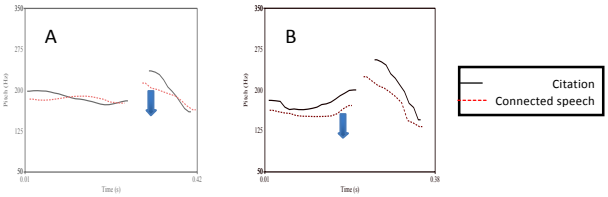
### 2.2. Results and analysis

The acoustic measurements of f0 values and duration of the monosyllabic words show that the tonal inventory of Nanjing Chinese includes: a high falling tone T41 (average duration =130ms), a rising tone T213 (average duration = 190ms), a low falling tone T21 (average duration of 100ms), a level tone T32 (average duration = 230ms), and a checked high tone T45 (average duration = 85ms), by using Chao's [3] five-point pitch scale. T-tests have found significant differences in 17 disyllabic combinations when the derived tones in the disyllabic environment are compared with the citation tones.

In 12 out of these 17 disyllabic sequences, a tone maintains its tonal contour, albeit deviating slightly from its canonical form. For example, in Figure 1A, the pitch contour of T41 is lowered ($p_{onset}$=0.02) when preceded by T32, whose low f0 offset triggers the progressive coarticulation process. However, the

surface contour of T41 does not seem to deviate much from its citation form. The same progressive lowering effect is also observed in the /T41 + T32/, /T41 + T45/, /T21 + T41/, /T21 + T213/, /T21 + T45/, /T32 + T21/ and /T32 + T45/ sequences.

**Figure 1:** Pitch drop does not lead to phonological tonal change in T41 in the word /ta32 -kə41/ 'oldest brother' (Panel A) or T213 (Panel B) in the word /wan213- xun41/ 'get married'.



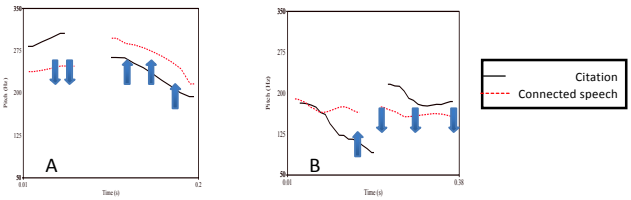

In terms of direction of coarticulation effect, the first tone in a disyllabic sequence could also undergo pitch change from its canonical form. In Figure 1B, the pitch contour of T213 is leveled before T41, with its f0 offset significantly lowered than its citation form ($p_{offset}$=0.006). Because of the relatively short time to realize two adjacent high pitch targets, the high pitch offset of the first rising tone delays and collapses with the f0 onset of the following high falling tone. This phonetic peak delay effect has also been reported in other studies of Chinese languages [6,7].

In some sequences, bidirectional assimilatory coarticulation occurs and causes noticeable pitch deviations from citation tones. In Figure 2A, a tonal coarticulation process under the influence of peak delay is observed in the /T45 + T41/ sequence: the pitch contour of T45 is lowered and leveled ($p_{mid}$=0.0003, $p_{offset}$=0.01) compared with its citation form. At the same time, the surface T41 shifts up ($p_{onset}$=0.003, $p_{mid}$=0.0003, $p_{offset}$=0.01) after T41. However, both T45 and T41 seem to maintain their canonical contours.

Figure 2B illustrates a phonetic compromising of pitch height difference in the sequence /T21 + T32/. The f0 offset of T21 is significantly raised ($p_{offset}$=0.0009) by the relatively high onset of the following T32. In the meantime, the entire pitch contour of T32 is significantly lowered ($p_{onset}$=0.0001, $p_{mid}$=0.0004, $p_{offset}$=0.0003) by the low offset of the preceding T21. However, the canonical tone shapes do not alter for both T21 and T32, despite the pitch variations.

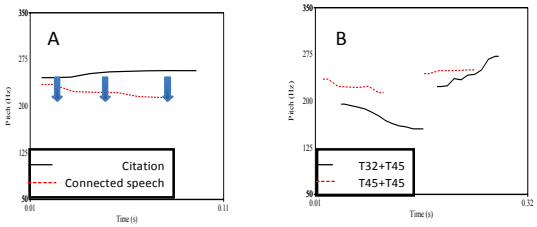
As shown in Figure 1&2, such pitch changes do not lead to phonological tonal alternation; that is, these contextual tonal variations are phonetic in nature and should be seen as tonal coarticulation [5,6,7].

**Figure 2:** Bidirectional contextual tonal variations do not lead to phonological tonal change in T45 and T41 in the word /pəʔ45-fən41/ 'north wind' (Panel A) or T21 and T32 in the word /wan21-fan32/ 'dinner' (Panel B).




Different from tonal coarticulation, a surface tone may deviate from its citation form and assimilate to another phonemic tone. As shown in Figure 3A, the entire pitch contour of the first T45 ($p_{onset}$=0.01, $p_{mid}$=0.003, $p_{offset}$=0.0001) is lowered and changes its direction into slightly falling before another T45. The surface pitch contour of the first T45 tone in the /T45 + T45/ sequence is similar to that of T32 in the /T32+T45/ sequence, as illustrated in Figure 3B. However, the surface T45 does not completely change its tonal identity into T32 due to the pitch height difference ($p_{offset}$=0.001) and duration difference ($p_{duration}$=0.003). Figure 3B shows that tonal contours of /T45 + T45/ and /T32 + T45/ do not overlap.

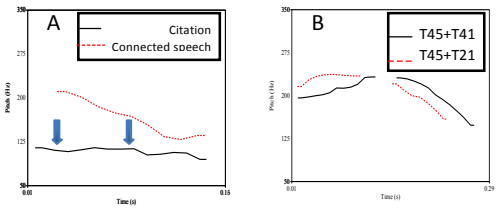
**Figure 3:** Peak delay effect on the first T45 in /T45 + T45/ (excised from the disyllabic word /xuaʔ45-ɕyeʔ45/ 'ski') results in direction change from rising to falling (Panel A). However, surface pitch contours of /T45 + T45/ in the word /xuaʔ45- ɕyeʔ45/ 'ski' and surface form of /T32 + T45/ in the word /xua32-ɕyeʔ45/ 'chemistry' do not overlap completely (Panel B).




In the sequence /T45 + T21/, the first half of T21 is significantly raised ($p_{onset}$=0.008, $p_{mid}$=0.01) by the high f0 offset of the preceding T45. The derived T21 has a high falling pitch contour (Figure 4A).

**Figure 4:** Progressive assimilation in /T45 + T21/ changes the pitch contour of T21 (excised /pəʔ45-jen21/ 'supercilious look') from low falling to high falling (Panel A). However, pitch contours of /T45 + T21/ in the word 'supercilious look' and /T45 + T41/

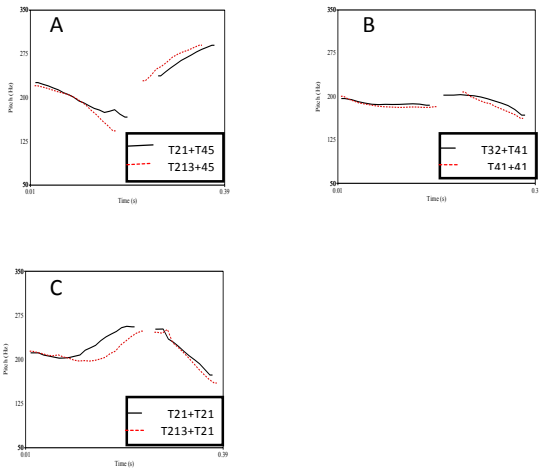
in the word /pəʔ45-jen41/ ‘white smoke’ do not overlap completely (Panel B).




In Figure 4A, the surface T21 resembles T41 in terms of tone shape. However, there is significant difference between T21 and T41 after T45 ($p_{onset}$=0.03, $p_{mid}$=0.04, $p_{offset}$=0.01, $p_{duration}$=0.03), despite the similarity in pitch contours. In Figure 4B, it can be seen that the surface forms of T21 and T41 differ in pitch height and duration when occurring after T45. Just as in the case of T32 and T45, which is shown in Figure 3 and discussed above, the contrast between T21 and T41 may be reduced but not neutralized after T45. Non-neutralizing sandhi, which is a tonal incomplete neutralization process, is proposed here to describe this type of contextual tonal variation.

Aside from the afore-mentioned tonal realization patterns, three disyllabic sequences in Nanjing Chinese exhibit yet another contextual tonal variations, as shown in Figure 5.

**Figure 5:** The two first tones in any of the three disyllabic pairs (Panel A: T213 and T21 before T45; Panel B: T41 and T32 before T41; Panel C: T21 and T213 before T21) are acoustically and perceptually indistinguishable from each other.




In Figure 5A, the pitch contours of the surface forms of /T213 + T45/ and /T21 + T45/ sequences overlap completely. The t-tests on f0 and duration measurements found no significant difference between the surface forms of T213 and T21. This complete overlap of two pitch contours is a result of the peak delay effect in the sequence /T213 + T45/. The f0 offset of T213 is significantly lowered ($p_{offset}$ < 0.00001) because the rising tone is followed by a high pitch target (i.e., the onset of T45). With the lowered f0 offset, the surface T213 has a low falling shape and becomes identical to T21.

Figure 5B shows that the pitch contours of the surface forms of /T41 + T41/ and /T32 + T41/ sequences overlap completely. The f0 offset of T41 is raised by another T41 ($p_{offset}$=0.04). The f0 onset is lowered ($p_{onset}$=0.003) and therefore the overall pitch contour overlaps with T32. T-test results support the finding of complete neutralization since there is no significant difference on f0 values and duration of the surface forms between T41 and T32 before T41. Due to the raising effect of the f0 offset and lowering effect of the f0 onset, the surface pitch contour of the first T41 is leveled and resembles T32 in the same tonal environment. The raising effect of the f0 offset may be attributed to an obvious cause: a phonetic interpolation between the low f0 offset of the first T41 and the high f0 onset of the second T41 raises the low offset of the first T41. The lowering of the f0 onset of the first T41 seems to lack an obvious phonetic explanation. One may speculate that this lowering of the f0 onset of the first T41 is very likely motivated by the perceptual confusion of its now less steep falling contour (after its f0 offset raising) with the contour of T32, which might have led the listener-turned-speaker to exaggerate the similarity and produce the same tone. This analysis is consistent the findings of the influence of speech perception on language sound system and sound change [9].

In the /T21 + T21/ combination, the f0 offset of the first T21 is raised ($p_{offset}$ < 0.00001). The derived T21 has an ascending contour, which is similar to T213. Comparing the surface forms of /T21 + T21/ and /T213 + T21/, as shown in Figure 5C, the two contours are essentially identical, ignoring the temporal difference due to speech rate. The t-test results show no significance difference in average f0 and duration between the two sequences. This type of complete neutralization is traditionally referred to as tone sandhi [4,10].

## 3. PRELIMINARY PERCEPTION EXPERIMENT

A preliminary discrimination perceptual test was conducted with one native speaker who also participated in the production experiment. The speaker was asked to judge whether three pairs of disyllabic words involved in neutralizing sandhi and two pairs of disyllabic words involved in non-neutralizing sandhi were the same or different. The speaker was able to distinguish two words involved in non-neutralizing sandhi, but unable to

discriminate the words in the neutralizing sandhi pairs. However, the listener reported a certain amount of perceptual confusion on perceiving non-neutralizing tone sandhi pairs. In other words, for non-neutralizing sandhi processes, perceptual distinction may be preserved to some extent, but perceptual confusion caused by the acoustic similarities may still affect the discrimination process.

## 4. CONCLUSION AND DISCUSSION

In sum, the 17 cases of contextual tonal variations in Nanjing Chinese are classified into three categories based on the degree of tonal deviations and completeness of neutralization: 12 patterns of phonetic tonal coarticulation, two patterns of non-neutralizing sandhi and three patterns of neutralizing sandhi. It has been suggested that tone sandhi differs from tonal coarticulation in two major ways: (1) tone sandhi is attributed to language-specific morphophonemic constraints, which cannot simply be explained by phonetic motivation, while tonal coarticulation is attributed to physical constraints of an utterance; and (2) tone sandhi involves categorical change and neutralization, while tonal coarticulation preserves tonal identity [10]. This study shows that the relationship between tone sandhi and tonal coarticulation is rather complicated. First, tonal coarticulation and tone sandhi have the same articulatory bases. For example, the peak delay process was found in all three of neutralizing sandhi, non-neutralizing sandhi and tonal coarticulation processes. Second, it was found that the derived tone in the non-neutralizing sandhi environment significantly differs from its canonical form and almost merges with another phonemic tone, demonstrating a possible transitional stage from sub-phonemic tonal coarticulation to category-changing and neutralizing sandhi processes. In this light, complete tonal neutralization as phonological alternation may be seen as originating in phonetic tonal coarticulation. It may have gone through an earlier stage where the degree of deviation in a surface tonal contour might have been comparable to what we have observed in the non-neutralizing sandhi processes in Nanjing Chinese. Such a degree of deviation in tonal contour may render a phonemic tone A's surface tone *A similar to another phonemic tone B (perhaps even more than the underlying tone A), paving the path to an incipient neutralizing tone sandhi. When the listener recategorized *A as B due to the perceptual similarity, the neutralizing sandhi process would be established in the phonology. The relation between phonetics and phonology suggested here is that phonological invariance cannot be completely separated from phonetic variance. When the phonetic variance crosses the perceptual boundary of phonological categories, a phonological alternation might start to occur, leading to the establishment of seemingly abstract phonological rules. Such a view on the phonetics-phonology relation is also consistent with the argument that phonological processes are phonetically detailed and the boundary between phonetic and phonological process is blurry [11].

## 5. REFERENCES


[1] Liu, D. 1995. *Nanjing Dialect Dictionary.* Nanjing: Jiangsu Education Publisher.

[2] Sun, H. 2003. A discussion on the tones of Nanjing dialect. *Nanjing Xiaozhuang Xueyuan Xuebao* (Journal of Nanjing Xiaozhuang University), 1, 34-40.

[3] Chao, Y-R. 1965. *A Grammar of Spoken Chinese.* Berkeley & Los Angeles: University of California Press.

[4] Chen, M. Y. 2000. *Tone Sandhi: Patterns across Chinese dialects*. Cambridge: Cambridge University Press.

[5] Gandour, J., Potisuk, S., Dechongkit, S. 1994. Tonal coarticulation in Thai. *Journal of Phonetics*, 22, 474– 492.

[6] Huang, T. 2011. Contextual and pitch range effect on tonal realizations in Rugao Chinese. *Journal of Chinese Linguistics*, 39(1), 149-176.

[7] Xu,Y. 1993. Contextual tonal variation in mandarin Chinese. Ph.D. dissertation, University of Connecticut.

[8] Boersma, P., Weenink, D. 2005. Praat: doing phonetics by computer (Version 4.3.01). [Computer program]. Retrieved from http://www.praat.org/

[9] Ohala, J. J. 1990. The phonetics and phonology of aspects of assimilation. In Kingston J., & Beckman, M. E. (Eds.), *Laboratory Phonology I: Between the grammar and the physics of speech*, 258-275.

[10] Shen, X. S. 1990. Tonal coarticulation in Mandarin, *Journal of Phonetics*, 18, 281-295.

[11] Steriade, D. 2000. Paradigm uniformity and the phonetics-phonology boundary. In Papers in laboratory phonology 5: acquisition and the lexicon, edited by M. B. Broe & J. B. Pierrehumbert, 313-334. Cambridge: Cambridge University Press.