An Acoustic Comparison between Two Pairs of Assimilatory and Dissimilatory Tone Sandhi Processes in Nanjing Mandarin in Categoricalness/Gradience

Xin Li, René Kager

Utrecht Institute of Linguistics, Utrecht University, The Netherlands
xin.li@uu.nl, r.w.j.kager@uu.nl

Abstract

The current study aims to investigate the categoricalness versus gradience of tone sandhi application in assimilatory and dissimilatory tone sandhi processes, within the setting of one language. Our hypotheses were that (a) assimilatory processes can be gradient, based on their articulatory motivation; and (b) dissimilatory processes should always be categorical, hence never show any gradient application, based on Ohala's "hyper-correction" theory. We selected in Nanjing Mandarin two pairs of comparable assimilatory and dissimilatory tone sandhi processes based on previous researchers' observations. Results show a near-categorical assimilatory Sandhi 1 and a gradient assimilatory Sandhi 3, congruent with the prediction that assimilatory processes are allowed to apply in a gradient fashion. Though we found that dissimilatory Sandhi 4 is a categorical process, our observation of gradience in dissimilatory Sandhi 2 suggests that dissimilatory changes can also occur in a gradient way, which contradicts the prediction by the "hyper-correction" account.

Index Terms: categoricalness versus gradience, tone sandhi production, assimilation and dissimilation, gradient dissimilation

1. Introduction

In tone languages, lexical tones may undergo alternations under the influence of their contextual tones. This phenomenon is generally known as tone sandhi [1]. Tone sandhi processes can be classified as "assimilatory" and "dissimilatory" based on whether the altered tone becomes more or less similar to its context tone. Assimilatory and dissimilatory processes are widely believed to have distinct origins, the former based in articulation [2, 3], and the latter not [4]. The general question we pursue in this study is whether assimilatory and dissimilatory tone sandhi rules result in different acoustic realizations in native speakers, in terms of categoricalness versus gradience. If a tone sandhi rule is *categorical*, the underlying tone is neutralized with another tone, whereas if a sandhi rule is *gradient*, pitch features of the underlying tone are only partially lost from the surface form.

Previous experimental studies have looked into the categoricalness/gradience of tone sandhi processes in native speakers in many tone languages/dialects. For example, [5] found in Taiwan Southern Min that a high level *yinping* (in traditional tone classification of this dialect) is categorically assimilated and neutralized to a mid level *yangqu* before a low rising *yangping* (55.24 → 33.24; the numbers indicate tone values on [6]'s five-level scale; the dot is used to separate tones); Beijing Mandarin has a dissimilatory tone sandhi process which was observed to involve a non-neutralizing change from underlying low T3 to rising surface T2 before another low T3 (LL.LL → LH.LL) [7-9]. However, no general conclusions can be drawn from comparing these studies about the relationship between on the one hand, assimilatory/dissimilatory tone sandhi and on the other, categoricalness/gradience of sandhi application, because cross-linguistic comparisons may be hampered by language-specific factors such as phonological processes or sound inventories. Also, the experimental and analytical methodologies adopted in these studies vary from each other.

A single-language setting is crucial to a fair comparison between assimilatory and dissimilatory processes. A few studies targeting neutralization in specific tone languages/dialects involved a number of tone sandhi processes. However, whether the sandhi processes in these studies are assimilatory or dissimilatory in nature is not always clear. For example, [10] reported a categorical tone sandhi rule which changes a mid falling T1 before another T1 to a neutralized high rising T2, with the onset of the second T1 raised as well due to assimilatory carryover effect. This process was transcribed as L.L → LH.L (hence, dissimilatory) in their study. Yet it may also be interpreted as an assimilatory process, because the offset of the first T1 is raised to the same pitch(f0) height as the onset of the following T1.

No earlier study compared the categoricalness/gradience of undisputed assimilatory and dissimilatory processes within a single tone language, as far as we are aware. Hence, a study of this type is needed to establish whether there are differences in tone sandhi categoricalness/gradience between pair(s) of sandhi processes that differ in their assimilatory/dissimilatory nature, but are otherwise maximally comparable in terms of f0 trajectories.

Predictions can be made regarding the categoricalness/gradience of assimilatory and dissimilatory processes, from the different origins of these processes. Assimilatory processes are mostly believed to be motivated by increasing ease of articulation [2, 3]. [11] made the prediction that assimilatory processes tend to start out as gradient, because sound changes due to articulatory processes are inherently gradient, and the changes become reinterpreted as phonological rules by listeners in a slow diachronic progress. With regard to dissimilatory processes, the most widely accepted view concerning their motivation was Ohala [4, 12]'s proposal that a dissimilatory sound change occurs because listeners are under pressure to preserve an underlying segment in the surface form, and misattribute a feature that is intrinsic to the surface form to coarticulation, erroneously applying corrective processes ("hyper-correction"). This view predicts dissimilatory processes to always be categorical, as claimed explicitly by Ohala [4][13] and later researchers who referred to the "hyper-correction" theory [11] [14] [15].

The current study will investigate the categoricalness/gradience of assimilatory and dissimilatory tone sandhi processes in Nanjing Mandarin, a tone language of China known to have both assimilatory and dissimilatory tone sandhi processes [16-18]. The specific goal is to select pair(s) of comparable assimilatory and dissimilatory tone sandhi processes from Nanjing Mandarin, and look into the categoricalness/gradience of the tone sandhi applications by native speakers, in order to test two core predictions, in particular that (a) assimilatory processes can be gradient, and that (b) dissimilatory processes should always be categorical, hence never show any gradient application.

2. Two pairs of comparable assimilatory and dissimilatory tone sandhi processes in Nanjing Mandarin

There have been a number of experimental studies on lexical tones and disyllabic tone sandhi patterns of Nanjing Mandarin, of which the most cited are [16-18]. They used listeners aged between 50 and 90 year old. The lexical tones and two pairs of assimilatory and dissimilatory disyllabic tone sandhi rules from their studies are shown as below.

2.1. Lexical tones

Nanjing Mandarin has five distinctive lexical tones (neutral tone excluded). Table 1 summarizes the transcriptions of the tones recorded in [16-18]. T1 in Nanjing Mandarin is recorded in these studies as a high/mid-falling tone; T2 is a low-rising tone; T3 is recorded as a low level/dipping tone; T4 is a high/mid level tone; T5 is a high and abrupt tone.

Table 1. *Transcriptions of the five lexical tones of Nanjing Mandarin in previous studies. (A single number in the tone value indicatess a short and abrupt tone.)*

T1	T2	T3	T4	T5	Reference
41	24	11	44	5	[16]
31	24	11	44	5	[17]
41	24	22/212	44/33	5	[18]

2.2. Two pairs of comparable assimilatory and dissimilatory tone sandhi processes

Consulting [16-18]'s studies on Nanjing Mandarin tone sandhi patterns, we found two pairs of comparable assimilatory and dissimilatory tone sandhi processes. In assimilatory Sandhi 1 (T1.T1 → T4.T1), the offset of a high-falling T1 in the first syllable is raised to the same pitch as the onset of the following T1, transcribed as HL.HL → HH.HL [19]. This sandhi process changes a high falling T1 preceding another T1 to a high level tone resembling a lexical T4. In dissimilatory Sandhi 2 (T4.T5 → T1.T5), the offset of a high-level T4 in the first syllable deviates from the high onset of the following T5 instead of approaching it, transcribed as HH.H → HL.H [19]. This sandhi process changes a high level T4 preceding a high T5 to a high falling tone resembling a lexical T1.Though the underlying-to-surface tonal changes are opposite in direction in the two sandhi rules, both sandhi rules involve a Nanjing Mandarin T1 vs. T4 contrast, with equal distance between the underlying tone and the surface tone. Thus these constitute a pair of assimilatory and dissimilatory tone sandhi processes that are comparable with each other. Similarly, assimilatory Sandhi 3 (T3.T1 →T2.T1) changes a low level/dipping T3 preceding a high falling T1 to a low rising T2, transcribed as LL.HL → LH.HL [19]. Dissimilatory Sandhi 4 (T3.T3 → T2.T3) changes the low level/dipping T3 preceding another T3 also to a low rising T2, transcribed as LL.LL → LH.LL [19]. Both sandhi rules involve an alternation between an underlying T3 and a surface T2, and constitute another comparable pair of assimilatory and dissimilatory tone sandhi processes. In sum, we will look into the categoricalness/gradience of Sandhi Pair 1): assimilatory Sandhi 1 versus dissimilatory Sandhi 2 and Sandhi Pair 2): assimilatory Sandhi 3 versus dissimilatory Sandhi 4, collecting new data from native Nanjing speakers.

3. Methodology

3.1. Participants

Eighteen native speakers of Nanjing Mandarin (9 females & 9 males) were recruited for the experiment. All speakers were aged between 18 and 30 years old. They were born and raised in Nanjing and had never left Nanjing for more than half a year. None of them had any speaking or hearing problems. They were paid for the experiment, and were not aware of the purpose of the experiment.

3.2. Stimuli

For each of the five lexical tones of Nanjing Mandarin, four monosyllabic words were selected. The monosyllabic words were all frequently used and easy-to-combine Chinese characters used in Nanjing Mandarin. The distribution of front-back and high-low vowels was balanced across the words. Since these monosyllabic words will be used in combining disyllabic words with other monosyllabic words, zero consonants were excluded as far as possible, for the convenience of splitting syllables. For each monosyllabic word Ta (T1/T2/T3/T4/T5), there were 10 disyllabic combinations, namely, 5 disyllabic combinations of Ta.Tx (X =T1, T2, T3, T4, T5), when Ta was in the first syllable, and 5 disyllabic combinations of Tx.Ta, when Ta was in the second syllable. Therefore, the word list contained 220 words (11 tone types (1 monosyllabic word + 10 disyllabic sequences +) * 4 monosyllabic words * 5 lexical tones).

Most of the disyllabic words in the word list were selected from [16], which documented the typical Nanjing Mandarin vocabulary in the 1990s. A few other words were common Mandarin vocabularies that were also used in the Nanjing Mandarin area. We had all the test words examined by a native female Nanjing Mandarin speaker living in Utrecht, and all of the words were reported to be common words that are natural to produce. All of the words were then randomized to work as the stimuli.

3.3. Recording

Recordings took place in a quiet room in Nanjing University, in Nanjing, China. Recordings were conducted with a headset microphone, using the software cool edit pro set up on the computer, with 44.1 kHz sampling rate and 16 bit rate in mono channel.

The stimuli were presented one by one in simplified Chinese characters on the computer screen. Speakers were seated in front of the screen and read each stimuli two times naturally in Nanjing Mandarin. They were able to control the speed of stimulus presentation, but generally 1-2 seconds between trials. Altogether 40 monosyllabic tokens and 400 disyllabic tokens were elicited from each of the 18 participants.

3.4. F0 analysis

The segmentation was manually conducted using Praat 4.3.09 [20]. The F0 values of the tone contour of each word were measured at 10 equidistant points (P1…P10) in the tone-bearing rhyme part of the syllable, resulting in a set of time-normalized f0 values, which eliminated the difference among durations of every rhyme part to get comparable f0 curve. These values were automatically extracted by a self-written script in Praat and saved together with their original time scale.

In order to eliminate the anatomical variation between speakers caused by their different sizes of vocal folds and to allow between-speaker comparison, f0 normalization was conducted by converting Hertz values to semitones relative to each speaker's average pitch [21].

4. Experimental results

4.1. Lexical tones produced in isolation

Figure 1 shows the 5 lexical tones of Nanjing Mandarin averaged across the 18 speakers (9 males and 9 females). T1 is a high-falling tone, with a pitch contour that falls from the higher to the lower end of the speaker's pitch range; T2 is a low-rising tone, whose pitch contour rises from the lower to the upper end of the pitch range; T3 is a low falling tone, which falls slightly from the lower pitch range to even lower; T4 is a mid-high level tone, which stays at the mid-high level of the pitch range; T5 is a salient high tone with a concave contour shape, and it is short and abrupt if duration is considered. The tonal shapes of the lexical T1, T2, T4 and T5 observed in our study mostly replicate those in previous studies, yet we find the overall falling tendency in lexical T3, which may be a new feature of the productions by the younger age group.

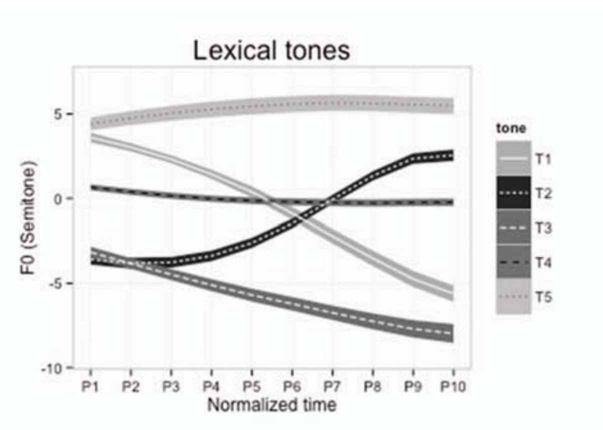

Fig. 1. *Five lexical tones in Nanjing Mandarin produced in isolation (with normalized time) by 18 speakers. Central lines represent the means. Shaded areas (ribbons) stand for ±1 standard error of the mean.*

4.2. Categoricalness/gradience of the two pairs of assimilatory and dissimilatory sandhi processes

Growth curve analysis [22] was employed to compare the sandhied tones against their claimed target surface tones in this study. The advantage of this analysis method is that it evaluates the overall f0 contours of the tones. It was successfully used by [23] and [24] to determine whether there is tone neutralization in Tianjin Mandarin and Shanghai Wu tone sandhi processes, respectively. In growth curve analysis, a non-linear f0 curve is fitted with the formula in (1), with x representing time, y representing f0, and different coefficients indicating different features in f0 curves. The intercept a indicates the overall f0 mean of a curve; b indicates the slope of the f0 change; c represents the steepness of the centered peak [24]. For all the tone sandhi processes in this study, we follow [23] and [24] to include up to second-order polynomials to adequately model for f0 changes of the tones, and meantime avoid overfitting.

$$y = a + bx + cx^2+dx^3+ex^4... \quad (1)$$

Each model was first built with only the fixed effects of linear and quadratic time terms, and the random effects of speaker and speaker-by-tone on the time terms (base model). Then the fixed effects of tone on the intercept, on the linear time term and on the quadratic time term were added in the model in a stepwise way. The significance of each effect was evaluated by log-likelihood model fit comparison at each step.

4.2.1. Sandhi pair 1: assimilatory Sandhi 1 vs. dissimilatory Sandhi 2

Figure 2 depicts the sandhied form of T1.T1 against its claimed target surface tone pattern T4.T1, averaged across 18 speakers. The growth curve analysis revealed that the surface T1 on syllable 1 (T1(syl1)) in T1.T1 and the underlying T4 in T4.T1 are marginally significantly different in overall f0 mean ($p < .05$ *). However, they are not significantly different in f0 slope ($p > .05$) or sharpness of the centered peak ($p > .05$). It seems that the surface T1(syl1) in T1.T1 is near-neutralizing with the underlying T4 in T4.T1, which suggests that assimilatory Sandhi 1 applies in a near-categorical way.

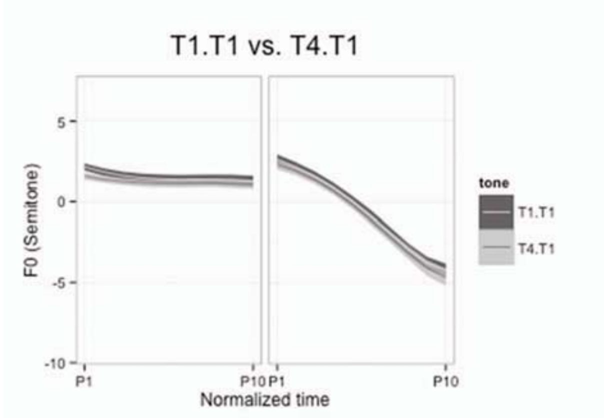

Figure 2. *Sandhied T1.T1 vs. its claimed surface T4.T1 in Nanjing Mandarin (with normalized time) by 18 speakers. Central lines represent the means. Shaded areas (ribbons) stand for ±1 standard error of the mean.*

Figure 3 depicts the sandhied form of T4.T5 against its claimed target surface tone pattern T1.T5, averaged across 18 speakers. The growth curve analysis revealed no significant difference between the T4 in T4.T5 and the T1 in T1.T5 in overall f0 mean (p > .05). But the two tones are significantly different in steepness of f0 slope (p < .001 ***) and sharpness of the centered peak (p < .05 *). The T4 in T4.T5 is significantly less steep in f0 slope and less sharp in the centered peak than the T1 in T1.T5, and thus the surface form of T4 in T4.T5 is non-neutralizing with the underlying T1 in T1.T5. Hence we conclude that dissimilatory Sandhi 2 applies in a gradient way.

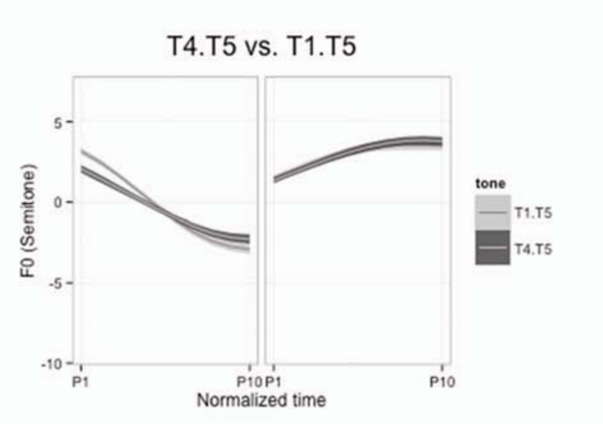

Figure 3. *Sandhied T4.T5 vs. its claimed surface T1.T5 in Nanjing Mandarin (with normalized time) by 18 speakers.*

4.2.2. Sandhi pair 2: assimilatory Sandhi 3 vs. dissimilatory Sandhi 4

Figure 4 depicts the sandhied form of T3.T1 against its claimed target surface tone pattern T2.T1, with averaged data across the 18 speakers. The growth curve analysis revealed a significant difference between the T3 in T3.T1 and the T2 in T2.T1 in overall f0 mean (p < .001 ***), f0 slope (p < .001 ***), but no significant difference in sharpness of the centered peak (p > .05). Hence the surface T3 in T3.T1 is evidently non-neutralizing with the underlying T2 in T2.T1, suggesting that assimilatory Sandhi 3 applies in a gradient way.

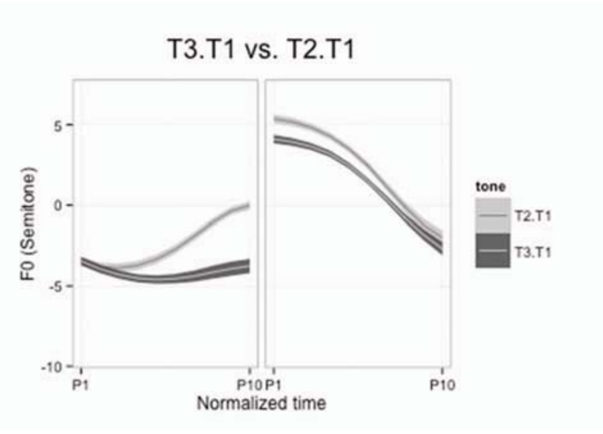

Figure 4. *Sandhied T3.T1 vs. its claimed surface T2.T1 in Nanjing Mandarin (with normalized time) by 18 speakers.*

Figure 5 depicts the sandhied form of T3.T3 against its claimed target surface tone pattern T2.T3, averaged across 18 speakers. The growth curve analysis revealed that the T3 on syllable 1 (T3 (syl1)) in T3.T3 and the T2 in T2.T3 are not significantly different in f0 mean (p > .05), f0 slope (p > .05) or sharpness of the centered peak (p > .05). Therefore there was no reason to reject the null hypothesis that the surface form of T3 (syl 1) in T3.T3 is neutralizing with the T2 in T2.T3, which suggests that assimilatory Sandhi 4 applies in a categorical way.

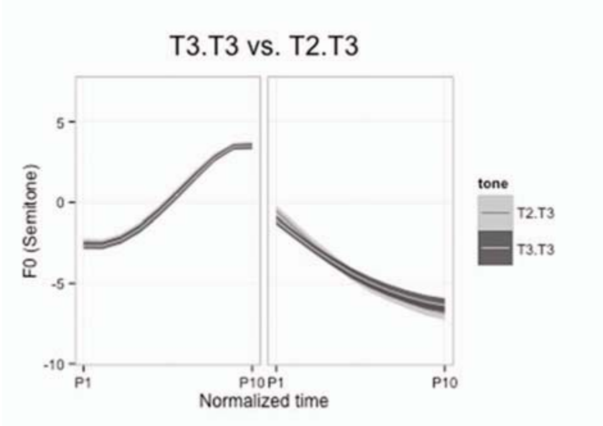

Figure 5. *Sandhied T3.T3 vs. its claimed surface T2.T3 in Nanjing Mandarin (with normalized time) by 18 speakers.*

5. General discussion and conclusions

The current study aimed to investigate the categoricalness/gradience of tone sandhi application in assimilatory and dissimilatory tone sandhi processes in a single language, in order to avoid interference from language-specific factors. We selected in Nanjing Mandarin two pairs of comparable assimilatory and dissimilatory tone sandhi processes, Pair 1 involving T1 vs. T4 tonal contrast and Pair 2 involving T3 vs. T2 contrast, based on previous researchers' observations.

In Sandhi Pair 1, assimilatory Sandhi 1 was observed to be a near-categorical tone sandhi process; dissimilatory Sandhi 2 to be a gradient process. In Sandhi Pair 2, assimilatory Sandhi 3 was found to be a gradient sandhi process; dissimilatory Sandhi 4 to be a categorical sandhi process.

The observations of near-categoricalness in assimilatory Sandhi 1 and gradience in assimilatory Sandhi 3 are both congruent with [10]'s prediction that assimilatory processes are allowed to apply in a gradient fashion. Though we found that dissimilatory Sandhi 4 is a categorical process, our observation of gradience in dissimilatory Sandhi 2 suggests that dissimilatory changes do not necessarily apply in a discrete way; instead, the dissimilatory changes can also occur in a gradient way, which contradicts the prediction by the "hyper-correction" account that dissimilatory changes should always be categorical.

In addition to the insights we provided on the motivations of assimilatory/dissimilatory tone sandhi processes and their categoricalness/gradience in a general sense, this study also presented an updated measurement of the lexical tones and the four selected sandhi rules of Nanjing Mandarin based on data elicited from younger speakers, compared with prior studies on this dialect [16-18].

A follow-up study will add the inter-speaker variation data to reveal more about the nature of the tone sandhi rules in discussion.

6. References

[1] Chen, M.Y., *Tone sandhi: patterns across Chinese dialects*. 2000, Cambridge: Cambridge University Press.

[2] Grammont, M., *Traité de phonétique*. 1933, Paris: Delagrave.

[3] Johnson, L., "Dissimilation as a natural process in phonology". *Stanford Occasional Papers in Linguistics*, 1973. 3: pp. 45-56.

[4] Ohala, J., "The phonetics of sound change", in *Historical Linguistics: Problems and Perspectives*, edited by C. Jones. 1993, London: Longman. pp. 237-278.

[5] Myers, J. and J. Tsay, "Neutralization in Taiwan Southern Min tone sandhi", in *Interfaces in Chinese Phonology: Festschrift in Honor of Matthew Y. Chen on His 70th Birthday*, edited by Y.-C.E. Hsiao, Hsu, H.-C., Wee, L.-H., Ho, D.-A.. 2008, Academia Sinica: Taiwan. pp. 47-78.

[6] Chao, Y.-r., "A system of tone letters". *Le Maitre Phonetique*, 1930. 45: pp. 24-27.

[7] Peng, S.-H., "Lexical versus 'phonological' representations of Mandarin sandhi tones", in *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, edited by M.B. Broe and J.B. Pierrehumbert. 2000, Cambridge: Cambridge University Press. pp. 152-167.

[8] Myers, J. and J. Tsay, "Investigating the phonetics of Mandarin tone sandhi". *Taiwan Journal of Linguistics*, 2003. 1(1): pp. 29-68.

[9] Yuan, J. and Y. Chen, "3rd tone sandhi in standard Chinese: A corpus approach". *Journal of Chinese Linguistics*, 2014. 42(1): pp. 218-237.

[10] Zhang, J. and J. Liu, "Tone sandhi and tonal coarticulation in Tianjin Chinese". *Phonetica*, 2011. 68: pp. 161-191.

[11] Kiparsky, P., "The phonological basis of sound change", in *The Handbook of Phonological Theory*. 1995, Oxford: Basil Blackwell. pp. 640-670.

[12] Ohala, J., "The listener as a source of sound change", in *Papers from the Parasession on Language and Behavior*, edited by Masek, Hendrick, Miller. 1981, Chicago: Chicago Linguistic Society. pp. 178-203.

[13] Ohala, J., "Experimental phonology", in *The Handbook of Phonological Theory*, edited by J. Goldsmith. 1995, Oxford: Blackwell. pp. 713-722.

[14] Garrett, A., "Sound change", in *The Routledge Handbook of Historical Linguistics*, edited by C. Bowern and B. Evans. 2015, London: Routledge. pp. 227-248.

[15] Jatteau, A. and M. Hejná, "Dissimilation can be gradient: evidence from Aberystwyth English". *Papers in Historical Phonology*, 2016. 1: pp. 359-386.

[16] Liu, D., ed. Nanjing Fangyan Cidian (Dictionary of the Nanjing Dialect). 1995, Nanjing: Jiangsu Education Publishing House.

[17] Jiangsusheng Difangzhi Bianzuan Weiyuanhui (Editorial Committee of Jiangsu Province Chorography), *Jiangsu Shengzhi Fangyanzhi (Jiangsu Province Dialect Records)*. 1998, Nanjing: Nanjing University Press.

[18] Song, Y., *Nanjing Fangyan Shengdiao Shiyan Yanjiu (Experimental Study on Tones of Nanjing Dialect)*. 2006, Master thesis, Nanjing: Nanjing Normal University.

[19] Ma, Q., "Nanjing Fangyan Liangzizu Liandu Biandiao de Youxuanlun Fenxi (An OT Solution to Bisyllabic Tone Sandhi in Nanjing Dialect)". in *Yuyan Yanjiu (Studies in Language and Linguistics)*, 2009. 29(1): pp. 27-32.

[20] Boersma, P. and D. Weenink, *Praat: Doing phonetics by computer* 2009.

[21] Andruski, J.E. and J. Costello, "Using polynomial equations to model pitch contour shape in lexical tones: An example from Green Mong". *Journal of the International Phonetic Association,* 2004. 34: pp. 125-140.

[22] Mirman, D., *Growth curve analysis and visualization using R*. 2014, Boca Raton: CRC Press.

[23] Li, Q. and Y. Chen, "An acoustic study of contextual tonal variation in Tianjin Mandarin". *Journal of Phonetics*, 2016. 54: pp. 123-150.

[24] Zhang, J. and Y. Meng, "Structure-dependent tone sandhi in real and nonce disyllables in Shanghai Wu". *Journal of Phonetics*, 2016. 54: pp. 169-201.