THE EFFECT OF WORD FREQUENCY AND NEIGHBOURHOOD DENSITY ON TONE MERGE

Yifei Bi, Yiya Chen, & Niels Schiller

Leiden University Center for Linguistics; Leiden Institute for Brain and Cognition
y.bi@hum.leidenuniv.nl, yiya.chen@hum.leidenuniv.nl, n.o.schiller@hum.leidenuniv.nl

ABSTRACT

This paper studies the effect of word frequency and neighbourhood density on lexical tone merge in Dalian Mandarin. Monosyllabic words with two lexical falling tones (i.e. Tone1 and Tone 4) are produced by 60 native speakers from two different generations (middle-aged vs. young). The stimuli consist of three conditions: high neighbourhood density with high word frequency (HDHF), high neighbourhood density with low word frequency (HDLF) and low neighbourhood density with high word frequency (LDHF). Syllable duration as well as the F_0 curve and F_0's velocity profile of tonal contours are quantitatively analysed through linear-mixed modelling and functional data analysis. Results show that Tone 1 and Tone 4 are near-merged in Dalian Mandarin. Word frequency and neighbourhood density show no effect on duration, but do affect the concave and convex of F_0 curves and the slope of F_0's velocity profile, which suggests their role in the tone merge process.

Keywords: word frequency, neighbourhood density, tone merge, Dalian Mandarin, functional data analysis

1. INTRODUCTION

Lexical tone merge, a process of sound change, is a tonal contrast neutralization phenomenon that has been investigated through word production in various languages [9, 11, 13, 15]. Earlier studies have examined from different aspects about the effect of word frequency on tone merge in languages such as Southern Min [15], Manange [9] and Cantonese [13]. Research on syllable contraction of Southern Min [15] showed that there was a positive correlation between this kind of tone merge and word frequency. [9] showed that word frequency affected the tone merge: high-frequency words had reductive pitch effects and low-frequency words had the retention of tones. [13] reported the effects of word frequency in Cantonese tone merge: low-frequency words were more affected in tone merge from the aspect of analogical sound change. However, it is not clear how languages may differ in the effect of word frequency on potential sound merge.

Neighbourhood density has also been considered to have effects on word production in tasks such as tip-of-the-tongue elicitation, tongue twister production and picture naming [6, 12, 14, 20]. [20] summarised that words with high neighbourhood density were produced with shorter latency and higher accuracy than those with low neighbourhood density. The duration of words in HDHF condition tended to be shorter than that in LDHF condition [6]. [14] reported that the duration of HDLF words was shorter than LDHF words. To our knowledge, the effect of neighbourhood density on tone merge has not yet been studied.

We further investigate the effect of these two known factors (word frequency and neighbourhood density) in lexical tone production. Our testing case is Dalian Mandarin, which has been reported to have two falling lexical tones that are claimed to be merging. Our specific research questions are whether the two falling lexical tones are merged and whether word frequency and neighbourhood density affect the tone merge process.

Dalian Mandarin belongs to the Jiao-Liao Mandarin group, which is one of the primary Chinese dialects. Compared with Beijing Mandarin which has a high-level tone (Tone 1) and a falling tone (Tone 4), Dalian Mandarin has a slightly different tonal system [7, 11, 19]: Tone 1 (T1) and Tone 4 (T4) both have a falling tonal contour. [11] argued that T1 and T4 were almost merged despite durational differences between the two, based on data from only one native speaker. [7] did an acoustic analysis of data from three-generation speakers (four for each generation reported) and showed that T1 and T4 had the tendency of merge. Note that in addition to the lack of sufficient number of participants, previous studies on Dalian Mandarin only considered high frequency data. And neighbourhood density was not considered at all.

This paper presents a relatively comprehensive study on the effect of word frequency and neighbourhood density on tone merge in Dalian Mandarin with data from 60 participants. We quantitatively analysed the tonal F_0 contours with both functional data analysis and linear-mixed modelling to reveal the effects of word frequency and neighbourhood density in tonal contrast neutralization.

2. METHODOLOGY

2.1. Participants

30 middle-aged (mean: 50; SD: 3.6) and 30 young (mean: 22; SD: 3.6) native speakers of Dalian Mandarin participated in the experiment. All have normal vision and no history of speech disorders according to their self-report. They were paid to participate in the experiment.

2.2. Stimuli

90 monosyllabic words were selected, classified into three conditions: HDHF, HDLF, LDHF. Each condition contained 30 monosyllabic words, which consisted of 15 T1 and 15 T4 words, respectively.

According to the corpus of spoken Chinese [3], the logged frequency [1] of high-frequency monosyllabic words in this experiment is between 2.81 and 4.9 and the average is 3.82. The logged frequency of low-frequency monosyllabic words is between 0.6 and 2.0 and the average is 1.53. According to the Modern Chinese Dictionary (6th version, published by The Commercial Press), the average number of homophones calculated manually for high and low neighbourhood density monosyllabic words are 17 and 4, respectively.

Take '安 an1' (1 means T1) and '猜 cai1' as an example. For HDHF monosyllabic word '安 an1', its neighbourhoods include 鞍, 氨, 胺, 庵, 谙, etc. All of these share the same sound 'an1' but they are different Chinese characters. For LDHF monosyllabic word '猜 cai1', it is the only one that is used in daily speech. Thus, '安 an1' is chosen as high neighbourhood density monosyllabic word and '猜 cai1' is chosen as low neighbourhood density monosyllabic word.

2.3. Experiment design and procedure

The experiment was carried out with E-prime 2.0 run on a laptop equipped with a Creative SBX-FI5.1 pro sound card. Participants were asked to read the stimuli in Chinese characters on the computer screen. They were recorded in a quiet room using a condenser microphone, and their recordings were directly stored into the computer's hard disk.

Participants were exposed to a learning phase with four trials using common monosyllabic words (not used in the test) to help them get used to the procedure before they proceeded to the testing phase. The 90 monosyllabic words were divided into six blocks and each block was composed of 15 trials. In the testing phase, participants took a self-pacing break between blocks. The order of the trials was randomised for each participant. There was no repetition for any of the trials.

[1] This value is based on $\log_{10}$ (FREQcount+1), in which FREQcount is the number of times the word appears in the corpus.

2.4. Data analysis

2.4.1. Data preparation

The analysis of all acoustic data was conducted in Praat [2]. All the sound files were manually segmented. The onset and offset of target vowels determined the relevant time intervals for the extraction of duration and F_0 values. F_0 values were sampled at 20 equidistant measurement points using a Praat script. F_0 values were converted into z-scores in order to reduce the participants' variability.

2.4.2. Linear-mixed modelling

The data were modelled using linear-mixed modelling using *R* [16], *lme4* [1] and *lmerTest* [10]. Condition and tone (with interaction) were included as fixed effects. By-subject slopes for the effect of tone and by-item intercept were included as random effects. Likelihood ratio tests were performed to decide the random terms of the model.

2.4.3. Functional data analysis

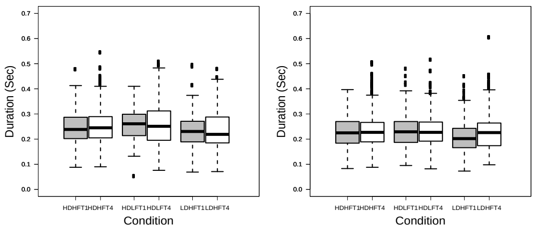
Functional data analysis (FDA) [18] was used to investigate the data of F_0. FDA provides a method to analyse the dataset consisting of entire curves with different durations. Two main procedures were conducted: smoothing with a linear time registration and functional principal component analysis (FPCA).

Smoothing (with roughness penalty) was realised by the B-spines [5]. A linear time registration scaled all the smoothed curves into normalised duration (i.e. 1), which was used for further analysis by FPCA. FPCA provided a model to approximate the (normalised) smoothed curves using the mean curve and a number of Principal Component (PC) curves and their weights (PC scores), based on the formula $f(t) \approx \mu(t) + \sum_{j=1}^{\propto} s_j * PC_j(t)$. $\mu(t)$ is the mean curve, s_j is the PC score (PCs) and $PC_j(t)$ is the corresponding PC curve. The first two PC scores (s_1 and s_2), which represent most of the variation of the smoothed curves, were used for performing linear-mixed modelling [8]. In addition, functional t-test was used for calculating the absolute value of t-statistic in each sampling point of the (normalised) smoothed curves [18]. All FDA analysis was carried out by the R package of '*fda*' [17].

3. RESULTS AND DISCUSSION

3.1. Duration

Figure 1: Duration of T1 and T4 produced by (a) middle-aged speakers and (b) young speakers in three conditions (HDHF, HDLF and LDHF).



(a) (b)

Figure 1 showed the boxplot of the duration of T1 and T4 produced by middle-aged and young speakers in three conditions (HDHF, HDLF and LDHF). We performed a linear-mixed effects modelling for the duration produced by middle-aged and young speakers separately. The results showed that there was no significant difference between T1 and T4 in any of the three conditions produced by either middle-aged or young speakers (p > 0.05). This suggests that T1 and T4 were merged in Dalian Mandarin in terms of duration. Word frequency and neighbourhood density showed no effect on the durational contrast of the two mergers, which is different from [7, 11].

3.2. F_0 curves and F_0's velocity profile

3.2.1. Average curves and functional t-test

To compare T1 and T4, apart from F_0 curves, their velocity (which is the speed of the two falling tones), also reflects useful acoustic features [4]. Therefore, we analysed both F_0 curve and F_0's velocity profile.

Figures 2 and 3 showed the average F_0 curves and their velocity profiles for T1 and T4 in three conditions and the results of functional t-test among young speakers. For the middle-aged speakers, the results were similar. In the figures of functional t-test, dotted lines represent the 0.05 critical values for the t-statistic. The higher one represents a more conservative critical value.

As shown in Figures 2 (a, c, e) and 3 (a, c, e), T1 and T4 had almost the same tendency in curves. The results of functional t-test in Figures 2 (b, d, f) and 3 (b, d, f) showed that there were significant differences in the middle parts of F_0 curves and in the initial parts of F_0's velocity profiles between T1 and T4 in three conditions.

Figure 2: The average of the (normalised) F_0 curves for T1 and T4 in three conditions (a: HDHF; c: HDLF; e: LDHF) and their functional t-test (b: HDHF; d: HDLF; f: LDHF) by young speakers.

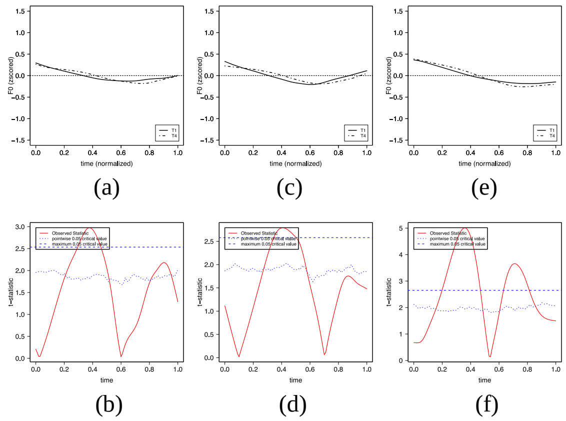
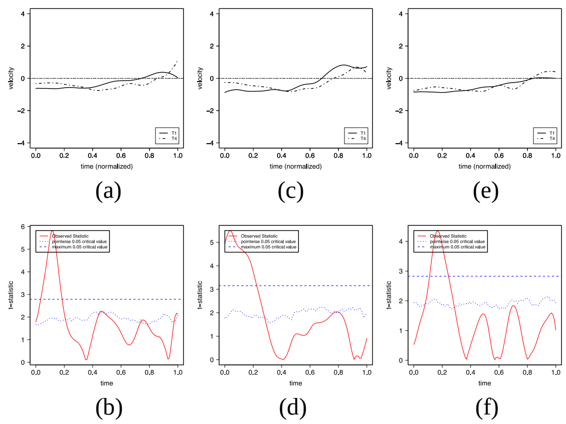


Figure 3: The average of the (normalised) F_0's velocity profile for T1 and T4 in three conditions (a: HDHF; c: HDLF; e: LDHF) and their functional t-test (b: HDHF; d: HDLF; f: LDHF) by young speakers.

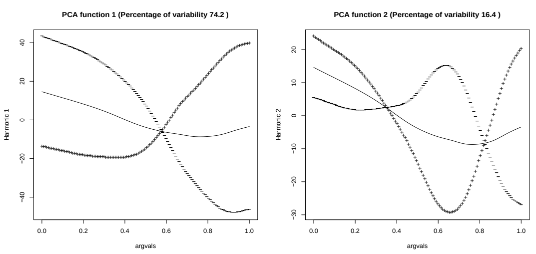


These results indicated that T1 and T4 had been partially merged. Their degree of neutralization varied as a function of word frequency and neighbourhood density. Since the functional t-test only considers the t-statistic of each sampling point in the smoothed curves, we also performed the linear-mixed modelling for PC scores (s_1 and s_2) to investigate tone merge from the whole range and at the same time, take into account the variation due to individual speakers and stimuli items.

3.2.2. Linear-mixed modelling for PC scores

Figures 4a and 4b showed an example of FPCA results for F_0 curves in the condition of HDHF by young speakers. Each panel shown in solid line is the mean curve $\mu(t)$. The ± curves are obtained by adding to or subtracting from $\mu(t)$ the curves (a) $\sigma(s_1) * PC_1(t)$ and (b) $\sigma(s_2) * PC_2(t)$. σ denotes standard deviation. As shown in Figure 4, s_1 mainly altered the slope of the F_0 curves and s_2 altered the concave and convex in the middle part of the curves. The outputs of FPCA indicated that s_1 and s_2 could explain most variation of the F_0 curves in HDHF condition (74.2% and 16.4%, respectively). This made it possible to conduct quantitatively analysis of the effect of s_1 and s_2 on tone merge using the linear-mixed modelling.

Figure 4: The results of FPCA for (a) PC_1 *(t)* and (b) PC_2 (*t*) of the F_0 curves in the condition of HDHF by young speakers.



(a) (b)

We performed a linear-mixed modelling of s_1 and s_2 for middle-aged and young speakers separately. Tables 1 and 2 showed only the significant results. In Table 1, for F_0 curves, opposite to the results of s_1, there was significant difference in s_2 for any of the three conditions among middle-aged and young speakers, respectively. Therefore, both word frequency and neighbourhood density affected s_2, i.e. the concave and convex of F_0 curves. We interpreted it that they influenced the turning point of production. Through the summary of the model of F_0 curves, the average s_2 of T4 was smaller than that of T1, which means F_0 curves of T4 were more concave.

For F_0's velocity profile, there was a significant interaction between condition and tone on both s_1 and s_2. Therefore, we analysed how each condition contributed to the effect on tone merge. Table 2 showed that high neighbourhood density affected s_1 (slope of the curve) more than low neighbourhood density. Both high and low-word frequency affected s_1 in F_0's velocity profile. In the summary of the model, in HDHF and HDLF conditions, s_1 of T4 was smaller than that of T1, which means T4 fell more steeply than T1 from the whole range.

In summary, T1 and T4 were partially merged from the aspect of F_0 in Dalian Mandarin. Both word frequency and neighbourhood density affected the degree of tone merge, in terms of both the concave and convex of the F_0 curves and the slope of the F_0's velocity profile.

Table 1: Summary of linear-mixed modelling for F_0 curves on any of the three conditions (which are not shown).

Generation	PCs	df	*F*	*p*
Middle	s_2	1	37.47	<0.001
Young	s_2	1	30.76	<0.001

Table 2: Summary of linear-mixed modelling for F_0's velocity profile.

Generation	Condition	PCs	df	*F*	*p*
Middle	HDHF	s_1	1	14.43	<0.001
Middle	HDLF	s_1	1	4.36	<0.05
Young	HDHF	s_1	1	5.64	<0.05
Young	HDLF	s_1	1	5.10	<0.05
Middle	LDHF	s_2	1	9.73	<0.01
Young	HDHF	s_2	1	5.98	<0.05

4. CONCLUSION

Results of the acoustic analysis of lexical tonal production data from 60 speakers show that T1 and T4 in Dalian Mandarin are near-mergers. Different from [7, 11], no significant durational difference was found between the two tones across different frequency and neighbourhood density conditions (HDHF, HDLF and LDHF). However, in terms of F_0, the two lexical tonal contrasts remained across the three conditions in F_0 contours' turning point. The two tones also showed a significant difference in the velocity profile of their F_0 contours although only in monosyllabic words with high neighbourhood density (i.e. HDHF and HDLF).

Previous studies on tonal neutralization emphasise the effect of word frequency in contrast maintenance [9, 13, 15]. Our results, however, showed that in Dalian Mandarin, it is the neighbourhood density that exerted an effect. Future study will investigate the possible effects of word frequency and neighbourhood density on T1 and T4 merge in disyllabic words.

5. ACKNOWLEDGEMENT

We would like to thank the China Scholarship Council for the fellowship to YB, the European Research Council for the ERC-Starting Grant (206198) to YC, as well as the support from Leiden University Center for Linguistics, Dalian No. 76 Middle School and all the participants.

6. REFERENCES

[1] Bates, D., et al. 2013. *lme4: Linear mixed-effects models using Eigen and S4. R package 1. 1-7. Online:* http://CRAN.Rproject.org/package=lme4.

[2] Boersma, P., Weenik, D. 2013. *Praat: Doing phonetics by computer [Computer program]. Version 5.4.04, retrieved 28 December 2014. Online:* http://www.praat.org/retrieved.

[3] Cai, Q., Brysbaert, M. 2010. SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS One*, 5(6): p. e10729.

[4] Cheng, C., Chen, J.-Y., Xu, Y. Year. An acoustic analysis of Mandarin Tone 3 sandhi elicited from an implicit priming experiment. *Proc. Fourth TAL* Netherlands.

[5] De Boor, C. 2001. A practical guide to splines.

[6] Gahl, S., Yao, Y., Johnson, K. 2012. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *J. Memory and Language*, 66(4), 789-806.

[7] Gao, Y. 2007. *Dalian Fangyan Shengdiao Yanjiu*. Liaoning Normal University Press.

[8] Gubian, M., Torreira, F., Boves, L. 2015. Using Functional Data Analysis for investigating multidimensional dynamic phonetic contrasts. *J. Phonetics*, 49(0), 16-40.

[9] Hildebrandt, K.A. 2007. *Manange tone: scenarios of retention and loss in two communities*. Ph.D. thesis, UMI Ann Arbor.

[10] Kuznetsova, A., Brockhoff, P., Christensen, R. 2013. *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package 2.0. Online:* http://CRAN.R-project.org/package=lmerTest.

[11] Liu, T.-h. 2009. The Phonology of Incomplete Tone Merger in Dalian. *UC Berkeley Phonology Lab Annual Report*, 435-461.

[12] Luce, P.A., Pisoni, D.B. 1998. Recognizing spoken words: The neighborhood activation model. *Ear and hearing*, 19(1), 1.

[13] Mok, P.P., Zuo, D., Wong, P.W. 2013. Production and perception of a sound change in progress: Tone merge in Hong Kong Cantonese. *Language Variation and Change*, 25(03), 341-370.

[14] Munson, B., Solomon, N.P. 2004. The effect of phonological neighborhood density on vowel articulation. *J. Speech, Language, and Hearing Research*, 47(5), 1048.

[15] Myers, J., Li, Y. 2009. Lexical frequency effects in Taiwan Southern Min syllable contraction. *J. Phonetics*, 37(2), 212-230.

[16] R Core Team 2014. *R: a language and environment for statistical computing* [Version 3.1.2]. Vienna, R Foundation for Statistical Computing.

[17] Ramsay, J., et al. 2013. *fda: Functional data analysis. R package 2.3.8. Online*: http://CRAN.R-projcet.org/package=fda.

[18] Ramsay, J.O. 2006. *Functional data analysis*. Wiley Online Library.

[19] Song, X. 1963. *Liaoning Yuyin Shoulue (A sketch of Liaoning Phonology).*

[20] Yao, Y. 2011. *The effects of phonological neighborhoods on pronunciation variation in conversational speech*. Ph.D. thesis, University of California, Berkeley.