# Analyzing Nanjing Tones and Sandhi: statistical modelling methods

Si Chen, Caroline Wiltshire, Bin Li, & Ratree Wayland

## The Questions:

Major issues in documenting/analyzing a tone language:

1) what is the tone inventory?
2) what are the tone sandhi rules?
   (& disentangle from tonal coarticulation)

Modelling **underlying pitch targets** can answer both.

## Introduction

### A. Nanjing Dialect

- Nanjing city, Jiangsu Province, east coast of China.
- Population over 5.3 million
- Hongchao subgroup within Jianghuai dialects.

### B. Nanjing Tone Descriptions

| Tone | Syllable | Liu (1995) | Sun (2003) | Meaning |
|---|---|---|---|---|
| T1 | /fu/ | 41 | 31 | 'skin' |
| T2 | /fu/ | 24 | 13 | 'hold' |
| T3 | /fu/ | 11 | 22/212 | 'rotten' |
| T4 | /fu/ | 44 | 44 | 'negative' |
| T5 | /fu/ | 5 | 55 | 'fortune' |

| Liu (1995) | Sun (2003) |
|---|---|
| T1→T4/_T1 (41→44/_41) | T1→T4/_T1 (31→44/_31) |
| T2→T3/_T5 (24→11/_5) | T2→T3/_T5(13→22/_5) |
| T3→T2/_T1 (11→24/_41) | T3→T2/_T1(22→13/_31) |
| T3→T2/_T3 (11→24/_11) | T3→T1/_T3(22→31/_22) |
| T4→T1/_T5 (44→41/_5) | None |
| T5→3/_T5 (5→3/_5) | T5→T4/_T5 (5→44/_5) |

### C. Purposes of Study

- Determine inventory/sandhi rules based on acoustic data using statistical modelling

## Methods

### Participants & Stimulus Materials

- 12 native speakers of Nanjing (6F/6M), age 35-65
- 660 mono-σ tones (11 σs*5 tones*12 speakers), 360 di-σ tones (5 di-σs*6 combos*12 speakers).
- Real words; most chosen from Liu (1995)

### Measurement and Normalization

- Digital recording, sampling rate 44.1kHz.
- Praat script extracted time-normalized F0 values, 20 time points in each V, 25.6ms analysis window.
- Logarithmic Z-score normalization on F0 values (Rose 1987; Zhu 1999)

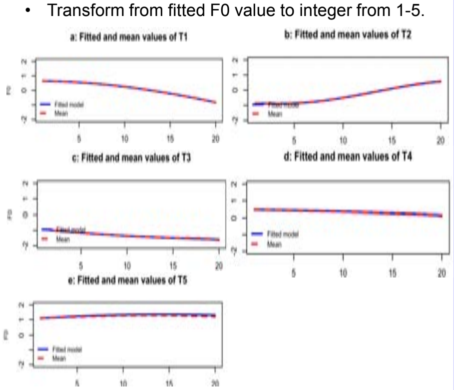
## Statistic Models & Results

### A. Modelling

- Underlying pitch targets (UPTs) are 'the smallest articulatorily operable units associated with linguistically functional pitch units such as tone and pitch accent' (Xu & Wang, 2001:321).
- Prom-On, Xu & Thipakorn (2009) offer a quantitative target approximation model.
- Chen et al. (2017) show this model can distinguish gradient, phonetic perturbation from categorical, phonologized changes.

### B. Results: Citation Tones

- To transform F0 to Chao numbers: build on previous methods (Shi 1990; Zhu, Shi & Wei 2012), but use UPTs.
- Calculate sample quantiles for all fitted values of mono-σ tones; find cut-off values for each
- Transform from fitted F0 value to integer from 1-5.




Fitted (solid) and mean (dotted) values of Nanjing T1-T5; fitted values are from the optimal model for each tone.

- Four tones best modelled by linear underlying target, but T2 requires quadratic term.
- Cut-off values corresponding to each quantile: 20%: -0.90; 40%: -0.02; 60%: 0.39; 80%: 0.97.

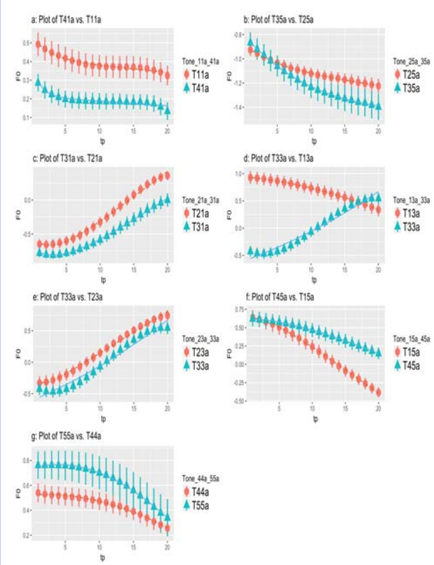
| Tone | Initial F0, normalized (Chao's #) | Final F0, normalized (Chao's #) | Values in literature | Our Tone values |
|---|---|---|---|---|
| 1 | 0.65 (4) | -0.82 (2) | 31/41 | 42 |
| 2 | -0.85 (2) | 0.94 (4) | 24/13 | 24 |
| 3 | -0.94 (1) | -1.59 (1) | 22/212/11 | 11 |
| 4 | 0.45 (4) | 0.15 (3) | 44 | 43 |
| 5 | 1.08 (5) | 1.31 (5) | 5/55 | 55 |

## Statistic Models & Results, cont.

### C. Results: Sandhi

- Compare UPTs of tone in sandhi position to potential tone it turns into in same di-syllabic context.
- Fit optimal model to obtain coefficients of UPTs for each speaker; compare UPT of each tone in sandhi position to that of citation tone in that position
- Non-significant difference in the coefficients is taken as evidence of neutralized sandhi tones.

Note: label T41a means T4 when it's first in a T4+ T1 sequence




- The UPT of T3 before T3 neutralized with that of citation T2 (Liu 1995), not that of T1 (Sun 2003).
- Modelling supports Sun (2003)'s proposed T5 → T4/_T5; UPTs of T4 and T5 neutralize before T5.

| Liu (1995) | Sun (2003) | Our Sandhi rules |
|---|---|---|
| T1→T4/_T1 (41→44/_41) | T1→T4/_T1 (31→44/_31) | **T1→T4/_T1 (42→43/_42)** |
| T2→T3/_T5 (24→11/_5) | T2→T3/_T5 (13→22/_5) | **T2→T3/_T5 (24→11/_5)** |
| T3→T2/_T1 (11→24/_41) | T3→T2/_T1 (22→13/_31) | **T3→T2/_T1 (11→24/_42)** |
| T3→T2/_T3 (11→24/_11) | T3→T1/_T3 (22→31/_22) | **T3→T2/_T3 (11→24/_11)** |
| T4→T1/_T5 (44→41/_5) | None | **T4→T1/_T5 (43→42/_5)** |
| T5→3/_T5 (5→3/_5) | T5→T4/_T5 (5→44/_5) | **T5→T4/_T5 (5→43/_5)** |

## Conclusions

We apply statistical techniques to Nanjing data to resolve discrepancies in earlier accounts of its tones & sandhi

We show that statistically testing underlying pitch targets reveals the categorical nature of tone sandhi.

Acoustic studies involving statistical testing and modelling of UPTs can provide a quantitative basis for a more precise transcription of tones and sandhi, and

These techniques can be applied to any undocumented or understudied language to provide a tonal analysis based on productions by the speech community.

### Selected References


Chen, S., C. Zhang, A. McCollum, & R. Wayland. 2017. Statistical modelling of phonetic and phonologised perturbation effects in tonal and non-tonal Languages. *Speech Communication* 88, 17-38.

Li, Q., & Y. Chen. 2016. An acoustic study of contextual tonal variation in Tianjin Mandarin. *Journal of Phonetics* 54, 123 – 150.

Liu, D.Q. 1995. *Nanjing Fangyan Cidian (A dictionary of the Nanjing Dialect)*. Nanjing: Jiangsu Education Publisher.

Liu, D. Q. 1997. *Nanjing Dialect Phonetic Profile*. Shanghai: Shanghai Education Publisher.

Prom-On, S., Y. Xu, & B. Thipakorn. 2009. Modeling tone and intonation in Mandarin and English as a process of target approximation. *JASA 125*(1), 405 – 424.

Song, Y.D. 2006. *Nanjing Fangyan Shengdiao Shiyan Yanjiu*. (The acoustic study of tones in Nanjing dialect). Master's thesis, Nanjing Normal University, Nanjing.

Sun, H.X. 2003. Nanjing fangyan shengdiao de ruogan wenti. (A discussion on the tones of Nanjing dialect). *Journal of Nanjing Xiaozhuang College, 1, 34 – 40.*

Xu, C.,M. Liu. C. Zhang, S. An, W. Yu & J.M. Chen. 2007. The spatiotemporal dynamics of rapid urban growth in the Nanjing metropolitan region of China. *Landscape Ecology 22*, 925-937.


## Further details on model & statistics

Prom-On et al. (2009) argue control of vocal fold tension reflects two antagonistic forces (activation of cricothyroid & thyroarytenoid muscles); should be modeled as 2nd order linear system at least.

Following Chen et al. (2017), we used a critically damped linear system and considered only 2nd and 3rd order linear systems. The optimized model was chosen from the following two models:

1) Second order linear system with linear underlying targets $at + b$

$$f_0(t) = \beta e^{-\lambda t} + at + b$$

where $f_0(t)$ stands for f0 values and $\lambda$ the rate of approaching the UPT; *a* is the slope of the UPT, and *b* is its intercept.

2) Third order linear system with linear underlying targets $at + b$

$$f_0(t) = (c_1 + c_2 t + c_3 t^2) e^{-\lambda t} + at + b$$
$$c_1 = f_0(0) - b$$
$$c_2 = f'_0(0) + c_1\lambda - m$$
$$c_2 = (f''_0(0) + 2c_2\lambda - 2c_1\lambda^2)/2$$

where $f_0(t)$ stands for f0 values, $\lambda$ the rate of approximating the UPT, *a* the slope of the UPT, *b* its intercept, $c_1$, $c_2$ and $c_3$ are transient coefficients determined by initial f0 values, velocity and acceleration.

Model with the least Akaike's Information Criterion (AIC) chosen as optimal, as least AIC indicates best fit (Kim & Timm, 2006).

For each tone, used nonlinear regression to fit the four models to find optimal one and to calculate coefficients of UPTs from f0 values extracted from our speech production data.

Used non-parametric Wilcoxon signed rank test to determine whether obtained coefficients of UPTs of tones were significantly different.

## Questions or Comments:

E-mail: sarahchen@polyu.edu.hk or wiltshir@ufl.edu