Tonal Assignment in English Loanwords in Mandarin Chinese

Eleanor Glewwe
eleanorglewwe@ucla.edu
University of California, Los Angeles
CLS 51 | 24 April 2015

Question: How are loanwords borrowed from English, a language without lexical tone, assigned tones in Mandarin Chinese, a language with lexical tone?

Approach: A corpus study of English loanwords in Mandarin

Principal Findings:
- Voicing in English is the most reliable predictor of tone in Mandarin
- English stress plays a secondary role, if any, in tonal adaptation

1. Background
- A common thread in English-to-Chinese tonal adaptation: the relationship between stress in English and lexical tone in Chinese languages

1.1 English Stress
- English stressed syllables have longer duration, higher F0, and greater amplitude than unstressed syllables (Cutler 2008)
- These properties can also distinguish Chinese tones → stress-to-tone?

1.2 Tonal Adaptation in Cantonese (Hao 2009)
- Overall, follows stress-to-tone principles:

English Stress	Cantonese Tone
Stressed (primary or secondary)	High
Pre-tonic unstressed	Mid
Post-tonic unstressed	Low

1.3 Tones of Mandarin

Tone	Chao Digits	Pinyin	Relative Duration in Citation Form (Whalen & Xu 1992)
First (high)	55	mā	shortest
Second (rising)	35	má	longer
Third (falling-rising)	214	mǎ	longest
Fourth (falling)	51	mà	shortest

- Mandarin has tonotactic gaps (e.g. no *dán*)
  - Syllables with unaspirated onsets tend not to appear with rising tone (Wu 2006)
  - With rare exceptions, Mandarin doesn't fill tonotactic gaps in loanword adaptation

1.4 Previous Work on Tonal Adaptation in Mandarin

- Corpus study by Wu (2006) of 100+ established loanwords still used in present-day Taiwan

Table 1: Tonal Assignment of English Loanword Syllables in Mandarin
(Based on Wu 2006)

Syllable Type	**Tone**	**Explanation**
Sonorant Onset	Rising	Stressed in English, but sonorant onsets lower F0 (Hombert, Ohala, & Ewan 1979)
Aspirated Stop/Affricate Onset	Rising	Stressed in English, but Mandarin aspiration lowers F0
Unaspirated Stop/Affricate Onset	High	Stressed in English

  - Generalizations confined to initial syllables of disyllabic adaptations of English trochees
  - Onsets refer to onsets of Mandarin adapted syllables, not of English syllables
  - Caveat: Wu's discussion of tonal assignment and aspiration is based on 35 syllables, only 7 of which have aspirated onsets
- Experiment by Chang & Bradley (2012)
  - Native Mandarin speakers asked to adapt disyllabic English nonce words (trochees and iambs) into Mandarin
  - English nonce words have expected Mandarin segmental adaptations: syllable structure (CV, CVN), onset sonority, and onset aspiration all match
  - Only stressed syllables analyzed
  - Falling tone most common overall
  - CVs with sonorant onsets less likely to get high tone than CVs with obstruent onsets
  - CVs with aspirated stops more likely to get rising tone than CVs with unaspirated stops

Table 2: Tonal Assignment of English Loanword Syllables in Mandarin

Syllable Type	Tone	
	Wu 2006	**Chang & Bradley 2012**
Sonorant Onset	Rising	**High less likely** relative to obstruent onsets
Aspirated Stop/Affricate Onset	Rising	**Rising** second most likely (after falling)
Unaspirated Stop/Affricate Onset	High	**High** second most likely (after falling)

2. The Present Study's Corpus

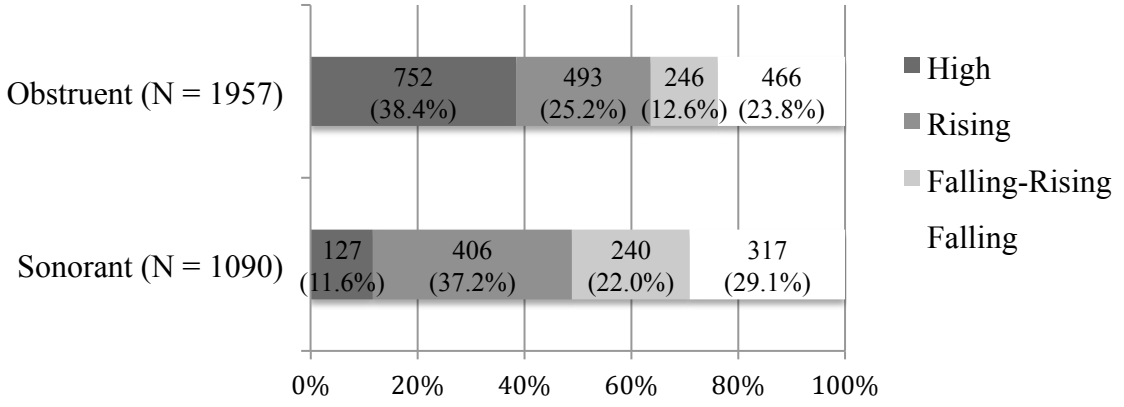
- 3,660 syllables from 1,372 English loanwords in Mandarin
- Sources:
  - A corpus from a dissertation on segmental adaptation (Dong 2012): 1,194 English borrowings including 292 place names and 577 first names (85% of the syllables in my corpus)
  - Corpus from Wu (2006) (contributed 58 syllables to my corpus)
  - Online Chinese dictionary MDBG: search for "loanword" (13% of the syllables in my corpus)
- Corpus originally included 1,551 loanwords, but I excluded 179 for a variety of reasons:
  - Semantic influence (e.g. 蹦极 *bèngjí* 'bungee jumping': 蹦 means 'jump', 极 means 'extreme')
  - Loanwords from languages other than English or borrowed from English into Mandarin through another Chinese dialect
- Also excluded 60 syllables subject to third tone sandhi (ˇ ˇ → ˊ ˇ), e.g. first syllable in 法老 *fǎlǎo* 'pharaoh')
- 10 syllables in corpus can only bear one tone (*ā, fó, hēi, lè, miù, rì, sè, sēn, tè,* and *téng*) → most analyses based on a reduced corpus of 3,456 syllables without these syllables

3. Testing Past Claims

3.1 Wu (2006), Chang & Bradley (2012): English stressed syllables whose Mandarin adaptations begin with sonorants less likely to get high tone, more likely to get rising tone than those whose Mandarin adaptations begin with obstruents

- My corpus reveals the same pattern across *all* syllables (considering sonority of English segment to which Mandarin onset corresponds)

Fig. 1: Tones of Loanword Syllables by Mandarin Onset's Sonority in English

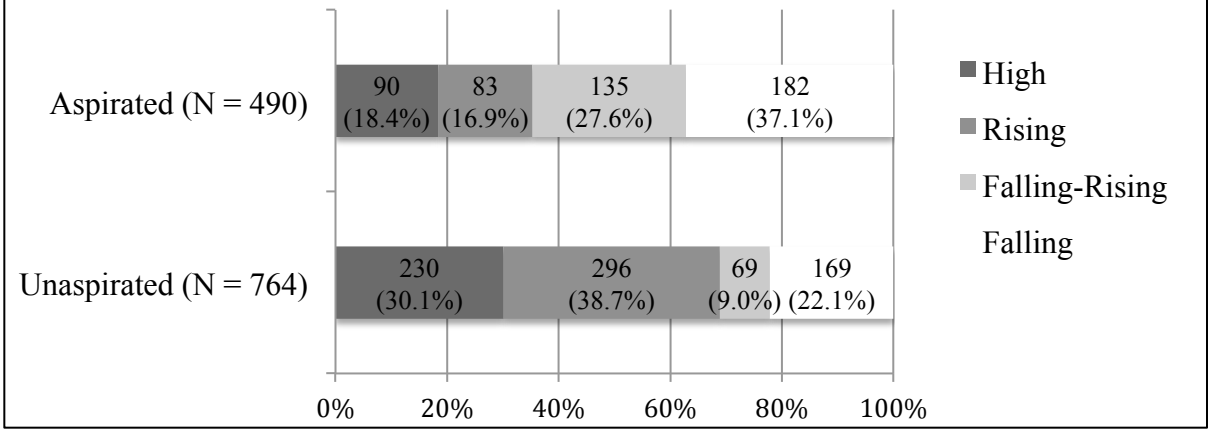


- Loanword syllables (including epenthetic syllables) whose Mandarin onsets correspond to English obstruents most often get high tone; loanword syllables whose onsets correspond to English sonorants most often get rising tone, rarely get high tone

3.2 Wu (2006), Chang & Bradley (2012): English stressed syllables more likely to get rising tone if their onset in Mandarin is aspirated than if it is unaspirated

- I was unable to replicate this finding
- Loanword syllables with aspirated onsets get rising tone *less* often than loanword syllables with unaspirated onsets

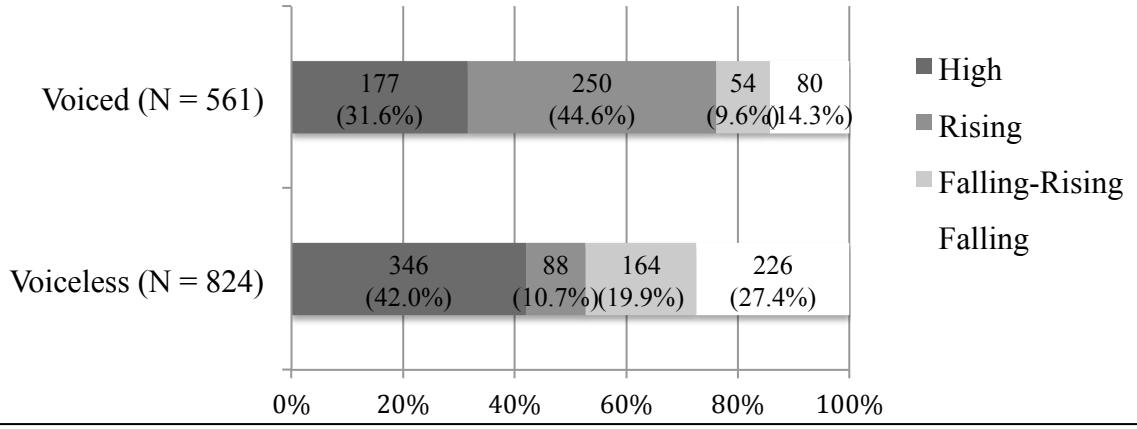
Fig. 2: Tones of Loanword Syllables by Onset's Aspiration in Mandarin



4. Voicing

- Underexplored in previous literature
- Turns out to play a role in tonal adaptation
- Examining the effect of voicing in non-epenthetic loanword syllables whose onsets come from English obstruents:

Fig. 3: Tones of Non-Epenthetic Loanword Syllables with Onsets from English Obstruents by English Voicing



- Large differences in tone distributions are significant: $\chi^2 = 179.8687$, $p < 0.001$
- English voiced obstruent → rising tone
- English voiceless obstruent → high tone, then falling tone (both start high)
- Pattern consistent with lowering effect of voiced obstruents on F0 (Hombert, Ohala, & Ewan 1979)
- Why the relatively high rate of high tone assignment in syllables from voiced English obstruents?
  - Voiced English stops and affricates tend to be adapted with Mandarin unaspirated sounds, and syllables with unaspirated onsets tend to have gaps in rising tone
  - That rising tone is still preferred for these syllables testifies to strength of lowering effect
- We saw a preference for high tone in loanword syllables with onsets from English obstruents (Fig. 1), but 61.2% of those obstruents were voiceless → association actually between *voiceless* obstruents and high tone

- Comparing Wu's and my findings:

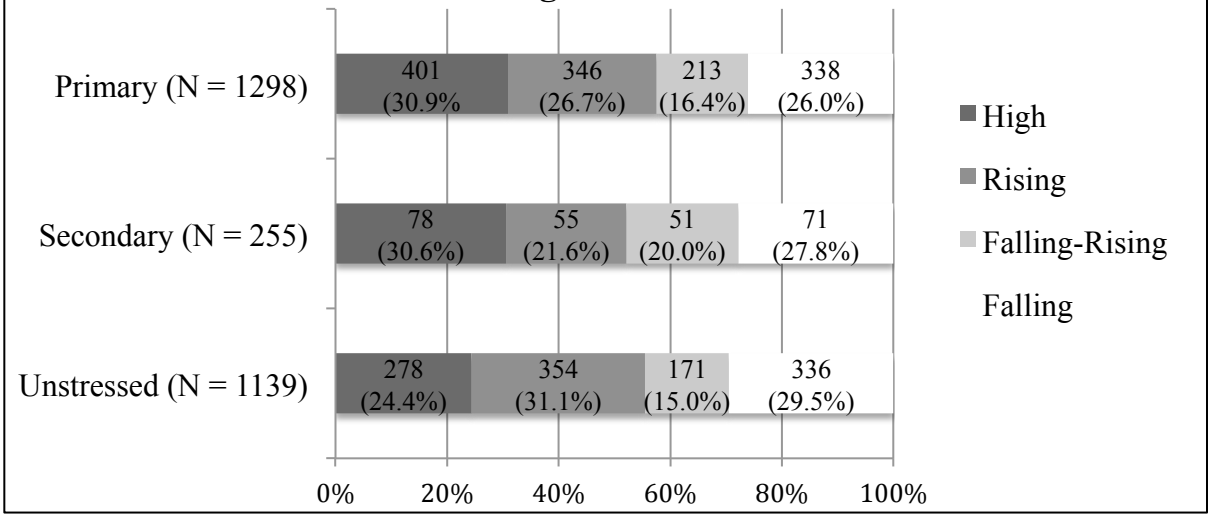
Table 3: Tonal Assignment of English Loanword Syllables in Mandarin

Wu 2006		Present Study			
Onset Type (in Mandarin)	**Tone**	**Onset Type (in Mandarin)**	**Tone**	**Onset Type (in English)**	**Tone**
Sonorant	Rising			Sonorant	Rising
Aspirated Stop/Affricate	Rising	Aspirated Stop/Affricate	Falling	Voiceless Obstruent	High
Unaspirated Stop/Affricate	High	Unaspirated Stop/Affricate	Rising	Voiced Obstruent	Rising

5. Stress-to-Tone Adaptation

5.1 Are the different English stress levels associated with particular tones?

Fig. 4: Tones of Non-Epenthetic Loanword Syllables by English Stress



- Overall significant differences: $\chi^2 = 24.0637$, $p < 0.001$
- Primary vs. secondary—no significant difference: $\chi^2 = 4.0221$, $p = 0.2591$
- Primary vs. unstressed—significant difference: $\chi^2 = 16.6695$, $p < 0.001$
- Secondary vs. unstressed—significant difference: $\chi^2 = 12.9904$, $p < 0.005$
- Primary- and secondary-stressed syllables pattern together, unstressed syllables pattern differently
- Stressed syllables prefer high tone—matching high pitch?

- Unstressed syllables get high tone less, get rising tone more—also matching pitch?

5.2 A different approach: Do segmentally identical syllables get different tones depending on their English stress level?

Table 4: Tones of Particular Mandarin Proper Syllables by English Stress

Mandarin Syllable	Stress	High	Rising	Falling-Rising	Falling
li	Primary (N = 37)	0	0	15	22
	Secondary (N = 1)	0	0	1	0
	Unstressed (N = 124)	0	0	50	74
la	Primary (N = 17)	17	0	0	0
	Secondary (N = 1)	1	0	0	0
	Unstressed (N = 41)	41	0	0	0
luo	Primary (N = 40)	1	28	0	11
	Secondary (N = 3)	0	2	0	1
	Unstressed (N = 16)	0	11	0	5
xi	Primary (N = 22)	22	0	0	0
	Secondary (N = 1)	1	0	0	0
	Unstressed (N = 32)	32	0	0	0
ni	Primary (N = 9)	0	9	0	0
	Unstressed (N = 45)	0	45	0	0
ai	Primary (N = 46)	26	0	0	20
	Secondary (N = 1)	1	0	0	0
	Unstressed (N = 3)	3	0	0	0
wei	Primary (N = 28)	14	10	4	0
	Secondary (N = 4)	1	0	3	0
	Unstressed (N = 18)	4	14	0	0
di	Primary (N = 13)	1	4	0	8
	Unstressed (N = 32)	0	21	1	10
ya	Primary (N = 5)	0	0	0	5
	Secondary (N = 4)	0	0	0	4
	Unstressed (N = 36)	0	0	2	34

- Little evidence for stress affecting tone in individual syllables
- Case of *li*: proportions of *li*s that receive falling-rising and falling tone about the same whether *li* adapts an English stressed syllable or an unstressed one
- Two syllables show evidence of sensitivity to stress:

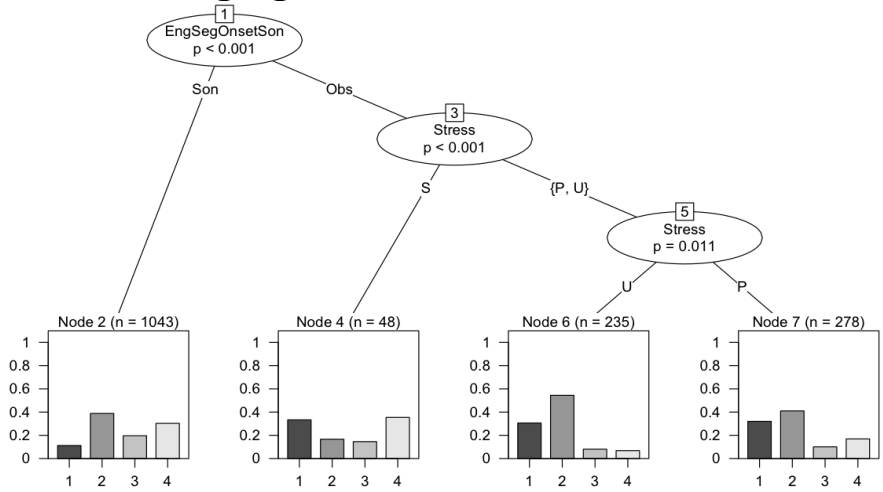
- *Wei*: tone distributions of *wei*s from primary-stressed, secondary-stressed, and unstressed syllables differ significantly (Fisher's Exact Test: $p < 0.001$), primary stress → high tone, unstressed → rising tone
- *Di*: tone distributions of *di*s from primary-stressed and unstressed syllables differ just significantly (Fisher's Exact Test: $p < 0.05$), primary stress → falling tone, unstressed → rising tone

- Results of two approaches to investigating effect of stress conflict: unclear how much English stress influences tonal assignment

6. Exploring the Determinants of Tonal Assignment With A Decision Tree

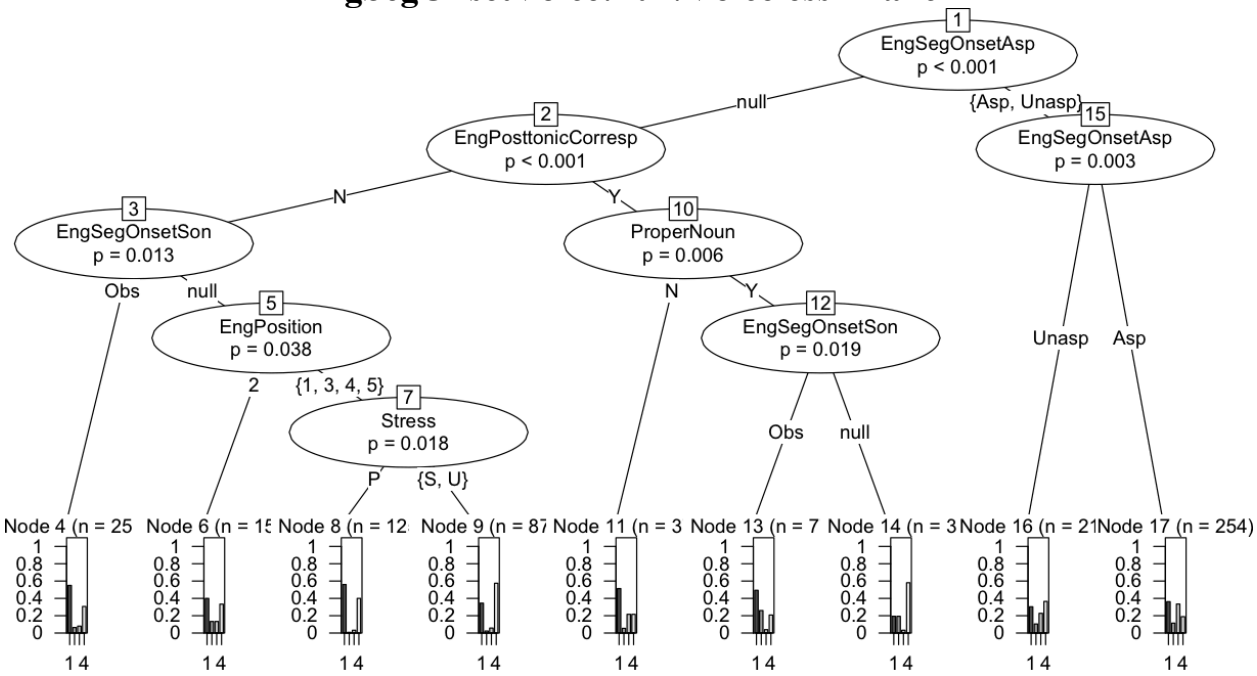
- Another way of uncovering which factors significantly affect Mandarin tonal adaptation
- `ctree()` function from the R package `party` (Hothorn, Hornik, & Zeileis 2006) generates binary-branching decision trees through recursive partitioning
- Potential uses:
  - Provide confirmation for my findings
  - Reveal other determinants of tone
  - Show that certain determinants only have an effect in a subset of the corpus
- Tree split in two parts because too large to show all at once
- First split by EngSegOnsetVoice (voicing of English segment to which adapted syllable's onset corresponds; values: Voiced, Voiceless, null)

Figure 5: Decision Tree for All Non-Epenthetic Loanword Syllables—EngSegOnsetVoice:Voiced Branch



- Recall: both sonorants and voiced obstruents in English associated with rising tone—all terminal nodes but one show rising tone preference, as expected
- Syllables whose onsets come from sonorants split off: rising tone preferred
- Syllables whose onsets come from voiced obstruents split by stress
  - Oddly, secondary-stressed syllables split off
  - Then primary-stressed and unstressed split: both prefer rising tone, but primary-stressed get falling tone more → interaction of voicing and stress

Figure 6: Decision Tree for All Non-Epenthetic Loanword Syllables — EngSegOnsetVoice:null/Voiceless Branch



- First split by EngSegOnsetAsp divides syllables with onsets from voiceless stops and affricates from syllables with onsets from voiceless fricatives and syllables from onsetless English syllables
- Syllables with onsets from voiceless stops and affricates split by aspiration: both rarely get rising tone
- Node 2 and below exhibit a series of puzzling splits
  - Terminal nodes show little rising tone, preference for high and/or falling tone
  - High tone preference probably driven by onsets from voiceless fricatives
  - Falling tone preference probably driven by loanword syllables from onsetless English syllables (many *aò*s from /əʊ/ and *yà*s from /ə/)

- Major contributions of decision tree:
  - Confirmation of findings about effects of English voicing and sonority
  - First split by English voicing → strengthens claim that voicing is single best predictor of tonal assignment
  - Suggests stress may only matter for loanword syllables with onsets from English voiced obstruents

7. Conclusion

- Voicing of English obstruents is best predictor of tonal assignment:
  - Voiceless obstruents associated with high tone, voiced obstruents with rising tone (even though Mandarin syllables with unaspirated onsets tend to have rising tone gaps)
  - Low-level phonetic details of English syllables affect their tonal adaptation and may override native tone-onset type associations
- While stress is almost the sole determinant of tonal assignment in English loanwords in Cantonese, its effect is weaker in Mandarin
  - Some overall significant differences in tones assigned by stress level and trends that make phonetic sense, but don't seem to hold for individual syllables
  - Decision tree suggests stress may only affect the tones of loanword syllables with onsets from English voiced obstruents
- Future directions:
  - Further exploration of relationship between tonal assignment in loanwords and tonal patterns in the native lexicon
  - Corroborating corpus findings with experimental evidence

Many thanks to Kie Zuraw for her advice on this project and to the UCLA Phonology Seminar audience for helpful feedback.

References

Chang, Zhuting & Dianne Bradley. (2012). Tonal Assignment for English Loanwords in Mandarin: An Experimental Approach. Talk at *CUNY Conference on the Segment*. CUNY Graduate Center.

Cutler, A. (2008). Lexical Stress. In D. B. Pisoni & R. E. Remez (Eds.), *The Handbook of Speech Perception* (pp. 264-289). Oxford: Blackwell.

Dong, X. (2012). *What borrowing buys us: A study of Mandarin Chinese loanword phonology*. Ph.D. Dissertation. University of Utrecht.

Hao, Y.-C. (2009). Tonal Adaptation of English Loanwords in Cantonese. *Proceedings of the 21st North American Conference on Chinese Linguistics (NACCL-21*), ed. by Yun Xiao, 42-54. Smithfield, RI: Bryant University.

Hombert, J.-M., Ohala, J. J., & Ewan, W. G. (1979). Phonetic Explanations for the Development of Tones. *Language, 55*(1), 37-58.

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, *15*(3), 651-674.

MDBG. (2014). *Chinese-English Dictionary.* www.mdbg.net.

Whalen, D. H. & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, *49*(1), 25-47.

Wu, H. I. (2006). Stress to Tone: A Study of Tone Loans in Mandarin Chinese. In. F.-F. Hsieh & M. Kenstowicz (Eds.) *MIT Working Papers in Linguistics* 52: *Studies in Loanword Phonology*, 227-253.