Lexical Tone Effects on Voice Onset Time in Cantonese

Holman Tse (University of Pittsburgh)
hbt3@pitt.edu

Workshop in Innovations in Cantonese Linguistics
The Ohio State University
March 16, 2012

Introduction

- Two major research areas on the relationship between phonetics and phonology.

1. Factors affecting Voice Onset Time (VOT) without a loss in phonological contrasts
2. Tone and consonant interaction

- Could tone be a factor in affecting VOT?
- How do tonal effects on VOT relate to consonant and tone interaction?

Specific Research Questions

1. Does tone have an effect on VOT in Cantonese?
2. If so, what kind of an effect? (i.e. would historically voiced stops correspond to lower VOT?)
3. Could these observed tonal effects be predicted by either static or dynamic measurements of F0?

Voice Onset Time (VOT)

- Defined as the time interval between the stop burst and the onset of voicing of the following vowel (Lisker and Abramson 1964)
- Acoustic measurement as a way of defining stop categories
- Three types of stops defined by VOT:
  - Voiced: VOT < 0
  - Voiceless Unaspirated (short-lag): small positive VOT
  - Voiceless Aspirated (long-lag): large positive VOT
- Even though stops may be defined by a range of VOT values, the VOT value required to maintain stop contrasts is not fixed
- Much research has explored factors that affect VOT values without causing a loss in phonological contrasts

What Factors affect VOT

- **Rate of Speech** (Kessinger and Blumstein 1997)
  - faster speaking rate → VOT approaches 0 (for voiced and long-lag stops only)
- **Place of Articulation** (Cho and Ladefoged 1999)
  - Cross-linguistically, velars have higher average VOT than labials and alveolars
- **Vowel Height** (Klatt 1975)
  - Stops preceding high vowels have longer VOT
- **Prosodic Position** (Cho & McQueen 2005)
  - Language specific effects
- **Lexical Tone**
  - Language specific effects????

Tonal effects on VOT

Language	Stop categories	Reported VOT effects
Kera (Chadic)	Voiceless (historically voiced/voiceless)	L < M < H
Mazatec	Pre-nasalized, short-lag, long-lag	L < H (for pre-nasalized, defined as positive)
Shanghainese	Voiced, short-lag, long-lag	LM, MM < H, HL (for long-lag)
Korean (younger speakers)	Short-lag, medium-lag, long-lag	L < H
Mandarin	Short-lag, long-lag	213, 25 > 55, 51
Hakka	Short-lag, long-lag	Short (checked) tones < long tones
Taiwanese	(Voiced?), Short-lag, long-lag	LL, LR, LF > HL, HF

Reference list at the end

Consonant and Tone Interaction

- Tonal effects on VOT suggest tone affecting consonants, contrary to Hyman and Schuh's (1974) claim: "Consonants affect tone but tone does not affect consonants"
- BUT this refers to diachronic changes that lead to changes in phonological categories
- Supported by synchronic articulatory studies showing how voiced stops lower F0 while voiceless stops have a raising effect in tonal and non-tonal languages
- This suggests that if there is an association btwn VOT and tone, lower tonal onsets would correspond to lower VOT (but tone ≠ F0)

Is this what we actually see?

- In Kera, Mazatec, Shanghainese, and Korean, yes. But these languages all have different stop categories.
- Mandarin and Taiwanese have higher VOT for lower onset tones
- For Hakka, checked vs non-checked is what matters
- Strongest relationship between voiced stops and low F0/tone
- What about another language that lacks voiced stops?

Modern Cantonese tone system

	Middle Chinese Tone A	Middle Chinese Tone B	Middle Chinese Tone C
High Register Tones (Yin) Voiceless Stops (unchanged)	55 both long-lag and short lag	25 both long-lag and short lag	33 both long-lag and short lag
Low Register Tones (Yang) *Voiced Stops > Voiceless Stops	21 Long-lag stops	23 Long-lag stops	22 Short-lag stops

Adapted from Haudricourt (1972)

- Would tones with historically voiced stops have lower VOT?
- Aspirated stops: maximum of 5 contrasting tones
- Unaspirated stops: maximum of 4 contrasting tones (with a few exceptions for reduplication and onomatopoeia)
- 6 contrasting tones unattested in words beginning with a stop

Methodology

6 subjects (5 male, 1 female)

- native Hong Kong Cantonese speakers in their early 20's
- Less than 4 years in the U.S. at the time of recording.

Recordings

- made with solid state recorder in sound-proof booth
- 20 words spoken in a carrier phrase:
- 10 repetitions each, 200 tokens per subject, but only 8 words used for present study, the rest used for different experiments/distraction

Stimuli

Tone	pʰa	pa
55	趴地 [pʰaː tei] 'lie down on the floor'	巴士 [paː si] 'bus'
25	(豬)扒飯 (t͡sy) [pʰaː faːn] 'pork chop with rice'	把手 [paː sɐu] 'handle'
33	怕怕 [pʰaː pʰaː] 'scared'	霸位 [paː wɐi] 'to hog seats'
21	爬山 [pʰaː saːn] 'climb a mountain'	爸爸 [paː paː] 'father'

Carrier Phrase

XY, 我會讀 X 俾你聽

XY, ŋɔ23 wui23 tuk22 X pei35 nei23 tɛŋ55

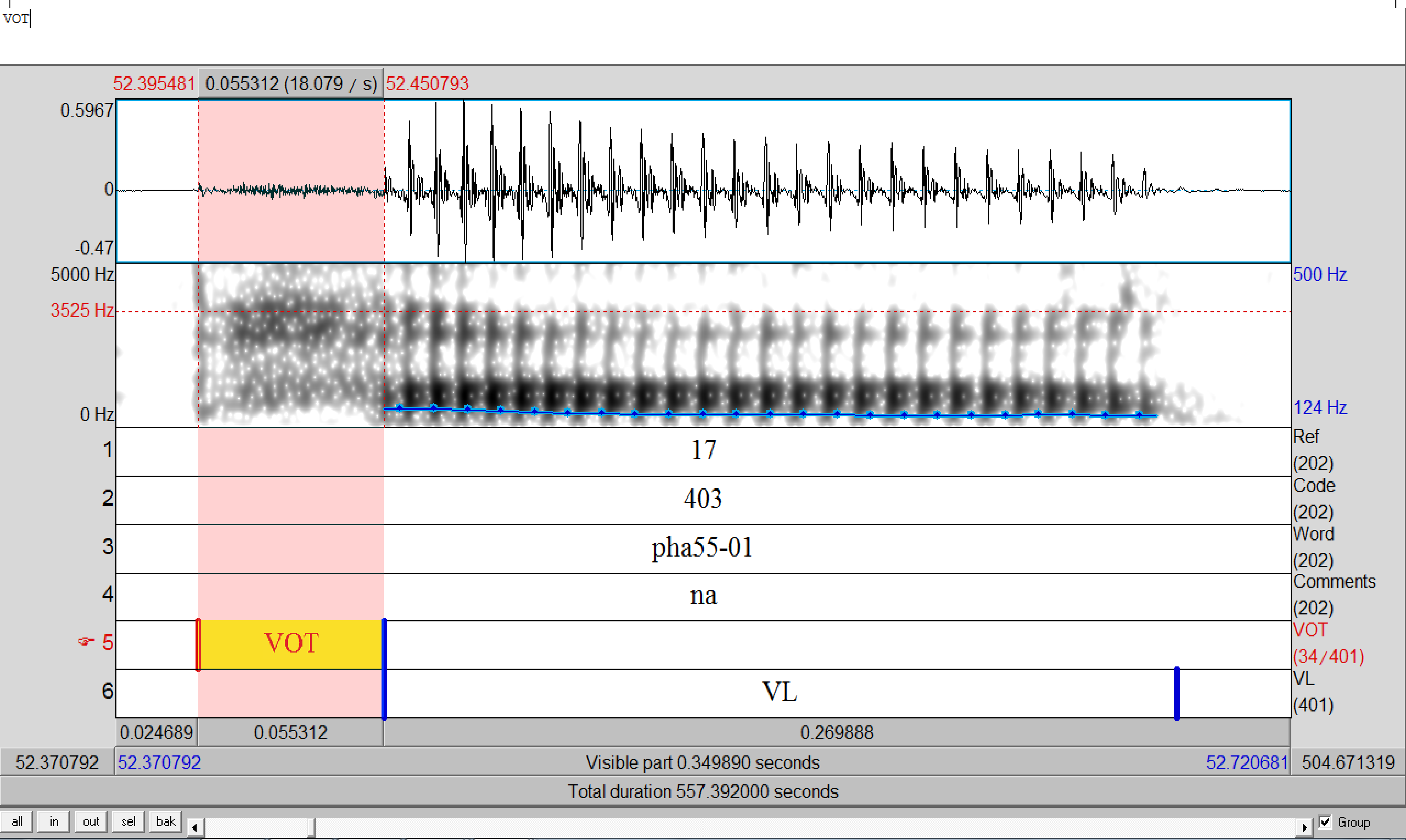
'XY, I will read X to you'

Note: 6 contrastive tones with the same stop unattested in Cantonese

Measurements

- VOT and vowel boundaries segmented for each token placed in the carrier phrase
- Researcher auditorily checked for correct tone pronunciation
- A few cases of speaker self-correction, context phrase token used when there was no self-correction
- PRAAT script used on segmented textgrids to measure VOT and vowel duration, separate PRAAT script used to measure time-normalized F0 values at 10 times points per token

Sample Textgrid labels



Analysis

- Word Duration = VOT + Vowel Duration
- Normalized VOT = $\frac{VOT}{Word\ Duration}$
- Normalized VOT used to control for rate of speech effects
- The difference between the maximum and minimum F0 values for each token was calculated (F0_difference)
- ANOVA performed to test the effects of tone category and subject on VOT
- Pearson's Correlation Test used for F0

Results

Short-lag stops

Tone	Normalized VOT Avg	Std. Deviation	N
55	0.0551	0.02382	60
33	0.0555	0.02834	60
25	0.0602	0.03353	60
21	0.0610	0.03357	60
All Tones	0.0580	0.03002	240

ANOVA test of tone and Normalized VOT for all tokens:
$p = 0.833$, n.s.

- Not surprising (cf. Kessinger and Blumstein 1997)

Long lag stops

Tone	Normalized VOT Avg	Std. Dev. of Normalized VOT	VOT average (in seconds)	Std. Dev. of VOT	N
55	0.2487	0.06648	0.0627	0.19569	60
33	0.2619	0.06912	0.0657	0.02152	60
25	0.3002	0.06110	0.0771	0.01932	60
21	0.3240	0.07782	0.0785	0.02346	60
All Tones	0.2837	0.07476	0.0701	0.02202	240

Repeated Measure ANOVA test for tone and Normalized VOT for all tokens:
$p < 0.001$
ANOVA test for tone and actual VOT:
$p < 0.001$

Pair-wise analysis: long-lag stops

Tonal Pairs	p =	Significance
55 & 25	0.1118	n.s.
55 & 33	0.9816	n.s.
55 & 21	0.0046	**
25 & 33	0.1680	n.s.
25 & 21	0.8133	n.s
33 & 21	0.0219	*

General grouping: 55, 33, (25) < (25), 21

Comparison of ANOVA and F0 Onset vs Normalized VOT Correlation test results

Subject	ANOVA p value	ANOVA significance	Pearson Coefficient	p value of correlation	Correlation Significance
1 (male)	< 0.000	**	-0.053	0.744	n.s.
2 (male)	0.001	**	-0.326	0.040	*
3 (male)	<0.000	**	-0.350	0.027	*
4 (male)	0.003	**	-0.001	0.997	n.s.
5 (male)	0.02	*	-0.313	0.049	*
6 (female)	0.469	n.s.	0.129	0.426	n.s.
All Subjects	<0.000	**	-0.210	0.001	*

Correlation test of actual VOT for all subjects: $p = 0.249$, n.s.

F0 Correlation Test for all time points

Subject	1	2	3	4	5	6	7	8	9	10
1	n.s.	n.s.	n.s.	n.s.	n.s.	*	n.s.	n.s.	n.s.	n.s.
2	*	*	*	*	**	**	**	**	**	**
3	*	*	*	**	**	*	*	*	*	*
4	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
5	*	*	*	**	**	**	*	**	*	*
6	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
All	*	**	**	**	**	**	**	**	**	**

* ($p < 0.05$)
** ($p < 0.01$)
n.s. (not significant)

F0 contour vs VOT correlation

- Correlation test of F0_difference vs Normalized VOT for all 240 long-lag tokens
  - Correlation coefficient= 0.022, p=0.738, n.s.
- Correlation test of F0 standard deviation vs Normalized VOT for all 240 long-lag tokens
  - Correlation coefficient = 0.015, p=0.820, n.s.

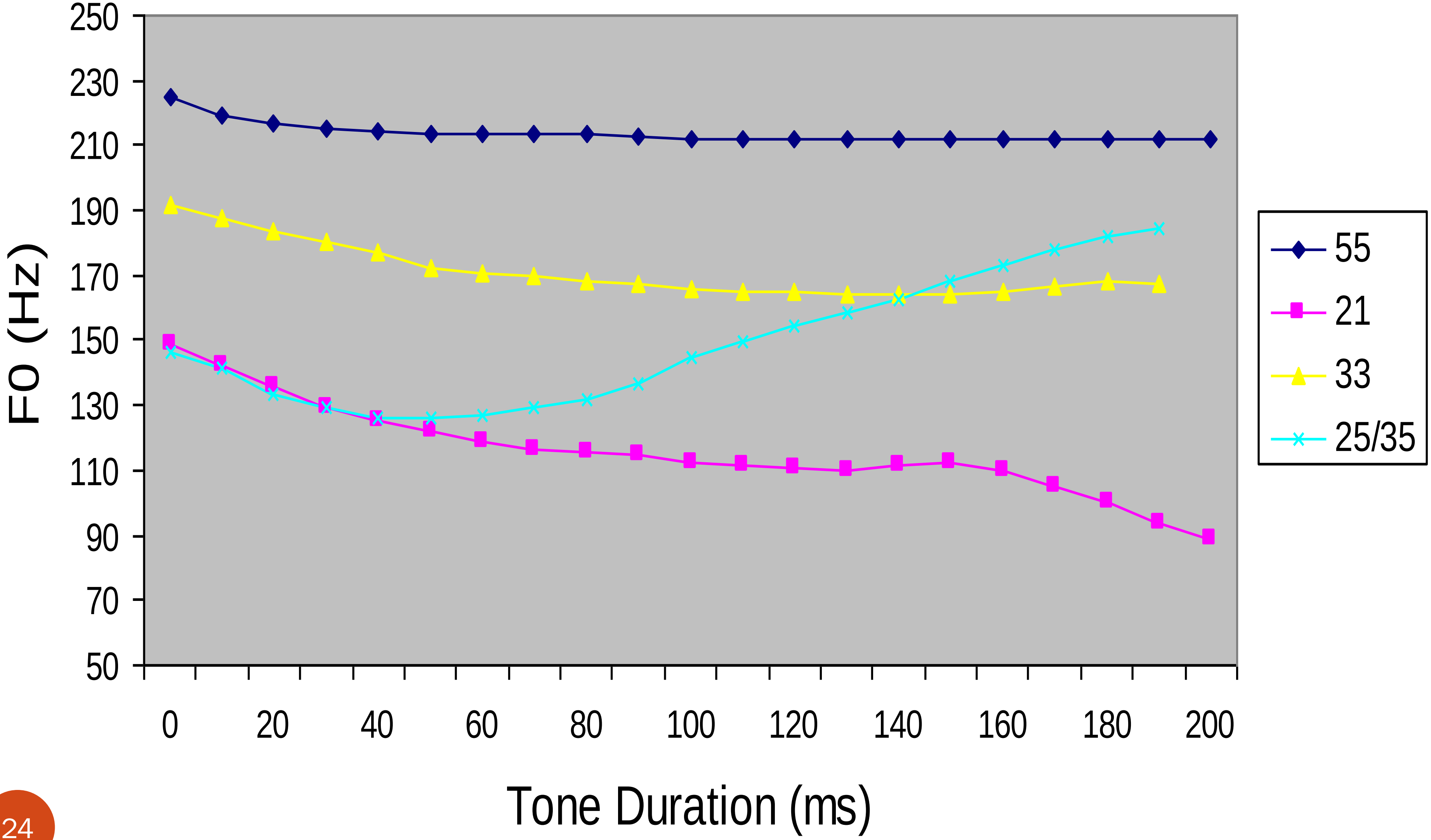
Discussion

- Tone has a significant effect on VOT for long-lag stops in Cantonese (w/o a loss in phonological contrasts between stops)
- The 21 tone is associated with higher VOT than for the other tones. Less certainty about the 25 tone.
- The effect is roughly (but not exactly) inversely correlated with F0
- The degree of contour F0 movement has no effect
- Tone is a stronger predictor of VOT than F0
  - If F0 were the actual cause, why not a significant difference btwn 55 and 33, but a significant difference between 33 and 25, 21.
  - This is the opposite of what would be predicted based on previous F0/tone consonant interaction studies.

Discussion (continued)

- Results from the Cantonese are the opposite of Korean, Kera, Shanghainese, but similar to Mandarin and Taiwanese.
- Voiceless stops are less constrained than voiced stops thus this may allow for more cross-linguistic variability in how they are produced
- VOT is a potential secondary cue in recognition of tonal distinctions (21 vs the other tones)
- VOT differences could be a consequence of creaky voice
- Consonant affects F0 but not the tone, F0 effects are local for voiceless stops and do not affect the actual tone category

Sample Tone Contours (taken from Subject 2)



Follow up research

- Perceptual Tests in which listeners are asked to identify tonal category for tokens in which VOT values are manipulated
- More subjects especially female speakers
- Expand token list to include different environments
- Measuring spectral tilt (is creaky voice involved for 21 tone)?
- Other tonal languages (can broader generalizations be made?)

References

- Chen, L., Peng, L., & Chao, K., (2009). The effect of lexical tones on voice onset time. *Proceedings of the 11th IEEE International Symposium on Multimedia, 552-557.*
- Cho, T. and Ladefoged, P. (1999). Variations and Universals in VOT: Evidence from 18 Languages. *Journal of Phonetics* 27, 207-229.
- Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics, 33(2), 121-157. doi: 10.1016/j.wocn.2005.01.001*
- Haudricourt, A. 1972. Two-way and three-way splitting of tonal systems in some Far-Eastern languages. (trans. C. Court). In J. Harris and R. Noss (eds.), Tai Phonetics and Phonology. Bangkok: Central Institute of English Language, Mahidol University 58-86.
- Herrera Z., E. (2003). Tono, VOT y sonoridad en el mazateco de Santa Clara, Oaxaca. In *La tonía: dimensiones fonéticas y fonológicas. Herrera Z., Esther and Pedro Martín Butragueño (eds.) El Colegio de México Centro de Estudios Lingüísticos y Literarios. 61-74.*
- Hyman, L. M., & Schuh, R. G. (1974). Universals of Tone Rules: Evidence from West Africa. *Linguistic Inquiry, 5(1), 81-115.*
- Kessinger, R. and Blumstein, S. (1997). Effects of speaking rate on Voice Onset Time in Thai, French and English. *Journal of Phonetics* 25: 143-168.
- Klatt, D. (1975). Voice Onset Time, Frication, and Aspiration in Word-Initial Consonant Clusters. *Journal of Speech and Hearing Research* 18: 686-706.
- King, L., & Schiefer, L. (1990). Tone Effect on Voice-Onset Time (VOT)? -Results from Shanghai. *Forschungsberichte -Institut Fur Phonetik Und Sprachliche Kommunikation Der Universitat Munchen. (28), 253-295.*
- Lai, Y. (2004). The effect of aspiration on fundamental frequency in Taiwanese syllables. Unpublished M.A. Thesis. University of Kansas.
- Pearce, M. (2009). Kera tone and voicing interaction. *Lingua 119: 846-864.*
- Silva, D. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* (23) 287–308.