AN ARTICULATORY, ACOUSTIC, AND AUDITORY STUDY
OF BURMESE TONE

A Dissertation
Submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Linguistics

By
James Frederick Gruber, M.S.

Washington, D.C.
June 15, 2011
© 2011 James Gruber
AN ARTICULATORY, ACOUSTIC, AND AUDITORY STUDY
OF BURMESE TONE

James Frederick Gruber, M.S.
Thesis Advisor: Elizabeth Zsiga, Ph.D.

ABSTRACT

This dissertation investigates the complex tonal contrast in the modern Burmese language. The four Burmese tones are reported to make multi-way distinctions in pitch, phonation, duration, intensity, vowel quality, and syllable structure, but the literature is frequently conflicted about the association of these qualities with each tone. Bradley (1982) has described Burmese as a register, rather than tonal, language due to the bundling of prosodic qualities headlined by contrastive phonation types. However, contrastive phonation in Burmese has eluded detection in numerous instrumental phonetic studies (Thein Tun 1982, Ladefoged et al 1988, Watkins 1997). Production and perception experiments were conducted in order to first clarify the phonetic description, and in turn build a phonological model of the contrast.

Audio, electroglottographic, and aerodynamic recordings of ten native speakers were collected and analyzed for Duration, F0, and Phonation type (as measured by Open Quotient of the glottal waveform, rates of oral airflow, and spectral tilt – all measured dynamically over the syllable rhyme). The perception experiment studied the forced-choice listener identifications of stimuli re-synthesized to controlled levels of each dimension.
Major findings include (a) the contrast in phonation type is two-way (two breathier lax tones vs. two creaky-voiced tones), (b) breathy and creaky phonation are temporally dynamic in Burmese – produced systematically only at the vowel offset, and (c) the phonation contrast is neutralized in juncture while pitch contrasts are maintained.

The results indicate that a purely Tone or Register analysis of Burmese is inappropriate, in the sense that pitch contours or phonation type cannot alone serve as the basis of contrast. Instead, the present state of affairs is argued to represent an intermediate stage of tonogenesis, where multiple phonetic properties exist side-by-side to reinforce suprasegmental categories. The analysis is cast in Optimality theoretic terms, which act to highlight the phonological dilemma introduced by a seemingly redundant distribution of properties. The study has implications for the typology of Tone and Register systems, as it portrays a language in a transitional state between the two – some intrinsic pitch distinctions have been phonologized, but a concomitant voice quality distinction remains integral to the grammar.
Acknowledgements

My name alone is credited with authorship of this dissertation. Yet countless others have worked hard and contributed dearly to the efforts that have produced this work. It would not have been possible without their knowledge, guidance, and support.

This dissertation would simply have not been possible without my Burmese friends in the Washington, D.C. and Seattle, WA areas. They participated out of kindness and an eagerness to share their language and culture. They accepted me into their homes, fed me, and happily worked with me through long hours of not very stimulating data collection. I owe a great thanks to each of them. In particular, I want to thank Merle and Eddie Nakhuda for their role and the many times they hosted me in their home.

I am greatly indebted to the efforts of my committee members, Drs. Elizabeth Zsiga, Marc Brunelle, and Robert Podesva. Lisa Zsiga has had a profound impact on both this work and my development as a linguist. The dissertation exists because of Lisa’s encouragement to pursue and expand the topic from an initial seminar project, and later, to seek funding for the research tools. Beyond this work, Lisa has had a tremendous role in my development as a linguist and cultivating my interest in phonetics, phonology, and tone. From the day I walked into her office as a first year graduate student who thought he wanted to study syntax, she has been nothing but supportive and committed to my cause. I have benefited greatly from her knowledge and her teaching.

This dissertation was greatly improved by the counsel of Marc Brunelle and Rob Podesva, and their individual expertise in voice quality research. Marc’s experience investigating similar phenomena in Southeast Asian languages was invaluable. Every discussion I had with Marc led to significant re-thinking or refinement of a claim. Rob
repeatedly drew my attention to details I had not considered, and valuably redirected me from details I was getting lost in. I never felt more confident about the next step I was taking than after meeting with Rob. I also owe thanks to Dan Loehr who, as the unofficial fourth member of my dissertation committee, worked with me on my early attempts at Burmese and continued to supervise and provide his thoughtful feedback at numerous steps along the way. I am also grateful for the advice of Maria Gouskova. I still struggle with writing, but I know that I am a better writer today because of Maria’s efforts.

I am thankful for the financial support provided by the National Science Foundation, Division of Behavioral and Cognitive Sciences. The research was funded by the NSF Doctoral Dissertation Research Improvement Grant, BCS 0844031.

This research would have been excessively more difficult without the excellent work and the willingness to discuss this work by Justin Watkins, Alexis Michaud, David Bradley, Christian DiCanio, and Pat McCormick. Mietta Lennes, Christian DiCanio, and Cedric Gendrot all generously shared Praat scripts which were critical to tackling my data. I am similarly indebted to my friends Zhaleh Feizollahi, Michael Diercks, Justin Kelly, Rebecca Sachs, Kjell Konis, and Melissa Martz for the many hours they have spent helping me improve this work.

For making my time at Georgetown fruitful in fields of study beyond phonetics and phonology, I want to thank Héctor Campos, Polyvia Parara, and Guiseppe Tosi.

For their friendship, and for proving to me that I could survive life on the East Coast, I would like to particularly thank my brother Joe and his family. Kind thoughts are similarly extended to Jen Sclafani and Steve Hamilton, Carlos Balhana, Chad McHenry, Rachael Albritten, Jon Philips, Kris Concepcion, Dom Parr and many others.
My parents and family deserve special recognition for enabling me to get to this position: Fred and Anne, Young-Wha, Joe, and my comforting canine friend Rocco. I want to express my deepest appreciation for my wife Shannon and the time and effort she has expended on my behalf during the writing of this dissertation. Every setback or trial that I endured, she also endured. With great love, Shannon, I thank you for both your patience and impatience, as both were necessary to either persevere through the long process or to occasionally light a fire. I also want to thank my son, Royal, for changing my life. I haven’t met Royal yet, but he will be born in early 2012 and I am certain he will be an amazing person.

This dissertation is dedicated to my mother, Anne, who passed away in November 2008. Her life, her enthusiasm, and her curiosity continue to inspire me.
Table of Contents

List of Figures xv
List of Tables xxi

Chapter One • Outline of the Problem and the Data 1

0. Introduction .................................................................................................................. 1
1.1 Goal of the Dissertation .............................................................................................. 1
1.2 Description of Burmese Phonology ............................................................................... 3
   1.2.1 Overview of the Language .................................................................................. 3
   1.2.2 Description of Burmese Phonology: Inventory of
       Burmese Phonemes ............................................................................................... 4
   1.2.3 Description of Burmese Phonology: Inventory of
       Burmese Tonemes ................................................................................................. 5
   1.2.3.1 Overview ....................................................................................................... 5
   1.2.3.2 The Low Tone [a] ......................................................................................... 6
   1.2.3.3 High Tone [á] ............................................................................................... 7
   1.2.3.4 Creaky Tone [à̰] ........................................................................................ 9
   1.2.3.5 Checked Tone [àʔ] ...................................................................................... 10
   1.2.3.6 Minor Syllables [ə] ...................................................................................... 12
   1.2.4 Description of Burmese Phonology: Syllable Codas ........................................... 13
   1.2.5 Description of Burmese Phonology: Vowel Quality ........................................... 18
1.3 Typological Interest of the Dissertation .................................................................... 23
   1.3.1 Burmese Dialects and the Language Background
       of Study Participants ............................................................................................... 23
   1.3.2 Tone or Register? .............................................................................................. 26
   1.3.3 Typology of Tone in Southeast Asia ................................................................... 28
   1.3.4 The Tonogenetic Roots of Typological Tone Patterns ....................................... 31
1.4 Outline of the Dissertation .......................................................................................... 37
Chapter Two • Phonetic Properties of Burmese Tone ............... 39
  2.1 Introduction ......................................................................................... 39
  2.2 Intensity ............................................................................................... 42
  2.3 Tone Duration ....................................................................................... 43
  2.4 Pitch .................................................................................................... 48
  2.5 Phonation Type ................................................................................... 54
    2.5.1 The Role of Phonation Type ............................................................ 54
    2.5.2 Creaky Tone, but “creaky” voice? ...................................................... 56
    2.5.3 Phonetic Studies: Impressionistic Descriptions ............................. 58
    2.5.4 Phonetic Studies: Acoustic Measures ............................................. 60
    2.5.5 Phonetic Studies: EGG and the Glottal Waveform ........................ 67
    2.5.6 Phonetic Studies: Aerodynamic Analysis ...................................... 70
  2.6 Summary of Phonetic Literature ....................................................... 72
  2.7 Tonogenesis ......................................................................................... 74
  2.8 Prior Phonological Accounts ............................................................. 76
    2.8.1 Bradley (1982): “Register in Burmese” ........................................... 77
    2.8.4 Green (2005): “Word, Foot, and Syllable Structure in Burmese” .... 81
    2.8.4.1 Overview .................................................................................... 81
    2.8.4.2 Meter .......................................................................................... 82

Chapter Three • Experimental Methodology ........................................ 84
  3.1 Overview and Hypotheses .................................................................... 84
  3.2 Methodology: Subjects ........................................................................ 86
  3.3 Methodology: Stimuli .......................................................................... 88
  3.4 Methodology: Equipment .................................................................... 96
  3.5 Methodology: Analysis, Phonetic Correlates of Phonological Tone ...... 98
    3.5.1 Overview ....................................................................................... 98
    3.5.2 Segmentation Procedure ............................................................... 99
    3.5.3 Duration ....................................................................................... 104
3.5.4 Fundamental Frequency ................................................................. 105
3.5.5 Spectral Tilt .................................................................................. 107
3.5.6 Open Quotient: Analysis of the EG waveform.............................. 111
  3.5.6.1 Calculating Open Quotient ......................................................... 111
  3.5.6.2 Common Errors in Automated DEGG Peak Detection ............ 114
  3.5.6.3 Open Quotient Statistics and Predictions .................................. 117
3.5.7 Airflow Measures ......................................................................... 118

Chapter Four • Results of Production Experiments 123

4.1 Introduction ...................................................................................... 123
4.2 Duration ............................................................................................ 127
  4.2.1 Overview of DURATION Results ................................................ 127
  4.2.2 DURATION of Isolated tokens – Carrier Phrase 1 ..................... 130
  4.2.3 DURATION of Phrase-Final tokens – Carrier Phrases 5, 6 .......... 132
  4.2.4 DURATION of Phrase-Medial tokens – Carrier Phrases 2, 3, 4 .... 134
  4.2.5 DURATION of Phrase-Medial tokens before Minor Syllables –
            Carrier Phrases 7, 8 ................................................................. 137
  4.2.5 Inter-Speaker Variation ............................................................... 141
  4.2.6 Summary of DURATION Data ..................................................... 143
4.3 Fundamental Frequency ..................................................................... 145
  4.3.1 F0 on Isolated Tokens – Carrier Phrase 1 .................................. 145
  4.3.2 F0 on Phrase-Final Tokens – Carrier Phrase 5, 6 ................. 148
  4.3.3 F0 on Phrase-Medial Tokens – Carrier Phrase 2, 3, 4 .......... 150
  4.3.4 F0 on Phrase-Medial Tokens before Minor Syllables –
            Carrier Phrase 7, 8 ................................................................. 156
  4.3.5 Statistical Analysis of F0 ............................................................ 158
  4.3.6 Summary of F0 Data ................................................................. 176
4.4 Spectral Profile: Comparative Amplitude of Harmonics ................... 180
  4.4.1 Overview of Statistical Analysis of Spectral Tilt ......................... 180
  4.4.2 H1-H2 ....................................................................................... 185
  4.4.3 H1-A1 ....................................................................................... 190
5.6.1 Overview of Statistical Analysis of F0 ........................................ 275
5.6.2 Anticipatory Co-articulation of F0:
   The Effects of a Following High Tone ........................................... 276
5.6.3 Perseverative Co-articulation of F0:
   The Effects of a Preceding High Tone ......................................... 278
5.6.4 Other Effects of a Preceding High Tone: Downdrift .................. 281
5.6.5 Differences Between the F0 Contours of Creaky and
   Checked Tones ................................................................. 285
5.6.6 Summary of Statistical Analysis of
   of the Effects of Context of F0 .............................................. 286
5.7 Effects of Syllable Juncture on F0: Statistical Evidence ................ 287
   5.7.1 Overview ....................................................................... 287
   5.7.2 Effects of Phrase Boundaries on Creaky, Checked Contours .... 289
   5.7.3 Association of Creaky Voicing and Glottalization with Low F0... 293
   5.7.4 Association of Creaky Voicing and a High F0 peak .............. 299
   5.7.5 Discussion of Findings ..................................................... 303

Chapter Six • Perception of Tone ................................................. 305
6.1 Background ........................................................................... 305
6.2 Methodology ......................................................................... 307
   6.2.1 Stimuli ............................................................................ 307
   6.2.2 Presentation ...................................................................... 311
6.3 Results ................................................................................... 312
   6.3.1 Overview .......................................................................... 312
   6.3.2 Natural and Control Stimuli .............................................. 315
   6.3.3 Re-synthesized Stimuli: Even Contours ............................ 317
   6.3.4 Re-synthesized Stimuli: Falling Contours ....................... 321
   6.3.5 Re-synthesized Stimuli: Rising Contours ....................... 321
6.4 Analysis and Discussion ......................................................... 322
Appendix I • Production Data by Carrier Phrase  384
   AI.1 H1-H2 Results.................................................................................................384
   AI.2 H1-A1 Results.................................................................................................392
   AI.3 H1-A3 Results.................................................................................................400
   AI.4 Open Quotient Results: ΔOQ........................................................................408

Appendix II • Consonant Durations Abridged in Chapter Five  411
   AII.1 Consonant Durations with an Initial High Tone.................................411
   AII.2 Consonant Durations with an Initial Creaky Tone .........................413

References  414
List of Figures

1.1 Multi-channel display of canonical production of High tone ..................9
1.2 Multi-channel display of canonical production of Low tone .................9
1.3 Multi-channel display of canonical production of Creaky tone ..............12
1.4 Multi-channel display of canonical production of Checked tone ..........12
1.5 Waveform of the two syllable utterance [tam.bá] ................................17
1.6 Plots of Open-syllable and Closed-syllable Vowel Sets,  
   from Watkins (2001) ........................................................................18

2.1 Mean Duration by Tone in Citation form, three prior studies .............45
2.2 Mean Duration by Tone in Phrase-medial forms, three prior studies ....45
2.3 Comparison of F0 traces for four speakers from three prior studies .....50
2.4 Four-way Register Hypothesis ..........................................................55
2.5 Spectra from representative (a) breathy, (b) modal, (c) creaky vowels ..61

3.1 The EG2-PCX Electroglottograph (Glottal Enterprises, Inc.) ............97
3.2 Photograph of the aerodynamic measuring system ...............................97
3.3 Acoustic pressure and EGG waveforms and spectrogram of [ti] ........101
3.4 Acoustic pressure and EGG waveforms and spectrogram of [t̃i] ........102
3.5 Acoustic waveform and spectrogram of Creaky [t̃e] ..........................103
3.6 Gradual vs. Abrupt F0 rises ..............................................................106
3.7 Longer open phase of modal voicing .................................................113
3.8 Shorter open phase of creaky voicing ..............................................113
3.9 Open Quotient trace of Low tone [ta] .............................................114
3.10 A single cycle of the EGG waveform .............................................115
3.11 The EGG and DEGG signals for a Phrase 1 (Isolated) utterance  
   of Creaky [a] ..................................................................................116
3.12 The OQ trace for a Phrase 1 (Isolated) utterance of Creaky [a] ........116
3.13 Unfiltered airflow waveform illustrating post-vocalic 75ms span…….120

4.1 DURATION (z-score) for tokens in Carrier Phrases (1) – (4)………………..128
4.2 DURATION (z-score) for tokens in Carrier Phrases (5) – (8)………………..128
4.3 Mean DURATION by Tone in Carrier Phrase 1 for speakers A - F …….130
4.4 Mean DURATION by Tone in Carrier Phrase 1 for speakers G - I……….130
4.5 Mean DURATION by Tone in Carrier Phrase 6 for speakers A - F……….133
4.6 Mean DURATION by Tone in Carrier Phrase 6 for speakers G - I……….133
4.7 Mean DURATION by Tone in Carrier Phrases 2 - 4 for speakers A - F 135
4.8 Mean DURATION by Tone in Carrier Phrases 2 - 4 for speakers G - I..135
4.9 Mean DURATION by Tone in Carrier Phrases 7, 8 for speakers A - F ..138
4.10 Mean DURATION by Tone in Carrier Phrases 7, 8 for speakers G - I…138
4.11 Mean F0 (z-score) at deciles in Carrier Phrase 1…………………………..146
4.12 Mean F0 (z-score) at deciles in Carrier Phrase 5…………………………..148
4.13 Mean F0 (z-score) at deciles in Carrier Phrase 6…………………………..148
4.14 Mean F0 (z-score) at deciles in Carrier Phrase 2…………………………..151
4.15 Plot of Peak F0 (z-score) against ΔF0 in Carrier Phrase 2………………..152
4.16 Mean F0 (z-score) at deciles in Carrier Phrase 3…………………………..153
4.17 Mean F0 (z-score) at deciles in Carrier Phrase 4…………………………..153
4.18 Mean F0 (z-score) at deciles in Carrier Phrase 7…………………………..157
4.19 Mean F0 (z-score) at deciles in Carrier Phrase 8…………………………..157
4.20 Mean F0_{MAX}, ΔF0, and slope by Tone in Carrier Phrase 1……………….160
4.21 Mean F0_{MAX}, ΔF0, and slope by Tone in Carrier Phrase 2……………….161
4.22 Mean F0_{MAX}, ΔF0, and slope by Tone in Carrier Phrase 3……………….162
4.23 Mean F0_{MAX}, ΔF0, and slope by Tone in Carrier Phrase 4……………….163
4.24 Mean F0_{MAX}, ΔF0, and slope by Tone in Carrier Phrase 5……………….164
4.25 Mean F0_{MAX}, ΔF0, and slope by Tone in Carrier Phrase 6……………….165
4.26 Mean F0_{MAX}, ΔF0, and slope by Tone in Carrier Phrase 7……………….166
4.27 Mean F0_{MAX}, ΔF0, and slope by Tone in Carrier Phrase 8……………….167
4.28 Continuum of phonation mode as defined by glottal state …………180
4.29 H1-H2 at late deciles, Carrier Phrase 5…………………………………….187
4.30 H1-H2 at Early and Late portions of Vowel, Carrier Phrase 5 ............ 188
4.31 ΔH1-H2, Carrier Phrase 5 ........................................................................... 188
4.32 H1-H2 at Early and Late portions of Vowel, Carrier Phrase 1 .......... 189
4.33 H1-A1 at Early and Late portions of Vowel, Carrier Phrase 1 .......... 192
4.34 ΔH1-A1, Carrier Phrase 1 ........................................................................... 192
4.35 H1-A1 at Early and Late portions of Vowel, Carrier Phrase 2 .......... 193
4.36 ΔH1-A1, Carrier Phrase 2 ........................................................................... 194
4.37 H1-A1 at Early and Late portions of Vowel, Carrier Phrase 5 .......... 195
4.38 H1-A1 at late deciles, Carrier Phrase 5 .................................................. 195
4.39 ΔH1-A1, Carrier Phrase 5 ........................................................................... 195
4.40 H1-A1 at Early and Late portions of Vowel, Carrier Phrase 6 .......... 196
4.41 H1-A1 at late deciles, Carrier Phrase 6 .................................................. 196
4.42 ΔH1-A1, Carrier Phrase 6 ........................................................................... 196
4.43 H1-A3 at Early and Late portions of Vowel, Carrier Phrase 5 .......... 199
4.44 H1-A3 at late deciles, Carrier Phrase 5 .................................................. 199
4.45 ΔH1-A3, Carrier Phrase 5 ........................................................................... 199
4.46 H1-A1 (z-score) at deciles in Carrier Phrase 1, All Subjects .......... 201
4.47 H1-A1 (z-score) at deciles in Carrier Phrase 1, Individual Subjects .... 202
4.48 Mean OQ at deciles in Carrier Phrases 1 – 8 ........................................... 214
4.49 Overall, Early, and Late mean OQ in Carrier Phrase 1 ................. 216
4.50 ΔOQ in Carrier Phrase 1 ........................................................................... 217
4.51 Overall, Early, and Late mean OQ in Carrier Phrase 2 ................. 218
4.52 Overall, Early, and Late mean OQ in Carrier Phrase 3 ................. 218
4.53 ΔOQ in Carrier Phrase 3 ........................................................................... 219
4.54 Overall, Early, and Late mean OQ in Carrier Phrase 5 ................. 220
4.55 ΔOQ in Carrier Phrase 5 ........................................................................... 220
4.56 Overall, Early, and Late mean OQ in Carrier Phrase 6 ................. 221
4.57 ΔOQ in Carrier Phrase 6 ........................................................................... 221
4.58 Overall, Early, and Late mean OQ in Carrier Phrase 7 ................. 223
4.59 Overall, Early, and Late mean OQ in Carrier Phrase 8 ................. 223
4.60 Mean Oral Airflow (z-scale) temporally, overall, and at peak in Carrier Phrase 1 ................................................................. 228
4.61 Mean rate of Oral Airflow (z-score) separating Low tone tokens with a late intonational rise and those without, Phrase 1 .......... 229
4.62 Mean Oral Airflow (z-scale) temporally, overall, and at peak in Carrier Phrase 2 .................................................................. 230
4.63 Mean Oral Airflow (z-scale) temporally, overall, and at peak in Carrier Phrase 3 ................................................................. 231
4.64 Mean Oral Airflow (z-scale) temporally, overall, and at peak in Carrier Phrase 6 ................................................................. 233
4.65 Mean rate of Oral Airflow (z-score) separating Low tone tokens with a late intonational rise and those without, Phrase 6 .......... 233

5.1 Waveform demonstrating a single consonant duration .......... 261
5.2 Waveform demonstrating a two consonant duration .............. 261
5.3 Mean Vowel Duration comparing CV and CVN Syllables ....... 267
5.4 Mean Vowel Duration and Coda Duration for each speaker .... 270
5.5 F0 and OQ traces for Isolated Creaky and Checked tokens ....... 294
5.6 Blow up scatterplot of data points composing means at 70% decile point ................................................................. 295
5.7 Scatterplot of F0M and OQLATE in Checked tokens embedded in Carrier Phrase 4 ......................................................... 301

6.1 Base token for 36 synthesized stimuli of Creaky tone /l̩ɛ / .......... 309
6.2 Screenshot of forced-choice orthographic options ................. 311
6.3 Responses to Even-contoured Stimuli ................................... 318
6.4 Responses to Falling-contoured Stimuli .................................. 319
6.5 Responses to Rising-contoured Stimuli ................................. 320
7.1 Phonetic interpretation of the Thai middle tone in three tone sequences, adapted from Morén and Zsiga (2006) ........................................... 348
7.2 F0 traces for Low tone tokens following a High tone ............. 350
7.3 High Tone F0 in context of an adjacent H ................................ 352

A.1 H1-H2 at Early and Late portions of the Vowel in Phrase (1) ....... 384
A.2 ΔH1-H2 in Phrase (1) ......................................................... 384
A.3 H1-H2 at Early and Late portions of the Vowel in Phrase (2) ...... 385
A.4 ΔH1-H2 in Phrase (2) ......................................................... 385
A.5 H1-H2 at Early and Late portions of the Vowel in Phrase (3) ..... 386
A.6 ΔH1-H2 in Phrase (3) ......................................................... 386
A.7 H1-H2 at Early and Late portions of the Vowel in Phrase (4) ..... 387
A.8 ΔH1-H2 in Phrase (4) ......................................................... 387
A.9 H1-H2 at Early and Late portions of the Vowel in Phrase (5) ..... 388
A.10 ΔH1-H2 in Phrase (5) ......................................................... 388
A.11 H1-H2 at Early and Late portions of the Vowel in Phrase (6) ..... 389
A.12 ΔH1-H2 in Phrase (6) ......................................................... 389
A.13 H1-H2 at Early and Late portions of the Vowel in Phrase (7) ..... 390
A.14 ΔH1-H2 in Phrase (7) ......................................................... 390
A.15 H1-H2 at Early and Late portions of the Vowel in Phrase (8) ..... 391
A.16 ΔH1-H2 in Phrase (8) ......................................................... 391
A.17 H1-A1 at Early and Late portions of the Vowel in Phrase (1) ..... 392
A.18 ΔH1-A1 in Phrase (1) ......................................................... 392
A.19 H1-A1 at Early and Late portions of the Vowel in Phrase (2) ..... 393
A.20 ΔH1-A1 in Phrase (2) ......................................................... 393
A.21 H1-A1 at Early and Late portions of the Vowel in Phrase (3) ..... 394
A.22 ΔH1-A1 in Phrase (3) ......................................................... 394
A.23 H1-A1 at Early and Late portions of the Vowel in Phrase (4) ..... 395
A.24 ΔH1-A1 in Phrase (4) ......................................................... 395
A.25 H1-A1 at Early and Late portions of the Vowel in Phrase (5) ..... 396
A.26 ΔH1-A1 in Phrase (5) ......................................................... 396
A.27  H1-A1 at Early and Late portions of the Vowel in Phrase (6)........397
A.28  ΔH1-A1 in Phrase (6).................................................................397
A.29  H1-A1 at Early and Late portions of the Vowel in Phrase (7)......398
A.30  ΔH1-A1 in Phrase (7).................................................................398
A.31  H1-A1 at Early and Late portions of the Vowel in Phrase (8)......399
A.32  ΔH1-A1 in Phrase (8).................................................................399
A.33  H1-A3 at Early and Late portions of the Vowel in Phrase (1).....400
A.34  ΔH1-A3 in Phrase (1).................................................................400
A.35  H1-A3 at Early and Late portions of the Vowel in Phrase (2)....401
A.36  ΔH1-A3 in Phrase (2).................................................................401
A.37  H1-A3 at Early and Late portions of the Vowel in Phrase (3)....402
A.38  ΔH1-A3 in Phrase (3).................................................................402
A.39  H1-A3 at Early and Late portions of the Vowel in Phrase (4)....403
A.40  ΔH1-A3 in Phrase (4).................................................................403
A.41  H1-A3 at Early and Late portions of the Vowel in Phrase (5)....404
A.42  ΔH1-A3 in Phrase (5).................................................................404
A.43  H1-A3 at Early and Late portions of the Vowel in Phrase (6)....405
A.44  ΔH1-A3 in Phrase (6).................................................................405
A.45  H1-A3 at Early and Late portions of the Vowel in Phrase (7)....406
A.46  ΔH1-A3 in Phrase (7).................................................................406
A.47  H1-A3 at Early and Late portions of the Vowel in Phrase (8)....407
A.48  ΔH1-A3 in Phrase (8).................................................................407
A.49  ΔOQ in Phrase (1).......................................................................408
A.50  ΔOQ in Phrase (2).......................................................................408
A.51  ΔOQ in Phrase (3).......................................................................408
A.52  ΔOQ in Phrase (4).......................................................................409
A.53  ΔOQ in Phrase (5).......................................................................409
A.54  ΔOQ in Phrase (6).......................................................................409
A.55  ΔOQ in Phrase (7).......................................................................410
A.56  ΔOQ in Phrase (8).......................................................................410
## List of Tables

(1.1) Consonant Inventory .................................................................5  
(1.2) Phonetic Qualities of the Four Burmese Tones .......................6  
(1.7) Vowel Quality Difference Between Open and Closed-syllable Vowel Sets .................................................................20  
(1.8) Fifty-one Syllable Rhymes of Burmese ...................................21  
(1.12) Tonogenetic Stages in Risiangku Tamang (from Mazaudon and Michaud 2008) .................................................................34  
(1.13) Tono/Registro-Genesis in Risiangku Tamang .........................36  
(2.1) Overview of Prior Instrumental Studies of Burmese Tone ..........41  
(2.2) Intensity by Tone (adapted from Watkins 1997: 3) ....................42  
(2.5) Survey of Experimental Findings and Researcher Observations for Tone Duration .................................................................44  
(2.6) Summary of Descriptions for Pitch across Tone ......................53  
(2.7) Summary of Findings Supporting or Refuting the Four-way Register Hypothesis .................................................................71  
(2.8) Chest vs. Head Registers (adapted from Henderson 1952, and Matisoff 1973) .....................................................................77  
(2.11) Tableau. Neutralization to [H] in Checked Syllables (Lee 2007) ...80  
(3.1) Biographical Data of Study Participants ....................................87  
(3.2) Set of Stimuli Tokens: Eight Syllable Rhymes per Carrier Phrase...94  
(4.30) Effect of Carrier Phrase on DURATION • Statistical Groupings ......129  
(4.31) DURATION in Isolated Utterances, means and Scheffé results ......131  
(4.33) DURATION in Phrase-Final Tokens, means and Scheffé results ......134  
(4.34) DURATION in Phrase-Medial Tokens, means and Scheffé results ..136  
(4.35) Comparison of High, Low > Creaky DURATION by Position ......137  
(4.36) DURATION in Phrase-Medial, Pre-Minor syllable Tokens,
means and Scheffé results.................................................................138
(4.37)  Random Effect of Speaker by ANOVA of DURATION...............141
(4.38)  Random Effect of Speaker on DURATION, excluding
Speakers A & E.............................................................................142
(4.39)  Variation by Speaker of Phrase-Medial Tone Contrast
in vowel DURATION.....................................................................142
(4.43)  Subsets by post-hoc Scheffé test in Phrase 1: Isolated Sylable....160, 168
(4.44)  Subsets by post-hoc Scheffé test in Phrase 2: [Low __ Low] ......161, 169
(4.45)  Subsets by post-hoc Scheffé test in Phrase 3: [High __ Low]......162, 170
(4.46)  Subsets by post-hoc Scheffé test in Phrase 4: [Low __ High].....163, 171
(4.47)  Subsets by post-hoc Scheffé test in Phrase 5: [High __ #].........164, 172
(4.48)  Subsets by post-hoc Scheffé test in Phrase 6: [Low __ #].........165, 172
(4.49)  Subsets by post-hoc Scheffé test in Phrase 7: [High __ minor σ]...166, 173
(4.50)  Subsets by post-hoc Scheffé test in Phrase 8: [Low __ minor σ]...167, 174
(4.58)  F0 Cues to Tone In Isolated and Phrase-Final Tokens..............178
(4.59)  F0 Cues to Tone in Phrase-Medial Tokens.............................178
(4.60)  Key to Presentation of Spectral Tilt Statistical Results..........183
(4.61)  Sample F-test Results...........................................................184
(4.62)  F-test Results, Mixed Model Analyses of Five Statistics of H1-H2 186
(4.63)  F-test Results, Mixed Model Analyses of Five Statistics of H1-A1 191
(4.64)  F-test Results, Mixed Model Analyses of Five Statistics of H1-A3 198
(4.65)  Differences in H1-H2 by Speaker Sex ...................................204
(4.66)  Differences in H1-A1 by Speaker Sex ...................................205
(4.67)  Differences in H1-A3 by Speaker Sex ...................................205
(4.68)  Summary of Mixed Model Analyses of Five Statistics of H1-H2 ...210
(4.69)  Summary of Mixed Model Analyses of Five Statistics of H1-A1 ...210
(4.70)  Summary of Mixed Model Analyses of Five Statistics of H1-A3 ...210
(4.71)  Effect of Tone on OQ in Isolated Syllable Tokens..................216
(4.72)  Effect of Tone on OQ in Phrase (2) Tokens............................217
(4.73)  Effect of Tone on OQ in Phrase (3) Tokens............................217
(4.74)  Effect of Tone on OQ in Phrase (4) Tokens............................217
(4.75) Effect of Tone on OQ in Phrase (5) Tokens ........................................ 220
(4.76) Effect of Tone on OQ in Phrase (6) Tokens ........................................ 221
(4.77) Effect of Tone on OQ in Phrase (7) Tokens ........................................ 222
(4.78) Effect of Tone on OQ in Phrase (8) Tokens ........................................ 223
(4.79) Summary of Tone Differences in OQ by Carrier Phrase Context .......... 225
(4.80) Effect of Tone on Measures of Oral Airflow in Isolated Tokens ........ 228
(4.81) Effect of Tone on Measures of Oral Airflow in Phrase (2) Tokens .... 230
(4.82) Effect of Tone on Measures of Oral Airflow in Phrase (3) Tokens .... 231
(4.83) Effect of Tone on Measures of Oral Airflow in Phrase (6) Tokens .... 233

(5.5) Underlying and Surface Forms of Vowels in Green (2005) .......... 245
(5.10) Orthographic Evidence for Allophony During WB Era ............... 252
(5.14) CODA DURATION following CVO vs. CV Syllables, Phrase (2) ...... 259
(5.15) CODA DURATION following CVO vs. CV Syllables, Phrase (3) ...... 260
(5.16) Low Tone CVN vs. CV Medial Consonant DURATION, Phrase (2) . 263
(5.17) Low Tone CVN vs. CV Medial Consonant DURATION, Phrase (3) . 263
(5.18) Summary of Consonant DURATION Results for Creaky and High Tones ................................................................. 265
(5.20) Isolated Syllables: Differences in VOWEL DURATION .............. 268
(5.21) Phrase-Final Syllables: Differences in VOWEL DURATION ......... 268
(5.22) Phrase-Medial Syllables: Differences in VOWEL DURATION ....... 268
(5.23) Summary of Speaker Similarity for VOWEL DURATION in Open CV and ĊV Syllables ....................................................... 271
(5.32) Effect of a Following Low vs. High Tone on F0-LATE .............. 277
(5.33) Effect of a Preceding Low vs. High Tone on F0-EARLY .......... 279
(5.34) Effect of a Preceding Low vs. High Tone on F0-LATE ............ 282
(5.35) Scheffé Comparison of Creaky vs. Checked F0 Contours Across Phrases ................................................................. 285
(5.37) Effect of Phrase Position on ΔF0 and F0\_LATE for Creaky Tones .... 291
(5.38) Effect of Phrase Position on ΔF0 and F0\_LATE for Checked Tones ... 292
(5.39) Correlation of F0 and OQ at Late Decile Points ....................... 296
(A.1) Consonant Durations with an Initial High Tone .................................. 411
(A.2) Consonant Durations with an Initial High Tone ............................... 412
(A.3) Consonant Durations with an Initial High Tone ............................... 413
(A.4) Consonant Durations with an Initial High Tone ............................... 413
Chapter One. Outline of the Problem and the Data

0. Introduction

The Burmese language presents a pattern that is interesting to linguists because pitch alone cannot describe the language’s four way tonal contrast. Instead, “tone” refers to a complex of phonetic qualities that includes pitch height, pitch contour, phonation mode, intensity, duration, and syllable structure. The largest confound for phonological analysis, however, is that the prior literature describing and quantifying the phonetics of the tones and syllables is quite limited, and where it is not sparse it is conflicting. Following a review of the existing gaps in the phonetic and phonological understanding of Burmese tone, this dissertation describes, conducts and provides the results of a series of experimental phonetic studies which not only fill these gaps in knowledge, but in turn provide the appropriate data for a more informed, integrated phonological analysis. This final analysis is “integrated” in the sense that articulatory, acoustic, auditory, and diachronic data are evaluated and reflected in the final model.

1.1 Goal of the Dissertation

Linguistic tone refers to the use of pitch to convey meaning at the lexical level. For Pike (1948), “a tone language may be defined as a language having lexically significant, contrastive but relative pitch on each syllable” (cited in Sprigg 1977: 1). Yip (2002) broadens the definition of “tone” to the “phonological category that distinguishes two words or utterances” on the basis of pitch. However, the Burmese language utilizes a
phonemic contrast on vowels that is widely recognized as tonal, yet is difficult to reconcile with this understanding of contrast provided solely by pitch. While pitch is prominent in classifying the four tones of Burmese, it seems necessary to also refer to, at least, mode of phonation and syllable structure. This state of affairs is not unusual in related and geographically proximate languages where pitch often intersects with vowel quality, tenseness, duration, and phonation type to form sets of lexically contrastive features which are, for convenience, still referred to as tonal inventories (Matisoff 1970, Egerod 1971, Mazaudon 1973, 1977, Maddieson and Ladefoged 1985, Sagart 1986, Diffloth 1989, Huang 1995, Yip 1995, 2002, Andruski and Ratliff 2000, Edmondson et al. 2001, Svantesson and House 2006, Fulop and Golston 2008, Tang 2008, DiCanio 2009).

Henderson (1967), writing primarily about Mon-Khmer languages, defines the characteristic as a regional one, writing that “phonological tone is in our area (SE Asia) very frequently a complex of other features besides pitch—such as intensity, duration, voice quality, final glottal constriction and so on.”

For Burmese, the situation is further complicated in that many impressionistic phonetic descriptions of the tones disagree with one another, and more recent acoustic studies have done little to clarify this confusion. Discrepancies in the literature concern the realization of pitch in the tones, whether certain tones are marked by breathy or creaky phonation, if vowel quality differs between the tones, and even over how many tones there are. In the capsules and phonetic samples of each tone provided in the following section, the reader can see the range of phonetic properties different researchers have called upon to distinguish the tones, as well as the lack of consensus.
between these many authors\(^1\). Bear in mind that this survey is intended to introduce the basic array of properties associated with each tone and that the sources listed here are primarily impressionistic accounts. Chapter Two provides a fuller profile of these properties, their varying descriptions in the literature, and the findings of acoustic and other instrumental studies that are sometimes at odds with these accounts.

### 1.2 Description of Burmese Phonology

#### 1.2.1 Overview of the Language

Burmese belongs to the Burmish branch of the Lolo-Burmese division of the Tibeto-Burman language family. It is spoken natively by approximately 32,000,000 people almost exclusively within modern Myanmar, where it is the national language (Ethnologue: Lewis 2009). Ethnic Burmese populations are centered around the Irrawaddy basin, in central and southern Myanmar. Within Myanmar, the Burmese language is also the second language, and frequently the language of literacy, for roughly another ten million people. These are generally Karen, Kachin, Mon, Shan, and other peoples. Pockets of mostly multilingual Burmese speakers exist in Bangladesh, Malaysia, Thailand, and the United States\(^2\).

Burmese has a substantial written history, dating back to the Myazedi inscriptions from AD 1113, a stone inscription found near Pagan with parallel texts written in Old

---


\(^2\) According to the U.S. 2000 Census, 16,720 respondents classified their race as “Burmese” (Census 2000 Brief, 2002). As detailed in Kiviat (2009), Burmese community and religious leaders within the U.S. are skeptical of this tally and suspect underreporting. One explanation for an artificially low count was that identification of Burmese ethnicity required a write-in response following the selection ‘Other Asian’.
Burmese, Pali, Pyu, and Mon. A Burmese civilization in present-day Myanmar was established by the 9th Century AD, within a few centuries replacing earlier Pyu (Tibeto-Burman) and Mon (Mon-Khmer) as the dominant regional culture, centered at Pagan. The script is Indic in origin but unique to Burmese, having been borrowed and adapted from the Indic-inspired Mon script. It is written left to right, and is mostly alphabetic, though some vowel and all tone marking uses diacritics written above, below, and adjacent to their carrier symbols.

Related languages in the Lolo-Burmese and Tibeto-Burman families are regularly tonal. These include Burmish Achang and Atsi, the Naxi language, and Loloish Akha, Lahu, Lisu, and Yi, all tonal to some degree and all spoken mainly in Myanmar and/or Southwestern China (see Lewis 1973, Matisoff 1973, Wannemacher 1996, Bradley 2003, Michaud 2006, Michaud and He 2007, Ethnologue 2009).

1.2.2 Description of Burmese Phonology: Inventory of Burmese Phonemes

The focus of this dissertation is on a suprasegmental contrast realized on the vowel in open syllables and the combined vowel and syllable coda in closed syllables. Despite a highly restricted set of possible coda consonants, the four suprasegmentals function to produce fifty-one possible rimes in the language. The distinction is traditionally referred to as four “Tones” (Armstrong and Pe Maung Tin 1925, Cornyn 1944, Okell 1969): Low, High, Creaky and Checked. The tones could be said to occur with eight different vowel types (roughly (i, e, e, u, o, a, a) though a more precise picture is offered in the ensuing sections. The possible rimes are paired with thirty-four possible initial consonants, shown in the phoneme chart in (1). Of these, only [w] and [j] may occur non-
initially. With the data elicited for this dissertation, the tokens of interest used an onset of only [t], the voiceless unaspirated stop. The following syllables in the frame sentences used the singleton onsets [k], [b], [m], or [l], so that only tone-bearing vowels between these consonants or phrase-finally were examined. Beyond these phonemes, little more will be said of the consonantal onset inventory in this dissertation. The coda inventory is discussed further in §1.2.4.

<table>
<thead>
<tr>
<th>(1) Consonant Inventory³</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Labial</strong></td>
</tr>
<tr>
<td><strong>Stops</strong></td>
</tr>
<tr>
<td><strong>Fricatives</strong></td>
</tr>
<tr>
<td><strong>Affricates</strong></td>
</tr>
<tr>
<td><strong>Nasals</strong></td>
</tr>
<tr>
<td><strong>Liquids</strong></td>
</tr>
<tr>
<td><strong>Glides</strong></td>
</tr>
</tbody>
</table>

1.2.3 Description of Burmese Phonology: Inventory of Burmese Tonemes

1.2.3.1 Overview

The present section briefly introduces the four tones, beginning with the table in (2), a summary of the range of phonetic qualities across which the four tones are regularly said to differ. Bradley (1982: 121) claimed that the each tone had a unique level for each

³ Information in the table in (1) borrows mainly from Watkins (2001) “Illustrations of the IPA” with two additional notes.
   a. The palatal affricate series /ʧʰ, ʧ/, ʤ/ is sometimes regarded as the palatal stop series /cʰ, c, j/ (as in Maran 1971: 27).
   b. Liquid /r/ only occurs in non-native loanwords such as [re sitiou] “radio” (Wheatley 1982).
quality such that, “The parameters involved in their realization (would) enable us to regard any of the six parameters as the contrastive one, if one had to choose only one.”

<table>
<thead>
<tr>
<th>Tone</th>
<th>Duration</th>
<th>Intensity</th>
<th>Pitch</th>
<th>Phonation</th>
<th>Syll. Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>long, moderate</td>
<td>low</td>
<td>low level</td>
<td>breathy or plain</td>
<td>CV or CVN</td>
</tr>
<tr>
<td>High</td>
<td>long</td>
<td>moderate</td>
<td>high rise or fall</td>
<td>plain, modal</td>
<td>CV or CVN</td>
</tr>
<tr>
<td>Creaky</td>
<td>short</td>
<td>high</td>
<td>high sharp fall</td>
<td>creaky, weak glottal closure</td>
<td>CV or CVN</td>
</tr>
<tr>
<td>Checked</td>
<td>extremely short</td>
<td>highest</td>
<td>initial high fall</td>
<td>abrupt, complete glottal closure</td>
<td>CVO</td>
</tr>
</tbody>
</table>

1.2.3.2 The Low Tone [a]

The Low tone is invariably described as low-pitched, but with a contour that is either level, bears a slight fall (Richter 1967), or has a final rise (Cornyn 1944, on occasion according to Okell 1969). Firth (1936), Stewart (1955), Okell (1969) and others note that the tone is not low-pitched so much as low in relation to the other tones. Along these lines, Bradley (1982) labels it the “Even” tone.

Low tones are also produced without the stress ascribed to the other tones. Low tone syllables are moderately long and produced with regular, modal glottal vibration. In the orthography, on most vowels it is denoted by the absence of a diacritic or trailing character. Low-toned syllables may be either open (CV) or closed with a sonorant nasal segment (CVN).
1.2.3.3 The High Tone [â]

Many of the contradictory phonetic claims regarding Burmese tones concern the production of the High tone. In particular, whether it is breathy-voiced or not and whether it is consistently a falling tone or simply high-pitched. High tone vowels are long, being either longer or equally long to Low vowels. They are also produced with greater intensity than Low and reduced vowels, but comparison with Creaky and Checked tone intensity is not as straightforward. Sources are consistent regarding both duration and intensity – uniformly calling the High tone the longest or “very long” (Bradley 1982: 120, Watkins 1997: 323) and also “stressed” (Firth 1936), “heavily stressed” (Stewart 1955), “fortis” (Becker 1964) or produced with emphasis. It is sometimes referred to as the “Heavy” tone due to this long, stressed syllable (as in Bradley 1982). Sources are less clear regarding pitch and phonation properties.

Evidently high-pitched, the High tone is usually characterized as bearing a high onset and then falling over the duration of the vowel (Armstrong & Pe Maung Tin 1925, Firth 1933, 1936, Cornyn 1944, Stewart 1936, 1955, Richter 1967, Yip 1995). Acoustic phonetic studies (Watkins 2005a,b, and the author’s own, Gruber and Feizollahi 2006) have often found the tone to bear a gradually rising pitch. Both acoustic studies elicited data in a context preceding a Low tone word, indicating a rising contour may be specific to this context. Also regarding pitch, it is often noted that the “high” pitch of the High tone is not as high as that seen on the Creaky or Checked tones. Bernot (1963) relates a “pitch between low and creaky tones”, Bradley uses a qualifier, “fairly high”, rather than just “high” pitch. Bradley (1982: 125) also reports one speaker from Mandalay for whom the High tone was produced with a lower pitch than even the Low tone, both vowel
initially and finally. Like the Low tone, High tone syllables may be either open (a CV syllable) or closed with a sonorant nasal segment (CVN).

The voice quality attributed to High tone syllables represents the most substantial contradiction between descriptive sources, and between descriptive and instrumental sources. Most authors indicate that the tone is breathy or frequently produced with breathy voicing (Thurgood 1978) while others make no mention of phonation properties. Okell (1969: 11) writes that the High tone has a “normal voice-quality, sometimes sounding almost breathy in contrast to creaky and stop tones”.

The data recorded for this dissertation generally confirm these descriptions, but find the breathy voice quality only in contrast to Creaky and Checked tones. Further, this distinction is limited to contexts where the tone-bearing syllable is not followed by another syllable. Multi-channel displays of the canonical pronunciation of each tone in isolation are given and described in Figures 1 – 4. In Figures 1 and 2, the Low and High syllables both display clear signs of vowel-final breathiness, particularly in the unfiltered and low-pass filtered airflow traces, where the rates of oral airflow ramp upwards at the vowel offset.
1.2.3.4 The Creaky Tone [à]

The Creaky tone bears a pitch contour which starts high and falls, though the steepness of this fall is not always agreed upon. Bradley (1982) states it is “slightly falling” in comparison to the High or possibly Checked tones, while Okell (1969) describes a Creaky tone pitch that “falls more sharply than Heavy tone”. Okell’s description is more in line with the findings of acoustic studies (Thein Tun 1982, Javkin and Maddieson 1983, Watkins 2005a). The Creaky tone also bears a high peak intensity early in the vowel and a brief duration (or, at least shorter than Low or High syllables). Creaky tone syllables may also be either open (CV) or closed with a sonorant nasal segment (CVN).
The Creaky tone is, in accordance with its name, consistently identified with creaky-voicing or a slow glottal closure. For some authors, this laryngealization is a “weak glottal stop” (Richter 1967) or “gradually applied glottal stop” (Okell 1969: 11) in contrast to the sharp glottal stop of the Checked tone. The slow glottal closure is often cited as the cause of the shorter Creaky syllable duration. Each of these qualities can be seen in the waveforms and displays in Figure 3, as compared to those for the Low and High tones. The falling pitch of the Creaky tone is not so apparent in the F0 trace in Figure 3, but investigation of the wider spaced periodic cycles confirms that the latter portion of the vowel sees a drastically slower rate of vibration which is unavailable to detection by auto-correlation (Praat 5.1.03, Boersma and Weenink 2009) due to the irregularity of the heavily laryngealized periodicity.

Phonetic descriptions of the Creaky tone are rather consistent concerning all major qualities: pitch contour, phonation, duration, and intensity. Acoustic studies, as will be shown in Chapter Two, have been less successful identifying these attributes.

1.2.3.5 The Checked Tone [ʔà]

Checked tone syllables are the only syllables which may have an obstruent coda, precisely, only /ʔ/ though different voiceless obstruent surface forms are possible according to context. This has lead some (Bernot 1963, Bradley 1982, Yip 1995) to not consider Checked as a tone at all, but solely as a syllable type. Historically, the Checked tone and words using the Checked tone were syllables closed with a final stop. These stops /p, t, c, k/ have since debuccalized, but are frequently retained in the writing system as a cue to the Checked tone status of the vowel.
Checked syllables bear the shortest vowel of all the tones, one that is cut off by the quick glottal closure of the final stop. A common alternative label for the tone used in Burmese grammars is the “Killed tone”, in reference to this abrupt closure and to the orthographic character used with final stops, {•}, which has the name /ʔəʔ/ “the killer” (Roop 1972: 62, Soe 1999: 9). The vowel bears a very high pitch and high peak intensity. The high pitch has been classified as level (Becker 1964, Richter 1967, Mesher 2006), slightly falling (Bradley 1982), or sharply falling (Okell 1969, Watkins 2001). The voice quality of Checked syllable vowels is not often discussed in the literature, as the tone is typically contrasted with the other tones only in terms of the glottal stop coda, and not the preceding vowel. Firth (1936) describes Checked vowels as using a “bright voice” and Watkins (1997: 3) states that the Checked tones are “tense”.

As mentioned above, the obstruent coda of the Checked tone is only a glottal stop phrase-finally and is frequently realized as many of the allowable onset consonants in juncture (see §1.2.4). The phonetic properties of Checked vowels in these cases are not described in the literature, and are a novel finding of this dissertation (see §4.3-4.5).
Figure 3, 4. Multi-channel display of canonical production of Creaky, Checked tones, spoken in isolation. At top is the acoustic waveform, followed by the glottal waveform and spectrogram from the same utterance of a tone-bearing vowel. An unfiltered and low-pass filtered trace of the oral airflow from a separate recording by the same speaker are presented at the bottom. Accordingly, the airflow signals are not time-aligned or scaled with the top three graphs. That is, despite superficial appearances, the airflow and acoustic recordings yielded utterances with similar durations.

1.2.3.6 Minor Syllables [ə]

Finally, vowels may occur in non-tone-bearing Minor syllables, in contrast to Major syllables with Low, High, Creaky, or Checked tones. Minor syllables are considered by some to be a fifth tonal possibility, but are generally regarded as having no tonal specification. Following Green’s (2005) analysis, minor syllables are viewed in this dissertation as unstressed or reduced syllables in weak prosodic positions. The nature of this reduction is multifaceted and the Major vs. Minor distinction is characterized by the list of distinctions in (3), adapted from Sprigg (1957), Okell (1969), and Green (2005):
None of the phonetic qualities found with the other tone-bearing syllables are reliably found with minor syllables. Bradley (1982) reports that minor syllables have no consistent pitch qualities and are generally produced with modal phonation. Regarding duration, minor syllables are considerably shorter, in part because they lack a coda consonant (noted in (3d)), but primarily because minor vowels are extremely reduced (3a, 3c). Acoustic phonetic studies have reported the duration of minor syllable vowels to be dramatically shorter than major vowels, even Checked, at around 40-50 ms, or roughly half the length of Checked vowels. Minor syllables were recorded for the present study, but their phonetic qualities do not represent a major component of this dissertation.

### 1.2.4 Description of Burmese Phonology: Syllable Codas

Three syllabic templates can describe all Burmese syllables: CV, CVO, and CVN. The letter N may represent either a nasal consonant or simply a nasal quality on the vowel nucleus. The possible obstruent codas found in CVO syllables are highly restricted — only a glottal stop or an obstruent identical to a following onset consonant are allowed.
The obstruents are best treated as a single underlying segment, either surfaceing as [ʔ] in word-final position or fully assimilating to the following onset in close juncture. The demonstration of this assimilation in (4) is adapted from Watkins (2001) (though for more complete descriptions of the alternation, refer to Cornyn (1944), Sprigg (1957), or Okell (1969)):

(4) Assimilation of glottal stop coda to a following onset consonant

a. [lɛʔ] ‘hand’

b. [sʰaʊ] ‘carry along’

c. [lɛːsʰaʊ] ‘gift’ from [lɛʔ] + [sʰaʊ]

The forms in (4) show the two realizations of the obstruent coda on Checked syllables: [ʔ] in isolation (4a) or as part of a geminate over a syllable/morpheme boundary (4c).

Nasal codas behave similarly, but do not fully assimilate to the following onset. Instead, a nasal coda adopts just the onset’s place of articulation, forming a homorganic NC trans-syllable cluster, as in [eɪm:e] in example (5a).

(5) Assimilation of nasal stop coda to the Place of a following onset consonant.

Examples from Okell (1969: 6)

a. /eɪN/ + /me/ → [eɪm:e]
   ‘house’ locative ‘at home’

b. /θaN/ + /jeɪʔ/ → [θaŋ.jeɪʔ]
   ‘iron’ ‘hook’ ‘iron hook’

---

4 See Chapter 5, Section 3 for further discussion of the segmental alternation in Checked syllables and other properties of the glottal stop concerning its status as a coda consonant.
In both examples above, the Place of the surface nasal form is determined entirely by context, [m] before another labial onset and [n] before palatal [j]. Underlyingly, the nasal codas are best treated as all being a single nasal consonant without specified Place features, a single “placeless nasal consonant”, N, (Green 2005: 9) just as all obstruent codas are underlyingly a single placeless glottal consonant. These two possible coda segments, /Ɂ/ and /N/, represent Modern Burmese reflexes of two series of historical obstruent and nasal codas in Proto Lolo-Burmese (Thurgood 1981) and earlier forms of Burmese as evidenced by archaic inscriptions and the modern orthography. The historical coda contrasts were depleted to a single glottal and nasal coda sometime between Written Burmese and the modern language (Maran 1971, Luce 1985, Bradley 1982, Bradley 2010). Before or during this time, the suprasegmental features in CVO syllables were neutralized to a single tone, while the sonorant codas of CVN syllables did not neutralize tonal oppositions. In this way, a Checked tone syllable and an obstruent-closed syllable are one and the same, and the High, Low, and Creaky tones are "restricted" to syllables without obstruent codas (either open CV or sonorant CVN).

As mentioned earlier, nasal codas do not always have a segmental instantiation, but may realize N solely as vowel nasality. The pair in (6) lists the alternative forms, as described in the grammars of Okell (1969) and Cornyn and Roop (1968):
(6) Contextually determined phonetic forms of the nasal syllable /paʊN/

(a) [paʊn] In juncture in casual speech: Coda is a nasal stop segment, homorganic in place to the following onset.

(b) [paʊ]⁵ In careful speech, citation form, or phrase-final position: There is no coda segment. Nasalization is realized on the vowel.

Watkins (2005a) uses the term “orthographically closed syllables” for the CVN syllables, indicating that, although always written, no coda segment may be produced in speech. In (6), one syllable (a) is phonetically closed and the other (b) is open, but both may bear the High, Low, or Creaky tones and the same set of vowels. The nasal coda segment surfaces when another syllable immediately follows in casual, connected speech, as in (6a) or example (5) above. In these instances, the coda is clearly more than “orthographic”, but when another syllable does not immediately follow, the syllable is open and the orthographic nasal coda is realized solely as vowel nasalization (as in 6b).

Acoustic and aerodynamic data collected for this dissertation were compatible with this observation. Figure 5 is the waveform of a Low tone syllable with a nasal coda followed by a [b] initial syllable in the word [bá], a politeness marking particle. The Low tone vowel [a] is comparatively short, 133ms. By comparison, when there was no coda segment, the same speaker produced Low tone vowels with an average duration of 215.6 ms over 12 samples.

⁵ Bennett and Lehman (1994) describe a nasal off-glide in the production of open CVN syllables – coronal [v̚] after monophthong vowels, velar [v̚] after diphthong vowels (see §1.2.5 for vowel differences).
Figure 5. Waveform of the two syllable utterance [tam.bá]. The nasality of the first syllable coda surfaces as a distinct [m] segment. Vowel duration in the phonetically-closed syllable is shorter than in comparable open syllables.

In this example, the nasal coda is clearly segmental as a labial nasal stop and not simply the presence of nasality on the vowel, as in the hypothetical form *[tābā]*. Secondly, the consonantal nasal coda affects the length of the preceding vowel, making it shorter. The single example in Figure 5 does not prove this effect, but it is illustrative of a pattern that is statistically confirmed in Chapter Five. The data in Chapter Five address another outstanding issue in the literature: the acoustic consequences of syllables with non-segmental nasal codas, such as [paõ] in (6b), in particular for duration. These data compare the vowel duration of phonetically-open, nasal syllables (CVN as [CṼ]) and open oral syllables (CV) and find a mostly negligible difference. Putting the codas aside, there is also a distinction between the vowels themselves in each of the syllable shapes. The next section turns to this difference in vowel quality.
1.2.5 Description of Burmese Phonology: Vowel Quality

Vowels in Burmese fall into two distinct groups, labeled here as the “Open-syllable Set” and the “Closed-syllable Set” of vowels. The diagrams in Figure 6, courtesy of Watkins (2001), illustrate the two sets.

![Open-Syllable Vowels](image1) ![Closed-Syllable Vowels](image2)

**Figure 6. Plots of Open-syllable (left) and Closed-syllable (right) vowel qualities, from Watkins (2001).** Schwa is more appropriately considered a reduced-syllable vowel, but is presented in the left-hand diagram since all minor syllables are open.

The division into the Open and Closed Vowel sets of Figure 6 follows the standard phonetic description of sixteen vowel phones in Burmese (including [ə]). Though the syllables closed by nasality and closed with a glottal stop are typically transcribed with the same set of vowels, it also makes sense to organize the vowels into three sets with an additional 'nasalized' set of vowels (see Thurgood, 1978). In this case, the division would be:

<table>
<thead>
<tr>
<th>Open syllables</th>
<th>Nasal syllables</th>
<th>Checked syllables</th>
</tr>
</thead>
<tbody>
<tr>
<td>[i]</td>
<td>[ɨ]</td>
<td>[iʔ]</td>
</tr>
<tr>
<td>[e]</td>
<td>[eɪ]</td>
<td>[eɪʔ]</td>
</tr>
<tr>
<td>[ε]</td>
<td></td>
<td>[ɛ]</td>
</tr>
<tr>
<td>[a]</td>
<td>[ã]</td>
<td>[aʔ]</td>
</tr>
<tr>
<td>[ɔ]</td>
<td>[aʊ]</td>
<td>[aʊʔ]</td>
</tr>
<tr>
<td>[u]</td>
<td>[ʊ]</td>
<td>[ʊʔ]</td>
</tr>
<tr>
<td>[ə]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Burmese orthography represents the closed Nasal and Checked vowels the same (i.e. Nasal [aʊ] and Checked [aʊʔ] are both written with the vowel {e•}, where {•} represents possible consonants occurring with the vowel). However, Thurgood’s analysis of 23 distinct vowels picks...
It is immediately noticeable that Open-syllable vowels are “pure” vowels, conveying a single sustained vowel quality, while the Closed-syllable set contains some diphthong vowels. Another generalization is that a majority of the Open-syllable vowels are transcribed with symbols for “tense” vowels ([i], [u], [e], [o]), while the Closed-syllable vowels either have a lax or centralized vowel quality ([ɪ], [ɛ], [o]) or possibly just movement towards a lax vowel quality ([eɪ], [aɪ], [au], [ou]). Additionally, two vowel characters are included in both sets, [a] and [ɛ], though their placement by Watkins in the diagrams of Figure 6 warns against reading too much out of the traditional transcriptions. Spectrographic data in Watkins (2005a) confirms these vowel qualities, showing that [a] and [ɛ] of the Closed-syllable vowel set attest a distinct vowel quality, not unlike the more transparent (through transcription) differences in vowels from Checked and nasalized syllables.

Watkins (2005a) reports that [a] and [ɛ] have systematically different formant structures in phonologically open and closed syllables. In his study, the formant structures of Checked vowels indicate that a Checked [ɛ] is slightly lower and more back (on average 300 Hz lower F2) than a Creaky [ɛ], while a Checked [a] was higher (150-200 Hz lower F1) and more front (100-150 Hz higher F2) than a Creaky [a] (Watkins 2005a, figures 7,8). Mehnert and Richter (1976) transcribe the Closed-syllable low vowel as [ʌ], indicative of a similar vowel quality to that reported by Watkins.

up on the fact that it is not entirely accurate to depict CVO and CVN vowels as oral and nasal counterparts of the same set of vowel qualities. This is primarily because the vowel [ɛ] occurs in closed CVO syllables but has no nasalized CVN counterpart. It is also the case that they are no active phonological processes supporting the treatment of certain vowels as “counterparts” of one another – see §5.2 for further discussion.
Taking stock, it is clear that the two vowel sets are different and do not share even one or two vowels with superficially identical transcriptions. However, it is not clear how the sets are different, or if the difference is in any way consistent. If one treats the Open and Closed-syllable sets as allophonic counterparts (i.e. \( i \sim \text{i} \), \( e \sim e\text{i} \), \( u \sim \text{u} \), \( o \sim oo \), \( \text{o} \sim \text{au} \)), there should exist some common thread to the phonetic difference between each set of counterpart vowels. The differences are examined in (7), where an individual look at each pair does not offer a cohesive pattern, but rather two complementary patterns: diphthongization or a lax/centralized quality.

| (7) Vowel Quality Difference Between Open and Closed-syllable Vowel Sets |
|-----------------|-----------------|-----------------|
| Open-syllable   | Closed-syllable | Description     |
| vowels          | vowels          |                 |
| [i]             | [i]             | lower, more central |
| [e]             | [e\text{i}]     | diphthongized; movement to higher, more back |
| [ɛ]             | [ɛ]             | slighter lower and more back |
| [a]             | [a]             | higher and fronter, more central |
| [ɔ]             | [aʊ]            | diphthongized; movement to higher, more central |
| [ʊ]             | [ʊ]             | diphthongized; movement to higher, more central |
|                 |                 | lower and less back, more central |

Watkins (2001) describes the closed syllable vowels as centralized. Though the transcriptions do not always represent what one would call “centralization” it does seem to be the case that the front vowels are less front in closed syllables, back vowels are slightly more forward, high vowels are lower and low vowels higher. The diphthongization only coincides with this centralization in that the latter vowel state uses such a quality.

A complication posed by the differing vowel qualities is their distribution amongst the four tones and the three syllable types. The two sets of vowels are
predictably distributed between the CV, CVO, and CVN syllable types. Despite the proliferation of possible combinations brought about by four phonemic tones on each vowel, the restrictions on allowable codas limits Burmese to only fifty-one possible syllable schema, all with an obligatory onset. These are given below in a table adapted from Watkins (2005a) and Thein Tun (1982):

The table in (8) demonstrates how the distribution of tone to vowel quality is restricted. Significantly, only three tones (Low, High, and Creaky) appear in CV syllables with the “monophthongal” vowels, while “diphthongized” vowels occur with all four tones in “orthographically closed” syllables — with Checked tones in CVO syllables and the others in CVN syllables.

The distributions are incongruous however — the distributions of tone and vowel quality amongst the CV, CVN, and CVO syllable types do not align. As the diagram in (9)
illustrates, the source of the misalignment is the dichotomous behavior of the nasal-coda syllables, which may be phonetically open or closed according to phrase position.

(9) Incongruous Distribution of Tone and Vowel Quality

CVN syllables behave parallel to CV syllables regarding pitch, phonation, intensity, and duration (i.e. regarding tone) but share a set of vowels instead with the CVO syllables. Additionally, the phonetic veracity of the label “Closed-syllable Vowel” is refuted by some surface forms; in example (6) above, in either phonetically closed (6a) or open (6b) syllables the “Closed-syllable” diphthong vowel [aʊ] is used. Since CVN syllables can be phonetically open or closed, the “Closed-syllable Vowel Set” does not only surface in closed syllables. Rather, the vowels are restricted to phonologically closed syllables.

Phonetically open CVN syllables also contradict otherwise plausible phonetic motivations for the differing vowel qualities – it is difficult to maintain that the diphthongized vowel quality is a phonetic effect created by the presence of a coda since the coda does not even exist in citation forms (e.g. diphthongal [paʊ] in (6b)). Green (2005), in fact, proposes the reverse alternation: that the closed-syllable vowel set represents the underlying vowel quality of the Burmese vowel phonemes, and only in the context of open syllables is the Open-syllable Vowel set used. Either direction indicates an allophonic account that treats the Closed-syllable vowels as diphthongized/centralized.
counterparts to Open-syllable vowels. Yet there is no cohesive phonetic explanation for the allophony between the sets as a productive, synchronic process. A speculative take on this process is that the lingual articulation loses importance and moves toward a more central vowel as an incomplete shift toward schwa accompanying the quick transition to a glottal stop (or other nasal or obstruent stops). This logic is challenged by the range of vowel qualities reported in closed syllables and by the different possible coda realizations of phonologically-closed syllables. Regardless of whether the phonological coda is realized as a glottal or supra-glottal stop or solely as vowel nasalization, the shared vowel quality of nasal syllables and the non-nasal Checked syllables is difficult to reason in a productive synchronic account.

A more plausible explanation for this pattern likely lies in the history of the Burmese language and a diachronic origin of the divide between the vowel sets at a stage in the language where such a change was phonetically motivated. In Chapter Five, historical evidence is presented which demonstrates this prior stage of allophony, one that naturally explains the split-alternation between diphthongization and centralization in phonologically closed syllables.

1.3. Typological Interest of the Dissertation

1.3.1. Burmese Dialects and the Language Background of Study Participants

For this dissertation, the production and perception of Burmese was tested with native speakers hailing from the Rangoon Region of Myanmar. It is therefore more precise to state that all ensuing findings and arguments pertain not to “the Burmese language”, but to “Rangoon Burmese”. However, due to the prominence of the Rangoon dialect in
modern Burman culture and media and to the common identification of Rangoon speech as Standard Burmese (Maran 1971, Wheatley 1987) and also for the purpose of brevity, the shorter label “Burmese” will be used in this dissertation following the present section. The remainder of this section outlines the dialect situation in Myanmar and the position of Rangoon Burmese within that setting.

Regional dialects of Burmese are spoken throughout Myanmar, some varying considerably from Standard Burmese in their lexical items and less so in phonology. The two most widely-spoken and influential dialects are those used in the two largest population centers, Rangoon and Mandalay. Both may be considered representative of Modern Standard Burmese. Wheatley (1987) states that “the speech of Mandalay in Upper Burma…is indistinguishable from that of Rangoon”, while others describe their differences as solely lexical (Patrick McCormick, personal communication). As concerns phonology, and in particular dialectal variation in the tones, Mandalay and Rangoon Burmese are not markedly different (contra to a few of the claims in Maran 1971).

Nonetheless, until the 20th Century, the Mandalay dialect was the most representative of a national standard in accordance with the city’s role as sovereign seat of the Burmese kingdom at the end of the Third Anglo-Burmese War in 1885. From the British Colonial era through last century, the speech of Rangoon has emerged as the modern de facto standard as the result of increased urbanization of the population and the subsequent pervasiveness of Rangoon-based media and popular culture. Other major dialects, many of which are sometimes referred to as distinct languages, include Arakan.

---

7 Maran’s (1971) comparison of Burmese dialects labels the Mandalay variety as Central Burmese and Standard Burmese 1, while calling the Rangoon variety Standard Burmese 2.
(in Rakhine Region), Intha, and Tavoyan (Dawei) (see Okell 1995), the conservative Yaw dialect described in Okell (1989), and Mergui.

While not always stated explicitly\(^8\), the prior phonetic research on Burmese has mostly described the Rangoon variety of speech or a standard variety with features common to Rangoon and Mandalay speech. For this reason, the Rangoon dialect was studied in this dissertation to enable the most relevant comparison with prior observations and findings. Further, as the most widely spoken variety of Burmese, the Rangoon dialect presented the most feasible group of native speakers for recruitment within the U.S.A.

The final point above raises another qualification that is worth accounting for concerning the data. As every speaker was an expatriated Rangoon native, the study was arguably limited to a diaspora dialect. To counter this objection, a few considerations point toward the subjects’ data being representative of Rangoon speech. For one, the recruited subjects could scarcely be characterized as a cohesive language community – some speakers were long-term residents of the Washington, D.C.-Baltimore metropolitan area, others to the Seattle, Washington area, and a few were students (or their acquaintances) who had lived in the United States for less than two years (and in one case, less than three months). Secondly, the experimental findings in Chapter Four are reconcilable with the characteristics of Standard Burmese identified in prior studies using subjects from Rangoon. Finally, the data collected from each speaker were also

---

8 • Okell (1969) describes “Colloquial Burmese” in contrast to more formal, literary styles which are held in high regard. He states that “Burmese is remarkably free from regional dialect problems…” (ibid: xii). Colloquial Burmese for Okell can thus be understood to mean the “Standard Burmese” spoken by Burmans and lacking potential regional dialect features.
• Bradley (1982) periodically refers to recordings of two Mandalay speakers in his text, but offers no quantitative data. Otherwise, his descriptions of the tones are appropriate for Rangoon Burmese.
• Thein Tun (1982) recorded at least a dozen subjects, for whom no regional background is mentioned.
• Watkins (1997) collected data from two speakers: a Rangoon and an Arakan speaker.
• Watkins (2005a, b) both use a speaker from Rangoon.
reconcilable with each other – that is, the major findings of the study are reflected in each speaker’s data despite the dispersion of their backgrounds. This relative homogeneity suggests that the speakers could reasonably be assumed to speak a similar dialect to one another\(^9\), and that this dialect was presumably that of Rangoon as each speaker identified a native association to Rangoon.

1.3.2 Tone or Register?

Understanding the suprasegmental contrast in Burmese cuts to the question of how linguists want to define “Tone”. Regarding pitch contrasts, the Burmese tones fit the main criterion for Tone classification – differences in pitch can carry a distinction between lexical items. On the other hand, the distinction in Burmese nicely matches most understandings of register contrasts, particularly as they are used in Southeast Asia – each category represents a cluster of phonetic qualities indicating laryngeal and supralaryngeal differences in articulation. The previous section demonstrated how the four tones yield a multi-way distinction across numerous phonetic parameters. That pitch can also differ along with the rest of these qualities is hardly surprising, and certainly not unique within register systems. Bradley (1982) points out that the contrast in Burmese is suggestive of the influence of Austro-Asiatic languages (particularly the historical Mon population of modern Myanmar) rather than what one expects in the generally more tonal Tibeto-Burman languages.

Regardless of any quantitative results and subsequent conclusions set forth here, this dissertation alone cannot resolve the dilemma of whether Burmese is best classified

\(^9\) Or at the minimum, all subjects shared a dialect background whereby the production of tone qualities was similar.
as either a Tone language with register-type differences or as a Register language with robust pitch differences. The decision has been made here to use the label “Tone” instead of “Register”. This was primarily done for two reasons which have little descriptive or theoretical consequence: first, one should be selected the sake of consistency throughout the present text, and second, to adhere to the tradition within the literature concerning Burmese. However, less superficial arguments also arguably apply.

Arguments against the use of “Register” are that (1) systems with contrast between more than two registers are typologically rare (particularly four-way register contrasts), (2) the pitch differences offer the most robust three-way\(^\text{10}\) distinction across possible contexts, and (3) the pitch differences do not fall out from the phonation or vowel quality contrasts in any straightforward way. That is, in register languages the *breathy* register is typically the low-pitched variant while the Burmese High tone is reported to be both breathy and high-pitched (see references in §1.2.3.3 and §2).

In favor of “Tone”, it has been shown that prototypical Tone languages such as Mandarin and Vietnamese regularly utilize glottalization, coda segments, and other qualities to enhance or further differentiate tonally-distinct lexical items (Egerod 1971, Davidson 1991 for Mandarin Chinese, Michaud 2004 for Hanoi Vietnamese). If these quintessentially tonal languages use phonation, whether contrastively or not, then there are likely very few languages which strictly fit Pike or Yip’s definition of Tone. This conclusion advocates an expanded use of “Tone” to describe categorical contrast between sets of suprasegmental features, including pitch, realized at the syllable or word level.

\(^{10}\)“Three” ways because the Creaky and Checked tones have essentially identical pitch behavior (see Watkins (2005a) as well as §4.3 of this dissertation).
1.3.3 Typology of Tone in Southeast Asia

Regardless of the appropriation of the term Tone or Register, it is evident that phonological pitch contours and phonation types frequently interact. Looking just at a handful of Southeast Asian languages, one can see that this interaction takes many forms.

To begin with, it is possible that phonological phonation and pitch contrasts exist side-by-side in a language without intersecting. The White Hmong language (Hmong-Mien family) arguably presents such a case, as it contrasts seven different pronunciations of each vowel, five pitch variants (high ɔ́, mid ɔ, low ɔ̀, falling ɔ̂, rising ɔ̌) and two phonation types (creaky ɔ̰, breathy ɔ̤) (Fulop and Golston 2008). These specifications do not overlap such that the creaky and breathy phonation types function equally with pitch in producing phonemic contrast.

When the interaction is intersective, it may be described as either independent or dependent according to whether the presence of one feature is entailed or not by the presence of another (e.g. breathy tones are always low-pitched = dependent). Yip (2002) describes two TB languages with systems referred to as tone that also employ phonation types: Bai (citing Edmondson and Li, 1994), which divides a complex eight-way contrast between lexically specified pitch patterns and tenseness; and Jingpho where three tonal patterns are pronounced on either tense or lax vowels for a six-way contrast. The interaction between pitch and phonation in these languages is intersective and independent. The features intersect (as in a lax, high vowel), but are independent of one another in that knowledge of high or low pitch on a vowel does not entail knowledge of the tense/lax dimension on that vowel.
Coupe (2003) describes the pattern in Ao Mongsen\textsuperscript{11}, where the interaction is also \textit{intersective} and \textit{independent}, but with a curious wrinkle. In Ao Mongsen, a set of modal or creaky vowels may occur with any of the three tones (High, Mid, or Low) but is \textit{dependent} upon a low, mid vowel quality /a/. As the creaky /a/ bears contrastive meaning and is shown to be “independent of both suprasegmental and segmental influence” (Coupe 2003: 44), Coupe (\textit{ibid}: 45) concludes that “…there does not appear to be any justifiable alternative to the proposition that creaky voice be analyzed as the somewhat idiosyncratic realization of a separate vowel phoneme.”

An \textit{intersective}, but \textit{dependent} interaction of pitch and phonation is found in the tone system of Sgaw Karen, a TB language spoken in Myanmar. Brunelle and Finkeldey (2011) present data for the tone system of Sgaw Karen which demonstrate six tones bearing a combination of three phonation types and at least three pitch contours. The two breathy-voiced tones bear similar high falling pitch contours and the two glottalized tones bear similar low falling contours, meaning the pitch contour can be predicted by the presence of phonation in these cases. The dependence is only partial however, as pitch is not predictable by modal voicing. The two modal tones can be either high rising or low falling. The Sgaw Karen example demonstrates how an interaction is either \textit{independent} or not, but there are different degrees to which it may be \textit{dependent}. A fully \textit{dependent} interaction implies a fully redundant contrast.

Like Sgaw Karen, the Burmese tone system has an \textit{intersective} and \textit{partially dependent} interaction between pitch and phonation distinctions. Other dimensions of contrast are also unevenly distributed so that it seems necessary to define the tones by at

\textsuperscript{11} More precisely, Ao Mongsen is the Mongsen dialect of the Ao Naga language spoken in northeast India. Coupe refers to the language in his study as Waromung Mongsen, as his data was collected in the Waromung village among Mongsen speakers of Ao.
least (i) pitch, (ii) mode of phonation, (iii) duration, and (iv) syllable structure. Ideally, a four-way contrast is definable by two binary features (suggested for Burmese by Green, 2005), yet such an account of Burmese tones looks to be impossible. Even disregarding either the pitch or phonation qualities (the two main archetypal correlates of Tone or Register) does not allow one to describe the tonal system on the basis of pitch alone, since syllable codas still necessarily define the Checked tone category. In contrast to this, Bai, as mentioned above, for the most part allows a syllable bearing any tone to be either \([+\text{TENSE}]\) or \([-\text{TENSE}]\), either nasal or oral, and of these three qualities (pitch, tenseness, nasality) most possible combinations are available.

Comparison of these interactions is compelling because the synchronic typology of phonation in tone can offer insight into the historical origins of tone/register systems, a much discussed issue in the last half-century of work in Southeast Asian linguistics. The reverse of this statement is also true: an understanding of the historical origins of tone/register features can offer crucial insights into the synchronic patterning of such systems. In fact, this is a central argument of the dissertation, that even for a phonologist solely concerned with productive synchronic behavior, something of the nature of a synchronic contrast can be found by unraveling the series of diachronic changes which yielded that system. The observation is pertinent for Burmese and Tibeto-Burman languages where redundant or uneven distributions of distinctive features that define tone can confound an economical synchronic analysis. An example of such a distribution in the Tamang language is provided in the following section, but the discussion first turns to outlining a general understanding of tonogenesis.
1.3.4 The Tonogenetic Roots of Typological Tone Patterns

Coinciding with some of the early work on Tonogenesis, Maran (1973) presented a general model of the stages of tone development. The five stages listed in (10) demonstrate how, “the manner in which a tonal system takes over is gradual…” (Maran 1973: 107).

(10) Stages of Tonogenesis: Maran (1973)12 “On Becoming a Tone Language”

Stage 1: Atonal
Stage 2: “Attendant pitch characteristics” without functional load
Stage 3: Redundancy of pitch and conditioning environment – codas and pitch perturbation exist side-by-side
Stage 4: Tone is primary contrast. An “advanced depletion of finals” (i.e. coda distinctions) makes tone unpredictable, determined lexically.
Stage 5: Tonal

Although Maran (1973) was charting the development of “Tone”, abstracting away from some of the specifics allows one to easily apply the progression in (10) to other diachronic changes, such as the genesis of a register distinction (see footnote 12). Essentially, any suprasegmental distinction conditioned by another, incrementally depleted, feature can fit the model. Specifically, Maran focused on the transfer between “finals” (coda segments) and tone. With a discussion of Burmese and Tibeto-Burman tone ultimately in mind, his focus represents more than a preference, but a specific claim of his work – that a phonologization of prosodic features from an “initial-based

---

12 Huffman (1976) proposes a model for registrogenesis from onset voicing distinctions which is similar in most other respects to Maran’s. Huffman uses four stages (below), which roll Stage 2 and 3 of Maran together, since both refer to the redundant coincidence of segmental and suprasegmental distinctions.

Stage 1: Onset distinctions, vowels are the same
Stage 2: Onset distinction with redundant register split on vowels
Stage 3: Distinctive registers with optional onset distinction
Stage 4: Distinctive registers with no onset distinction
mechanism does not seem to apply without (a) final-based mechanism” having first applied (Maran 1973: 111). In other words, coda consonants produce tone reflexes chronologically first, and then, only with the coda contrast lost, may onset distinctions potentially influence the pitch contour. This claim is not followed here, but Maran’s account follows a common emphasis on segmental effects in the literature on historical tone, particularly at the time of his work (Burling 1967, Matisoff 1970, Gandour 1974, Huffman 1976, Mazaudon 1977, Ferlus 1979, Weidert 1987). Further, his argument coincides with the position in Matisoff (1973), in the same volume (Hyman 1973), that monosyllabicity is a necessary precursor to tone-inducing consonantal depletion.

Thurgood offers another perspective – that tone never develops directly from consonantal voicing distinctions in the onset or coda, but rather through phonation types realized on the vowel. In some cases, the phonation types could themselves be the result of earlier consonant distinctions. Thurgood (1978) suggested such an approach based off of Burmese data, and more formally put forth his hypothesis in Thurgood (2002), concerning the tonogenesis in Vietnamese. According to this hypothesis, prior accounts connecting historical consonant distinctions to modern tone reflexes ought to be reanalyzed with an intervening stage carrying a phonation distinction, as in the diagram in (11).

(11) Tonogenesis-via-phonation in a Hypothetical Language

<table>
<thead>
<tr>
<th>Proto-Language</th>
<th>Intermediate Stage</th>
<th>Modern Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>pat</td>
<td>pat</td>
<td>pát</td>
</tr>
<tr>
<td>bat</td>
<td>pat</td>
<td>pát</td>
</tr>
<tr>
<td>onset voicing contrast /p/ vs. /b/</td>
<td>phonation contrast [TENSE] vs. [BREATHY]</td>
<td>pitch-based contrast High vs. Low tone</td>
</tr>
</tbody>
</table>
Ohala (1974, 1981, 1993) has long argued that diachronic sound changes are grounded in physical phonetic actions, specifically actions which create perceptually confusable structures. This can be termed the Hypo-correction account as it relies on listeners to *under correct* incidental phonetic variance. For tone, this incidental phonetic variance is the intrinsic F0 values on an adjacent vowel caused by certain consonant classes, reported in numerous studies (Lehiste and Peterson 1961, Gandour 1974, Abramson 1977, Kim et al 2002). Hombert, Ohala, and Ewan (1979) and others (Hombert 1975, Hyman 1976, Silva 2006) have examined these F0 patterns while highlighting the role of the listener in tonogenetic processes. The key step for tonogenesis in a Hypo-correction account is when incidental F0 perturbations become recognized and expected by listeners (from Stage 2 to Stage 3 for Maran 1973).

A strict interpretation of Thurgood’s hypothesis by way of Hypo-correction calls for two separate listener “errors”: (1) a segmental voicing distinction is mistaken for a phonation contrast which (2) is in turn mistaken for a pitch contrast. Overlapping the listener errors presents a more probable take on Tonogenesis-via-phonation: intrinsic vowel F0 perturbation caused by consonants is a product of the consonantal phonatory properties, so that linked F0 and phonation differences are concurrently perceived by listeners and the order in which one diachronically becomes the most salient distinction in a contrast is not universal.

All of the above accounts emphasize the incremental instantiation of tone features, whether they concern sound change generally (Ohala 1993), tonogenesis generally (Thurgood 2002), or tonogenesis in Burmese (Maran 1973, Thurgood 1978). If, following Maran (1973), the initial stage is “Atonal” and the final stage is a fully “Tonal”
language, then it becomes apparent that most of the languages discussed above fall somewhere in between, at an intermediate stage. A survey of tone systems comparing these interactions and their historical origin is beyond the scope of this dissertation. Instead, Standard Burmese, with its reported mix of inter-dependent suprasegmental qualities, offers a test case for exploring this issues in-depth within a single language. A series of incremental changes have depleted the consonant coda system, but have not yet produced an entirely pitch-based Tone contrast. Therefore, it is these intermediate stages, and the nature of the synchronic contrast at these stages, that are of interest to this dissertation.

An insightful example of tonogenetic processes can be seen in a recent study by Mazaudon and Michaud (2008) on a dialect of the Tibeto-Burman language Tamang. In Risiangku Tamang, a voicing distinction on initial consonants conditioned the bipartition of a 2-tone system to a 4-tone contrast, shown in the diagram in (12). The subsequent loss of the voicing distinction resulted in a fully supra-segmental system, with the four tones shown on the far right of the diagram.

(12) Tonogenetic Stages in Risiangku Tamang (Mazaudon & Michaud 2008)

<table>
<thead>
<tr>
<th>Early Tone</th>
<th>2-way Contrast</th>
<th>Redundant Stage</th>
<th>4-way Contrast (modern)</th>
</tr>
</thead>
<tbody>
<tr>
<td>* Tone 1 CV₁</td>
<td>* $C_{ph} V_1$ → * $C_{ph} V_1$</td>
<td>* $C_{ph} V_1$ (modal)</td>
<td></td>
</tr>
<tr>
<td>* Tone 2 CV₂</td>
<td>* $C_{ph} V_2$ → * $C_{ph} V_2$</td>
<td>* $C_{ph} V_2$ (modal)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* $C_p V_1$ → * $C_p V_1$</td>
<td>* $C_p V_2$ (modal)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* $C_b V_1$ → * $C_b V_3$ (breathy)</td>
<td>* $C_b V_4$ (breathy)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* $C_p V_2$ → * $C_p V_2$</td>
<td>* $C_p V_3$ (breathy)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* $C_b V_2$ → * $C_b V_4$ (breathy)</td>
<td>* $C_b V_4$ (breathy)</td>
<td></td>
</tr>
</tbody>
</table>
In the intermediate stage, the voicing contrast on initials and the tonal contrast are redundantly associated. Voiced stop onsets gave rise to a breathy, lower-pitched tone counterpart, denoted as either Tone “3” or “4”. By the final stage, representing the modern language under analysis by Mazaudon and Michaud, the breathy voice quality and lower pitch have been phonologized. In the table above, the “Redundant Stage” resembles ‘Stage Three’ of Maran’s model and the modern “4-way Contrast” is effectively ‘Stage Five’. In theory, there is another intermediate stage – Maran’s ‘Stage Four’ – where the distinctive tone carries the contrast, but the conditioning environment still occurs optionally or in certain contexts.

On this point, Mazaudon and Michaud (2008) reported an interesting finding with Risiangku Tamang speakers. In their data, there was an occasional voicing distinction on initial stops. While onsets with Tone 1 and 2 vowels were never voiced, those with Tones 3 and 4 (which had [b]-initial proveniences) frequently were. This production difference reveals a not-quite-complete transfer of laryngeal contrasts from the consonant to the following vowel, which is argued to reflect the optional redundancy of Stage Four, a “missing link in tonogenesis” (Mazaudon and Michaud 2008: 231).
### (13) Tono/Registro-Genesis in Risiangku Tamang (Mazaudon & Michaud 2008)

<table>
<thead>
<tr>
<th>Tone</th>
<th>2-way Contrast</th>
<th>4-way Contrast (modern)</th>
<th>Initial C Voicing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone 1</td>
<td>* C&lt;sub&gt;ph&lt;/sub&gt;V&lt;sub&gt;1&lt;/sub&gt; ↘</td>
<td>C&lt;sub&gt;p,ph&lt;/sub&gt;V&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Tone 1: Never [b] initially</td>
</tr>
<tr>
<td></td>
<td>* C&lt;sub&gt;p&lt;/sub&gt;V&lt;sub&gt;1&lt;/sub&gt; ↘</td>
<td>C&lt;sub&gt;p,ph&lt;/sub&gt;V&lt;sub&gt;1&lt;/sub&gt;</td>
<td>Tone 1: Never [b] initially</td>
</tr>
<tr>
<td></td>
<td>* C&lt;sub&gt;b&lt;/sub&gt;V&lt;sub&gt;1&lt;/sub&gt; ↗</td>
<td>C&lt;sub&gt;p,ph&lt;/sub&gt;V&lt;sub&gt;3&lt;/sub&gt;</td>
<td>Tone 3: May be fully voiced [b]</td>
</tr>
<tr>
<td>Tone 2</td>
<td>* C&lt;sub&gt;ph&lt;/sub&gt;V&lt;sub&gt;2&lt;/sub&gt; ↘</td>
<td>C&lt;sub&gt;p,ph&lt;/sub&gt;V&lt;sub&gt;2&lt;/sub&gt;</td>
<td>Tone 2: Never [b] initially</td>
</tr>
<tr>
<td></td>
<td>* C&lt;sub&gt;p&lt;/sub&gt;V&lt;sub&gt;2&lt;/sub&gt; ↘</td>
<td>C&lt;sub&gt;p,ph&lt;/sub&gt;V&lt;sub&gt;2&lt;/sub&gt;</td>
<td>Tone 2: Never [b] initially</td>
</tr>
<tr>
<td></td>
<td>* C&lt;sub&gt;b&lt;/sub&gt;V&lt;sub&gt;2&lt;/sub&gt; ↗</td>
<td>C&lt;sub&gt;p,ph&lt;/sub&gt;V&lt;sub&gt;4&lt;/sub&gt;</td>
<td>Tone 4: May be fully voiced [b]</td>
</tr>
</tbody>
</table>

Risiangku Tamang demonstrates Maran’s Stage Four, at least when describing the transition in contrastive function from a consonant voicing distinction \( \rightarrow \) tone. The remaining relation between phonation and pitch complicates the generalization (2 breathy vs. 2 modal tones), and points out the inadequacy of a single continuum for a multivalent difference. More importantly for the present study, the Tamang example provides a proof-of-concept that incremental changes can be detected in synchronic data. In this case, intra-speaker variation revealed a prior conditioning feature that no longer bore a functional load. In his dissertation, Brunelle (2005) sought to find incipient tonogenesis in Eastern Cham by looking for inter-speaker variation between age groups (he found none). Turning to Burmese, the modern language presents an opportunity to study an intermediate tonogenetic stage and how this state of affairs is realized in actual speaker data. That is, what are the differences between the phonemic and sub-phonemic distinctions in a multivalent categorical contrast and how are they detected? Can they be diagnosed in intra-speaker or inter-speaker variation, or in the varying ways that redundant cues are produced between the tones or in different phrasal contexts? Answers
to these questions address cross-linguistic concerns that look beyond an improved phonetic and phonological description of Burmese (itself a worthy goal). These data also provide a penetrating look into the process of tonogenesis and advance the current state of knowledge regarding the typology of Tone systems.

Lastly, it should be clarified that positioning Modern Burmese at an “intermediate” stage is not an indication of ongoing tonogenesis. Future developments in Burmese tone form an interesting question, but are not a significant concern for this dissertation. Rather, the discussion of diachronic tone concerns how the present state was arrived at, as clearly some process resembling tonogenesis helped produce the modern system of suprasegmental contrast. Whether this ultimately leads to a strictly pitch-based tonal system is a matter of prophesy, not phonology. In his discussion of possible tonogenesis in Cham, Brunelle (2005) is careful to emphasize that a possible or precedentented outcome is not a necessary outcome. To extend the analogy of tonogenesis as a “path” or “route” to tone, it is just as driving on the proper road towards a destination does not require arriving at that destination. There are plenty of stops, exits, and opportunities to u-turn along the way.

1.4 Outline of the Dissertation

The present chapter has presented the general problem of classifying Burmese tone. In Chapter Two, previous impressionistic and instrumental studies of the phonetic nature of the distinction are reviewed, which reveal numerous discrepancies that have either hindered or obfuscated previous analyses of the tones’ contrastive nature. Chapter Two then turns to the literature regarding the phonological representation of tone generally and
then, of tone in Burmese. In Chapter Three, production experiments intended to clarify these discrepancies are described. The experiments use the same subjects and elicitation stimuli, but investigate different modes of production (acoustic pressure, vibration at the glottal source, and oral airflow), simultaneously in some cases. In Chapters Four and Five, the results of these experiments are then described as they address the phonetic question of acoustic correlates to each tone (Chapter Four) and phonological concerns such as the analysis of meter and permissible syllable structures in Burmese (Chapter Five). Chapter Five tackles the pertinent phonological issues via new statistical analyses of data presented in Chapter Four, as well as new data, and historical data concerning the diachronic development of the Burmese syllables. Chapter Six describes the methods and results of a simple perception study testing the same set of subjects sensitivity to the multiple phonetic dimensions composing tone. Finally, in Chapter Seven, the various sources of data are collectively considered in the presentation of a phonological model of both the underlying contrast on lexical items and the common alternations found in the production experiment.
Chapter Two. Phonetic Properties of Burmese Tone

2.1 Introduction

The following chapter reviews the literature regarding the phonetic correlates of the four Burmese tones, as reported in Burmese and in relevant studies for other languages. The literature here includes a variety of sources, from phonological explications (Green 2005) to experimental phonetic studies (Thein Tun 1982) to linguistic grammars (Okell 1969) and second-language learning manuals (Mesher 2006). Interestingly, where experimental phonetic studies have been the least clear, the impressionistic sources tend to agree – the association of distinct phonation types with each tone. Notable works on the Burmese grammar and sound system (Stewart 1955, Cornyn 1945, Okell 1969, Sprigg 1977, Thurgood 1978, Bradley 1982, and Watkins 2001) are consistent in their labeling of voicing patterns during the Low, Creaky, and (usually) High tones. However, the acoustic studies in Watkins (1997, 2005), Ladefoged, Maddieson, and Jackson (1988), and Gruber and Feizollahi (2006) have all met inconclusive results in that phonation type has proven not only difficult to associate with the tones, but difficult to detect at all in Burmese speech by acoustic means. This discrepancy between descriptive phonation qualities and experimental results serves as a central motivation for this dissertation’s phonetic research. While the chapter does review claims regarding duration, intensity, and pitch contours produced with each tone, the focus concerns claims about phonation types and the experimental designs which have yielded these conclusions.

Instrumental studies which have examined Burmese tone are Watkins (1997, 2005a, 2005b), Mehnert and Richter (1972, 1973, 1976, 1977), Thein Tun (1982), Javkin and Maddieson (1983), and the author’s own collaborative work (Gruber and Feizollahi
A few other works (Ladefoged, Maddieson, and Jackson 1988, Ladefoged and Maddieson 1996, and Lee 2007) offer small samples of acoustic data concerning Burmese – these are included when relevant as they pattern with others’ data, reaffirming that the observations hold with other speakers and for other researchers. It will be shown that the range of experimental studies mentioned here in fact rarely conflict genuinely with one another. While the various researchers’ findings often do not align, this is often for the superficial reason that the different experiments measured different types of data – i.e., different words or contexts.

As many of these works are cited repeatedly throughout this chapter, each study and its data collection – speakers, stimuli, context - is reviewed briefly in the table in (1). These sources tend to focus on a single type of data (pitch or modes of phonation, but not both). Accordingly, the sections which follow are divided not by author, but by each phonetic quality that together constitute the “suite” of features that compose lexical tone in Burmese. In order of increasing complexity, these are intensity (§2.2), duration (§2.3), pitch (§2.4), and phonation (§2.5). The final portion of the chapter reviews discussion of the origin of Tone distinctions (§2.7) and phonological analyses of the composition of tone features (§2.8.1 – 2.8.3) and syllable structure (§2.8.4) in Burmese.
## 1. Overview of Prior Instrumental Studies of Burmese Tone

<table>
<thead>
<tr>
<th>Study</th>
<th>Subjects</th>
<th>Measured Variables</th>
<th>Stimuli</th>
<th>Context</th>
<th>n per subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thein Tun (1982)</td>
<td>4 (2f, 2m)</td>
<td>F0</td>
<td>[h] onset, 50 vowels(^1)</td>
<td>Bisyllabic phrase: <em>[da]</em></td>
<td>50</td>
</tr>
<tr>
<td>Experiment 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thein Tun (1982)</td>
<td>12 (m)</td>
<td>DURATION F1 – F3</td>
<td>[h] onset, 51 possible vowels(^1)</td>
<td>Bisyllabic phrase: <em>[da]</em></td>
<td>102</td>
</tr>
<tr>
<td>Experiment 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Javkin &amp; Maddieson (1983)</td>
<td>1 (m)</td>
<td>GLOTTAL WAVEFORM (inverse filter) final 20 cycles</td>
<td>15 lexical items: [la, ṭa, ma, ṭa, sa] (unbalanced)(^2)</td>
<td>Citation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>72</td>
</tr>
<tr>
<td>Mehnert &amp; Richter (1976)</td>
<td>2 (1f, 1m)</td>
<td>DURATION F1 – F3</td>
<td>[ʔ] onset, 51 possible vowels(^1)</td>
<td>Citation</td>
<td>51</td>
</tr>
<tr>
<td>Watkins (1997)</td>
<td>2 (m)</td>
<td>F1 – F5 (Hz, dB) at mid-vowel</td>
<td>[p, t, k, m, n, ŋ] onsets, vowel [a]</td>
<td>Citation</td>
<td>48</td>
</tr>
<tr>
<td>Watkins (2005a)</td>
<td>1 (m)</td>
<td>F0 F1 – F3 at mid-vowel EGG Closed Quotient</td>
<td>[p, t, k, m, n, ŋ] onsets, vowel [a]</td>
<td>Low __ Low varies</td>
<td></td>
</tr>
<tr>
<td>Watkins (2005b)</td>
<td>≤ 4</td>
<td>F0 at onset + offset EGG (CQ) at onset + offset</td>
<td>not known</td>
<td>16 bisyllabic combinations not known</td>
<td></td>
</tr>
<tr>
<td>Gruber &amp; Feizollahi (2006)</td>
<td>4 (2f, 2m)</td>
<td>DURATION, F0 H1-H2 at mid-vowel</td>
<td>[m] onsets, 5 vowels</td>
<td>Low __ Low</td>
<td>80</td>
</tr>
</tbody>
</table>

\(^1\) See example (1.8) for the fifty-one possible Vowel-Tone combinations in Standard Burmese. Experiment One of Thein Tun (1982) excluded the minor syllable vowel [ŋ].

\(^2\) Stimuli list was “unbalanced” because every tone was not paired with every onset. There were five stimuli each of the Low and Creaky tone, but only three High tone and two Checked tone words were used.
2.2 Intensity

The relative intensities between the tones are uncontroversial amongst the many researchers providing phonetic descriptions. A typical description is seen in (2):

<table>
<thead>
<tr>
<th>Tone</th>
<th>Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘checked’</td>
<td>very high</td>
</tr>
<tr>
<td>‘creaky’</td>
<td>high</td>
</tr>
<tr>
<td>‘high’</td>
<td>medium</td>
</tr>
<tr>
<td>‘low’</td>
<td>low</td>
</tr>
<tr>
<td>‘minor’</td>
<td>lowest, very low</td>
</tr>
</tbody>
</table>

Others authors indicate no distinction between the intensity of Creaky and Checked syllables, and also High in some cases, but all sources identify a pattern of decreasing intensity from these three tones to Low and Minor syllables. The rankings either in (3a) or (3b) represent any impressionistic account from any of the nearly twenty authors cited in this chapter:

(3a) Creaky and/or Checked Tone > High Tone > Low Tone > Minor Syllable

(3b) Creaky, Checked, and High Tone > Low Tone > Minor Syllable

Acoustic studies of Burmese have generally chosen to not measure intensity, but there is some evidence indicating that the consensus cited in (3) reflects only citation form utterances. Thein Tun (1982: 79) justifies his choice to ignore intensity (and measure

---

only F0 and duration as correlates of tone) on the grounds that “…the intensity display trace of the tones does not show any consistency or any common factor by which the contrastive features of the tones can be distinguished.” Confusingly, he then argues that the stimuli context used to record his data was too static, causing “the intensities of the tone (to) not behave as regularly as they do in sentences” (Thein Tun 1982: 79). It is highly unlikely that the tones behave more regularly across different sentences than in a controlled short phrase since focus, sentential stress patterns, and other variables affect the quantifiable intensity of a single syllable (Watkins 2005b). Regardless of the author’s interpretation, Thein Tun’s lack of meaningful findings suggests that intensity distinctions which are highly salient in citation form are more likely to be leveled in medial positions.

2.3 Tone Duration

As with intensity, all sources generally agree on the relative duration of the four tones: both Low and High tone syllables are pronounced with a long duration, Creaky syllables tend to be shorter, and Checked syllables are consistently the shortest. There are minor disagreements between descriptions, but none that defy the order presented in (4).

(4) High, Low >> Creaky >> Checked >> Minor syllable

For example, either Low or High tones (or both equally) are considered to be the longest, and the duration of the Creaky tone is implied by some to be rather short and by others to be of “medium” duration and much nearer the length of High and/or Low than Checked
syllables (Firth 1936, Wheatley 1987). The survey of authors in (5) demonstrates this consensus and suggests that the scale in (4) holds across a number of phonetic environments as well.

(5) Survey of Experimental Findings and Researcher Observations for Tone Duration

<table>
<thead>
<tr>
<th>Source</th>
<th>Scale: Longest $\rightarrow$ Shortest</th>
<th>Context or Type of Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Javkin &amp; Maddieson (1983)</td>
<td>High &gt; Low &gt; Creaky &gt; Checked</td>
<td>Isolated (V only)</td>
</tr>
<tr>
<td>Mehnert &amp; Richter (1976)</td>
<td>Low &gt; High &gt; Creaky &gt; Checked</td>
<td>Isolated (V only)</td>
</tr>
<tr>
<td>Lee (2007)</td>
<td>Low &gt; High &gt; Creaky &gt; Checked</td>
<td>Isolated (onset+V)</td>
</tr>
<tr>
<td>Thein Tun (1982)</td>
<td>High &gt; Low &gt; Creaky &gt; Checked</td>
<td>Phrasal</td>
</tr>
<tr>
<td>Watkins (2005a)</td>
<td>Low, High &gt; Creaky &gt; Checked</td>
<td>Phrasal</td>
</tr>
<tr>
<td>Gruber &amp; Feizollahi (2006)</td>
<td>High &gt; Low &gt; Creaky &gt; Checked</td>
<td>Phrasal</td>
</tr>
<tr>
<td>Armstrong &amp; Tin (1925)</td>
<td>High, Low &gt; Creaky &gt; Checked</td>
<td>Impressionistic</td>
</tr>
<tr>
<td>Firth (1933, 1936)</td>
<td>High, Low &gt; Creaky &gt; Checked</td>
<td>Impressionistic</td>
</tr>
<tr>
<td>Cornyn (1945, 1968)</td>
<td>High, Low &gt; Creaky &gt; Checked</td>
<td>Impressionistic</td>
</tr>
<tr>
<td>Bernot (1963)</td>
<td>High &gt; Low &gt; Creaky</td>
<td>Impressionistic</td>
</tr>
<tr>
<td>Bradley (1982)</td>
<td>High &gt; Low &gt; Creaky &gt; Check &gt; Minor</td>
<td>Imprsn., Citation$^5$</td>
</tr>
<tr>
<td>Wheatley (1987)</td>
<td>High &gt; Low, Creaky &gt; Checked</td>
<td>Imprsn., Citation$^5$</td>
</tr>
<tr>
<td>Watkins (1997)</td>
<td>High &gt; Low &gt; Creaky &gt; Checked</td>
<td>Imprsn., Citation$^5$</td>
</tr>
</tbody>
</table>

Scale of Duration by Tone as reported across multiple studies, representing both the findings of acoustic experiments (in white) and personal observations of other authors typically in undefined phonological contexts (in shading).

The abridged experimental results at the top of the survey are reproduced more fully in Figures 1 (Isolated tokens) and 2 (Phrase-medial tokens) below. Duration data for all six

$^4$ Other than for Gruber and Feizollahi (2006), rankings given in (5) are not statistically significant, as no inferences can be made regarding other researchers’ data without a standard error for the reported means. Watkins (2005a) is listed with High, Low tones unranked because it was determined to be highly unlikely that Watkins’ reported means for duration of High and Low vowels were different from one another. It is recognized here that this is not certifiably shown with Watkins’ figures, which included only the mean duration and $n$ count.

$^5$ These three authors all explicitly note that they are strictly describing the tones as uttered in citation form and that the phonetic realization of each tone can vary greatly in other cases.
studies shown conform to the relative scale in (4), notably exhibiting no clear distinction between High and Low tones across studies. One dissenting claim comes from Thein Tun (1982), who reported that even High and Low tones strongly contrasted for his speakers (see Figure 2). Buoyed by the consistent four-way duration distinction in his data,

![Mean Duration of Tones in Citation Form](chart1)

**Figure 1.** Duration (in ms) of the vowels in each tone from three acoustic studies which recorded tone-bearing syllables produced in isolation. Javkin and Maddieson (1983) conforms to the standard descriptive order High > Low > Creaky > Checked. Both Mehnert and Richter (1976) and Lee (2007) found Low > High (though with small samples, i.e. $n = 3$ per tone for Lee (2007)).

![Mean Duration of Tones in Phrase-medial context](chart2)

**Figure 2.** Duration (in ms) of the vowels in each tone from three studies recording tone-bearing syllables within a frame sentence. Watkins’s (2005a) measures tend to be much longer, but all studies fit the relative order High > Low > Creaky > Checked. TT $n = 24$ per tone (12 speakers); G&F $n = 80$ (4 speakers); Watkins $n = 12$ (1 speaker).

---

6 Reported data in Thein Tun (1982) are insufficient for a retrospective statistical test for significance. He does share the fact that of 1200 measured tokens, just 34 do not follow the pattern High > Low > Creaky > Checked, and the majority of these discrepant tokens were cases of High shorter than Low.
though heeding caution without having conducted a suitable auditory experiment, he surmises that “total length is the most important phonetic cue by which the hearers can distinguish the phonological tones” Thein Tun (1982: 93). This claim does not withstand further investigation, both in this dissertation and by other researchers.

Comparison of the charts in Figures 1 and 2 suggests that the duration difference between Creaky and High/Low syllables is much more pronounced in isolated or citation utterances. The durations in Figure 2 indicate some leveling of this difference when the tones are pronounced phrase-medially. The same cannot be said for Checked syllables, which are always the shortest tone regardless of context.

These observations are supported by statistical analysis of the data provided in Gruber and Feizollahi (2006), for whom only one of four speakers had significantly different durations for all four tones (High > Low > Creaky > Checked). The three other speakers in their study produced High and Low syllables without significantly different durations, and for two speakers there was also no statistical significance between High, Low and Creaky tone syllable durations with only Checked syllables being significantly shorter. These results across a number of studies are quite congruous to those given in Chapter Four of this dissertation, where phrase-medial tone-bearing syllables are examined in greater capacity than has heretofore been seen. Creaky syllables are, on the whole, of shorter duration than High and Low syllables, but in many instances the difference is minimal in comparison to the much shorter Checked syllables.

The pattern noted above calls into question the presumed causes of the (partial) neutralization in medial position. The effect seen in embedded tokens may alternatively be due to (i) adjacent words and phonetic material or simply to (ii) a quicker speaking
rate in casual, connected speech. An increased rate of speech presents a likely cause for
duration shortening, however, Creaky and Checked tones were not markedly shorter in
studies examining medially embedded tokens. Rather, the reduction in duration contrast
across tones is best characterized as the compression of High and Low syllable length,
whereas Creaky and Checked syllables are already too short to be further compressed by
a faster rate of speech. Conversely, the difference between contexts could also be
understood as a hyper-performance of High and Low tone distinctions in citation form
utterances. Regardless of the cause, the differences found on syllables uttered sentence
medially at a casual speaking rate are far-reaching, and will be discussed further
regarding every phonetic property under investigation. Duration though, plays a crucial
role with other properties, as a shorter vowel has less time available for articulatory
differences to manifest themselves. In the discussion of pitch and phonation results, the
quantifiable durational effects of ‘speaking rate’ are more easily distinguished from
stylistic effects brought on by adjacent syllables and their segments.
2.4 Pitch

Sprigg (1977) offers the most detailed account to date of pitch and tonal pitch alternations in Burmese. Drawing on the pronunciation of compound lexical items with which he was familiar, he illustrates single examples of most possible sequences of three sequential tones. Sprigg acknowledges, as others do (Thurgood 1978, Bradley 1982, Watkins 2005a, 2005b), that this effort just begins to account for the myriad of effects caused by tone sandhi, semantic focus, and intonational patterns in Burmese phonology. For experimental purposes, these tonal and prosodic phenomena, or the current state of knowledge concerning them, conspire to make naturalistic data appear impenetrable to investigation of pitch contours. Yet the descriptions of pitch by the many researchers looking at Burmese tone are felicitously consistent with one another. This consistency is of course not due to an unexpectedly transparent system of pitch contrasts in the language, but rather to (i) the careful control of pitch conditions in the few instrumental studies and (ii) impressionistic accounts’ penchant for describing only the tones’ citation forms. In fact, each of the few striking differences between accounts (e.g. High tone listed as either falling or rising) can be attributed to the varying contexts that served as the basis for description.

Looking just at the pitch peak, either the Creaky and Checked (Bradley 1982, Watkins 1997) tones are generally considered to reach the highest F0 values, while others indicate no meaningful difference between Checked, Creaky, or High tone pitch peaks. Greater disagreement is found concerning the pitch contours associated with each tone: the two most discrepant claims are whether the High tone rises or falls and whether the Checked tone is a level high or a high falling tone.
Descriptions of pitch for the Low and Creaky tones are for the most part uniform; Low as low pitched and even, Creaky as High pitched but falling. It is worth pointing out that the Low tone’s “low pitch” is only “relative to the other tones” (Stewart 1955) – a trait alluded to in some of the labels applied to the tone. Firth (1933) calls it a “Mid tone”, Bradley (1982) uses the name “Even” and Wheatley says “Level”, all suggestive that production involves the speaker’s regular pitch range rather than a marked low pitch target. A few researchers note an additional phenomenon. Javkin and Maddieson (1983) describe the Low tone as a “Rising” tone and Okell (1969) and Cornyn both offer that the Low tone “sometimes gently rises toward the end” (Cornyn 1945: 5). Okell (1969: 11) offers that the final rise may occur “before a pause”, but not in phrase final position. This pre-pausal rise was found regularly for four of the twelve speakers who contributed speech data for this dissertation (see Chapter 4), in both isolated syllable utterances and in final positions.

In Figure 3 are F0 traces from the same studies as the duration data in Figure 2, for which the measured syllable was produced before a Low tone syllable. The two speakers shown from Gruber and Feizollahi (2006) maintain similar pitch contours as one another. The mean F0 values from the one speaker in Watkins (2005a) produce contours consistent with these two speakers, the differences over the initial portion being attributable to the inclusion of the nasal onset in Gruber and Feizollahi’s (2006) stimuli. The Low syllable bore no distinguishable pitch peak and gradually fell or remained relatively level. Both the Creaky and Checked syllables had early pitch peaks of roughly the same height. The High syllable for all speakers rose in pitch gradually to a later peak of nearly the same frequency as Creaky or Checked peaks and then fell slightly (except in
Figure 3. Comparison of F0 traces for four speakers from three prior studies.
Values from Gruber and Feizollahi (2006) (Fig. 3a, 3b) represent means of 40+ tokens. Figure 3c displays F0 means ($n = 12$) for each tone in Watkins (2005a). Figure 3d displays mean F0 ($n=28$ for each tone, but 16 Checked tokens) at the vowel onset and offset for two female subjects in Thein Tun (1982).

Watkins (2005a)) at the syllable end. Accounting for the sonorant onset in Gruber and Feizollahi (2006), the F0 tracks in Figure 3 are sufficiently congruous to leave little doubt about the expected pitch contour of the four tones in this single context.

A primary reason the impressionistic descriptions conflict regarding the tones’ distinctive pitch contours is because the tones and the pitch contours are frequently modified by their syntactic, morphological, and phonological context. Descriptions of the High tone typically say it is long and gradually falling in pitch, but some accounts (Watkins 2005a) show it to be long and rising. Bradley (1982) points out that it is indeed
both: long with a rising F0 when before another High or Low tone, but with a gradually falling pitch before a pause or a Creaky or Checked tone. More evidence comes from Watkins (2005a), who states that in a 2000 study, he found High tone syllables to rise when preceding a Low or another High tone in close juncture. Acoustic data in Thein Tun (1982), Watkins (2005a), and Gruber and Feizollahi (2006) further confirm this observation.

The same studies provide little evidence of the variation reported in Checked tones, showing only a falling contour and never an even high pitch. Since impressionistic accounts do not elaborate on the association, it is not clear to what extent or in what circumstances these alternate contours occur. However, plausible explanations are easily conjured for a tone/syllable noted for its brevity and segmental alternations in close juncture (cf. (1.4)). Simplifying a bit, a missing fall could be attributed to an abbreviated duration or to the quality of the assimilated coda consonant, understandings that retouch upon the competing interpretations of medial position effects as durational or positional. For German and dialects of English, Grabe (1998) and others have shown that languages can have different strategies for realizing a set pitch contour over appreciably different durations, by either (a) compressing (or stretching) the contour or (b) maintaining the rate of change and thereby clipping off late pitch targets for which there is not enough time to achieve (see also Grabe et al 2000, Nolan 2006). Similar truncation of contours has been found for a Southeast Asian lexical tone contrast in Morén and Zsiga’s (2006) observations concerning Thai. Many Thai tones were shown in connected speech to realize only the earlier pitch targets of their canonical form: “Falling” tones were produced without a fall, “Rising” tones without a rise or with just a slight, very late rise
(see also Nitisaroj 2006). Syllogistically, one might conclude that Burmese falling tones may also not realize their fall in connected speech. Whereas the truncated contours in Morén and Zsiga (2006) were argued to result from a positional lack of prominence, evidence here suggests any neutralization of an F0 fall in Burmese is at least partially ascribed to an adjoining syllable and its onset. It is not yet determined how the falling contour of Checked syllables is realized in adjoined syllables such as [CVp.pV], [CVl.lV], [CVm.mV], etc… Additionally, data from Watkins (2005b) and Sprigg (1977) indicate that the tone of an adjoining syllable may also bear a meaningful effect on pitch contours.

In Watkins’ study, one speaker pronounced a set of bisyllabic tokens, placing a syllable of each tone in the first and second position to create sixteen two-syllable sequences. No further words or tones preceded or followed this bisyllabic token. The resulting pitch contours over two consecutive toned syllables offer an unsurprising finding — syllables bearing the same phonological tone will have different pitch contours in different tonal contexts. Some of the bisyllabic pitch contours appear indicative of phonetic co-articulation between adjacent pitch targets: a Low syllable rising slightly in anticipation of the early pitch peak in High, Creaky, or Checked syllables. Others seem less phonetically motivated: the typically very high-pitched Creaky tone showing a reduced pitch peak and fairly even contour when preceding a High syllable. The tri-syllabic examples in Sprigg (1977) report similar alternations. One example, the Checked-Creaky-High sequence in [ʧɛm.mə.lɛ:] “little hen”, claims a low even pitch through the first two Checked and Creaky syllables when preceding a final High tone (Sprigg 1977: 17). Further examination of contextual effects is warranted, with an eye to identifying phonologized alternations and those which are phonetically motivated.
To summarize the findings regarding pitch, acoustic studies of Burmese have found similar pitch patterns as one another but importantly, they have also all studied (with the exception of Watkins, 2005b) toned-syllables only in citation form or in a null tonal context of [Low __ Low]. It is clear that different contexts alter the pitch contour, a fact that readily explains the inconsistent impressionistic descriptions made by authors who do not necessarily refer to a single context. Despite this confusion, a loose picture of pitch behavior emerges which is consistent with the F0 findings across multiple sentential contexts in this dissertation, provided in Chapter Four. The summary table in (6) arranges a suitable consensus\textsuperscript{7} of descriptions into just two contexts: isolated or sentence medial positions.

<table>
<thead>
<tr>
<th>Tone</th>
<th>Citation or isolated syllables</th>
<th>In sentential context</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Height</td>
<td>Contour</td>
</tr>
<tr>
<td>Low</td>
<td>moderately low</td>
<td>even</td>
</tr>
<tr>
<td>High</td>
<td>initial high</td>
<td>gradual fall</td>
</tr>
<tr>
<td>Creaky</td>
<td>initial high</td>
<td>steep fall</td>
</tr>
<tr>
<td>Checked</td>
<td>initial high</td>
<td>steep fall</td>
</tr>
</tbody>
</table>

= no contextual difference ≠ different pitch contours ≈ some differences reported between contexts reported between contexts

Burmese uses three contour tones and one even, relatively low tone. Two of the contour tones, Creaky and Checked, have indistinguishable falling pitch contours in both the existing descriptions and the experimental findings of multiple phonetic studies. In the cases when these tones are not falling, they begin high and remain high, failing to realize

\textsuperscript{7} Various discrepancies existed of course, primarily with the Low and Checked tones. See throughout §2.4 above.
a pitch fall. In other words, Creaky and Checked tone syllables are almost always found with a high pitch peak early in the syllable, and never with a rise to that peak. The other tone with a contour (High) is found to either rise or fall depending on the context (according to the three most in-depth descriptions of Burmese tone in the literature – Sprigg 1977, Bradley 1982, Watkins 2005a) and to do so more gradually than the other two contour tones. Lastly, the even tone (Low) is noted for its absence of pitch targets, high or low. It is reported to occasionally rise in pre-pause positions, a context which suggests a possible phrase-level intonational melody as opposed to an underlying quality of the tone.

2.5 Phonation Type

2.5.1 The Role of Phonation Type

The link between phonation type and tone poses a central question of this dissertation: Is there a difference in phonation between the four tones of Burmese and if so, does it bear contrastive significance? Gordon (2001) and Gordon & Ladefoged (2001) model phonologically relevant phonation types as a continuum of the degree to which the glottis is open: from wide open – permitting high rates of airflow, to completely shut – producing a glottal stop. Breathy phonation (with an abducted glottis) and creaky phonation (with an adducted glottis) occupy opposite sides of this continuum, with modal

---

8 More precise modeling of phonation type to glottal configuration can be found in Catford (1964, 1968), Laver (1980), and Esling & Harris (2005), which utilize more phonatory distinctions than any language is known to contrast. Gordon and Ladefoged’s (2001) use of modal, creaky, and breathy should be understood to be a 3-way division of a “complex articulator” for which accurate modeling would require more than one linear dimension. While it may be a matter of controversy whether Burmese tones reflect phonatory differences at all, there is little doubt that a 3-way distinction recognizing modal voicing, a state more open than model, and another less open than modal will be sufficient to capture the major contrast present in Burmese.
phonation occupying an intermediate position. Following some descriptions of the tone contrast in Burmese as a four-way contrast in mode of phonation (à la Bradley 1982), the four tones represent four points along the continuum illustrated in Figure 4.

![Figure 4: Four-way Register Hypothesis](image)

Figure 4. Four-way Register Hypothesis set on a continuum of phonation modes as defined by glottal state. Continuum adapted from Gordon and Ladefoged (2001).

Figure 4 represents a strong hypothesis about the role of phonation type which will be referred to as the “Four-way Register Hypothesis”. The literature on this matter has been described above as highly conflicting, but note that the conflict is essentially between impressionistic descriptions in grammars or philological accounts and the results of multiple instrumental phonetic studies. On the one hand, impressionistic sources either make the same claims as one another or make none at all about certain tones. For example, the High tone is sometimes associated with breathy phonation (as in Firth 1936, Thurgood 1981, or Bradley 1982), but is just as often described without any reference to phonation type (Cornyn 1945, Okell 1969). On the other hand, acoustic and other instrumental phonetic studies unhappily agree that results indicating a specific phonation type are murky at best. In the case of the High tone’s association with breathy voicing, no study to date has reported any consistent acoustic correlate of breathiness, increased airflow, or a more open glottal configuration. After a necessary discussion of key terms in
the following section, subsection §2.5.3 briefly address the impressionistic accounts and §2.5.4 - §2.5.6 detail the numerous negative findings from phonetic studies.

2.5.2 Creaky Tone, but “creaky” voice?

Not only do the tenuous experimental results cast some doubt on how reliable the phonation contrast is, but descriptions of pitch (Figure 1.3 and §2.4) and acoustic data from Watkins (1997, 2005a) and Gruber and Feizollahi (2006) reveal phonetic qualities that do not fit well with the definition of creaky phonation found in the literature on glottal states (Hollien et al. 1966, Ladefoged 1971, Laver 1980, Laver 1991, Esling and Harris 2005, Pennington 2005, Edmondson and Esling 2006). Creaky phonation is produced by increased glottal tension and compression (Ní Chasaide and Gobl 1997) and (likely) a lowered larynx (Laver 1980, Hirose 1997) which has the acoustic consequence of lowering pitch and substantial amounts of jitter and shimmer, period-to-period variation in frequency and amplitude respectively (Childers and Lee 1991). In contrast to this, the Burmese Creaky tone is uniformly described as bearing a very elevated pitch and an ensuing fall, often without mention of the uneven or irregular vocal fold vibration.

Investigation of Creaky syllable waveforms recorded by this author for Gruber and Feizollahi (2006) also reveals little of the irregularly spaced glottal pulses expected in a creaky vowel. Yet, as seen in the production data collected for this dissertation (see earlier examples in Figures 1.1 – 1.4), more constricted or more open laryngeal states do occur with the predicted tones in at least some cases. The discrepancy is perhaps terminological. The term “creaky” has a long tradition within Burmese linguistics, where it has acquired a customized meaning that does not necessarily take into account how it is
used in the greater phonetics community, although the different uses are certainly related in the sense of increased laryngeal tension of some kind. Ladefoged, Maddieson, and Jackson (1988) describe the Burmese Creaky articulation as one where the larynx “becomes increasingly tense” (italics added) over the duration of the vowel. Kingston (p.c.) and Fulop (p.c.) have similarly suggested “tense” voice is perhaps more suited to the Burmese production. Tense phonation, as with Creaky, refers to a tighter and more compressed glottal state, but differs in that the laryngeal tension is a result of more stiffness in the vocal fold cover, which may be produced by a raised larynx in contrast to the laryngeal lowering often associated with creak (Edmondson et al. 2001, Edmondson and Esling 2006). This tense state produces an increased rate of regular, periodic vibration.

Terminology is certainly part of the confusion, but the data informing this opinion are not without problems. In fact, one point that can be drawn from the following survey (§2.5.3 - §2.5.6) is that it is not entirely clear that the phonetic evidence is so irreconcilable with creaky voicing. As noted in the quote above by Ladefoged, Maddieson, and Jackson (1988) and throughout this work, a key aspect of the laryngealization on Creaky syllables is timing. If creaky voicing is only produced terminally on Creaky tone syllables and not during the initial portion of the vowel, then the mismatch with heightened pitch is less puzzling. High pitch occurs early and is followed by a sharp fall aligned with increasing laryngeal tension over the second half of the vowel. Additionally, many Creaky tone recordings collected for this study, unlike previously reported data, do show “irregularly spaced glottal pulses” – particularly at the vowel offset. The reasons for this inconsistency are discussed with these data.
Details aside, it is clear that Burmese “creak” relates some kind of glottal adduction that shares properties with Laver’s (1980) and later descriptions of creaky voice. The classification of the particular Burmese production under one of the canonical labels “creaky”, “tense”, or “harsh” voice poses an interesting and worthwhile question, but more finely re-labeling Burmese “Creaky tone” is not a necessary goal of this dissertation. Numerous philologists, grammarians, and phoneticians have used the title “Creaky” so that there is neither enough motivation nor evidence, without laryngoscopic data for example, to be more precise about the physiological nature of the constriction. Accordingly, the terms “creaky” or “laryngealized” are sufficiently descriptive for present purposes and will be used throughout this dissertation to refer to the laryngeal constriction found on Creaky or Checked syllables.

2.5.3 Phonetic Studies: Impressionistic Descriptions

Thurgood (1978: 224) offers a revealing comment on the predictability of marked phonation types in Burmese, claiming that while every tone “may occur with a clear voice quality, the breathy voice is found only with the heavy (High) tone, the creaky voice is found only with the creaky tone, (and) the level tone always occurs with the unmarked clear voice quality.” So a word produced with breathy voicing is almost certainly a High tone word, but a High tone word is not certain to be produced with breathy voice. Possible interpretations of Thurgood’s comment are that the phonation distinction is frequently neutralized to modal voicing or that it is in free variation and phonetic properties other than phonation serve as cues to the Tonal identity.
Context-sensitive neutralization seems likely in light of other claims by Thurgood (1978) that Creaky syllables are less creaky when followed by another syllable in close juncture. A similar, and more widely recognized, process in Burmese phonology is the loss (via complete segmental assimilation) of glottal stop codas in Checked syllables also in close juncture. This alternation, described in §1.2.4, is one sense the complete assimilation of one obstruent to another, but may also be seen as a loss of glottal constriction in connected speech, not unlike the loss of constricted creaky voicing for Creaky tones reported above. The observations by Thurgood (1978) should be held in mind when looking at the inconclusive and generally negative findings of experimental studies.

Including Thurgood’s description, impressionistic accounts present a fairly cohesive picture of phonation in Burmese that rarely deviate from the set of associations making up the Four-way Register Hypothesis (see §1.4 and Figure 5 above) other than in cases where no specific phonation type is attributed to a tone. The most common deviation states Checked syllables are produced with normal modal phonation (as in Bradley 1982). Other variants include a single author (Firth 1933) ascribing a breathy quality to Low tone words and Watkins (1997, 2005) referring to Checked syllables as “tense” rather than creaky (or modal). More often though, the Checked tone is not explicitly affiliated with a phonation type, though a glottal state (full adduction) at the end of the vowel is presupposed in the production of a glottal stop.

Similarly for the other tones, many authors do not expressly mention phonation. The resulting ambiguity can be difficult to interpret, particularly for the High tone where silence may be suggesting the tone is produced with regular modal vibration and specifically not breathy voicing, or it is possibly just making no claim regarding
phonation. Likewise, a few authors do not specifically mention creaky voicing with the Creaky tone, although it is more likely these authors felt creakiness did not warrant mention, being implicit already in the label “Creaky tone”. The Low tone is regularly not attributed any phonation type, which is likely not a meaningful omission for a toneme lacking any marked phonation quality. Authors who do mention phonation for the Low tone uniformly call its pronunciation modal, “normal” (Bradley 1982) or “plain” (Yip 1995) voiced. The sole exception to this being Firth’s (1933) observation (see above) which is noteworthy not just for being unique, but also for reflecting one of the major findings of Chapter Four: that Low tone syllables are often just as breathy as the High tone.

2.5.4 Phonetic Studies: Acoustic measures

Acoustic measures of phonation type look at the distribution of spectral energy, which can indirectly reflect qualities of the glottal source prior to vocal tract filtering (Rothenberg 1981, Ladefoged, Maddieson, and Jackson 1988, Klatt and Klatt 1990, Kirk et al. 1993, Ní Chasaide and Gobl 1997, Ladefoged 2003). Most of the spectral measures for phonation examine the amount of energy in higher harmonics, where voicing produced with a tenser, more constricted larynx tends to have comparatively greater energy. Rothenberg (1981) states this correlation is the result of the different closing velocities of the vocal folds: rapid vocal fold closing during tenser creaky phonation creates an excitation of the upper harmonics that does not occur with breathy or lax voicing. Usually, the strength of higher-band energy is quantified as the dB difference between the amplitude of a lower harmonic (normally H1, the first harmonic) and that of
a higher harmonic (either H2 or the harmonic nearest F1 or F3, written as A1 or A3). The measures (H1-H2, H1-A1, H1-A2, H1-A3) are known alternatively as spectral slope, spectral skew, or spectral tilt as they roughly portray how steep the drop-off in energy is from low to high-range frequencies. In relation to the phonation types under study in this dissertation, breathy voice quality is known to produce a steep spectral tilt as the drop-off to the lower amplitude high frequencies is greater than in modal or creaky voicing. Creaky voicing results in a flatter roll-off to higher energy upper harmonics. Modal voicing should be intermediately skewed (Gordon and Ladefoged 2001, Ladefoged 2003). This contrast is illustrated with data from Burmese in the spectra in Figure 5, taken from a single speaker in this study (Speaker I, male).

![Figure 5. Spectra from representative (a) breathy, (b) modal, (c) creaky vowels.](image)

For many languages other than Burmese, measures of H1-H2 or H1-F1 have successfully differentiated creaky or breathy modes of phonation. Ladefoged and Maddieson (1985) found H1-H2 to distinguish tense and lax registers in Wa and three Tibeto-Burman languages (Jingpho, Yi, Hani) spoken in Southwestern China and Myanmar. Other studies and languages with significant H1-H2 differences reported between registers include Huffman (1987) for Hmong, Traill and Jackson (1987) for !Xóõ, Cao and

Particularly relevant data are given in DiCanio (2009), who looks at register in the Takhian Thong dialect of Chong, a Mon-Khmer language spoken in Thailand and Cambodia. Two aspects of DiCanio’s study are highly germane to Burmese phonatory properties. First, acoustic and EGG analysis successfully detect a three-way contrast in Chong phonation (breathy, modal, creaky/tense) that is also reported to exist in Burmese. Secondly, DiCanio underlines the fact that while studies of pitch regularly look at temporally dynamic pitch changes represented as F0 contours, studies of phonation have rarely taken into account the possibility of temporally complex phonation patterns. This was required for Takhian Thong Chong because the language uses a Breathy-Tense contour register, a distinction that would be impossible to capture with a static measurement. To identify this complex phonation, DiCanio (2009) measured H1-H2 and H1-A3 at twelve evenly-spaced time indices over the course of the vowel and was able to clearly capture the complex Breathy-Tense change, but also found some temporal
alignment of phonation even in the “simple” registers. Regarding measures of spectral tilt, H1-H2 best distinguished the Tense from non-Tense registers and H1-A3 best distinguished the Breathy register from the non-Breathy registers. This split is potentially consequential to the analysis of Burmese in demonstrating that a three-way contrast may be reflected in a single phonetic quality (spectral energy), but still need reference to more than one metric of that quality. Also, the data from Chong support the generalization by Esposito that measures of higher harmonics (e.g. H1-A3) are more suited to capturing modes from the breathy end of the continuum of phonation types.

In acoustic phonetic studies of Burmese however, spectral tilt has not had even modest success in detecting phonation types. Watkins (1997) compares measures of the spectral profile in data from two speakers each of Wa and Burmese, less with the intent to detect and analyze specific phonatory settings and more so to test the suitability of spectral data to make claims about phonation in these languages. He concludes that higher frequency information is a suitable acoustic diagnostic for making a three-way distinction between phonation modes (modal, creaky, breathy) but not a four-way distinction aligning nicely with Burmese’s four tones. The lack of a four-way contrast is not particularly troubling as the glottal state associated with Checked tones is not necessarily anticipated in the voicing patterns of the vowel. Rather, the robustness of the data is problematic. Watkins conclusion concerns “higher frequency information” because comparative amplitudes of F4 were his most reliable diagnostic. Watkins found measures of the lower harmonics such as the amplitudes of H1, H2, H1-H2 to be inconclusive (as well as the amplitudes of F1, F2, and F3). For one speaker, F4 amplitudes were significantly different between Checked vs. Creaky vs. High/Low tones,
while the other speaker’s data showed significantly different F4 amplitude on just Checked syllables. Statistical tests showed differences between tones to be insignificant for both H1-H2 and H1-A1. Although H1-A1 did significantly distinguish Low from the other three tones, it did not distinguish between the allegedly breathy High and the expectedly creaky Creaky and Checked tones, representing opposing ends of the phonation spectrum. In summary, from an array of spectral measures, Watkins found only the amplitude of F4 to have a strong correlation to the reported phonation contrast, and just for one of two speakers. These are not robust results, a fact Watkins interprets as a failure of the diagnostic (spectral information) rather than a lack of truly creaky or breathy voice in the subjects’ speech.

A possible cause of this failure is specific power spectra that Watkins was measuring. In order to avoid the effect of both initial consonants and “the effects on phonation of final glottal consonants” (Watkins 1997: 5), Watkins inspected spectra “calculated from a window of 100ms near the middle of each vowel” (ibid.). In doing so, these measures may have precisely avoided the very effects of the weak or strong glottal closure that the study was trying to detect. What can be ascertained from Watkins (1997) is that he found multiple spectral measures (specifically the comparative amplitude of higher and lower harmonics) to be poor predictors of phonation at the vowel mid-point. This conclusion does not rule out either possibility put forth in the Four-way Register Hypothesis – it neither confirms that the phonation contrast is consistently present on the tone-bearing vowels nor does it conclusively refute the adequacy of the metrics used with acoustic data.

---

9 The study also consisted of a fairly small n as well (48 tokens per speaker, 12 per tone), while measures of harmonic amplitude have been shown to need relatively large n values.
Later studies have also encountered this difficulty with acoustic analysis. Gruber and Feizollahi (2006) measured H2-H1\textsuperscript{10} for each tone, from spectra taken at the vowel midpoint on five vowels, and met mixed results. Results indicating breathiness were nil. No consistent difference was found between High and Low tone syllables, and High tones often had the lowest H1-H2 value of all four tones – precisely opposite of the expected result should the values accurately reflect breathy voicing. Creaky and Checked syllables however were associated with a shallower spectral slope across all four speakers in the study for certain vowels\textsuperscript{11,12}. While these results are not contradictory to the Four-way Register Hypothesis, neither are they an endorsement. Gruber and Feizollahi (2006) conclude that pitch contour was a more robust correlate of tone than phonation type, but remain open to the possibility that H1-H2 is a poor metric for the Burmese-specific production of phonation qualities. Considered with the similar results in Watkins (1997), it is safe to say that H1-H2 and other measures of spectral tilt are indeed a poor metric when only the middle portion of syllables or vowels is examined. In Chapter Four, experiments for this dissertation show that tilt measures taken from spectra windowed at

\textsuperscript{10}Note the use of H2-H1, the inverse of H1-H2. Subsequent discussion of Gruber and Feizollahi (2006) converts all findings and values in this study to the inverse H1-H2 measures to remain consistent with the rest of the chapter.

\textsuperscript{11}For [e], three of four speakers produced Creaky and Checked syllables with significantly lower (p<0.05) H1-H2 values than High or Low, and significantly lower than Low syllables for the fourth speaker (but not High, which had even lower H1-H2 values). Another front vowel [i] saw similar results for just two speakers. One speaker produced [ma] tokens following the predicted order High > Low > Creaky, Checked. Other vowels revealed no particular pattern for any speaker.

\textsuperscript{12}Preferably, calculations of spectral slope involving the second harmonic should avoid vowels with low first formants as well as speakers with high F0s as the increased amplitude in the area of the first formant can obscure the spectral profile of the glottal source. Watkins (1997) measures only [a] syllables for this reason, noting that checked [a] will have slightly different resonances. The study in Gruber and Feizollahi (2006) looked at five vowels of varying height [i, e, a, o, u], but analyzed vowels separately, not pooled. Confusingly, of the five vowels, [e] produced the strongest results supporting the Four-way Register Hypothesis. As the interference from F1 on H2 in high vowels would apply equally to all tokens of that vowel regardless of tone, the spectral measurements would remain relative within a vowel. I.e., a creaky [e] might still be expected to have a stronger H2 than a breathy [e], leaving intact the scale of higher H2-H1. Still, it isn’t clear why [e] and [i] would produce more fitting results for G&F (2006) than other vowels.
the vowel offset are more reliable than those described above, but they are still less robust indicators of a phonation contrast than other acoustic and physiological measures.

Ladefoged, Maddieson, and Jackson (1988) discuss an inverse filtered set of Burmese data. Their treatment is brief and does not include quantified details of the glottal waveform, but it is noteworthy for its insight into why measures of harmonic amplitude may be inappropriate for the Burmese phonation contrast. Examining the filtered glottal waveform of each tone’s production, they found “no difference between smooth (High tone) and creaky phonations in the rising or falling slopes of the glottal pulse”. They attribute this negative finding to the overall greater amplitude of the specific creaky “sample” recorded for the experiment. However, given the reported intrinsic intensity difference between the tones (c.f. §2.2), this negative finding is likely not limited to their single recording, but bound to be replicated in the majority of cases, as the signal amplitude of Creaky and Checked syllables are expected to be greater. Citing data from Jalapa Mazatec in Kirk et al. (1993), H1-H2 values are asserted as an acceptable correlate of phonation type, just not for Burmese because comparisons of the first two harmonics are “not very reliable when the vowels to be compared… have different pitches” (LM&J 1988: 301). Ladefoged, Maddieson, and Jackson (1988) contend that measures of H1-F1 are likely preferable to H1-H2 for this reason, but their analysis points to a larger problem for all acoustic measures comparing harmonics. Namely, so many phonetic features systematically vary between the tones that have some acoustic consequence on the harmonic structure. Distinctive pitch values affect the position of H1, a distinctive vowel quality in the Checked tone alters the distribution of higher-band energy, and consistently different intensity levels alter the scale of the amplitude
differences. Even if other phonetic measures were to clearly indicate a strong phonation contrast between the tones in Burmese, it would hardly be surprising if the indirect (and highly obfuscated) method of comparing of harmonics failed to capture it.

2.5.5 Phonetic Studies: EGG and the Glottal Waveform

Javkin and Maddieson (1983) also examine inverse filtered glottal waveforms for each tone, presenting a complete study highlighted by a couple methodological points of interest. Their statistical analysis uses five measures (unreplicated elsewhere) of the filtered glottal cycle: the duration (1) and slope (2) of the rising portion of the cycle, the duration (3) and slope (4) of the falling portion of the cycle, and jitter (5), the amount of period to period variation. More importantly, Javkin and Maddieson choose to measure these factors over only the final twenty cycles of each sample, likely a more fruitful approach than looking at the middle of the vowel as other studies have unsatisfactorily discovered. Another noteworthy feature of the study is the stated intent of the analysis to determine the metric which best differentiates the group of High/Low tones from Creaky/Checked tones. This (somewhat prescient) decision to pursue a single divide in the data pool rather than a four-way contrast both simplifies the task of analysis and heeds the similarity in production of Creaky and Checked tones (noted extensively by Watkins 2005a, see below).

Results were mostly positive: Creaky and Checked tone glottal cycles had a statistically shorter duration and steeper slope for both the rising and falling segments of the cycle. The measures of the falling segment correspond to the velocity of the vocal fold closure, where a higher velocity closing (i.e. steep slope) is a key laryngeal
articulation composing creaky phonation (Childers and Krishnamurthy 1985, Michaud 2004). The findings strongly indicate at least a two-way contrast between the tighter/tenser tones and the more open ones. Despite the encouraging results, Javkin and Maddieson repeatedly lament that other phonetic differences produced with each tone (duration, amplitude, and F0 differences) created a serious confound for examination of the glottal cycle. Dynamic pitch differences affect the duration of each cycle and therefore the duration and slope of component branches. Greater amplitude signals create steeper slopes. The considerable bimodal duration difference\textsuperscript{13} entailed that the twenty measured glottal cycles represented about a final third of a Low or High sample, but nearly all of a Creaky or the entirety of a Checked sample. While the segment duration and slope measures confirmed the hypothesized contrast, measures of jitter were less reliable. Javkin and Maddieson attribute this to the interference of pitch movement creating more cycle-to-cycle variation than the irregular periodicity of creaky phonation that the measure intends to detect.

The previous two studies mentioned investigated the glottal waveform through inverse filtering, a process which filters out the estimated vocal tract resonances from the recorded acoustic pressure signal. Only two studies by Watkins (2005a, 2005b) directly measure the glottal source via electroglottography, and though both present limited data, they firmly recommend EGG and measures of Closed Quotient (or the inverse, Open Quotient) as an instrumental technique for analyzing phonation type in Burmese.

\textsuperscript{13} Recall that Javkin and Maddieson (1983) recorded citation form tone-bearing syllables in isolation, where High and Low tone syllables are known to be exceptionally long and Creaky syllables shorter than in medial contexts. For Javkin and Maddieson’s data, Creaky and Checked syllables were approximately only one-third the duration of High or Low syllables.
Watkins (2005a) pursues the idea that Creaky and Checked syllables invoke a similar glottal tightening. This is corroborated in data that show how acoustically similar the tones are, revealing similar pitch peaks, pitch contours, and syllable durations\(^{14}\) between the two, as well as a similar laryngeal configuration through the vowel. Laryngograph traces (see Watkins’ Fig. 3, pg.5) reveal a closed quotient (CQ) of roughly 60% near the end of either vowel. However, where the CQ steadily increases (from initial 50%) to this value through the entire Checked vowel, the Creaky tone has a CQ in line with High and Low tones (approx. 55%) until midway through the vowel before sharply rising, indicating a rapid increase in glottal tension potentially inducing or accompanying creak.

Subsequent research by Watkins (2005b) also includes EGG data and expands the stimuli to four carrier phrases per tone. Presented as CQ scores (calculated twice per syllable), the EGG readings from Watkins (2005b) hint at some of the effects of connected speech on the realization of laryngeal features. For instance, Creaky and Checked syllables had comparatively open glottal readings when another syllable followed, and a much higher CQ when phrase-final except after High syllables. It is possible that these variations not only support Thurgood’s (1978) generalization about the unreliable presence of breathy or creaky voicing features (but predictable alignment with High or Creaky tone respectively), but that further study could reveal a systematic behavior explaining the occurrence of more lax or constricted glottal states in Burmese speech.

\(^{14}\) Duration differences between the two are more pronounced in other studies (e.g. Gruber and Feizollahi 2006). Watkins’ (2005a) data aside, vowel durations are arguably tied to the most distinguishing impressionistic characteristic between the two, the rate at which the terminal glottal closure is articulated.
2.5.6 Phonetic Studies: Aerodynamic Analysis

No previous aerodynamic study of Burmese tone production exists\textsuperscript{15}, but Maddieson and Ladefoged (1985) do look for similar phonation properties in the oral airflow rates of the tense and lax registers of related and nearby languages. In Jingpho, Hani, and Wa, they found significantly higher rates of airflow accompanied the lax vowels than the tense vowels (i.e. average 112mL/s for \textit{lax} tokens, 98mL/s for \textit{tense} in Hani), a result they consider to indicate a less constricted glottis during lax vowel production. Conversely, a more constricted glottis with tenser vocal fold cover should impede airflow. This inference assumes that, given a consistent vocal tract shape (i.e. supralaryngeal shaping of airflow), rates of oral airflow measured at the mouth should roughly reflect airflow at the glottis. In a fourth language they examine, Yi, the correlation with higher airflow was found with only one out of three speakers, even though other metrics (H1-H2) indicated a true tense–lax distinction. Maddieson and Ladefoged attribute the unclear results to the higher intensity used with the Yi tense register – increased respiratory effort leading to greater supraglottal airflow leads to similarly high rates of airflow for both tense and lax voicing. The parallels to Burmese Creaky and Checked tones are obvious in the counter-action of a strong stress, with additional subglottal pressure, accompanying a phonation type that entails glottal impedance, and thus reduced airflow. Interpretation of airflow readings in Burmese for Creaky and Checked (and stressed High) syllables should be attentive to the role of intensity in shaping the aerodynamic waveform.

\textsuperscript{15} Bhaskararao and Ladefoged (1991) measured oral and nasal airflow in production of voiceless nasal stops, a well-reported feature of Burmese phonetics. Voiceless nasals are used phonemically in the language as onset consonants and may occur before a vowel bearing any of the tones. The interaction between a voiceless nasal onset and each tone represents a worthwhile topic for future research, but one that is corollary to the present study rather than central.
Before summarizing all of the accounts reviewed in this chapter, the table in (7) looks at solely at the studies that provided explicit arguments (be they impressionistic or experimental results) concerning phonation types produced with each tone. Spectrographic measures showed very little evidence of systematic phonation differences listed in the descriptive accounts, while measures of the glottal waveform met qualified success, finding a two-way contrast in very limited data.

<table>
<thead>
<tr>
<th>Tone : Phonation</th>
<th>Descriptive</th>
<th>Spectral Tilt</th>
<th>Glottal Waveform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low = modal</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>High = breathy</td>
<td>%</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Creaky = creaky</td>
<td>%</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Check = creakiest</td>
<td>n/a</td>
<td>×</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Symbol Legend

✓ association found
% assoc. not found consistently
✗ association not found
n/a: author does not discuss

Diagnostics of the glottal waveform in Javkin and Maddieson (1983) and the EGG signal in Watkins (2005a, b) successfully captured a two-way distinction between the tones. High and Low tones were distinct from one another in none of these studies, but were always distinct from Creaky (and Checked) tones. Since it was not clear whether the results indicated that High and Low tones employed modal or breathy voicing, the chart in (7) marks both jointly as ✓, as they successfully revealed an association with a more lax phonation.
2.6 Summary of Phonetic Literature

This chapter has related the current state of knowledge regarding the phonetics of the lexical tone contrast in Burmese. Despite the multitude and variety of studies investigating tonal phenomena, most evidence of the tones’ production exists only for citation form utterances. Section 2.3 revealed uncertainties about the duration of Creaky tone syllables, which were more similar to Checked syllable length in isolation but nearer the High and Low syllables when embedded in a phrase. Section 2.4 showed the relatively clear picture available concerning pitch peaks and contours associated with each tone, but within just one controlled context. Sandhi processes and the effects of stress or intonation on the phonetic qualities of each tone are scantly covered in the literature (and not at all outside of Sprigg 1977 and comments throughout various works by Watkins). Questions about phonation type are a current topic of research, but little has been resolved to this point regarding how reliably the Burmese High tone can be linked to breathy phonation or Creaky tone to creaky phonation, either in production or in the perception of native Burmese speakers. These inconsistencies between the qualitative and quantitative data suggest that prior phonetic experiments have somehow examined tone production inadequately, an understanding for which the review above points to a two possible causes:

- Neutralization of phonation distinction in the examined contexts
- Acoustic measures of phonation problematically taken at vowel mid-point

First, carrier phrases utilized in the experiments may have provided carefully controlled phonetic contexts, but detrimentally, the embedded positions chosen possibly neutralized the phonation distinction. Secondly, measures of phonation looking at the vowel mid-
point were off target, as all descriptive and acoustic evidence points to the vowel offset as
the locus of phonation effects. Working from these observations, a few requisite design
features are apparent for any further phonetic clarification: more temporally precise
measures, data embedded in phonetic contexts more likely to elicit phonation differences,
and other streams of data capable of more directly reflecting the glottal state – both EGG
and airflow recordings. Chapter Three describes the collection and testing of these data
and Chapter Four gives the results of these tests. Prior to the experimental methods and
results, the remainder of this chapter surveys the phonological literature regarding the
tone contrast and its development.

Many of the “problems” addressed in Chapters Two, Three, and Four revolve
around a lack of clear phonetic data for accurate description of the Burmese lexical tone
contrast, a central obstacle to phonological analysis of the language’s tonal system. In
addition to clarifying the role of laryngeal settings associated with each tone, the other
major problem this dissertation sets out to resolve is the distribution of tone to syllable
type when compared to that of vowel quality to syllable type. If the prior part of this
chapter reviewed questions in the literature that set the stage for the experiments and
results given in Chapters Three and Four, then the remainder of this chapter frames the
main phonological problems for analysis that are tackled by a finer look at select
production data (Chapter Five), a perception experiment (Chapter Six), and ultimately, a
phonological account (Chapter Seven) of the underlying contrast and the contextual
forms found in Chapter Four.

Section 2.7 discusses the historical interaction of pitch and phonation qualities in
Southeast Asia and Section 2.8 reviews the arguments of the few phonological analyses
of Burmese with an eye toward their featural representation of the contrast. The second major phonological issue, the complex distribution patterns of vowel quality, syllable codas, and tone, is presented in the context of Green (2005), an Optimality Theoretic analysis of Burmese metrical structure. Green’s analysis is introduced in §2.8.4, though assessed more fully in §5.2. when data can be brought to bear on critical issues.

2.7 Tonogenesis

The diachronic influence of phonation properties on pitch is well-established in the literature (Matisoff 1973, Diffloth 1989, Thongkum 1990, Thurgood 1993, 2002 – all of which address such phenomena solely in SE Asian languages, and Yip 2002), but many of the details of this transfer are less clear. It has been noted that the same laryngeal feature can correspond to directly opposing tone reflexes across languages – final glottal (or breathy) proveniences have been connected to a rising tone in many languages, but a falling tone in others (Mazaudon 1977, Yip 1982, Kingston 1985, 2005, Mithun 1999).

Kingston (2005) offers a striking insight in his account of divergent tonogenesis in the Athabaskan language family, attributing opposing reflexes to superficially similar but physiologically different articulations of the source laryngeal segment. Kingston argues that two groups of Athabaskan languages employed different articulatory strategies for achieving the same proto-language glottalic segment. Specifically, he attributes the High tone reflex of Chipewyan and nearby languages to coordinated contraction of the cricothyroid and thyroarytenoid muscles. Both are adductive gestures which together create a tense voice quality, such that stiffened vocal folds create airflow impedance and increase the rate of vibration. On the other hand, thyroarytenoid
contraction by itself typically increases F0, but when applied strongly enough can limit the vibration to the vocal fold cover and “lower F0 extremely” (Kingston 2005: 153). Kingston contends that Gwich’in and other Athabaskan languages, distinct from the Chipewyan set, recruited this articulation for glottalic coda segments, and in turn produced the preceding vowels with a lower F0. In this manner, divergent articulatory paths to glottal closure yielded different F0 by-products for phonologization.

In East and Southeast Asian tonogenesis, glottalization has generally prompted high tone reflexes (Matisoff 1973). A final glottal stop has conditioned a rising F0 on the preceding vowel nucleus for the rising tone in the Tibeto-Burman language Lahu (Matisoff 1970), the Chinese rising tone (Egerod 1971), and the rising tones of Vietnamese (Haudricourt’s (1954) seminal work). In a review of tonogenesis in the Tibeto-Burman language family, Mazaudon (1977) discusses the case of Hayu, where final obstruents with an accompanying glottal occlusion have developed a rising tone.

Mazaudon (1977) extends the pattern to Burmese, stating that the final [ʔ] of Checked syllables conditioned a rising high tone. Similarly to Kingston (2005), she also explains divergent F0 patterns with articulatorily different types of glottalization as their source. In contrast to the pitch rise before the *sharp* glottal closure of [ʔ], Mazaudon contends that the *gradual* glottal closure of the Creaky tone has led to a falling contour. However, a wide-range of phonetic studies contradict the claim that Checked tones are produced with a rising or even high tone (Javkin and Maddieson 1983, Watkins 2005a, §4.3 of this study17).

---

17 Figures 4.11 – 4.19 show that not only did Checked syllables consistently have a falling pitch contour, but that they had nearly identical rates of F0 decline as Creaky tone syllables. Creaky tokens fell further on average due to the longer vowel duration on which they could achieve the fall.
Turning to the evidence for tonogenesis in Burmese, it is not clear at which stage the consistent pitch perturbations became phonologized, but Maran (1971) considers the development of contrastive pitch to be a fairly recent event in Burmese. It is generally understood that phonological pitch evolved on top of at least three existing three proto-tones which already bore creaky and breathy voicing distinctions and contrastive finals (Maran 1971, Thurgood 1978). Bradley (1982) proposes a different course however, with Tibeto-Burman proto-tones developing phonation contrasts via contact with Mon peoples and more Austro-Asiatic register systems. Maran (1969) offers extraordinary observational evidence from non-standard dialects which contradict Bradley’s theory. The Northern (and to a lesser degree, Mandalay) dialects that Maran describes possess a more conservative tonal phonology, retaining laryngeal coda segments and realizing tone features redundantly. Specifically, High tone lexical items use a final [-h], Creaky lexical items a final [-Ɂ], and Checked items a final obstruent. Low tone lexical items occur in open syllables in these dialects.

2.8 Prior Phonological Accounts

This section surveys four phonological descriptions of tone-related phenomena in Burmese. Bradley (1982) describes the contrast as register (§2.8.1), Yip (1995) as a tone contrast (§2.8.2), and Lee (2007) as a cross-classifying mixed tone-register system (§2.8.3). Finally, Green (2005) abstracts away from the tone contrast in his discussion of the Burmese metrical system, treating High, Low, and Creaky as three contrasting suprasegments (§2.8.4).
2.8.1 Bradley (1982): “Register in Burmese”

Bradley explicitly opts for the term *register* over *tone* on the following grounds:

“...In terms of the orthographic and terminological tradition, (register) seems best; almost all of the acoustic parameters involved show characteristics that fit as secondary results of a register contrast, but not of a pitch-based tone system.” (Bradley 1982: 122)

He frames his discussion as a case of a Tibeto-Burman language with a suprasegmental contrast that resembles Mon-Khmer register systems, where two or three-way contrasts are common between suprasegmental qualities which reflect a bundle of interrelated phonetic qualities related to the configuration of the larynx, tongue, and pharyngeal cavity (Gregerson 1976, Huffman 1976). The table in (8) provides a typical division of the qualities. Bradley further speculates that the development of a Proto-Burmo-Lololish creaky tone and significant contact and assimilation with Mon culture triggered a global switch of a tonal system to one of register.

(8)\(^\text{18}\) Chest vs. Head Registers (adapted from Henderson 1952, Matisoff 1973)

<table>
<thead>
<tr>
<th>low register, also “chest” or “sepulchral”(^\text{19})</th>
<th>default, or absence of register</th>
<th>high register, also “head”</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ ATR</td>
<td>0 ATR</td>
<td>- ATR</td>
</tr>
<tr>
<td>lower pitch</td>
<td>normal</td>
<td>higher pitch</td>
</tr>
<tr>
<td>breathy phonation</td>
<td>modal</td>
<td>creaky/tense phonation</td>
</tr>
<tr>
<td>higher, more closed vowel set</td>
<td>fronter, more open vowel set</td>
<td></td>
</tr>
<tr>
<td>longer vowels</td>
<td>shorter, diphthongized Vs</td>
<td></td>
</tr>
<tr>
<td>[-h] coda</td>
<td>[ʔ] coda</td>
<td></td>
</tr>
<tr>
<td>earlier voiced obstruent onset</td>
<td>earlier voiceless obs. onset</td>
<td></td>
</tr>
</tbody>
</table>

\(^{18}\) The features listed here represent a general grouping of phonetic qualities rather than any single register language. A given register language will likely utilize not all, but some subset of the features listed in (8).

\(^{19}\) The term “chest” register is found the literature on voice quality (Laver 1980 for one). “Sepulchral” is used by Henderson (1952: 157) regarding Khmer.
It is not surprising that registers, which so intimately correspond to laryngeal settings, should also attest distinct pitch patterns. The presumption in Bradley then is that a register-based model for Burmese could capture the systematic pitch distinctions in the language. However, the specifics of Burmese tone (and in particular, pitch) align very poorly with the register prototypes in (8). Primarily, the High tone is mismatched with the chest register, which is problematically associated with a lower F0.

Following Gregerson (1976), Bradley relates a phonetic difference in the tone inventory not found in other descriptions – one based on an Advanced Tongue Root scale tied to different modes of phonation. The Creaky tone being [+ ATR] with creaky voicing, Low as [0 ATR] with modal phonation, and the High tone as [- ATR] and with breathy voicing (Checked is not a distinct register in Bradley’s system, but simply a closed syllable20). Bradley’s ATR values are assigned on the basis of small vowel quality differences (specifically, those reported in Thein Tun 1982) between the tones. Interestingly, they seem to contradict an understanding that ties some degree of pharyngeal constriction to constriction of the glottis. Video laryngoscopy studies discussed in Esling et al (1998), Edmondson et al. (2001, 2004), Edmondson and Esling (2006) have highlighted the role of supralaryngeal shaping in some phonation contrasts, in particular the action Edmondson and Esling (2006) label epigloto-pharyngeal constriction. Along with a retracted tongue root, a [+ATR] setting entails an enlarged

---

20 The treatment of Checked syllables raise another difficulty with Bradley’s ATR continuum – it does not account for the strong phonetic resemblance of Checked and Creaky syllables. This is the case whether Checked vowels are regarded as underlyingly distinct or as closed-syllable counterparts of the Creaky register. If the latter, one would expect vowel qualities to be fairly stable between the two, sharing the [+ATR] quality identified with the Creaky tone. Contra this prediction, the vowel quality split described in §1.2.5 suggests the opposite — Checked syllables are more generally associated with -ATR vowels.
pharyngeal cavity and lowered larynx, both characteristic of breathy phonation rather than the creaky phonation associated with Creaky tone vowels.


Yip (1995) focuses on the pitch distinctions between the tones. Bearing in mind that her brief analysis is admittedly speculative and given for illustrative purposes, Yip’s proposal (in (9)) offers a valuable point of comparison.

(9) Distribution of H, M, L autosegments on each tone (Yip 1995)

<table>
<thead>
<tr>
<th>high</th>
<th>low</th>
<th>creaky</th>
<th>checked</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>OR</td>
<td>L</td>
<td>H</td>
</tr>
<tr>
<td>σ</td>
<td>σ</td>
<td>σ</td>
<td>σ</td>
</tr>
</tbody>
</table>

First, Checked syllables are set aside from the tone system, treated as simply a closed syllable. It is not clear whether they are regarded as the closed-syllable counterparts of any tone or strictly the Creaky tone, nor is there discussion of how consistent pitch contours are realized with Checked syllables. Significantly, the analysis in (9) requires three phonological tone distinctions: H, M, and L autosegments. These autosegments are associated to syllables and multiple tones may link to a single syllable to form contour tones. The High tone bears either an M tone or an HM contour. With either alternative, the contrast with the HL contour of Creaky syllables is noteworthy for its attempt to capture minor pitch distinctions in the phonological specification, whereas other proposals regard the pitch peak of High and Creaky tones as phonologically the same (i.e., both associate with an H).

Lexical specification of features is handled straightforwardly in Lee (2007). High tone syllables bear an H tone, Low tone syllables an L tone, Creaky tone syllables have a lexically creaky vowel, and Checked syllables a [ʔ] coda. Assignment of an H tone to Creaky and Checked syllables is triggered by their respective forms of glottal constriction. Likewise, their falling contour is an effect of the same glottalization, though it is argued that this fall is strictly phonetic and does not warrant the insertion or linking of an L tone.

(10) H is not lexically specified on Creaky, Checked syllables (Lee 2007)

<table>
<thead>
<tr>
<th>high</th>
<th>low</th>
<th>creaky</th>
<th>checked</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>L</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>CV</td>
<td>CV</td>
<td>CV?</td>
<td></td>
</tr>
</tbody>
</table>

Lee contends that the association of an H tone with Creaky and Checked syllables is pressured by two distinct markedness constraints militating against an L tone on a Creaky vowel or on a syllable with a [ʔ] coda, *ʔ/L and *V̰/L. To avoid violation of *ʔ/L, an L tone on a Checked input /kʰaʔ/ becomes H, as in the tableau in (11).

(11) Tableau. Neutralization to [H] in Checked syllables (Lee 2007)

<table>
<thead>
<tr>
<th>/kʰaʔ/</th>
<th>SPECIFY-T</th>
<th>*ʔ/L</th>
<th>IDENT-T</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td></td>
<td>!</td>
</tr>
</tbody>
</table>

21 This is also the distribution of lexical features for each tone found in Green (2005). Green’s description of the contrast ends there however, as he focuses on the structure of the various syllable types rather than the realization of the tones.
This explanation is counter-intuitive in two ways. First, it is not clear why two separate constraints, *?/L and *V/L, are needed to produce what is the same effect: the dispreference for a single segment or syllable to contain both glottal constriction and low pitch. Secondly, glottalization is paradoxically argued to both raise and lower the pitch on these forms. It phonologically induces raising with the insertion of an H tone under pressure from the *?/L and *V/L constraints, while it phonetically lowers the pitch at the end of the syllable in order to explain the falling contour also found on these lexical tones.

2.8.4.1 Green (2005): Word, Foot, and Syllable Structure in Burmese

Green (2005) provides an explanatory account in Optimality Theory terms of the prosodic structure of Burmese, and thus, only deals with tone peripherally. Lexical tone is treated as a single quality that can be associated with any vowel, permitting Green to abstract away from the pitch, phonation, duration, and intensity of each tone. The central claim of his analysis is that only footed syllables can bear tone and that only certain syllables in Burmese may be footed. Therefore, his analysis is highly relevant to the present discussion on the grounds that any analysis of the tonal inventory of Burmese should take into account their variations on syllable structure.

One goal of Green (2005) is to explain the paradox in §1.2.4 – that is, why are diphthongs restricted to closed syllables? And why do open syllable vowels with nasalization behave the same as vowels in syllables closed by a glottal stop? At the same time, Green intends to show how his solution is tied to the behavior of minor, or schwa, syllables and the restriction of only glottal stop and placeless nasals in coda position. In this way, one “tonal” property is directly accounted for – the [?] coda of Checked tones.
This is mostly done with a reinterpretation of the Coda Condition (Itô 1989), which bans all place features from the right edge of syllable, whether on a consonant or a vowel.

2.8.4.2 Meter

Metrically, Green divides Burmese words into those with one foot, monopodic words, and those with more than one, polypodic words. Polypodic words are generally compound\(^{22}\) forms and are multi-syllabic, built from multiple morpheme constituents. A monopodic word has one final, heavy syllable which is footed, and all preceding syllables are minor, containing an /ə/ nucleus and other simplified consonantal features, either underlyingly or through reduction. Polypodic words contain multiple footed syllables as well as interspersed unfooted, reduced syllables, but notably the final syllable must still always be heavy. A vowel is underlyingly specified for High, Low, or Creaky tone. Stress is assigned to the output by \textsc{eval} and therefore, which tone actually surfaces (for Checked tone or due to phonological processes influencing tone, including a variety of sandhi rules) and which syllable with tone reduces to [ə] is assigned at the output level as well.

To describe the reduction to minor syllables in multi-syllabic, monopodic words\(^{23}\), Green holds that only the rightmost syllable is footed, and that unfooted syllables are not licensed for a range of phonetic features (see description in §1.2.3.6). The rightmost footing is derived from two undominated constraints, \textsc{align}-L(\textit{f}, \sigma)` and \textsc{align}-R(\textit{f}, \sigma)`

\(^{22}\) Compounding is prevalent throughout Burmese morphology and compounds are frequently monopodic as non-final morphemes are often reduced to minor syllables. Therefore, a word with multiple feet is always a compound, but a compound word need not contain multiple feet.

\(^{23}\) The exceptions to reduction are the polypodic words: compounds, borrowings, and rare “superlong” words. These forms exhibit that some prosodic structure must be stored lexically and that the relevant I-O constraints are highly ranked in order to maintain them in the output of these instances.
and a highly-ranked ALIGN-R(⟨f, ω⟩). The first two constraints require both the left and right edge of each foot to be aligned with a foot-head. This is only possible if the foot-head is the only syllable in the foot—thus all feet are monosyllabic. Because of ALIGN-Ft-R, only one such monosyllabic foot occurs on the right edge of the prosodic word. Unfooted syllables are not under pressure to have tone by a high-ranked markedness constraint, FOOTSALIENCE, which requires a foot to be associated with a tone. Furthermore, being unfooted forces a syllable to be monomoraic because of the ranking FOOTBINARITY >> NOHEAVY, which states that nuclei are preferably monomoraic but more importantly, a foot must be bimoraic. Footed syllables must be bimoraic, while unfooted syllables cannot violate NOHEAVY and are accordingly monomoraic. Along with the restriction against right-edge PLACE features on vowels, monomoraic syllables are restricted to toneless schwa nuclei – they must lack vowel [PLACE] features and cannot bear tone or have a coda. Green also derives the reduction of complex onsets in Minor syllables, for which the reader is referred to Green (2005, §3.4) as the account is unrelated to the moraic alignment of Burmese syllables.
Chapter Three. Experimental Methodology

3.1 Overview and Hypotheses

The previous chapter reviewed the phonetic literature on Burmese tones and highlighted the outstanding issues in description. The present chapter casts these unresolved issues as a set of concrete hypotheses concerning each tone, which are then tested with an array of acoustic, EGG, and aerodynamic metrics, particularly those that speak to the glottal state. In Chapter Four, the results of these experiments are shared with an eye toward the dissertation’s main objective, an integrated phonological description of Burmese tone and tone-related phenomena.

For the data collection, ten native speakers of Burmese recorded sets of words that contrast both in tone and syllable structure by reading aloud a script of over 450 phrases presented in the Burmese orthography. The entire script was read through twice, once into a microphone with a simultaneous EGG recording and a second time while wearing a pneumotach air mask (§3.4). Analysis of the phrases looked at a single vowel, [a], bearing each tone in open and closed syllables embedded in eight carrier sentences (see §3.3 for details). The hypotheses tested with these multiple channels of data are given in H1 – H10 below and are proposed according to the associations gathered from Chapter Two. In the methodological descriptions that follow, particular attention is paid to the data collection and analysis of Hypotheses H1 – H4, which concern modes of phonation

H1. High tone syllables are associated with breathy phonation
H2. Low tone syllables are pronounced with regular, modal phonation
H3. Creaky tone syllables are associated with creaky phonation
H4. Checked syllables are also associated with creaky phonation
H5. Tone-bearing syllables are bimoraic and coda consonants bear moraic weight
H6. The offset F0 of a vowel is not different whether the following syllable bears a High or Low tone
H7. The onset F0 of a vowel is greater after a High tone than a Low
H8. The offset F0 of a vowel is lesser after a High tone than a Low
H9. Creaky and Checked tones in juncture have less glottalization and a smaller drop in F0
H10. Creaky voicing and a glottal stop coda induce a raised pitch

The remaining sections of this chapter explain the methodological decisions and procedures for the subject pool (§3.2), phonetic instruments (§3.3), data collection (§3.4), as well as the methods of analysis (§3.5). In the description of Chapter Two, there was a fair, though far from precise, sense of what phonetic properties are expected with each tone. A main goal of this dissertation is to resolve not just which properties are associated with each tone, but how robust these associations are, and which, if any, property is the primary indicator of the four way contrast. The bulk of the production data addressing this problem is presented in Chapter Four, but note that all of the production experiments in this dissertation (i.e., the tests of H5 – H10 in Chapter Five) examine the same set of recordings, the design, elicitation, and organization of which is described in this chapter. All subsequent quantitative analyses refer to the methodology detailed below for the collection of their respective data sets, whether acoustic, electroglottographic, or aerodynamic.
3.2 Methodology: Subjects

Twelve subjects were recruited and recorded for this study, but in most cases only data for ten (Speakers A through J) are presented. Two speakers were unable to complete the recording session due to issues with the equipment or task. These subjects will be called Speakers K and L and their data are fully discarded from the pooled data to appear in this or any subsequent chapter regarding production.\(^1\)

All of the subjects were United States residents, living in either the metropolitan Washington, D.C. (2 subjects) or Seattle (10) areas. They were compensated $100 for their participation. All were former residents of Myanmar, and had spent their childhood and some portion of their adult lives in Rangoon. Each subject identified Burmese as his/her first language and the language spoken by his/her parents in the family home in Rangoon. One speaker (E) declared native-like fluency in Japanese as well. The subject pool was regarded as representative of a single dialect group, which could specifically be called Rangoon Burmese but will generally be simplified to Burmese or Modern Burmese in these pages\(^2\).

The subjects were aged 19 – 65 years, with a majority between the ages of 40 and 50. The table in (1) provides a breakdown by age as well as other pertinent biographical information. Subjects were dividedly evenly by gender, six male and six female, however both discarded speakers were male, yielding an unequal tally for the analyzed data: four male and six female. For the most part, the factors in (1) do not shape the analysis in this chapter or later. With the limited subject pool, it would be difficult to separate the

\(^1\) Speaker K and L are included in the perception experiment in Chapter Six. Their responses to this task were uncompromised and used equally with the other subjects.

\(^2\) See the discussion in §1.3.1 regarding the nomenclature of modern Burmese dialects and the identification of the Rangoon/Mandalay dialect as Standard Burmese.
influence of a single factor from individual speaker differences. Biographical factors are discussed in a few instances when linked to noteworthy patterns or behavior.

(1) Biographical Data of Study Participants

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>Age Group</th>
<th>Years in the U.S.</th>
<th>Other language proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Female</td>
<td>61+</td>
<td>30+</td>
<td>Chinese</td>
</tr>
<tr>
<td>B</td>
<td>Female</td>
<td>41-50</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Female</td>
<td>41-50</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Female</td>
<td>41-50</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>Female</td>
<td>41-50</td>
<td>5</td>
<td>Japanese</td>
</tr>
<tr>
<td>F</td>
<td>Female</td>
<td>21-30</td>
<td>&lt; 1</td>
<td>French</td>
</tr>
<tr>
<td>G</td>
<td>Male</td>
<td>51-60</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>Male</td>
<td>51-60</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>Male</td>
<td>41-50</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>Male</td>
<td>31-40</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>Male</td>
<td>&lt; 20</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>Male</td>
<td>61+</td>
<td>30+</td>
<td></td>
</tr>
</tbody>
</table>

Recall that Subjects K and L were recorded but are not included in the analysis stage of this study. Speaker K's production during the aerodynamic portion of the task was unproblematic, but equipment failure rendered the majority of his simultaneous acoustic and EGG recording unusable. Speaker L had repeated difficulty maintaining consistent volume and pace, yielding results which this researcher felt were task-based and had little or nothing to do with the stimuli. These performance difficulties had a too confounding (and arbitrarily applied) influence on the data, therefore all production data for Speaker L was completely discarded.
3.3 Methodology: Stimuli

This section reports the content of the stimuli script, both listing and justifying the inclusion (or exclusion) of certain sounds. The basic format of the script was a set of tokens, the main object of measurement, embedded in eight carrier phrases that supplied phonetic contexts relevant to confirming Hypotheses (H1) – (H10) in a variety of contexts. The main tokens consisted of a set of [t]-initial syllables with different combinations of vowel and tone.

All together, prior studies have looked at tone-bearing syllables in just two environments: those spoken in isolation or between two Low tone words. The eight carrier phrases, given below in (1) – (8), also include these two frames (1, 2), but additionally furnish phrase-medial frames with different adjacent tone-bearing (3, 4) or toneless syllables (7, 8) and frames which place the variable token phrase-finally (5, 6).

Phrase (1) *Token in isolation*  

Phrase (2) Low __ Low  

<table>
<thead>
<tr>
<th>Romanization</th>
<th>IPA transcription</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>cundo__kou yei:-ne-de</td>
<td>[ʧәndә __ ko jeɪәde]</td>
<td><em>I am writing</em> ___</td>
</tr>
</tbody>
</table>

Phrase (3) High __ Low  

<table>
<thead>
<tr>
<th>Romanization</th>
<th>IPA transcription</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>hta’pyo:__-ba</td>
<td>[tʰa’pjό __ ba]</td>
<td><em>Please say</em> ___ <em>again.</em></td>
</tr>
</tbody>
</table>

Phrase (4) Low __ High  

<table>
<thead>
<tr>
<th>Romanization</th>
<th>IPA transcription</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>da__la:</td>
<td>[da __ lά]</td>
<td><em>Is this a</em> ___</td>
</tr>
</tbody>
</table>

---

3 Romanization follows the transliteration method applied in the glossary of Cornyn and Musgrave (1958).
Phrase (5) High __ #

\[ \begin{aligned}
\text{kou:} & \quad [\text{kó __}] \\
\text{Nine} & \quad __.
\end{aligned} \]

Phrase (6) Low __ #

\[ \begin{aligned}
\text{hou'ke, da} & \quad [\text{ho'kè da __}] \\
\text{Yes, it is} & \quad __.
\end{aligned} \]

Phrase (7) High __ Minor Syllable

\[ \begin{aligned}
\text{kou: bama-lou pyo:} & \quad [\text{kó bəməlo pjó}] \\
\text{Say “nine ___” in Burmese.}
\end{aligned} \]

Phrase (8) Low __ Minor Syllable

\[ \begin{aligned}
\text{da mahou'-hpu:} & \quad [\text{da məhoʔ pú}] \\
\text{No, it isn’t} & \quad __.
\end{aligned} \]

The carrier phrases in (1) – (8) were devised to maintain some meaning regardless of the embedded stimuli token and to control the tones or segments on words adjacent to this variable syllable. As the stimuli tokens were composed of nonce forms and various parts of speech, some naturalness was plainly compromised, but the task itself could arguably only be so natural. The process of repeating eight sentences with a single word altered more than sixty times into a microphone or mask is a task that already promotes fairly self-conscious speech. Natural spontaneous speech was unlikely even with the most natural carrier phrases. To allow the subjects to acclimate to the nature of the stimuli, they were permitted to review and practice reading all of the carrier phrases with an initial warm-up period during which equipment was calibrated and levels for each
speaker were checked by the researcher. After seeing and reading a few example sentences of a token in each carrier phrase, no subject had difficulty construing the intended readings. Subjects were encouraged to read the sentences quickly and naturally with a conscious pause between repetitions.

Phrases (1) and (2) importantly replicated the two contexts investigated in prior studies. As familiar as the contexts were, there was scant EGG and no prior aerodynamic data recorded for them. The EGG and airflow signals could be compared with the acoustic record to provide further insight into the relationship between the glottal state and acoustic measures of phonation (e.g. H2-H1), a point of interest for Burmese as well as more generally. Regardless of the additional channels of data, replication was a worthwhile endeavor in itself. Most importantly, Phrases (1) and (2) provide a point of comparison for each speaker with their productions in the other carrier phrases. For example, different predictions would be made for the High tone pitch contour in Phrases (3) – (8) were a speaker found to produce High tone syllables without a fall in isolation or rise in [Low __ Low] contexts (as indicated by the descriptions in §2.4). Lastly, isolated syllables in this study corresponded to what have been called the citation form productions of each tone. Consequently, Phrase (1) tokens are referred to throughout by both terms, in “isolation” and “citation form”, with the choice dependent on the discussion at hand (i.e. “isolated” when the absence of surrounding phonetic content is most pertinent).

Carrier Phrases (3) and (4) were included to provide additional phrase-medial contexts to strengthen the associations intimated by Phrase (2). Moreover, the two phrases looked at the effects of an adjacent High tone syllable, which presumably has an
explicit high pitch target that may influence pitch contours and the timing of peaks on neighboring syllables.

Inclusion of Phrases (5) and (6) was critical to many points of the analysis. Recall from Chapter Two that laryngealization (whether creaky phonation or complete closure in a glottal stop) was reported primarily in citation form, but was said to not occur in close juncture and was difficult to detect when embedded in sentences. Yet no instrumental study to date had confirmed or even tested this observation by comparing phrase-medial and phrase-final (or citation) utterances. Proper testing of Hypotheses H3 and H4 required a block of Creaky and Checked\(^4\) tone syllables in phrase-final position. As creaky phonation was predicted in this environment if anywhere in Burmese, a lack of creaky phonation found here could not easily be explained by other factors, and would suggest that creaky phonation is simply not produced in the modern language.

Phrase-final contexts also offer important insights into the production of the High and Low tones. Following Thurgood’s (1978) dual claims of (i) creaky neutralization in close juncture and (ii) the erratic realization of both creaky and breathy phonation, the possibility exists that the difficulty detecting breathy voice when embedded in sentences (Watkins 2005, Gruber and Feizollahi 2006) was not a result of inadequate measures, but a regular positional neutralization of breathy phonation. This possibility – that phrase-medial laryngeal neutralization extends to cases of underlying breathy voicing – was tested. Comparison of phrase-medial and phrase-final productions of High tone permit

\(^4\) There is little doubt that glottal stop codas on Checked syllables are lost in juncture, assimilating to the following onset. But the presence of creaky voice quality during the vowel is not clear, even when the glottal stop is present. Therefore, phrase-final Checked syllables are necessary to enable comparison between phonation qualities found with [CVʔ] and those with supra-laryngeal obstruent codas, such as the [CVk] output from Carrier Phrase (1).
one to test whether High tones are associated with breathy phonation in all positions, just phrase-finally, or never at all.

Additionally, examining phrase-final utterances of all four tones can address the different effects of phrase-position and connected speech – functioning to isolate the phonetic characteristics attributable to having no immediately following words from those of running connected speech. Phrase-medial and phrase-final contexts were also necessary for later tests (§5.4) to elicit CVN and C\~V instantiations of nasally-closed syllables.

Lastly, the minor syllables following the framed token in Carrier Phrases (7) and (8) refined the characterization of phrase-medial effects. Minor syllables in this position could test whether claims made for close juncture applied regardless of the adjoining phonetic material, or if some effects were limited to cases where the following syllable is a major, tone-bearing syllable. The potential effects are not limited to neutralization of phonation type, but concern all phonetic characteristics of the tones. Duration, intensity, and pitch are all potentially less influenced by the following weak syllable than a full syllable bearing a tone and its accompanying features. This possibility alludes to a major consideration for phonological analysis, whether the Low tone specifically targets regular pitch and modal phonation or simply realizes the phonetic default of null values for these features. The behavior of each tone when adjacent to Low tones (recorded in Carrier Phrases 2, 3, and 4) was compared to those preceding Minor syllables in phrases (7) and (8). If, for example, the effects on F0 caused by a following Low tone were distinct from F0 patterns preceding a Minor syllable, this would evidence a specified pitch value or
target for Low tones. No significant difference between pre-Low and pre-Minor tones would indicate that Low tone syllables lack a specified pitch target.

The eight carrier phrases, with the inserted stimuli tokens (described below), composed the entirety of production data investigated in this chapter and the analyses in the rest of the dissertation. Sprigg (1977) endeavored to catalog at least one of each possible tri-syllabic sequence, providing the full range of possible tri-grams that was not used here in order to limit the data set to a size reasonable for both analysis and for the participants.

In a complete session, fifty-one rimes were recorded in at least one phrase, though analysis was limited to [a] and the toneless [ə]. Burmese uses seven vowels that occur in CV syllables with three of the tones (excluding Checked) and seven counterpart vowels occurring with all four tones in CVO/CVN syllables. These total fifty-one with the schwa used in all Burmese unstressed syllables and an additional Checked rime\(^5\) added. Each rime was paired with a voiceless unaspirated plosive [t] onset to form syllables which often, but not always, resulted in a noun, verb, or other lexical item in the Burmese language. The analyzed set of tokens consisted of the seven possible combinations of tone (four) and syllable structure (two: CV, CVC) with the vowel [a] as well as reduced, minor syllables using [ə]. This subset, shown in (2) below, provided eight iterations each for Phrases (1) – (8).

---

\(^5\) Recall that CV and CVN syllables can each realize seven vowel distinctions, while CV? syllables have eight, including both [ɛ] and [aɪ] instead of one or the other (c.f. (1.8)).
(2) Set of Stimuli Tokens: Eight Syllable Rhymes per Carrier Phrase

<table>
<thead>
<tr>
<th>Open Syllables</th>
<th>Closed Syllables</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ta/</td>
<td>/taN/</td>
</tr>
<tr>
<td>/tá/</td>
<td>/táN/</td>
</tr>
<tr>
<td>/tà/</td>
<td>/tàN/</td>
</tr>
<tr>
<td>/tɔ/(^6)</td>
<td>/tàʔ/</td>
</tr>
<tr>
<td>x 3 CV</td>
<td>x 3 CVN</td>
</tr>
</tbody>
</table>

Total: 4 open syllables 4 closed syllables

The carrier phrases above were formulated with the varied set of actual and nonce words in mind, being equally meaningful whether containing a noun, particle, or nonsense syllable in the variable position (e.g. “Say the word ‘bip’ again, please.”). Burmese orthography permits all of these combinations, with a unique composite character for each rime. For instance the characters \{.tc\} represents the syllable [tɪN], with an initial character \{t\} for onset [t] and a separate \{c\} representing the rime [ɪN]. Familiarity with the orthography prevented any confusion that might arise with the nonce forms regarding the intended vowel or tone\(^7\).

\(^6\) Production of Minor syllables in the embedded positions of each carrier phrase did not offer useful analysis of unstressed syllables, as speakers pronounced the syllable in these cases as though Creaky. In every phonetic sense under investigation, Minor syllable data in the embedded positions resembled Creaky tone data. Minor syllables produced elsewhere in the carrier phrases did not resemble Creaky tone syllables, but instead bore the anticipated qualities of Minor syllables: a very reduced duration, no high pitch peak, and a schwa vowel.

\(^7\) That said, there were a few instances when, it would seem, letters were simply misread. In one case, a subject read [tɔ] rather than [ta]. If identified at the time of the recording, the token was noted and cycled back to at the end of the sub-session and repeated with the proper vowel. If not identified until the data were processed by the researcher, the misread tokens were discarded.
The stop onset [t] was used with every rime to control for the effects of onset consonants on the following vowel. Specifically, an alveolar voiceless unaspirated plosive was chosen to minimize VOT, ease segmentation with a clear boundary between the stop’s silence and the vowel’s periodicity, and to perturb F0 at the vowel onset as little as possible. Voiceless aspirated stops are known to cause an initial rise in F0 on a following vowel and voiced stops suppress F0 (Lehiste and Peterson 1961, Hombert 1975). By using a voiceless, unaspirated [t] for all tokens, the characteristic influences of the initial consonant (formant resonances, VOT, and F0 perturbation) were kept as weak and consistent as possible across all stimuli.

Every sentence was read aloud three times iteratively under each condition (the microphone/EGG recording and airflow recording composing the two conditions for each subject). The third repetition was always discarded and only the first two readings were annotated and measured for analysis.

Stimuli were blocked by Carrier Phrase and Condition and syllables with all possible rimes were randomized within each block. Subjects were presented with 22 randomly ordered versions of a single Carrier Phrase (51 for Phrase 1) with a different embedded token in each. The order of Carrier Phrase blocks was also randomized. Every subject performed the microphone/EGG session first, before the reading a different randomization of the same stimuli while wearing the CV mask.

The tallies below in (3) represent an idealized count of recorded tokens. However, the tally of actual tokens available for analysis was slightly lower due to various causes of attrition during the process of recording, segmenting and analyzing (according to procedures detailed in §3.5) over 3,000 syllables.
(3) Ideal tally of all recorded tokens

<table>
<thead>
<tr>
<th>Phrase</th>
<th>Tokens</th>
<th>Syllables</th>
<th>Repetitions</th>
<th>Conditions</th>
<th>Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>960</td>
<td>Phrase (1) tokens</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>2,240</td>
<td>Phrase (2) – (8) tokens</td>
<td>8</td>
<td>2</td>
<td>2</td>
<td>10</td>
</tr>
</tbody>
</table>

320 x 7 Carrier Phrases

3,200 Total tokens

Recordings were occasionally cut-off or otherwise uninterpretable due to either speaker error or equipment failure. In some cases, acoustic records were available without a corresponding EGG track. Others were discarded because of subject readings errors that were unnoticed at the time of the recording. Finally, a few were simply lost for unknown reasons, most likely researcher errors such as mislabeled or unlabeled tokens which were not read correctly by automated analysis techniques.

3.4 Methodology: Equipment

All data were acquired using a DELL PC laptop connected to either the electroglottograph or airflow electronics interface. The stimuli script was presented on a separate laptop (Apple MacBook) as a randomly ordered PowerPoint slideshow that was advanced by keystroke controlled by the researcher. The acoustic pressure and EGG signal were obtained simultaneously with an EG2-PCX Electroglottograph data acquisition system (Glottal Enterprises, Inc., Syracuse, NY), the EGG electrode array as the channel B input and a Behringer ECM8000 omnidirectional condenser microphone as the channel A input. The equipment can be seen in the photographs of Figures 1 and 2 below.

---

8 Syllable counts shown in (2). Repetitions vary for Isolated tokens. Two conditions are the two readings through the script: (i) microphone/EGG and (ii) airflow.

9 All equipment (except Apple MacBook) sponsored by NSF Doctoral Dissertation Improvement Grant 06-605.
Recordings were made with the WaveView software program (Glottal Enterprises), and were digitized (both channels) at 44.1kHz and saved as stereo WAV files. All analysis of waveforms was performed in PRAAT 5.1.03 (Boersma and Weenink 2009).

Figure 1. The EG2-PCX Electroglottograph (Glottal Enterprises, Inc.)

Aerodynamic data were recorded in a separate session, using a similar configuration to that above. A Dell laptop recorded the aerodynamic signals using the AeroView software program (Glottal Enterprises, Inc.) while the speaker was presented
the randomized stimuli on a second laptop. Subjects wore a circumferentially-vented (CV) Rothenberg mask (Model S/T-1, Glottal Enterprises, Inc), which they held firm to their face during recording. Two pressure transducers were plugged into the mask: a wideband PT-2E (Glottal) transducer to the oral chamber and a mid-band PT-25 (Glottal) transducer to the nasal chamber. The transducers and system were calibrated before and after the practice session and during recording, the calibration was checked after every quarter of the script was completed (roughly 45 recordings). Channels from each transducer were acquired, digitized, and saved in AeroView through the MS-110 transducer electronics interface. The nasal signal was used as a guide to segmentation and for analysis in Chapter Five. The oral airflow signal was segmented and analyzed in Praat (Boersma and Weenink 2009) for duration and F0. Airflow rates in mL/s were measured in the WaveView software.

Sessions with either treatment (EGG or Aerodynamic) took between 60-90 minutes. Subjects were permitted to stop and take breaks whenever they pleased, but never did so within a session. The only regular pauses during a recording session were for the purposes of calibration or to otherwise verify the proper functioning of the equipment.

3.5 Methodology: Analysis, Phonetic Correlates of Phonological Tone

3.5.1 Overview

Once acquired from each speaker, the recordings were digitized and analyzed for nearly thirty different measures, described individually in sections 3.5.3 – 3.5.7 below. Before analysis, every recording was annotated in Praat 5.1.03 (Boersma and Weenink 2009) so
that each stimuli token and the two to zero adjacent words (zero in the case of Phrase (1) isolated tokens) were labeled and further segmented for vowel and coda boundaries in the case of closed syllables (§3.5.2 describes the segmentation criteria). Values for acoustic characteristics and open quotient of the EGG signal were then extracted from the annotated sound files using Praat scripts. Measures for airflow rates were taken manually using the WaveView program. Procedures for each phonetic quality (pitch, spectral tilt, …) are explained in a separate sections that states how measures were calculated, provides examples of any recordings that were problematic for measurement, and concludes with predictions for that metric as to the reported qualities of the tones.

3.5.2 Segmentation Procedure

Prior to explaining each metric, it is important to establish how the forms under analysis were selected from the continuous waveform of the carrier phrase. In some ways the following section is an account of how duration measures were taken, yet establishing segment boundaries entails the inclusion or exclusion of certain material in the analysis for every quality along with duration. In particular, criteria for the end boundary were crucial as the vowel offset exhibited some of their strongest and most interesting behavior in fundamental frequency, spectral characteristics, the glottal waveform, and rates of airflow.

Three types of syllables comprise the data: CV, CVN$^{10}$, and CVO. Except for the ‘O’ obstruent coda, calculations of duration, FO, harmonic, and closed quotient values used all three components: the VOT of the consonantal onset (C), the nucleus (V), and

---

$^{10}$ Results presented for the CVN data are limited, featuring primarily in the segmental duration measures of Chapter Five.
any sonorant coda (N). Checked syllable obstruent codas were excluded for reasons given shortly.

The beginning of the syllable was aligned with the beginning of the stop burst. Because every carrier phrase preceded the variable token with an open syllable and every token had a voiceless [t] onset, the stop closure and ensuing burst was always clearly defined in both the acoustic waveform and spectrogram. Voice onset time for the voiceless unaspirated stop onset was typically 6 – 10ms. For measures that were taken in deciles, this meant that the first decile, at 0%, fell during the burst and not during a periodic signal. Therefore, the 0% time-point almost uniformly has null values for F0, OQ, and harmonic amplitudes.

Marking the vowel closure was a less straightforward task, for both open and closed syllable types. The boundaries of Checked syllables did not include the [?] coda since it could not be evaluated consistently across the various phonetic contexts. The voiceless stop segment [?] was typically not identifiable in either the acoustic or aerodynamic record in phrase-final tokens, as the boundary was indistinct between the silence accompanying the consonantal glottal closure and the ensuing silence of utterance finality. Phrase-medially, Checked codas typically assimilated to the following onset. Affiliating this geminate consonant, either entirely or partially, with the Checked token was undesirable, not just for consistency across tokens but because F0, OQ, and harmonic measures during a stop consonant are not informative. Measurements taken during the stop closure would have created a misrepresentative contrast in the data when compared with vowel values. Segmentation of the nasal coda in CVN syllables was considerably simpler, the boundary between the nasal segment and following onset being clear in the
acoustic record in all cases except for the [m] onset employed in Carrier Phrase (8). Data for CVN syllables in phrase (8) were only considered for the analysis of duration in Chapter Five.

Turning to the end boundary of open CV syllables, many High and Low open syllables terminated with multiple cycles of periodic flow extending beyond any formant structure. The post-vocalic span was made up of weak, near sinusoidal cycles of slightly increasing frequency, often for 50ms or more. Conversely, Creaky or Checked vowels (and a few phrase-final High and Low) have an equally problematic span composed of a final few glottal pulses that are further and more irregularly spread out. Figures 3 and 4 below offer two such examples and demonstrate the criteria for final boundaries.

![Figure 3. Acoustic pressure and EGG waveforms and spectrogram of [ti], Low tone form. Figure illustrates additional cycles at vowel offset. The late breathy span not included in the vowel segmentation is shown in the shaded box.](image)

Segmentation of open syllables excluded this extra periodicity. For tokens such as that shown in Figure 3, measurements of the syllable and vowel durations were cut-off at the
point where the glottal waveform dies and the visible formants weaken. Oral airflow traces from similar tokens reveal that these extra cycles correspond with significantly greater rates of airflow, the increase of air passing through the glottis creating an excitation signaled by the simple, near sinusoidal, acoustic waveform seen at the end of Figure 3. As the periodicity lacked any spectral characteristics of the preceding vowel, these cycles were not counted as part of the vowel for any measures (duration, F0, etc…). This post-vocalic laryngeal excitation was not ignored however, as it was an important (and consistent) finding of the study accounted for in manners other than the acoustics of the vowel. Aerodynamic analysis of the timespan before and after the vowel closure addressed this issue (see §3.5.7 regarding airflow measures).

![Figure 4. Acoustic pressure and EGG waveforms and spectrogram of [t̥], Creaky tone form. Figure illustrates vowel terminal creakiness.](image)

Many Creaky syllable tokens, on the other hand, were heavily glottalized at the vowel terminus, marked with irregular periodicity. While posing a dilemma for measurement, the glottalization was a welcome finding in the acoustic waveform, in line
with expectations for creaky voicing as defined in Laver (1980), Edmondson and Esling (2006) and others. The various measures comprising the analysis should effectively capture this laryngeal quality. Increased time between pulses is reflected in a lowered F0 and weakened first harmonic (thus smaller differences between H1 and other harmonics). Longer periods of closure in the glottal cycle result in lower OQ scores. To select an endpoint for the vowel amid the late irregular pulses, a cut-off point was chosen at the end of the first pulse that was greater than a predetermined distance from the prior pulse. This distance was set at >40ms for male speakers and >25ms for female speakers\(^\text{11}\). The offset boundary criterion is demonstrated in Figure 5.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Acoustic waveform and spectrogram of Creaky [tɛ]. The shaded section demonstrates a cycle-to-cycle period under the 25ms threshold for female speakers. The following cycle far exceeds the threshold and was not segmented within the vowel. Including this final burst would dramatically extend the vowel duration (160.5ms to 205ms instead) and influence any acoustic readings taken from the later deciles.}
\end{figure}

\(^{11}\) Different thresholds were used to approximately accommodate pulses with a period three to four times that of regular vibration at the expected frequency. To illustrate, a male speaker produces a Creaky syllable with a fall to a low late F0 of 80Hz, yielding an expected period of 12.5ms. A pulse at three times this distance (37.5ms) is just under the threshold of 40ms.
This threshold excluded the occasional late pulse which was far enough removed from the fully periodic portion of the vowel so as to skew many of the key measures, and at the same time included most of the wider and irregularly spaced pulses which might best represent the constricted phonation sought in measurements. Checked vowel transitions into the glottal stop coda were segmented according to the same procedure described above, when applicable. In most cases however, this boundary was considerably clearer than found with Creaky vowel offsets.

This section has explained and justified with examples, the manner in which duration measurements were taken, a procedure relevant to the measures comprising the rest of this chapter.

3.5.3 Duration

One result of the segmentation procedure was that Checked tone durations in this chapter (and the ensuing analysis in Chapter Four) only concern the duration of the vowel. Claims about the duration of Checked syllables are reserved for the analysis in Chapter Five.

Regular differences in duration between the tones were expected, fitting the generalization formed from all prior accounts: High, Low > Creaky > Checked. In citation form, this contrastive pattern has been consistently found in prior work (with the possibility of an added distinction between High and Low tones). Comparison of the durations in Phrases (2) – (8) will test the accuracy of this generalization in seven sentence frames.
3.5.4 Fundamental Frequency

Measurements of fundamental frequency were taken at eleven evenly spaced points over the rime duration (deciles from 0% to 100%) using a modified version of a Praat script by Lennes (2003). The script also extracted F0_{MAX} and F0_{MIN}, the maximum and minimum F0 values and the positions of these extrema, T_{MAX} and T_{MIN} (time points within the span of the token’s DURATION). All calculations of F0 were made in Praat using an auto-correlation algorithm.

Inferential statistical tests looked at the following three measures of F0:

(4) F0_{MAX} The maximum F0 at any point in the rime.

(5) ΔF0 The change in F0 from an early to a late extreme, formulated as:

\[ \Delta F0 = (F0_{MIN} - F0_{MAX}) \left( \frac{(T_{MIN} - T_{MAX})}{(|T_{MIN} - T_{MAX}|)} \right) \]

(6) m_{F0} The slope of this change calculated as ΔF0 divided by the difference in milliseconds between the extrema, formulated as:

\[ m_{F0} = \Delta F0 / (|T_{MIN} - T_{MAX}|) \]

As with duration, the findings for F0 on tones pronounced in isolation or [Low__Low] contexts are well anticipated and would only be noteworthy if they diverged from the consensus seen in §2.4. The data in Carrier Phrases (2) – (8) offer the possibility of novel quantitative observations regarding F0 behavior in the tones.

Turning to the predicted results for each statistic, F0_{MAX}, by all previous accounts, is not expected to be a strong indicator of Tone. The peak-less Low tone should be distinct from the group of High, Creaky, and Checked tones, but there is little in the literature to indicate that these three should bear significantly different F0 maxima.
ΔF0 is formulated to have a negative value correspond to samples where the F0 maximum precedes the F0 minimum (i.e. a fall), and for the opposite order (rising pitch), a positive value. Low tones would therefore be expected to show a value near zero (little change over course of rime) and Creaky and Checked tones would have a decisively negative value, expressing a falling F0. Predictions for the High tone behavior are less clear, and are likely dependent on the carrier phrase. Note though that ΔF0 simply marks the difference in extrema of a contour and therefore fails to capture the distinction between two even pitch contours regardless of the disparity in their height (e.g. a flat 300Hz and flat 100Hz track would both yield a ΔF0 of 0). Nor does ΔF0 note how quickly the change in F0 occurs, a distinction captured by looking at the rate of change in F0 between the extrema, calculated in (6) as \( m_{F0} \). The simple comparison in Figure 6 demonstrates how sharply changing contours yield higher absolute \( m_{F0} \) values than evenly changing contours.

![Gradual vs. Abrupt F0 rise](image)

**Figure 6. Gradual vs. Abrupt F0 rises.** In either case, the ΔF0 is +100Hz. A steady rise from 100Hz to 200Hz rises 100Hz, as does a trajectory that is mostly flat with an abrupt rise to 200Hz only for the last few periods. On the other hand, \( m_{F0} \) would reveal a steeper incline for the dashed F0 track.
In particular, $m_{F0}$ was measured with the express intent of capturing a possible distinction between the gradual falls reported with High tones and the rapid F0 fall described on Creaky and Checked tones.

To summarize, each of the statistical values for F0 were tested for each tone in each carrier phrase. The predictions in (7) assume that the patterns reported for isolated citation form utterances hold for utterances in each controlled context.

\begin{enumerate}
\item[(7) a.] $F_{0\text{MIN}}$: Checked, Creaky, High > Low
\item[(b)] $\Delta F0$: Low > High, Creaky, Checked
\item[(c)] $m_{F0}$: Low > High > Creaky, Checked
\end{enumerate}

3.5.5 Spectral Tilt

While measures of the spectral profile have previously failed to provide meaningful results (§2.5), revisiting the diagnostic was worthwhile in order to ascertain if the puzzling conclusion reached in other studies was owing to spectral tilt being an insufficient measure of phonation in Burmese tones or because the location of the measures in prior studies (vowel middle, rather than end) was insufficient. In other words, was the measure wrong or the way it was taken wrong? If different measures of spectral tilt are found to still be inadequate, then perhaps some understanding can be gained of what phonetic properties of Burmese tone are obfuscating the data.

The spectral profile of each tone was evaluated with three comparisons of harmonic amplitudes, $H1-H2$, $H1-A1$, and $H1-A3$, was measured from spectra at multiple points in the vowel. The comparisons were defined in the following ways:
(8) H1-H2: difference in amplitudes of the 1st and 2nd harmonics

(9) H1-A1: difference in amplitudes of the 1st harmonic and the peak harmonic of F1, the 1st vowel formant

(10) H1-A3: difference in amplitudes of the 1st harmonic and the peak harmonic of F3, the 3rd vowel formant

FFT spectra over a 10ms Gaussian-like window were extracted at eleven evenly spaced points over the rime duration, and then interpreted using a modified version of a Praat script by DiCanio (2007) to give a set of amplitude values at each decile, 0% to 100%. First, the frequency of the harmonics was calculated as the F0 via auto-correlation algorithm and the integer multiples of the fundamental (i.e. H2 = F0 x 2). The formant frequencies, F1 and F3, were derived by LPC analysis, and taken as the mean of the filter coefficients over the decile interval. The script then found the maximum amplitude within +/-10% of each harmonic or formant frequency (i.e. the peak amplitude between 108Hz – 132Hz for a H1 at 120Hz). This procedure computed the dB values for H1, H2, A1, and A3 from which the differences were easily calculated. From this pool of data, five statistics were selected to capture the change in spectral profile over the course of the vowel.

(11) Difference at 80%: H1-H2_{80\%}, H1-A1_{80\%}, H1-A3_{80\%}

(12) Difference at 90%: H1-H2_{90\%}, H1-A1_{90\%}, H1-A3_{90\%}

(13) Mean of EARLY values (20-30%): H1-H2_{EARLY}, H1-A1_{EARLY}, H1-A3_{EARLY}

(14) Mean of LATE values (80-100%): H1-H2_{LATE}, H1-A1_{LATE}, H1-A3_{LATE}

(15) Change in EARLY and LATE mean differences: ΔH1-H2, ΔH1-A1, ΔH1-A3
The H1-H2, H1-A1, and H1-A3 values at the 80% and 90% deciles were selected for statistical evaluation, along with averaged early and late values, taken as the mean of the 10%-30% (EARLY) and 80% - 100%\textsuperscript{12} (LATE) percentile data points, respectively. A mean of early and late ranges was used instead of an individual percentile because many of the individual data points could not be defined by the automated pitch, formant, and spectral analysis detailed above. For example, out of 160 [a] vowel tokens in Speaker E’s data, 15 (9.4%) H1-A3 scores at the 80\textsuperscript{th} percentile were not available, along with 24 (15%) scores at the 90\textsuperscript{th} percentile\textsuperscript{13}. Therefore, a mean improved the likelihood of acquiring some valid reading for each token without significantly altering the data pool from the readings that would have been attained at a single percentile point. A test of H1-H2\textsubscript{LATE} values across all speakers found no significant difference between data points which were the mean of three values, means of two values, and those where only one percentile data point was available.

Finally, the statistic $\Delta$H1-H2 was found by subtracting H1-H2\textsubscript{EARLY} from H1-H2\textsubscript{LATE} to capture changes to the spectral profile over the duration of the vowel. $\Delta$H1-A1 and $\Delta$H1-A3 were calculated in the same manner. Similarly to $\Delta$F0, these measures of change do not reveal the magnitude of the differences being compared, meaning a $\Delta$ of 0 could as likely signify a vowel that was breathy from start to finish as one that was creaky from start to finish. Such differences were captured by any of the four static measures of spectral energy: H1-Hx\textsubscript{80\%}, H1-Hx\textsubscript{90\%}, H1-Hx\textsubscript{EARLY}, and H1-Hx\textsubscript{LATE}.

\textsuperscript{12} However, the 10% decile was scarcely used due to interference from the onset consonant. Automated spectral measurements at this point were erratic. On the other hand, automated measurements at the 100% decile were often $\mathcal{N}$A, but when defined were consistent adjacent values.

\textsuperscript{13} From the sets of 15 and 24 missing values, thirteen of each set were from the same thirteen tokens, all of which were Creaky or Checked tone tokens. In fact, for Speaker E only one Low or High tone syllable had an undefined H1-A3 value at either late percentile: a phrase-final Low tone [a].
For all three spectral tilt diagnostics, H1–H2, H1-A1, and H1-A3, breathy voicing is expected to produce larger (positive) differences than modal and creaky vowels, as higher harmonics have less energy when vocal fold tension is reduced as in breathy phonation (§2.5 for description). Conversely, during more constricted modes of phonation, the increased energy in the higher harmonics that results from greater vocal fold closing tension should lead to smaller (or possibly negative) differences between H1-H2, H1-A1, and H1-A3. For Δ scores of all three metrics, a positive value indicates that the difference between the harmonic amplitudes has increased from early to late in the vowel, reflecting movement towards phonation with a steeper spectral roll-off and less energy at higher frequencies. A negative Δ indicates that the late difference is smaller and the change during the vowel is towards phonation with a shallower spectral roll-off. It is worth bearing in mind that these values only signal change towards comparatively weaker or stronger differences and not specific phonatory settings. That is, positive ΔH1-A3 does not mean the vowel offset is breathy any more than a 10mph increase in one’s driving does not indicate speeding, merely a higher speed than before.

Confirmation of the set of Tone-to-Phonation Hypotheses H1 – H4 requires at least a three-way division in phonation behavior (breathy vs. modal vs. creaky) between the tones. For every statistic representing the static or dynamic spectral profile, the anticipated division follows the same ranking: one where higher values indicate more open or movement towards more open phonation. The hypothesized ranking for the four tones is displayed in (16):
High > Low > Creaky, Checked

Ranking supporting Hypotheses (H1) – (H4) for:

- H1-H2 at all percentiles of rime duration
- H1-A1 at all percentiles of rime duration
- H1-A3 at all percentiles of rime duration
- ΔH1-H2, ΔH1-A1, ΔH1-A3

3.5.6 Open Quotient: Analysis of the EGG waveform

3.5.6.1 Calculating Open Quotient

Open Quotient (OQ) is the ratio of the open phase of the glottal cycle to the period of that cycle. The open phase marks the time between the glottal opening instant and the subsequent closing instant, as determined by landmarks in the EGG waveform considered to represent these moments. The proportional duration of the open phase is not a direct measure of, but it is highly dependent on how open or unconstricted the glottis is during voicing. It is therefore a useful assessment of glottal activity, but not one that can simply be equated to phonatory settings (i.e. OQ 60% = breathy, OQ 50% = modal, OQ 40% = creaky, OQ 35% = tense). Interpretation of open quotient should be mindful that the values are blind to qualitative categories of phonation. That is, contrasts of breathy or creaky phonation entail a more complex set of acoustic and physiological correlates than just a relatively long or short open phase. Also, rather plainly, the open and close phase durations are relative to a speaker’s idiosyncratic production and anatomy. Heeding the above concerns, longer open phases can generally be assumed to correspond with wider glottal states, and therefore OQ was used in this study to test the phonation modes.
examined in Hypotheses H1 through H4. Syllables predicted to bear breathy phonation (Hypothesis 1) should have the highest OQ, while those bearing modal phonation (Hypothesis 2) should be lower but remain comparatively higher than more laryngeally constricted syllables (Hypotheses 3 and 4). One might also anticipate that OQ further distinguish the partial constriction of Creaky tone forms from the complete glottal closure denoting Checked tone codas.

EGG signals were recorded as the right channel of a dual-channel WAV file, allowing segmentation of words and vowels to be done concurrently with the segmentation of the acoustic waveform on the left channel (§3.5.2). Prior to analysis, all EGG waveforms were band-pass filtered from 50Hz to 500Hz. Individual tokens were then extracted from the phrases and saved as mono WAV files which were in turn processed by a Praat script developed by Gendrot (2005)\(^\text{14}\) which output a continuous trace of OQ per period. A separate script devised by the researcher then read the OQ trace and extracted values from eleven evenly spaced points in the rime duration, to match the decile measurements of other phonetic properties (i.e. pitch, spectral tilt).

Open Quotient was calculated from the DEGG, the derivative of the EGG signal. The precise moment that the vocal folds open or close is not aligned with a peak in the EGG signal, but rather is thought to correspond to the moment of peak upward or downward velocity. Taking the derivative of the EGG signal matches positive or negative peaks to maximum slope in the EGG, therefore corresponding better to the instant of vocal fold opening or closure (see stroboscopic alignment with EGG in Baer et al. 1983, 14 The actual script used was nominally different from Gendrot’s Calcul OQ script, but the method of determining the DEGG signal and detecting the DEGG negative maximum were unchanged. Gendrot’s script smoothed the resulting OQ trace, while the present script generated a trace with an isolated value representing each cycle (see trace in Fig. 9 below).
Hess and Ludwigs 2000, discussed by Henrich et al. 2004). Parallel EGG and DEGG signals are demonstrated in the two samples in Figures 7 and 8, respectively revealing modal and more constricted patterns of vibration.

**Figure 7. Longer open phase of modal voicing**

**Figure 8. Shorter open phase of creaky voicing.**

**Figures 7, 8.** Corresponding EGG (dashed line) and DEGG (solid line) signals showing three cycles extracted from vowels with modal (Fig 7) and creaky (Fig 8) voicing. Tokens recorded from Speaker F. Opening and Closing instants for one cycle are marked with open phase in light shading, closed phase in darker shading.

The open phase was defined as the span between opening instant₁ and closing instant₁. To compute the OQ for a given cycle, this duration was divided by the period (opening instant₁ – opening instant₂), as in (17).

\[
OQ = \frac{DUR_{OPENPHASE}}{T} = \frac{t_{Closing} - t_{Opening1}}{t_{Opening2} - t_{Opening1}}
\]

In the modal voiced example highlighted in Figure 7 above, the open phase is 48.98% of the cycle’s period, while the open phase in Figure 8’s constricted sample is visibly shorter, at 36.37%.
3.5.6.2 Common Errors in Automated DEGG Peak Detection

The output of this automated process contained numerous errors and empty datapoints, a consequence of inaccurate peak detection resulting from either (a) the automated detection algorithm or (b) EGG waveforms which were simply difficult to obtain reliable data from, even by close inspection by the researcher. Each of these instances is demonstrated below, followed by a description of the procedure for handling such tokens.

Figure 9. Open Quotient trace of Low tone [ta] spoken in isolation (Carrier Phrase 1) by Speaker F. Data points at the intermittent spikes in OQ, with values approaching the 82% peak, corresponded to glottal cycles with erroneously identified DEGG negative peaks.

Figure 9 is of a trace plotting the OQ value for each cycle of a low tone vowel. Frequent inconsistent values are apparent as “spikes” up to 80% and higher. One erroneously measured cycle is featured in Figure 10, which shows the precise cycle in the EGG waveform that yielded the circled measurement in Figure 9: 80.9% at time 171ms of the sample. The unexpectedly high OQ calculation is shown to result from a misidentified negative maximum (misidentified positive peaks are unlikely), which leads to an under-calculation of the period.
Figure 10. A single cycle of the EGG waveform at 0.170ms of same token shown in Fig. 9. OQ is calculated at 80.9% for this cycle due to the misidentified negative peak reflecting OpeningInstant₂. The probable OpeningInstant₂ is marked above. For this cycle, the open phase duration was correctly calculated but the overall period was rendered too short by a misidentified second opening instant. Thus, the resulting OQ ratio was too high. A manual correction, using the probable “opening instant₂” marked in the figure, yields an OQ of 0.482\(^\text{15}\), in-line with the adjacent OQ values in Figure 9, 0.447 before and 0.491 after. When possible, other errors in the OQ trace were manually corrected in the same manner. However, other erroneous datapoints could not be recalculated because the proper DEGG negative peak helping to define the period \(T\) was not apparent. In these cases, the value was changed to or left as \(NA\), as was done for the deciles 70\% - 90\% for the Creaky vowel in Figures 11 and 12. In this example, the first raised plateau of miscalculated OQ scores was manually corrected as the DEGG landmarks were apparent.

\(^{15}\) As in (17), OQ is defined as \(DUR_{\text{OPENPHASE}} / T\). \(OQ_{170\text{ms}} = DUR_{\text{OPENPHASE}} / T\) \(OQ_{170\text{ms}} = 1.88\text{ms} / 3.9\text{ms}\) \(OQ_{170\text{ms}} = 0.482\)
In the latter portion of the vowel though, a plausible opening instant from the DEGG (or EGG) waveforms is not identifiable.

![Image of EGG, DEGG, and OQ trace for a Phrase 1 (Isolated) utterance of creaky [g]. Open phase and other landmarks are indicated for the first misidentified cycle, where OQ is miscalculated as 76.9%. Later cycles had no identifiable opening instants in the DEGG signal.]

The two types of errors detailed above are only distinguished from one another by their outcomes. Both occur when negative maxima are not properly identified in the automated script, but one is correctable upon inspection and the other is not. A systematic procedure was followed to select likely automated errors for individual inspection. Consecutive data points with a climb of greater than 25% from cycle to cycle were hand-checked and corrected when a more evident closing instant was found in the DEGG. When no plausible minimum could be identified, the data was relabeled as NA. Further, two or
more consecutive datapoints matching this criteria were also identified as NA rather than recalculated\textsuperscript{16}. The final set of corrected OQ scores was then normalized per each speaker’s mean and standard deviation and tallied for the statistical measures.

3.5.6.3 Open Quotient Statistics and Predictions

As with the spectral tilt data, statistical evaluation of OQ looked at means of OQ datapoints, thereby increasing the likelihood of a usable value for each token, rather than only comparing OQ at a single decile (i.e. 80% or 90%) where NA values were common. Three means were chosen to represent the data: OQ\textsubscript{MEAN}, OQ\textsubscript{EARLY}, and OQ\textsubscript{LATE}. Additionally, ΔOQ expresses the difference between OQ\textsubscript{LATE} and OQ\textsubscript{EARLY} to capture any change in the open phase over the duration of the rime. The following four statistics were found from the decile values listed below:

\begin{align*}
(18) \quad & \text{OQ\textsubscript{MEAN} = Mean of all decile data points} \\
(19) \quad & \text{OQ\textsubscript{EARLY} = (OQ\textsubscript{10\%} + OQ\textsubscript{20\%} + OQ\textsubscript{30\%}) / 3} \textsuperscript{17} \\
(20) \quad & \text{OQ\textsubscript{LATE} = (OQ\textsubscript{80\%} + OQ\textsubscript{90\%} + OQ\textsubscript{100\%}) / 3} \\
(21) \quad & \text{ΔOQ = OQ\textsubscript{LATE} – OQ\textsubscript{EARLY}}
\end{align*}

In the statistical evaluation of each OQ mean, breathy phonation should correspond to more open glottal configurations and creaky phonation or laryngeal articulation immediately prior to a glottal stop should correspond to lower OQ scores. Therefore, each mean was tested against the hypothesized order High > Low > Creaky,

\textsuperscript{16} While eleven data points were technically given for each token, the precise calculation entailed that the 0% and 100% decile points were always NA values; these were never checked manually and left as NA. Similarly, the 10% decile point was more often than not returned as an NA value and was not corrected manually, ostensibly limiting all OQ tracks to eight decile data points representing 20-90% of the rime.

\textsuperscript{17} Since OQ\textsubscript{10\%} and OQ\textsubscript{100\%} were usually valued NA, the OQ\textsubscript{EARLY} and OQ\textsubscript{LATE} scores were effectively means of the 20-30% and 80-90% range, respectively.
Checked. However, it is not expected that $OQ_{\text{MEAN}}$ or $OQ_{\text{EARLY}}$ are particularly strong indicators of tone. For $\Delta OQ$, the hypothesized order is also High $>$ Low $>$ Creaky, Checked, as positive $\Delta OQ$ signifies a rise in the degree of glottal opening, while negative $\Delta OQ$ indicates tightening of the glottis over the course of the rime.

3.5.7 Airflow Measures.

Aerodynamic data were collected primarily to investigate rates of oral airflow, which were presumed to correlate to the degree of glottal abduction and adduction. In this study, only rates of oral airflow were quantified and analyzed; the nasal airflow signal was used only as a guide to segmentation of nasal segments. Once segmented, oral airflow signals were measured as the acoustic recordings were for F0 and Duration (§3.5.3 – 4) and included in the analysis of these qualities in Chapter Four.

Airflow measured at the mouth is not the same as airflow at the glottis due to possible obstructions as well as air lost through the nasal passage (even in oral vowels). Still, overall rates of airflow should have a straightforward articulatory counterpart amongst phonation types if comparisons across tones are made with data using the same vowel (avoiding vowels with low first formants) and therefore providing roughly the same vocal tract airflow impedance (Rothenberg 1977). That is, greater rates of airflow should correspond to a more lax or open glottal state and lesser airflow to a more closed glottis (as for lax vowels in Jingpho or Wa according to Maddieson and Ladefoged (1985)). This is because a wider glottis allows more air to escape while a narrower or tense glottis restricts flow. The correlations of course also assume equal levels of respiratory effort, which is known to not be the case for Creaky and Checked tones.
However, this possible confound was effectively circumvented by the array of airflow measures examined ((22) – (27) below), whereby certain measures were obscured by intrinsic intensity differences and others were not.

Measures of airflow were taken in milliliters/second from a low-pass filtered airflow waveform. Filtering was done in Praat at 50Hz in order to remove all acoustic energy from the signal, including the vibration of the speaker’s fundamental frequency (assumed to be > 50Hz for all subjects). In Chapter Four, statistical analysis compares tokens bearing a single vowel, [a], thus controlling for the effects of vocal tract shaping. Using WaveView, the filtered airflow trace was measured both at precise moments or over windows as the average rate of airflow. This study analyzed airflow rates at the following points of the syllable tokens:

(22)  \( \text{airflow}_{\text{WHOLE}} \): mean airflow rate in mL/s over whole syllable
(23)  \( \text{airflow}_{\text{EARLY}} \): mean airflow rate in mL/s over the 1\textsuperscript{st} half of the syllable
(24)  \( \text{airflow}_{\text{LATE}} \): mean airflow rate in mL/s over the 2\textsuperscript{nd} half of the syllable
(25)  \( \text{airflow}_{\text{PEAK}} \): maximum mL/s value detected during the vowel
(26)  \( \text{airflow}_{\text{MID}} \): mL/s flow at the syllable mid-point
(27)  \( \text{airflow}_{\text{POST}} \): mean airflow rate during period following vowel closure, described below

Without prior studies, it was not necessarily clear how these measures would reflect the phonation qualities found on Burmese tones. For example, the average airflow over the entire vowel, \( \text{airflow}_{\text{WHOLE}} \), was likely a less precise metric than just the later airflow,
A facet of the data that was immediately noticeable in the airflow waveform was a sharp rise in airflow at the vowel offset in High and Low tone tokens, along with a sustained heightened rate of flow at the close of the vowel’s periodicity. Usually, much of this late peak in airflow fell outside of the syllable as segmented for this study. The High tone token shown in Figure 13 demonstrates the late rise in airflow and the segmented post-vocalic span.

**Figure 13. Unfiltered oral airflow waveform illustrating post-vocalic 75ms span**
Token is High tone [tá] pronounced in isolation by Speaker I.

Even by casual observation, it was clear that only some High and Low syllables (and only those without a following syllable – Phrases 1, 5, 6) contained this additional flow and while Creaky or Checked tones never did. In order to capture this observed distinction, it
was decided to measure a portion of the waveform after the token of interest. A consistent period of seventy-five milliseconds was chosen to measure this span, beginning at the vowel offset, as determined by the loss of visible formant structure and a simplified airflow waveform devoid of vowel components (see also segmentation in §3.5.2). For tokens in phrase-medial contexts, following consonants frequently interrupted this span, and for non-medial tokens, inhalation/exhalation by the subject for non-linguistic purposes also rendered it not useful to measure the full 75 millisecond span. In these cases, an abbreviated duration was used, being as suitable a measure as a 75ms span for mL/s of oral airflow since the measure was a mean rate of flow rather than a total. Mathematically, a shorter or longer measured interval would not incur different averages – a too brief span, if projected to a full 75ms would retain the same average rate in mL/s. An articulatory effect is conceivable, but the data did no bear out any difference; abbreviated intervals did not have significantly different oral airflow rates than 75ms intervals.

For following consonants, the span was measured only up to the rise in airflow signaling the onset (either the consonant beginning or VOT burst for stops), leaving an interval between 0 and 74ms. Zero ms spans were discarded. Exhalation was treated identically to following onset consonants, by using an interval shorter than 75ms. Inhalation, on the other hand, created highly negative airflow rates, equivalent to 0 mL/sec flow for the purposes of this study. Rather than truncating the 75ms period at the point of inhalation, negative values were treated as though zero and mL/s was recalculated manually as:
Similar to the other measures investigating airflow, a high rate of $\text{mL/s}_\text{post-vocalic}$ is associated with breathiness. In this case, the higher airflow is not indicative of a breathy voiced vowel, but of syllable final extremely breathy phonation. A low $\text{mL/s}_\text{post-vocalic}$ rate corresponds to a mostly closed glottis. Additionally, the stop consonants following many Checked syllables correspond to flow rates near or at zero, as no airflow passes the vocal tract occlusion.
Chapter Four. Results of Production Experiments

4.1 Introduction

This chapter presents the results from the acoustic, electroglottographic, and airflow recordings, primarily as they pertain to detecting phonation types and laryngeal articulations during tone production. Between-tone differences in Vowel Duration and Fundamental Frequency are examined first, followed by the results for the multiple modes of data collected from each speaker – measures of harmonic amplitude, open quotient (OQ) values over the duration of the vowel, and rates of airflow, particularly at the vowel offset. These data are used to test Hypotheses One to Four (H1 – H4) from Chapter Three concerning distinctive phonation types.

The preview data in Chapter One (Figures 1.1–1.4) demonstrated that the hypothesized associations between tone and phonation type can be found in at least some instances. In these examples, a High tone syllable bore characteristics of breathy voicing, such as a higher rate of airflow, and a Creaky tone syllable likewise clearly exhibited creaky voicing with irregular glottal pulses visible in both the acoustic record and EGG waveform. However, the data collected for this study reveal that these signs of breathy and creaky voicing were often not apparent in Burmese speech. Discussion in Chapter Two offered a few explanations why non-modal phonation types might not be produced or detected:

i. Distinctive phonation types are realized at the vowel terminus. Metrics which examine vowel mid-points will reveal an inconsistent association or potentially none at all.

ii. Distinctive phonation types are neutralized in junction or are simply not always realized (both claimed in Thurgood 1978).
iii. Typical metrics of creaky and breathy phonation are inappropriate or unreliable for detecting the distinctive phonation types produced with Burmese High or Creaky tones.

The data presented in this chapter support the first two arguments above and refute the possibility in (iii). Numerous significant differences denoting phonation were found, though mainly with measures taken from the latter portion of the vowel. Different sentential contexts also revealed a distinction in the data set, primarily between contexts where the variable token was phrase-medial and those where it occurred phrase-finally, regardless of which tone adjacent syllables bore. In other words, [L __ L] tokens behaved similarly to those in the frame [H __ L], but quite differently from [L __ #] embedded tokens. This finding strongly supports (ii) above in that phonation type cues were context dependent, being reduced in close juncture with a following syllable in the controlled phrase-medial positions.

There were also noticeable effects of tone on measures of Vowel Duration and Fundamental Frequency as well as strong interactions between these measures and Carrier Phrase. In general, these interactions mirrored the patterns found with measures of phonation – effects were muted or fully neutralized in phrase-medial tokens. Tone-bearing syllables with a closely following syllable adjoined in fast, connected speech were more likely to be roughly similar, regardless of tone, in levels of duration, glottal aperture, and (to a lesser extent) movement in F0.

For analysis, data was pooled over all ten subjects, after being converted to a standardized z-score to minimize the effects of physiological differences between speakers. In each environment, values were tabulated and analyzed by ANOVA
(Duration, F0 statistics) or by means of a linear mixed model (Harmonics, OQ, Airflow) with Tone as a fixed effect and Speaker as an error term. An F-test of this model provided a global evaluation as to whether the four group means (each tone) in a given statistic were different. Given the questions being asked about the lexical tone contrasts in Burmese, a statement that some difference exists, such as between the OQ means for each tone in phrase-final position, was only so interesting without an understanding of which individual tones were different from one another. Scheffé or Bonferroni-Dunn post-hoc tests conducted multiple comparisons of the six possible tone pairs for each measure that was found to show a significant effect of Tone. These post-hoc tests were able to disclose if an overall effect was the result of all four tones exhibiting unique behaviors or the result of some other relationship, such as three tones with nearly identical data but a fourth with radically different values. As it turns out, across all of the measurements in (1) – (29) below, the former was never the case within a Carrier Phrase, while the latter was quite common; Creaky or (most frequently) Checked tone syllables often yielded a unique range of data points.

The sections below follow the same order of phonetic qualities seen in the Literature Review (§2) and the Methodology chapter (§3). Section 4.2 shows the vowel durations associated with each tone and Section 4.3 looks at their distinct pitch contours. The associations listed in hypotheses H1 through H4 are not tested in earnest until Sections 4.4 – 4.6, which examine measures of phonation type between the tones. The twenty-nine dependent variables described in detail through Chapter Three and tested over the following five sub-sections are:
1. DURATION (vowel duration in ms)
2. F0\textsubscript{MAX} (Hz)
3. ΔF0 (Hz difference in peaks, early peak gives negative #)
4. \( m_{F0} \) (Hz/sec as ΔF0 – t between peaks)
5. H1-H2\textsubscript{80%} (dB difference at 80\textsuperscript{th} percentile)
6. H1-H2\textsubscript{90%} (dB difference at 90\textsuperscript{th} percentile)
7. H1-H2\textsubscript{EARLY} (mean dB difference of 10-30\textsuperscript{th} percentile)
8. H1-H2\textsubscript{LATE} (mean dB difference of 80-100\textsuperscript{th} percentile)
9. ΔH1-H2 (difference in dB difference, LATE – EARLY)
10. H1-A1\textsubscript{80%} (dB difference at 80\textsuperscript{th} percentile)
11. H1-A1\textsubscript{90%} (dB difference at 90\textsuperscript{th} percentile)
12. H1-A1\textsubscript{EARLY} (mean dB difference of 10-30\textsuperscript{th} percentile)
13. H1-A1\textsubscript{LATE} (mean dB difference of 80-100\textsuperscript{th} percentile)
14. ΔH1-A1 (difference in dB difference, LATE – EARLY)
15. H1-A3\textsubscript{80%} (dB difference at 80\textsuperscript{th} percentile)
16. H1-A3\textsubscript{90%} (dB difference at 90\textsuperscript{th} percentile)
17. H1-A3\textsubscript{EARLY} (mean dB difference of 10-30\textsuperscript{th} percentile)
18. H1-A3\textsubscript{LATE} (mean dB difference of 80-100\textsuperscript{th} percentile)
19. ΔH1-A3 (difference in dB difference, LATE – EARLY)
20. OQ\textsubscript{MEAN} (mean ratio of all deciles)
21. OQ\textsubscript{EARLY} (mean ratio of 10-30\textsuperscript{th} percentiles)
22. OQ\textsubscript{LATE} (mean ratio of 80-100\textsuperscript{th} percentiles)
23. ΔOQ (difference of means OQ\textsubscript{LATE} – OQ\textsubscript{EARLY})
24. \textit{airflow}\textsubscript{WHOLE}: mean airflow rate in mL/s over whole syllable
25. \textit{airflow}\textsubscript{EARLY}: mean airflow rate in mL/s over the 1st half of the syllable
26. \textit{airflow}\textsubscript{LATE}: mean airflow rate in mL/s over the 2nd half of the syllable
27. \textit{airflow}\textsubscript{PEAK}: maximum mL/s value detected during the vowel
28. \textit{airflow}\textsubscript{MID}: flow at the syllable mid-point
29. \textit{airflow}\textsubscript{POST-VOCALIC}: mean airflow rate during 75ms following vowel offset
4.2 Duration

4.2.1 Overview of Duration Results

All measures of Duration in this section were of only the vowel – measures of Checked syllables do not include the glottal stop coda and CVN syllables were not considered. A main conclusion reached here reflects an idea posited in Chapter Two – that duration distinctions are most pronounced in citation form utterances and are significantly neutralized in connected speech. A three-way contrast in Duration (High, Low > Creaky > Checked) was found regularly in the data. Except in phrase-medial tokens, Duration values shifted toward a two-way contrast (High, Low, Creaky > Checked), one representing a distinction between open CV and closed CVO syllables rather than between tones.

Duration was standardized for each speaker\(^1\), allowing data for every speaker to be pooled and analyzed statistically for effects of Tone in each context. The standardized and pooled means for each Carrier Phrase are presented in Figures 1 and 2 below. An ANOVA on Duration with Tone and Carrier Phrase as factors and Subject as a random effect was performed and found a significant effect for both Tone \((F(3,27) = 98.69, p = .000)\) and Carrier Phrase \((F(7,63) = 38.86, p = .000)\). Post-hoc Scheffé tests further revealed that each Tone constituted a distinct set (values for each tone were distinct from every other tone at \(p = .000\)). However this was not the case when data within each Carrier Phrase was inspected. Effects between the Phrases were strong for some (Phrase 1 was different from all others at \(p = .000\)), while many did not have statistically different

---

\(^1\) Duration values were transformed according to mean and standard deviation of all open syllables in all Phrases (i.e. not Checked tones). Therefore, the distribution of all transformed duration values was mostly normal but for the set of low-scored Checked vowel durations creating a second mode in the data.
effects from one another on the vowel DURATION. Looking at Figure 1, consider the resemblance between phrase-medial token DURATIONS in Phrases (2), (3), and (4).

![Figure 1. DURATION (z-score) for tokens in Carrier Phrases (1) – (4). At column bases are re-transformed standardized values in milliseconds.](image1)

![Figure 2. DURATION (z-score) for tokens in Carrier Phrases (5) – (8). At column bases are re-transformed standardized values in milliseconds.](image2)
The homogeneity of vowel duration data between these Phrases suggests that the presence of surrounding tone-bearing syllables affected duration, but that the tonal properties of these adjacent syllables was relatively unimportant. There were five different subsets of Carrier Phrases with statistically indistinguishable effects on duration, listed in the table in (30) with the mean standardized duration of all tokens in that phrase.

(30) | Group                        | Phrase # | Subsets - Standardized Duration in ms |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Isolated tokens</td>
<td>1</td>
<td>272.4</td>
</tr>
<tr>
<td>b. Phrase-final, w/ focus</td>
<td>5</td>
<td>252.6</td>
</tr>
<tr>
<td>c. Phrase-final tokens</td>
<td>6</td>
<td>211.5</td>
</tr>
<tr>
<td>d. Medially, between</td>
<td>4</td>
<td>195.8</td>
</tr>
<tr>
<td>tone-bearing syllables</td>
<td>3</td>
<td>191.9</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>187.8</td>
</tr>
<tr>
<td>e. Medially, followed by</td>
<td>7</td>
<td>172.3</td>
</tr>
<tr>
<td>minor syllable</td>
<td>8</td>
<td>165.2</td>
</tr>
</tbody>
</table>

The data for each of these environments (30a – 30e) was pooled for the descriptive summaries below as well as for one-way ANOVAs with only Tone as a factor. Each group in (30), identified as a different phonetic environment, is discussed separately in the sections below. The ANOVA results indicated some statistically significant difference in the mean duration across tones for each environment – the nature of these differences between tone groups was examined by post-hoc Scheffé tests.

---

2 Durations given in (2) represent the standardized values pooled over all subjects, but retransformed into milliseconds for illustrative purposes, using the normalization parameters for a single subject. The values therefore represent the behavior of all subjects as though a single speaker (in this case, Speaker I).
4.2.2 *Duration of Isolated tokens – Carrier Phrase 1*

For every speaker, the recorded vowel durations in Carrier Phrase (1) conformed to the expectations garnered from prior research (§2.3): considerably longer High and Low tones and very short Checked tones. Such findings offer a three-way *duration* contrast (High, Low >> Creaky >> Checked) which was anticipated in isolated syllable utterances if nowhere else. Figures 3 and 4 display the mean *duration* of isolated Phrase (1) tokens for individual speakers.

![Figure 3](image1.png)

*Figure 3, 4. Mean *duration* by Tone of isolated utterances for each speaker.*
The separate figures for female subjects A – F (Fig 3) and male subjects G – I (Fig 4) are not indicative of any gender difference in the data, but are used simply to accommodate the data of ten subjects to page width. Each mean represents $n = 12$. Error bars indicate +/-1 standard deviation.
High and Low tone vowels were the longest, but neither had reliably greater duration across speakers. A few speakers showed a statistically significant difference in duration between High and Low tones, but not in a direction consistent across speakers. Creaky tone vowels were regularly shorter (often half the duration of High or Low vowels) and every speaker produced Checked vowels as the shortest of the four tones by far. For every speaker, the longest Checked vowel spoken in isolation was still shorter than any Phrase (1) High or Low tone vowel and only longer than roughly a quarter of all Phrase (1) Creaky vowels (26.3%). While Creaky and Checked durations did overlap this little bit, the difference in mean duration between Creaky and Checked tokens was highly significant (t(22), p = 0.00 by independent samples t-test) for all ten speakers.

Statistical tests over all speakers’ data with normalized duration show an overall difference between the four tones (F(3,27) = 135.6, p = .000). Three distinct subsets, High, Low > Creaky > Checked, were all significantly different from one another by post-hoc Scheffé test at p < .01. As shown in (31), High and Low tone means were not significantly different by Scheffé test at p = .135.

(31) Duration in Isolated Utterances, z-score means and Scheffé results.

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Low</th>
<th>Creaky</th>
<th>Checked</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+1.579</td>
<td>+1.386</td>
<td>-.317</td>
<td>-1.111</td>
</tr>
<tr>
<td>p</td>
<td>.135</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

3 High tone vowels were significantly longer for two speakers by independent samples t-test t(22) (Speaker A p = 0.01; for F p < 0.01), and Low tone significantly longer for two other speakers (for Speaker D p = 0.05; E p = 0.01). The remaining six speakers showed non-significant (p > 0.05) larger means for High tone vowels.
4.2.3 **Duration of Phrase-Final tokens – Carrier Phrases 5, 6**

In phrase-final utterances (Phrases 5 and 6), the same three-way distinction (High, Low > Creaky > Checked) was upheld, though there was a great deal of variation between speakers as well as between the two carrier phrases. The different behavior of the two phrases was considered here to be more likely an effect of the overall length of the phrase preceding each token, rather than attributed to the preceding tones (High or Low). The two sentences are repeated for reference in (32):

(32) Carrier Phrases (5) and (6) with Phrase-Final Tokens

<table>
<thead>
<tr>
<th>Phrase (5)</th>
<th>Phrase (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OWER ___</td>
<td>OWER ___</td>
</tr>
<tr>
<td>[kû __ ]</td>
<td>[hû'kè da __ ]</td>
</tr>
<tr>
<td>Nine __.</td>
<td>Yes, it is __.</td>
</tr>
</tbody>
</table>

In Phrase (5), the embedded token was phrase-final following a High tone, but in a phrase of only two syllables. In this position, the tokens received a considerable amount of focus and produced results (at least for vowel duration) in line with Phrase (1) isolated tokens. Phrase (6) tokens were generally shorter and occurred finally in a tetra-syllabic phrase, following a Low tone syllable. More precisely, all possible open syllable tokens (High, Low, and Creaky CV) were produced in Phrase (6) with a shorter mean duration than Phrase (5) \(n = 4\) for eight of ten speakers. Therefore, the two phrases’ data were not pooled and only Phrase (6) tokens were considered for the “Phrase-Final” duration values and statistical analysis given below. The Phrase (5) duration data were not selected since it was similar to already examined Phrase (1) data.
Figure 5, 6. Mean DURATION by Tone in Phrase (6) tokens for each speaker. Each mean represents $n = 4$. Error bars indicate +/-1 standard deviation. High, Low, and Checked vowels are in-line with isolated token DURATION in Figures 3, 4. Creaky vowels are more variable, according to speaker.

In this phrase-final position, the distinction of High, Low > Checked DURATION remained very strong (highly significant at $t(6), p < 0.05$ for eight of ten speakers, barring D and J) but the DURATION of Creaky vowels is highly varied. For some speakers, Creaky vowels were not significantly longer than phrase-final Checked vowels (Speakers C, E, F and G$^4$), while for other speakers they grouped with the High and/or Low tones (Speakers B, E, F).

---

$^4$ Between-group difference by post-hoc Scheffé test for Creaky and Checked durations: Speaker C ($p = .964$), E ($p = .960$), F ($p = .949$), and G ($p = .059$).
D, H, I, and J). For only one speaker were Creaky vowels a distinct group in Duration as they had been in Isolated utterances (Speaker A at \( p < 0.05 \) from each other tone in pair-wise comparison).

Pooling normalized values for all speakers, there was an overall difference in vowel Duration between the tones (\( F(3, 27) = 29.72, \ p = .000 \)). Between group differences by post-hoc Scheffé test reveal a three-way distinction at \( p < .05 \) shown in (33), which approached a four-way distinction as the difference between High and Low Duration neared significance (\( p = .056 \)).

(33) Duration in Phrase-Final Tokens, z-score means and Scheffé results.

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Low</th>
<th>Creaky</th>
<th>Checked</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+.296</td>
<td>-.104</td>
<td>-.798</td>
<td>-1.212</td>
</tr>
<tr>
<td>( p = .056 )</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

4.2.4 Duration of Phrase-Medial tokens – Carrier Phrases 2, 3, and 4

Phrase medially, the Duration of High and Low tones was reduced considerably but those for Creaky tone vowels were not altered greatly. In fact, for four speakers (A, C, E, G) the mean Creaky vowel Duration in phrase-medial contexts was actually greater than in isolation\(^6\). The result was that High, Low, and Creaky vowels had more similar

---

\(^5\) Results of post-hoc Scheffé test: Speaker H \( (p = .172) \), I \( (p = .245) \), J \( (p = .706) \) had no difference between Low and Creaky means. Speaker D, no difference \( (p = .998) \) between High and Creaky means. Speaker B, Creaky different from both; \( (p = .501) \) with Low, \( (p = .218) \) with High.

\(^6\) For speakers A, C, E, G. Significantly so for Speaker A \( (t(22) = 6.87, p = .000) \); Speaker G \( (t(22) = 2.29, p = .032) \).
DURATION when phrase-medial (details below). Checked vowels, as in other contexts, remained markedly shorter than the other three tones in the three medial phrases.

Figure 7, 8. Mean DURATION by Tone in Phrases (2), (3), & (4) for each speaker. Each mean represents \( n = 12 \). Error bars indicate +/-1 standard deviation. High, Low, and Creaky vowels have more uniform DURATION than in other environments.

Figures 7 and 8 give the mean durations for the tones pooled over Phrases (2), (3), and (4), which all placed the variable token between two other tone-bearing syllables. Looking to the pooled DURATION values for all speakers, there was again a significant overall difference in vowel DURATION between the tones \( (F(3, 27) = 75.32 \ p = .000) \). However, the between group differences were considerably different than was seen for Isolated (§4.2.2) or Phrase-final (§4.2.3) tokens, as Creaky vowel DURATION became
more in-line with that of High and Low vowels. This similarity was reflected in the lower, but still very significant $F$-statistic and also in the results of post-hoc Scheffé tests, where Creaky and Low DURATION were not significantly different ($p = .107$), although Creaky and High DURATION values still were ($p = .000$). A comparison of the standardized means and the statistically different groups by Scheffé test are given in (34).

(34) Duration in Phrase-Medial Tokens; z-score means and Scheffé results.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>-.229</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>-.470</td>
<td></td>
</tr>
<tr>
<td>Creaky</td>
<td>-.597</td>
<td></td>
</tr>
<tr>
<td>Checked</td>
<td>-1.460</td>
<td></td>
</tr>
</tbody>
</table>

As in other environments, there were three distinct subsets of Duration values – though with the tones now grouped differently. High tone vowels remained statistically longer than Creaky vowels when phrase-medial, however the magnitude of the difference was dramatically reduced (< 0.5sd, rather than between 1-2sd). This proximity of Creaky to both High and Low Duration values is the key finding in this data, a distribution which was to some degree borne out in the Scheffé results but which is more apparent by putting aside the statistical probability of homogeneity between the means and reviewing a simple comparison of the means across sentence contexts, given in (35). There can be little doubt that the large difference in Duration seen in phrase-final and isolated tokens was mostly neutralized in medial positions.
To summarize, High and Low tone vowels had significantly greater duration in the Isolated and Phrase-final positions, a difference that was only maintained for High vowels (and less robustly) in medial positions. It is also worth noting that the grouping of Low and Creaky duration values is hardly definitive as the difference between their mean durations approached significance at $p < .05$. However, inferential statistics tell the story rather poorly. While statistical differences existed between tones in Phrase-medial duration, the effect was drastically different from the pattern in isolation. High and Low vowels in isolated syllables were typically 100ms longer than Creaky vowels, or more than 150% of their duration. Compare this gap with the much smaller significant difference in Phrase-medial vowels: one of roughly 10-30ms, or about 105-115% of the mean Creaky vowel duration.

4.2.5 Duration of Phrase-medial tokens before Minor syllables – Carrier Phrases 7, 8

The effects on vowel duration in Phrases 7 and 8 were similar to the effects for the Phrase-medial tokens described above, but with briefer values across the board. The effect of tone was significant ($F(3, 27) = 39.33$ $p = .000$), though much of the variance in these phrases was from the much shorter Checked tones – Creaky, Low, and High tone duration values were more closely distributed in this context than any other. Creaky
and Low tone DURATION were not statistically different by Scheffé test ($p = .285$), nor were Low and High tone DURATION ($p = .088$).

**Figure 9, 10.** Mean DURATION by Tone in Phrases (7), (8) for each speaker. Each mean represents $n = 12$. Error bars indicate $+/1$ standard deviation. High, Low, and Creaky vowels have more uniform DURATION than in other environments.

(36) DURATION in Phrase-Medial, Pre-Minor Syllable Tokens, $z$-score means and Scheffé results.

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Low</th>
<th>Creaky</th>
<th>Checked</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-.668</td>
<td>-.842</td>
<td>-.923</td>
<td>-1.637</td>
</tr>
</tbody>
</table>

$p = .088$  \hspace{1cm}  $p = .285$  \hspace{1cm}  -
On a speaker-by-speaker basis for tokens preceding a Minor syllable, Creaky vowels were longer than High or Low vowels for two (of ten) speakers (Subjects D, G), longer than only Low vowels for another two (B, H), and with a shorter mean duration for the remaining six subjects (A, C, E, F, I, J). This pattern closely resembled that found for other Phrase-medial tokens (§4.2.3), with the difference being a stronger neutralization of the contrast in CV syllables as all durations were shortened.

The nature of the phrase-medial effect seen above (in both types of medial frames) warrants some discussion of why this context yielded a shorter vowel duration, as well as other phonetic consequences detailed in §4.3 - 4.6. In a cursory comparison of the study’s elicited Carrier Phrases, “phrase-medial” meant precisely that: the embedded stimulus token was uttered in the middle of the phrase rather than at its end. When comparing the effects of sentence position, it is worth clarifying that regular differences in phrase-medial pronunciation are not necessarily claimed to be the product of “mediality”, in the sense that the presence of adjacent phonetic material induces some kind of phonetic or phonological change. In most cases, it is the manner in which adjacent phonetic material was seamlessly strung together with an embedded token in running speech which contributed to the effects of “phrase-medial” position. The elicitation of all data requested subjects to speak at a casual and conversational pace, and data for one subject (Speaker L) was entirely disqualified because the subject repeatedly failed to maintain such a pace (see §3.2).

The claim is not as simple as that the reduced phrase-medial duration times were a product of a faster rate of speech in these utterances, mainly for two reasons. One,
medial Creaky and Checked syllables were only slightly shorter than when spoken in isolation (see §4.2.4 and footnote 6), a reduction that cannot compare to that seen between medial and isolated High or Low tone syllables. This asymmetry speaks to a floor vowel duration which tone-bearing vowels do not sink below under normal conditions\(^7\). Quite simply, Creaky and Checked vowels could not reduce to a short vowel length because they were already short.

Secondly, the data arguably better represents the *lengthening* of isolated and final forms rather than the *shortening* of medial forms. Overall lower mean DURATION for all tones implies some de-emphasis medially, but the interaction between tone and context on DURATION data more strongly suggests that Phrase (1) isolated tokens were produced with the full emphasis and focus of citation forms. Similary, longer vowels in the phrase-final position of Carrier Phrases (5) and (6) are interpretable as phrase-final lengthening, a noted effect of phrase-level boundaries on prosody in English and cross-linguistically (Lehiste 1973, Klatt 1975, Cooper and Danly 1981, Wightman *et al.* 1992). The divide between the two phrase-final carrier sentences further illustrates the effect of emphasis – a disyllabic counting phrase (Phrase 5: “Nine ___”) placed more emphasis on the stimulus token than final position in a longer phrase (Phrase 6: “Yes it is a ___”). Accordingly, significantly longer vowels were attested with Phrase (5). Lastly, the observations here regarding phrase-mediality and connected speech do not only pertain to the present duration data, but also to the numerous variables quantifying pitch and phonation types that are addressed in the upcoming sections.

\(^7\) Note though that Minor “toneless” syllables are regularly much shorter than this “floor” for every speaker. Minor syllable vowel DURATION typically fell between 25ms – 75ms, about half the length of the shortest Checked syllable vowels.
4.2.6  \textit{Inter-Speaker Variation}

The charts in sections 4.2.2 – 4.2.5 showed \textit{duration} to be mostly consistent across speakers. Speaker differences certainly existed though, and were not limited to trivial variation by speaker in absolute \textit{duration}, but also in the relative length of the tone-bearing vowels – i.e., Creaky vowel \textit{duration} matched Low vowel \textit{duration} for some speakers, but was significantly shorter for others. In the ANOVAs for each phrasal environment using the standardized $z$ \textit{duration} values, the random effect of speaker was found to have a significant effect on \textit{duration} in both Medial contexts (37a, b). Speaker differences were not significant in Final contexts (37c), nor for Isolated utterances (37d). Standardized $z$-scores allowed the test to examine if overall patterns in the data were different by speaker instead of absolute values in milliseconds, which were certain to be different by speaker.

<table>
<thead>
<tr>
<th>Context</th>
<th>Effect of Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Medial – between tones</td>
<td>$F(9,27) = 4.13$</td>
</tr>
<tr>
<td>b) Medial – pre-minor syllable</td>
<td>$F(9,27) = 4.79$</td>
</tr>
<tr>
<td>c) Final (Phrase 6)</td>
<td>$F(9,27) = 0.91$</td>
</tr>
<tr>
<td>d) Isolated (Phrase 1)</td>
<td>$F(9,27) = 1.94$</td>
</tr>
</tbody>
</table>

The effect found in the medial contexts but not others, seen in (37a, b), was not surprising given the variation described earlier for the context-dependent compression of High and Low vowel \textit{duration} (Example (35), §4.2.4). Still, much of the inter-speaker variation was found in the data of just two speakers, A and E, who produced atypically longer vowels phrase-medially (cf. Figures 7, 9), and shorter vowels (particularly Checked) in citation form (cf. Figure 3). Running the same one-way ANOVA, but with the data for Speaker A and E excluded, the effect of speaker was reduced in every phrasal context.
This is denoted in the smaller $F$-statistic values and accordingly higher $p$-values in (38a-d). Even with the two most irregular speakers removed from the data pool, a significant effect of Speaker was still found with one set of Medial forms and a near significant effect with the other set of Medial forms.

(38) Random Effect of Speaker on DURATION excluding Speakers A, E

<table>
<thead>
<tr>
<th>Context</th>
<th>Effect of Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) MEDIAL – between tones</td>
<td>$F(7,21) = 3.98$</td>
</tr>
<tr>
<td>b) MEDIAL – pre-minor syllable</td>
<td>$F(7,21) = 2.47$</td>
</tr>
<tr>
<td>c) FINAL (Phrase 6)</td>
<td>$F(7,21) = 0.85$</td>
</tr>
<tr>
<td>d) ISOLATED (Phrase 1)</td>
<td>$F(7,21) = 1.39$</td>
</tr>
</tbody>
</table>

The conclusion is that DURATION behavior was consistent across speakers for words that were not sentence-medial, but that some vowel durations were compressed in juncture and speakers varied as to which tones were shortened and by how much. Primarily, the variance resulted from a split between speakers for whom High and Low tones were reduced to a DURATION comparable to their Creaky tone data, and those for whom High and Low tone vowels had a statistically greater DURATION. This division of speakers is demonstrated in (39).

(39) Variation by Speaker of Phrase-Medial Tone Contrast in DURATION

<table>
<thead>
<tr>
<th>Contrast</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-way</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>High = Low = Creaky</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three-way</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High, Low &gt; Creaky</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8 Labels do not include Checked tones because all speakers produced Checked tones medially with a significantly shorter vowel DURATION.
9 Actually High > Low, Creaky in the case of Speakers E and J. Since this distribution still presents a three-way contrast, E and J are presented on the same row in (39) for ease of interpretation.
4.2.7 **Summary of Duration Data**

Given the robust duration distinction reported in Thein Tun’s (1982) multi-speaker study and the well-understood brevity of Checked vowels, the finding that Tone had an effect on the measure of **Duration** in all Carrier Phrases was not surprising. More specifically, the duration differences between the tones matched the distinctions listed in the descriptive literature on Burmese (e.g. Okell 1969, Wheatley 1987, Watkins 2001) with a few mostly anticipated context-dependent alternations. Checked syllables always had a distinctly shorter **Duration** regardless of context. On the other hand, differences between High, Low, and Creaky vowel **Duration** were context-dependent. The contrast in **Duration** values was strongest in citation, reflected in a three-way distinction (four-way for some speakers) in the isolated position of Phrase 1 (§4.2.2). In embedded contexts every tone had a shorter mean **Duration**, though a further distinction was found with High and Low tone vowels – **Duration** was more reduced (from citation form lengths) in phrase-medial positions than in phrase-final positions. The trend towards more uniform durations in High, Low, and Creaky tone vowels is shown in (40), with double brackets ‘>>’ denoting greater contrast than a single bracket ‘>’.

\[(40)\]

a. High, Low >> Creaky >> Checked \hspace{1cm} Isolated utterances
b. High, Low > Creaky >> Checked \hspace{1cm} Phrase-final position
c. High, Low, Creaky >> Checked \hspace{1cm} Medial positions, in juncture

The pattern in (40) reveals a context-sensitive neutralization. In phrase-medial position, a distinction (found in isolation) between High, Low, and Creaky **Duration** is neutralized.
Analysis of individual speakers’ data showed that for some speakers the neutralization was complete, while for others the High and Low tones were shortened but still of greater duration than Creaky tones. Since these three tones are produced only in open syllables (CV) or sonorant-coda syllables (CVN), and the consistently shorter Checked tone is restricted to obstruent-closed syllables (CVO), it is arguable that the duration contrast found in medial forms is in fact a product of syllable structure rather than Tone. Open syllable vowels had a moderate length and CVO syllable vowels were very short. Of course, CVN syllables were not considered in the statistical analyses and only vowel length, rather than the duration of the rhyme or entire syllable, was measured. Open and closed syllables (both CVN and CVO) may indeed be equally long if vowel and coda durations were totaled, entailing that all tone-bearing syllables are of similar length when uttered phrase-medially at a casual pace. Issues of syllable and segment duration in every Burmese syllable type (CV, CṼ, CVN, and CVO) are put aside until Chapter Five. The phonological analysis later in the dissertation pursues this idea though – that medially, durations by tone are roughly uniform while being categorically different in citation.

In phrase-final position, the neutralization was also found, but only as a partial neutralization. High and Low tone syllables were still compressed such that their mean duration was similar to, but still significantly greater than, the shorter duration of Creaky syllables.
4.3 Fundamental Frequency

Many of the Carrier Phrases were constructed specifically to test the effects of neighboring tones on the pitch profile of a tone. Therefore, for the investigation of F0 similar results amongst contextual frames were disregarded and data for each Carrier Phrase was considered individually. For instance, the difference in the realization of pitch between [L__H] and [L__L] tokens or [L__#] and [H__#] tokens was noteworthy regardless of how strong or weak it was found to be.

The set of all F0 values was normalized for each speaker, providing a $z$-score scale reflecting F0 height within that speaker’s attested range. Data for every speaker was pooled in this format. Figures 11 – 19 below chart the mean value of $z_{F0}$ for all speakers to show the relative height and movement in F0 over each decile of the syllable duration, which has been normalized over the average duration of each tone.

4.3.1 F0 of Isolated Tokens (Carrier Phrase 1)

The production of Burmese tones spoken in isolation have been well-studied in the literature and the expected findings in this context were mostly clear. The normalized F0 traces presented in Figure 11 were easily countenanced with the descriptions from prior research, but for one conspicuous pattern found with the Low tone.

A majority of the speakers (7 of 10) produced a late, steep F0 rise on Low tone syllables, matching the identification of the tone as the “Rising Tone” in Javkin & Maddieson (1983) or as rising at times in descriptions by Cornyn (1944), Okell (1969) and Bradley (1982). In the ensuing section on phrase-final data, it is shown that the same sets of speakers mostly produced accordingly similar contours in the final position of
Phrases (5) and (6). The split in behavior (Speakers C, H, I vs. A, B, D, E, F, G, and J) was limited to isolated and phrase-final utterances as the shape of Low tone contours was more or less homogenous, and non-rising, across speakers in medial contexts.

Watkins (p.c.) has indicated that this final rise is a politeness marker, common in citation forms. This understanding nicely explains the inter-speaker variation: some speakers chose to employ this form and others did not, their decision likely shaped by their response to the data collection setting. One might conclude that the rise designates the implementation of a Burmese sentence-level intonational melody, yet the effect is curiously not generalized to all forms in the language, but seems restricted to only Low tone forms. A fuller analysis of this phenomenon would capture the fact that the marker is in some way compatible with only the Low tone, or possibly even a select set of Low tones.
tone lexical items. However, the issue is not resolved in this dissertation and it is only discussed further within the scope of the following justification behind the decision to exclusively use the non-rising F0 data in the chapter’s ensuing statistical analyses.

To begin with, pooling both the rising and non-rising data together was not an option. A mean F0 trace of all speakers would not have reflected the typical Burmese speaker’s production of the Low tone. Why not then use the more common, rising contoured data? Though the non-rising tokens represent the minority in the isolated syllable utterances, the division in all cases was not reliably different from 50/50\(^{10}\). It was felt that the non-rising tokens were more representative of the Low tone’s typical phonetic representation, while the rising Low tone samples contained some effect of an auxiliary prosodic influence, regardless of the origin of this influence in the analysis.

Looking to the contours found with the other tones in isolation, High tones saw a gradual pitch fall over the course of the syllable. A steeper fall was found on Creaky and Checked tone syllables, for which the falling F0 traces were nearly identical. While the High, Creaky, and Checked tones all had early F0 peaks, the High tone early peak was lower on average, meaning that F0 in Creaky and Checked tones typically had farther to fall, as well less time in which to achieve that fall. Comparisons of F0 slope (\(m_{F0}\)) in the statistical analysis (§4.3.5) were able to capture the distinction between steeper Creaky and Checked falls and those of High and Low tones, but there was no affirmation in any

\(^{10}\) The two speakers whose data was generally removed from the study (speakers K and L) did not produce Low tones with this final rise in any Carrier Phrase. Including these subjects, the set of speakers producing the late rise comprised 7 of 12 speakers in Phrases (1) (citation form) and (5) [High __ #].

The late F0 rise was less frequent in the final position of Carrier Phrase (6) [Low __ #], where three fewer speakers produced this final rise. Therefore 4 of 12 speakers incorporated a late rising F0 in this context. From those who had rising Low tones in Phrase (1) and (5), Speakers D, E, J joined C, H, and I in producing Low tones with low, slightly and slowly falling F0 traces in Phrase (6).

The split was perhaps not entirely random though – only one of the speakers without the late rise was female. The proportion of speakers using the rise in Phrases (1), (5), and (6) is then 5 of 6 female speakers and just 2 of 6 male speakers.
Carrier Phrase context of a three-way distinction in the tones denoting an even vs. a gradual fall vs. a steep fall.

4.3.2 F0 of Phrase-Final Tokens (Carrier Phrases 5, 6)
Results for phrase-final tokens in Phrase (5) and (6) are provided in Figures 12 and 13.

Figures 12 (above); Figure 13 (below). Mean F0 (z-score) at deciles in Phrases 5, 6 by Tone across all speakers. Y-axis values are arbitrary scale normalized for each speaker. Duration on x-axis is mean duration for each tone across all speakers in milliseconds. $n = 40$ for each decile data point. Two F0 traces are shown for the Low tone: the dotted line for the data of Speakers C, H, and I (and D, E, J in Fig. 11); the solid line represents the normalized mean F0 values for Speakers A, B, D, E, F, G, and J.
The split in speaker behavior between rising and level Low tone pitch patterns is evident in each figure. The same three speakers (C, H, I) who produced isolated Low tones without a rise, did the same for both phrase-final Carrier Phrases. Additionally, three other speakers (D, E, J) had no final rise in their Phrase (6) data. F0 contours in phrase-final positions were similar to those of isolated tokens in other ways too, sharing the same three-way distinction outlined in description of Phrase (1) tokens above (§4.3.1):

(41) Phrase-final F0 Patterns

a. **Low tone** contours fell slightly initially and remained low-pitched (or frequently had a rise due to intonational considerations).

b. **High tone** contours had high onset F0 followed by a gradual fall over the course of the syllable.

c. **Creaky** and **Checked tone** contours had the highest onset F0, before a sharp fall to the lowest F0 offset among the tones.

Also as seen before, Creaky and Checked tones bore virtually indistinguishable F0 traces. This likeness was not particular to isolated and phrase-final positions however, and will be seen to hold in most environments throughout the rest of section 4.3.

To compare the effects of the two Carrier Phrases (5) and (6) on the embedded token, it was apparent that Phrase (5) [High __ #] tokens had a much higher F0 at the vowel onset than those embedded after a Low tone in Phrase (6) [Low __ #].

It is tempting to attribute the raised F0 entirely to the tone of the preceding syllable: the initial raised F0 induced by a preceding High tone either as a result of a
phonological alternation or simply the limitations of the physical system achieving and reversing a heightened rate of vocal fold vibration. However, as was seen in the section concerning vowel DURATION, the difference between Phrases (5) and (6) was more than minimal – that is, the environment in which the embedded token was placed did not differ solely in having either a preceding [H] or [L] tone. For the short noun phrase that constituted Phrase (5), every speaker systematically produced a more emphatic stimulus token. In the less emphatic Phrase (6) context, tokens were shorter (§4.2.3) and had lower Open Quotient (§4.5.2) values at all decile timepoints, to accompany the lower F0 values seen in Figure 13 above. Two contrasts could explain the large difference in F0 between Figures 12 and 13: either (a) the preceding tone, High vs. Low, or (b) the difference in emphasis of the embedded position, the focused terminal position of Phrase (5) vs. the weak terminal position of Phrase (6). Statistically, both factors were independently found to affect F0 (as explored in §5.6.3 – 5.6.4), indicating that the exceedingly low F0 found with Phrase (6) tokens may best be explained as the compounded result of two separate F0 lowering effects – a prior Low tone and sentence-terminal position.

4.3.3 F0 of Phrase-Medial Tokens (Carrier Phrases 2, 3, and 4)

*PHRASE (2) [LOW __ LOW]*

Figure 14 below shows that the results for Phrase (2) tokens matched the prior results from the studies in Watkins (2001) and Gruber and Feizollahi (2006)\(^\text{11}\) (cf. Chapter 2, Figure 3). There were also numerous similarities to the F0 traces found on tokens in isolated and phrase-final frames. Low tones began with a slight fall and then leveled out.

\(^{11}\) The [Low __ Low] frame of Carrier Phrase (2) did not only match the phonetic context of Gruber & Feizollahi’s data, it used the exact same carrier sentence: [ʃɔˈndə __ ko jeɪnədə]. ‘I am writing __.’
Creaky and Checked tones were high-pitched early and falling, though the fall in medial cases tended to be much less precipitous. A very different contour was found for High tones though, which bore a gradually rising F0 in Phrase (2) tokens. In addition to matching previous instrumental studies, this finding fit observations concerning the High tone by Bradley (1982) and Watkins (2005a).

The traces in Figure 14 show High and Low tones with a similar initial F0. Likewise, but at a higher F0, Creaky and Checked tones shared similar values early at the vowel onset. In fact, this similarity extended to all decile data points for Creaky and Checked tokens, though the brevity of Checked syllables typically interrupted the fall, whereas the longer Creaky vowel (and periodicity) allowed for a continued drop to lower average F0 values at vowel offset. Figure 15 provides a comparison of Creaky and Checked F0 attributes by plotting the height and change in F0 for every individual token in Phrase (2). This display shows the greater variability in Checked data, in both height
and the F0 fall measured for each sample. While every individual Creaky (and every Low) token had a negative change in F0, a few Checked tokens did not even fall, showing a positive $\Delta$ in normalized F0.

![Figure 15. Plot of normalized MAX F0 against the change in F0 ($\Delta$F0) for Creaky and Checked tokens in Phrase 2. Negative $\Delta$F0 (most datapoints above) indicates a rising F0 trace. Only a few tokens displayed a positive $\Delta$F0. Low tokens (in green) are plotted for purposes of comparison of density with the more scattered Checked data points.](image)

Phrases (3) and (4) also embedded the stimuli in a phrase-medial position, with either a preceding High tone (Phrase 3) or a following High tone (Phrase 4). Results for F0 in these tonal frames bore many similarities to the [Low __ Low] frame of Phrase (2). Figures 16 and 17 illustrate.
Figures 16, 17. Mean F0 (z-score) at deciles in Phrase 3 and 4 by Tone across all speakers. Y-axis values are arbitrary scale normalized for each speaker. Duration on x-axis is mean duration for each tone across all speakers in milliseconds. n = 40 for each decile data point.

**Phrase (3) [High __ Low]**

The medial tokens in Phrase (3) yielded F0 means that were scarcely different from those in Phrase (2) at the syllable onset, but were considerably lower at the offset. The mean normalized F0 of all four tones was lower in Phrase (3) than any other phrase-medial
This is curious given that the two phrases were followed by the same tone (Low) but were differentiated by the tone preceding the variable token (High in one case, Low in the other). The effect was such that falling tones fell further, the even tone dropped in pitch, and the High tone (rising in other medial contexts) did not rise, but maintained a somewhat elevated pitch. Of the forty total High tone samples recorded by all ten speakers in Phrase (3), in only a single instance was there a rise in F0. The other thirty-nine fell to some, albeit slight, degree. Again, tokens of the non-High tones were all distinctly falling – Creaky and Checked tones fell as was typical and Low tones fell for longer and to lower values than in the other medial frames. The effect across all tones was one of F0 depression following a High tone, possible causes for which are discussed in depth in §5.6.4 examining F0 co-articulation, downstep, and sandhi effects on pitch.

Even if the across-the-board lower pitch was entirely the product of co-articulation, downstep, or sandhi processes, the effect did not color all of the F0 information shown in Figure 16 for Phrase (3) data. The mean F0 traces here demonstrate a number of interesting generalizations regardless of the lower offset F0: High tones fall slightly after a High tone, Low tones fall and stay low, and Creaky/Checked tones fall as usual.

\[ t\text{-test results for each tone provided here.} \]

<table>
<thead>
<tr>
<th>Comparison of Phrase (2) vs. Phrase (3) means</th>
<th>Comparison of Phrase (2) vs. Phrase (4) means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Checked • ( t(78) = 4.377, p &lt; .001 )</td>
<td>Checked • ( t(78) = 4.377, p &lt; .001 )</td>
</tr>
<tr>
<td>Creaky • ( t(78) = 4.377, p &lt; .001 )</td>
<td>Creaky • ( t(78) = 4.377, p &lt; .001 )</td>
</tr>
<tr>
<td>High • ( t(78) = 4.377, p &lt; .001 )</td>
<td>High • ( t(78) = 4.377, p &lt; .001 )</td>
</tr>
<tr>
<td>Low • ( t(78) = 4.377, p &lt; .001 )</td>
<td>Low • ( t(78) = 4.377, p &lt; .001 )</td>
</tr>
</tbody>
</table>
**Phrase (4) [Low __ High]**

For Phrase (4) tokens, in a [Low __ High] frame, the pitch tracks for the Low tones (Figure 17) were nearly identical to those in Phrase (2), thus indicating that an ensuing High tone syllable has little to no anticipatory effect on a Low tone pitch. Similarly, Creaky and Checked tones bore similar falls to their Phrase (2) counterparts, but were in fact significantly lower (see §5.6.2) at the syllable offset – contrary to what might be expected were the following High tone to have any anticipatory effect. These lower F0 offsets could be at least partially explained by the uninterrupted periodicity in this carrier phrase due to the following syllable’s sonorant [l] onset. As opposed to the abrupt transition to voiceless [k] in the Creaky and Checked Phrase (2) tokens, the extended periodic portion saw a longer F0 fall with a rise for the following High tone beginning during its sonorant onset.

**Summary of Phrase-Medial F0 Data**

A couple points stand out when reviewing results for the three phrase-medial carrier phrases collectively. As has been noted numerous times, the average F0 traces for Creaky and Checked syllables were very similar in all three phrases, differing primarily in the duration of the vowel. The other point concerns the pitch alternation seen with High tones between a rising contour in Phrase (2) and a generally level but elevated pitch in Phrases (3) and (4). The diagrams in (42) illustrate.
(42) Mean F0 Trace of High Tones in Phrase-Medial Carrier Phrases

![Graphs showing F0 traces for different phrases](image)

The understanding adopted in this dissertation, and detailed more fully in Chapter Seven, is that a peak F0 was realized during the middle High tone in all three contexts, but was simply more distinct when there are no comparable F0 peaks on neighboring syllables. A High pitch target between Low tones forms a more identifiable peak and therefore a more identifiable contour – a fall with an early peak, and a rise with a late peak. When an adjacent syllable also bears a high pitch, the peak associated with the High tone may not be discernable from a sustained trans-syllable elevated pitch. This observation speaks to the phrase-medial data in a number of cases; for both the perseverative and anticipatory effects of preceding and following High tones, respectively, and regarding the implementation in any context of the high pitch targets associated with any of the tone categories.

4.3.4 F0 of Phrase-Medial Tokens Before Minor Syllables (Carrier Phrases 7, 8)

Lastly, the fundamental frequency produced with each tone in the two medial phrases followed by a minor syllable are given in Figures (18) and (19). The F0 traces here effectively highlight the same contrasts seen between Phrases (2) and (3), which also differed according to a High or Low preceding tone.
Figures 18, 19. Mean F0 (z-score) at deciles in Phrases 7 and 8 by Tone across all speakers. Y-axis values are arbitrary scale normalized for each speaker. Duration on x-axis is mean duration for each tone across all speakers in milliseconds. $n = 40$ for each decile data point.

In Phrase (7) the preceding High tone appeared to raise the F0 at the onset of the embedded token, since the vowel-initial mean F0 for every tone was greater in Phrase (7) than in Phrase (8). As when embedded medially in Phrases (2) and (3), the single rising tone in a post-Low context (High) was falling after a High tone, while typically falling tones were produced with a more precipitous F0 drop when following the High tone in
Phrase (7). These data, as the other medial forms did, also suggested the effect of a preceding High tone an initially higher but progressively lower F0 towards the syllable end. Potential explanations of this effect are explored in Chapter Five.

This concludes the presentation and description of the F0 responses found on the stimuli embedded in the eight carrier sentences. The decile-by-decile tracking of F0 height is an apt visualization of F0 movement, but comparisons at ten timepoints do not facilitate a statistical evaluation of the tonal differences. For this evaluation, the F0 traces in Figures 11 – 19 were described by three statistics capturing their key differences, $F_{0_{\text{MAX}}}$, $\Delta F_0$, and $m_{F_0}$ calculated from the set of normalized F0 values. The following section applies inferential statistical tests to these data.

4.3.5 Statistical Analysis of F0

For each Carrier Phrase, the figures below give the mean peak F0 ($F_{0_{\text{MAX}}}$), the mean difference between F0 maxima and minima ($\Delta F_0$), and the mean slope between the maxima and minima ($m_{F_0}$). All means represent the pooled data of all subjects after each subject’s data was normalized to his/her F0 range. Figures (20) – (27) also provide values retransformed into Hertz for illustrative purposes, offering a familiar and intuitive scale that depicts a hypothetical average speaker. The pooled z-score data was transformed along the parameters $M = 185$Hz, $SD = 30$Hz, which approximate the standardization values used for Speaker B (female). In essence, the hypothetical speaker is one with the overall average F0 and range of Speaker B, but the tone-by-tone use of this range that
represents the average behavior of all Speakers, A – J. These data are displayed on the left y-axis of each Figure and in **bold font** within the graph.

Following each figure is a chart detailing statistical differences between the tone means for each variable. In nearly all cases, there was a significant global difference by ANOVA at $p < .05$ between the tones. Of more interest were the results of post-hoc Scheffé tests identifying homogenous subsets of the four tones for each statistic. The statistical groupings determined here confirmed many of the observed differences in the above sections.

---

13 For the measure $\Delta F_0$, there was a significant overall difference between tones in all eight Carrier Phrases. All but one phrase had an overall difference by ANOVA for $F_{0\max}$ (Phrase 6), as well as a single phrase for the $n_{f_{0}}$ measure (Phrase 3).
Figure 20. Mean $F_{0}^{\text{MAX}}$, $\Delta F_{0}$, and slope ($m$) by Tone in Carrier Phrase 1. Central boxplot gives mean and distribution of $F_{0}^{\text{MAX}}$ data for each Tone in normalized $F_{0}$ for all speakers. Box contains Q1-Q3 of all data with whiskers marking the minimum and maximum values excluding outliers beyond 1.5x the inter-quartile range. Mean $\Delta F_{0}$ is shown at level of $F_{0}^{\text{MIN}}$ with +/- value. Accompanying line indicates $m_{F_{0}}$, with precise values provided in the lower box.

Subsets by post-hoc Scheffé test in Phrase 1: Isolated Syllable.

\[
\begin{array}{cccc}
\text{Tone} & F_{0}^{\text{MAX}} (\text{Hz}) & \Delta F_{0}: \text{Max-Min} & m_{F_{0}} (\text{Hz/sec}) \\
\hline
\text{ʔ} & \text{C} & \text{L} & \text{H} \\
\text{L} & \text{C} & \text{H} & \text{ʔ} \\
\text{H} & \text{L} & \text{ʔ} & \text{ʔ} \\
\hline
\text{Difference in pair, } p & .823 & .075 & .828 & .798 & .940 & - \\
\end{array}
\]

F(3, 20)=10.85, $p = .000$  
F(3, 19.7)=39.66, $p = .000$  
F(3, 19.8)=10.63, $p = .000$
Figure 21. Mean F0\textsubscript{MAX}, ΔF0, and slope ($m$) by Tone in Carrier Phrase 2. Central boxplot gives mean and distribution of F0\textsubscript{MAX} data for each Tone in normalized F0 for all speakers. Mean ΔF0 is shown at level of F0\textsubscript{MIN} and accompanying line indicates $m_{F0}$; precise $m_{F0}$ values provided in lower box.

(44) Subsets by post-hoc Scheffé test in Phrase 2: [Low __ Low]

<table>
<thead>
<tr>
<th>F0\textsubscript{MAX} (Hz)</th>
<th>ΔF0: Max-Min</th>
<th>$m_{F0}$ (Hz/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>C 234</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>H 209</td>
<td>?</td>
<td>C</td>
</tr>
<tr>
<td>193 L</td>
<td>.069</td>
<td>.201 .286 .364</td>
</tr>
</tbody>
</table>

Difference in pair, $p$

Tone $F(3, 27) = 47.7, p = .000$ $F(3, 27) = 38.77, p = .000$ $F(3, 27) = 14.24, p = .000$
Figure 22. Mean $F_{0\text{MAX}}$, $\Delta F_0$, and slope ($m$) by Tone in Carrier Phrase 3. Central boxplot gives mean and distribution of $F_{0\text{MAX}}$ data for each Tone in normalized $F_0$ for all speakers. Mean $\Delta F_0$ is shown at level of $F_{0\text{MIN}}$ and accompanying line indicates $m_{F_0}$; precise $m_{F_0}$ values provided in lower box.

Subsets by post-hoc Scheffé test in Phrase 3: [High __ Low]

<table>
<thead>
<tr>
<th>Tone</th>
<th>$F_{0\text{MAX}}$ (Hz)</th>
<th>$\Delta F_0$: Max-Min</th>
<th>$m_{F_0}$ (Hz/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>233</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>228</td>
<td>H</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>210</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>L</td>
<td>199</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Difference in pair, $p$

<table>
<thead>
<tr>
<th>Tone</th>
<th>$F(3,27)$ = 24.86, $p = .000$</th>
<th>$F(3,27) = 18.42, p = .000$</th>
<th>$F(3,27) = 2.36, p = .094$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>.687</td>
<td>.989</td>
<td>.251</td>
</tr>
<tr>
<td>H</td>
<td>.139</td>
<td>.516</td>
<td>.108</td>
</tr>
<tr>
<td>L</td>
<td>.989</td>
<td>.251</td>
<td>.108</td>
</tr>
<tr>
<td>?</td>
<td>.251</td>
<td>.108</td>
<td>.251</td>
</tr>
</tbody>
</table>
Figure 23. Mean $F_{0\text{MAX}}, \Delta F_0$, and slope ($m$) by Tone in Carrier Phrase 4. Central boxplot gives mean and distribution of $F_{0\text{MAX}}$ data for each Tone in normalized $F_0$ for all speakers. Mean $\Delta F_0$ is shown at level of $F_{0\text{MIN}}$ and accompanying line indicates $m_{F_0}$; precise $m_{F_0}$ values provided in lower box.

(46) Subsets by post-hoc Scheffé test in Phrase 4: [Low ___ High]

<table>
<thead>
<tr>
<th>$F_{0\text{MAX}}$ (Hz)</th>
<th>$\Delta F_0$: Max-Min</th>
<th>$m_{F_0}$ (Hz/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low 237</td>
<td>H L</td>
<td>H C</td>
</tr>
<tr>
<td>C 236</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H 216</td>
<td></td>
<td></td>
</tr>
<tr>
<td>194 L</td>
<td>.988</td>
<td>.111</td>
</tr>
</tbody>
</table>

Difference in pair, $p$

| Tone | $F(3, 27) = 52.76, p = .000$ | $F(3, 27) = 42.83, p = .000$ | $F(3, 27) = 26.49, p = .000$ |

$F_{0\text{MAX}}$ Hz, $\Delta F_0$ Hz, $m_{F_0}$ Hz/sec
Figure 24. Mean F0\text{MAX}, ΔF0, and slope (m) by Tone in Carrier Phrase 5. Central boxplot gives mean and distribution of F0\text{MAX} data for each Tone in normalized F0 for all speakers. Mean ΔF0 is shown at level of F0\text{MIN} and accompanying line indicates m_{F0}; precise m_{F0} values provided in lower box.

Subsets by post-hoc Scheffé test in Phrase 5: [High ___ #]

<table>
<thead>
<tr>
<th>Difference in pair, $p$</th>
<th>$F_{(3,19.7)}=9.05, p = .001$</th>
<th>$F_{(3, 20)} = 8.96, p = .001$</th>
<th>$F_{(3,20)} = 15.03, p = .000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone</td>
<td>$F_{(3,19.7)}=9.05, p = .001$</td>
<td>$F_{(3, 20)} = 8.96, p = .001$</td>
<td>$F_{(3,20)} = 15.03, p = .000$</td>
</tr>
<tr>
<td>C</td>
<td>1.415</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>C</td>
<td>1.415</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>C</td>
<td>1.415</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>L</td>
<td>0.832</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

(47)
Figure 25. Mean $F_{0\text{MAX}}$, $\Delta F_0$, and slope ($m$) by Tone in Carrier Phrase 6. Central boxplot gives mean and distribution of $F_{0\text{MAX}}$ data for each Tone in normalized $F_0$ for all speakers. Mean $\Delta F_0$ is shown at level of $F_{0\text{MIN}}$ and accompanying line indicates $m_{F_0}$; precise $m_{F_0}$ values provided in lower box.

(48) Subsets by post-hoc Scheffé test in Phrase 6: [Low __ #]

<table>
<thead>
<tr>
<th>Tone</th>
<th>$F_{0\text{MAX}}$ (Hz)</th>
<th>$\Delta F_0$: Max-Min</th>
<th>$m_{F_0}$ (Hz/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-0.114</td>
<td>-0.521</td>
<td>-0.403</td>
</tr>
<tr>
<td>?</td>
<td>-1.51</td>
<td>-1.122</td>
<td>-0.570</td>
</tr>
<tr>
<td>H</td>
<td>-0.414</td>
<td>-0.569</td>
<td>-1.339</td>
</tr>
<tr>
<td>L</td>
<td>-1.15</td>
<td>-1.60</td>
<td>-2.004</td>
</tr>
</tbody>
</table>

Difference in pair, $p$

<table>
<thead>
<tr>
<th>Tone</th>
<th>$F(3, 22) = 6.07, p = .004$</th>
<th>$F(3, 22) = 3.45, p = .034$</th>
<th>$F(3, 22) = 4.66, p = .011$</th>
</tr>
</thead>
</table>
Figure 26. Mean F0\textsubscript{MAX}, ΔF0, and slope (m) by Tone in Carrier Phrase 7. Central boxplot gives mean and distribution of F0\textsubscript{MAX} data for each Tone in normalized F0 for all speakers. Mean ΔF0 is shown at level of F0\textsubscript{MIN} and accompanying line indicates m\textsubscript{FO}; precise m\textsubscript{FO} values provided in lower box.

(49) Subsets by post-hoc Scheffé test in Phrase 7: [High \_ minor σ].

<table>
<thead>
<tr>
<th>Tone</th>
<th>F0\textsubscript{MAX} (Hz)</th>
<th>ΔF0: Max-Min</th>
<th>m\textsubscript{FO} (Hz/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
<td>244</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>C</td>
<td>C 237</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>H</td>
<td>230H</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>224</td>
<td>L</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>.386</td>
<td>.373 .607</td>
<td>- .633 .091</td>
<td>- .122 -</td>
</tr>
</tbody>
</table>

\(F(3, 27) = 4.83, p = .008\) \(F(3, 27) = 7.41, p = .001\) \(F(3, 27) = 6.22, p = .002\)
Figure 27. Mean F0\textsubscript{MAX}, ΔF0, and slope (m) by Tone in Carrier Phrase 8. Central boxplot gives mean and distribution of F0\textsubscript{MAX} data for each Tone in normalized F0 for all speakers. Mean ΔF0 is shown at level of F0\textsubscript{MIN} and accompanying line indicates m\textsubscript{F0}; precise m\textsubscript{F0} values provided in lower box.

(50) Subsets by post-hoc Scheffé test in Phrase 8: [Low ___ minor σ].

<table>
<thead>
<tr>
<th>Tone</th>
<th>F0\textsubscript{MAX} (Hz)</th>
<th>ΔF0: Max-Min</th>
<th>m\textsubscript{F0} (Hz/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>237</td>
<td></td>
<td></td>
</tr>
<tr>
<td>?</td>
<td>232</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>214</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>?</td>
<td>196</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>196</td>
<td>L</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>.839</td>
<td>.768</td>
<td>-.779</td>
<td></td>
</tr>
</tbody>
</table>

Difference in pair, \(p\)

Tone \(F(3,27) = 19.70, p = .000\) \(F(3, 28.2) = 19.3, p = .000\) \(F(3,29.2) = 7.19, p = .001\)
**Carrier Phrase (1) – Isolated Tokens** • Isolated utterances of each tone contained a strong two-way difference in all three statistical variables between the Creaky/Checked tones and High/Low tones. Tone had a significant effect for each variable by ANOVA ($p < .001$ for all three). The means and Scheffé results verify that the Creaky and Checked tones systematically had higher peaks as well as greater and steeper falls in F0. For all three measures, Creaky and Checked tone syllables were indistinguishable from one another.

High and Low tone means were visibly different, though statistical differences between the two were limited, perhaps due to the limited sample size and greater variation in the Low token data in isolated utterances. High tone syllables had a higher F0_{MAX} and a steeper negative $m_{F0}$ than the Low syllables, but not significantly so. The difference in mean ΔF0 between High ($M = 1.85z, SD = 1.39$) and Low tones ($M = 0.85z, SD = 1.77$) was still very significant ($p = .003$ by Scheffé pair comparison).

(43) Repeated from (43) above. Carrier Phrase (1) Statistical Results

Phrase 1: Isolated Syllable. Subsets by post-hoc Scheffé test.

<table>
<thead>
<tr>
<th>F0_{MAX} (Hz)</th>
<th>ΔF0: Max-Min</th>
<th>$m_{F0}$ (Hz/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
<td>L H</td>
<td>H L</td>
</tr>
<tr>
<td>C</td>
<td>H ?</td>
<td>? C</td>
</tr>
<tr>
<td>.823 .075</td>
<td>- - .828</td>
<td>.798 .940 -</td>
</tr>
</tbody>
</table>

Carrier Phrase (2) – [Low ___ Low] • Tokens in Carrier Phrase (2) also showed the difference by Tone to be statistically different in all three variables (at $p < .001$ for all

---

14 Recall that Low tone F0 values were not included for the speakers who used the late intonational rise ($§4.3.1$).
three). Scheffé test results revealed a strong distinction between Creaky/Checked and the High tone means – High tones rose in F0 while all others fell. The variation in Low tone falling contours created some surprising findings. As would be expected, Low tone had the lowest F0\text{MAX} by significant proportions, but had a negative slope that was indistinct from that of the Creaky tone fall, and a fall in F0 (ΔF0) that was not significantly different from Checked tones. The similarity in means is not indicative of a large fall in Low tones so much as an abbreviated F0 fall for Creaky and Checked tones. Both Creaky (\(M = -1.88\ z, -56hz\)) and Checked tones (\(M = -1.32\ z, -40hz\)) had firmly negative, but significantly different ΔF0 values (\(p = .002\) by Scheffé pair comparison).

(44) Repeated from (44) above. Carrier Phrase (2) Statistical Results
Phrase 2: [Low ___ Low]. Subsets by post-hoc Scheffé test.

<table>
<thead>
<tr>
<th>F0\text{MAX} (Hz)</th>
<th>ΔF0: Max-Min</th>
<th>(m_{F0}) (Hz/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>? 234</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>C 230</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>H 209</td>
<td>?</td>
<td>C</td>
</tr>
<tr>
<td>193</td>
<td>C</td>
<td>?</td>
</tr>
</tbody>
</table>

\(\text{Difference in pair, } p\)  
\(\text{Tone} \quad F(3, 27) = 47.7, p = .000\)  
\(F(3, 27) = 38.77, p = .000\)  
\(F(3, 27) = 14.24, p = .000\)

Carrier Phrase (3) – [High ___ Low] • Phrase (3) had similar effects to the Phrase (2) frame, but with a slightly higher initial mean F0 yielding greater F0 falls for all tones. The preceding High tone in Carrier Phrase (3) elevated the initial F0 on the following embedded tokens of all four tones. For this reason, slope was inadequate for differentiating statistically between the tones – an ANOVA did not find an overall difference in \(m_{F0}\) between the four tones \((F(3,27) = 2.36, p = .094)\). As expected, Creaky and Checked tones had the steepest negative slope, but not significantly more so than the Low tone.
One similarity to the data from [Low __ Low] frames (Phrase 2) was that the greatest $\Delta F_0$ (negative value) belonged to Creaky tones, while the mean $\Delta F_0$ of Low and Checked tones were not significantly different. Additionally, $\Delta F_0$ was the only variable that was significantly different between the Creaky and Checked tones, indicating that their overall pitch height and contour were again very similar. The $\Delta F_0$ difference was a product of the tones’ unequal durations, as an equivalent negative slope realized over the longer Creaky tone duration resulted in a greater overall drop.

(45) Repeated from (45) above. Carrier Phrase (3) Statistical Results
Phrase 3: [High __ Low]. Subsets by post-hoc Scheffé test.

<table>
<thead>
<tr>
<th>Tone</th>
<th>$F_{0\text{MAX}}$ (Hz)</th>
<th>$\Delta F_0$: Max-Min</th>
<th>$m_{F_0}$ (Hz/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>233</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>?</td>
<td>228</td>
<td>?</td>
<td>L</td>
</tr>
<tr>
<td>H</td>
<td>210</td>
<td>L</td>
<td>?</td>
</tr>
<tr>
<td>L</td>
<td>199</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>.687</td>
<td>.139</td>
<td>.989</td>
<td>.251</td>
</tr>
</tbody>
</table>

$F(3,27) = 24.86, p = .000$  $F(3, 27) = 18.42, p = .000$  $F(3, 27) = 2.36, p = .094$

**Carrier Phrase (4) – [Low __ High]** • In the [Low __ High] context of Phrase (4), all four tones averaged a falling F0 contour, however the falls here were statistically very distinct. Figure (23) above and the Scheffé results in (46) (repeated below) demonstrated.

All three variables found highly significant group differences by ANOVA ($p < .001$ for each), each metric clearly separating the steep fall of Creaky and Checked tones from the shallow fall found on High and Low tones. Differences were found between these steep and shallow subsets as well. The High and Low tones had significantly different mean $F_{0\text{MAX}}$ and $\Delta F_0$, but not $m_{F_0}$. Conversely, Creaky and Checked tones had very similar results for $F_{0\text{MAX}}$ and $\Delta F_0$ but a significantly different
The falling F0 of Creaky tones had a statistically greater negative mean slope. Overall, no single statistic found a four-way contrast in Phrase (4) data, though Maximum F0 and ΔF0 drew a clear distinction, grouping together the similar Creaky and Checked tones and regarding High and Low as significantly distinct.

(46) Repeated from (46) above. Carrier Phrase (4) Statistical Results
Phrase 4: [Low __ High]. Subsets by post-hoc Scheffé test.

<table>
<thead>
<tr>
<th>Tone</th>
<th>F0\text{MAX} (Hz)</th>
<th>ΔF0: Max-Min</th>
<th>m_{F0} (Hz/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>236</td>
<td>L</td>
<td>C</td>
</tr>
<tr>
<td>L</td>
<td>216</td>
<td>?</td>
<td>C</td>
</tr>
<tr>
<td>C</td>
<td>194</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

\[ \text{Difference in pair, } p = .988 \]

\[ \text{Tone } F(3, 27) = 52.76, p = .000 \]
\[ \text{ΔF0: Max-Min } F(3, 27) = 42.83, p = .000 \]
\[ m_{F0} (Hz/sec) F(3, 27) = 26.49, p = .000 \]

**Carrier Phrase (5) – [High __ #]** • Data for Carrier Phrase (5) was unevenly distributed since just three speakers (see §4.3.1-2) were tabulated for the Low tone values. Due in part to the scant Low tone data, statistical findings for Phrase (5) data were murky beyond the Checked tone, which had the highest F0, and the largest and steepest F0 drop. Because of the extreme Checked tone data, ANOVA results for Phrase (5) tokens showed a highly significant effect of Tone for each variable, F0\text{MAX} \( p = .001 \), ΔF0 \( p = .001 \), m_{F0} \( p < .001 \).

Excluding the Checked tone data, the three other tones were scarcely distinct in F0\text{MAX} and ΔF0. Only slope \( m_{F0} \) distinguished the shallow falls of Low and High tones from the more abruptly falling contour of the Creaky tone. The unclear boundaries between F0 results speak to the pitch raising effect of the preceding High tone – initial F0 was high on embedded tokens of all four tones, thereby precipitating a comparatively steeper fall from this early high. As in other contexts, the Creaky and Checked tones in
Phrase (5) were not significantly different by any variable. The table below repeats the statistical details seen above in Figure 24 and (47).

(47) Repeated from (47) above. Carrier Phrase (5) Statistical Results
Phrase 5: [# High __ #]. Subsets by post-hoc Scheffé test.

<table>
<thead>
<tr>
<th></th>
<th>$F_{0\text{MAX}}$ (Hz)</th>
<th>$\Delta F_0$: Max-Min</th>
<th>$m_{F_0}$ (Hz/sec)</th>
<th>Tone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>?</td>
<td>1.951</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>1.415</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>1.123</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>0.832</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Difference in pair, $p$</td>
<td>.165</td>
<td>.112</td>
<td>.249</td>
<td>.761</td>
</tr>
<tr>
<td></td>
<td>.115</td>
<td>-</td>
<td>.295</td>
<td>.196</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>.212</td>
<td>.512</td>
<td>-</td>
</tr>
</tbody>
</table>

$F(3,19.7)=9.05, p =.001$ $F(3, 20) = 8.96, p = .001$ $F(3,20) = 15.03, p=.000$

**Carrier Phrase (6) – [# Low __ #]** As in Isolated and the other phrase-final data, Phrase (6) analysis was likewise limited to a subset of the speaker pool for Low tone means – five of the ten speakers were counted (see §4.3.2). An ANOVA on each of the three statistical variables with Tone as the main factor and speaker as a random effect showed a significant effect of Tone on each variable (details in (48)). Between-group distinctions by Scheffé tests were muddled by the across-the-board depression of $F_0$ in this frame. The means in Figure 25 revealed a prominent contrast from Phrase (5) data in that initial $F_0$ values were dramatically lower, consistent with the generally lower $F_0$ throughout the embedded token.

(48) Repeated from (48) above. Carrier Phrase (6) Statistical Results
Phrase 6: [# Low __ #]. Subsets by post-hoc Scheffé test.

<table>
<thead>
<tr>
<th></th>
<th>$F_{0\text{MAX}}$ (Hz)</th>
<th>$\Delta F_0$: Max-Min</th>
<th>$m_{F_0}$ (Hz/sec)</th>
<th>Tone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>-.014</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>?</td>
<td>-.106</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>-.414</td>
<td>?</td>
<td>-1.49</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-1.15</td>
<td>C</td>
<td>-1.60</td>
</tr>
<tr>
<td>Difference in pair, $p$</td>
<td>.115</td>
<td>-</td>
<td>.125</td>
<td>.625</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>.212</td>
<td>.512</td>
<td>-</td>
</tr>
</tbody>
</table>

$F(3,22) = 6.07, p =.004$ $F(3, 22) = 3.45, p = .034$ $F(3, 22) = 4.66, p = .011$
Still, the same tendencies from other contexts were noticeable in the ranking for each statistic. Creaky and Checked tones had the highest peak F0, the greatest negative Δ, and the steepest slope, but were at times not significantly greater than the High tone, which also fell from a depressed initial F0. Finally, the Creaky and Checked tones were again not different from one another in Phrase (6) for any of the F0 variables examined.

**Carrier Phrase (7) – [High __ minor σ]** • Phrase (7) results show a large effect of a preceding High tone on the initial F0 and overall contour for every tone – similar to what was found for other [High __ …] initial frames (Phrases 3, 5). Statistical results in Figure 26 revealed a similarity between the Creaky, Checked tones and the Low tone, where the fall from the early peak is similar in size (ΔF0) and rate (mF0). High tone means however, portrayed a much shallower fall that was statistically distinct. Accordingly, an ANOVA on each variable with Tone as the main factor and Speaker as a random effect found an overall difference between the four tones (results in (49)).

\[(49)\] Repeated from (49) above. Carrier Phrase (7) Statistical Results
Phrase 7: [High __ minor σ]. Subsets by post-hoc Scheffé test.

<table>
<thead>
<tr>
<th>Tone</th>
<th>F0MAX (Hz)</th>
<th>ΔF0: Max-Min</th>
<th>mF0 (Hz/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
<td>244</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>C</td>
<td>C 237</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>H</td>
<td>230H</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>224</td>
<td>L</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>.386</td>
<td>.373</td>
<td>.607</td>
<td>.633 .091</td>
</tr>
<tr>
<td>.091</td>
<td>.122</td>
<td></td>
<td>.002</td>
</tr>
</tbody>
</table>

The average Creaky contour possessed the greatest negative change in F0 (mean Δ = 2.85Hz, -86Hz) and the steepest downward slope (mean mF0 = -2.163, -65Hz/100ms). Still the Creaky and Checked samples did not have not significantly different means for
either of these variables, nor for F0_{MAX}. The High tone means were distinct for all three variables, having the smallest and most gradual F0 fall, reflecting the High tone tendency to remain fairly high-pitched while Creaky, Checked, and Low tokens all fell considerably from their early F0 peaks\textsuperscript{15}.

The behavior of Low tones in Phrase (7) was conspicuously different from in other frames – for Low tone tokens in any Carrier Phrase, the mean F0_{MAX} was the greatest, and ΔF0 and m_{F0} means had the largest negative values. This change was most evident when Phrase (7) and Phrase (8) Low tone data were compared. Without a preceding High tone, Phrase (8) data showed a smaller F0 drop and slope.

**Carrier Phrase (8) – [Low __ minor σ]** • In Phrase (8) data, Creaky and Checked tones were distinct from the High tone in all measures of F0, but not from the Low tone. The table in (50) reveals how the variables reflecting contour (ΔF0 and m_{F0}) did not distinguish Low from Checked and Creaky tone values, which seemed to fall more moderately (smaller F0 drop, shallower slope) in this environment.

\textsuperscript{15} The mean F0 for each tone in Phrase (7) peaked at the vowel onset. Inspection of the contours in Figure 18 suggests that not only was the F0 maximum early in the syllable, but that F0 was actively falling from a higher peak in the prior syllable.
As in other Phrases, the Creaky and Checked tones were nearly identical but for the overall change in F0, for which the same rate of change over a longer period produced a greater average fall for Creaky tokens.

The data for both Carrier Phrases (7) and (8) illustrate a point about the inability of the statistical variables $\Delta F_0$ and $m_{F0}$ to capture a qualitative difference between contours. Visually, the falling contour of the Low tones in Figures 18-19 was not the same as those of the falling Creaky and Checked tones. Low tone mean F0 dropped smoothly from an initial high which occurred early in the vowel before flattening out over the remaining portion of the rhyme. In Phrase (7), co-articulation from the preceding High tone likely heightened the initial F0 (explored further in §5.6.3), while the onset F0 of Phrase (8) data raised only slightly before quickly dissipating. Small F0 perturbation such as this is likely incurred by the initial voiceless stop onset, rather than neighboring tones. By comparison, the average fall attested in Creaky and Checked tones was more linear, occurring initially and then continuing over the course of the syllable. This contrast points to a difference in the nature of the early high: high F0 on Low tone syllables was pinpointed at the vowel onset, while Creaky and Checked tones maintained a relatively high F0 beyond the onset for roughly the first half of the syllable. Statistically though, the overall change in F0 and the slope between the maxima could not capture this difference.
4.3.6  Summary of F0 Data

Reviewing all data from every Carrier Phrase, a number of generalizations can be made concerning F0 production that fall into two sets. The first set, listed in (51 – 53) reflects the properties of each tone, while the second set (54 – 57) concerns the effects that the different sentence contexts had on F0, whether induced by the tone of an adjacent syllable, by speaking rate, or by phrase boundaries.

<table>
<thead>
<tr>
<th>F0 Properties of Each Tone</th>
</tr>
</thead>
<tbody>
<tr>
<td>(51) <strong>Low tones</strong> have a primarily flat contour at a comparatively low F0. Elevated pitch is only found with Low tones…</td>
</tr>
<tr>
<td>i) in the transition from the syllable onset.</td>
</tr>
<tr>
<td>ii) initially following a High tone syllable.</td>
</tr>
<tr>
<td>iii) finally with a certain intonational melody.</td>
</tr>
<tr>
<td>(52) <strong>High tones</strong> have an F0 peak with a context-dependent locus.</td>
</tr>
<tr>
<td>High tone F0 either…</td>
</tr>
<tr>
<td>i) falls in isolation or phrase-finally.</td>
</tr>
<tr>
<td>ii) rises medially between Low or toneless syllables.</td>
</tr>
<tr>
<td>iii) remains high before another High tone.</td>
</tr>
<tr>
<td>iv) has a lower overall F0 after another High tone.</td>
</tr>
<tr>
<td>(53) <strong>Creaky and Checked tones</strong> always had a falling F0, and always bore the steepest falls.</td>
</tr>
<tr>
<td>i) With a single exception(^\text{16}), Creaky and Checked tones had the greatest negative ΔF0 in every context.</td>
</tr>
<tr>
<td>ii) In all medial contexts the Creaky and Checked tones still bore the steepest negative slope – though not significantly in every Phrase.</td>
</tr>
<tr>
<td>iii) The F0 falls were weakened in juncture, particularly for Checked syllables.</td>
</tr>
</tbody>
</table>

\(^{16}\) In Phrase (3), ΔF0 for Low tones \((M = -2.03z, -61Hz)\) was not significantly lower than for Creaky tones \((M = -1.97z, -59Hz)\).
Effects of Context on F0

(54) **Anticipatory Co-articulation** was minimal. Tokens embedded before a High tone syllable exhibited little to no co-articulation to the following F0 peak.

(55) **Perseverative Co-articulation** was robust at the vowel onset. The elevated F0 of a preceding High tone syllable consistently raised the following syllable’s onset F0, regardless of tone. However, overall F0 on post-High syllables tended to be lower (falling contours fell further). In summary, after a High tone…
   i) The onset F0 of the following syllable is higher.
   ii) The offset F0 on the following syllable is lower.

(56) **Similarity between Creaky and Checked Tone data.** Across all Carrier Phrases, the Creaky and Checked syllables exhibited the same F0 behaviors.

(57) **Partial Neutralization of F0 fall in medial forms.** Falling contours were greater – both in total F0 drop and in the rate of the fall – when followed by a pause (either phrase-final or in isolation). When followed in close juncture, the falls were considerably weaker.

The generalizations in (54) – (57) are explored further with more specific statistical tests in the analyses of Chapter Five.

Looking at the statistical variables used to measure F0 contours, it is evident in the ANOVA results above that no single metric captures the four-way distinction between tones. In many Phrases, even the most discerning statistic would not be able distinguish Creaky from Checked F0 patterns, since by all appearances they are the same. Still, accounting for a three-way divide between Low, High, and Creaky/Checked tones is
possible, and is best done using the variables $F_{0\text{MAX}}$ and $\Delta F_0$ as criteria. This is the case across phonetic environments, which can generally be separated into two groups regarding $F_0$: phrase-medial forms and phrase-final forms (including isolated utterances). Two tables (58, 59) demonstrate the distinct $F_0$ qualities that form a three-way split in each group.

(58) F0 Cues to Tone in Isolated and Phrase-Final Tokens

<table>
<thead>
<tr>
<th>$\Delta F_0$</th>
<th>Low-Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive or slightly negative</td>
<td>Low Tone</td>
<td></td>
</tr>
<tr>
<td>Firmly negative</td>
<td>High Tone</td>
<td>Creaky or Checked Tone</td>
</tr>
</tbody>
</table>

(59) F0 Cues to Tone in Phrase-Medial Tokens

<table>
<thead>
<tr>
<th>$\Delta F_0$</th>
<th>Low-Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive or slightly negative</td>
<td>High Tone</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>Low Tone</td>
<td>Checked Tone</td>
</tr>
<tr>
<td>Firmly negative</td>
<td>High Tone</td>
<td>Creaky or Checked Tone</td>
</tr>
</tbody>
</table>

In most cases, the $F_0$ maximum distinguishes the very high Creaky and Checked tones from the less high Low and High tones. $\Delta F_0$ differences can then separate the individual tones in these subsets, though the interpretation of $\Delta F_0$ relies on the phrase position of the tone. The different context-based conditions were the product of two major findings of this chapter: the alternation between a falling High tone phrase-finally and even or

---

$^{17}$ Creaky and Checked tone samples had a statistically greater (and distinct) mean $F_{0\text{MAX}}$ by Scheffé post-hoc test in Phrases 1, 2, 3, 4, and 8. In the three phrases they were not statistically greater than the High tone (Phrases 5, 6, and 7) they still had the highest $F_{0\text{MAX}}$ mean values.
rising High tone phrase-medially (see 52), and the partial neutralization of falling contours in medial position, particularly for Checked tones (see 53iii).

A key characteristic of all of the F0 data was how generally consistent the contours were for Low, Creaky, and Checked tones. Creaky and Checked tones always began high and fell. Low tones (other than the intonational rise found with some speakers) were effectively low and even, after an initial descent from an onset for which F0 height was conditioned by the tonal pitch peaks of the prior syllable. Only the High tone exhibited wholesale changes to the contour shape depending on context, rising after Low tones in Phrases (2), (4), and (8) and gradually falling in citation form (Phrase 1) and after another High tone in Phrases (3), (5), and (7). This is not to say that differences were not found between phrase contexts in Low, Creaky, and Checked tone data, but that they were fairly minor differences which never unambiguously altered the general pitch profile of each tone: Low as low-pitched and flat, Creaky as a sharp fall, and Checked also as a sharp fall.
4.4 Spectral Profile: Comparative Amplitude of Harmonics

The following three sections present data addressing issues of phonation. It was shown above that the four tones are reliably produced with context-dependent distinctions in both Duration and Fundamental Frequency. However, the association of phonation properties with each tone, restated in the continuum in Figure 28, is the concern of the first four hypotheses proposed in Chapter Three (H1 through H4). These associations are at last tested with the data below from the acoustic record (this section), EGG signal (§4.5), and airflow trace (§4.6).

![Figure 28. Continuum of phonation mode as defined by glottal state](image)

4.4.1 Overview of Statistical Analysis of Spectral Tilt and Presentation of Results

Looking first at acoustic analysis of phonation, measures of spectral tilt in this study were unsuccessful at distinguishing the four tones for all but a limited set of cases. Specifically, only temporally specific measures of H1-A1 and H1-A3 from late in the vowel detected differences linked to phonation, and only in non-medial positions. All temporal measures of H1-H2 were unable to distinguish the contrastive tones. These findings matched the weak and inconsistent results for spectral tilt analysis in previous studies, even though potential methodology flaws had been accounted for: measures were taken vowel-terminally rather than at the mid-point and more than one (potentially
neutralizing) context was examined. Context proved to be critical as successful H1-A1 and H1-A3 results were found just in phrase positions that also produced the firmest confirmation of contrastive phonation in the EGG and airflow data. Simply put, the measures were most successful when it seems the contrast was most apparent. For all measures of phonation, these were syllables in isolation or embedded phrase-finally. In phrase-medial tokens where EGG results less consistently signaled phonation differences, all fifteen measures of spectral tilt were inconclusive.

The set of all harmonic values was normalized for each speaker, providing a z-score scale that reflected variation from each individual’s average H1, H2, A1, and A3 amplitudes. For each normalized amplitude difference, $z_{H1-H2}$, $z_{H1-A1}$, $z_{H1-A3}$, five temporally determined measures represented each token. Data were pooled for all speakers and analyzed for each Carrier Phrase. Three types of charts presenting these data are provided in select cases to illustrate noteworthy results – a full array of the charts for each Carrier Phrase can be found in the appendix. A scatterplot gives every 80th and 90th percentile data point as well as the mean “Late” amplitude difference (the by-token average of spectral tilt values at the 80%, 90%, and 100% deciles). This display gives a clear sense of the overlapping distribution of spectral tilt values between tones, since the coefficient of variation was often rather large with these figures.18 Two sets of boxplots give the median and distribution of values for the averaged “Early” and “Late” harmonic differences. Finally, the change (as $\Delta$) in spectral profile19 between these time frames is shown for each tone.

---

18 E.g. $M = -0.09$, $SD = 1.02$, $CV = 1.133\%$ for H1-H2 on Checked syllables in Phrase 2 [Low __ Low]).
19 Calculated on a by-token basis rather than simply the difference between the mean EARLY and mean LATE amplitude differences.
Each spectral measure within each Carrier Phrase was analyzed as the dependent variable in a linear mixed model with Tone as a factor and Speaker as an error term. Viewed collectively, these data express the effects of tone and sentence context on the spectral profile of different tone-bearing vowels.

Including Checked tone spectral tilt results with the other tones of course introduced the confound of a distinct vowel quality on one tone. This confound was minimized to the extent it could be by the stimuli selection of the vowel [a]. The closed-syllable, Checked low vowel (reported variously as IPA [a] (Watkins 2001) or IPA [ʌ] (Mehnert & Richter 1972-1977) was regarded as more approximate to its open-syllable counterpart (cf (1.7) in Chapter 1; compare open [ɔ] ~ closed [ ao]). Summaries of the statistical findings for spectral tilt are presented in such a way that the influence of the Checked vowel quality is partially suppressed.

In the tables of statistical results throughout §4.4 all significant differences are reported for each variable in each Carrier Phrase context, with the intent of highlighting the measurements and phrases which best distinguish High, Low, and Creaky (and not Checked) tone data. The charts flag cases where a global statistical difference is unquestionably the result of a shallower spectral tilt with the Checked tone – High, Low, and Creaky tones are indistinguishable by pair-wise comparison in these instances.

To demonstrate the presentation of statistical results, a key is supplied in (60) followed by a sample portion of data in (61) from the H1-H2 results for two phrase contexts. Each cell in the table tells whether there was a significant effect of Tone for a measure within a select Carrier Phrase. Non-significant effects are shown in an unshaded cell with only the $p$-value listed. Significant effects are shaded differently according to
how strongly the given measure supported the three-way phonation type divide between breathy, creaky, and modal voicing. Generally, the stronger the results of the statistical test, the darker the cell is shaded. According to convention, a single asterisk * denotes statistical significance at the $p < .05$ level, and a double asterisk ** denotes significance to the level of $p < .01$. The key in (60) lists each possibility and other aspects of notation.

<table>
<thead>
<tr>
<th>(60)</th>
<th>Key to Presentation of Spectral Tilt Statistical Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>* Dark Grey</td>
<td>Significant difference involving all four tones.</td>
</tr>
<tr>
<td>* Light Grey</td>
<td>Significant difference involving subset of tones. High, Low, or Creaky tone not included.</td>
</tr>
<tr>
<td>* Diagonal</td>
<td>High, Low, and Creaky tones are not different.</td>
</tr>
<tr>
<td>* Cross out</td>
<td>Significant difference exists, but is antithetical to $H1$-$H4$.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>H</th>
<th>High</th>
<th>*</th>
<th>$p &lt; .05$</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>Low</td>
<td>**</td>
<td>$p &lt; .01$</td>
</tr>
<tr>
<td>C</td>
<td>Creaky</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>Checked</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Only significant differences in a dark grey shaded cell reflect a statistically identifiable three-way split in spectral tilt between a breathy High tone, a modal Low tone, and a creaky-voice Creaky tone. A light grey cell indicates that the means for all four Tones fit the expected hierarchy of (High > Low > Creaky, Checked), but that statistical significance was only found for some of the differences, and not each pair. A diagonally-crossed cell indicates that while there was a significant effect of Tone, it was exclusively due to an exceptional Checked tone mean.

---

20 Therefore, a light grey cell will not list all four tones. The unlisted tones were not significantly higher or lower by pair-wise comparison than those listed in the cell. For instance, in the table in (61), the measure $H1$-$H2$ at the 90% decile reports "** $H, L > K". In this case, the Creaky tone was not significantly different from either group in this context.
Additional information may also appear in some cells, in order to demonstrate the precise nature of the statistical difference when a three-way split was not the case. The letters H, L, C, and K are used to indicate the tones and “>” to indicate a significantly greater mean. For example, in (61) below, the Phrase (5) data for H1-H2<sub>EARLY</sub> report that Low and High tone means were greater than Creaky and Checked means, but only significantly greater than the Checked tone (notated as K), and not the Creaky (notated as C). The cell also states that Low tone means were higher than High tone means (by virtue of the order “L, H”), though not by a significant margin. Statistical groupings by mean are only provided when the significant difference supported the hypothesized Tone-to-Phonation associations.

(61)  Sample F-test Results, Mixed Model Analysis of H1-H2

<table>
<thead>
<tr>
<th></th>
<th>H1-H2&lt;sub&gt;EARLY&lt;/sub&gt;</th>
<th>H1-H2&lt;sub&gt;80%&lt;/sub&gt;</th>
<th>H1-H2&lt;sub&gt;90%&lt;/sub&gt;</th>
<th>H1-H2&lt;sub&gt;LATE&lt;/sub&gt;</th>
<th>ΔH1-H2</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.</td>
<td>[Low__High]</td>
<td>.169</td>
<td>.236</td>
<td>* .002</td>
<td>* .001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C, L &gt; H</td>
<td>C, L &gt; H</td>
<td>.080</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>[High__#]</td>
<td>** .001</td>
<td>* .033</td>
<td>* .026</td>
<td>.079</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L, H &gt; K</td>
<td>L &gt; K</td>
<td>H, L &gt; K</td>
<td>.830</td>
</tr>
</tbody>
</table>

Also marked separately are significant differences that were incompatible with the four Tone-to-Phonation hypotheses. For these measures, a statistically significant finding is reported, but the cell is left unshaded and it is crossed out. In the example in (61), the means for H1-H2<sub>LATE</sub> in Phrase (4) were significantly different between tones ($F(3,72) = 6.27, \ p = .001$) primarily because of the difference between the small H1-H2 score for the High tone and the dramatically steeper spectral tilt for the Creaky tone. These results fail to support either Hypothesis One (High tone is breathy) or Hypothesis Three (Creaky tone is creaky).
To review, fifteen measures of spectral tilt (listed in (5) – (19) at the beginning of this chapter) were analyzed in a linear mixed model so as to define the effect of Tone on the dynamic spectral profile of tone-bearing vowels. Results of the statistical tests are given in the following three sections in tables that state (i) the $p$-value that there was an overall difference between the four Tone means, as well as (ii) the ranking of individual Tone means against one another by pairwise comparison. The key in (60) and description above is intended to facilitate the interpretation of these data in sections 4.4.2 – 4.4.4.

4.4.2 $H1$-$H2$

A comparison of $H1$-$H2$ scores for each tone did not present a cohesive difference, neither overall nor within each tested context. Only in a single context (Carrier Phrase 5), did the data trend in the proper direction for the four Tone-to-Phonation hypotheses ($H1$ to $H4$): High > Low > Creaky, Checked. Yet statistical tests scarcely supported this three-way division. In other contexts, the highest or lowest $H1$-$H2$ scores were inconsistently distributed amongst the tones and the differences were rarely significant. Often the highest mean $H1$-$H2$ values belonged to the Creaky tone (though not significantly), running counter to the proposed alignment of tone to phonation types, which would predict lower scores with greater glottal constriction. Results of the mixed model analysis for each Carrier Phrase appear in (62) below.

Tokens in Carrier Phrase (5) provided the results most in-line with the associations of **Hypotheses One** ($H1$) through **Four** ($H4$). Checked tone syllables had the lowest $H1$-$H2$ scores for all statistics, and were significantly lower than the non-
glottalized tones for H1-H2<sub>Early</sub>, H1-H2<sub>80%</sub>, and H1-H2<sub>90%</sub>. Furthermore, while H1-H2 values for the Creaky tone were not statistically different from the High and Low data, the Creaky mean was lower. Thus, the phrase-final embedded position of Phrase (5) yielded the best-fitting H1-H2 data for the hypothesized phonation associations – a lesser H1-H2 for Creaky and Checked syllables suggests greater laryngeal constriction with their production in this phrase-final context.

In other Carrier Phrases, four of the five statistics were insignificant, with only the Early measures of H1-H2 routinely showing an effect of Tone. However, the pairwise comparisons show that the effect was mostly the result of the distance between Checked tone values and the High and Low tone means. This finding is better attributed to the Checked tone vowel quality as Checked vowels regularly yielded the lowest F1 and corresponding smallest amplitude differences across all contexts (differences which were even larger in the terminal positions of Phrases 1, 5, and 6).

\[(62)\]  

<table>
<thead>
<tr>
<th>Phrase</th>
<th>H1-H2&lt;sub&gt;Early&lt;/sub&gt;</th>
<th>H1-H2&lt;sub&gt;80%&lt;/sub&gt;</th>
<th>H1-H2&lt;sub&gt;90%&lt;/sub&gt;</th>
<th>H1-H2&lt;sub&gt;Late&lt;/sub&gt;</th>
<th>ΔH1-H2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Isolated</td>
<td><strong>p = .001</strong></td>
<td>.513</td>
<td>.321</td>
<td>.248</td>
<td>.091</td>
</tr>
<tr>
<td>2. [Low__Low]</td>
<td>*.042</td>
<td>.125</td>
<td>.383</td>
<td>.417</td>
<td>.145</td>
</tr>
<tr>
<td>3. [High__Low]</td>
<td>*.039</td>
<td>.943</td>
<td>.430</td>
<td>.721</td>
<td>.073</td>
</tr>
<tr>
<td>4. [Low__High]</td>
<td>.169</td>
<td>.236</td>
<td>*.002</td>
<td>*.001</td>
<td>.080</td>
</tr>
<tr>
<td>5. [High__#]</td>
<td><strong>.001</strong></td>
<td>*.033</td>
<td>*.026</td>
<td>.079</td>
<td>.830</td>
</tr>
<tr>
<td>6. [Low__#]</td>
<td>.052</td>
<td>.696</td>
<td>.325</td>
<td>.773</td>
<td>.153</td>
</tr>
<tr>
<td>7. [High_minor]</td>
<td>*.039</td>
<td>.139</td>
<td>*.024</td>
<td>*.009</td>
<td>.426</td>
</tr>
<tr>
<td>8. [Low_minor]</td>
<td>*.001</td>
<td>.227</td>
<td>.837</td>
<td>.503</td>
<td>.091</td>
</tr>
</tbody>
</table>

**F-test Results, Mixed Model Analyses of Five Statistics of H1-H2**
The distribution of data in Phrase (5) tokens is offered in Figure 29 and the means for H1-H2<sub>EARY</sub>, H1-H2<sub>LATE</sub>, and ΔH1-H2 appear in Figures 30 and 31. Though Phrase (5) tokens yielded the most robust statistical results for H1-H2, these data hardly support any division according to phonation type, whether as a tripartite Breathy > Modal > Creaky divide or simply a two-way split in behavior between the non-glottalized High/Low tones and the Creaky/Checked tones on the other side. As a point of comparison, the EARLY and LATE means for Phrase (1) data are included in Figure 32. No pattern is evident in any of these figures, for Phrase (1) or (5).

![Figure 29](image.png)

**Figure 29.** Plot of all normalized H1-H2 values at 80<sup>th</sup> and 90<sup>th</sup> decile of vowel duration. Mean H1-H2<sub>LATE</sub> is plotted adjacent to individual data points. The left y-axis is z-scaled; the right y-axis is retransformed to dB for illustrative purposes according to a single speaker’s standardization parameters<sup>21</sup>.

---

<sup>21</sup> Speaker I (male), for all H1-H2 data points: $M = 2.78$ dB, $SD = 3.96$.  

187
To briefly add to the discussion of unsubstantial findings, consider also the results posted in (62) for Phrase (6) data, where no consistent pattern is to be found. The mixed model analysis offers no reason to suspect that tones bore different effects on H1-H2 readings. Only H1-H2\textsubscript{EARLY} approached significance ($F(3,74) = 2.695, p = .052$) because Low tone tokens had a distinctly high H1-H2 mean. Otherwise, no differences are significant and
the order of means ran counter to the proposed ranking. That is, the High tone was not the next highest, but instead had the lowest average $H1-H2_{\text{EARLY}}$.

![Figure 32. Distribution of Early and Late H1-H2 values across all tokens for all speakers.](image)

To summarize, H1-H2 was unable to correlate to any type of phonation distinction in the data. The data for Phrase (5) were the most promising, where High and Low tones had greater H1-H2 means than the Creaky and Checked tones across all measures. However, only Checked H1-H2 values were statistically different, as the difference between Creaky and High/Low means was not great or consistent enough in any measure of the Phrase (5) data. In Phrases (4), (6), and (7), a number of statistics showed a significant effect of Tone, but not in the direction that would indicate creaky voicing on Creaky and Checked tokens or modal or breathier voicing with High or Low tone tokens. Rather, for these data the Creaky tokens often had the greatest mean difference in harmonic amplitudes – indicating the least creaky phonation during the tone
most associated with creakiness. Finally, $\Delta H_1-H_2$ did not correspond in any way to an increase in $H_1-H_2$ over the course of High and Low tone syllables – neither in Carrier Phrase (5) nor in any other recorded environment.

Recall that prior studies examining spectral profiles either reported $H_1-H_2$ amongst many measures which did not correlate well with phonation type (Watkins 1997) or used $H_1-H_2$ exclusively without success (Gruber & Feizollahi 2006). The above findings suggest that the problem encountered in these studies was not the result of measuring the wrong part of the vowel or only collecting tokens in a phonation neutralizing Low __ Low medial context. Rather, it seems that $H_1-H_2$ is simply a poor metric for detecting the Burmese tonal distinction. This discrepancy is explored at greater length after the review of $H_1-A_1$ and $H_1-A_3$ data in §4.4.6.

4.4.3 $H_1-A_1$

Spectral measures using higher frequency landmarks (F1, F3) were more successful, with one foreseeable drawback – the vowel quality in Checked syllables. Checked [a] vowels regularly had a lower F1, which increases the likelihood of a greater amplitude $A_1$ and in turn promotes a smaller $H_1-A_1$ difference. Comparisons of $H_1-A_1$ values are therefore relatively uninformative about Checked tone phonation, as consistently lower Checked values merely reflect the tone’s unique vowel quality.

Fortunately, a number of other strong contrasts are apparent in the $H_1-A_1$ data in certain sentence contexts, as tested by a linear mixed model with Tone as a fixed factor and Speaker as an error term. Consider the findings in (63) for isolated tone-bearing
syllables (Phrase 1) – all five measures of H1-A1 report a statistically significant
difference between group means. Pair-wise comparisons of the tone group means show a
division into two groups: High and Low tones together and consistently lower values
found with Creaky and Checked tones. The Creaky tone was lower than the High and
Low tones, but not significantly so, for both H1-A1_{EARLY} and H1-A1_{80\%}. The charts in
Figures 33 – 34 permit a closer inspection of the means providing the Phrase (1) results.
Similar effects of Tone were found in the phrase-final position of Phrases (5) and (6),
which are discussed further below.

(63)  *F*-test Results, Mixed Model Analyses of Five H1-A1 Statistics

<table>
<thead>
<tr>
<th>Phrase 1. Isolated</th>
<th>H1-A1_{EARLY}</th>
<th>H1-A1_{80%}</th>
<th>H1-A1_{90%}</th>
<th>H1-A1_{LATE}</th>
<th>ΔH1-A1</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.  [High–Low]</td>
<td>* .045 L &gt; K</td>
<td>** .005 L &gt; H, K</td>
<td>** &lt; .001 C &gt; K, H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.  [High–#]</td>
<td>.377</td>
<td>** &lt; .001 L,H &gt; C,K</td>
<td>** &lt; .001 L,H &gt; C,K</td>
<td>** &lt; .001 L,H &gt; C,K</td>
<td>** &lt; .001 L,H &gt; C,K</td>
</tr>
<tr>
<td>6.  [Low–#]</td>
<td>** .003 L &gt; K, C</td>
<td>** &lt; .001 H,L,LC&gt;CK</td>
<td>** &lt; .001 H,L &gt; C,K</td>
<td>** &lt; .001 H,L &gt; C,K</td>
<td>** &lt; .001 H &gt; C, K</td>
</tr>
<tr>
<td>7.  [High_minor]</td>
<td>** .004 HLC &gt; LCK</td>
<td>** .002 H,L,C &gt; K</td>
<td>** &lt; .001 HLC &gt; CK</td>
<td>** &lt; .001 HLC &gt; K</td>
<td>.303</td>
</tr>
<tr>
<td>8.  [Low_minor]</td>
<td>* .032 H &gt; K</td>
<td>.385</td>
<td>.099</td>
<td>* .031 All pairs ns</td>
<td>.490</td>
</tr>
</tbody>
</table>

The results for Phrase (1) match a number of predictions from Chapter Three and support
some claims of hypotheses H1 – H4. Note however that the High and Low tones are
never significantly different, which could be taken to support either Hypothesis One
(High tone is breathy, as is the Low tone) or Hypothesis Two (both High and Low use
modal voicing), but not both. Without taking into account other phonetic correlates, this
movement to significantly higher late values points to a late breathy voice quality on vowels with these tones. In the terms of a hypothetical speaker representative of the pooled data\textsuperscript{22}, the change is from a roughly 2dB difference early in the vowel to 9dB difference at the 80-90% deciles. The increase reflects a change to weaker high

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{frame1.png}
\caption{Distribution of Early and Late H1-A1 values across all tokens for all speakers.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{frame2.png}
\caption{Mean $\Delta$H1-A1 for each tone. $\Delta$ represents the difference between H1-A1\textsubscript{LATE} and H1-A1\textsubscript{EARLY} in each token.}
\end{figure}

\textsuperscript{22} Speaker I (male), for all H1-A1 data points: $M = 4.915$dB, $SD = 4.1$. 

192
harmonics (around F1) caused by a laxer or more open glottal state late in the vowel. Regardless of whether this late quality is breathy or modal, it is evident that Creaky and Checked tones produce a different spectral quality – in both static and dynamic measures.

Tone-bearing syllables produced in phrase-medial contexts lost much of this contrast. Data for Carrier Phrases (2), (3), and (4) showed none of the differences between High, Low, and Creaky H1-A1 values in any measure. There was a weak tendency for Checked tones to have the smallest (non-significant) H1-A1 score. As this was nearly an across-the-board trait in every Carrier Phrase context, this marginal difference is better attributed to Checked ‘A1’ amplitude readings being taken from a lower frequency formant than for the other three tones. Figures 35 and 36 give the means for H1-A1_{EARLY}, H1-A1_{LATE}, and ΔH1-A1 as a demonstration of the lack of contrast between the Tone effects in the multiple phrase-medial contexts.

![Figure 35. Distribution of Early and Late H1-A1 values across all tokens for all speakers.](image-url)
The phrase-final tokens in Carrier Phrases (5) and (6) revealed a pattern much like that seen for isolated syllables (Phrase 1), with similar statistical results in tow. There was a strongly significant two-way difference in H1-A1₀%, H1-A1_LATE, and ΔH1-A1 between the very large amplitude differences for High and Low tones and the small H1-A1 values found on Creaky and Checked tones.

In the full results for these two phrases in Figures 37 – 42, it is worthwhile to compare the scale of these between-group differences with those in Figures 35 and 36. The striking difference seen in the set of late vowel H1-A1 means in Figure 38 is plainly of a different type than in Figure 35. The amplitude differences seen in Phrase (1, 5, 6) H1-A1_LATE values are larger than those seen at any point in other contexts. These strong statistical differences seen in both sets of phrase-final tokens (Phrase 5, 6) stem from a categorically different behavior in the production of High and Low tone tokens – a drop in higher frequency energy is unattested in other tones or in other phrases.

The datapoints and means for LATE H1-A1 values and the mean Δ scores for the two Phrase-final embedded Carrier Phrases are now given.
Figure 37. Distribution of Early and Late H1-A1 values across all tokens for all speakers.

Figure 38. All normalized H1-A1 values at 80th and 90th decile of vowel duration. Mean H1-A1_{LATE} is plotted adjacent to individual points. Left y-axis is z-scaled. Right y-axis is retransformed.
Figure 40. Distribution of Early and Late H1-A1 values across all tokens for all speakers.

Figure 41. All normalized H1-A1 values at 80th and 90th decile of vowel duration. Mean H1-A1_{LATE} is plotted adjacent to individual data points. Left y-axis is z-scaled. Right y-axis is retransformed to dB.
Finally, phrase-medial contexts with following minor syllables (Phrases 7, 8) produced results akin to the other phrase-medial contexts with two adjacent major, tone-bearing syllables. For Phrase (7), the static measures of H1-A1 all showed a significant effect of tone, but one that was invariably the product of vastly different amplitude differences in Checked tones, as High, Low, and Creaky tones were indistinguishable in this context. In fact, for the four static measures of H1-A1 in (63), the means for High, Low, and Creaky tones were evaluated in all pair-wise comparisons to be almost certainly not different from one another ($p = 1.000$). Checked tone means were lower than each of the three other tones at significance or near significance. Results for Phrase (8) were less uniform than for Phrase (7), but also showed no pattern of consistently higher or lower H1-A1 mean values with any one tone.

4.4.4 $H1-A3$

Findings for spectral tilt quantified as H1-A3 resemble those for the H1-A1 amplitude difference. The effect of Tone and its interaction with Carrier Phrase are more or less repeated from the above section. In medial positions, there was little reliable information to be found between spectral profiles of High, Low, and Creaky tone vowels. Checked vowels, with an inherently lower F3, generally had the smallest amplitude differences. However, in terminal utterances, those without an adjoining word to follow the embedded token, the H1-A3 data were cleanly divided according to Tone. The table in (64) presents the results of a linear mixed model analysis of the five H1-A3 measures with Tone as a
fixed factor and Speaker as an error term. Discussion of Phrase-by-Phrase results below is minimal since the pattern repeats substantially from descriptions in §4.4.3 for H1-A1.

(64)  $F$-test Results, Mixed Model Analyses of Five Statistics of H1-A3

<table>
<thead>
<tr>
<th>Phrase</th>
<th>H1-A3$_{\text{EARLY}}$</th>
<th>H1-A3$_{\text{80%}}$</th>
<th>H1-A3$_{\text{90%}}$</th>
<th>H1-A3$_{\text{LATE}}$</th>
<th>ΔH1-A3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Isolated</td>
<td>* &lt; .001 L, H &gt; C, K</td>
<td>* &lt; .001 H, L, C &gt; K</td>
<td>** &lt; .001 H, L, C &gt; K</td>
<td>** &lt; .001 H, L &gt; C, K</td>
<td>** &lt; .001 H &gt; C, K</td>
</tr>
<tr>
<td>2. [Low__Low]</td>
<td>** .022 H &gt; K</td>
<td>.408</td>
<td>.597</td>
<td>.213</td>
<td>.155</td>
</tr>
<tr>
<td>3. [High__Low]</td>
<td>** .002 LH &gt; HC &gt; CK</td>
<td>** &lt; .001 L &gt; H, K</td>
<td>** .004 L &gt; C, H</td>
<td>** &lt; .001 L &gt; C</td>
<td>** .006 C &gt; K, H</td>
</tr>
<tr>
<td>4. [Low__High]</td>
<td>** &lt; .001 L &gt; H, K, C</td>
<td>.331</td>
<td>.183</td>
<td>.220</td>
<td>** .043 C &gt; H</td>
</tr>
<tr>
<td>5. [High__#]</td>
<td>** .007 L, H &gt; K</td>
<td>** &lt; .001 L, H &gt; C, K</td>
<td>** &lt; .001 L &gt; H, C, K</td>
<td>** .002 H, L &gt; K</td>
<td></td>
</tr>
<tr>
<td>6. [Low__#]</td>
<td>** &lt; .001 LH &gt; HC &gt; CK</td>
<td>** .003 L &gt; C, K</td>
<td>** .003 H &gt; C, K</td>
<td>** .002 HL &gt; LK &gt; KC</td>
<td>.121</td>
</tr>
<tr>
<td>7. [High_minor]</td>
<td>.054</td>
<td>.119</td>
<td>.192</td>
<td>.051</td>
<td>.629</td>
</tr>
<tr>
<td>8. [Low__minor]</td>
<td>.183</td>
<td>.366</td>
<td>.130</td>
<td>.087</td>
<td>.626</td>
</tr>
</tbody>
</table>

Statistical differences were found in the isolated and embedded phrase-final positions of Carrier Phrases (1), (5), and (6). In these positions, the strongest evidence for distinct phonation behavior by Tone using H1-A3 scores was found in comparisons of the H1-A3 taken at decile points late in the vowel – either H1-A3$_{\text{90\%}}$ or H1-A3$_{\text{LATE}}$. Even with these most robust results, the statistical division of tones firmly points to only a two-way contrast: the High and Low tones being produced one way and the Creaky and Checked tones produced in another. Figures 43 – 45 show this contrast in the data of Phrase (5), with the distribution$^{23}$ and means of H1-A3 scores. High and Low tone tokens

---

$^{23}$ Normalized dB values on the right y-axis in Figure 44 were re-transformed according to parameters representative of Speaker I (male), for all H1-A3 data points: $M = 26.67$dB, $SD = 7.5$. 

198
saw a conspicuous rise in H1-A3 values that was absent from Creaky and Checked tokens. Charts of H1-A3 data for all other Carrier Phrases can be found in the appendix.

Figure 43. Distribution of Early and Late H1-A3 values across all tokens for all speakers.

Figure 44. All normalized H1-A3 values at 80th and 90th decile of vowel duration. Mean H1-A3 is plotted adjacent to individual data points. Left y-axis is z-scaled. Right y-axis is retransformed to dB.

Figure 45. Mean ΔH1-A3 for each tone. Δ represents the difference between H1-A3_LATE and H1-A3_ EARLY in each token.
As shown in the table in (64), data from the phrase-medial embedded positions of Carrier Phrases (2), (3), (4), (7), and (8) offered few statistically significant differences, none of which fit the Tone-to-Phonation associations of hypotheses H1 through H4. For example, Phrase (3) data showed a significant effect of Tone that was the product of generally large H1-A3 values for Creaky-tone tokens and comparatively shallower spectral tilt among High tone tokens. These findings indicate a greater degree of laryngealization on High tone than on Creaky tone syllables, contra to hypotheses H1 (High tone is breathy) and H3 (Creaky tone is creaky).

4.4.5 Inter-Speaker Variation

A main characteristic of the spectral data was the necessity to pool all ten subjects to form any recognizable patterns by Tone. Individually, the variation in each speaker’s data was too great to permit meaningful conclusions. This degree of variation was anticipated with these data for a few reasons, described briefly here.

Generally, statistical analysis of harmonic amplitude measurements requires a fairly large $n$ to extract useful statistical inference to compensate for considerable variation inherent in these data (Hanson 1995, 1997; Ni Chasaide and Gobl 1997). Even if disregarding the nature of the data, the sample size involved in an investigation of a single subject’s behavior within a Carrier Phrase and by Tone (effectively $n = 2$) was simply too small for reliable inferences.
Recall that for spectral profile data, the aerodynamic recording was not used as it did not supply a meaningful signal for spectral analysis. Thus, each tone-bearing target syllable was produced just twice per Carrier Phrase and per Speaker in the one operative recording session. Furthermore, a sizeable amount of decile datapoints (particularly at the syllable edge) were uninterpretable by automated script analysis. So \( n \) was only 2 for each decile datapoint, and only in the ideal cases when no data were discarded.

The individual traces in Figure 47 illustrate these concerns about the data. The traces convey the dynamic change over ten decile measures of H1-A1 in tone-bearing syllables pronounced in isolation by each speaker. While some speakers reflected the overall tendencies seen in §4.4.3, others’ data could be quite noisy and show little to no difference at the vowel terminus. Figure 46 first provides traces of the dynamic H1-A1 mean for all ten subjects. Recall that in isolation, the four tones had similar H1-A1 values over the early portion of the vowel, but a statistically significant two-way divide was found in values over LATE deciles. The divide was manifest in measures of H1-A1\textsubscript{LATE} and \( \Delta H1-A1 \), and is apparent in the composite traces of Figure 46.

![Figure 46. Mean H1-A1 (z-score) at each decile (of the average rhyme duration) in Carrier Phrase 1. Pooled over all subjects. All four tones show a rising H1-A1 at the vowel terminus, but the High and Low tones rose to a significantly higher. \( n = 20 \) at most points, though occasional values were missing.](image-url)
Figure 47. Mean H1-A1 (z-score) at each decile (of the average rhyme duration), for each subject. The two-way contrast in phonation is less apparent in individual speaker data.
Looking to the individual speaker data, one can see that the traced means of some speakers perfectly demonstrated the group two-way distinction (see Speakers C and H in Figure 47), others partially produced the distinction (Speakers A, B, and G – discussed below), and a few show no conclusive distinction between the four tones (Speaker F in Figure 47). Those speakers who “partially” produced the distinction typically had Creaky tone values more in-line with the High and Low tones. Consider the traces in Fig. 47a for Speaker A. For this speaker, the trace of Creaky tone H1-A1 means follows the course of High and Low tone data, while the Checked trace begins with and maintains a lower H1-A1 difference throughout the vowel. Thus, a two-way distinction still existed, but as a 3-to-1 contrast rather than 2 by 2 Tones.

Another noteworthy aspect of the between-speaker results was the division of the spectral tilt values according to speaker sex. Numerous researchers (Holmberg et al. 1988, Klatt and Klatt 1990, Hanson 1997, Hanson and Chuang 1999) have reported physiological sex-based differences in measures of glottal qualities like spectral tilt. Namely, Hanson and Chuang (1999) found that male speakers had on average, a flatter spectral tilt, particularly as captured by H1-A3 measures. Additionally, male data tended to have “less interspeaker variation for all measures” (Hanson & Chuang 1999: 1064).

By comparison, glottalization in Burmese is not a consequence of anatomy or phrase-level prosody, but is performed with contrastive linguistic intent at the lexical level. The data collected for this study could therefore investigate the interaction between Speaker Sex and linguistically contrastive Phonation properties. Data for each spectral tilt measure taken over the latter portion of the vowel (H1-H2_{LATE}, H1-A1_{LATE}, H1-A3_{LATE})
was separated between the set of female (six) and male (four) subjects. Female and male results in each Carrier Phrase are presented for comparison in (65) – (67) below, with the male speaker values listed in shading below the corresponding female values. Along with the mean and standard deviation of the standardized spectral tilt values, the results of an F-test for an effect of Speaker Sex is listed in the far right column. A level of significance at $p < .05$ is indicated by a single asterisk (*) and $p < .01$ by dual asterisks (**).

<table>
<thead>
<tr>
<th>Context</th>
<th>PHRASE 1</th>
<th>PHRASE 2</th>
<th>PHRASE 3</th>
<th>PHRASE 4</th>
<th>PHRASE 5</th>
<th>PHRASE 6</th>
<th>PHRASE 7</th>
<th>PHRASE 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td>Creaky</td>
<td>Checked</td>
<td>F-test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isolation</td>
<td>$f$</td>
<td>0.655 (0.54)</td>
<td>0.287 (0.85)</td>
<td>0.508 (1.01)</td>
<td>0.07 (0.71)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$m$</td>
<td>-0.062 (0.46)</td>
<td>0.164 (0.76)</td>
<td>0.201 (0.94)</td>
<td>0.077 (0.82)</td>
<td>$p = .005$**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHRASE 2</td>
<td>$f$</td>
<td>0.013 (0.37)</td>
<td>0.215 (0.47)</td>
<td>0.277 (0.95)</td>
<td>-0.573 (1.02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[L_L]</td>
<td>$m$</td>
<td>-0.979 (0.23)</td>
<td>-0.824 (0.47)</td>
<td>-0.543 (0.75)</td>
<td>-0.304 (0.92)</td>
<td>$p &lt; .001$**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHRASE 3</td>
<td>$f$</td>
<td>-0.362 (0.71)</td>
<td>0.491 (0.65)</td>
<td>0.108 (1.04)</td>
<td>0.202 (0.71)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[H_L]</td>
<td>$m$</td>
<td>0.313 (0.88)</td>
<td>-0.333 (0.35)</td>
<td>0.055 (0.85)</td>
<td>-0.194 (0.84)</td>
<td>$p = .444$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHRASE 4</td>
<td>$f$</td>
<td>-0.883 (0.51)</td>
<td>0.322 (0.64)</td>
<td>0.536 (0.37)</td>
<td>-0.35 (0.57)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[L_H]</td>
<td>$m$</td>
<td>-0.261 (0.62)</td>
<td>-0.269 (0.91)</td>
<td>-0.107 (0.8)</td>
<td>0.07 (0.6)</td>
<td>$p = .775$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHRASE 5</td>
<td>$f$</td>
<td>0.546 (1.0)</td>
<td>0.098 (1.26)</td>
<td>0.028 (0.92)</td>
<td>-0.466 (0.68)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[H_#]</td>
<td>$m$</td>
<td>-0.092 (1.0)</td>
<td>0.413 (0.81)</td>
<td>0.087 (0.57)</td>
<td>-0.317 (0.71)</td>
<td>$p = .848$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHRASE 6</td>
<td>$f$</td>
<td>0.548 (0.58)</td>
<td>0.425 (0.65)</td>
<td>-0.089 (0.83)</td>
<td>-0.111 (1.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[L_#]</td>
<td>$m$</td>
<td>-0.143 (0.3)</td>
<td>-0.344 (0.19)</td>
<td>0.037 (1.68)</td>
<td>0.467 (0.98)</td>
<td>$p = .191$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHRASE 7</td>
<td>$f$</td>
<td>-0.178 (0.44)</td>
<td>-0.107 (0.39)</td>
<td>0.194 (0.6)</td>
<td>-0.698 (1.07)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[H_min.]</td>
<td>$m$</td>
<td>0.137 (1.42)</td>
<td>-0.182 (0.85)</td>
<td>0.688 (0.84)</td>
<td>-0.319 (0.96)</td>
<td>$p = .157$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHRASE 8</td>
<td>$f$</td>
<td>-0.155 (0.57)</td>
<td>0.482 (0.52)</td>
<td>0.367 (0.87)</td>
<td>-0.264 (0.74)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[L_min.]</td>
<td>$m$</td>
<td>0.224 (1.39)</td>
<td>0.042 (0.83)</td>
<td>-0.183 (1.31)</td>
<td>0.035 (1.01)</td>
<td>$p = .818$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tables in 65 - 67. Overall differences between female and male speakers were only significant across measures in data from Carrier Phrases (1) and (2). H1-A3 found a consistent difference across-the-board with lower values in male subjects. Male data are listed in shaded cells.

---

24 Analysis by linear mixed model with factors of Tone and Speaker Sex and individual Speaker as a random effect. Ideally, the F-test had (3,75) degrees of freedom as $n = 12$ for each tone for the female speaker set and $n = 8$ for each tone in the male set. Occasional missing values reduced the second term by approximately 10 degrees of freedom in most cases.
## Differences in H1-A1 by Speaker Sex

<table>
<thead>
<tr>
<th>Context</th>
<th>High</th>
<th>Low</th>
<th>Creaky</th>
<th>Checked</th>
<th>F-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHRASE 1</td>
<td>1.035 (0.51)</td>
<td>0.683 (0.71)</td>
<td>0.212 (1.07)</td>
<td>-0.452 (0.65)</td>
<td></td>
</tr>
<tr>
<td>Isolation</td>
<td>1.119 (0.66)</td>
<td>1.025 (0.71)</td>
<td>0.181 (0.77)</td>
<td>0.153 (0.75)</td>
<td>( p = .025 \ast )</td>
</tr>
<tr>
<td>PHRASE 2</td>
<td>0.419 (0.36)</td>
<td>0.482 (0.29)</td>
<td>0.422 (0.73)</td>
<td>-0.219 (0.87)</td>
<td></td>
</tr>
<tr>
<td>[L_L]</td>
<td>-0.628 (0.87)</td>
<td>-0.442 (0.82)</td>
<td>-0.178 (1.23)</td>
<td>0.048 (0.85)</td>
<td>( p = .001 \ast \ast )</td>
</tr>
<tr>
<td>PHRASE 3</td>
<td>-0.166 (0.66)</td>
<td>0.53 (0.44)</td>
<td>0.092 (0.53)</td>
<td>-0.183 (0.71)</td>
<td></td>
</tr>
<tr>
<td>[H_L]</td>
<td>-0.208 (0.42)</td>
<td>0.056 (0.45)</td>
<td>-0.012 (0.49)</td>
<td>-0.73 (0.58)</td>
<td>( p = .025 \ast )</td>
</tr>
<tr>
<td>PHRASE 4</td>
<td>-0.511 (0.61)</td>
<td>-0.121 (0.49)</td>
<td>-0.028 (0.45)</td>
<td>-0.212 (0.67)</td>
<td></td>
</tr>
<tr>
<td>[L_H]</td>
<td>-0.536 (0.63)</td>
<td>0.131 (0.81)</td>
<td>0.008 (1.12)</td>
<td>-0.355 (0.67)</td>
<td>( p = .833 )</td>
</tr>
<tr>
<td>PHRASE 5</td>
<td>0.773 (0.97)</td>
<td>0.657 (0.82)</td>
<td>-0.538 (0.66)</td>
<td>-1.247 (0.89)</td>
<td></td>
</tr>
<tr>
<td>[H_]</td>
<td>1.112 (0.63)</td>
<td>1.265 (0.46)</td>
<td>0.062 (1.09)</td>
<td>-0.433 (0.88)</td>
<td>( p = .003 \ast \ast )</td>
</tr>
<tr>
<td>PHRASE 6</td>
<td>1.321 (0.55)</td>
<td>1.21 (0.80)</td>
<td>0.3 (0.54)</td>
<td>0.489 (0.88)</td>
<td></td>
</tr>
<tr>
<td>[L_]</td>
<td>1.679 (0.54)</td>
<td>1.389 (0.63)</td>
<td>0.29 (0.39)</td>
<td>0.313 (0.51)</td>
<td>( p = .469 )</td>
</tr>
<tr>
<td>PHRASE 7</td>
<td>-0.011 (0.59)</td>
<td>0.121 (0.68)</td>
<td>0.023 (0.77)</td>
<td>-1.035 (0.53)</td>
<td></td>
</tr>
<tr>
<td>[H_min.]</td>
<td>0.162 (0.77)</td>
<td>-0.241 (0.51)</td>
<td>-0.105 (0.56)</td>
<td>-0.542 (0.99)</td>
<td>( p = .749 )</td>
</tr>
<tr>
<td>PHRASE 8</td>
<td>0.049 (0.69)</td>
<td>0.157 (0.40)</td>
<td>0.27 (0.35)</td>
<td>-0.413 (0.52)</td>
<td></td>
</tr>
<tr>
<td>[L_min.]</td>
<td>0.009 (0.62)</td>
<td>0.041 (0.68)</td>
<td>-0.226 (0.97)</td>
<td>-0.484 (0.78)</td>
<td>( p = .266 )</td>
</tr>
</tbody>
</table>

## Differences in H1-A3 by Speaker Sex

<table>
<thead>
<tr>
<th>Context</th>
<th>f</th>
<th>High</th>
<th>Low</th>
<th>Creaky</th>
<th>Checked</th>
<th>F-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHRASE 1</td>
<td>m</td>
<td>1.17 (0.39)</td>
<td>0.981 (0.63)</td>
<td>0.679 (0.84)</td>
<td>-0.357 (0.64)</td>
<td></td>
</tr>
<tr>
<td>Isolation</td>
<td>m</td>
<td>1.008 (0.87)</td>
<td>0.723 (0.75)</td>
<td>-0.098 (1.24)</td>
<td>-0.158 (1.13)</td>
<td>( p = .032 \ast )</td>
</tr>
<tr>
<td>PHRASE 2</td>
<td>f</td>
<td>0.632 (0.42)</td>
<td>0.71 (0.33)</td>
<td>0.648 (0.67)</td>
<td>-0.136 (1.13)</td>
<td></td>
</tr>
<tr>
<td>[L_L]</td>
<td>m</td>
<td>-0.547 (0.47)</td>
<td>0.047 (0.59)</td>
<td>0.156 (.5)</td>
<td>0.114 (0.96)</td>
<td>( p = .001 \ast \ast )</td>
</tr>
<tr>
<td>PHRASE 3</td>
<td>f</td>
<td>0.242 (0.74)</td>
<td>0.871 (0.89)</td>
<td>0.726 (0.67)</td>
<td>-0.141 (0.61)</td>
<td></td>
</tr>
<tr>
<td>[H_L]</td>
<td>m</td>
<td>0.025 (0.88)</td>
<td>0.463 (0.91)</td>
<td>0.169 (0.74)</td>
<td>-1.247 (1.3)</td>
<td>( p = .005 \ast \ast )</td>
</tr>
<tr>
<td>PHRASE 4</td>
<td>f</td>
<td>0.053 (0.81)</td>
<td>0.111 (0.65)</td>
<td>0.189 (0.512)</td>
<td>-0.202 (0.57)</td>
<td></td>
</tr>
<tr>
<td>[L_H]</td>
<td>m</td>
<td>-0.963 (0.41)</td>
<td>-0.055 (0.69)</td>
<td>-0.446 (0.71)</td>
<td>-0.52 (0.64)</td>
<td>( p = .001 \ast \ast )</td>
</tr>
<tr>
<td>PHRASE 5</td>
<td>f</td>
<td>0.814 (0.77)</td>
<td>0.257 (1.14)</td>
<td>-0.083 (0.98)</td>
<td>-0.881 (0.91)</td>
<td></td>
</tr>
<tr>
<td>[H_#]</td>
<td>m</td>
<td>1.092 (0.62)</td>
<td>1.075 (0.41)</td>
<td>0.442 (0.56)</td>
<td>-0.202 (0.9)</td>
<td>( p = .043 \ast )</td>
</tr>
<tr>
<td>PHRASE 6</td>
<td>f</td>
<td>1.61 (0.40)</td>
<td>1.31 (0.69)</td>
<td>0.598 (0.76)</td>
<td>0.77 (0.693)</td>
<td></td>
</tr>
<tr>
<td>[L_#]</td>
<td>m</td>
<td>0.713 (0.84)</td>
<td>0.605 (0.91)</td>
<td>-0.988 (2.13)</td>
<td>-0.546 (1.50)</td>
<td>( p &lt; .001 \ast \ast )</td>
</tr>
<tr>
<td>PHRASE 7</td>
<td>f</td>
<td>0.222 (0.51)</td>
<td>0.502 (0.83)</td>
<td>0.355 (0.83)</td>
<td>-0.386 (0.64)</td>
<td></td>
</tr>
<tr>
<td>[H_min.]</td>
<td>m</td>
<td>0.049 (1.3)</td>
<td>-0.181 (0.52)</td>
<td>-0.622 (1.33)</td>
<td>-0.685 (1.19)</td>
<td>( p = .011 \ast )</td>
</tr>
<tr>
<td>PHRASE 8</td>
<td>f</td>
<td>0.504 (0.53)</td>
<td>0.539 (0.5)</td>
<td>0.726 (0.59)</td>
<td>0.091 (0.8)</td>
<td></td>
</tr>
<tr>
<td>[L_min.]</td>
<td>m</td>
<td>0.131 (0.73)</td>
<td>0.243 (0.56)</td>
<td>-0.294 (0.94)</td>
<td>-0.598 (1.55)</td>
<td>( p = .004 \ast \ast )</td>
</tr>
</tbody>
</table>
The findings for this study partially support Hanson’s observation that H1-A3 showed the strongest contrast between the female and male speakers, and that male speakers had a consistently shallower spectral tilt as measured by H1-A3. This was the case in all but one context, Carrier Phrase (5) where the Speaker Sex difference was statistically significant, but male values were actually higher for all four tones. Contra to Hanson’s finding, there was not markedly greater variation in the male data.

In comparing just the values for Creaky tone data, one can see that the male speakers regularly had a lower H1-A3 in Creaky tone tokens (except again for the curious Phrase 5 findings). Although not as consistently across contexts as with the Creaky tone, a shallower tilt was found with the four male speakers for tokens bearing any tone. For example, mean H1-A3 in Phrase (2) embedded tokens was considerably lower amongst males for the High and Low tone in addition to Creaky tone tokens. This was not the case for syllables produced in isolation (Phrase 1), where High and Low tone standardized means were indistinct between male and female subjects.

The analysis of the overall pooled data led to the generalization that realization of the contrastive phonation with the Creaky tone was context-dependent. Yet male subjects regularly produced the Creaky tones with more glottalization in all contexts according to H1-A3 measures, suggesting the possibility that creaky voicing was a consistent phonetic property of the Creaky tone in male speech, but was context-dependent in female speech. A review of the male-subject only data in (67) refutes this hypothesis, but further supports the notion that a phonation contrast is context-dependent, with Creaky tones distinct in just the terminal positions of Phrases (1), (5), and (6). The male subject standardized mean H1-A3_LATE was higher in Creaky tokens than both High and Low
tokens in Phrase (2) [Low__Low], and higher than High tone tokens in Phrases (3) [High__Low] and (4) [Low__High].

Findings for H1-H2_LATE and H1-A1_LATE were less different by Speaker Sex. Significant differences in the sets of means were found mainly in two contexts – Phrase (1) [Isolated syllables] and Phrase (2) [Low __ Low]. One pattern in the H1-H2_LATE data (65) was a lower mean for males, though not always significantly, in most of the Carrier Phrase environments (Phrases 1 – 6). H1-A1_LATE, data (in 66) showed few consistent patterns according to Speaker Sex. Male means were lower for High, Low tones in Phrase (2) [Low__Low], but higher in the other phrases embedding in a terminal position, both Phrases (5) [High__#] and (6) [Low__#]. In fact, male and female means for H1-A1_LATE were scarcely different in Phrases (4), (6), (7), and (8).

4.4.6 Summary of Spectral Tilt Results

Measures of spectral tilt have been the source of considerable confusion in the literature on Burmese tones. Despite robust associations with phonation types in a variety of languages (see §2.5.4), findings that linked shallower or steeper spectral tilt to the Creaky or High Tones in Burmese ranged from inconclusive (Watkins 1997) to firmly negative (Gruber and Feizollahi 2006). It has not been clear why acoustic analysis of spectral profiles produced such erratic results. A few possible causes were discussed in Chapter Two:
(a) There simply was no contrastive phonation between the tones of Burmese.

(b) A phonation contrast did exist, but the language-specific phonetic implementation of this contrast was such that spectral tilt was poorly suited to capture it.

or (c) Methodological problems in previous studies hindered the strength of their findings.

The data presented in sections 4.4.2-4.4.4 resolve many of the issues creating this confusion.

To begin with, spectral tilt did indeed find a robust distinction between the tones that aligned with predicted phonation properties. The findings were limited though to measures of H1-A1 and H1-A3 and were only found when the measured syllable was not followed by another word in connected speech. In a sense, all three possibilities listed above explain the previous difficulties with spectral tilt measures. It is possible that there was no contrastive phonation to detect (possibility a), in tones occurring sentence-medially. Where the contrast was reliably produced, H1-H2, a well-accepted indicator of phonation type, failed to detect it (possibility b). Finally, prior studies reporting no linkage between tone and phonation type only looked at phrase-medial forms (Gruber and Feizollahi 2006), a methodological problem in a sense (possibility c). Another potential methodological issue, the portion of the vowel from which power spectra are evaluated, is not fully resolved here. Findings for H1-A1 and H1-A3 at late decile timepoints (LATE, 80th and 90th percentiles) nicely detected a distinction, but significant differences were also found between “EARLY” means of all three spectral tilt measures with surprising
frequency; the Creaky and Checked tokens often had significantly shallower spectral tilt over the range of early decile datapoints. Additionally, this significance held across multiple contexts, rather than only in terminal utterances.

The tables in (68) – (70) consolidate the findings presented through the above sections in order to give a broader picture of the contrast with streamlined statistical results. To review the notation used in the tables: the darker the cell, the more strongly the Tone-to-Phonation associations of hypotheses H1 – H4 are supported. Dark grey cells always confirm at least a two-way distinction between greater spectral tilt for the modal/breathy voiced Low and High tone and a shallower tilt for the more constricted Creaky and Checked tones. This presentation makes it plainly visible that significant findings were limited to Carrier Phrases (1), (5), and (6). Lighter-shaded cells are infrequent outside of these phrase contexts, except for EARLY means where the Low tone (but not High) was often significantly higher than the other tones.

In addition to the juncture-driven alternation, another key finding of this section pertains to the phonation associations composing **Hypotheses One** and **Two**. Putting aside the Checked tone, which is distinct from the other three in a variety of highly salient qualities, the hypothesized phonation contrast is tripartite: Breathy vs. Modal vs. Creaky. Even the most robust spectral tilt findings did not support this division of the tones. High and Low values were rarely distinct – in only three cases reviewed above were they significantly different (Phrase 3 H1-A1\textsubscript{LATE}, Phrase 4 H1-H2\textsubscript{90%}, and Phrase 4 H1-H2\textsubscript{LATE}) and in all three instances, the Low tone had significantly greater average spectral tilt than the High tone – reverse of the predicted order if the High tone were indeed breathier.
### (68) Summary of Mixed Model Analyses of Five Statistics of H1-H2

<table>
<thead>
<tr>
<th>H1-H2</th>
<th>Early</th>
<th>80% decile</th>
<th>90% decile</th>
<th>Late</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrase 1</td>
<td>**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>**</td>
<td>-</td>
<td>-</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>5</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>*</td>
<td>-</td>
<td>*</td>
<td>**</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### (69) Summary of Mixed Model Analyses of Five Statistics of H1-A1

<table>
<thead>
<tr>
<th>H1-A1</th>
<th>Early</th>
<th>80% decile</th>
<th>90% decile</th>
<th>Late</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrase 1</td>
<td>** L, H &gt; K **</td>
<td>** H, L &gt; K</td>
<td>** HL &gt; CK</td>
<td>** HL &gt; CK</td>
<td>** HL &gt; CK</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>* All pairs ns</td>
<td>**</td>
<td>-</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>4</td>
<td>* L &gt; C, K</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>** LH &gt; CK</td>
<td>** LH &gt; CK</td>
<td>** HL &gt; CK</td>
<td>** LH &gt; CK</td>
</tr>
<tr>
<td>6</td>
<td>** L &gt; K, C **</td>
<td>** H &gt; C, K</td>
<td>** HL &gt; CK</td>
<td>** HL &gt; CK</td>
<td>** H &gt; C, K</td>
</tr>
<tr>
<td>7</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>*</td>
<td>-</td>
<td>-</td>
<td>* All pairs ns</td>
<td>-</td>
</tr>
</tbody>
</table>

### (70) Summary of Mixed Model Analyses of Five Statistics of H1-A3

<table>
<thead>
<tr>
<th>H1-A3</th>
<th>Early</th>
<th>80% decile</th>
<th>90% decile</th>
<th>Late</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrase 1</td>
<td>** L,H &gt;C,K **</td>
<td>**</td>
<td>**</td>
<td>**HL &gt;C &gt;K</td>
<td>** H &gt; C, K</td>
</tr>
<tr>
<td>2</td>
<td>*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>** L &gt; C, K **</td>
<td>*</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>4</td>
<td>** L &gt; HKC **</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>*</td>
</tr>
<tr>
<td>5</td>
<td>** L, H &gt; K **</td>
<td>** L, H &gt; K</td>
<td>** L,H &gt;C,K</td>
<td>** L,H &gt;C,K</td>
<td>**</td>
</tr>
<tr>
<td>6</td>
<td>** L &gt; C, K</td>
<td>** L &gt; K, C</td>
<td>** H &gt; C, K</td>
<td>** H, L &gt; C</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
In short, H1-A1 and H1-A3 effectively distinguished phonation properties, but only a two-way distinction and only in terminal contexts (isolated, phrase-final). Measures of spectral tilt taken late in the vowel were more successful in these contexts. Even for H1-A1 and H1-A3, the statistic tracking spectral tilt change during the vowel ($\Delta$) was only moderately effective at contrasting High, Low tones from constricted Creaky and Checked tones.

The non-significant differences are worthy of examination as well, and for H1-A1 and H1-A3 it is noteworthy that the phrase-medial data do not even trend in the appropriate direction. Values for High, Low, and Creaky tone tokens were scarcely different and were inconsistently ordered with the steepest spectral tilt occurring alternately with any of the three. In the fifty phrase-medial cells in (69) and (70), only seventeen (34%) had Creaky tones with the lowest mean of the three. On the other hand, thirty out of thirty of the cells for isolated and phrase-final data showed higher means for High and Low tones.

By comparison, no meaningful trends present themselves in the H1-H2 data. It has already been discussed how the array of phonetic qualities which distinguish the lexical tones work to obscure the already indirect spectral cues to phonation. The distinct formants of Checked vowels, different levels of intensity, and dynamically variant pitch levels collude against compatible comparisons of spectral information across tones. But it is not clear why these interweaving effects should hinder H1-H2 responses but allow for meaningful contrasts in H1-A1 and H1-A3 readings.

One critical difference between H1-H2 and H1-A# measures for this matter is not that lower or higher frequency harmonics are less or more instructive as to Burmese
phonation differences, but that the measure looks at multiple integers of a single component, F0. Vastly different F0 values (which are not just possible between the tones, but expected) therefore impact both terms in the formula $H_1 - H_2^{25}$. On a smaller scale, $H_1 - H_2$ could be failing the Burmese data as it tries to compare the amount of energy at rather different frequencies. Assuming a stable vowel quality across tone-bearing samples, comparisons of energy at formant frequencies could understandably be more reliable.

Regardless, some context-dependent contrast between the tones was evident in the spectral data. The findings for $H_1 - A_1$ and $H_1 - A_3$ are corroborated by the EGG and aerodynamic data in the upcoming sections, which support both the two-way split (High, Low > Creaky, Checked) and the environments in which it is seen. In fact, these more direct measures of glottal aperture uncover a stronger contrast in the terminal positions and indicate a moderate difference in some medial phrases.

---

25 If, for some reason, one compared $H_1 - H_5$ in order to include information from higher frequency energy, the error would only be compounded. Large F0 differences would be multiplied five times over and the second amplitude term would measure inconsistent areas of the power spectrum.
4.5 **Open Quotient Measures**

4.5.1 **Description of OQ data**

Open quotient measures of the EGG signal presented the same context-dependent pattern found in the more successful of the spectral tilt data. Between-tone differences were prominent in isolated or phrase-final utterances and are best characterized as a late rise on the Low and High syllables. As with F0 and spectral tilt, temporally dynamic measures were necessary to draw distinctions between the tones of Burmese. In medial positions, with the embedded token in juncture, the two-way distinction was mostly neutralized, though a weak tendency persisted for lower OQ scores with Creaky and Checked syllables. Though it is not obvious from surveying the means, open quotient measures were more successful at differentiating tones than spectral tilt, where intra- and inter-speaker variation was greater. For the noticeable OQ patterns, the impediment to a cohesive body of data was missing and discarded values, particularly late in Creaky and Checked syllables. The data that were available offered a fairly stable view between speakers. This section begins by displaying the mean OQ trace over the syllable duration for each tone in each frame. A number of patterns emerge through this visual inspection that are confirmed in the following subsection’s statistical analysis. As with the other modes of data, each decile OQ datapoint represents the mean OQ score of all subjects’ pooled data, standardized for individual speaker differences.

Putting aside minor differences, there are two apparent patterns in the figures. In Phrases (2, 3, 4, 7, 8), the OQ trace remained relatively level for all tones, and all tokens were at roughly similar values (though Creaky and Checked trended towards lower values).
Figure 48. Mean OQ at deciles in Phrases (1) - (8) by Tone across all speakers. Y-axis values are z-scaled, normalized for each speaker. Decile data points for each chart a – h represent \( n = 20 – M \), where M is the count of missing values.
In Phrases (1, 5, 6) there was a clear cut two-way distinction in the data whereby Low and High tones achieved a more open glottal state and Creaky and Checked tones maintained or closed further to a more constricted state. Two sets of phrases \{1, 5, 6\} and \{2, 3, 4, 7, 8\} reflected an obvious distinction in context: those containing a medial embedded token in juncture with a following syllable and those where the embedded token terminated the utterance. This bimodal behavior of the OQ data is entirely in line with findings throughout this study. For duration, spectral profile, and (to a lesser extent) F0, sentence position bore a much greater effect on the phonetic character of the embedded token than did the tonal content of adjacent syllables.

4.5.2 Statistical Results

The Open Quotient traces in Figure 48 represented the mean at individual decile timepoints. For statistical analysis, an average of early and late timepoints for each token was calculated to generate OQ_{EARLY} and OQ_{LATE}. Results by Carrier Phrase for these two statistics are presented in boxplots next to bars giving the average open quotient over the entire rhyme of each token, OQ_{MEAN}. A graph of the mean ΔOQ for each tone in each Phrase accompanies. The data were analyzed by Carrier Phrase for each statistic, MEAN, EARLY, LATE, and Δ, as a dependent variable in a linear mixed model with Tone as a fixed effect and Speaker as a random effect. In cases where the global null hypothesis was rejected, Bonferroni corrected multiple comparisons tested the means for all tones pair-wise, revealing how individual tones differed.
Looking at the values for the isolated tokens from Phrase (1), the bimodal split between High/Low and Creaky/Checked tones (also seen in spectral tilt measures) is clearly visible in Figures 49 and 50, and shown to be statistically significant in (71). Particularly in $OQ_{LATE}$ and $\Delta OQ$, there were robust statistical differences between these tone groups with post-hoc comparisons showing that High and Low tone means were not different from one another, but were so individually from the Creaky and Checked means. In fact, all four statistics showed a significant effect of Tone in isolated utterances (see 71), as Checked tones also had a significantly lower $OQ$ than Low and High tones in both $OQ_{\text{MEAN}}$ and $OQ_{\text{EARLY}}$.

(71) Effect of Tone on Open Quotient In Isolated Syllable Tokens

<table>
<thead>
<tr>
<th>Phrase 1</th>
<th>Main effect</th>
<th>Differences by multiple comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>$F(3,209) = 6.77$ $p &lt; .01$</td>
<td>$L, H, C &gt;&gt; C, K$</td>
</tr>
<tr>
<td>EARLY</td>
<td>$F(3,209) = 4.40$ $p &lt; .01$</td>
<td>$C, L, H &gt;&gt; K$</td>
</tr>
<tr>
<td>LATE</td>
<td>$F(3,178) = 23.00$ $p &lt; .01$</td>
<td>$L, H &gt;&gt; C, K$</td>
</tr>
<tr>
<td>$\Delta$</td>
<td>$F(3,178) = 15.23$ $p &lt; .01$</td>
<td>$L, H &lt;&lt; K, C$</td>
</tr>
</tbody>
</table>

![Figure 49. $OQ_{\text{MEAN}}$ and boxplots depicting $OQ_{\text{EARLY}}$ and $OQ_{\text{LATE}}$ data distribution. Bar marker and value represent mean, not median, values. Right y-axis reflects a z-scale, normalized for speaker. Left y-axis is retransformed from z for illustrative purposes to inverse vocal fold contact area as a percentage.](image-url)
Figure 50. Mean ΔOQ for each tone in Carrier Phrase (1). Δ taken as the token-by-token difference between OQ_{EARLY} and OQ_{LATE}. Positive value indicates increasing vocal fold contact over course of the syllable rhyme.

Statistical analysis of data from Phrases (2), (3), and (4) showed few differences between the effects of tone on OQ in phrase-medial syllables. There were no differences to report in tokens embedded between two Low tone syllables (Phrase 2, see (72)) as well as with a following High tone (Phrase 4, see (74)). In Phrase (3) [High ___ Low], data were consistently ordered High > Low > Creaky > Checked but only a few contrasts were statistically significant (see 73).

(72) Effect of Tone on Open Quotient In Phrase (2) Tokens

<table>
<thead>
<tr>
<th>Phrase 2</th>
<th>Main effect</th>
<th>Differences by multiple comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>F (3,78) = 1.22, p = .321</td>
<td></td>
</tr>
<tr>
<td>EARLY</td>
<td>F (3,78) = 1.58, p = .212</td>
<td></td>
</tr>
<tr>
<td>LATE</td>
<td>F (3,72) = .829, p = .482</td>
<td></td>
</tr>
<tr>
<td>Δ</td>
<td>F (3,72) = .98, p = .407</td>
<td></td>
</tr>
</tbody>
</table>

(73) Effect of Tone on Open Quotient In Phrase (3) Tokens

<table>
<thead>
<tr>
<th>Phrase 3</th>
<th>Main effect</th>
<th>Differences by multiple comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>F (3,70) = 4.22, p &lt; .05</td>
<td>H, L &gt;&gt; L, C &gt;&gt; C, K</td>
</tr>
<tr>
<td>EARLY</td>
<td>F (3,76) = 2.10, p = .107</td>
<td></td>
</tr>
<tr>
<td>LATE</td>
<td>F (3,67) = 3.61, p &lt; .05</td>
<td>H, L, C &gt;&gt; C, K</td>
</tr>
<tr>
<td>Δ</td>
<td>F (3,67) = 1.49, p = .226</td>
<td></td>
</tr>
</tbody>
</table>

(74) Effect of Tone on Open Quotient In Phrase (4) Tokens

<table>
<thead>
<tr>
<th>Phrase 4</th>
<th>Main effect</th>
<th>Differences by multiple comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>F (3,71) = 2.15, p = .117</td>
<td></td>
</tr>
<tr>
<td>EARLY</td>
<td>F (3,72) = 1.63, p = .191</td>
<td></td>
</tr>
<tr>
<td>LATE</td>
<td>F (3,66) = 0.46, p = .711</td>
<td></td>
</tr>
<tr>
<td>Δ</td>
<td>F (3,65) = 0.72, p = .546</td>
<td></td>
</tr>
</tbody>
</table>
These findings signal a tendency toward the familiar two-way phonation split in some phrase-medial forms, though produced neither as strongly nor reliably as in isolated or phrase-final syllables. The distribution of OQ means yielding the partially significant findings in (73) are charted in Fig. 52 - 53. By comparison, Fig. 51 shows no differences.

Figures 51 (above) and 52 (below). OQ\text{MEAN} and boxplots depicting OQ\text{EARLY} and OQ\text{LATE} data distribution in Phrases (2) and (3). Bar marker and value represent mean, not median, values. Right y-axis reflects a z-scale, normalized for speaker. Left y-axis is retransformed from z for illustrative purposes to inverse vocal fold contact area as a percentage.
Figure 53. Mean ΔOQ for each tone in Carrier Phrase (3). Δ taken as the token-by-token difference between OQ_{EARLY} and OQ_{LATE}. Positive values indicate increasing contact over course of the syllable rhyme.

The findings for Carrier Phrase (3) data signaled a tendency toward the familiar two-way phonation split, between two tones associated with less constriction and laxness and two more constricted tones. Though produced neither as strongly nor as reliably as in isolated or phrase-final syllables, the results suggest that glottalization in Creaky and Checked tones is a tendency (albeit weak) even in phrase-medial forms. Figures 52 and 53 (above) display the distribution of OQ means that yielded the partially significant findings in this context.

Phrase-Final Contexts: Carrier Phrases (5) and (6)

It was apparent in the decile-by-decile traces in Figure 48 that Open Quotient varied greatly by tone in the two phrase-final embedded contexts. This distinction was borne out by the statistics which stated significant differences between the group means for OQ_{LATE} and ΔOQ in both Phrase (5) and (6). Data for both Phrases were clustered into two groups: (a) the High and Low tones and (b) the Creaky and Checked tones. The groups differed significantly in OQ_{LATE} and ΔOQ by post-hoc multiple comparisons in Phrase (5)
Effect of Tone on Open Quotient In Phrase (5) Tokens

### Phrase 5: Main effect

<table>
<thead>
<tr>
<th></th>
<th>F (3,69)</th>
<th>p</th>
<th>Differences by multiple comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>2.92</td>
<td>.056</td>
<td></td>
</tr>
<tr>
<td>EARLY</td>
<td>2.41</td>
<td>.074</td>
<td></td>
</tr>
<tr>
<td>LATE</td>
<td>6.47</td>
<td>&lt; .01</td>
<td>H, L &gt;&gt; C, K</td>
</tr>
<tr>
<td>Δ</td>
<td>2.87</td>
<td>&lt; .05</td>
<td>H, L &lt;&lt; K, C</td>
</tr>
</tbody>
</table>

---

**Figure 54.** OQ\_MEAN and boxplots depicting OQ\_EARLY and OQ\_LATE data distribution. Bar marker and value represent mean, not median, values. Right y-axis reflects a z-scale, normalized for speaker. Left y-axis is retransformed from z for illustrative purposes to inverse vocal fold contact area as a percentage. Both OQ\_MEAN and OQ\_LATE show an effect of Tone: High, Low > Creaky, Checked.

**Figure 55.** Mean ΔOQ for each tone in Carrier Phrase (5). Δ taken as the token-by-token difference between OQ\_EARLY and OQ\_LATE. Positive values indicate increasing vocal fold contact over the course of the syllable rhyme.
(76) Effect of Tone on Open Quotient In Phrase (6) Tokens

<table>
<thead>
<tr>
<th>Phrase 6</th>
<th>Main effect</th>
<th>Differences by multiple comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>$F(3,72) = 1.91$</td>
<td>$p = .152$</td>
</tr>
<tr>
<td>EARLY</td>
<td>$F(3,66) = .527$</td>
<td>$p = .665$</td>
</tr>
<tr>
<td>LATE</td>
<td>$F(3,50) = 3.82$</td>
<td>$p = .015$ L, H $&gt;&gt;$ K, C</td>
</tr>
<tr>
<td>Δ</td>
<td>$F(3,48) = 3.99$</td>
<td>$p = .013$ L, H, K $&gt;&gt;$ K, C</td>
</tr>
</tbody>
</table>

Figure 56. OQ MEAN and boxplots depicting OQ EARLY and OQ LATE data distribution. Bar marker and value represent mean, not median, values. Right y-axis reflects a $z$-scale, normalized for speaker. Left y-axis is retransformed from $z$ for illustrative purposes to inverse vocal fold contact area as a percentage. Both OQ MEAN and OQ LATE show an effect of Tone: High, Low $>$ Creaky, Checked.

Figure 57. Mean ΔOQ for each tone in Carrier Phrase (6). Δ taken as the token-by-token difference between OQ EARLY and OQ LATE. Positive values indicate increasing vocal fold contact over the course of the syllable rhyme.
and nearly so in Phrase (6), shown in the tables of (75) and (76). Both phrases corresponded to a noticeably lower OQ\textsubscript{EARLY} score as well (approaching significance in Phrase 5), indicating that glottal constriction could begin as early as the first 1/3\textsuperscript{rd} of the tone-bearing vowel. The group means of each OQ statistic for Phrases (5) are charted in Figures 54 and 55, and for Phrase (6) in Figures 56 and 57.

\textit{Contexts Preceding a Minor Syllable: Phrases (7) and (8)}

Open Quotient on medial syllables preceding a minor syllable exhibited an effect of Tone which was a balance between the phonation-neutralizing medial environments (cf. 72 - 74) and the phonation-contrasting terminal environments (cf. 71, 75, 76). In the Phrase (7) and (8) data, signs of creaky voicing existed, but were not as robust as in tokens embedded phrase-finally. Indications of an effect of Tone were scarcely distributed amongst the quantified variables – no OQ variable besides OQ\textsubscript{LATE} showed a between-group difference by the linear mixed model analysis. Open Quotient readings late in the vowel were regularly lower in Creaky and Checked syllables, statistically so for Creaky in Phrase (7) and Checked in Phrase (8). The distribution of means yielding these results are displayed in Figures 58 (Phrase 7) and 59 (Phrase 8). Figures for \Delta\text{OQ} can be found in the appendix.

<table>
<thead>
<tr>
<th>(77) Effect of Tone on Open Quotient In Phrase (7) Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phrase 7</strong></td>
</tr>
<tr>
<td>MEAN</td>
</tr>
<tr>
<td>EARLY</td>
</tr>
<tr>
<td>LATE</td>
</tr>
<tr>
<td>(\Delta)</td>
</tr>
</tbody>
</table>

\textsuperscript{26} In Phrase (6), Open Quotient fell only slightly late in Checked vowels, a fact which produced a \(\Delta\text{OQ}\) not significantly different from the OQ rise found with High and Low tones. See Figure 55.
(78) Effect of Tone on Open Quotient In Phrase (8) Tokens

<table>
<thead>
<tr>
<th>Phrase 8</th>
<th>Main effect</th>
<th>Differences by multiple comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>$F(3,72) = 3.29$ $p &lt; .05$</td>
<td>H, L, C $&gt;&gt;$ C, K</td>
</tr>
<tr>
<td>EARLY</td>
<td>$F(3,70) = 1.18$ $p = .325$</td>
<td></td>
</tr>
<tr>
<td>LATE</td>
<td>$F(3,67) = 5.26$ $p &lt; .01$</td>
<td>H, L, C $&gt;&gt;$ C, K</td>
</tr>
<tr>
<td>$\Delta$</td>
<td>$F(3,66) = 2.08$ $p = .111$</td>
<td></td>
</tr>
</tbody>
</table>

Figures 58 (above) and 59 (below). OQ\textsubscript{MEAN} and boxplots depicting OQ\textsubscript{EARLY} and OQ\textsubscript{LATE} data distribution. Bar marker and value represent mean, not median, values. Right y-axis reflects a $z$-scale, left y-axis is retransformed from $z$ for illustrative purposes to inverse vocal fold contact area as a percentage.
4.5.3 Discussion of EGG Findings

The electroglottographic data in Section 4.5 provided the clearest depiction yet in the literature regarding the phonation contrast produced with Burmese tones. The findings of Watkins (2005) for tones spoken in isolation were reproduced and expanded to allow a characterization of tone-based laryngeal constriction across contexts. The four tones were visibly and statistically split into two different marked behaviors – either constriction increased throughout the vowel or a normal constriction generating modal vibration followed by a late rapid increase in aperture. This two-way contrast was extremely robust in isolated and phrase-final contexts when vowel final qualities were not obscured by an adjoining segment. However, in phrase-medial contexts the increased constriction was also apparent and sometimes statistically significant, while there was no indication of the late rising OQ on medial High and Low syllables. Open Quotient values dropped in Creaky and Checked syllables when leading into a minor syllable (Phrase 7, 8) or when following a High tone and preceding a Low (Phrase 3). In the two phrase-medial contexts which followed a Low tone, the mean OQ traces of all four tones were nearly identical, indicating a degree of glottal closure consistent with modal voicing without any signs of late constriction or widening.

Properties of the dynamic OQ traces were quantified in four ways for analysis. One was ill-suited for distinguishing the tones (OQ\textsubscript{EARLY}) as the OQ scores from the first half of each vowel were comparable and consistently signaled regular, modal levels of constriction for tokens of all four tones. Three measures captured the two-way laryngeal contrast to some degree, but OQ\textsubscript{LATE} was the most effective. The table in (79) shows how a significant effect of Tone predicted by the Tone-to-Phonation hypotheses was found in
OQ_{LATE} in all contexts but for the medial Carrier Phrases (2) and (4). Even so, the OQ_{LATE} means in Phrases (2) and (4) were higher for the High and Low tones, just not by a statistically significant proportion. Another noteworthy aspect of the data captured in (79) was that, excluding Phrase (1) results, OQ_{EARLY} means offered no meaningful patterns either statistically or simply by their order.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Carrier Phrases with Significant Difference ( (p &lt; .05) )</th>
<th>Carrier Phrases with Means Ordered ( (ns) ) {High, Low} &gt; {Creaky, CVO}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Iso : Medial : Final : Minor</td>
<td>Iso : Medial : Final : Minor</td>
</tr>
<tr>
<td>MEAN</td>
<td>1 : 3 : 7 : 8</td>
<td>1 : 2 : 3 : 5 : 6 : 7 : 8</td>
</tr>
<tr>
<td>EARLY</td>
<td>1 : 3 : 5 : 6</td>
<td>1 : 2 : 3 : 4 : 5 : 6 : 7 : 8</td>
</tr>
<tr>
<td>LATE</td>
<td>1 : 3 : 5 : 6</td>
<td>1 : 3 : 4 : 5 : 6 : 7 : 8</td>
</tr>
<tr>
<td>( \Delta )</td>
<td>1 : 5 : 6</td>
<td>1 : 3 : 4 : 5 : 6 : 7 : 8</td>
</tr>
</tbody>
</table>

A review of the Figures in §4.5.2 suggests that distinctions found in the overall mean OQ (OQ_{MEAN}) and in the change in OQ values (\( \Delta \)OQ) were primarily the consequence of divergent OQ values late in the vowel. That is, there was no meaningful distinction according to mixed model analysis in the data (for any one measure in any context) for which the distinction in OQ_{LATE} was not as or more significant. Both OQ_{MEAN} and \( \Delta \)OQ yielded contrasts in line with each other and OQ_{LATE}, but added no decisive power to the single linear mixed model of OQ_{LATE}. It can be concluded that the degree of glottal aperture over the end of the vowel constitutes a better articulatory goal for Burmese speakers producing tones than other measures of the EGG waveform, whether taken at a static point or measuring dynamic changes in quality.
4.6 Aerodynamic Data

4.6.1 Overview

Aerodynamic data of the production of the four tones were analyzed in four of the eight Carrier Phrases: Isolated Phrase (1), Phrase-medial Phrases (2) and (3), and Phrase-final Phrase (6). Phrases were not pooled in the statistical analysis, as each Phrase represented an environment with a potentially unique effect on the rates of oral airflow (as above, with the distinction between the two medial Phrases being the potentially breathy preceding High tone in Phrase 3). Along with the reduced subject pool of eight (see §3.5.7), fewer Phrases meant fewer tokens comprising each mean, sixteen in this case\(^{27}\). Statistical inferences were still quite strong with this reduced \(n\), possibly because there were very few discarded tokens. Since measures were made manually, no data points were undefined by an automated script. Only twelve tokens (out of a possible 288) were not included in the final calculations\(^ {28}\), mainly due to high rates of airflow maxing out the pressure transducers’ range and creating an unreliable, clipped signal.

For each speaker, all data points were transformed to a \(z\)-score scale, thereby controlling not just for inter-speaker differences in oral airflow capacity, but also for the possibility of any variation introduced by equipment configuration for each recording session. Since the same calibration was performed before each subject session (see §3.5.7 for calibration procedure), this was likely not a large effect. A mean and standard deviation for the \(z\)-scale was built from each speaker’s pool of values for \(\text{airflow}_{\text{WHOLE}}\),

\(^{27}\) Sixteen for each tone in each context, meaning 64 Phrase (2) tokens to compare in a linear mixed model of Phrase (2) data. Eight subjects x two repetitions yields an \(n\) of 16. Compare to \(n = 20\) for spectral or EGG data (ten subjects) or \(n = 40\) for F0 and duration data (ten subjects x two treatments (acoustic and airflow recordings) x two repetitions). For Phrase (1), \(n\) is 1.5x greater. Further, all \(n\) described in this footnote represent idealized counts, not accounting for missing data,

\(^{28}\) Ten for Speaker H, two for Speaker I. Missing tokens were not restricted to a certain tone either, but were evenly distributed – 2 High, 2 Low, 2 Checked, and 4 Creaky (Speaker H) and 2 Low (I).
airflow_{EARLY}, airflow_{LATE}, and airflow_{MID} values. Every speakers’ data were then pooled in this format for presentation in the Figures 60 – 63 and for statistical analysis.

Results for the airflow experiment did indicate a two-way split between the four tones that conforms to the phonation-based contrast identified in previous sections with acoustic and EGG data. High and Low tones were associated with significantly greater amounts of airflow than the more constricted Creaky and Checked tones. As seen in the data above, this difference was only found in certain contexts and was detectable only at the end of the tone-bearing vowel. The data for syllables uttered in isolation illustrated this contrast very well. These are shown first in the following section.

4.6.2 Results for Airflow Experiment

Phrase (1) isolated syllables with all four tones had roughly similar airflow rates in the early portion of the vowel. By the 2\textsuperscript{nd} half of the vowel the two-way contrast was apparent, and even later the difference was dramatic during the 75ms subsequent to the vowel offset (Figure 60a). Late vowel measures revealed lower airflow in Creaky and Checked syllables ($M = -0.83z, 61.1 \text{ ml/s}$ for Creaky; $M = -0.79z, 63.7 \text{ ml/s}$ for Checked\footnote{cf Figure 62a, 63a for medial tokens with airflow rates between 0 and 1 z, 115 - 180 ml/s.}) and similarly low rates for the post-vocalic span. By contrast, High and Low tones saw a large difference between the mean values for airflow_{LATE} and airflow_{POST-VOCALIC}. Figure 60b displays the distribution of airflow_{WHOLE} and airflow_{MAX} data points in Phrase (1). The table in (80) states that the difference seen in airflow_{WHOLE} was significant.

\footnote{This calculation intentionally does not include 1) the PEAK airflow, avoiding an extreme right-ward skewing of the data, nor 2) the POST-VOCALIC average, which was thoroughly bi-modal with values either near zero or approaching PEAK rates.}

\footnote{Values in \textbf{bold} represent those re-transformed from a pooled z-scale to ml/s.}
but that $\text{airflow}_{\text{peak}}$ and $\text{airflow}_{\text{max}}$ were unable to draw a firm two-way difference between the group means.

<table>
<thead>
<tr>
<th>Phrase 1</th>
<th>Main effect</th>
<th>Differences by multiple comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>1\textsuperscript{ST} \text{HALF}</td>
<td>$F(3,66) = 2.07$</td>
<td>$p = .112$</td>
</tr>
<tr>
<td>2\textsuperscript{ND} \text{HALF}</td>
<td>$F(3,66) = 41.22$</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>\text{POST-VOCALIC}</td>
<td>$F(3,66) = 19.33$</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>\text{ENTIRE VOWEL}</td>
<td>$F(3,66) = 16.71$</td>
<td>$p &lt; .001$</td>
</tr>
<tr>
<td>\text{MID-POINT}</td>
<td>$F(3,66) = 6.11$</td>
<td>$p = .001$</td>
</tr>
<tr>
<td>\text{PEAK}</td>
<td>$F(3,66) = 1.94$</td>
<td>$p = .131$</td>
</tr>
</tbody>
</table>

The statistics used for both the graphs and mixed model analysis above included all Low tone tokens pooled. Recall from §4.3.1 that isolated Low tones were produced in two distinct fashions by the subject pool: three with an even or slightly falling F0 and seven with a late, sharp rise in F0. Separating these subject groups did not reveal a meaningful difference in oral airflow rates. In Figure 61, the \text{EARLY}, \text{LATE}, and \text{POST-VOCALIC}
airflow rates for each group is given, demonstrating that there was no correlation between rising pitch and an increase in airflow.

![Figure 61. Mean rate of Oral Airflow (z-score) separating Low tone tokens with a late intonational rise and those without, Isolated syllable data. Y-axis in z-scale representing milliliters/second over three time spans: EARLY, LATE, and POST-VOCALIC. Regardless of intonation, airflow increased dramatically at the vowel terminus.](image)

Phrase (2) medial tokens all had similar rates of airflow regardless of tone. The averaged airflow readings in Figure 62a show an increased rate of oral airflow later in the vowel, higher in fact for the High and Low tones than in isolated syllables where this elevated rate was contrastive. Measurement of the post-vocalic span generally consisted of the overlapping onset of the following word, the voiceless stop [k]. Duration of this span was on average 58.9 ms, and not the full 75 ms. Airflow during this span was understandably low, frequently at or approaching 0 ml/s. A linear mixed model analysis of each statistical variable representing the changing airflow, with Tone as a fixed effect and Speaker as a random effect, found no differences between the results of the four tones. See (81) below for details. This lack of a difference is apparent not only in the similar trace in Figure 62a, but in the data distribution in 62b. The only significant difference found was for Checked syllables to regularly have higher airflow over the first half of the vowel, a
tendency which also influenced the statistically higher oral airflow at the vowel midpoint.

**Figure 62a.** Average rate of Oral Airflow in z-scale representing milliliters/second over three time spans: EARLY, LATE, and POST-VOCALIC. Re-transformed scale in ml/s is given on y-axis in blue. The airflow rates of the four tones were not statistically different.

**Figure 62b.** Distribution by token of the average rate of Oral Airflow taken over the entire vowel on the left, and the peak rate of airflow attained during the vowel. Values in converted z-scale given on right y-axis. Neither statistic effectively captured a phonation contrast.

<table>
<thead>
<tr>
<th>Phrase 2</th>
<th>Main effect</th>
<th>Differences by multiple comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>1ST HALF</td>
<td>$F(3,43) = 12.15$, $p &lt; .001$</td>
<td>K &gt;&gt; C, H, L</td>
</tr>
<tr>
<td>2ND HALF</td>
<td>$F(3,43) = 0.36$, $p = .782$</td>
<td>No pairwise differences: H, C, L, K</td>
</tr>
<tr>
<td>POST-VOCALIC</td>
<td>$F(3,34) = 0.75$, $p = .531$</td>
<td>No pairwise differences: H, C, L, K</td>
</tr>
<tr>
<td>ENTIRE VOWEL</td>
<td>$F(3,43) = 1.79$, $p = .164$</td>
<td>No pairwise differences: K, H, C, L</td>
</tr>
<tr>
<td>MID-POINT</td>
<td>$F(3,43) = 4.19$, $p = .011$</td>
<td>K, H &gt;&gt; H, L, C</td>
</tr>
<tr>
<td>PEAK</td>
<td>$F(3,43) = .49$, $p = .694$</td>
<td>No pairwise differences: K, H, C, L</td>
</tr>
</tbody>
</table>

**Phrase (3),** also a phrase-medial context, produced similar statistical results – the mean rates of airflow over each time frame were indistinguishable between the tones. EARLY and POST-VOCALIC results compared well with those in Phrase (2), but the LATE means were rather lower, an effect not clearly attributable to the preceding High tone in Phrase (3). Recall that Phrase (3) embedded tokens were the penultimate syllable in the utterance, calling into question whether decreased airflow over the second half of the
vowel was a product of following a High tone syllable or simply the effect of declination at a later sentence position. Statistical results from the mixed model analyses are given in (82) and show that the four tones were decidedly not different across the range of statistics. The single pair-wise difference found in the data was between Checked and Creaky means for $\text{airflow}_{\text{Early}}$.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure63a.png}
\caption{Average rate of Oral Airflow in $z$-scale representing milliliters/second over three time spans: Early, Late, and Post-Vocalic. Re-transformed scale in ml/s is given on y-axis in blue. The airflow rates of the four tones were not statistically different.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure63b.png}
\caption{Distribution by token of the average rate of Oral Airflow taken over the entire vowel on the left, and the peak rate of airflow attained during the vowel. Values in converted $z$-scale given on right y-axis. Neither statistic effectively captured a phonation contrast.}
\end{figure}

(82) Effect of Tone on Measures of Oral Airflow in Phrase (3) Tokens

<table>
<thead>
<tr>
<th>Phrase 3</th>
<th>Main effect</th>
<th>Differences by multiple comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Half</td>
<td>$F(3,44) = 3.27$</td>
<td>$p = .030$ K, H, L &gt;&gt; H, L, C</td>
</tr>
<tr>
<td>2nd Half</td>
<td>$F(3,44) = 1.75$</td>
<td>$p = .171$ No pairwise differences: H, L, K, C</td>
</tr>
<tr>
<td>Post-Vocalic</td>
<td>$F(3,44) = 2.40$</td>
<td>$p = .080$ No pairwise differences: L, H, C, K</td>
</tr>
<tr>
<td>Entire Vowel</td>
<td>$F(3,44) = 2.56$</td>
<td>$p = .067$ No pairwise differences: H, K, L, C</td>
</tr>
<tr>
<td>Mid-Point</td>
<td>$F(3,44) = 1.48$</td>
<td>$p = .234$ No pairwise differences: H, L, K, C</td>
</tr>
<tr>
<td>Peak</td>
<td>$F(3,44) = .65$</td>
<td>$p = .585$ No pairwise differences: H, L, K, C</td>
</tr>
</tbody>
</table>
Phrase (6) data, representing phrase-final embedded tokens in the airflow experiment, revealed the distinct two-way contrast also seen in isolated syllables. The only statistical difference found in airflow over the first half of the vowel (\(\text{airflow}_{\text{EARLY}}\)) was between a higher mean airflow rate in Checked syllables and a low rate in Creaky ones. The LATE and POST-VOCALIC measures revealed the same striking difference found in Phrase (1), with High and Low tones accompanied by a surge of air after the vowel periodicity has ended. This divergence was confirmed by the statistical tests (see 83). A mixed model analysis of \(\text{airflow}_{\text{LATE}}\) and \(\text{airflow}_{\text{POST-VOCALIC}}\) showed a strongly significant difference (\(**\), \(p < .001\) for both) between the large amounts of airflow in High and Low tones and the reduced flow in Creaky and Checked. The mean airflow over the entire vowel was also significantly different (\(p < .001\)), but grouped Checked with the High and Low tones because the greater airflow early in Checked vowels counteracted the lower rates later in the vowel to produce a higher overall mean. Tests of the static airflow measures of the PEAK and vowel MID-POINT could not distinguish the group means.

As with Low tones in Phrase (1), no distinction could be found in the oral airflow averages between speakers who produced a rising F0 on Low tones in the Phrase (6) final position (Speakers A, B, F, G) and those who used a low and fairly level F0 (Speakers C, D, E, H, I, J). The graph in Figure 65 gives the mean airflow rates over each time frame for each group of speakers, again showing no correlation between a rising F0 and rates of oral airflow.
Figure 64a. Average rate of Oral Airflow in $z$-scale representing milliliters/second over three time spans: EARLY, LATE, and POST-VOCALIC. Re-transformed scale in ml/s is given on y-axis in blue. High and Low tones had much higher rates of airflow late and after the vowel.

Figure 64b. Distribution by token of the average rate of Oral Airflow taken over the entire vowel on the left, and the peak rate of airflow attained during the vowel. Values in converted $z$-scale given on right y-axis. Creaky tones were significantly lower in for airflow$_{\text{whole}}$. Otherwise, the statistics were unhelpful in finding a contrast.

(83) Effect of Tone on Measures of Oral Airflow in Phrase (6) Tokens

<table>
<thead>
<tr>
<th>Phrase 6</th>
<th>Main effect</th>
<th>Differences by multiple comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>1$^{\text{ST}}$ HALF</td>
<td>$F (3,37) = 3.35$ \hspace{1em} $p = .029$</td>
<td>K, H, L &gt;&gt; H, L, C</td>
</tr>
<tr>
<td>2$^{\text{ND}}$ HALF</td>
<td>$F (3,37) = 9.48$ \hspace{1em} $p &lt; .001$</td>
<td>H, L &gt;&gt; K, C</td>
</tr>
<tr>
<td>POST-VOCALIC</td>
<td>$F (3,37) = 11.08$ \hspace{1em} $p &lt; .001$</td>
<td>L, H &gt;&gt; C, K</td>
</tr>
<tr>
<td>ENTIRE VOWEL</td>
<td>$F (3,37) = 8.75$ \hspace{1em} $p &lt; .001$</td>
<td>H, L, K &gt;&gt; C</td>
</tr>
<tr>
<td>MID-POINT</td>
<td>$F (3,37) = 3.23$ \hspace{1em} $p = .033$</td>
<td>No pairwise differences: L, H, K, C</td>
</tr>
<tr>
<td>PEAK</td>
<td>$F (3,37) = .25$ \hspace{1em} $p = .865$</td>
<td>No pairwise differences: L, H, K, C</td>
</tr>
</tbody>
</table>

Figure 65. Mean rate of Oral Airflow ($z$-score) separating Low tone tokens with a late intonational rise and those without, Phrase (6) data. Y-axis in $z$-scale represents ml/s over three time spans. Regardless of intonation, airflow increased dramatically at the vowel terminus.
4.6.3 Summary of Airflow Results

Overall, the airflow data display many of the trends noted in the other modes of data that looked at phonation type contrasts (H1-A1, H1-A3, OQ). Mainly, that a clear phonation distinction was made by Burmese speakers but that it was…

a) Only a two-way contrast: \{High, Low\} vs. \{Creaky, Checked\}

b) Context-sensitive: Found in citation and phrase-final forms, and neutralized phrase-medially.

c) Temporally-sensitive: The entire vowel was not realized with contrastive phonation. The distinction was produced only at the vowel terminus, and no contrast was seen during early periods of the vowel.

The timing of the effect on oral airflow was not only found late on the vowel, but was particularly pronounced after the vowel, as revealed in the immense contrast in the 75ms post-vocalic span between High and Low tones on the one hand, and the Creaky and Checked tones on the other.

Results from the airflow portion of the study did more than complement the EGG and acoustic findings with another mode of data. Rates of oral airflow are a direct measure of a physiological difference in the production of the tones, and thus the results did not merely indicate a difference in a measure of phonation, but provided an insight into the nature of that difference. Namely, that the increase in airflow at the end of High and Low vowels was far too overwhelming to be regarded simply as the absence of creak, and accordingly that these forms employed some type of marked laryngeal state, with a widened glottis. As a result of greater vocal fold abduction, breathy or lax phonation types prompt greater rates of airflow (Ladefoged and Maddieson 1996, Pennington
2005). Such a portrayal of the laryngeal articulation involved with Burmese High and Low tones was reinforced by the simultaneous decrease in vocal fold contact, identified in the EGG signal in §4.5 (cf. Figure 48). Conversely, Creaky and Checked tones restricted airflow, a finding consistent with the late or syllable-final glottalization indicated by other data (lower OQ scores, irregular periodicity in the acoustic waveform). Alternatively, the glottal widening and increased airflow could be a product of the phrase position. In this understanding, a phrase-final burst of air could be neutralized by the constricted glottis of Creaky and Checked tones, therefore only being found with the unmarked Low and high-pitched High tones when uttered phrase-finally.

Another important facet of the aerodynamic findings was the apparent independence of the breathy quality from the pitch contour of High tones. Most obviously, because the breathy quality found with High tones also occurred with the Low tone tokens. Closer inspection shows that the breathiness was realized simultaneous with nearly any level of pitch: with a large fall (High tone), a level low (Low tone), or a sharp rise (intonationally altered Low tone). There was some evidence that rising pitch weakened the breathy quality, but they were by no means incompatible. Such an effect would be predicted phonetically as breathy phonation is typically associated with lowered F0 (Gregerson 1976, Hombert 1979, Titze 1994, Yip 2002) and increased airflow. Hombert et al. (1979: 48) state that the lower subglottal pressure generated by the lack of vocal fold tension and additional airflow would “by itself … lead to a somewhat lower F0”. As stated above, breathiness and a high F0, though articulatorily antagonistic, could co-occur in Burmese, as attested by the sharply rising Low tones, where vowel-final breathiness coincided with a ramped up F0. In this case, the lack of glottal tension
produced with Low tones (and the resulting airflow) is mitigated by the need to achieve faster rates of vocal fold vibration.

4.7 Summary

The array of experimental data in this chapter has shed light on the numerous phonetic issues in Burmese Tonal phonology reviewed in Chapter Two. One can definitively say that the tonal categories in Modern Burmese have a multi-valent phonetic realization. There was a consistent, statistically-backed distinction in each of the properties examined: DURATION, FUNDAMENTAL FREQUENCY, SPECTRAL PROFILE, OPEN QUOTIENT, and ORAL AIRFLOW. While Chapter Four demonstrated that all of these distinctions exist, their realization was perhaps not as convoluted as the literature made them appear. One key generalization that can be made is that many of the distinguishing phonetic qualities were strongest in citation form, nearly as distinct in phrase-final forms, and attenuated or fully neutralized in phrase-medial forms in running speech.

For one, the DURATION contrast was particularly strong in citation form. For many speakers, there was a four-way contrast in citation tokens corresponding to the reported vowel lengths in Thein Tun (1982) and traditional grammars (e.g. Okell 1969): High > Low > Creaky > Checked. In sentence medial tokens though, the contrast was weakened to one between CV and CVO syllables with somewhat shortened Creaky tone vowels.

The contrast in PITCH could be reduced to a three-way contrast between a low-pitched tone (Low), a falling tone (Creaky, Checked), and a tone bearing a high pitch
target that may occur early or late in the syllable (High). Of all the phonetic qualities investigated in this study, the F0 contours associated with each tone remained the most distinct regardless of phonetic context. The eight Carrier Phrases influenced the canonical pitch contours in several ways, but in nearly every Phrase context some three-way contrast was maintained between a sharp fall (Creaky, Checked), a very slight and gradual fall (Low), and a moderate fall/rise (High).

The phonation contrast captured separately by Spectral Profile, Open Quotient, and Oral Airflow rates was found to be a simple two-way contrast between ‘laxer’ and ‘more constricted’ phonation types that was most often unrealized in medial positions. The neutralization in medial position occurred even when phonation was not a property of the vowel, but of the coda segment in Checked syllables.

These findings partially confirm the set of hypotheses H1 – H4, repeated below.

**H1. High tone syllables are associated with breathy phonation**
**H2. Low tone syllables are pronounced with regular, modal phonation**
**H3. Creaky tone syllables are associated with creaky phonation**
**H4. Checked syllables are also associated with creaky phonation**

The confirmation is “partial” because any positive associations between Tone category and a non-modal phonation type were context-sensitive. In Phrase (2) for example, all hypotheses were rejected except Hypothesis Two, though in this case the Low tone was

---

32 Though it should be noted that Syllable Type, which was not quantified, does clearly distinguish Checked tones from the other three in any context.

33 In the sense that Checked syllable codas lose their glottalized qualities when the glottal stop coda assimilates fully to the following onset consonant.
hardly unique in using modal voicing. With the qualification of a proper context, only **Hypothesis Three** and **Four** were unreservedly confirmed, while **Hypothesis One** and **Two** were exclusive of one another. The High and Low tones were shown repeatedly to be indistinguishable in all measures of phonation type, entailing that both **Hypothesis One** and **Two** cannot hold true. Looking to the absolute results for each measurement, the High and Low tone in terminal utterances were not simply *different* from the creaky voiced tokens, they were also *different* from the regular modal values found in medial samples and in early portions of their own vowel. At, during, or after the vowel offset, spectral tilt became steeper, OQ increased, and appreciably greater amounts of air passed through the lips (and presumably the glottis). Based on this observation, it is appropriate to say that Burmese High and Low tones are associated with breathy phonation. Therefore **Hypothesis One** is confirmed for isolated and phrase-final High tones with the caveat that the vowel is not breathy-voiced so much as the syllable contains a late breathy quality.

This concludes the analysis of phonetic correlates of the Burmese tones. Going forward, the crucial part of this conclusion for phonological analysis is simply that the much-reported three-way (or even four-way) contrast in phonation does not exist between the tones. It is henceforth assumed in this dissertation that Burmese contrasts just two modes of phonation: a more lax and a more constricted mode. A simple two-way opposition is typologically and perceptually more plausible, and was shown to best fit the production data of this study’s subjects. In many ways, the question of phonation type was at the center of any debate over the primary suprasegmental contrast in Burmese. The experiments in this chapter have offered considerable insight into this debate – surely
modern Burmese is not a “register” language, where voice quality distinctions may or may not realize incidental pitch patterns, and yet register-like differences nicely separate the tones when spoken in certain contexts. So, numerous questions remain, including those posed by hypotheses H5 through H10 in Chapter Three. The ensuing chapters further examine acoustic and articulatory data to test these hypotheses, which turn to the phonological claims about the targets associated with each tone.
Chapter Five. Evidence for Integrated Phonological Analysis

5.1 Overview of Evidence Employed in the Study

The remaining chapters of this dissertation turn to the dilemma of how to phonologically represent the differences between the four Burmese lexical tones. For the integrated phonological account sought by this dissertation, the review of prior phonological accounts from Chapter Two, along with the phonetic evidence presented in the previous chapter, has set the stage for new analyses capable of addressing shortcomings in the present understanding of Burmese. The following chapters lay out these analyses and resolve many of these shortcomings with further examination of the production data from Chapter Four, diachronic evidence, and a perception study (in Chapter Six) of the acoustic cues to the tone contrast.

To begin, one of the most puzzling patterns presented in Chapter Two concerned the interaction of vowel quality and syllable type, for which the distribution has so far defied adequate phonological description. Section §5.2 returns to the analysis of Burmese meter and syllable structure in Green (2005). The critique of Green’s proposal offered here sets up the investigation of the following two sections. In §5.3, the historical origins of the vowel quality distribution are shown to assemble a simple explanation for an otherwise confusing and unintuitive pattern. Section 5.4 then compares measures of vowel duration before each of the coda types to construct a model of moraic weight in Burmese syllables à la Broselow et al. (1997). Little is known about what the obstruent and nasal codas have in common and where the two behave differently, particularly in regard to their effects on the preceding vowel duration. All of the production data collected for this dissertation, acoustic and aerodynamic, were able to provide the
durations of the vowels and coda segments necessary to inform a model of Burmese syllable structure and weight that is argued to improve upon that in Green (2005).

Following an interim discussion of the findings informing models of Burmese syllable structure (§5.5), Sections 5.6 and 5.7 turn to the analysis of the F0 patterns seen in Chapter Four, with particular focus on the systematic F0 changes brought about by adjacent tones. The F0 analyses fall within the theme of the chapter as a statistical examination of phonologically relevant aspects of the data. In this case, the phonological question concerns whether a neighboring tone influences F0 as a sandhi effect or via co-articulation, and if so, in what way?

Finally, despite the clearer picture in Chapter Four of the tones’ production, the featural representation of the contrast is debatable. F0 contours provided the strongest distinction between High, Low, and Creaky across contexts, but duration and acoustic correlates of phonation type offered robust distinctions in citation form tokens and phrase-finally. No prior literature exists concerning the perception of Burmese tones. Therefore, even a preliminary perception study offers great potential to disentangle the converging phonetic dimensions. This study is included in a separate chapter (§6) to follow.
5.2 Distribution of Vowel Quality: Critical Review of Green (2005)

5.2.1 Vowel Quality and Syllable Structure in Green (2005)

The central argument in Green (2005) is that the facts in (1i–iv) can be attributed to the banning of \([\text{PLACE}]\) features “from the right edge of a syllable.”

<table>
<thead>
<tr>
<th>(1) Generalizations about the Composition of a Burmese Major Syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. Only debuccalized segments are permitted in coda position, ([?]) and a placeless nasal ([\text{N}]).</td>
</tr>
<tr>
<td>ii. Mid-vowels ([e, \text{ɔ, o}]) only occur in open syllables.</td>
</tr>
<tr>
<td>iii. Diphthongs ([ai, \text{ei, ou, au}]) only occur in closed syllables with one of the two possible codas.</td>
</tr>
<tr>
<td>iv. All possible rimes occur with the Low, High, and Creaky tones, but checked tone only appears on vowels in a syllable closed by ([?]).</td>
</tr>
</tbody>
</table>

Green (2005) extends the Coda Condition on consonantal place as formalized in Itô and Mester (1994) to “place” features on the vowel. He labels this constraint \(* \text{PLACE}]_\sigma\), stated explicitly in (2), which will henceforth be called the Extended Coda Condition, or \(\text{EXTENDED CODA CON} \).

(2) **EXTENDED CODA CONDITION (\(* \text{PLACE}]_\sigma\):**

\[ * \mu]_\sigma \quad \text{“The rightmost mora of a syllable does not dominate} \]
\[ \mid \quad \text{Place features.”} \quad \text{(Green 2005: 9).} \]

\(\text{PLACE}\)

It is necessary to define \(* \text{PLACE}]_\sigma\) as a moraic restriction in order to conflate the requirement of a placeless \(V\) or placeless \(C\) at the syllable right edge. Furthermore, Green stipulates that the coda consonants \([?]\, [\text{N}]\) are considered moraic and the first mora is the head mora of a bimoraic foot. Therefore, an open syllable such as \([k^h\text{æ}]\) is permitted,
since the head mora dominates the vowel’s place feature while the second, rightmost mora shares it as in (3a). Conversely, an open diphthong [kʰaɪ, kʰei, kʰəʊ, kʰoʊ] as in (3b) is not possible because the second mora dominates the place feature of [i] or [ʊ].

![Diagram](image)

When a diphthong occurs without a coda in the input, its allophonic variant [e, ɔ, o] is realized due to the ranking ExtCodaCon » Max(Place), which prefers the removal of vowel Place features to their realization at the right edge.

Green’s analysis of Burmese prosodic structure nicely handles a number of phenomena in the data and does so with familiar, well-grounded constraints. It is argued in the following section however that the effort to justify the Extended Coda Condition, a major feature of the analysis, introduces critical inconsistencies in the prosodic alignment of the language’s syllables.

5.2.2 Difficulties for Green (2005): Inconsistent Treatment of Diphthongs

Both nasal and obstruent coda consonants are considered to have moraic weight by Green. Open major syllables contain bimoraic vowels and the same vowel becomes monomoraic with a coda. Diphthong vowels are necessarily of the form V_µ V_µ in order to be banned in open syllables, yet contradictorily, these same vowels are required to be monomoraic (VV_µC_µ) to be licit with a coda. If a diphthong may share both vowel
qualities as a single moraic head when permitted, why is such a structure not permitted in an open syllable, allowing a perfectly licit\textsuperscript{1} structure like that in (4)?

![Diagram](image)

(4)

The distribution of diphthong vowels in Green's analysis seems to follow this theoretically unsatisfactory maxim: diphthongs may be monomoraic when they are permitted, but are bimoraic (and therefore banned) in the syllables that do not permit them.

The treatment of diphthongs is also responsible for other inconsistencies. In two tableaux (examples (25) and (26) in Green 2005: 11), two similar input diphthongs derive opposite outputs: one output as a closed-syllable diphthong formed by the epenthesis of a final [ʔ], and the other an open-syllable monophthong output resulting from deletion of the second vowel segment. In the former, the input syllable /daɪ-/ is forced to realize an epenthetic coda consonant (i.e. [daɪN]) in order to rescue itself from having \textsc{place} in the second mora. This is because when [ɪ] heads the second mora, it bears \textsc{place} features visible to the \textsc{extended coda condition}. Quite oppositely, the latter example derives the output [tʃʰo] from the input /tʃʰou/ because, again, \textsc{place} is banned in the

\textsuperscript{1} Recall that the \textsc{extended coda condition} militates against the right-edged mora \textit{uniquely} dominating place features. In (4), the \textsc{place} features under the rightmost mora are shared.

244
second mora. The inputs are effectively the same (diphthong vowels) as are the triggers (violation of ExtCODACon), yet somehow the repairs are opposite: one inserts a PLACE-less consonant and the other deletes the offending PLACE-bearing vowel quality.

5.2.3 Difficulties for Green (2005): Underlying Vowel Phonemes

Behind the distribution of the diphthong vowels in Green (2005) is the assumption that only eight underlying vowels provide the assortment of possible rimes in Burmese, and that the underlying forms are the diphthongal/closed-syllable set. The alternation of vowel qualities in his analysis is given in (5).

<table>
<thead>
<tr>
<th>VOWEL UR</th>
<th>Open-σ</th>
<th>Nasal Closed-σ</th>
<th>Glottal Closed-σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/²</td>
<td>[i]</td>
<td>[i]</td>
<td>[i]</td>
</tr>
<tr>
<td>/ei/</td>
<td>[e]</td>
<td>[ei]</td>
<td>[ei]</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>[ɛ]</td>
<td>-</td>
<td>[ɛ]</td>
</tr>
<tr>
<td>/ai/</td>
<td>-</td>
<td>[ai]</td>
<td>[ai]</td>
</tr>
<tr>
<td>/a/</td>
<td>[a]</td>
<td>[a]</td>
<td>[a]</td>
</tr>
<tr>
<td>/au/</td>
<td>[ɔ]</td>
<td>[au]</td>
<td>[au]</td>
</tr>
<tr>
<td>/ou/</td>
<td>[o]</td>
<td>[ou]</td>
<td>[ou]</td>
</tr>
<tr>
<td>/u/</td>
<td>[u]</td>
<td>[u]</td>
<td>[u]</td>
</tr>
</tbody>
</table>

One motivation for the alternations presented in (5) is that any account where “input monophthongs /e, ɔ, o/ have diphthongal output allophones would have grave difficulties motivating diphthongization in closed syllables” (Green 2005: 15). Green may be saying that it is problematic for vowel lengthening to be conditioned by heavy final obstruents,

² Vowel symbols used in (5) differ from elsewhere in this dissertation in order to match those used by Green (2005). Green presents /a, i, ei, ɛ, u/ as the same vowel in open and closed syllables, while phonetic transcriptions in the literature prefer [a, ɪ, eɪ, ɛ, ʊ] in closed syllables. Additionally, [ʌ] is sometimes used rather than [a] in closed syllables (Mehnert and Richter 1976).
but this objection is removed if the diphthongs are not treated as long vowels, but simply as contour vowels. That is, a vowel with the phonetic transcription [ʊʊ] is not necessarily long with the moraic weight of two short vowels, an understanding confirmed by duration evidence reported in Thein Tun (1982) and in the data collected for this study.

Alternatively, the problem posed by Green may be that it is difficult to conceive of a uniform vowel change that would simultaneously alternate [i] with [i] and [a] with [a], but [o] with [ʊʊ] and [e] with [ei]. The confusion here may be that the vowel representations Green assumes (in (5)) do not recognize the other variations in vowel quality in closed syllables other than the diphthongs. This stability of the pure vowel phonemes [a, i, e, ε, u] in open or closed syllables is not to be found in the phonetic literature. The “grave” problem with monophthong inputs becomes much less grave when one acknowledges that the entire “diphthongized” set of vowels differs from the open-syllable “monophthong” set. Watkins (2005a) argues that not only do the counterpart [i] and [ʊ] occur in closed syllables, but that [ε] and [a] in orthographically-closed syllables differ from those in open syllables. Bradley (1982), Okell (1969), and Cornyn and Roop (1968) all recognize or describe the difference in pronunciation between [i] and [i], [u] and [u], and the open and closed [ε] and [a]. Furthermore, the Burmese orthography also distinguishes the closed vowel set from the open vowels in nearly every case (Wheatley 1987). Notably, open [a] and closed [ɑʔ]/[ʌʔ] differ (ʊ vs. ʊ), as do open [ε] closed [ɛʔ], (ʊ, ʊ vs. ʊ). Only the mid-vowels [ɛ] and [auʔ] are represented with characters indicating a shared vowel quality (ʊ vs. ʊ).
Regardless, the point Green raises has validity – it is still unclear how the alternated sets [i]~[ɪ], [a]~[ʌ], [o]~[ʊ], and [e]~[ei] represent a uniform effect on vowel quality. In the following section, it is argued that the most suitable explanation is to simply treat the different vowel sets as different underlying vowels. There is in fact a transparent phonetic motivation for the diphthongization of [e], [ɔ], and [o] in closed syllables and not all vowels, but it is one that was active in the history of the language and not at present. The orthographic distinctions described above are noteworthy because they reflect pronunciation during the era in which the relationship between the open and closed vowel qualities was allophonic. The discussion now turns to the evidence for this relationship.

5.3 Distribution of Vowel Quality: Diachronic Evidence

Vowel Quality and Tone interact in Burmese phonology in such a way that one could propose between eight (Green 2005), fifteen (Watkins (2001), and twenty-two (Thurgood 1978) underlying vowel qualities. In Green’s model, the diphthongal vowel set represented the underlying vowel quality, which surfaced in closed syllables or was altered in open syllables under pressure to satisfy an extended coda condition that was argued to affect even vowels. Alternatively, an underlying set of monophthong vowels could be subject to an allophonic alternation in closed syllables such that diphthongal or more-centralized vowel qualities are induced by a final consonant. Mehnert and Richter (1972 – 1977), for one, suppose this allophony.

It is argued here that the most suitable analysis of the vowel quality distribution involves no active alternation, but rather one that occurred in the history of the language.
In this account, fifteen underlying vowels are assumed (à la Watkins, *cf* Tables 7, 8, (2001: 7)) and the restricted distribution of certain vowel qualities is a fossilized, static phonological association between syllable type and the bisected vowel inventory. To begin with, there are no evident synchronic alternations between the two sets of vowels to indicate any active derivation of one set from the other. Following the accounts of Maran (1971) and Bradley (2011) and aided by the diachronic comparisons of forms in Thurgood (1981) and Luce (1985), this section presents the historical evidence for the origin of the split vowel qualities. Most significantly, the explanation demonstrates that during the period of an active allophonic alternation there was no one-to-one correspondence between open and closed-syllable vowel qualities, but that the *eight* closed-syllable forms were primarily derived from just *five* open-syllable vowels. For example, Old Burmese /i/ has two MB closed-syllable reflexes, both /ɪ/ and /eɪ/, while /e/ and /o/ have no corresponding closed-syllable form.

The diachronic account nicely resolves two issues with the allophonic analysis of Green (2005) and others. To begin with, allophony presupposes there is some shared quality to the closed-syllable vowel set, such that one can claim that vowels are diphthongized or centralized when shortened or otherwise influenced by a syllable coda. Instead, half are diphthongs and the other four, monophthongs /i, ɛ, a, u/, fit a description as centralized variants of the open-syllable vowels /i, ɛ, a, u/. An analysis which cohesively treats the eight closed vowel qualities under a single transformation is unlikely, and to date, has not been provided in the literature.

A further issue is the mismatch in the number of vowels found in each environment: seven in open syllables and eight in closed. Deriving one set from the other
entails leaving one vowel out, [ε] or [ai] being the most conspicuous candidates. The
diphthong [ai] lacks an obvious open-syllable counterpart, but problematically [ε] is only
found with a glottal, and not nasal, coda. Neither is an ideal allophonic variant of /ɛ/ in
both CV and CVN syllables.

Maran’s (1971) description focuses on the transition from the language of the 18th
Century, frequently called Written Burmese (WB), which is as indicated by name
preserved to an extent in the modern orthography. At a stage prior to Written Burmese,
coda consonants only occurred with three vowels, /a/, /i/, and /u/ and were themselves
limited to the stop series /p, t, c, k/ and /m, n, ɳ, ŋ/. Simplifying a bit for the diphthongs
/ai/ and /au/, the eight MB closed-syllable vowels are divided between reflexes of WB
[-p/m] or [-t/n] rimes and those from WB [-k/ŋ] rimes. Diphthongization was an
allophonic alternation affecting certain vowels only before the velar consonant codas.
Vowel quality was therefore not only connected to the presence of a coda, but was
sensitive to the place features of that coda, an association that became less apparent with
the impoverishment of the coda distinctions. The diagram in (6) provides the two most
straightforward cases demonstrating how the closed-syllable set is composed of pairs of
former velar and non-velar rimes.
The origins of closed-syllable vowels /ɛ/, /ai/, /ao/ are more complicated, but are ultimately similar in being the product of velar consonant codas inducing vowel allophony before and during the era of Written Burmese. The diagram in (7) outlines the brief discussion below. For a fuller explanation of the incremental developments involved see Chapter Two in Maran (1971).

(7) Diachronic Development of Checked Vowel Qualities /a, i, e, ao, ai/

Recall that only the Old Burmese vowels /a/, /i/, and /u/ were followed by consonants. Similar to the /i/ and /u/ provenience of four closed-syllable vowel qualities (in 7), OB /a/ had multiple WB and Modern Burmese reflexes in closed syllables, depending on the consonantal place of the coda. Following changes in the vowel prior to the WB stage, [-k]
final syllables with the original /a/ vowel yielded the modern Checked /eʔ/ rime, while most other closed /a/ rimes more directly yielded the modern /aʔ/ after the simplification of all codas to /ʔ/. This diachronic development offers insight into the paradigmatic gap in the set of possible Burmese syllable rimes where no nasally-closed /eN/ rimes exist. While the other velar-induced vowel allophones developed concurrent oral and nasal rimes (e.g. /i/ → [ei] before oral and nasal velars), the formation of /e/ rimes relied on the prior diachronic change of /a/ → /e/ regardless of a velar coda. Maran (1971) provides evidence that /eN/ was not a possible rime at the time of the velar-induced allophony, surprisingly not because no such change had occurred, but because Old Burmese /aŋ/ had a Written Burmese [iN] reflex, via the stages OB /aŋ/ → /en/ → WB /iN/.

The remaining two closed-syllable diphthongs /aɪ/ and /au/ are shown by Maran (1971) and Bradley (2011) to be a later development, not from /a/, /i/, or /u/ but from Old Burmese diphthongs /ui/ and /ea/. By the era of Written Burmese, these two OB diphthongs had become the two mid-vowels /ɛ/ and /ɔ/. Since at this stage of the language, coda consonants could follow other vowels than only the OB monophthongs, the two mid-vowels occurred with the codas /k/ and /ŋ/ and surfaced as in (8) and (9) (from Maran 1971: 107):

(8) /ɛ/ → [aɪ] / __ velars; thus [aɪk], [aɪŋ], [aɪ́ŋ], [aɪ̰́ŋ]

(9) /ɔ/ → [au] / __ velars; thus [aʊk], [aʊŋ], [aʊ́ŋ], [aʊ̰́ŋ]

There is accompanying orthographic evidence for many of the changes noted above as well, with the relation between now-phonologized allophonic pairs being
directly captured in their written characters. Such evidence is worthwhile for Burmese given the influence of the pronunciation norms from accordingly-named Written Burmese on much of the modern writing system. In (10), the adjoining characters that signal the Modern Burmese open-syllable vowel /ɔ/ and the closed-syllable /au/ are shown side by side. The diphthong /au/ has no independent symbol, but is written with the same character for /ɔ/ followed by the velar consonants which are no longer pronounced. The orthography reflects an era when the seven syllable rimes in (10) were seen as sharing a single vowel quality, with diphthong [au] as a closed-syllable allophone of open-syllable /ɔ/.

<table>
<thead>
<tr>
<th>(10) Orthographic Evidence for Allophony During WB Era</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV</td>
</tr>
<tr>
<td>Low</td>
</tr>
<tr>
<td>Diacritic</td>
</tr>
<tr>
<td>Combined</td>
</tr>
<tr>
<td>Form</td>
</tr>
<tr>
<td>[kɔ]</td>
</tr>
<tr>
<td>Other Consonants</td>
</tr>
</tbody>
</table>

In this way, the open vs. closed syllable vowel quality distinction resembles the allophony envisioned by Green (2005): [ɛ] and [ɔ] in open syllables, complementary [ai] and [au] in syllables closed by ? or N codas. However, considering the diachronic derivations of the other closed-syllable vowel qualities, there is little reason to argue that (8) and (9) represent isolated synchronic processes rather than a prior process for which the output forms have created fossilized rimes.
Bennet and Lehman (1994) state an observation which suggests some productive association between diphthong vowels and velar consonants. They claim that pre-pausal CVN syllables are often realized with the N quality not just as vowel nasality, but as a faint excrescent nasal coda consonant. The noteworthy part of their observation is that this excrescent consonant is produced as the velar [ŋ] after the diphthong vowels (ai, ei, au, ou), but as coronal [n] after the monophthong closed-syllable vowels (i, a, u).

Similarly, Maran (1971) reported that the conservative Northern Burmese dialect in his description retained the coda consonants [k], [n], and [ŋ], but that, “Diphthongs occur only if the final is a velar consonant (Maran 1971: 104).” Examples include: [anŋ] ‘lake, pond’, [aŋk] ‘to be hot, humid’, [taŋ] ‘hill’, [pʰauŋk] ‘to open or bore a hole’ (ibid.).

Also of note in (10) is that [au] rimes, unlike other vowels, occur in writing only with velar codas. Closed syllable rimes with other vowel nuclei in Burmese often retain the full range of coda consonants, such that a MB /ŋ/ or /N/ phonological coda has a one-to-many correspondence to orthographic codas, i.e. {•δ, -p}, {•δ, -t}, {•δ, -c}, and {•δ, -k} all represent /ŋ/; while {•δ, -m}, {•ʃ, -n}, {•δ, -ŋ}, and {•δ, -ŋ} represent /-N/. For one, this orthographic evidence limits the historical allophony of [ɔ ~ au] to velar-closed rimes, putting the alternation in line with the pre-velar diphthongization of other vowels. Additionally, the restricted orthographic codas for [au] are evidence for the late stage of the diachronic change from Old Burmese diphthong /ea/ → Middle Burmese monophtong /a/, since the change would have to have occurred after the reduction of codas to only velar /k/ and /ŋ/.
To conclude this section’s discussion, the history of the Burmese open and closed-syllable vowels demonstrated how multiple asynchronous changes created the present distinction, thereby providing a natural explanation of asymmetries in the vowel inventory. Many of the vowel pairs bear surface similarities were shown to have little relation to one another, in particular /e/~/eɪ/ and /o/~/oʊ/, which were assumed to be allophones by Mehnert & Richter (1976) and Green (2005). Even at the historical stage of the language when diphthongization was actively triggered by velar codas, these pairs did not constitute allophonic counterparts as closed-syllable /eɪ/ and /oʊ/ developed not from /e/ and /o/, but from earlier forms of /i/ and /u/ respectively.

It is understood here that diachronic records do not definitively rule out synchronic interpretation by Burmese learners and speakers. That is, because /e/~/eɪ/ are not historically related does not preclude a speaker of MB from associating the two complementary vowels. However, they have very little reason to make such an association given the following three facts:

a) There are no morpho-lexical clues suggesting one is derived from the other. Green’s (2005: 11) analysis of /tʰou/ → /tʰo/ in open syllables offers no logical reason, such as morphological alternates, for the underlying vowel to not simply be /o/.

b) The set of alternations supposed by an allophonic account cannot be adequately described as a single effect. That is, [a]~[a], [ɛ]~[eɪ], and [ɔ]~[au] are not easily characterized as a single alternation of features between the first
vowel in each pair (in open syllables) and the latter vowel quality induced by any coda segment or vowel nasality.

c) The asymmetry of closed /ɛ/ and /au/ vowels to a single open-syllable counterpart offers no clear synchronic understanding.

The phonological analysis of all tone and tonal phonetic properties that follows (Chapter Seven) assumes the underlying set of vowels argued for in this section: fifteen underlying vowels with distributions restricted to either open or closed syllables.

In addition to the restricted distribution of diphthong vowels to closed syllables, another major piece of the phonological puzzle concerning syllable types is the length of the open and closed vowels, and by extension, of the modern coda segments /N/ and /ʔ/. The following section describes an acoustic analysis of these durations which addresses questions of their phonological representation.

5.4 Prosodic Structure of Burmese Syllables: Acoustic Evidence

5.4.1 Introduction to Statistical Tests of Prosodic Structure

In this section, phonetic data concerning the realization of CVN and CVO syllables is brought to bear on the account of syllable structure proposed by Green (2005) and alternatives to this model. Green’s (2005) account offered a concise use of existing Optimality Theoretic machinery to derive the restrictive syllabic and metrical structure seen in Burmese. There were, however, a number of shortcomings in the analysis concerning the representation of diphthongs and the nasal codas, particularly the moraic
weights assigned to them. The central hypothesis tested in this section (Hypothesis Five below) tackles how coda segments (nasal and glottal) are realized and whether they have moraic weight. These issues, in turn, speak to the interaction between syllable structure and vowel quality and the phonological representation of the diphthong vowel set.

**H5. Tone-bearing syllables in Burmese are bimoraic and phonetically realized codas bear moraic weight.**

Hypothesis Five (H5) is born from both the discussion of Green’s (2005) model and the robust vowel duration findings reported in §4.2. While Creaky vowels in connected speech were consistently shorter than High and Low tone vowels, the difference was roughly 15-30 ms whereas Checked vowels were frequently half the duration of the other tones (or about 100 ms shorter). Broadly, the distribution for oral vowels was a two-way split, (High, Low, Creaky) > (Checked), so that extremely short vowels and closed syllables were exclusive to one another. It is argued here that the duration difference between the tones is not well characterized as a phonological distinction in vowel length, but is better defined by syllable type (open or closed), explaining most of the difference as the influence of weight-bearing coda consonants. Hypothesis H5 extends the open vs. closed contrast to the nasal vowel set and predicts that nasal coda segments have moraic weight, but that open nasal CV̂ syllables have bimoraic vowels and no weight-bearing coda, contra Green (2005). In a language where codas are weightless, vowel length is unchanged by the presence or absence of coda consonants (see the analysis of Malayalam vowel length in Broselow et al 1997). Accordingly, phonological N codas can be seen as weightless and not heading a mora if vowel length in phonetically-open CVN syllables is
not substantially different from that in CV syllables with the same tone. On the other hand, occurrence of short vowels only in phonetically-closed CVO and CVN syllables suggests that Burmese coda consonants have moraic weight. Hypothesis H5 predicts that vowel durations are shorter in phonetically-closed syllables while the overall duration of all bimoraic syllables is roughly equal. These predictions are explicitly stated in (11) regarding measures of duration of the vowel nuclei and overall syllable length.

\[
\begin{align*}
(11) \text{Vowel Duration:} & \quad CV, C\text{̃} > CVN, CVO \\
\text{Syllable Duration:} & \quad CV = CVN = CVO
\end{align*}
\]

Implicitly, the relationships in (11) also indicate that coda consonants have a significant measurable length in order to equalize durations of the entire syllable. Measures of the CODA as well as Vowel and Syllable Duration prove crucial to the confirmation or rejection of Hypothesis H5. Each type of duration is introduced in the sections that follow (§5.4.2 – 5.4.3) and brought to bear on the alignment of morae amongst the syllable types found in Burmese.

5.4.2 Phonetic Realization of Coda Segments

Before showing that CVO and CVN codas have weight, it is necessary to demonstrate that they simply have a phonetic realization at all, an issue not as trivial as it may sound. Consider that in isolation, no nasal (N) or glottal (O) coda is apparent in the acoustic record, while in medial-position the N or O is rarely pronounced as a distinct segment. Rather, it is adjoined or assimilated to the following syllable’s onset, as attested in (12) below for the Carrier Phrases 2, 3, and 7 where the embedded position is followed by an obstruent stop onset, either [k] or [b]. The forms illustrate how the consonant juncture
presents a challenge for the measurement of CODA DURATION and how it must necessarily be examined with measures of the coda-onset inter-nucleic span [CVC.CV].

(12) Phonetic Forms of Glottal and Nasal Codas in Juncture

a. Phrase (2) [ʧəndə CV(O/N) ko jeinəde]

/taʔ/ + /ko/ → [tak.ko] Complete Assimilation of Checked Coda

/taN/ + /ko/ → [tan.go] Nasal Place Assimilation

b. Phrase (3) [tʰaʔ pjó CV(O/N) ba] || Phrase (7) [kó CV(O/N) bəməlo pjó]

/taʔ/ + /ba/ → [tap.pa] Complete Assimilation of Checked Coda

/taN/ + /ba/ → [tam.ba] or [tam.ma] Nasal Place Assimilation

Therefore, reference to the measure of CODA DURATION in this chapter and subsequently refers to the length of this consonant sequence, whether formed as a singleton, geminate, or homorganic nasal-obstruent cluster. The data examined below establish that in positions where phonological codas are reported to surface in Burmese, closed syllables have a greater consonantal duration than open syllables. The comparison was performed using data from two Carrier Phrases, (2) and (3), where the following onset consonant was [k] or [b], respectively. The measured portion of each token is given in bold in (13) along with the predicted findings assuming that phonological [ʔ] and N codas are indeed realized. For glottal codas, the comparison was performed against all open syllables in the data set, bearing Creaky, High, and Low tones (or rather, the duration of the ensuing onset following these open syllables). For the CVN data, CODA DURATION was compared only within Tone (i.e. Creaky CVN vs. Creaky CV).
Glottal Coda Coda Duration

The consonant span was routinely longer in Checked syllables than the onset consonant alone after open syllables, demonstrating the physical reality of Checked syllable codas. By $t$-test comparing the mean consonant duration for every speaker, the CVO.OV internal consonant duration was significantly longer than the CV.ko consonant for 9 of 10 speakers and CV.ba for all ten speakers.

### Table

Comparison by speaker of mean duration in milliseconds (given with standard deviation (in parentheses) and n (italicized)) for each consonant sequence. Nine of ten speakers produced a significantly longer intervocalic consonant /k/ span in Phrase 2 when the preceding syllable bore a glottal coda. For all but Speaker J (in shading), the difference between CVk.kV and CV.kV durations was roughly 2:1.
The means for each speaker and t-test results are provided in the tables in (14) and (15). The comparisons in these tables show that the consonant spans including a coda were not just significantly longer, but were roughly twice the length of a solitary onset consonant, quite literally reflecting the difference between a singleton and geminate. This was the case but for speaker J, who had equal length [k.k] and [k] spans in Phrase 2. The same speaker, however, did produce geminate-like durations for the [p.p] sequence in Phrase 3.

| (15) CODA DURATION following CVO vs. CV Syllables, Carrier Phrase 3 |
|---------------------|-------|-----|------------------|-------------------|
| CV.p.pV | CV.bV | Sig? | t Test comparing means |
| Speaker A: 104.0 (35.8), 4 | 55.3 (10.9), 12 | ✓ | t(14) = 4.40 p = .001 |
| Speaker B: 116.6 (13.2), 2 | 31.3 (3.5), 6 | ✓ | t(6) = 16.69 p = .000 |
| Speaker C: 103.6 (21.0), 4 | 53.9 (10.2), 12 | ✓ | t(14) = 7.20 p = .000 |
| Speaker D: 110.8 (19.2), 4 | 38.0 (11.6), 12 | ✓ | t(14) = 9.27 p = .000 |
| Speaker E: 126.0 (23.3), 4 | 44.6 (20.7), 12 | ✓ | t(14) = 6.63 p = .000 |
| Speaker F: 103.1 (25.9), 4 | 39.2 (11.7), 8 | ✓ | t(10) = 6.06 p = .000 |
| Speaker G: 107.0 (24), 4 | 46.8 (22.7), 11 | ✓ | t(13) = 4.49 p = .001 |
| Speaker H: 58.1 (13.5), 4 | 32.1 (11.2), 9 | ✓ | t(11) = 3.65 p = .004 |
| Speaker I: 89.1 (2.8), 4 | 39.5 (15.1), 12 | ✓ | t(14) = 6.39 p = .000 |
| Speaker J: 86.9 (8.7), 4 | 39.7 (10.5), 12 | ✓ | t(14) = 8.08 p = .000 |

**Table.** Comparison by speaker of mean duration in milliseconds (given with standard deviation (in parentheses) and n (italicized)) for each consonant sequence. All ten speakers produced a significantly longer intervocalic consonant /b/ span in Phrase (3) when the preceding syllable bore a glottal coda. The difference between CV.p.pV and CV.bV durations was roughly 2:1.

The waveforms in Figures 1 and 2 offer an illustrative example of this difference with two sequences from Phrase (1) produced by Speaker A. Figure 1 shows a 61.2 ms stop, the onset of the following syllable [ko], while in Figure 2 the same position reveals a 125.4 ms closure coinciding with the /ʔ/ coda and ensuing /k/ onset.
Figure 1. **Waveform demonstrating a single consonant duration.** Segmentation marks the 61.2 millisecond duration of the obstruction formed by the onset [k] in the sequence /ta/ + /ko/, as spoken by Speaker A in Carrier Phrase 2.

Figure 2. **Waveform demonstrating a two consonant duration.** Segmentation marks the 125.4 millisecond consonantal obstruction formed by the coda and onset stops of the sequence /tàʔ/ + /ko/, not including the burst in [ko]. Sample uttered in Carrier Phrase 2 by Speaker A.

In addition to the duration difference, there was an evident voicing distinction between the groups, visible in Figures 1 and 2: Checked codas produced voicelessness during the obstruent sequence (Fig. 2) and the open syllables yielding a voiced onset (Fig.1). For Phrases (2) and (3) specifically, the Checked-initial sequence was realized [tàk.ko] or [tàp.pa] respectively, while the onset in CV-initial sequences was voiced, [ta.go] or [ta.ba]. These data match the descriptions of intervocalic voicing of initial consonants in juncture detailed by Sprigg (1957) and Okell (1969). The voicing
distinction was made by all speakers with both [k~g] and [p~b] onsets, including the one case where no duration difference was found – Phrase (2) data for Speaker J. While duration measures presented in (14) did not indicate a geminate [kk] in Speaker J’s data, there was still a trace of the [ʔ] coda as Checked tone sequences were produced with a voiceless [k] whereas glottal vibration continued through the stop onset [g] which followed open syllables.

To summarize, Checked syllable codas had two identifiable phonetic manifestations realized with the following onset: an extended consonantal duration and voicelessness with obstruent onsets.

*Nasal Coda CODA DURATION*

Turning to the nasally-closed syllables, it was likewise found that nasal codas were associated with an increased CODA DURATION, though the effect was of lesser magnitude than seen with glottal codas. The durations for each speaker are on display in (16) and (17), where Low tone CV and CVN syllables are compared in Phrases (2) and (3). Summarized results for High and Creaky tone syllables are given in the table in (18).

For Low tone productions by all ten speakers, the CVN.CV consonant sequence had a greater duration than the singleton CV.CV, though not significantly so for two of the ten speakers with Phrase 2 (\([n.g]\) vs. \([g]\)) and for a single speaker with Phrase 3 (\([m.b]\) vs. \([b]\)). This was a less robust difference than that found above between CVO and CV CODA DURATION, a result that could be partly attributed to less precise segmentation between the Vowel-Nasal transition in CVN syllables, in contrast to the distinct boundary.

---

\(^3\) Full results for High and Creaky tone CVN durations can be found in Appendix II.
### Table. Comparison by speaker of mean duration in milliseconds for each NC consonant sequence with an initial Low tone in Phrase (3). Only Speaker H, in shading, with very limited data did not have significantly longer consonant closures following a CVN syllable.
formed by the complete closure of the obstruent stops following CVO syllables in juncture. However, looking at (16) one can see that the speakers with non-significant differences had means which trended towards longer CVN sequences and approached significance. Truly weak differences between the CV and CVN syllables were only found with the Creaky tone (cf. 18).

High tone CV and CVN syllables yielded similar results to the Low tone seen in (16) and (17). For Phrase (3), all ten had significant \text{CVm.pV > CV.bV} duration differences by \( t \) test at \( p < .05 \), while the \text{CVŋ.kV > CV.kV} difference in Phrase (2) was significant for seven of ten speakers, with the three non-significant comparisons trending or approaching significance for greater CVŋ.kV durations.

However, the difference was somewhat reduced with Creaky tone syllables, being both smaller in millisecond durations and less consistent across speakers. While the mean duration of CVN.CV spans was greater than the singleton duration for every speaker in both Phrase (2) and (3), it was only significantly greater by \( t \) test for three of ten showing \text{CVŋ.kV > CV.kV} in Phrase (2) and for seven of ten for \text{CVm.bV > CV.bV} in Phrase (3). For some speakers, the Creaky N was \( 1\frac{1}{2} - 2 \) times the duration of the singleton onset, as was found for High and Low CVN codas, but other subjects showed a much slighter difference in \text{CODA DURATION}. For example, Speaker J had a mean \text{CODA DURATION} of 45.6 (s.d. 11.3, \( n = 4 \)) for CV syllables and only slighter longer 53.6 ms (s.d. 6.8, \( n = 4 \)) for CVN syllables.
(18) Summary of Consonant DURATION Results for Creaky and High Tones

Number of Speakers for whom CVN codas had greater duration and $t$ test $p$-values for speakers with non-significant results.

<table>
<thead>
<tr>
<th>Phrase (2)</th>
<th>CVŋ.kV &gt; CV.kV</th>
<th>High Tone</th>
<th>Creaky Tone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td># of 10</td>
<td>$p$-values of $ns$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7/10</td>
<td>.055</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.071</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Phrase (3) | CVm.bV > CV.bV | 10/10 | 7/10 | .147 | .054 |
|            |                |       |       | | |

Results for CODA DURATION can be summarized as follows. Overall, the codas of CVO and CVN syllables are phonetically realized as segments with a measurable duration. However, this is only the case when a following syllable provides an onset consonant to which the coda can assimilate. Without this juncture, the O and N codas have no discernable acoustic manifestation separate from the vowel.

Checked syllable codas conclusively produced a longer consonantal closure than in open syllables. The assimilated obstruent coda and following onset were between 30 – 80 ms longer than the onset alone. CVN syllable nasal codas were similarly longer, though somewhat tenuously with the Creaky tone, particularly for the CVŋ.kV cluster in Phrase (2). This was because Creaky nasal codas were generally shorter than for the other tones, making their mean durations less distinct from the CV.kV onset alone following open Creaky syllables. Reasons for this pattern in the data are discussed further in §5.4.4.
5.4.3 Vowel Length: Bimoraic, Monomoraic, and Shared Mora Vowels

Finally, measures of Vowel Duration in open and closed syllables offer a clear picture of coda weights in CVO and CVN tokens. For CVO syllables, it was shown in §4.2 that vowels in Checked syllables were consistently shorter, for all speakers in all environments: citation forms, phrase-finally, between tone-bearing syllables, and before multisyllabic words with initial reduced syllables. In the present section, Vowel Duration in CVN syllables is revealed to be similarly curtailed in frames where the phonological N coda is phonetically realized as a nasal segment. On the other hand, when the CVN syllable type’s nasality is realized solely as a quality of the vowel, CṼ, the Vowel Duration is in line with vowels in open CV syllables.

Recall that the main hypothesis of this section proposes that Coda segments have weight, and are therefore predicted to affect the length of their nucleus. This prediction was stated in (11) and is repeated below:

\[(19) \text{Vowel Duration: } CV, CṼ > CVN, CVO\]

To confirm the ranking in (19), Vowel Duration was compared in three environments which elicit both CṼ and CV[n] forms of Burmese nasal coda syllables:

- Isolation (Phrase 1 tokens)
- Phrase-finally (Phrase 5 and 6 tokens pooled)
- Phrase-medially (Phrase 2, 3, and 4 tokens pooled)

Means for each tone in each environment were compared within and across speakers.

To pool all speakers, the same standardization procedure and parameters used in §4.2 were employed to convert ms measures to z-scores, which were in turn
retransformed for illustrative purposes as though from a single hypothetical speaker resembling Speaker I’s production. Pooled subject data for each environment is given in Figure 3. The durations of Checked vowels, the only vowels in CVO syllables, are provided with the Creaky CV and CVN means.

**Figure 3. Mean Vowel Duration comparing CV (Open) and CVN (Nasal Coda) syllables.** Compares tone as an isolated syllable, in phrase-final and medial position. Vowels in CVN syllables have a shorter duration when the N coda segment is phonetically realized. Durations are in milliseconds, standardized across speaker data.

The graphs in Figure 3a and 3b show the citation form and phrase-final productions of CVN syllables where N is realized not as a segment, but as vowel nasality. In these cases, the duration of vowel nuclei in open CV and phonologically-closed, but phonetically-open CVN syllables were not substantially different. Statistical results of t tests comparing the means (shown in both z and retransformed ms) reveal that there was no statistical difference in Vowel Duration between the two for either Creaky or Low tone syllables. These are given in (20) and (21) below. Contrastively, the results in (22) demonstrate that in juncture, with an evident nasal stop coda, CVN vowels of every tone were significantly shorter than their open syllable counterparts.
(20) **Isolated Syllables: Differences in Vowel Duration**

<table>
<thead>
<tr>
<th></th>
<th>CV</th>
<th>CVN</th>
<th>t Test comparing means</th>
</tr>
</thead>
</table>
| **Creaky:** | -0.575 (0.80) | -0.573 (0.53) | CV ≈ CṼ  
204.8 ms 204.9 ms  
|         | t(149) = 0.02  p = .986 |        |
| **Low:**  | 1.234 (0.67)     | 1.031 (0.90)     | CV ≈ CṼ  
331.4 ms 317.2 ms  
|         | t(161) = 1.62  p = .108 |        |
| **High:** | 1.482 (0.66)     | 1.127 (0.76)     | CṼ > CṼ  
348.7 ms 323.9 ms  
|         | t(151) = 3.07  p = .003 |        |

(21) **Phrase-Final Syllables: Differences in Vowel Duration**

<table>
<thead>
<tr>
<th></th>
<th>CV</th>
<th>CVN</th>
<th>t Test comparing means</th>
</tr>
</thead>
</table>
| **Creaky:** | -0.656 (0.68) | -0.666 (0.54) | CV ≈ CṼ  
199 ms 198.4 ms  
|         | t(151) = 0.10  p = .923 |        |
| **Low:**  | 0.335 (0.82)     | 0.302 (0.97)     | CV ≈ CṼ  
268.5 ms 266.1 ms  
|         | t(145) = 0.23  p = .822 |        |
| **High:** | 0.602 (0.73)     | 0.303 (0.62)     | CṼ > CṼ  
287.2 ms 266.2 ms  
|         | t(142) = 2.64  p = .009 |        |

(22) **Phrase-Medial Syllables: Differences in Vowel Duration**

<table>
<thead>
<tr>
<th></th>
<th>CV</th>
<th>CVN</th>
<th>t Test comparing means</th>
</tr>
</thead>
</table>
| **Creaky:** | -0.620 (0.33) | -1.225 (0.44) | CṼ > CṼ  
201.6 ms 159.3 ms  
|         | t(231) = 11.89  p = .000 |        |
| **Low:**  | -0.468 (0.34)     | -1.123 (0.37)     | CV > CVN  
212.2 ms 166.4 ms  
|         | t(223) = 13.64  p = .000 |        |
| **High:** | -0.233 (0.37)     | -0.932 (0.45)     | CV > CṼN  
228.7 ms 179.7 ms  
|         | t(223) = 12.68  p = .000 |        |

Note that the duration difference between open and nasal syllable *High* tone vowels was significant in all three measured environments. In fact, for all three tones in CVN
syllables, it is apparent in the pooled means that CVN vowels were shorter in all environments. Clearly, there was some tendency for nasal vowels to be slightly shorter, but this tendency was weaker than the predictable shortening of vowels in phonetically closed syllables as conveyed by statistical significance for the Low and Creaky tokens. For the High tone, while significance was found in all positions, the magnitude of the difference was considerably greater in medial position (see below). Furthermore, in speaker-by-speaker data, no trend holds regarding open oral and nasal syllable lengths. On an individual basis, speakers were as likely to produce a longer duration vowel with Creaky and Low C̃ as with CV tokens (Figure 4 below).

Concerning the significance of High tone differences in Isolated and Final position, it is not surprising to find a consistent, though relatively small difference between two different vowels given more than seventy samples of each. However, Isolated and Final nasal C̃ High vowels were shorter than oral CV vowels by a fairly small amount (25 ms in the retransformed standardized scale), while the difference between Medial C̃N and CV vowels was greater (roughly 50 ms) and more consistent (reflected in a substantially higher t score, >12 medially, but 3.07 (Isolated) and 2.64 (Phrase-Final) or see details in (20-22)). This difference indicates a durational effect of N on High tone vowels equivalent to that distinguished by statistical significance on Low and Creaky vowels.

**Inter-speaker Variation**

The pattern presented in Figures 3a-c above for all pooled VOWEL DURATION data was also upheld individually by speakers. In medial position, CV vowels were significantly
longer by \( t \) test than nasal-coda CVN vowels for all but one speaker (C) with Creaky tone vowels and another speaker (D) for Low vowels\(^4\). Within High tone data, CV vowels were longer for every speaker. The speaker-by-speaker breakdown of phrase-medial VOWEL and CODA DURATIONS for Checked CVO and Creaky CV/CVN syllables is shown in Figure 4.

![Figure 4. Mean Vowel Duration and Coda Duration for each speaker. Values are compared for Checked (CVO), Creaky Open (CV) and Creaky Nasal Coda (CVN) syllables produced in phrase-medial position, where O and N codas are phonetically realized in juncture with the following onset.](image)

\(^4\) VOWEL DURATION was not different in medial positions for…
Creaky CV vs. CVN with Speaker C (\( t(22) = 1.041, p = .309 \)).
Low CV vs. CVN with Speaker D (\( t(22) = 1.80, p = .085 \)).
The lack of difference between phonetically-open oral CV and nasal C\(\tilde{V}\) durations was also consistent across speakers. There was a trend towards longer oral CV than nasal with High tones across speakers, but the difference was only significant for two speakers with CV, C\(\tilde{V}\) syllables produced in isolation. No such pattern existed for Creaky and Low tone vowels which speakers were as likely to have a greater Vowel Duration when open and nasal as when oral. Statistical results by speaker are summarized in (23), showing how many of the ten subjects produced CV and C\(\tilde{V}\) vowels with similar durations and the t test results for the few cases where there was a significant difference in duration. Note though that significant differences for the Creaky tone data were not in the same direction – C\(\tilde{V}\) > CV for one speaker and CV > C\(\tilde{V}\) for another.

(23) Summary of Speaker Similarity for Vowel Duration in Open CV and C\(\tilde{V}\) Syllables

<table>
<thead>
<tr>
<th></th>
<th>Creaky Tone</th>
<th>Low Tone</th>
<th>High Tone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isolated</td>
<td># of 10</td>
<td>*</td>
<td># of 10</td>
</tr>
<tr>
<td>CV &gt; C(\tilde{V})</td>
<td>8/10</td>
<td>F (p = .002)</td>
<td>10/10</td>
</tr>
<tr>
<td>CV &gt; CV</td>
<td>E (p = .026)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Phrase-Final  

CV > C\(\tilde{V}\)  
D \(p = .001\)  
CV > CV  
I \(p = .011\)  

This concludes the presentation of findings for segment durations. Results not listed in this chapter for reasons of space are given in Appendix II, including Nasal CODA Durations by speaker for High and Creaky syllables.
5.4.4 Evaluation of Hypothesis Five

The previous sections have shown that when coda segments are present, the Vowel Duration of CVO and CVN syllables was reliably reduced. The reduction was found to be greater in CVO than in CVN syllables. Further, amongst CVN syllables, the vowel was shortened the most in Creaky tone data. When no coda segment was present, Vowel Duration did not differ between oral and nasal syllables in Creaky and Low tone forms, though High tone CV vowels were slightly shorter than open oral CV vowels. Putting aside the High tone interaction with open syllables, the findings fit the proposed hierarchy in (11) with the added distinction that obstruent-closed vowels were predictably shorter than nasal-closed vowels.

(24) **Vowel Duration**:  CV, ĈV > CVN > CVO

If, following hypothesis H5, all tone-bearing vowels are bimoraic and coda segments bear weight, a two-way distinction might be expected: open syllable bimoraic vowels vs. closed syllable monomoraic vowels. The three-way distinction in (24) is accommodated to the model by assuming different moraicity for obstruent and sonorant codas. The representations in (25a - c) demonstrate how obstruent codas head a mora while sonorant codas share a mora with the vowel nucleus.

(25)  a) Bimoraic Vowel  b) Sonorant Coda  c) Obstruent Coda

\[
\begin{align*}
\text{(25a)} & \quad \text{Bimoraic Vowel} \quad \sigma \\
& \quad \mu \quad \mu \\
& \quad C \quad V \\
\end{align*}
\[
\begin{align*}
\text{(25b)} & \quad \text{Sonorant Coda} \quad \sigma \\
& \quad \mu \quad \mu \\
& \quad C \quad V \quad N \\
\end{align*}
\[
\begin{align*}
\text{(25c)} & \quad \text{Obstruent Coda} \quad \sigma \\
& \quad \mu \quad \mu \\
& \quad C \quad V \quad O \\
\end{align*}
\]
The representations in (25) model a direct relationship between phonetic duration and associations to morae. The shared morae of sonorant codas should correspond to a lesser Vowel Duration than in bimoraic vowels and a lesser Coda Duration than found for codas heading their own mora. The results above revealed precisely this relationship. Vowels preceding sonorant codas had a Vowel Duration roughly halfway between those of open syllable vowels and obstruent-closed vowels, while the sonorant codas themselves had briefer Coda Duration measures than obstruent codas.

Another general pattern in the data was that Creaky tone sonorant codas were consistently shorter than with the other tones. However, the Vowel Duration results in §5.4.3 revealed that the vowel alone in a CVN syllable was not significantly shorter when Creaky than when bearing High and Low tones. As Creaky vowels are known to be shorter in open syllables, it seems that this shortness is realized instead by the sonorant coda, in this case with a shorter nasal stop segment. If the brevity of Creaky CV syllables is tied to the second mora (via a [+Constricted Glottis] feature, as in §7), then it is expected that the same effect would be also found during the second mora of CVN syllables. The moraic alignment of (25b) places this effect on the nasal coda segment (i.e. the second mora).

5.5 Discussion of Syllable Shape Findings – Vowel Quality and Prosody

Sections 5.3 and 5.4 have examined two mostly unassociated issues of Burmese phonology: the historical development of vowel quality and the duration of syllabic components. Both topics addressed the phonological representation of the possible
Burmese rimes, providing insight which offers the most accurate, and therefore interesting, model of how syllabic shape varies according to tone and vowel quality.

Assuming all major syllables (stressed and tone-bearing) to be bimoraic, the duration of vowels and glottal and nasal codas portrayed a system whereby, in connected speech, all tone-bearing syllables had roughly equal durations composed of varying vowel and coda lengths. An achievement of the present chapter, along with the vowel duration data of Chapter Four, was to derive the well-known duration contrasts between the tones from the syllable shape rather than treat is as an intrinsic quality of the tone. By demonstrating the moraic weight of nasal codas, the brief duration of Checked vowels was effectively argued to be due to monomoraicity, since the second mora was headed by the obstruent coda. The understanding of possible syllabic shapes advocated here is employed in Chapter Eight, serving as a template upon which the variety of other tonal contrasts are realized. The underlying representation of CV, CVN, and CVO templates was provided by the trees in (25), which are adapted in (26) with /təʊN/ ‘South’ and (27) with /təʊʔ/ ‘flame, blaze’ to illustrate the proposed alignments of nasal and glottal codas with and without an adjoining syllable onset (in this case, a following /ko/).
5.6 Effects of Context on F0: Statistical Evidence

5.6.1 Overview of Statistical Analysis of F0

The analysis of F0 tests four effects of Carrier Phrases noted in §4.3, summarized in (4.54) – (4.56), and repeated here. An additional effect of context, the neutralization of F0 falls phrase-medially, is explored separately in §5.7 with accompanying Open Quotient data.

(28) Anticipatory co-articulation of F0 was minimal.

(29i) Perseverative co-articulation of F0 following a High tone was robust.

(29ii) Downstepped pitch was found following a High tone (after the co-articulated vowel onset, post-High tokens were generally lower-pitched).

(30) Creaky and Checked Tone F0 contours were not distinguishable.

Statistical tests of the contextual effects compared pairs of minimally different Carrier Phrases. It was decided not to pool data from all phrases with a relevant characteristic (e.g. a preceding High tone) in order to avoid confounding effects of other phrasal differences. That is, to test the effects of a preceding High vs. preceding Low tone, the following sets of means were compared separately:

(31a) Phrase 2 [Low __ Low] vs. Phrase 3 [High __ Low]

b) Phrase 5 [High __ #] vs. Phrase 6 [Low __ #]

Blocking the data as such separated the potential overlapping effects of adjacent tones and phrase-final boundaries, both hypothesized to induce more precipitous F0 falls and
thus greater negative $\Delta F_0$ and $m_{F_0}$\textsuperscript{5}. If both effects actively lower F0, then pooled phrase-medial and phrase-final data would confound the analysis.

Finally, due to the Carrier Phrases composing the data set, co-articulation of adjacent tones was tested only with a preceding High vs. Low tone contrast. The co-articulation between consecutive Creaky or Checked tones was not examined.

The overall claim made here is that phonetic co-articulation in Burmese is generally perseverative and not anticipatory. Over the early portion of a syllable, F0 results showed the influence of the prior syllable’s F0 targets. On the other hand, late F0 values rarely fell or rose in anticipation of a following syllable’s tone.

5.6.2 Anticipatory Co-articulation of F0: The Effects of a Following High Tone

In the F0 traces of Chapter Four, a following High tone appeared to have little to no influence on the pitch of the prior syllable. The ensuing pitch rise accompanying the High tone was not anticipated during the prior syllable. Hypothesis Six tested for such an anticipatory effect.

**H6. The offset F0 of a vowel is not different whether the following syllable bears a High or Low tone.**

Offset F0 levels were represented by the set of F0\textsubscript{LATE} values, which were compared from tokens embedded in the minimally contrastive Carrier Phrases (2) and (4):

- Phrase 2 [Low __ Low] vs. Phrase 4 [Low __ High]

\textsuperscript{5} For example, it was noted in §4.3.2 that tokens embedded in Carrier Phrase (6) [Low __ #] were exceedingly low-pitched because they both (i) followed a Low tone and (ii) were phrase-final at the end of a longer sentence where declination generally produced lower F0 results. It was therefore not appropriate to examine the F0 effects of a preceding Low tone by pooling Phrase (6) data with other Low-initial frames.
The comparison was run as a linear mixed model with the two frames above as levels in the factor Phrase and with Speaker as a random effect. The results in (32) partially confirm the hypothesis, indicating a difference for some tones (Creaky, Checked) but not others (High, Low). $P$-values for the High and Low tone data supported the null hypothesis that $F_{0\text{LATE}}$ was not different under the two conditions, therefore supporting Hypothesis Six (H6) that adjacent tones did not affect F0 in this way.

\begin{table}[h]
\begin{center}
\begin{tabular}{llll}
\hline
\textbf{Phrase} & \textbf{2 [L ___ L]} & \textbf{4 [L ___ H]} & \textbf{Difference by F-test} \\
\hline
\textbf{LOW} & -0.551 & -0.623 & $F(1,78) = .816$ & $p = .369$ \\
\textbf{HIGH} & 0.528 & 0.447 & $F(1,78) = .283$ & $p = .597$ \\
\textbf{CREAKY} & -0.192 & -0.561 & $F(1,77) = 5.32$ & * $p = .024$ \\
\textbf{CHECKED} & 0.560 & -0.219 & $F(1,64) = 19.14$ & ** $p < .001$ \\
\hline
\end{tabular}
\end{center}
\caption{Comparison of mean $F_{0\text{LATE}}$ in tokens preceding Low vs. High tones. Significantly different means are shaded. * for $p < .05$, ** for $p < .01$.}
\end{table}

Results for the Creaky and Checked tones rejected the null hypothesis, but curiously the difference did not portray higher F0s leading into a High tone, but rather the inverse. The low F0 at the end of a Creaky or Checked tone falling contour was significantly lower before a High than before a Low tone. Since these findings do not support the notion that an upcoming pitch target elevated F0 at the offset of the preceding syllable, they are mostly put aside. However, to address and dispel the concern that the lower offset pitch is in fact a kind of perceptually-driven anticipation of the following high pitch, a likely explanation points to a segmental, rather than tonal, difference between the following syllables in Phrases (2) and (4): Phrase (2) followed the embedded token with [ba] while Phrase (4) followed with [lá]. The sonorant [l] onset of Phrase (4) meant that periodicity was maintained between the syllables allowing for more accurate automatic calculation of F0 at the embedded token’s offset. This is not to say that pitch...
was significantly lower because of a segmental distinction, but that the F0 autocorrelation method was generally more capable of reflecting F0 at the offset when it did not abruptly transition into an aperiodic obstruent (and a geminate obstruent in the case of Checked vowels, where the effect was strongest).

5.6.3 Perseverative Co-articulation of F0: The Effects of a Preceding High Tone

It was observed in Chapter Four that tokens embedded after a High tone had noticeably higher mean F0s at the vowel onset. To test statistically whether this was the case, early F0 scores were compared between the Carrier Phrases with High-initial frames and their counterparts with a Low-initial frame. These were:

- Phrase 3 [High __ Low] vs. Phrase 2 [Low __ Low]
- Phrase 5 [High __ #] vs. Phrase 6 [Low __ #]
- Phrase 7 [High __ minor] vs. Phrase 8 [Low __ minor]

The present section tests **Hypothesis Seven** concerning the F0 measures at the rhyme onset.

**H7. The onset F0 of a vowel bearing any tone is greater after a High tone syllable than after a Low tone syllable.**

The hypothesis was tested using a linear mixed model with a two-level factor of Phrase (High vs. Low-initial frame) and Speaker as a random effect. The data was blocked by tone, so four separate tests were run rather than including Tone as a factor in the model. Since the change in initial F0 was found right at the onset, the mean F0 at the 10% decile
point was used as the dependent variable to test **Hypothesis Seven** (H7). Results are given in (33).

![Table](image)

**Table.** Comparison of mean \( F_{0.10}\) in tokens following High vs. Low tones. \( P\)-values indicating significantly different means are shaded, while near significant differences which trend in the predicted direction are given diagonal backgrounds. * for \( p < .05\), ** for \( p < .01\).

For every tone, tokens embedded in a [High __ …] phrase had a higher \( F_{0.10}\), but the distinction was not always significant, particularly for Creaky and Checked tokens. Putting aside the robust difference for all tones found between Phrases (5) and (6) in (33c) (which were previously shown in findings for both F0 and Duration to differ greatly in focus), there was no difference found on Creaky and Checked syllables, but High and Low tones showed a significant or near significant effect of the preceding syllable’s tone on their onset F0. Looking at the means in (33a) and (b), the nature of this split in the
tones’ behavior is straightforward. High and Low tones were low pitched in the environment [L_] and moderate-to-high pitched in the environment [H__]. Creaky and Checked tones, on the other hand, were always high at their onset – in fact higher in either phrase than High or Low tones in [H__] frames.

These findings do not support hypothesis H7 for all tones, but the hypothesis is partially confirmed for the Low and High tones. The raised pitch of a preceding High tone consistently had a carry-over effect into Low or High tone embedded tokens, elevating F0 at their onset. An explanation for this split between the tones delves into the phonological specifications of each tone. Generally, the effect of F0 co-articulation is discussed here as a physical limitation and therefore a phonetic process. This process acts in accord with the phonetic implementation of the phonological inputs for each tone to create a continuous interpretation of a sequence of timed pitch targets. Accordingly, the data can yield conclusions about both phonetic (i.e. the “continuous interpretation”) and phonological matters (i.e. the timing of the pitch targets).

A phonological understanding which nicely interprets these findings posits an early pitch target in Creaky and Checked tones and not in High or Low tones. Since Creaky and Checked tones already have a pitch specification at or near the vowel onset, there is little room for interference from an adjacent tone. The onset of High and Low tone syllables, lacking a specification, are more dependent on the surrounding context for their surface realization. The resulting patterns of higher or lower pitch at onset can be attributed to different sequences of timed targets:

- The highest onset F0 values are linked to an early syllable high target (as with Creaky or Checked examples).
• The relatively high values are linked to transition periods to/from a high target, but not precisely a High target at the onset.

• Low onset F0s are associated with peak-less spans (such as [LLL] or [LLә] in Phrases (2) and (8) respectively).

This discussion forms a building block to the proposal in Chapter Seven, which again considers the surface realization of F0 over specified and unspecified spans.

5.6.4 Other Effects of a Preceding High Tone: Downdrift

In Carrier Phrases consisting of a High-initial frame, the embedded token consistently followed a syllable offset with an elevated pitch. In this position, it was not peculiar that an embedded High tone had an attenuated pitch rise or that an embedded Low tone saw a greater pitch fall (from the high initial value). In the case of a [High __ Low] frame, an embedded High tone forms a High-High-Low sequence. If an F0 peak were attained at the end of the first High in this sequence, a rise on the second High would be unexpected. Furthermore, an embedded Low tone offers no targets for maintaining a high pitch after the first syllable in a High-Low-Low sequence.

At issue though, is that the pitch drops seen in the traces of mean F0 in Chapter Four were greater than would be predicted from just interpolation in a [High __ Low] frame. Rather, it seems that after the initial carry-over of a sustained high pitch, a High tone has the tendency to lower the pitch of a following syllable in the data of Chapter Four. A test of Hypothesis Eight looks to confirm the statistical strength of this tendency by looking at F0 measures at the rhyme offset.
H8. The offset F0 of a vowel bearing any tone is lesser after a High tone than after a Low tone syllable.

Specifically, Hypothesis 8 (H8) examined differences in F0_{\text{LATE}}, the F0 for each token averaged over the 80-100% decile datapoints. It is predicted that F0_{\text{LATE}} is lower in a syllable which follows a High tone than in one following a Low tone. The hypothesis was tested using the same linear mixed model and set of paired data used for hypothesis H7 (a two-level factor of Phrase (High vs. Low frame) and Speaker as a random effect). Results are in (34).

<table>
<thead>
<tr>
<th></th>
<th>Phrase 3 [H __ L]</th>
<th>Phrase 2 [L __ L]</th>
<th>Difference by F-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOW</td>
<td>-1.373</td>
<td>-0.551</td>
<td>$F(1,78) = 137.3$ ** $p &lt; .001$</td>
</tr>
<tr>
<td>HIGH</td>
<td>-0.214</td>
<td>0.528</td>
<td>$F(1,79) = 30.59$ ** $p &lt; .001$</td>
</tr>
<tr>
<td>CREAKY</td>
<td>-1.043</td>
<td>-0.192</td>
<td>$F(1,77) = 43.18$ ** $p &lt; .001$</td>
</tr>
<tr>
<td>CHECKED</td>
<td>-0.354</td>
<td>0.560</td>
<td>$F(1,61) = 33.48$ ** $p &lt; .001$</td>
</tr>
<tr>
<td>Overall</td>
<td>-0.743</td>
<td>0.084</td>
<td>$F(1,301) = 179.2$ ** $p &lt; .001$</td>
</tr>
</tbody>
</table>

(34) Effect of a Preceding Low vs. High Tone on F0-{\text{LATE}} (as z-score)

<table>
<thead>
<tr>
<th></th>
<th>Phrase 3 [H __ L]</th>
<th>Phrase 2 [L __ L]</th>
<th>Difference by F-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrase 5 [H __ #]</td>
<td>Phrase 6 [L __ #]</td>
<td>$F(1,22) = 14.04$ ** $p = .001$</td>
<td></td>
</tr>
<tr>
<td>LOW</td>
<td>-0.932</td>
<td>-1.771</td>
<td>$F(1,78) = 54.06$ ** $p &lt; .001$</td>
</tr>
<tr>
<td>HIGH</td>
<td>-0.391</td>
<td>-1.477</td>
<td>$F(1,57) = 0.63$ $p = .431$</td>
</tr>
<tr>
<td>CREAKY</td>
<td>-1.366</td>
<td>-1.546</td>
<td>$F(1,69) = 2.71$ $p = .104$</td>
</tr>
<tr>
<td>CHECKED</td>
<td>-1.139</td>
<td>-1.487</td>
<td>$F(1,232) = 33.04$ ** $p &lt; .001$</td>
</tr>
<tr>
<td>Overall</td>
<td>-0.979</td>
<td>-1.583</td>
<td>$F(1,69) = 2.71$ $p = .104$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Phrase 3 [H __ a]</th>
<th>Phrase 2 [L __ a]</th>
<th>Difference by F-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOW</td>
<td>-0.852</td>
<td>-0.546</td>
<td>$F(1,78) = 10.94$ ** $p = .001$</td>
</tr>
<tr>
<td>HIGH</td>
<td>0.267</td>
<td>0.528</td>
<td>$F(1,79) = 3.80$ $p = .055$</td>
</tr>
<tr>
<td>CREAKY</td>
<td>-0.708</td>
<td>0.153</td>
<td>$F(1,70) = 16.86$ ** $p &lt; .001$</td>
</tr>
<tr>
<td>CHECKED</td>
<td>-0.080</td>
<td>0.570</td>
<td>$F(1,72) = 12.82$ ** $p = .001$</td>
</tr>
<tr>
<td>Overall</td>
<td>-0.337</td>
<td>0.170</td>
<td>$F(1,305) = 40.84$ ** $p &lt; .001$</td>
</tr>
</tbody>
</table>

Table. Comparison of mean F0_{\text{LATE}} in tokens following High, Low tones. Significant $p$-values are shaded. Nearly significant differences in means which trend in the predicted direction have diagonal backgrounds. * for $p < .05$, ** for $p < .01$. 
**Hypothesis Eight** was confirmed by the differences between phrase-medial contexts, but the phrase-final position of Phrases (5, 6) actually found strongly significant differences indicating the reverse association – a lower $F_{0\text{LATE}}$ after Low tones. It has been discussed previously that Phrases (5) and (6) did not minimally differ by just the High vs. Low initial tone in the carrier frame, but that they presented differing levels of emphasis which has confounded other Duration and F0 findings. The results in (34b) above express a similar confound and are thus disregarded concerning the present hypothesis.

Comparisons of Phrase (2) vs. (3) and Phrase (7) vs. (8) reject the null hypothesis that the contexts have no effect for each tone and thus confirm the observation tested by H8. Fundamental frequency at the syllable offset was indeed statistically more likely to be lower when the preceding syllable bore a High tone. This finding is interesting because the phonetic motivations are much less clear than for the perseverative coarticulation effect seen in §5.6.3. By using $F_{0\text{LATE}}$, rather than $\Delta F_{0}$ or $m_{F0}$, a potential explanation can be ruled out – that the difference was simply a consequence of the onset co-articulation in that a higher onset $F_{0}$ yielded a greater fall to the default level. Instead, two possible causes are discussed here that can broadly be labeled declination and downdrift.

*Declination* could explain the low offset pitch range as an artifact of the particular stimuli sentence. However this can only explain the $F_{0}$ difference between phrases in (34a) and not that in (34c). In Phrase (3), the following syllable was not only Low-toned, but also phrase-final, making the embedded token the penultimate syllable (i.e. $[\sigma^{\text{HIGH}} \sigma^{\text{LOW}} \#]$). This was not the case in Phrase (2), with which Phrase (3) was being compared. Section 4.3.5 demonstrated that phrase-finality typically lowered F0 and it
would not be surprising if the intonational effects of phrase-finality extended to the penultimate phrase position. In other words, the effect seen with Phrase (3) data was not just anticipation of a following Low tone, but the anticipation of a phrase-final Low tone.

A far different approach is to consider the F0 depression as a form of downstep, a cross-linguistically common tonal process (though mostly documented in African languages) by which consecutive tone bearing units do not employ the same pitch register (Schuh 1978, Clements 1979, Odden 1995). In this case, the Burmese High tone acts as a trigger to lower the overall F0 range in use, so that high pitch targets that follow a prior High are relatively lower after the downstep in range (or downdrift here in the terminology of Yip (2002) since the trigger is another overt lexical tone). More typically, Low tones trigger downdrift, but Odden (1995) discusses a case of downdrift between consecutive High tones. Since the downdrift effect seen in the Burmese data in Figure 4.16 applied to the pitch range employed for all tones and not just H pitch targets, it could be difficult to separate from the general effect of declination, essentially a physiological explanation for the penultimate phrase position account given above.

To summarize, the statistical evidence for downdrift was entirely lacking in final positions (34b). While the effects seen in (34a) are potentially confounded by declination to penultimate utterance position, similar effects were found in the comparison of (34c) which lacked the penultimate position confound. Still, the pattern is noted here but not explored further as the evidence necessary to posit downstep or some process of phonologized declination in Burmese extends beyond the scope of this study.
5.6.5 Differences Between the F0 Contours of Creaky and Checked Tones

Finally, it was noted repeatedly through Section 4.3.4 that the F0 contours of Creaky and Checked syllables were frequently similar. An analysis of the Scheffé results in (4.20) – (4.27) indicated that the two tones were not statistically different in ΔF0 nor \( m_{F0} \) in four of the eight Carrier Phrases: in (1), (5), and (6), representing all three frames with an empty following position, and in Phrase (7) with a following minor syllable. The summary in (35) separates these Phrases from those in which Creaky and Checked tokens were statistically different in at least one metric.

<table>
<thead>
<tr>
<th>Phrase</th>
<th>Frame</th>
<th>( p ) for ΔF0</th>
<th>( p ) for ( m_{F0} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>[# _ _ #]</td>
<td>.828</td>
<td>.940</td>
</tr>
<tr>
<td>5)</td>
<td>[H _ _ #]</td>
<td>.761</td>
<td>.099</td>
</tr>
<tr>
<td>6)</td>
<td>[L _ _ #]</td>
<td>.981</td>
<td>.512</td>
</tr>
<tr>
<td>7)</td>
<td>[H _ _ minor]</td>
<td>.091</td>
<td>1.000</td>
</tr>
<tr>
<td>2)</td>
<td>[L _ _ L]</td>
<td>** .002**</td>
<td>.364</td>
</tr>
<tr>
<td>3)</td>
<td>[H _ _ L]</td>
<td>** .000**</td>
<td>1.000</td>
</tr>
<tr>
<td>4)</td>
<td>[L _ _ H]</td>
<td>.210</td>
<td>** .005**</td>
</tr>
<tr>
<td>8)</td>
<td>[L _ _ minor]</td>
<td>* .011</td>
<td>.994</td>
</tr>
</tbody>
</table>

As seen in the lower half of (35), even when the two tones showed a statistical difference within a Carrier Phrase context, it was limited to a single measure. Differences in ΔF0 (found in medial Phrases 2, 3, and 8) were the result of a greater fall (i.e., a greater negative ΔF0) attested on the Creaky tone, though in these cases the mean negative slope for both was nearly identical. In only one of the eight Phrases (Carrier Phrase 4) was the mean \( m_{F0} \) for Creaky and Checked tones significantly different. In all, this comparison indicates that the F0 patterns for Creaky and Checked tones were indeed statistically similar, but also that this similarity was strongest in an isolated syllable or in phrase-final
position. Given the categorical segmental changes and additional F0 variability seen with these tones in juncture, the less robust findings in phrase-medial contexts were not unexpected.

5.6.6 Summary of Statistical Analysis of the Effects of Context on F0

The High tone was the only tone argued to undergo a contextually-driven pitch alternation. In citation form the High tone had a falling pitch contour, and phrase-final High tones also fell. Phrase-medial High tones held a high pitch throughout the syllable before another High tone, but rose before a Low tone syllable. For the other three tones, sentential context did yield consistent (and often statistically significant) effects, but the effects were incremental, rather than wholesale, changes. The general pitch contour of Low, Creaky, and Checked tones never altered and the means for the F0 variables (a) $F_{0\text{MAX}}$, (b) $\Delta F_0$, (c) $m_{F_0}$ held the same relative values across Carrier Phrases. That is, although the absolute values for $m_{F_0}$ systematically changed by context, the $m_{F_0}$ means were ranked Low $>>$ Creaky, Checked in every context. Therefore, the present analysis regards these contextual effects on F0 as gradient phonetic differences, grounded in the physical co-articulation of F0. The tests in §5.6.2-5.6.3 concluded that this gradient co-articulation between adjacent tones was more likely to be perseverative than anticipatory, though this observation is limited to the Low and High tones at present, as no Carrier Phrase in this study employed surrounding words bearing a Creaky or Checked tone.

An excellent portrayal of this gradient change in F0 can be seen with the onset F0 figures in Low tone tokens (cf. Fig 4.11 – 4.19). Initial F0 was strongly conditioned by the tone of the prior syllable, such that Low tone tokens in Phrases 5 ([High __ #]) and 7
([High __ minor]) had a markedly high F0 at onset. Yet, regardless of the magnitude of the F0 height difference the character of the Low tone pitch profile was preserved – the majority of a Low tone syllable was low-pitched and the spiked F0 occurred solely at the point of transition from the prior syllable.

Likewise for Creaky and Checked tones, the magnitude of the pitch fall varied according to context, but it nevertheless always fell from an early high peak. In final position, pitch fell greatly or to very low rates of vibration, while in medial position the fall was smaller or more gradual.

Finally, the F0 traces in Figures 4.11 – 4.19 reveal another possible effect not yet explored in this section, that the steep pitch fall of Creaky and Checked tones was attenuated when an adjoining syllable followed. This observation is tested in the following section, separate from the other F0 analyses, as it warrants comparison with the parallel changes in glottal constriction that have been attested in phrase-medial tokens.

5.7 Effects of Syllable Juncture on F0: Statistical Evidence

5.7.1 Overview

Nearly every prior description of Burmese phonology recounted the effects of close juncture on the coda of Checked tone closed syllables. Sprigg (1964) and Thurgood (1978) added that a number of effects are likewise seen with Creaky tones in juncture, many of which have already been observed in the data of Chapter Four. Effects range from morphological to segmental to suprasegmental.
Based on the F0 data explored in depth in the previous chapter, another suprasegmental effect is suggested – that the falling pitch contour of Creaky and Checked tones was weaker when another syllable closely followed. This section asks whether, statistically, F0 falls were attenuated in a medial position. As this is shown to be the case, it is further explored whether the shallower F0 fall was correlated to the suprasegmental effect noted above, the reduction of glottalization observed in phrase-medial tokens. These research questions are stated as **Hypothesis Nine**, further split into sub-hypotheses 9a and 9b.

**H9. Creaky and Checked tones in juncture have less glottalization at the syllable offset and coincidingly, less of a pitch drop over the latter half of the vowel.**

a. When phrase-medial, Creaky and Checked tone tokens have a lesser $\Delta F_0$ and shallower $m_{F_0}$.

b. There is a correlation between the F0 and OQ of phrase-medial Creaky and Checked tone tokens at the late decile timepoints.

The sub-hypothesis in (b) seeks to confirm that the forms with the strongest glottalization late in the vowel are also those with greatest F0 falls or lowest late F0 values. The articulatory basis for such a correlation is straightforward, as increased glottal constriction is commonly linked to slower, irregular vocal fold vibration (Laver 1980,
Childers and Lee 1991, see §2.5.2). If Hypothesis 9 (H9) is confirmed, then a complex suprasegmental context-sensitive alternation is demonstrated by the data – concomitant pitch and phonation qualities found in isolation or phrase-finally are both neutralized in juncture.

Conversely, a constricted glottis has also been linked to pitch raising. Section 2.7 detailed how historical accounts have tied glottalized finals to high tone reflexes (e.g. Kingston 2005) and how articulatory accounts acknowledge that a constricted glottal state with a reduced vocal fold aperture can yield a high pitch, if there is increased longitudinal tension (Titze 1994). For Burmese, Lee (2007) has explicitly argued that a phonological H tone autosegment is forced by the presence of a [CREAKY] feature specified on Burmese Creaky tones. This claim is tested with Hypothesis Ten.

**H10. Creaky voicing and a glottal stop coda induce a raised pitch in their containing syllable.**

Hypothesis Ten (H10) was tested by correlation. If late glottalization in a syllable produced an earlier high pitch peak, then a negative correlation should exist between the OQLATE and the F0MAX of the same syllable. That is, a lower OQ would correspond to a higher F0 maximum. Results of the correlation tests of H10 are provided in §5.7.4.

5.7.2 Effects of Phrase Boundaries on Creaky, Checked Falling Contours

Hypothesis H9a (Creaky and Checked tone tokens have lower ΔF0 and shallower $m_{F0}$ when phrase-medial than in other positions) was tested for Creaky and Checked tones in
a linear mixed model with CARRIER PHRASE as a fixed effect and SPEAKER as a random effect. Two related measures of the pitch contour, $\Delta F0$ and $F0_{LATE}$, for each Creaky and Checked syllables were dependent variables in individual $F$-tests, so TONE was not a factor in the model. Data from three sets of Carrier Phrases were compared: Isolated tokens, Phrase-final tokens, and Phrase-medial tokens. In order to disentangle the perseverative effects on F0 from the prior syllable’s tone (§5.6.3), the three sets were compared within two conditions according to whether the Carrier Phrase frame had an embedded [High ...] or [Low ...] position. Isolated tokens, which of course had no preceding syllable, were also compared with the other tokens in each condition. The grouping between conditions is shown in (36) and the results for [High ...] and [Low ...] conditions are provided separately in (37) for Creaky data and (38) for Checked data. All means are listed as $z$-score values normalized for each speaker and pooled across all subjects.

The results for Creaky tone data offer two points: 1) that $\Delta F0$ as calculated in this dissertation does not effectively capture the F0 dynamics of Burmese tones, and 2) the data again show that Phrases (5) and (6) do not effectively depict the simple contrast between a [High __] and [Low __] phrase-final frame, but reflect other differences that complicate the comparison.

<table>
<thead>
<tr>
<th>(36) Comparison of Phrase Position within Prior Tone Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Across High-initial Frames</strong></td>
</tr>
<tr>
<td>Isolated:</td>
</tr>
<tr>
<td>vs.</td>
</tr>
<tr>
<td>Phrase-Final:</td>
</tr>
<tr>
<td>vs.</td>
</tr>
<tr>
<td>Phrase-Medial</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Effect of Phrase Position on $\Delta F_0$ and $F_{0\text{LATE}}$ for Creaky Tones

a. $\Delta F_0$

<table>
<thead>
<tr>
<th></th>
<th>Isolated</th>
<th>Final</th>
<th>Medial</th>
<th>Difference by F-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Low... ]</td>
<td>-3.210</td>
<td>-1.602</td>
<td>-2.211</td>
<td>$F(2,227) = 28.98$</td>
</tr>
<tr>
<td>[High... ]</td>
<td>-3.210</td>
<td>-2.816</td>
<td>-2.904</td>
<td>$F(2,191) = 1.39$</td>
</tr>
</tbody>
</table>

**Between group differences by Post-hoc test**

<table>
<thead>
<tr>
<th></th>
<th>{Isolated}</th>
<th>{Medial}</th>
<th>{Final}</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[Low... ]</td>
<td>&gt;&gt;</td>
<td>&gt;&gt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- $p < .001$
- $p = .044$

<table>
<thead>
<tr>
<th></th>
<th>{Final}</th>
<th>{Isolated}</th>
<th>{Medial}</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[High... ]</td>
<td></td>
<td>&gt;&gt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- $p = 1.000$
- $p = .010$

Iso >> Medial

b. $F_{0\text{LATE}}$

<table>
<thead>
<tr>
<th></th>
<th>Isolated</th>
<th>Final</th>
<th>Medial</th>
<th>Difference by F-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Low... ]</td>
<td>-1.475</td>
<td>-1.546</td>
<td>-0.374</td>
<td>$F(2,202) = 46.62$</td>
</tr>
<tr>
<td>[High... ]</td>
<td>-1.475</td>
<td>-1.366</td>
<td>-1.043</td>
<td>$F(2,167) = 4.39$</td>
</tr>
</tbody>
</table>

**Between group differences by Post-hoc test**

<table>
<thead>
<tr>
<th></th>
<th>{Final}</th>
<th>{Isolated}</th>
<th>{Medial}</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[Low... ]</td>
<td></td>
<td>&gt;&gt;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- $p = 1.000$
- $p < .001$

<table>
<thead>
<tr>
<th></th>
<th>{Isolated}</th>
<th>{Final}</th>
<th>{Medial}</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[High... ]</td>
<td>{Final}</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- $p = 1.000$
- $p = .241$

Hypothesis H9a predicts that $\Delta F_0$ and $F_{0\text{LATE}}$ should be higher (i.e. lesser negative values) in phrase-medial tokens. While Isolated tokens showed the greatest F0 fall (negative $\Delta F_0$) amongst Creaky tones, when following a High tone the difference was not significantly greater than in Medial or Final tokens. This leveling is plausibly attributed to co-articulation from the preceding High tone, which could produce greater F0 maxima and accordingly, a greater F0 fall, regardless of sentence position. While $\Delta F_0$ results were not concise, comparison of the mean $F_{0\text{LATE}}$ values more clearly supported H9a. Post-hoc Bonferroni pairwise comparisons showed that the late F0 was significantly higher in both conditions for phrase-medial Creaky tokens than for isolated tokens, and medial tokens after a Low tone had a significantly higher late F0 than phrase-final tokens following a Low tone.
As discussed previously in Chapter Four, the phrase-final data drawn from Carrier Phrases (5) [High__#] and (6) [Low__#] should be interpreted cautiously. Phrase (6) tokens showed an overall lowering and compression of their F0 range, which reduced values for F0 and the change in F0 (ΔF0). The next table turns to the changes in F0 by phrase position for the Checked tone data.

<table>
<thead>
<tr>
<th>(38) Effect of Phrase Position on ΔF0 and F0\textsubscript{LATE} for Checked Tones</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textbf{a. ΔF0}</td>
</tr>
<tr>
<td>[Low__...]</td>
</tr>
<tr>
<td>[High__...]</td>
</tr>
</tbody>
</table>

\textbf{Between group differences by Post-hoc test}

[Low__...] \{Isolated\} \(>>\) \{Medial\}, \{Final\}  
\(p < .001\) \(p = .709\)

[High__...]
\{Final\}, \{Isolated\} \(>>\) \{Medial\}  
\(p = 1.000\) \(p < .001\)

<table>
<thead>
<tr>
<th>\textbf{b. F0\textsubscript{LATE}}</th>
<th>\textbf{Isolated}</th>
<th>\textbf{Final}</th>
<th>\textbf{Medial}</th>
<th>\textbf{Difference by F-test}</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Low__...]</td>
<td>-1.163</td>
<td>-1.489</td>
<td>0.126</td>
<td>(F(2,205) = 51.63) (*\ p &lt; .001)</td>
</tr>
<tr>
<td>[High__...]</td>
<td>-1.163</td>
<td>-1.139</td>
<td>-0.354</td>
<td>(F(2,168) = 10.20) (*\ p &lt; .001)</td>
</tr>
</tbody>
</table>

\textbf{Between group differences by Post-hoc test}

[Low__...]
\{Final\}, \{Isolated\} \(>>\) \{Medial\}  
\(p = .235\) \(p < .001\)

[High__...]
\{Isolated\}, \{Final\} \(>>\) \{Medial\}  
\(p = 1.000\) \(p < .001\)

In the Checked tone data, similar patterns were found, and even more strongly. Phrase-medial Checked tokens had statistically shallower falls and a higher F0 late in the syllable than did isolated tokens. Phrase-final data and ΔF0 again presented a muddled picture, but comparisons of F0\textsubscript{LATE} in both sentence conditions supported hypothesis H9 in that phrase-final tokens behaved similarly to those in isolation: phrase-final and isolated
forms were not statistically different from one another and both had significantly lower mean $F_{0\text{LATE}}$ than the phrase-medial Checked forms.

To summarize, both Creaky and Checked tones were generally falling tones. However, the fall was consistently more moderate when the tone-bearing syllable was in close juncture with a following syllable. Overall, $F_{0\text{LATE}}$ provided a clearer picture than $\Delta F_{0}$ of the effects of phrase position on the dynamic F0 contour associated with each tone. Lastly, the findings were more consistent in the Checked tone data.

5.7.3 Association of Creaky Voicing and Glottalization with Low F0

Recall from §4.5.2 that a more substantial difference was found in the Open Quotient readings of utterance final tokens, indicating a consistent neutralization of the voice quality distinction in medial positions. In his dissertation, Thurgood (1978) describes the loss of creaky voicing in juncture, and the loss of a glottal closure at the coda of checked syllables in juncture is well-reported by countless others (see §1.2.4). The second part of Hypothesis Nine asks if the shallower falls in pitch and loss of glottalization are connected. Hypothesis H9b made the claim that, “There is a correlation between the F0 and OQ of phrase-medial Creaky and Checked tone tokens at the late decile timepoints.”

Importantly, a test for correlation was run between OQ and F0 scores on each decile data point. In doing so, a more low-level phonetic effect was examined – whether increased glottal closure meant a greater fall in pitch. It is known that low pitch and a high degree of glottal constriction generally coincide in Burmese speech (i.e. the canonical phonetic description of Creaky and Checked tone qualities), so some statistical correlation would undoubtedly be found between F0 and OQ if a pattern was sought
across every data point for every tone. Therefore, the correlation was not run as such. The wider scope correlation is of course meaningful, and is understood to reflect a phonological pattern whereby glottalization and low pitch frequently co-occur in Burmese. The present section instead examines those forms which are already known to contain late glottalization and a concurrent pitch fall, and asks if greater glottal constriction and a lower F0 are linked or simply two independent phonetic qualities realized on the same token. To illustrate the difference, consider the pair of charts in Figures 5 and 6.

In Figure 5, the F0 drop and accompanying increase in glottal adduction typical of Creaky and Checked tones in isolation can be seen. Figure 6 illustrates the distribution of OQ and F0 for each token which composed the mean value at the 70% decile point for isolated Creaky tone data. The scatterplot in Figure 6 indicates that OQ scores were generally below average at this timepoint (70%) and that there was a small tendency for F0 to be lower when the glottal state was less open.

![Figure 5. F0 and Open Quotient traces for Isolated Creaky and Checked Tokens. Repeated from Figures 4.11 and 4.48.](image)}
Figure 6. Blow up scatterplot of data points composing means at 70% decile point. F0 and OQ traces of Figure 5 are underneath, displaying 70% decile means.

Correlation results between OQ and F0 are provided in the table in (39). The tests were run across all ten subjects and for each tone in two phrase-medial environments (39a-b) and two utterance-final environments (39c-d). A positive correlation indicates that the greater the constriction, the lower the corresponding F0 was at that timepoint, while a negative correlation would indicate that the creakier a vowel offset was, the higher the corresponding F0. A negative correlation is generally unexpected. A weak correlation or lack of a correlation would mean the realization of F0 and OQ were independent. More specifically, a lack of a correlation at a late timepoint of a Creaky or Checked syllable would indicate that although F0 dropped over the course of the syllable, it fell to varying heights that were independent of the amount of glottal constriction.

The correlation coefficients were rarely significantly different from 0 by a two-tailed test. Creaky tokens at the 70% decile in a [Low__Low] environment and the 80%
decile point of Low tones in a [High__#] environment showed correlations that were regarded as statistically unlikely to occur by chance\(^6\).

Though statistical significance was low, a few patterns are evident in the data. Primarily, the correlations were greater with more consistency for Creaky and Checked tone tokens in the phrase-medial contexts – where glottalization was neutralized or partially neutralized in juncture. Correlations were fairly weak and often negative in the Isolated or Phrase-final data (39c, d).

<table>
<thead>
<tr>
<th>(39) Correlation of (F0) and (OQ) at Late Decile Points (n) listed in parentheses</th>
<th>Tone</th>
<th>(R) at 70%</th>
<th>(R) at 80%</th>
<th>(R) at 90%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phrase (2):</strong> [Low__Low]</td>
<td><strong>Tone</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checked</td>
<td>.368 (11)</td>
<td>.012 (11)</td>
<td>.223 (6)</td>
<td></td>
</tr>
<tr>
<td>Creaky</td>
<td>.551* (19)</td>
<td>.424 (20)</td>
<td>.334 (18)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>-.353 (20)</td>
<td>-.345 (20)</td>
<td>-.155 (20)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>.316 (20)</td>
<td>.437 (19)</td>
<td>-.043 (19)</td>
<td></td>
</tr>
<tr>
<td><strong>Phrase (4):</strong> [Low__High]</td>
<td><strong>Tone</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checked</td>
<td>.166 (9)</td>
<td>-.117 (10)</td>
<td>-1.000** (2)</td>
<td></td>
</tr>
<tr>
<td>Creaky</td>
<td>.412 (15)</td>
<td>.351 (17)</td>
<td>.162 (11)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>.479* (19)</td>
<td>.312 (17)</td>
<td>.356 (17)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>-.123 (19)</td>
<td>-.115 (19)</td>
<td>-.049 (19)</td>
<td></td>
</tr>
<tr>
<td><strong>Phrase (1): Isolation</strong></td>
<td><strong>Tone</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checked</td>
<td>.274 (32)</td>
<td>-.163 (21)</td>
<td>.764 (6)</td>
<td></td>
</tr>
<tr>
<td>Creaky</td>
<td>.121 (40)</td>
<td>-.216 (20)</td>
<td>-.100 (9)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>.019 (55)</td>
<td>.116 (51)</td>
<td>.257 (48)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>-.043 (56)</td>
<td>-.145 (52)</td>
<td>-.008 (49)</td>
<td></td>
</tr>
<tr>
<td><strong>Phrase (5):</strong> [High__#]</td>
<td><strong>Tone</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checked</td>
<td>-.439 (10)</td>
<td>-.063 (6)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Creaky</td>
<td>-.217 (11)</td>
<td>.603 (7)</td>
<td>1.000** (2)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>.004 (20)</td>
<td>-.024 (20)</td>
<td>.000 (18)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>.348 (20)</td>
<td>.496* (17)</td>
<td>-.152 (18)</td>
<td></td>
</tr>
</tbody>
</table>

\(^6\) Two other cells reveal an extremely strong correlation, \(r = \pm 1.000\), but with only two datapoints composing the data for these cells, their findings cannot be given much weight.
At first glance, it might seem a surprising result that a correlation was less consistent in the contexts that saw the strongest co-occurrence of glottalization and sharply falling pitch, but the discrepancy can be explained in a fairly simple way. First, the stronger correlation in phrase-medial data represents the fact that the glottalization was not *always* neutralized, and on the occasion that a token was still glottalized, it likely coincided with a lower pitch. In other words, speakers occasionally produced, in juncture, Creaky tone syllables which were still creaky and Checked tone syllables that still contained a glottal stop. The relatively robust correlations (in comparison to 39c, d) in phrase-medial data are thus, not only explainable, but support the notion that a constricted glottis and low pitch accompany one another.

Secondly, the *n* counts in (39) also reveal that a large portion of isolated and phrase-final Creaky and Checked data was not taken into account by the correlations because either an F0 or (more likely) OQ value was not interpretable at the given decile datapoint. In (39a, b, d), *n* represents the number out of a possible 20 tokens which had both an identifiable F0 and OQ score. In (39c), for citation form tokens, *n* is out of a possible 60. For example, looking at the Creaky tone tokens in Phrase 5, a majority of the possible datapoints (66%) were not represented as only 11/20 at 70%, 7/20 at 80%, and 2/20 at 90% were usable. Compare this to the data for the same tone in Phrase (2), where the correlations were fairly strong: 19/20, 20/20, and 18/20 (only 5% of datapoints unavailable). As seen in Chapter Three, it was often not possible to calculate OQ scores (either by automated means or individual inspection of the waveform) because of irregular vibratory patterns attributable to creaky voicing or some type of glottal catch. If
the missing data represent many of the most severely glottalized datapoints, then the weak (or even negative correlation) findings are less surprising.

Looking at the F0 values for the datapoints with uninterpretable OQ scores, it is confirmed that these tokens had a lower offset F0 than the other datapoints tested in the correlations in (39). The mean F0 of the measurable datapoints (those with a measured F0 and OQ) is compared in the table in (40) with the F0 from the datapoints for which OQ was could not be defined in the glottal waveform. The “Undefinable OQ” datapoints had a consistently lower F0 in each environment and had particularly low values in the utterance final environments of Phrase (1) and (5), where a majority of the data did not have valid OQ readings. For Creaky tone syllables, the difference was statistically significant by $t$ test in all environments but for the medial position in Phrase (2) where the difference approached significance ($p = .053$, see 69a). The difference was not as consistent for Checked data, but the unavailable datapoints always had a lower mean F0, and were significantly lower in isolated tokens and near significant in the phrase-final Phrase (5) tokens.

<table>
<thead>
<tr>
<th>(40)</th>
<th>Mean F0 of Datapoints with Valid vs. Unmeasurable Open Quotient in Four Carrier Phrase Environments</th>
<th>Aggregated Data from 70%, 80%, and 90% Decile Points</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Tone</strong></td>
<td><strong>Tokens with a valid OQ score</strong></td>
</tr>
<tr>
<td><strong>Phrase (2):</strong> [Low_Low]</td>
<td>Checked</td>
<td>.356 (28)</td>
</tr>
<tr>
<td></td>
<td>Creaky</td>
<td>-.241 (54)</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>.270 (57)</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>-.640 (60)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Phrase (4):</strong> [Low_High]</th>
<th><strong>Tone</strong></th>
<th><strong>Valid OQ</strong></th>
<th><strong>Undefinable OQ</strong></th>
<th><strong>$t$ Test</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Checked</td>
<td>.102 (38)</td>
<td>-.466 (14)</td>
<td>$t(50)=.689 \ p=.019$</td>
<td></td>
</tr>
<tr>
<td>Creaky</td>
<td>-.089 (38)</td>
<td>-.879 (13)</td>
<td>$t(49)=3.82 \ p=.000$</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>.450 (53)</td>
<td>.514 (7)</td>
<td>$t(58)=-.274 \ p=.785$</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>-.652 (60)</td>
<td>none (0)</td>
<td>n/a</td>
<td></td>
</tr>
</tbody>
</table>
The results in (40) can be summarized as indicating that in tokens where the late glottalization was strong enough to interfere with accurate measurement of the glottal waveform, F0 was systematically lower. This finding offers two insights into the correlation results in (39): one, there is some association between glottal constriction and pitch late in Creaky and Checked syllables, and two, the correlation results were ineffective at finding this association because the correlation could only test a subset of the data which were unlikely to reflect an association.

5.7.4 Association of Creaky Voicing and a High F0 Peak

The high pitch regularly reported with Creaky and Checked tones has been attributed in phonological models to the presence of creaky phonation or a glottal stop in the same syllable. In Green (2005), the high pitch of Creaky and Checked tones was held to be a phonetic consequence of the phonologically-specified glottal features, [c.g.] or [ʔ], though no pitch targets were explicitly posited in his model. Lee (2007) likewise argued that an underlying [CREAKY] feature or [ʔ] coda induced the high pitch. For Lee, a set of constraint interactions forced an H tone in the output of these syllables.
Contrary to these claims, the data do not support the hypothesis in H10 (repeated below) that syllables bearing greater degrees of late glottal constriction were more likely to also bear higher F0 peaks.

**H10:** Creaky voicing and a glottal stop coda induce a raised pitch in their containing syllable.

The table in (41) presents the findings of correlations run between the OQ\(_{\text{LATE}}\) (mean OQ over the 70% - 90% deciles) and F0\(_{\text{MAX}}\) (peak F0 achieved at any point in the syllable).

<table>
<thead>
<tr>
<th></th>
<th><strong>Tone</strong></th>
<th><strong>R</strong></th>
<th><strong>Mean F0(_{\text{MAX}}):</strong></th>
<th><strong>Mean F0(_{\text{MAX}}):</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Correlated n</strong></td>
<td><strong>Undefined OQ</strong></td>
</tr>
<tr>
<td><strong>(a)</strong> Phrase (2): [Low__Low]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checked</td>
<td>.042 (15)</td>
<td>1.172</td>
<td>1.803 (5)</td>
<td></td>
</tr>
<tr>
<td>Creaky</td>
<td>.062 (19)</td>
<td>1.155</td>
<td>.975 (1)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>-.498* (19)</td>
<td>.514</td>
<td>.026 (1)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>.065 (19)</td>
<td>.223</td>
<td>.362 (1)</td>
<td></td>
</tr>
<tr>
<td><strong>(b)</strong> Phrase (4): [Low__High]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checked</td>
<td>-.346 (16)</td>
<td>.480</td>
<td>.655 (4)</td>
<td></td>
</tr>
<tr>
<td>Creaky</td>
<td>-.124 (16)</td>
<td>1.727</td>
<td>1.338 (4)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>.001 (19)</td>
<td>1.011</td>
<td>1.166 (1)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>-.087 (19)</td>
<td>.764</td>
<td>-.059 (1)</td>
<td></td>
</tr>
<tr>
<td><strong>(c)</strong> Phrase (1): Isolation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checked</td>
<td>-.113 (39)</td>
<td>1.345</td>
<td>1.225 (19)</td>
<td></td>
</tr>
<tr>
<td>Creaky</td>
<td>.043 (44)</td>
<td>1.273</td>
<td>1.768 (15)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>.220 (58)</td>
<td>.846</td>
<td>.835 (2)</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>.053 (57)</td>
<td>1.077</td>
<td>.425 (2)</td>
<td></td>
</tr>
<tr>
<td><strong>(d)</strong> Phrase (5): [High__#]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checked</td>
<td>.276 (11)</td>
<td>2.593</td>
<td>1.614 (9)</td>
<td></td>
</tr>
<tr>
<td>Creaky</td>
<td>-.193 (11)</td>
<td>1.240</td>
<td>1.746 (9)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>-.250 (19)</td>
<td>1.102</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>.123 (19)</td>
<td>1.187</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
A notably different effect was tested here than with the correlations used to test H9b in the preceding section. Where the correlations in §5.7.3 looked at the simultaneous interaction of the glottal state and pitch, the correlations in (41) generally concern F0 maxima from early in the syllable – asynchronous to the OQ readings at the vowel offset.

Very little correlation was found between the OQ at the syllable offset and the same syllable’s peak F0, particularly for the Creaky and Checked tones tested by hypothesis H10. In one case – Checked tones in medial Phrase (4) – does the data yield a moderately high negative correlation, $r = -.346$. The scatterplot in Figure 7 displays the distribution composing this correlation.

Figure 7. Scatterplot of $F_{0\text{MAX}}$ by $OQ_{LATE}$ in Checked tones embedded in Carrier Phrase (4). Three data points with considerably higher $F_{0\text{MAX}}$ values are encircled.

Three high-pitched data points, circled in Figure 7, appear to be the basis for the negative correlation. Furthermore, the correlation coefficients in other contexts do not suggest a
tendency for Checked syllables to bear a higher pitch peak when the vowel is creakier, nor do the F0 maxima of the tokens which were not evaluated in the correlation. The two rightmost columns in (41) compare the mean $F_{0_{\text{MAX}}}$ and $n$ between the tokens with a identifiable OQ and those without. Recall that in §5.7.3, in cases where heavy glottalization was expected, the number of unidentifiable OQ scores rose considerably, and their F0 was predictably lower. The table in (41) likewise shows that the number of unavailable datapoints was much higher in the utterance-final Creaky and Checked tones (41c,d), but the F0 maxima were not reliably higher or lower at these datapoints. Assuming that the majority of unidentifiable OQ scores were the product of heavy glottalization hindering interpretation of the glottal waveform, these data offer no reason to deduce that creaky and closed glottal articulations are responsible for the canonical high pitch of these two tones. High pitch peaks occur in Creaky and Checked tones regardless of glottal constriction; the height in F0 of this peak is not affected by the severity of the constriction.

Values for the High and Low tone data are also included in (41) for the purposes of comparison, though they test no proposed hypothesis. The High tone in the medial context of Carrier Phrase (2) revealed a high negative correlation that was significantly different from zero. However, since High tones had moderate or fairly high OQ ratios, this result says little about hypothesis H10 and the acoustic effects of glottal constriction. In this case, the correlation indicates a pattern whereby tokens with breathier offsets (i.e., a higher $OQ_{\text{LATE}}$) had lower F0 maxima. Beyond this noteworthy finding, High and Low tone data showed little to no correlation between the degree of glottal constriction and the height of F0 peaks.
To conclude, the height of a pitch peak on High, Creaky, or Checked tone syllables was independent of whether it ended with a nearly closed glottis, some creak, or a wide-open glottis. Open Quotient over the final third of a vowel had no discernable interaction with the F0 height of the rest of the syllable.

5.7.5 Discussion of Findings

The statistical analyses of Section 5.7 provided three main observations concerning the falling F0 contours of Burmese Creaky and Checked tones. Both Creaky and Checked tones had a statistically higher offset F0 in phrase-medial position than finally or in isolation (§5.7.2). That is, falling tones did not fall as greatly in medial position, though they still fell. Secondly, increased glottal constriction and a lower pitch were often linked in Burmese speech, but more specifically, there was a small statistical tendency for more heavily glottalized syllable offsets to have a lower F0 at that offset (§5.7.3). Finally, there was no association between the high pitch peak found with Creaky and Checked tones and the incidence of glottalization (§5.7.4). Creaky and Checked tones (and High tones for that matter) were regularly found with a high F0 peak that was not dependent on the presence of any phonatory properties or coda consonants within the syllable.

For phonological analysis, these data address critical issues of how to treat the contrasts and alternations in pitch – whether the different contours are underlyingly distinct, whether they are the result or cause of phonation distinctions, and why close juncture invokes changes in pitch (and glottalization). Taken together, the findings in §5.7 suggest that glottalization in the production of these tones certainly augments a fall in F0, but is not the sole source of this fall. This stance is detailed below.
An explanation that may be ruled out is that juncture with a following syllable yields a smaller fall in pitch because there is less time to realize a targeted low offset. The duration data simply do not bear out such an account of “clipping”. Both in isolation and phrase-finally, Creaky and Checked tones did not have a longer duration than when the same tokens were phrase-medial. The vowel in these forms was short, being abruptly cut-off by creakiness at the offset or a distinct glottal stop coda.

The explanation preferred here is that underlying falls in pitch are steep when co-occurring with vowel terminal glottalization and less steep in juncture when the glottal constriction is also less severe. The data show this in two ways. First, the two alternations overlap. Constricted glottis qualities were neutralized and falling pitch was more gradual in the same position – phrase-medially with an immediately following syllable. Second, in Creaky and Checked tokens there was a low-level phonetic tendency during identifiable creaky voicing for F0 to be lower as glottal constriction increased (as quantified by OQ).

However, it was repeatedly shown in the production data for this study (§4.3) that there was a pitch fall on Creaky and Checked tone syllables regardless of the sentential context – crucially, even when glottalization was neutralized. This leads to the conclusion that a low fundamental frequency at the vowel terminus does not give rise to creaky voicing or a more constricted glottal state. Likewise, a constricted glottis (as captured by a low OQ) is not the sole cause of falling pitch on Creaky and Checked tones, though it does seem to promote it. The phonological implications of this stance are explored further in the next chapter.
Chapter Six. The Perception of Tone.

6.1 Background

Hypotheses One – Ten concerned the properties of tone production. The present chapter and Hypothesis Eleven herein examine the perception of tone by Burmese native speakers. While the perception experiment described is not as comprehensive as the production studies of Chapters Four and Five, perception of the tonal contrasts is regarded as being as significant to modeling the phonological composition of the tone contrast. To date, no perception study of Burmese has been conducted, but tone perception experiments in Gandour (1981), Brunelle (2005, 2009), Svantesson & House (2006), and Zsiga & Nitisaroj (2007) have all assessed how speakers of SE Asian languages (Cantonese, Cham, Vietnamese, Kammu, and Thai) distinguish or fail to distinguish between different tonal cues. In the case of Zsiga & Nitisaroj (2007) for example, the cues were the pitch peak and the timing of the pitch peak. Prior phonological accounts of Burmese have contended that the tones represent a contrast in register (Bradley 1982), pitch (Yip 1995), or a hybrid of the two (Lee 2007) (§5.2 for review). Data from this dissertation has clearly demonstrated that many of the register and pitch distinctions between tones are both systematically produced. However, it does not necessarily follow that listeners use both distinctions as a cue to tone identification. It is not known how these same distinctions are perceived or if one contrast takes precedence over the others.

In pursuit of this question, a perception experiment was conducted using a forced choice classification task wherein listeners provided a toneme-based response to two-hundred forty citation form syllables, many of which were not naturally spoken Burmese, but were re-synthesized to controlled values for Duration, F0 Height, F0 Contour,
H11: Each phonetic parameter in the study (Duration, F0 Height, F0 Contour, Phonation Type) serves as an acoustic cue to Tone identification, but a single parameter is appealed to in instances of ambiguous stimuli.

While the hypothesis above is agnostic to the strength of specific parameters, a few general expectations can be outlined here given the production findings of Chapter Four. To begin with, no acoustic dimension should be divisible into four distinct ranges corresponding to the tones, since no such distinction was found quantitatively in the production data. It would be unexpected for listeners to systematically categorize a distinction that they did not produce even roughly. For example, if listener responses were hypothetically predicated entirely on the F0 contour, it would be unlikely that Creaky and Checked tones could be perceptually distinct since these two tones had virtually identical F0 contours in production. Instead, a more likely finding is that acoustic parameters which are attended to offer binary contrasts which isolate a tone (e.g. F0 Height separating Low vs. other three tones) or split the four categories (e.g. Phonation Type separating Low/High vs. Creaky/Checked). Any such distinctions are of considerable interest to the phonological model of Chapter Seven, which, if adequate, must account for any duration, F0, or phonation contrast which has been reliably found in
production and perception. The remaining sections of this chapter describe the design (§6.2), results (§6.3), and analysis (§6.4) of the experiment.

6.2 Methodology

6.2.1 Stimuli

The stimuli set presented to listeners was composed of 240 tokens re-synthesized to different extents. All tokens were recorded from a single speaker, a female in her 40’s with 15 years of U.S. residency and an occupational background in singing (Speaker C). Recordings were made with an omnidirectional microphone (Behringer ECM8000) and digitized at 44 kHz using an EG2-PCX data acquisition system (Glottal Enterprises, Inc.) and a Windows XP laptop running the WaveView software (Glottal). Re-synthesis of intensity, duration, and F0 was performed in Praat 5.1.03 (Boersma and Weenink 2009) on a Macbook laptop.

For resynthesis, 216 tokens were created from six recorded samples, each of two carrier syllables (/la/, /le/) produced with breathy, creaky, and modal voicing. In this way, PHONATION TYPE was not synthesized, but controlled for by these prototype tokens. The speaker was instructed to produce the syllables [la] and [le] in isolation with the Low tone and with emphatically breathy and creaky voicing (the phonation and not the tones, High and Creaky). After supervised practice and training, the speaker recorded multiple iterations of each syllable, from which the six prototype tokens were hand-selected by the

---

1 Note that a base Checked token was not used to construct the synthesized stimuli set. Checked tone samples were included in the set of natural control tokens and a Checked tone option was one of the four possible responses in the task.
researcher for voice qualities confirmed through spectrographic analysis and an accompanying EGG trace. A sample spectrograms of one of these samples is provided in Figure 1. This approach to controlling for phonation type was pursued for two reasons.

First, the perception study was conducted simultaneously with the production tasks in Chapter Three. Details of the phonetic realization of breathy and creaky voicing in Burmese were less clear at this time. In fact, prior to the results of this dissertation’s experiments, the phonetic literature had given the impression that the Burmese Creaky tone used a mode of phonation that was qualitatively different than the slowed, irregular vibration of constricted vocal folds associated with creaky voicing (see discussion in §2). Synthesis of the PHONATION TYPE parameter would have injected assumptions about the acoustic shape of Burmese creakiness or breathiness into the stimulus. While categorization of a synthesized phonation continuum poses an interesting question about what acoustic properties are heard as breathy or creaky by Burmese listeners, such a task would reveal less about the perceptual interaction between PHONATION TYPE and the other phonetic factors. Furthermore, the present state of knowledge demands that it first be established that creaky and breathy voice qualities are even systematically associated with a tone category. The prototype tokens of this study used naturally produced phonation that was more likely to be included by listeners within their native category boundaries².

Secondly, with or (particularly) without the production data, digitally altering the correlates of phonation type is a more complicated process than altering F0, duration, or intensity. Brunelle (2005) reports this same difficulty, explaining that for his study of

---

² This method did, however, assume that the breathy and creaky qualities were produced in a canonical fashion by the speaker who recorded the stimuli.
Eastern Cham, he chose to re-synthesize for pitch and duration, but not voice quality after his “attempts at modifying voice quality...resulted in stimuli that did not sound natural to my listeners.”

Intensity profiles for these tokens were standardized across each vowel: [la] syllables to 77dB, [lɛ] syllables to 75dB, levels which roughly matched the intensity of the prototype modal tokens. Once standardized for intensity, each of the six prototypes was re-synthesized for three duration levels and twelve pitch contours, yielding thirty-six distinct stimuli from each base token and 216 tokens overall (2 vowels x 3 phonation types x 3 durations x 12 F0 contours). Details of the parameters for re-synthesis and the listeners’ task are provided below.

![Figure 1. [lɛ̰̀] Base token for 36 synthesized stimuli of Creaky tone /lɛ/.](image)

Vowel duration was rescaled to three lengths: 175ms, 250ms, 325ms. These values were selected according to the typical durations found in the Speaker C data. Her High and Low tone citation utterances were generally 325ms or more and Creaky tone

---

3 Each of the intensity-standardized phonation prototype syllables was also included in the stimuli. Results for those tokens are described in §6.3.2.
syllables had durations around 175 ms. A duration of 250 ms then represents a mid-point where stimuli were not decisively associated with any tone on the basis of duration.

Fundamental frequency was re-synthesized to fit twelve contours: five even pitch tracks at 20 Hz intervals from 160 Hz to 240 Hz, and seven rising or falling pitch tracks with “gradual” or “late” changing slopes (see 1). The selected pitch range reflected that employed in the speaker’s elicited production data, where the even-pitched portions of Low tone samples had an average F0 of roughly 180Hz and Creaky and Checked F0 peaks were approximately 240-250 Hz. F0 minima in the stimuli were set at 20 Hz below the speaker’s “average” F0 and maxima expressed these typical peak values. Contoured tracks either fell to 175 Hz from early peaks (220 Hz, 240 Hz) or rose from 160 Hz to late peaks (220 Hz, 240 Hz). A single continuously high-pitched “rising” contour was also added, with a 220 Hz onset and 240 Hz offset. The table in (1) illustrates the arrangement of the varying slopes within the designated F0 range.

<table>
<thead>
<tr>
<th>(1)</th>
<th>Falling contour pitch tracks</th>
<th>Rising contour pitch tracks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gradual</strong></td>
<td>Onset MID-pt Offset</td>
<td>Onset MID-pt Offset</td>
</tr>
<tr>
<td>a) 240</td>
<td>207.5 175</td>
<td>e) 220 230 240</td>
</tr>
<tr>
<td>b) 220</td>
<td>197.5 175</td>
<td>f) 160 190 220</td>
</tr>
<tr>
<td><strong>Late</strong></td>
<td>Onset 2/3rd-pnt Offset</td>
<td>Onset 2/3rd-pnt Offset</td>
</tr>
<tr>
<td>c) 240</td>
<td>240 175</td>
<td>g) 160 160 220</td>
</tr>
<tr>
<td>d) 220</td>
<td>220 175</td>
<td></td>
</tr>
</tbody>
</table>

“Gradual” slopes were a straight-line interpolation from onset to offset F0 while the “Late” abrupt slopes maintained the onset F0 until the 2/3rd point of the vowel before rising or falling over the final 1/3rd.

Lastly, the twenty-four control tokens were included in order to establish a baseline for correct identification of natural tone-bearing syllables. These tokens were composed of eighteen natural citation-form productions (four of Creaky, High, and Low;
six Checked syllables) and the remaining six tokens were the prototype phonation-bearing tokens with standardized intensity.

All stimuli, re-synthesized and control tokens, were presented as monosyllables in isolation.

6.2.2 Presentation

The experiment was conducted with the DMDX software (Forster & Forster 2002) run on a Windows XP laptop. The participants were presented a visual stimulus offering a forced-choice selection between four Burmese orthographic symbols representing the /la/ or /le/ syllable with each tone (see Figure 2). Audio stimuli were played over audio-technica ATH-M40 headphones and subjects were permitted to replay stimuli as many times as they wished\(^4\). Before the experiment, a brief familiarization session with sixteen

---

\(^4\) For all but two listeners, replayed stimuli were rare. Subjects A and E replayed approximately \(\frac{1}{4}\) of all stimuli. Other subjects all together replayed only 4.3% of the stimuli tokens they were presented. More than one replay of a single sample was very rare – occurring for 1% of the stimuli.
natural tokens was conducted to instruct participants and set speaker volumes appropriately. The 240 stimuli were presented in six blocks of 40 tokens, each ordered randomly. Short breaks were permitted between each block, though none were taken.

Participants in the perception experiment were nearly the same as those listed in Chapter Three for the production experiment. Subjects A and L only performed the production tasks. Subject C did not participate because she contributed the stimuli utterances. For all nine remaining subjects, the perception experiment was conducted after the aerodynamic and acoustic production tasks detailed in this dissertation. In addition to these nine subjects, another subject (M) performed only the perception experiment. Subject M was a female university student in her 20s, who declared proficiency in French. In all, ten listeners performed the task (5 male, 5 female).

6.3 Results

6.3.1 Overview

The identification of the synthesized breathy-voiced and modal tokens relied predominantly on F0 HEIGHT and DURATION, while creaky-voiced stimuli were predominately identified as the Creaky tone at all levels of F0 or DURATION. Following sections present the results for the unsynthesized tokens (§6.3.2), even contoured stimuli §6.3.3), falling F0 stimuli in (§6.3.4), and rising F0 stimuli (§6.3.5). Before continuing, it is necessary to briefly discuss the forced choice options and the listeners’ interaction with them.
In particular, one choice, the creaky syllable လာ [là], was not a known Burmese lexical item. The eight possible choices are listed in (2), along with glosses for the other seven words.

(2) Orthographic Stimuli: Not all lexical items

<table>
<thead>
<tr>
<th>Burmese</th>
<th>/la/</th>
<th>Gloss</th>
<th>Burmese</th>
<th>/lε/</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>လာ</td>
<td>“come”</td>
<td>လယ္</td>
<td>“field”</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>လား</td>
<td>question particle</td>
<td>လဲ</td>
<td>“falling”</td>
<td></td>
</tr>
<tr>
<td>Creaky</td>
<td>လာ့</td>
<td>n/a</td>
<td>လဲ့</td>
<td>“sparkle, crystal”</td>
<td></td>
</tr>
<tr>
<td>Checked</td>
<td>လာပ္</td>
<td>particle</td>
<td>လက္</td>
<td>“hand”</td>
<td></td>
</tr>
</tbody>
</table>

As discussed previously in Chapter Three, familiarity with the Burmese writing system permits the researcher to assume that all subjects understand the phonetic form of the missing lexical item to be pronounced with a creaky [à] rime. Still, without an identifiable match in the lexicon, it was possible that listeners would avoid selecting the လာ့ item or, conversely, favor it for the numerous re-synthesized tokens which disobey the phonological grammar of the language.

A chi-square test for homogeneity between the distributions of responses for /la/ and /lε/ stimuli found a strong difference for every listener ($p < .001$, rejecting the null hypothesis that counts were the same). This difference could be attributed to the lexical gap of လာ့ or to the vowel quality difference of /a/ vs. /ε/, and is argued here to be the result of the latter. Inspection of the counts (in (3) below) indicated that the difference lay
between the Creaky and Checked tone identifications; /la/ tokens saw a greater proportion of Creaky identifications and /lɛ/ tokens were far more likely to be assigned a Checked identification. If there was a response bias with the lexical gap, it was in favor of the non-word over Checked [làʔ]. However, the trend is perhaps better stated as an extreme aversion of Checked /la/ identifications – three listeners heard and selected ཁོ་ [làʔ] zero times, and nine of the ten used it in response to less than 10% of all /la/ stimuli.

<table>
<thead>
<tr>
<th>(3)</th>
<th>Count of Responses to Re-synthesized Stimuli, by Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPA</td>
<td>[la]</td>
</tr>
<tr>
<td>Burmese</td>
<td>ཁོ་</td>
</tr>
<tr>
<td>Subj. A</td>
<td>26</td>
</tr>
<tr>
<td>D</td>
<td>30</td>
</tr>
<tr>
<td>E</td>
<td>16</td>
</tr>
<tr>
<td>F</td>
<td>23</td>
</tr>
<tr>
<td>H</td>
<td>38</td>
</tr>
<tr>
<td>I</td>
<td>15</td>
</tr>
<tr>
<td>J</td>
<td>41</td>
</tr>
<tr>
<td>K</td>
<td>47</td>
</tr>
<tr>
<td>L</td>
<td>22</td>
</tr>
<tr>
<td>M</td>
<td>40</td>
</tr>
<tr>
<td>Total</td>
<td>298</td>
</tr>
</tbody>
</table>

Table. Counts of each subject’s responses by toneme. Results of a chi-square test comparing the distribution of counts for /la/ and /lɛ/ sets are given in the far right column. The column of Checked tone responses is lightly shaded for ease of comparison.

A likely cause of this bias was the vowel distinction between open syllable quality /a/ and the Checked syllable quality /aʔ/, sometimes reported as /ʌʔ/ (Mehnert and Richter 1976). Since the base tokens for all re-synthesized stimuli were produced with open syllable /la/, the grounds for confusion with [làʔ] are small regardless of the duration and
pitch contour. Further, listeners had no aversion to selecting [làʔ] when played the natural unmodified Checked syllables, which undoubtedly contained the proper Checked vowel quality – 28 out of 30 possible responses were Checked. For these reasons, the chi-square results are interpreted not as a bias towards selecting the non-lexical item, but rather a bias against hearing the Checked token when presented an open syllable vowel /a/. Notably, the closed syllable /le/ was evidently more confusable with the open syllable counterpart.

6.3.2 Natural and Control Stimuli

Correct identification of tone for the unaltered recordings was high, with confusion between Creaky and Checked syllables yielding a majority of the misidentifications. The six intensity-controlled phonation prototypes were less successfully matched to the expected tones.

(4) Identification of Natural Stimuli ($n = 180$, 18 tokens x 10 subjects)

<table>
<thead>
<tr>
<th>Stimuli</th>
<th>Low</th>
<th>High</th>
<th>Creaky</th>
<th>Check</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>38</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>95%</td>
</tr>
<tr>
<td>High</td>
<td>4</td>
<td>35</td>
<td>1</td>
<td>0</td>
<td>87.5%</td>
</tr>
<tr>
<td>Creaky</td>
<td>0</td>
<td>2</td>
<td>32</td>
<td>6</td>
<td>80%</td>
</tr>
<tr>
<td>Check</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>55</td>
<td>91.7%</td>
</tr>
</tbody>
</table>

Overall, successful identification of natural stimuli was made 89% of the time (160/180). The two most common errors, involving Creaky and Checked tones, were primarily made by just two listeners. Speaker H was responsible for three of the six errors identifying
Creaky stimuli as Checked syllables. Speaker F made three of the five errors for the reverse misidentification. No other error was disproportionately supplied by any speaker: the two Low tone misidentifications were by two listeners, and the High tone identifications as Low were made by four different listeners. Only two subjects, I and M, had 100% success rates for all 18 natural tokens.

Identification of the phonation prototypes which served as a base for all re-synthesized stimuli met mixed results. Recall that these stimuli were not necessarily Low, High, or Creaky-toned, but were modal, breathy, or creaky-voiced with the restrained production\(^5\) of the other qualities characterizing the tones (F0, duration, and intensity). The two modal tokens were identified as Low-toned by all listeners, but identification of the breathy and creaky-voiced tokens was more variable. Breathy-voiced stimuli were as likely to be heard as a Low tone as a High. One creaky-voice prototype, [là], was always identified as a Creaky tone, but the other, [lέ], was mistaken for both High and Checked tones. The table in (5) provides the full results by token.

(5) Identification of Phonation Prototype Tokens \((n = 60)\)

<table>
<thead>
<tr>
<th>stimuli</th>
<th>Low</th>
<th>High</th>
<th>Creaky</th>
<th>Check</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modal</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>Breathy</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>50%</td>
</tr>
<tr>
<td>Creaky</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>2(^6)</td>
<td>50%</td>
</tr>
</tbody>
</table>

\(^5\) Speaker C recorded the base stimuli tokens and sought to produce the modal, breathy, and creaky stimuli with equivalent F0, duration, and intensity. Her attempts were mostly successful, but the breathy prototype still bore a somewhat higher F0, the creaky prototype token was somewhat shorter and fell slightly. Therefore, these qualities are said to have been produced in a “restrained” manner.

\(^6\) The 50% rate of misidentification for the Creaky [lέ] prototype provides further confirmation that the difference in counts between the /a/ and /ε/ stimuli is the result of skewing towards Checked [ε] identifications rather than any unusual pattern with the data for /a/.
6.3.3. Re-synthesized Stimuli: Even Contours

The even F0 stimuli display the overall pattern of the data solidly, with F0 Height, Duration, and Creaky Phonation all playing a visible role in shaping listener responses. Results are presented graphically in Figure 3 following the present discussion. Figure 3 (as well as ensuing charts of perception results) give the response counts out of 20 (pooling all subjects) for each combination of Duration and F0 Height/Contour. Breathy and modal stimuli produced similar results as one another, implying that the phonation distinction between the two did not influence responses. F0 had an easily discernable effect with a threshold of 200 Hz, such that stimuli above 200 Hz were typically identified as High tones while below, Low tone identifications predominated. Precisely at 200 Hz, the even contour stimuli were fairly ambiguous, being almost evenly split between High and Low tone responses. At the shortest Duration (175 ms), breathy and modal stimuli at or over 200 Hz frequently prompted Creaky or Checked tone responses. This effect of Duration in only higher Hz tokens (200 +) was as likely with breathy as with modal stimuli.

By contrast, creaky-voiced stimuli were mainly heard as the Creaky tone, though responses for other tones were always possible. Checked tone identifications occurred at all three Duration levels, but were most frequent with 175 ms tokens (and of course with primarily the vowel [ɛ]). High tone identifications made up between 10-20% of the responses to 325 ms long creaky-voice stimuli. Regardless of Duration and F0 Height, 66.3% of all even pitch stimuli with creaky-voicing were identified as the Creaky tone.
Figure 3. Listener responses to Stimuli with Even F0 Contours at 20 Hz intervals. Breathy and Modal stimuli elicit similar response matrices. Responses to creaky-voiced stimuli were overwhelmingly Creaky tone identifications (in red).
Responses to Falling-contoured Stimuli
Each pie (n = 20 judgments)

Figure 4. Listener responses to Stimuli with Falling F0 Contours. Four falling contours (listed in the leftmost column) constitute the stimuli – two gradual falls and two “abrupt” falls in which the entire F0 fall occurs after the mid-point of the vowel. Breathy and modal stimuli elicit similar response matrices. Responses to creaky-voiced stimuli were overwhelmingly Creaky tone identifications (in red).
Figure 5. Listener responses to Stimuli with Rising F0 Contours. Three contours (listed in the leftmost column) constitute the stimuli. Breathy and Modal stimuli elicit similar response matrices. Creaky stimuli with a gradual rise were mainly identified as Creaky or Checked tones, while those bearing an abrupt F0 rise drew a variety of responses.
6.3.4 Re-synthesized Stimuli: Falling Contours

Identification of falling contours revealed a different pattern. Low tone identifications were rare. As with even contoured data, breathy and modal stimuli yielded results similar to one another – in this case, responses were divided between a majority of Creaky tone identifications for 175 ms tokens and a majority of High tone identifications for longer (250, 325 ms) DURATION stimuli. The high F0 HEIGHT early in the syllable deterred Low tone identifications, but was compatible for listeners with either the High or Creaky tone (in which case, DURATION acted as a criterion). Interestingly though, modal or breathy PHONATION TYPES did not likewise deter Creaky tone identifications.

Falling creaky-voiced stimuli were overwhelmingly identified as Creaky (69.6%, 167/240) or Checked tones (19.2%, 46/240), with increasing occasional High tone identifications as DURATION lengthened. Figure 4 (separate page) depicts the pooled listener responses to all falling contour stimuli.

6.3.5 Re-synthesized Stimuli: Rising Contours

Rising contours are found naturally with just High and Low tones, and furthermore, with just Low tones in isolation (§4.3). Not surprisingly, rising contour stimuli elicited the most confused and ambiguous responses. More precisely, “gradual” rises were met by ambiguity, while “abrupt” late rising stimuli were interpreted as Low tone syllables across all PHONATION TYPES. It was not clear whether the preponderance of Low tone identifications for the 160-160-240 Abrupt Rising tokens can be attributed to the subjects using the late F0 rise as a cue to the Low tone or possibly not using it and simply hearing a syllable which is low-pitched for most of its duration.
Once again, identifications of breathy-voiced and modal stimuli were highly similar to each other (Figure 5a and 5b), indicating no distinct response by listeners to the breathy phonation. Gradual Rising breathy and modal stimuli showed little to no pattern, with High, Low, or Creaky tone identifications as likely as one another but for two trends. The High-to-Highest 220-240 Rising tokens had a significant majority of High tone responses and the 175 ms Duration tokens tended to elicit a greater number of Creaky and Checked responses.

While a rising F0 and Creaky or Checked tone were incompatible in production, this did not prevent a majority of Creaky tone identifications for creaky-voiced stimuli with a rising F0 Contour (see Figure 5c).

6.4 Analysis and Discussion

The results in §6.3 reveal an effect on listener responses for all three parameters of the synthesized data: Duration, F0, and Phonation Type. A Duration of 175 ms was more highly correlated with Creaky and Checked tone identifications across all F0 patterns and phonations, and Duration of 250 and 325 ms were more likely to elicit High and Low tone responses, even with falling creaky-voiced stimuli.

For F0, the presence of some F0 peak was more important than the direction of the F0 contour. Even-contoured stimuli were highly likely to be identified as Creaky or Checked tones if they bore a high pitch and short duration, despite the absence of any fall in pitch. Stimuli which maintained a low F0 throughout the syllable were likely identified
as a Low tone at any Duration, unless overruled in some cases by the influence of (creaky) PHONATION TYPE.

While the results firmly indicate phonation as an acoustic cue to Tone, this effect was notably limited to the distinction between creaky and non-creaky voicing. Listeners did not differentiate breathy stimuli from those with modal voicing.

To analyze these data, a stepwise multinomial regression was run, which determines which factor from a selected array explains the greatest amount of variation in the response data, as well as the next best predictors. The test proceeds “stepwise” from an initial model of simply choosing the mean (in this case, the most frequent response ‘Creaky’, which would be correct 36.6% of the time) and determining which factor improves the most upon this model in steps to also determine the 2nd, 3rd, … nth best predictor. This process is repeated for each factor until the model is not significantly improved by adding another factor.

For the perception experiment, factors tested by stepwise multinomial regression were DURATION, ONSET F0, F0 MID-POINT, F0 OFFSET, Δ F0, PHONATION TYPE, SUBJECT, and VOWEL QUALITY (/a/ vs /ɛ/). The table in (6) shows that the first step found PHONATION TYPE to be the factor that predicted the greatest amount of variation, by itself correctly predicting 46.9% of all responses, a 10.3% improvement from the baseline. Further iterations of the regression showed the next best predictors to be, in order, F0 MID-POINT, SUBJECT, DURATION, VOWEL QUALITY, and then Δ F0. Adding F0 ONSET and F0 OFFSET did not significantly improve the model’s predictive power. Following PHONATION TYPE, only the F0 MID-POINT substantially increases the percentage of correctly predicted responses, up to 57.7% of the data. After this second step, successive
iterations, although improvements according to a significant chi-square effect selection test, saw just an incremental increase in the percentage of predicted responses. Each significant effect, the chi-square effect selection test, the Nagelkerke $R^2$ (a pseudo-$R^2$, estimating improvement from the null model (Step 0)), and the overall model accuracy in predicted all responses are listed in (6).

<table>
<thead>
<tr>
<th>Step</th>
<th>Cue</th>
<th>$\chi^2$ Sig. of Step</th>
<th>Nagelkerke $R^2$</th>
<th>Predicted by Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>All “Creaky”</td>
<td>0.000</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>PHONATION TYPE</td>
<td>0.000</td>
<td>.277</td>
<td>46.9%</td>
</tr>
<tr>
<td>3</td>
<td>+ F0 MID-POINT</td>
<td>0.000</td>
<td>.449</td>
<td>57.7%</td>
</tr>
<tr>
<td>4</td>
<td>+ SUBJECT</td>
<td>0.000</td>
<td>.543</td>
<td>61.2%</td>
</tr>
<tr>
<td>5</td>
<td>+ DURATION</td>
<td>0.000</td>
<td>.608</td>
<td>62.8%</td>
</tr>
<tr>
<td>6</td>
<td>+ VOWEL</td>
<td>0.000</td>
<td>.630</td>
<td>63.6%</td>
</tr>
<tr>
<td>7</td>
<td>+ $\Delta$ F0</td>
<td>0.000</td>
<td>.645</td>
<td>63.8%</td>
</tr>
<tr>
<td>7</td>
<td>+ F0 OFFSET</td>
<td>$p &gt; .05$</td>
<td>.650</td>
<td>-</td>
</tr>
</tbody>
</table>

Looking in-depth at a few of the steps reveals how each factor affected listener responses. With only PHONATION TYPE included, the model made only a binary decision, captured in (7), predicting either a Creaky tone response for creaky stimuli or a High tone response for all other stimuli. The regression has only two outcomes at this stage because, without considering any other factor, a High tone response was calculated to be more likely for both breathy and modal-voiced stimuli. This finding is a statistical reflection of the similarity in results for breathy and modal stimuli noted repeatedly in §6.3.
By including the second step, the F0 MID-POINT, in the regression model, High and Low tone classifications were distinguished by F0, enabling adequate prediction of the listeners’ Low tone responses. A comparison of the classification table in (8) with that of (7) reveals the predicted Creaky responses to be exactly the same 740 tokens. Adding an F0 criterion only led to finer distinctions drawn between the non-creaky stimuli.

(8) Classification by Multinomial Regression after Step 2: F0 MID-POINT

<table>
<thead>
<tr>
<th>Observed Response</th>
<th>Low</th>
<th>High</th>
<th>Creaky</th>
<th>Checked</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>423</td>
<td>88</td>
<td>49</td>
<td>0</td>
<td>75.5%</td>
</tr>
<tr>
<td>High</td>
<td>185</td>
<td>362</td>
<td>71</td>
<td>0</td>
<td>58.6%</td>
</tr>
<tr>
<td>Creaky</td>
<td>105</td>
<td>212</td>
<td>495</td>
<td>0</td>
<td>61.0%</td>
</tr>
<tr>
<td>Checked</td>
<td>47</td>
<td>58</td>
<td>125</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>Overall</td>
<td>34.2%</td>
<td>32.4%</td>
<td>33.3%</td>
<td>0.0%</td>
<td>46.9%</td>
</tr>
</tbody>
</table>

It is perhaps curious that the mid-point proved to be a more suitable predictor of F0 response than the onset, offset, or Δ. This result encourages the understanding that the location of the F0 peak is not as important to identification as the simple presence of a peak somewhere in the syllable. By the design of the stimuli, the ONSET or OFFSET F0
values only reflected a peak roughly half of the time and $\Delta F_0$ only distinguished high F0 values after a rise or before a fall. In contrast, the vowel mid-point was never the location of an F0 peak, but syllables with a high F0 peak, whether rising, falling, or evenly high, would subsequently have a higher F0 MID-POINT than those which stayed below the 200 Hz threshold. In other words, of the test F0 factors, F0 MID-POINT best captured overall F0 HEIGHT.

The next two factors, SUBJECT and DURATION, refine the response prediction amongst the High, Low, and Creaky tone to accommodate inter-subject preferences and the association of longer DURATION with High and Low tones and briefer DURATION with Creaky and Checked tone identification. Interestingly, the regression model makes very few Checked tone predictions at this point (0.3% of all responses). Only after adding VOWEL QUALITY are Checked predictions made, since (as seen in §6.3.1) Checked responses differed so greatly between /a/ and /ɛ/. Still only 24.8% of the listeners’ Checked tone identifications are successfully predicted at this stage of the model. The classification matrix after Step 5 (VOWEL QUALITY) is given in (9), followed by the classifications for the next, and final significant, stepwise iteration ($\Delta F_0$) of the regression (in 10) which only marginally improved the model.

<table>
<thead>
<tr>
<th>Observed Response</th>
<th>Predicted Response by Regression Model</th>
<th>Low</th>
<th>High</th>
<th>Creaky</th>
<th>Checked</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Low</td>
<td>373</td>
<td>114</td>
<td>70</td>
<td>3</td>
<td>66.6%</td>
</tr>
<tr>
<td>High</td>
<td>High</td>
<td>89</td>
<td>402</td>
<td>122</td>
<td>5</td>
<td>65.0%</td>
</tr>
<tr>
<td>Creaky</td>
<td>Creaky</td>
<td>68</td>
<td>122</td>
<td>581</td>
<td>41</td>
<td>71.6%</td>
</tr>
<tr>
<td>Checked</td>
<td>Checked</td>
<td>22</td>
<td>27</td>
<td>124</td>
<td>57</td>
<td>24.8%</td>
</tr>
<tr>
<td>Overall</td>
<td>Overall</td>
<td>24.9%</td>
<td>30.0%</td>
<td>40.4%</td>
<td>4.8%</td>
<td>63.6%</td>
</tr>
</tbody>
</table>
(10) Classification by Multinomial Regression after Step 6: $\Delta F0$

<table>
<thead>
<tr>
<th>Observed Response</th>
<th>Predicted Response by Regression Model</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Low</td>
<td>377</td>
<td>105</td>
</tr>
<tr>
<td>High</td>
<td>92</td>
<td>398</td>
</tr>
<tr>
<td>Creaky</td>
<td>69</td>
<td>115</td>
</tr>
<tr>
<td>Checked</td>
<td>25</td>
<td>23</td>
</tr>
<tr>
<td>Overall</td>
<td>25.4%</td>
<td>28.9%</td>
</tr>
</tbody>
</table>

Returning to perception hypothesis H11, the data and regression analysis confirm that, in order, PHONATION TYPE, F0 HEIGHT, and DURATION were significant acoustic cues to lexical Tone in Burmese. The results for F0 CONTOUR were less clear. On the one hand, the Even F0 stimuli produced essentially the same responses as Falling F0 stimuli. On the other hand, the effects of Rising vs. Falling stimuli were considerably different. For the second half of the hypothesis, a single parameter across all tones was not found.

The perception experiment results offer two key conclusions for the present study. One, creaky voice quality is integral to the phonological identity of the Creaky tone, and possibly the Checked tone. Models of Burmese tonal phonology that treat creakiness as a secondary feature or an enhancement of a pitch contrast are disputed by the strength of creaky voicing as a cue in isolation found here. Secondly, it is evident that a single feature fails to define not only the four-way contrast, but also any individual category within that system. Rather than a single, primary cue for a tone, the evidence for Burmese implies that listeners utilize multiple aspects of the acoustic signal, such that the most reliable cue is likely context dependent. Suitable identification of the High and Low tones required knowledge of F0 and non-creakiness at the least. Creaky identification was
assigned to 66.9% of all creaky-voiced stimuli, but 39% of all Creaky identification by listeners were assigned to non-creaky stimuli. **Duration** and high F0 **Height** also influenced listener responses. Still, the qualities that elicit either a Creaky or Checked identification were unclear. **Vowel Quality** in this study intentionally looked at two vowels which were believed to be fairly similar in open and closed syllables, /a/ and /ε/. It remains to be tested, but it is hypothesized that a perception study employing a stronger vowel quality contrast between open and checked vowels, e.g. /ε/~/εu/ and /ɔ/~/au/, would find as strong or stronger of a division between Creaky and Checked responses than was for /a/~/ʌ/ above.

Additionally, the perception data reinforced a finding of production data in this dissertation and elsewhere (Watkins 2005a) that the tones’ behavior falls into two subsets: (1) the longer, non-falling High and Low tones, and (2) the shorter, falling, laryngealized Creaky and Checked tones. This binary opposition is argued to be more plausible from a perceptual standpoint than the three-way phonation opposition (breathy~modal~creaky) presumed in other accounts (Bradley 1982, Green 2005), and is also more in-line with two-way register contrasts frequently reported in languages of Southeast Asia (Henderson 1952, Gregerson 1976).

As strongly as creaky phonation appeared to influence listeners’ forced choice responses, a couple caveats are in order regarding the scope of the stimuli composing the experiment. To begin with, not only were all stimuli recorded from a single speaker, but all stimuli controlled for F0 and **Duration** were produced from only six recordings from that speaker. A strict interpretation of the results at present is that they only confirm with certainty that breathiness as produced in two specific speech acts was regarded by
listeners as equivalent to modal voicing (and likewise regarding the relevance of creaky voicing beyond the two creaky prototype tokens). Generalizing these findings to other speech acts and other listeners is not a terrible stretch, but certainly more stimuli (and different speakers) are warranted in future study. A second limitation to the stimuli is the lack of sentential contexts. Since production findings for this dissertation indicate that phonation qualities present in isolation in Burmese are systematically weakened or even absent in juncture, testing the perceptual cues in sentential environments could offer conclusions which better accompany the production data of Chapter Four.
Chapter Seven. The Phonology of Burmese Lexical Tone

7.1 The Phonological Composition of the Burmese Tones

7.1.1 Overview

This chapter presents a phonological account of the lexical tone contrast which integrates the findings of the acoustic, articulatory, and auditory studies conducted for this dissertation. Patterns revealed by these experiments are shown to be incompatible with prior analyses, particularly concerning the interchange of pitch and phonation distinctions. In production, the lexical tones were shown to utilize features of both Tone (pitch-based) and Register (voice quality-based) systems of contrast that overlap in certain contexts. In perception, native speakers in one such context (isolated syllables) employed both of these systems in order to classify the tonemes. Given these findings, a suitable analysis of Modern Burmese lexical tone should reflect the concurrent pitch and phonation qualities, as opposed to focusing on a single suprasegmental contrast. The present section outlines the proposed phonological representation of the tonal contrast. The discussion in §7.1.3 further establishes the requisite attributes for a model of the contrast, progressing from broad concerns of theoretical alignment to the individual phonological details requiring explanation. Section 7.1.4 outlines the set of distinctive laryngeal features widely-accepted in the literature on tone and register and their application to the Burmese contrast. Following this, §7.1.5 – 7.1.6 describe the Burmese-specific implementation and moraic alignment of these features. The remainder of the chapter applies this account of the underlying contrasts to the properties and alternations found in the production data, using Lee’s (2007) Optimality Theoretic analysis as a starting point.
7.1.2 The Proposal

The proposed representation functions within an autosegmental model that requires only the specification of an H tonal autosegment, a [CONSTRICTED GLOTTIS] feature, and the coda segments /N/ or /Ɂ/ to describe the necessary contrasts. The forms in (1) and diagrams in (2) demonstrate the proposed input and output representations of each tone.

(1) Input Forms: underlying contrast between the four lexical tones

(a) **High**  

CV^H  

/kʰa^H/  

(b) **Low**  

CV  

/kʰa/  

(c) **Creaky**  

CV^{H,[C,G]}  

/kʰa^{H,[C,G]}/  

(d) **Checked**  

CV^H  

/kʰa^{H}/

(2) Output Forms: phonological outputs of citation form utterances

(a) **High**  

H  

µ µ  

/kʰa/  

(b) **Low**  

H  

µ µ  

/kʰa/  

(c) **Creaky**  

H^{[C,G]}  

µ µ  

/kʰa^{H}/  

(d) **Checked**  

H  

µ µ  

/kʰa^H/  

Crucially, Chapter Five showed that every tone-bearing syllable in Burmese is best understood as bimoraic (also argued by Green 2005) and demonstrated that CV, CVN, and CVO vowel and coda durations matched this analysis. Bimoraic High, Creaky, and Checked syllables are all understood to lexically bear an H tone and no L tone exists in the analysis, either with Low tone syllables or compositionally within a complex, contour tone\(^1\). Rather, the Low tone is not signaled by a tonal feature, but by the absence of high

---

\(^1\) Notation here and throughout this chapter disambiguates the different uses of “low” and “high” as follows: Lowercase “low” or “high” is used as a descriptive adjective denoting low or high pitch, but not specifically a tone autosegment or one of the four Burmese tones. Tonal autosegments are always referred to with only an isolated capital letter L or H, while “Low” or “High” fully written out in title case is reserved for representing the two tone categories. The labels and diagrams in (1) and (2) demonstrate this usage.
or low pitch targets. This approach specifically differs from prior analyses (Yip 2002, Green 2005, Lee 2007) that posit phonological L tones or contour tones (such as HL) or derive pitch features from underlying phonation contrasts. By instead aligning tonal features in Burmese to the mora rather than to the syllable, timed pitch peaks can be captured more elegantly.

7.1.3 Desiderata for a Featural Representation of Burmese tone

The analysis herein assumes that phonological features and representations correspond to phonetic qualities (whether articulatory or acoustic) and that categorical phonetic processes are best explained within the phonology. Therefore, every consistent phonetic difference should be taken seriously in regards to its featural representation. The strongest or most salient differences signaling a contrast are deemed distinctive, and weaker differences identified as redundant or possibly tied to enhancement features (Stevens and Keyser 1989). Redundant features, reliant on the presence of a primary, are preferably excluded from the underlying representation and treated as either the product of phonological demands or phonetic implementation – an often intricate difference that can be difficult to tease apart.

This view of the phonetics-phonology interface is not a given – a more divided approach would separate the computation of abstract phonological representations and processes from their phonetic implementation. In the case of Burmese, each lexical tone could be represented as a single suprasegmental, [HIGH TONE], [LOW], [CREAKY], and [CHECKED], which carries with it a complex of phonetic qualities. In this understanding, the contrasting suprasegmentals can represent large-scale interactions between the tones.
(e.g. ‘induced Creaky tone’ in (3) below), and the varying phonetic forms of the tones are treated as specific implementations governed by the phonetics. The ‘induced Creaky tone’ of Burmese can in this way be expressed simply, via the re-association of the altered syllable to a single phonological unit (3b) instead of the wholesale transfer of multiple autosegments grounded in phonetic characteristics (3c).

(3) Association of Creaky features in ‘Induced Creaky Tone’ (Okell 1969: 18)

- **a. Base verb form**
  
  ![Diagram for Base verb form]
  
  ‘(he) came’

- **b. Re-linking to Abstract Tone Unit**
  
  ![Diagram for Re-linking to Abstract Tone Unit]
  
  ‘the man who came’

- **c. Re-linking to Compositional Tone Unit**
  
  ![Diagram for Re-linking to Compositional Tone Unit]
  
  ‘the man who came’

However, such an approach can fail to capture generalizations related to isolated phonetic qualities. For Burmese specifically, reference to an independent \([\text{CONSTRICTED GLOTTIS}]\) feature on the Creaky tone, rather than part of the phonetic implementation of a monolithic suprasegment, permits one to treat the loss of laryngeal features in juncture (both \([\text{C.G.}]\) and Checked \([^?]\)) as a uniform process.

At the other end of the spectrum from highly abstract tone units is an explicit approach that fully specifies each phonetic quality associated with the tones (i.e. Creaky tone as \([[\text{HL}], \text{[CONSTRICTED GLOTTIS]}, [-\text{LONG}], [+\text{STRESS}]]\), High tone as \([[\text{H}], \text{[SPREAD GLOTTIS]}, [+\text{LONG}], [+\text{STRESS}]]\)). Parsimony, of course, represents a competing demand in phonological explanation, making such an explicit approach disfavored. The model in (1) aims to employ as few phonemic specifications as necessary to describe the numerous phonetic differences between the tones, and in turn be able to explain the
asymmetric distributions of component features (from tonal autosegments to vowel quality to possible syllable shapes) and their known alternations in a grounded manner.

Looking toward the details of the analysis, the basic phonetic characteristics of duration, pitch, and phonation type found with each tone should be accounted for, whether in the underlying representation, the phonological output, or the mapping to the phonetics. Prior analyses have achieved this to a point, but the abundance of new data collected for this dissertation has both refined and expanded the challenge. A number of phonetic facts were clarified in earlier chapters, most notably regarding phonation type. The list in (4) reviews the findings seen as integral to an analysis.

(4) Phonetic Data to be accounted for by an analysis

a. **Phonation** forms a robust two-way distinction (High/Low vs. Creaky/Checked) in citation, isolated, and phrase-final forms. This distinction is neutralized in medial positions.

b. **Duration** differences are prominent in citation forms of the tones, but compressed medially and in connected speech.

c. **Pitch** can be reduced to three pitch profiles: An even, moderate pitch (Low tone), an early high pitch with an ensuing fall (Creaky, Checked tone), and a high pitch with varying location (High tone).

d. **Perception** in an identification task was dominated by a two-way phonation contrast and pitch height.

These data alone do not constitute the entirety of the puzzle posed by Burmese tonal phenomena. A focus of Green’s (2005) work was the curious distribution of diphthong
vowels only in closed syllables. Lee (2007) sought to explain why creaky voice and obstruent codas were restricted to falling tones. The phonological patterns requiring explanation are outlined in (5). Each of these concerns is addressed in the feature composition of the lexical tones argued for in §7.2 or in the Optimality Theoretic analysis of their distribution and evaluation in §7.3.

(5) Phonological Patterns to be explained
   a. **Asymmetry of Pitch to Syllable Type and Phonation Type.**
      Open, clear syllables may have a low or high pitch, but syllables with creaky voice and closed syllables are always produced with a high pitch.

   b. **Sonorant-Closed Syllables (CVN)** align with CVO syllables regarding vowel quality, but with CV regarding tone.

   c. **Vowel Quality** in CVN syllables reflects the vowel set associated with closed syllables, even when the syllables surface as open CṼ.

   d. **Reduction of Laryngeal Contrasts** in juncture was a prominent effect between contexts, ideally captured as a uniform operation.

(5a) represents a central phonological dilemma posed by the distributional restrictions on the four tones – how to account for the co-occurrence of high pitch and glottal constriction, as well as the lack of a tone with glottal constriction but no high pitch. The diagram in (6) depicts this asymmetric distribution. The present proposal resolves this issue via the underlying specification of an H tone autosegment with the Creaky and Checked tones, but alternative explanations are examined where tone autosegments are derived rather than underlying.
7.1.4 The Phonology of Laryngeal Features

Distinctive features for laryngeal contrasts used in this dissertation follow the feature set of Halle and Stevens (1971), incorporating ensuing insights from Lombardi (1991) and Avery and Idsardi (2001). The system of laryngeal features proposed in Halle & Stevens (1971) defines laryngeal articulations via four binary features which reflect the width and tension of the glottis. Importantly, the features define possible configurations of the larynx rather than the consequences of these configurations (Keating 1988). For example, the feature \([+\text{SPREAD GLOTTIS}]\) indicates a wide glottal aperture rather than the aspiration which could result from such an aperture. The feature \([+/- \text{ CONSTRICTED GLOTTIS}]\) is used to distinguish consonants produced with glottalic airstreams and creaky, tense, or otherwise glottalized voice qualities on vowels. This generality to the features permits their extension to both vowel and consonant distinctions, enabling a more streamlined understanding of patterns that encompass both categories. On consonants, the features define phonetic qualities of voicing, aspiration, and breathiness or glottalization, while on vowels the same distinctive features have consequences for pitch and voice qualities. One such pattern is the relation between consonant voicing and F0 height realized on adjacent vowels, the grounding for some of the core tonogenetic analyses (see §2.7). In Halle and Stevens’ system, one can nicely capture these F0 effects by assuming the transfer of a
single feature from the consonant to the vowel, such as [+STIFF V.F.] from an unvoiced stop onset to a vowel where the feature is implemented as a heightened F0. Following Lombardi (1991), all four features are taken here to be privative rather than binary – elements are either marked as bearing a [FEATURE] or the feature is absent. The table in (7) presents the possible privative feature combinations and the resulting laryngeal quality realized either consonantly (with a labial stop) and vocalically (with the low vowel /a/).

| (7) Distinctive Features for Laryngeal Settings, Adapted from Halle and Stevens (1971: 203) |
|---------------------------------|---------------------------------|---------------------------------|-----------------|
| Glottal Tension | Glottal Width | Consonantal | Vocalic |
| Ø | Ø | /b/ | /a/ |
| [STIFF V.F.] | Ø | /p/ | /á/ |
| [SLACK V.F.] | Ø | /b/ | /à/ |
| Ø | [SPREAD GLOTTIS] | /p/ | /ã/ |
| Ø | [CON. GLOTTIS] | /b/ | - |
| [STIFF V.F.] | [SPREAD GLOTTIS] | /pʰ/ | - |
| [SLACK V.F.] | [SPREAD GLOTTIS] | /bʰ/ | /ã/ |
| [STIFF V.F.] | [CON. GLOTTIS] | /p’/ | /a*/ |
| [SLACK V.F.] | [CON. GLOTTIS] | /b’/ | /ã/ |

Physiologically, a range of laryngeal articulations can raise or lower pitch. For example, raising the larynx often accompanies the raising of pitch (Hirose 1997) while lowering the larynx induces a vocal fold setting that tends to reduce pitch (Ohala 1973, see Honda et al. 1999 for details). While different articulations (and combinations of articulations)
can yield similar F0 consequences, targeted modulation of F0 is generally considered to
be achieved via the cricothyroid and thyroarytenoid muscles, which function to stretch or
(1994) explains the role of separate parts of the vocal folds in different rates of vibration:
tension of the vocal fold cover is responsible for high to medium range F0, while the
greater mass of the vocal fold body (composed of ligaments and the thyroarytenoid and
vocalis muscles) is mostly involved in lower range F0 production. Because the
thyroarytenoid is itself part of the muscle structure of the vocal folds, contraction clearly
increases vocal fold tension. However, Titze points out, this tension in the vocal fold
body has the effect of relaxing the vocal fold cover and therefore diminishing the tension
necessary for rapid vibration of the folds.

Looking at phonological laryngeal features, the observations above are captured
by treating F0 as primarily a product of [STIFF]/[SLACK]/[Ø] settings of vocal fold
tension. Definitions of glottal width ([SPREAD/CONSTRUCTED]) on the other hand, do not
control F0, but provide an indirect influence. Phonation type, even more so than pitch, is
the product of dynamic laryngeal configurations represented featurally by combinations
of both Glottal Tension and Glottal Width (and Larynx Height as well, in Avery and
Idsardi’s (2001) model). In the table in (7), “creaky voice” is the result of mapping two
laryngeal features to the phonetic component: [SLACK VOCAL FOLDS] and [CONSTRUCTED
GLOTTIS]. [CONSTRUCTED GLOTTIS] signals glottal adduction, though the manner in which
this adduction is achieved varies according to the presence of other features (STIFF or
SLACK V.F.) and language specific implementation. [SLACK VOCAL FOLDS] is the primary
contributor to a lower F0, since a laryngeal state where longitudinal and medial
compression is great but the vocal fold cover is lax will still produce lower F0 values. Prior phonological analyses of Burmese (see §2.8) have used \([\text{C.G.}]\) or simply \([+\text{Creaky}]\) to represent the glottalization associated with the Creaky tone. The present analysis maintains this usage of just a \([\text{Constricted Glottis}]\) feature to capture the phonation type of Creaky tone syllables. In this case, it is not necessary that \([\text{Slack Vocal Folds}]\) be phonologically specified, since the language-specific implementation of the \([\text{C.G.}]\) feature is thought to entail a slack vocal fold cover.

### 7.1.5 The Mora as the Tone Bearing Unit in Burmese

The mora is well-established as a landing site for tone in phonology both generally (Hyman 1985, Pierrehumert & Beckman 1988, Hayes 1989, Clements 2000) and in Southeast Asian languages (Duanmu 1990, 1994 for Chinese dialects, Morén & Zsiga 2006 for Thai). A key motivation for invoking morae as the TBU in most analyses (from Hyman 1985 onward) has been to permit a sequence of level tones which compose a contour tone. Note however, that although the mora is the TBU in (5), the syllable is cast as the unit of contrast. Creaky, Checked, and High tones have phonological specifications at the syllable level which are not contrastive by their Tone segment alone. Notably, the analysis of Burmese here supposes that compositional contours that mix pitch and phonation qualities are possible, and that the Burmese Creaky and Checked syllable templates are examples of such a structure. The existence of tonal pitch contours in language is uncontroversial (though there is some debate as to how to best represent them). Furthermore, Thongkum (1988) and DiCanio (2009) have shown dialects of the Chong language (Mon-Khmer) use a distinctive register contour. In Takhian Thong
Chong (DiCanio 2009: 162), a breathy-tense dynamic register contrasts with simple modal, breathy, and tense registers. The proposal for Burmese assumes a similar contour of laryngeal qualities, but of both pitch and phonation properties, rather than just one or the other. For the Creaky tone the contour is entirely suprasegmental [H-c.g.], while the structure of Checked tones employs a segmental component, [Hʔ].

Morén & Zsiga’s (2006) analysis of Thai tones, in particular, has been influential to the present account regarding both the underspecification of tone and the timing of pitch peaks linked to morae. In their study, an H or L pitch target was tied to the right-edge of a mora. Alignment of the H tone to the first mora produced a pitch peak at the syllable mid-point, while alignment of H to the second mora produced a peak at the syllable offset. Importantly, empty TBUs were understood to be unspecified for tone, denoting no active laryngeal gesture. Thus, “phonologically toneless moras are phonetically mapped to the neutral/default pitch range (of that speaker)” (Morén & Zsiga 2006: 134). Details of their data, particularly the realization of unspecified TBUs, are discussed in depth in §7.2.2.

Likewise for Burmese, it is assumed that a) empty TBUs are permitted and b) an unspecified tone value corresponds to a default F0. Falling contours, phonologically specified as HM or HL in Yip (1995), can be re-analyzed as a passive return to a neutral F0 from a preceding H tone (in the case of HM) or a similar passive fall augmented by the effects on periodicity of a glottalized coda (instead of a specified HL). This re-analysis further renders an L tone unnecessary, as the rising or falling contours of the non-Low tones can result from moraic timing of peaks rather than phonologically specified L tones.
7.1.6 Pitch and Other Correlates of Glottal Constriction

For phonologically contrastive purposes, the feature [CONSTRICTED GLOTTIS] is critical to the distinction between Creaky and High tones. Assuming that a /Ɂ/ consonant is specified as bearing [CONSTRICTED GLOTTIS], the feature furthermore splits the tone inventory into two groups, High/Low vs. Creaky/Checked, which each share a number of common phonetic features. The phonetic interpretation of the feature [CONSTRICTED GLOTTIS] is also significant to the analysis. Following the discussion in §7.1.4, the analysis holds that [CONSTRICTED GLOTTIS] segments in Modern Burmese are regularly implemented with [SLACK] and not [STIFF] vocal folds. This position is supported by much phonetic evidence. Inspection of the acoustic signal in Creaky tone samples revealed a protracted vowel offset with multiple irregular and widely-spaced glottal pulses. This acoustic profile reflects a slack vocal fold setting and creaky voicing, where an adducted, but fairly relaxed state does not produce vibration damping. Furthermore, it was determined in Chapter Five that late decreases in open quotient on either Creaky or Checked tones was correlated with lower F0s -- that is, the more glottalized a vowel’s offset was, the lower the offset F0 tended to be. Contrary to these findings, tense phonation, with an adducted glottis but stiff vocal fold cover, would exhibit quickly damped and more frequent offset pulses. While the vowel-final glottalization in Modern Burmese is clearly not tense voiced, a historically tense glottal constriction (i.e. [STIFF VF], [C.G.]) could be the source of the high tone with Creaky and Checked syllables. However, given the imprecise nature of the Creaky tone’s historical origin (§2.7), it would not be prudent to presume which direction the phonologization occurred – whether glottalization sourced high pitch or vice versa. The diachronic development is not clear
enough to state a causal relationship in either direction. Still, it is worthwhile to note that slightly different forms of glottalization in previous stages of the language could be more compatible with a high pitch. In this understanding, compatible with the chronologies in Maran (1969, 1971) and Thurgood (1978), glottalization once acted to raise the Tone’s pitch but now occurs at the syllable offset as a pitch-lowering creak voice quality. A low offset pitch could be a phonetic effect of the creak, or a reinforcing feature of *enhancement*, strengthening the salience of vowel-final glottalization.

The analysis can also reflect the connection between tone and intensity (§2.3) by treating the phonetic interpretation of an H autosegment similarly to stress. Watkins (2001: 293) states that 'stress is subordinate to tone in Burmese'. That is, stress is not assigned separate from tone, but as part of tone. Bernot (1972) describes High, Creaky, and Checked as the 'strong' tones, and it is common for the Low tone to be described as 'unstressed'. The position taken here is not that an [H] tone is stress, but rather, some form of prominence. Along with pitch, it has acoustic correlates that resemble prominence in many languages. By attributing intensity differences to the phonetic implementation of an H tone, the present analysis captures at least a simple two-way distinction between High, Creaky, Checked vs. Low tones without complicating the phonology. Without further data on the realization of intensity differences in sentence contexts, it is not meaningful to elaborate further than this simple approach.

The proposed alignment of features offers this simple generalization: Burmese uses open, glottalized, and closed syllables. Open syllables may or may not be produced with prominence, but glottalized and closed syllables always are. The primary acoustic correlates of this prominence are a heightened F0 and increased intensity.
Reviewing the findings of Chapter Four and Five, the phonetic evidence supporting the proposed representations in (1) and (2) can be separated into three main arguments. These are detailed in the following subsections concerning syllable structure (§7.2.1), the absence of specified L tones (§7.2.2), the temporal alignment of features (§7.2.3), and the interaction of pitch and phonation type (§7.2.4).

7.2 Evidence for the Proposed Analysis

7.2.1 Evidence for Possible Syllable Structures, Possible Rhymes

The syllable structure assumed in the analysis generally follows that of Green (2005), but for the treatment of the diphthong vowels. Chapter Five presented the arguments and phonetic data in support of the structures in (8).

(8) Surface Form Structure of Bimoraic Syllables in Burmese

The two forms in (b) represent the same underlying syllable in phonetically open (b') and closed (b'″) contexts.
Vowel and coda segment durations reflected different moraic weights for CV/C\(\tilde{\text{V}}\), CVN, and CVO vowels, and these weight classes account for the majority of the duration differences between the tones in connected speech. Therefore, a phonologically specified vowel length distinction between the tones is unnecessary, since it can fall out from the syllable type distinctions. Marginally shorter Creaky tones were common, but are seen to be the (sometimes statistically significant) effect of syllable final glottalization.

A few problematic details of Green’s (2005) analysis are also resolved. Chapter Five established that differences in vowel quality between the tones were better treated as underlying rather than produced by any active alternation. There are three sets of Burmese vowels aligned with the appropriate syllable types in (8) above: oral vowels, nasal vowels, and closed vowels, of which the sets of nasal and closed vowels mostly overlap. Superficially similar vowel pairs such as open /e/ and closed /eu/ were shown to not be allophonic variants of one another, neither at present nor ever. Historically, closed /ei/ was a velar-closed allophone of /i/ – not in other closed syllables nor as a variant of /e/. It is therefore not appropriate to derive open syllable vowels by banning bimoraic diphthongs as Green does. In fact, the diphthong vowels should not be excluded from bimoraic structures at all, since as shown in (8b’) no such restriction exists (i.e [taʊ] or [teɪ]). Thus, the moraic weight of vowels is in no way shaped by the vowel quality and is determined solely by the containing syllable type (CV vs. C\(\tilde{\text{V}}\) vs. CVN vs. CVO).
7.2.2 Evidence for a Lack of [L] Specifications

The following section argues against the use of an L tone autosegment in Burmese, and argues for the representation of contrastive pitch as an opposition between H and Ø, a null setting. Importantly, the null tone is phonologically, but not phonetically unspecified, a difference illustrated in (9) below following the discussion in Keating (1988). The diagrams in (9a-c) demonstrate three possible outcomes for a form underspecified for a feature which is surrounded by units fully specified for that feature. The discussion is specifically shifted to concern tone, with an idealized resulting pitch contour set below each autosegmental sequence.

Example (9a) depicts a phonologically underspecified form which receives the feature specification H via phonological spreading (left-to-right), while (9b′–b″) demonstrate forms where the intervening TBU is also phonetically unspecified. In these cases, a pitch target is entirely absent and the pitch contours are a straightforward interpolation between the closest adjacent targets. Both pitch contours show the absence of an intervening pitch target, as the direct transition between High and Low targets (9b′) or the lack of a contour between same targets (9b″).

<table>
<thead>
<tr>
<th>(9) Strategies to Resolve Phonological and Phonetic Underspecification of Pitch</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Phonological Underspecification</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Phonological Spreading</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

![Pitch Diagrams](image-url)
Keating (1988), in an influential proposal, argued that another outcome was possible – the phonological output may remain unspecified but language-specific phonetic rules interpret the null value in a predictable fashion. This is shown in (9c), where a null specification is realized as a value mid-way between the High and Low tones. The contour in (9c) differs notably from the phonetic underspecification accounts (9b) in the appearance of an intermediate steady state reflecting the target supplied by the phonetics. Note though that this example intentionally offers a clear difference in contour, while phonetic interpretations quite similar to (9a) and (9b) are in principle possible, depending on the phonetic strategies for co-articulation preferred by the language in question. This fact does not undermine the arguments for pitch values being possibly assigned in the phonetics, it simply requires other data to diagnose the type of underspecification as well as the fill-in strategy. Unspecified spans longer than a single TBU are preferred in this case, because regardless of length, pitch interpolation will extend over multiple TBUs that are phonetically underspecified. On the other hand, for a language that assigns phonetic values to phonologically unspecified TBUs, a passive return to a mid-range F0 is achieved during the unspecified span, the shape of the fall being dependent on the strength of perseverative and/or regressive co-articulatory effects in the given language.

Data from Morén and Zsiga’s (2006) analysis of Thai tones demonstrate some of the critical issues here. These data show that any unspecified span between two tones does not fully assimilate to the adjacent tone qualities, but gradually transitions between

---

2 α represents any TBU. Since (9) demonstrates options available to any language, a specific tone-bearing unit is not stated.
the adjacent syllable’s specified F0 and a default level. Examples (10a-b) illustrate the realization of unspecified TBUs in isolated tone-bearing syllables in Thai.

(10) Unspecified Tone in Thai per Morén and Zsiga (2006)

a. Phonological representations of Thai High and Low tones produced in isolation (M&Z 2006: 114, example 3).

i. Low Tone

\[ \overset{\mu}{\sigma} \]

ii. High Tone

\[ \overset{\mu}{\sigma} \]

b. Phonetic interpretation of Thai High and Low tones produced in isolation (M&Z 2006: 131, Figure 4).

i. Low Tone pitch contour

![Pitch vs. Rhyme Duration](image1)

ii. High Tone pitch contour

![Pitch vs. Rhyme Duration](image2)

Both cases concern an unspecified mora preceding a tone on the following mora, due to the highly ranked alignment of tone to the right-edge of the syllable. If the pitch characteristics of a tone were to spread to unspecified TBUs (i.e. $\emptyset H \rightarrow HH$), a fairly even high or low pitch would be expected. Instead the pitch contours in (10b) reflect a rise or fall from an early mid-range value corresponding to the $\emptyset$ TBU.

Looking to sequences of tones longer than an isolated syllable, single or multiple $\emptyset$ tones between two H tones are not expected to remain high in Thai, since an unspecified TBU denotes an unmarked, default pitch level. Morén and Zsiga (2006)
found precisely this pattern for a single Mid tone between two High tones, a sequence which would presumably entail the string of TBUs in (11).

(11) Phonological representation of a High-Mid-High sequence in Thai

\[
\begin{array}{c}
H & \Ø & \Ø & \Ø & H \\
\sigma & \sigma & \sigma & \sigma \\
\end{array}
\]

Figure 1. Phonetic interpretation of the Thai middle tone in select three tone sequences, adapted from Morén and Zsiga (2006: 155, Fig. 7). The Mid tone of a High-Mid-High sequence shows initial co-articulation with the preceding High, and then passively falls to a moderate F0 range. The Low-Mid-Low sequence does not have a commensurate rise.

Figure 1 shows a null span, the two toneless TBUs of a Thai Mid Tone, between (a) two H tones, (b) two L tones, and (c) between two other Mid tones. The assimilatory effect of neighboring tones on F0 is evident. Between High tones, unspecified TBUs are expected to drop partially or fully to a neutral level depending on the length of the null span and language-specific phonetic rules, and mutatis mutandem between Low tones\(^3\). In Thai, with stronger perseverative than anticipatory co-articulation (Gandour et al. 1992, Gandour et al. 1994, Morén and Zsiga 2006), the pitch tends to fall from a preceding high or rise from a preceding low, but gradually over the course of the entire syllable in either case.

\(^3\) The asymmetry between High _ High and Low _ Low frames is notable, but aside in the present paper. As no L tone is claimed for Burmese, the focus here is on the interaction of H and Ø tones.
Taking stock, three types of underspecification were discussed above, of which type (c) describes the representation of the Thai Mid tone in Morén and Zsiga (2006) as well as the present proposal for Burmese.

(12) Varieties of Underspecification

a. Phonological underspecification with a phonological resolution (e.g. spreading)
b. Phonetic underspecification with a phonetic resolution, in the sense that a physical system indeed does something during the transition between two intentful gestures.
c. Phonological underspecification with a phonetic resolution.

In Burmese, it is argued that only an H tone is specified in the phonology and that surface pitch characteristics of syllables without an underlying H are interpreted in the phonetics. Multiple arguments support this position, the first being that all instances of phonologically “low” pitch in the data are not low, but roughly mid-range for each speaker. The data in support of these claims were discussed at length in Chapter Four.

Still, an [L] autosegment need not signify truly low pitch, but simply lowness in phonological opposition to another defined value. It is therefore critical that in addition to not being low-pitched, Low tones incur no known phonological alternations on the segmental or suprasegmental content of neighboring syllables. The Burmese Low tone seems to be phonologically null regarding its effect on other tones.

Finally, acoustic evidence matches an account with no purposeful lowering of pitch in Burmese that is independent of simultaneous creakiness. Low tones in all
contexts held a neutral pitch throughout the token after a smooth initial fall, the height of which was dependent on the preceding tone. The statistical comparison in §5.6.3 confirmed this perseverative F0 co-articulation from a just prior peak; Low tone syllables in a post-H context had a much higher initial F0 than those in a post-L context. In nearly every case, the drop from the higher onset F0 was gradual, extending across the entire duration of the syllable. Figure 2 displays the individual F0 traces of the Low tone tokens produced by all speakers in Phrases 3 and 7 (High __Low and High __ minor, respectively), standardized and re-transformed to Hz.

![F0 Traces for Low Tone tokens, when following a High tone](image)

**Figure 2.** F0 traces for Low tone tokens following a High tone. The passive return to neutral F0 occurs gradually over the entire vowel duration.

En masse, the data show a passive F0 decline from the preceding High tone pitch peak that encompasses the entire syllable, and not the deliberate lowering of pitch. To be clear, these data do not definitively rule out an analysis with lowering signaled by an L tone. Instead, they are argued to show that an account lacking this phonological command can ably capture the same patterns. Combined with the lack of evidence that Low tones are
phonologically active, the need to posit an L tone which offers no additional explanatory power is called into question.

A potentially effective test for the deliberateness of pitch lowering found on Low tones would be to look at the production of Low tones in a [High __ High] frame. This context however, was not included in the eight Carrier Phrases recorded for the present study. Since Chapter Four found F0 co-articulation to mainly be progressive, the present account would predict that a Burmese Low tone be realized in this frame similarly to the Thai Mid tone in such a position (cf. (11) above) – falling initially from the prior tone’s peak, only rising minimally in anticipation of the following High tone. Sprigg (1977: 17) describes the tri-syllabic pitch contour of a High-Low-High sequence in [pʰoóndò디], “a monk”. He represents the tone realization as \_
\_
\_ (falling-low-high), a characterization which mostly matches the expected result described here, though with an earlier fall, on the first syllable rather than across the second.

Given a suitable duration as in the Figure 2 examples, a passive return to a mid-range F0 is found in the data. Over a short unspecified span however, a single mora for instance, the effects of co-articulation could strongly color the pitch contour, possibly to the point of maintaining a fairly high pitch throughout an H\_Ø\_H sequence. Figure 3 illustrates the F0 effects of an adjacent High tone, in this case on an embedded High tone syllable. Recall that the High tone has one mora associated with an H and another null mora. The representations in (13) give these associations for three Carrier Phrases, and are paired with three F0 traces in Figure 3.
(13) Phonological Representation of Consecutive High Tones

<table>
<thead>
<tr>
<th>Preceding Token</th>
<th>Embedded High Tone Token</th>
<th>Following Token</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø H</td>
<td>Ø H</td>
<td>Ø Ø</td>
</tr>
<tr>
<td>Ø Ø</td>
<td>Ø H</td>
<td>Ø H</td>
</tr>
<tr>
<td>Ø H</td>
<td>Ø H</td>
<td>Ø Ø</td>
</tr>
</tbody>
</table>

In notation above: “Ø H” = ((Ø)_{μ}(H)_{μ})_{σ}.

Figure 3. High Tone F0 in context of an adjacent H. F0 traces for High tone syllables in Phrases 3, 4, and 7 show little to no fall. The mean Low tone trace across all three Phrases is provided for comparison.

The underlined spans in (13) illustrate why there is no substantial fall in the F0 trajectories of Figure 3. The gap between the H of the embedded token and that of the adjacent syllable is not linked to an L tone (which are not posited), but is unspecified (the bold Ø). With no operative intervening L target, the pitch remains relatively high because a single mora is too brief a period for F0 to naturally fall to a default level between two specified H targets. Recall that High tone forms in other phrases either had a clear early F0 peak (Phrase 1 – isolated, Phrases 5 and 6 – phrase-finally) or had a rising F0 to a late peak (Phrase 2 – [Low__Low], Phrase 8 – [Low__Minor]). Only medially with an adjacent High tone are High tones produced with a relatively even high pitch.
It has been argued here that the lack of low-targeted trajectories allows an account to forego phonological specification of L. A phonologically unspecified mora with default fill-in rules explains the F0 patterns just as capably, and does so more simply. But what of the reverse approach – underspecification at even the phonetic level? Phonetic underspecification of tone would predict pitch on unspecified syllables to be directed by context – one finding described above. However, while the contextual effects on Low tones’ onset F0 were statistically significant, the same perseverative F0 co-articulation was found broadly across Burmese, for tokens of any tone (including Minor syllables). Furthermore, F0 determined by interpolation between specified targets is clearly irreconcilable with an analysis with only H specifications – as the only anchors for interpolation would be H tones. Phonetic underspecification is only plausible if underlying L tones are assumed, in which case phonetic underspecification of Low tones is a moot point.

7.2.3 Evidence for Temporal Alignment of Tones and Laryngeal Features

One of the major findings of the experiments in this dissertation was that, like F0, the reported phonation distinctions in Burmese were systematically realized temporally on the vowel. Specifically, creaky and breathy voicing were perceptible almost exclusively at the vowel offset. The proposed representations in (2) offer an improvement on previous models by reflecting the timing of the phonetic manifestation of the [CONSTRIC\(\)TED \(\)GLOTTIS\(\)] feature (and trivially, the glottalization of the glottal stop coda).
The pitch contours and alternations depicted in §4.3 are also more accurately accounted for by the tone alignments of the proposed model. The association of an H autosegment to the first mora yields an early pitch peak – matching the production of Creaky and Checked tones in all environments and High tones in isolation (cf. Figure 4.11) and phrase-finally (Figures 4.12, 4.13). When associated to the second mora, a gradual rise to a late peak is produced, as with High tones preceding a Low tone (Phrase 2, cf. Figure 4.14) or Minor syllable (Phrase 8, cf. Figure 4.19).

7.2.4 Evidence for the Independence of Pitch and Phonation Properties

A crucial detail of Lee’s (2007) analysis was that the high pitch of Creaky and Checked tones was induced from an underlying [c.g.] feature or [ʔ] segment. This section demonstrates that, while their distributions overlap in the tonal inventory, the H tone of Creaky and Checked syllables functions independently of the glottal constriction in the phonology and should therefore be treated as an underlying property of these tones. The falling pitch identified with these tonemes is argued to be the phonetic consequence of the late glottal constriction after an early H, but no L tone is necessarily posited for the fall. Other accounts (Green 2005, Yip 2002) link Checked syllables to no tone at all, an approach not considered here in the face of ample acoustic data to the contrary (cf. Figures 4.11 – 4.19).

To begin with, the open quotient and aerodynamic results of sections 4.4 and 4.5 confirmed that the end of Creaky and Checked vowels were produced with a tighter glottal configuration than High and Low tone vowels. However, it was also shown that this difference was regularly neutralized in juncture, evident in larger harmonic
differences for H1-A1 and H1-A3, higher OQ ratios, and equivalent airflow rates across tones. For Checked syllables the neutralization was even more apparent in the full assimilation of [ʔ] codas to a following onset consonant, such that /tàʔ/ + /ko/ → [tàːk:o].

Section 5.7 demonstrated that falls in pitch were likewise neutralized to some extent in juncture and that there was a low-level negative correlation between the degree of glottal constriction and F0 height (§5.7.3).

The correlation raises the question of whether glottalization or a falling F0 can be treated as a primary goal of Creaky and Checked toneme, and the other quality as an incidental phonetic by-product of that goal? That is, does the implementation of heavy glottalization at the vowel offset induce a lowering of F0, or is it the case that a lower rate of glottal vibration induces glottalization? On this point, it is helpful to consider the case of the Checked tone.

The loss of laryngeal features with glottal stop coda assimilation offers an opportunity to study pitch behavior without the hypothesized trigger of glottalization. The Checked tone is described at times in the literature as bearing a high and even pitch (§2.4). It is tempting to suppose that these authors were basing their description on medial Checked syllables with an assimilated (non-glottal) obstruent coda, with the assumption that high-pitched Checked tones do not fall when they lack the sharp glottal closure of the glottal stop coda segment. However, in every case (for the medial positions of Phrases 2, 3, 4, 7, 8), a high F0 and a falling contour was still associated with Creaky and Checked tones, regardless of whether the syllable being measured was [tàk], [tàb], [tàm], or [tàʔ] with a glottal stop coda. Therefore, raised pitch cannot be a phonological consequence of a tightened glottis since the H tone remains regardless of the coda’s
segmental form. Lee’s (2007) constraint ranking (in (14)) incorrectly predicts that no H tone occurs without the trigger of a coda [ʔ] or [c.g.] feature.

(14) Tableau. H Tone Dependent on Glottalization in Lee (2007)

<table>
<thead>
<tr>
<th>/taʔʔ]/+/ko/</th>
<th>*COMPLEX</th>
<th>*ʔ/L</th>
<th>IDENT-T</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>t a k.ko</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

L is maintained on Checked syllables: *ʔ/L unable to trigger H tone when coda is assimilated to /k/ onset.

To summarize, the H tone is not induced by [c.g.] or [ʔ] or any other features in the present model. Nor is there a phonetic association between high pitch and the glottalization of Creaky and Checked tones – section §5.7 demonstrated that late glottalization had no correlation with the height of the early F0 peak on these tones. In fact, it was correlated with lower F0 values at the vowel offset. However, the two properties, glottal tension and a high pitch, are related historically and more importantly, synchronically in Modern Burmese in their coinciding distribution, a fact that is problematic for an Optimality Theory analysis and a core OT tenet, Richness of the Base. This distribution restriction is addressed further in the OT analysis in §7.3.5. At this moment, the principal understanding is that the H autosegment is simply part of the
lexical form of these tonal categories. Models that have no underlying tone with Creaky or Checked syllables cannot account for the regular presence of a high, falling pitch on these forms in all contexts.

7.3 Distribution of Tones

7.3.1 Motivation for an Optimality Theory Account of Burmese Lexical Tone

An appealing aspect of the proposal lain out and defended in the previous two sections is the simple underlying contrast between the four lexical tones. A syllable can be lexically stored with or without an H tone. Those without a tone are Low tones. Those with an underlying H are classified as (i) High if they are otherwise unspecified, (ii) Creaky if they are accompanied by a [C.G.] feature, or (iii) Checked if they are in a closed syllable.

In Optimality Theory, the alignment of these features, and therefore the moraic timing integral to the analysis, results from the evaluation of a constraint ranking. The Burmese data pose a dilemma for an OT analysis because of the restricted distribution of pitch contours found with creaky voiced and obstruent-closed syllables. This dissertation has argued that neither quality phonologically triggers the high pitch and has demonstrated statistically that there is certainly no phonetic causation (except that which lowers F0). Yet the distribution restriction is an undeniable part of the language’s phonological system. Richness of the Base (ROTB) forces one to answer in the phonology why there is no low-pitch creaky tone or a low-pitched checked syllable in Burmese? While it is convenient to discuss the distribution with reference to ROTB, it is important to note that these issues are not specific to Optimality Theory, and the data form a notable pattern for
any theory. The contrast in Burmese is noteworthy because, as Mazaudon and Michaud (2008) showed for Tamang, it offers an example of what a transitory state in the development of Tone can look like. Details of the formalism aside, the Burmese data question how phonologists should represent a state of affairs where seemingly redundant qualities operate independently in the phonology. If a feature A is always found with feature X, ideally the relationship is capturable by causation: either A falls out from the phonetic interpretation or phonological computation of a form with X. Yet in Burmese, creaky voicing is always found with an early high pitch target even though (a) there is no phonetic correlation between the two and (b) there is no natural phonological entailment that can explain their variable manifestation on medial and final Creaky and Checked tone forms.

The solution within OT put forth here argues for the limited inclusion of distributional entailment constraints which are language specific and reinforce existing inventories. Specifically for Burmese, a constraint \([\text{LARYNGEAL}]_{\text{rhyme}} \Rightarrow \text{H}\) reinforces the distributional restriction which only licenses laryngeal features on the rhyme of syllables which also bear the prominence of an H tone.

\[
(15) \quad [\text{LARYNGEAL}]_{\text{rhyme}} \Rightarrow \text{H}_o: \quad \text{If a syllable rhyme is specified for a laryngeal feature, then the containing syllable must bear an H tone. That is, assign a violation to any non H syllable with marked LARYNGEAL features on the rhyme (GLOTTAL WIDTH, GLOTTAL TENSION, LARYNX HEIGHT).}
\]

This formulation assumes a default state for vowels and syllable rhymes without STIFF/SLACK vocal folds and without a SPREAD or CONSTRICTED glottis. In this way, the
sonorant N of CVN syllables is permitted with Low tone syllables, while the obstruent coda of Checked syllables requires an H tone.

Similar constraints are often regarded as unsatisfactory in phonological theory due to their lack of phonetic grounding and seemingly ad hoc nature. The present chapter contends that such constraints, while not phonetically grounded in a synchronic sense, may be motivated by documented diachronic processes that were themselves results of phonetically motivated alternations. For the Burmese Creaky and Checked tonemes, an H tone autosegment which was once induced in prominent glottalized syllables has come to be a fundamental part of that toneme’s representation.

Cross-linguistically, such a class of entailment constraints potentially explains various unnatural phonological patterns. Importantly though, explanatory power is curbed by permitting the entailment constraints in limited circumstances like those described above. Namely, when (a) the entailment reinforces an existing contrast and (b) it is supported by historical evidence of a prior stage when the distribution was the result of a phonetically grounded alternation.

In the history of the Burmese tones, increased glottal tension and aerodynamic effort in an earlier pronunciation of emphatic Creaky and Checked tones would very naturally have had raised F0 as a phonetic consequence. Speakers of an early form of Burmese reinterpreted the incidental F0 patterns as integral to the contrast, producing a prosodic unit where a distinctive voice quality co-occurs with only a high tone. By the modern era, the restriction of high pitch with an emphatic creaky vowel or obstruent-closed syllable was a fossilized distribution, and vowel-final laryngealization which lowered F0 was a vestigial property of certain tones which would not impair the
simultaneous production of high pitch.

The remainder of this chapter describes an OT analysis of the modern Burmese grammar. The account looks to not only explain the asymmetric distribution of glottal features between tones, but derives the forms for isolated, medial, and final forms of each tone set forth in Chapter Four. Along with the constraint in (15), the critical constraints and ranking to derive these forms are established.

7.3.2 Low Tone

Low tone forms should surface with no phonological tone associations in any context, neither L nor H⁴. Since there is no indication of articulatorily intended pitch lowering in Burmese tonal phenomena, a ban against L tone segments is proposed.

(16) *L: No low tones

The constraint *L prevents input low tones from surfacing when ranked above faithfulness constraints that preserve tone inputs (i.e., MAX-TONE in (17)). Since the Low tones are argued to have a null tonal content, *L must also outrank markedness constraints enforcing that some tone be realized on the output (18).

(17) MAX-TONE: Input tones have a corresponding output tone.

(18) SPECIFY-TONE: A TBU must be associated with a tone.

⁴ Elevated F0 was found at both the onset and offset of Low tone samples. In each case, this was not the result of a phonological tone; high onsets were a residual phonetic consequence of an immediately preceding F0 peak and high offsets were found only in certain intonational contexts.
The ranking in (19), demonstrated in the tableau in (20), bans L tones and permits morae to be unspecified for tone, both desired effects for Burmese. Regardless of whether the input possesses an L tone (20a) or is bare (20b), the output will have no tone association.

(19) *L >> MAX-TONE, SPECIFY-TONE

(20) Tableau. Low Tone [ta]: L tone fails to surface

<table>
<thead>
<tr>
<th>/ta^L/</th>
<th>*L</th>
<th>MAX-TONE</th>
<th>SPECIFY-TONE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>L</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td></td>
<td>µ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>Ø</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>µ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t a</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7.3.3 High Tone

The High tone possesses an unassociated input H. The output TBU for this H tone can vary however, since High tone syllables are found with both early and late pitch peaks according to context. The rankings which drive this distribution resolve conflicts with adjacent tones that either crucially violate or satisfy the alignment constraint (McCarthy
and Prince 1993) that enforces the preferred rightward alignment of tone units in Burmese, ALIGN-RIGHT(H,σ) in (21). Markedness constraints are potentially violated by adjacent tones which are either too similar (drawing from the Obligatory Contour Principle, Goldsmith 1976) or too close ((22) and (23) respectively).

(21)  ALIGN-R(H,σ): Align every H tone with the right-edge of a syllable

(22)  *T-CLUSTER: Identical tones “clustered” on consecutive TBUs are banned.

For Burmese specifically, *T-CLUSTER militates against consecutive morae linked to H tones.

(23)  TONE:TBU: A TBU may not be associated with more than one tone.

Unhindered alignment of the H for a High tone syllable is seen in the tableau in (24), whereas the tableaux in (25) and (26) depict forms where violation of ALIGN-R(H,σ) is forced. In (25), the violation is forced by the crucial ranking *T-CLUSTER >> ALIGN-R (H,σ). Phrase-finally (26), the introduction of a L% boundary tone acts to push the H tone back to the first mora. Together, these tableaux and their rankings explain the variability of the High tone pitch peak described in Chapter Four.
(24) Tableau. High Tone in Juncture: H tone aligns on second mora

<table>
<thead>
<tr>
<th>/kʰaH+/+/ta/</th>
<th>*T-CLUSTER</th>
<th>ALIGN-R(H,σ)</th>
<th>*H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>kʰa + t a</td>
<td>µ µ µ µ</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(25) Tableau. High Tone with Following H: *T-CLUSTER pushes H back to first mora

<table>
<thead>
<tr>
<th>/kʰaH+/+/σH/</th>
<th>*T-CLUSTER</th>
<th>ALIGN-R(H,σ)</th>
<th>*H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H H</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>kʰa + σ</td>
<td>µ µ µ µ</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(26) Tableau. High Tone in Final Position: Boundary tone pushes H to first mora

<table>
<thead>
<tr>
<th>/kʰáH+/ + L%</th>
<th>TONE:TBU</th>
<th>ALIGN-R(H,σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H L%</td>
<td></td>
</tr>
<tr>
<td>kʰa</td>
<td>µ µ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H L%</td>
<td></td>
</tr>
<tr>
<td>kʰa</td>
<td>µ µ</td>
<td></td>
</tr>
</tbody>
</table>
7.3.4 Creaky and Checked Tones

Creaky and Checked tone forms present two issues for which a theory of the tones’ composition must account. First, why are [CONstricted GLOTTIS] forms only found with H tones? Second, assuming a general pressure for tone segments to align rightwards, why does H routinely associate to the first leftmost mora in Creaky and Checked tone syllables, in violation of ALIGN-R(H,σ)?

Tackling the second matter, the association to the first mora is argued to be underlying for Creaky tones and forced onto the only tone-bearing segment for Checked tones (a vowel as opposed to aperiodic [ʔ]). Resolving the first issue addresses previous claims (such as in Lee 2007) that the H tone is the product of the increased glottal tension associated with creaky voicing or a glottal stop. The argument for the independence of pitch and phonation properties in §7.2.4 presented the crucial data for this discussion, seen again in (27). The alternations in (27) have been described by Okell (1969), Thurgood (1978), and others and were explicitly confirmed in the production data of Chapter Four.

In the two examples, when the [c.g.] feature is lost in juncture, the H tone is still realized and remains aligned with the first mora. This is the case for both [c.g.] representing creakiness on the Creaky tone vowel (27a) and the [c.g.] feature of a glottal stop segment (27b).
(27) Neutralization of Glottal Constriction in Juncture

a. Creaky Tone maintains H
   /tä\H[^C.G]/ + /ko/ \rightarrow [tä\H[^G].go]

b. Checked Tone maintains H
   /tä\H[^C.G]/ + /ko/ \rightarrow [tä\H[k.k:o]]

Gussenhoven and Teeuw (2008) and Frazier (2009) report a similar pattern for the glottalized tone in Yucatec Maya, a Mayan language of Mexico. The table in (28) reproduces Table I from Gussenhoven and Teeuw (2008).

<table>
<thead>
<tr>
<th></th>
<th>Tone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Short</td>
</tr>
<tr>
<td>Final</td>
<td>Low</td>
</tr>
<tr>
<td>Initial</td>
<td>Low</td>
</tr>
</tbody>
</table>

Like Burmese, the Yucatec Maya Glottalized tone always bears a falling pitch contour, but the vowel-final glottalization does not always surface\(^5\).\(^6\) Also like Burmese, the High tone falls in final position but rises elsewhere. In Gussenhoven and Teeuw’s analysis, an

---

\(^5\) Frazier (2009) presents the neutralization pattern far less categorically, occurring roughly in 40% of phrase-final tokens and 10-20% of medial or initial tokens in her data.

\(^6\) The pattern has been found widely across languages, with both contrastive (as in Burmese and Yucatec Maya data) and non-contrastive glottalization, suggesting some phonetic grounding for the claim that glottalization is more easily produced in phrase-final position. Multiple studies of variable glottalization in varieties of American English have shown any effect promoting vowel glottalization is neutralized medially (Pierrehumbert and Talkin 1992, Redi and Shattuck-Hufnagel 2001, Huffman 2005, and Roberts 2006). Vowel glottalization near or at the vowel offset was stronger and more common phrase-finally and was found less frequently on phrase-medial tokens. Stevens (2000) ascribes this phenomena to the reduced aerodynamic supply and lax glottal state that accompany utterance finality. Regardless of the precise laryngeal configuration used to create them, non-modal glottal states show a preference for phrasal boundaries.
H-tone is attached at the syllable level to the *Long High* tone, but “is firmly anchored on the first mora” of the *Glottalized* tone. In this way, the *Long High* H is able to reposition itself according to different boundary tones while the *Glottalized* H is always syllable initial.

A similar analysis is employed for the Burmese Creaky tone. The tableaux in (30) and (31) demonstrate the isolated (with creaky voice) and medial (without creaky voice) forms of the tone. In (29) is a formulation of the markedness constraint which drives the attested neutralization of creaky voicing in juncture with a following syllable.

(29) \textsc{agree[lar]}: Laryngeal features [\textsc{c.g}] and [\textsc{s.g}] at the end of a syllable must agree with those of the next syllable onset.

Framed as agreement between portions of adjacent syllables, the constraint taps into the generalization that Burmese onsets undergo an assortment of transformations in juncture (detailed in Sprigg 1957) that revolve around voicing contrasts. Further, the assimilation of glottal stop codas similarly forces agreement between consonants of adjacent syllables not just in laryngeal features, but \textsc{place} and \textsc{manner} of articulation as well.
(30) Tableau. Creaky Tone in Isolation: Association of H to first mora.

<table>
<thead>
<tr>
<th>/k^h_a H,[C.G]/</th>
<th>IDENT[+LAR]?</th>
<th>MAX-T</th>
<th>*ASSOC</th>
<th>ALIGN-R(H, σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H [C.G.]</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>k^h a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[C.G.]</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>k^h a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k^h a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(31) Tableau. Creaky tone in Juncture: High tone surfaces; associated to first mora even when [CONstricted GLOTTIS] segment is removed.

<table>
<thead>
<tr>
<th>/k^h_a H,[C.G]/ + /ta/</th>
<th>AGREE[LAR]</th>
<th>*ASSOC</th>
<th>IDENT[+LAR]</th>
<th>ALIGN-R(H, σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H [C.G.]</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k^h a + ta</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>[C.G.]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k^h a + ta</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k^h a + ta</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7 IDENT[+LAR] protects the correspondence of laryngeal features from the input to the output (following Lombardi 1991).
*ASSOC disallows new associations in the output. It is violated by both movement or spreading of tones from an input association.
ALIGN-RIGHT is similarly dominated in Checked tone evaluation by a constraint against the realization of tone on obstruents. This constraint, TONESONORITY, is listed in (32). Note that sonorant N codas with High, Low, and Creaky tones all surface with tonal contours – ruling out the alternative analysis that tone is aligned to the vowel nucleus. For Checked syllables, the unavailability of the second mora as a landing site forces the tone to surface earlier in the syllable, whether as an isolated syllable (33a) or in juncture with an assimilated coda (33b). Note too that AGREE[LAR], as with Creaky tones in juncture, forces the assimilation\(^8\) of the glottal stop coda.

(32) TONESONORITY: A TBU must be headed by a sonorant segment.

(33) Tableau. Checked Tone in Juncture: High tone surfaces with assimilated [ʔ].

\[\begin{array}{|c|c|c|}
\hline
/CV^[ʔ]/ (+ ko) & AGREED[LAR] & TONE SONORITY & ALIGN-R(H, σ) \\
\hline
\text{a)} & & & * \\
\hline
\text{b)} & *! & & \\
\hline
\end{array}\]

\(^8\) The highly ranked CODA CONDITION (Itô 1986, as an OT constraint in Itô and Mester 1994) initially forces the restrictive coda inventory for Burmese, yielding only coda segments without independent PLACE features. The re-linking of these features in (33b), where a Checked syllable has a [k] coda, does therefore not violate CODACONDITION since the PLACE feature is headed by the following onset.
7.3.5 Discussion

Thus far, the argument has sought to demonstrate that the H tone of the Creaky and Checked tones is underlying and not a product of the tonemic glottalization. However, it is certainly recognized that the two properties, increased glottal tension and high falling pitch, are intertwined – both historically and synchronically. Phonetically, productions of Creaky and Checked tones with lower open quotients were associated with steeper falls in F0 over the second half of the vowel (§4.3.6). In the model presented here, this phonetic effect contributes greatly to, but is not the sole source of the falling contour identified with the Creaky and Checked tones. Moraic timing alone was shown to capture much of the F0 patterns.

The Creaky and Checked tones not only had matching pitch contours, but matching phonation contours as well – glottalization was reliably a feature of the end of the syllable, where F0 was low. Speakers did not produce a “tense” glottal constriction on Creaky vowels that might be connected to elevated pitch, but the observed non-modal phonation was consistently creaky, characterized by irregular slowed vibration. The conclusion is that creaky phonation, as manifested in Burmese on the Creaky and Checked tones, acts to lower F0. Lee (2007) also attributes a falling F0 to the laryngealization as a phonetic effect, since Creaky and Checked syllables only bear a phonological H tone in his analysis. The H tones are forced to surface on [+creaky] and CV? syllables by the high-ranked constraints */L and */L⁹. Consequently, his analysis poses something of a paradox: glottalization raises pitch phonologically but then lowers pitch phonetically. The origin of these contradictory outcomes is interpreted here to be

⁹ Note though that no mechanism is provided to ensure that the H is realized early and is followed by the [c.g.] feature with concomitant F0 lowering. The temporal alignment of the specifications is resolved with the constraint discussed in (28) below.
In OT terms, markedness constraints like those proposed by Lee were the source of the restricted distribution. For the present discussion, they can be conflated into a single constraint, formalized as $[\text{GLOTTAL WIDTH}] \Rightarrow \text{H}\sigma$ in (34). Bear in mind that (34) is proposed not for the modern distribution of features, but to capture the historical link between phonation and pitch in the four tones' phonetic forms.

(34)$[\text{GLOTTAL WIDTH}] \Rightarrow \text{H}\sigma$: If a syllable bears a $[\text{GLOTTAL WIDTH}]$ feature (c.g. or s.g.), then it must bear an H tone.

On the understanding that an H tone is the primary instantiation of prominence, the constraint above restricts certain articulatorily difficult outputs to prominent syllables. The constraint interaction in Lee’ (2007) posed the relationship as causation, i.e. creaky voice begets a high tone. Conversely, the licensing constraint in (34) taps into the cross-linguistic generalization that more marked structures are permitted in prominent positions (deLacy 2002). Highly-ranked, $[\text{GLOTTAL WIDTH}] \Rightarrow \text{H}\sigma$ would have ensured that any $[\text{CONSTRICTED GLOTTIS}]$ output occurred only on H tone bearing syllables. Further, by casting the net wider to include increased glottal aperture as well as constriction, the link between phonation and pitch is extended to the formulation of the High tone, where breathy voicing was historically likely.

When the H tone became grammaticalized as a specified feature of the High, Creaky, and Checked tonemes, the constraint could become demoted in the grammar without altering output forms dramatically. However, the effect of $[\text{GLOTTAL WIDTH}] \Rightarrow \text{H}\sigma$ is still evident in the modern language since the distribution
holds – there is no contrastive \([CV^{c.g.}]\) tone. The constraint was thus broadened to \([\text{LARYNGEAL}]_{\text{rhyme}} \Rightarrow H_o\), the constraint employed by this account and repeated in (35). Note that this to include any kind of marked laryngeal feature, thereby enforcing the distribution in Creaky syllables and in Checked syllables with either [?] or assimilated codas.

(35) \([\text{LARYNGEAL}]_{\text{rhyme}} \Rightarrow H_o\): If a syllable rhyme is specified for a laryngeal feature, then the containing syllable must bear an H tone. That is, assign a violation to any non H syllable with marked LARYNGEAL features on the rhyme (GLOTTAL WIDTH, GLOTTAL TENSION, LARYNX HEIGHT).

The fixed association of H to the first mora of a Creaky tone syllable can be explained in a similar fashion. Given a production of "creakiness" that conditions pitch lowering, creaky voice and an H specification are incompatible. The constraint in (36) is violated by output forms that prescribe their simultaneous production. Both Yip (2002) and Morén and Zsiga (2006) propose similar constraints tying glottalization to low pitch.

(36) \(*H[C.G.]_{\text{TBU}}\): an H tone and [C.G.] feature specification may not be associated with the same TBU.

Possible output forms in MB never violate (36), so the constraint could be undominated. Given the constancy for H on the left mora in Creaky syllables, the grammaticalization of this link would be expected. Akin to the principle of LEXICON OPTIMIZATION (Prince and Smolensky 1993), Burmese learners subjected to only the surface forms of Creaky syllables would postulate underlying forms with a left-bound H.
7.3.6 Accomodating Richness of the Base, Evidence from Chang (2008)

Underlying representations form a significant part of the present analysis. While a number of phonological facts are accounted for by a simple distinction of [H] or no [H] between the High and Low tones, Creaky and Checked tone phenomena require input specifications which may seem more than minimal: the H tone must be specified and must be further specified on the Creaky tone to associate with the first mora. Modern phonological theory places an emphasis on output evaluation, given that stipulations on the input often provide little insight. The tableaux below demonstrate that this analysis does not limit the possible suprasegmental inputs, in accord with Richness of the Base, to derive the four lexical tones of Burmese. Furthermore, the two constraints in (35) and (36) are shown to to interact minimally with modern grammar, needing only to be ranked above IDENTLAR\textsuperscript{10} to yield the modern asymmetrical distribution of tone to phonation types to syllable types.

The basic contrasts between the four lexical tones rely on three qualities: an H autosegment, a [c.g.] autosegment, and a [?] coda. Under Richness of the Base, all possible combinations of these features must be accounted for. These are listed in (37) with those that match regular allotones shaded out.

\begin{equation}
(37) \quad \text{Possible Suprasegmental Inputs}
\begin{align*}
a. /CV^{(\Theta)}/ & \rightarrow \text{Low} \\
b. /CV^{H}/ & \rightarrow \text{High} \\
c. /CV^{H,c.g.}/ & \rightarrow \text{Creaky} \\
d. /CV^{?H}/ & \rightarrow \text{Checked} \\
e. /CV^{c.g.}/ \\
f. /CV^{?}/ \\
g. /CV^{?c.g.}/ \\
h. /CV^{H,c.g.}; H \text{ is unlinked} \\
i. /CV^{(\Theta)H,(H,c.g.)H}/ \\
\end{align*}
\end{equation}

\textsuperscript{10} And by extension, ALIGN-RIGHT(H,σ), though no necessary ranking is demonstrated.
To account for the possible inputs in (37e-g) requires an explanation as to why there are no creaky or checked syllables without an H tone autosegment. The two additional forms in (37h) and (37i) posit inputs where the H is either unlinked or pre-specified for the second mora, thereby causing the analysis to answer whether pre-specification of H to the first mora is necessary for one tone (as hypothesized). To begin with, only evaluation of these inputs in isolated or phrase-final position warrants more than minimal attention. The expected outputs in (38) provide the only discussion of medial forms.

(38) Output Forms of Select Inputs, AGREE[LAR] dominates, forcing neutralization of [c.g.] or glottal stop coda in juncture.

a. / CV^{c.g.} / \rightarrow [CV]
b. / CV? / \rightarrow [CVO]
c. / CV^{c.g.} / \rightarrow [CVO]
d. / CV^{H,c.g.}; H is unlinked \rightarrow [CV^H]
e. / CV^{(O)}d(H,c.g.)_H / \rightarrow [CV^{O}\ H] \text{ or } [CV^{H\ O}]

In positions where the laryngeal features are not neutralized, it is only necessary that \([\text{LARYNGEAL}_{\text{rhyme}} \Rightarrow H\sigma, *H[c.g.]_{\text{TBU}} >> \text{IDENT[LAR]} \text{ or } \text{DEP-T.} \) These rankings permit the possible repairs of deleting a feature responsible for the conflict or inserting an H to form a licit Creaky or Checked syllable. Evidence from the adaptation of English loanwords (Chang 2008) supply intriguing evidence that \([\text{LARYNGEAL}_{\text{rhyme}} \Rightarrow H\sigma >> \text{DEP-T is the appropriate ranking to repair the inputs of (37e-g). Chang’s data indicate that an obstructuent coda is re-analyzed as a Checked syllable (cf. 39), while non-nasal sonorant codas /l/ or /r/ are met with High or Low tone adaptations (cf. 40).} \]
(39) Obstruent Final Loanwords (from Chang 2008)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Philippines</td>
<td>→</td>
<td>[pʰi.liʔ.paɪ]</td>
</tr>
<tr>
<td>b. jacket</td>
<td>→</td>
<td>[dʒeʔ.keʔ]</td>
</tr>
<tr>
<td>c. brake</td>
<td>→</td>
<td>[bə.roiʔ]</td>
</tr>
<tr>
<td>d. club</td>
<td>→</td>
<td>[kə.ləʔ]</td>
</tr>
<tr>
<td>e. mart</td>
<td>→</td>
<td>[məʔ]</td>
</tr>
<tr>
<td>f. October</td>
<td>→</td>
<td>[ʔaʊʔ.too.ba]</td>
</tr>
</tbody>
</table>

(40) Non-nasal Sonorant Final Loanwords (from Chang 2008)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ball pen</td>
<td>→</td>
<td>[bɔ.pi]</td>
</tr>
<tr>
<td>b. April</td>
<td>→</td>
<td>[ʔeɪ.pi]</td>
</tr>
<tr>
<td>c. four</td>
<td>→</td>
<td>[pʰou]</td>
</tr>
<tr>
<td>d. rubber</td>
<td>→</td>
<td>[ra.ba]</td>
</tr>
</tbody>
</table>

The tableau in (41) shows how the ranking \([\text{LARYNGEAL}]_{\text{rhyme}} \Rightarrow \text{Hσ} \gg \text{DEP-T}\) captures this pattern. Obstruent-final inputs will violate DEP-T in order to preserve the coda.

(41) Tableau. Closed Syllable Loanword Rendered as Checked Tone.

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>/ mart /</td>
<td>CODACOND</td>
<td>[LARYNGEAL] (\Rightarrow) Hσ</td>
<td>MAX-T</td>
<td>DEP-T</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H H</td>
<td>H</td>
<td>m λ t</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>H H</td>
<td>H</td>
<td>m λ ?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H H</td>
<td>H</td>
<td>m λ ?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H H</td>
<td>H</td>
<td>m λ ?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For creaky inputs without a tone, \([\text{LARYNGEAL}]_{\text{rhyme}} \Rightarrow \text{Hσ}\) can force deletion of the \([\text{C.G.}]\) segment or insert an H. Again, loanword data indicate that at least some English
loanwords map to the Creaky tone, so H insertion is chosen as the repair.

(42) Tableau. Closed Syllable Loanword Rendered as Creaky Tone.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[C.G.] H [C.G.]</td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>[C.G.] n [C.G.]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[C.G.] m [C.G.]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Chang (2008: 50) lists the example count [kaʊ̃], a word prone to final glottalization in English, particularly the British English dialects with which Burmese bilinguals are potentially more familiar. Presumably, other demands could pressure a /CV[^C.G.]/ form to delete the [C.G.] feature, contra the evaluation in (42), and give a toneless [CV] output. This matter is not resolved here, as the more crucial point has been demonstrated – that a /CV[^C.G.]/ input yields a licit Burmese output.

Finally, for the possible input forms with an unlinked H and [C.G], the constraint *H[^C.G.\]TBU need only dominate IDENT[LAR] to yield a licit output resembling a High tone via deletion of [C.G.]. The same ranking resolves inputs with a prespecified H on the second mora\(^{11}\).

\(^{11}\) Alternatively, a Creaky tone would provide the most harmonic output if the ranking *H[^C.G.\]TBU >> ALIGN-RIGHT, *ASSOC held, pressuring violations of alignment and reassociation in order to separate the H and [C.G.] feature.
7.4 Conclusion

This chapter has proposed a model for the lexical representation of each tonal complex that specifies syllable structure, and the presence or absence of an H tone and \texttt{[CONSRCTED GLOTTIS]} feature. Evidence from instrumental phonetic studies supported many of the unique features of this proposal, such as the lack of a phonologically specified L tone and the moraic timing both of pitch peaks and of the \texttt{[C.G.]} on the second half of \textit{creaky} syllables. The analysis above also provides a simple understanding of the distribution of tones and laryngeal features in Burmese tonology. All tones prefer the right edge, but conflicting laryngeal features push underlying tones leftward to the first mora on Creaky, Checked, and (in some contexts) High syllables. This understanding not only explains the proper distribution of surface features amongst the four tones, but also the absence of tones that might otherwise be predicted, such as (i) a rising, glottalized tone and (ii) a falling, modal tone, and (iii) low or even pitch glottalized or checked tones.

This model of the tonal contrast was compared to alternatives which held a greater economy of representation, but did not reflect the multiple modes of data examined in the study. The proposed representations of tone-bearing syllables in Burmese have been informed by patterns revealed in \textit{acoustic} and \textit{articulatory} production data, by the \textit{auditory} salience of creaky-voicing (seen in citation forms), and by historical insights into the \textit{diachronic} relationship between vowel qualities, tones, and coda consonants. Because each of these data are integrated into the model, the representations and constraint interactions set forth in this section are argued to best capture the synchronic contrast across contexts in Burmese speech.
Chapter Eight. Conclusion.

This dissertation has resolved a number of phonetic questions about the tones of Burmese, and in turn, used these data to inform an explanation of phonological patterns of contrast between the tones and possible syllable types which had heretofore eluded a satisfactory analysis.

While descriptions of the four tones often focused on the reported phonation differences, it had not been clear if there was even a consistent association between certain tones and non-modal phonation types. Burmese was generally regarded as a “Register” language, and yet phonetic studies questioned whether High tones were produced with breathy-voicing or even whether Creaky tones were produced with creaky-voicing. Similar uncertainty existed concerning the consistent realization of pitch contours with certain tones. The phonetic investigation of this dissertation (Chapter Four) revealed systematic context-sensitive differences in Duration, Pitch, and Phonation Type. Primarily, this was a distinction between medial forms and utterance final forms (including isolated syllables). The second half of the dissertation (Chapters Five through Seven) explained two phonological puzzles: the distribution of vowels to syllable types (thus, overlapping tone) and the seemingly redundant co-occurrence restriction of glottal constriction and high pitch.

Chapter One presented the general problem of classifying the suprasegmental contrast in Burmese. It was shown that while linguists discussing the Burmese sound system tend to use the label “Tone”, descriptions of the tones tend to revolve around the differences in phonation rather than pitch. Sample tokens of each tone demonstrated that
both pitch and phonation differences were evident, and that the question going forward was not whether the distinctions existed, but to what degree and in which environments? Another problem of phonological analysis was outlined – the behavior of sonorant and glottal codas and the asymmetric distribution of the possible vowel qualities with the possible codas.

The phenomena in Burmese were then placed within the wider context of Tonogenesis (or Registrogenesis) in Southeast Asian languages. A typology of languages from the region was discussed along with stages of tonogenesis ranging from fully redundant phonation and pitch distinctions to contrasts along just one of these dimensions. A thorough investigation of the overlapping and inter-dependent qualities in Modern Burmese was argued to be an ideal case study of the synchronic and diachronic interplay of these phonetic properties.

**Chapter Two** reviewed the phonetic and phonological descriptions offered in the literature on Burmese Tone. Impressionistic phonetic descriptions could vary considerably from one another but were at even greater odds with the findings of instrumental phonetic descriptions, notably over some of the most fundamental properties of the tones (i.e. is the Creaky tone creaky-voiced?). The phonetic review pulled together these various sources to highlight the aspects of the phonetic description most in need of clarification. The chapter then considered potential explanations for the discrepant portrayals of tone in Burmese, and outlined the nature of the data that phonetic experiments could bring to bear on the matter. Two central limitations of prior studies were (a) the examination of only citation form utterances of each tone and (b) reliance on
indirect acoustic measures of phonation. The review suggested the importance of including glottographic and aerodynamic data as well as production data, more generally, in a range of sentential contexts.

Lastly, the final portion of Chapter Two reviewed the handful of analyses which have aimed to capture the phonological representation of the four tones’ contrast. Focal points were Lee’s (2007) analysis of the interaction between pitch-based tonal autosegments and laryngeal features, and Green’s (2005) analysis of syllable shape and vowel quality in Burmese.

Chapter Three described the methodology of a series of production experiments collecting the acoustic, EGG, and aerodynamic data employed in the statistical analyses of Chapters Four and Five. Interpretation of these data presented a number of issues for both the segmentation of sound waves and the calculation (manual or automated) of useful acoustic (F0, harmonic amplitudes) and articulatory correlates (OQ, airflow) of the phonetic properties under investigation. Each problematic issue was described and illustrated, followed by an explanation of how measurements were calculated in order to account for potential skewing that issue might introduce.

Chapter Four presented the findings of the production experiment for measures of (in order, and by section) Duration, F0, Open Quotient, and Airflow in ml/sec. Statistical analyses tested the strength of each measure as a correlate of a four-way Tone contrast. Statistically significant distinctions were found in each of these measures in some but not all sentential contexts. Primarily, a context sensitivity was revealed between
medial forms and utterance final forms (including isolated syllables). It was also shown that even in positions where the phonation contrast was most manifest, it was best described as a two-way opposition between a more open glottal state and a glottal constriction realized late on the tone-bearing vowel. The late onset of the marked laryngeal quality points to a central insight of the dissertation – the need to investigate phonation properties dynamically, just as studies of F0 have long done. Another clarification of the phonetic study was that certain acoustic measures were simply not an accurate metric of phonation type in Burmese (H1-H2 in particular).

The phonetic claims above are made with the full knowledge that a limited set of contexts were examined in a non-conversational laboratory setting. Numerous systematic permutations of tone production undoubtedly occur in other phrase positions, with non-focused elements, or as a result of myriad tone sandhi rules. Acknowledging that undescribed complexities exist in Burmese tonal behavior does not diminish the consistency of the findings which informed the present conclusions.

While Chapter Four addressed a primarily phonetic question of correlates to tone, **Chapter Five** tested the same set of production data for patterns or correlations which spoke to the phonological concerns of syllable moraicity, weight-bearing codas, and the effects on pitch and phonation during sequences of tones on adjacent syllables. It was shown that coda segments in Burmese were heavy, a finding which was considered to support the analysis of all syllables bearing any tone as bimoraic. Phonological N codas in phonetically-open CVN syllables had no effect on the duration of the preceding vowel, evidence which ran counter to the analysis of Green (2005). F0 co-articulation between
syllables was found to be strongly perseverative and not anticipatory, and the F0 contours of Creaky and Checked tone syllables were statistically indistinguishable from one another across contexts (when normalized for durational differences). Furthermore, the canonical falling pitch of these two tones occurred regardless of the degree of glottalization realized on a token. It could not be argued that creaky-voicing or constricted glottal features conditioned the falling pitch. Conversely, there was no statistical correlation between the early high F0 of tone-bearing syllables (High tones included) and the phonetic instantiation of non-modal phonation. Taken together, the statistical findings indicate that high pitch was a feature independent of phonation in High, Creaky, and Checked tone syllables. This finding is interpreted in the phonological model of Chapter Seven as an underlying phonological H target associated with syllables bearing these tones.

Additionally, accounts of the historical development of the vowel system from the earliest written records of Burmese (Old Burmese) to the language of a few centuries ago (Written Burmese) to the modern language undermined the prevailing analysis of the Burmese vowel system found in Mehnert and Richter (1972) or Green (2005). Namely, the asymmetric sets of open-syllable and closed-syllable vowels cannot be characterized as an allophony conditioned by coda segments.

Chapter Six described the methodology and results of a simple perception study that tested the salience of the multiple phonetic dimensions composing Tone. Listener responses to natural tokens and tokens re-synthesized for controlled levels of Duration, F0, and Phonation Type signaled a key role for phonation differences in the perception
grammar. Importantly, listeners only made a two-way distinction in phonation, matching the findings of the production data collected for this dissertation. As the perception study only tested isolated tone-bearing syllables, the opposition was between CREAKY stimuli and BREATHY and MODAL stimuli, which elicited similar responses.

Finally, in Chapter Seven, the various sources of data (duration, F0, spectral tilt, OQ, airflow, perceptual, and historical) are collectively considered in the presentation of a phonological model of both the underlying contrast on lexical items and the common alternations found in the production experiment. The phonological account offered a few notable approaches to the Burmese data which are potentially applicable to tone and register phenomena in other Southeast Asian languages as well as more generally. Borrowing an insight from Morén and Zsiga (2006), the account used no targeted Low tone in the mapping to the phonetics. Tone-directed F0 movement was described by a series of timed H tones mapped onto a default low F0 surface. According to this approach, a falling F0 may be the result of no active H target rather than an explicitly targeted L.

The central question of the phonological analysis however concerned the synchronic treatment of overlapping features within a single system of contrast. If lexical tone in Burmese is understood as a four-way contrast realized on syllables, then it is clear that “Tone” in this case refers to a system that must be defined by more than pitch. Likewise, analysis solely by voice quality cannot differentiate the Low and High tones, which were found in this study to use the same mode of phonation. Burmese tone then, is a mixed system, where pitch and glottal features operate concurrently.
This suite of distinct values that each tone bears for pitch, phonation, intensity, duration, and syllable structure creates a kind of paper trail of the diachronic processes which have shaped the present tonal system. While a system of contrast may be diachronically transitional, the synchronic contrast at any chronological point represents part of a well-formed grammar that ought to be explainable by phonological theory. Specifically looking at Optimality Theory, the tenet of Richness of the Base effectively necessitates a single dimension of contrast, yet intermediate stages of tonogenesis should theoretically have competing, or redundant, dimensions of contrast. The integrated analysis of Chapter Seven offered an explicit example, with Modern Burmese, of redundant contrastive features functioning simultaneously within a synchronous grammar.
Appendix I. Production Data by Carrier Phrase

1. H1-H2 Results

Carrier Phrase (1): Isolated Utterances

Figure 1. H1-H2 at Early and Late portions of the Vowel in Phrase (1).
The distribution of Early and Late H1-H2 values across all tokens for all speakers.

Figure 2. ΔH1-H2 in Phrase (1). The mean difference between $H1-H2_{\text{LATE}}$ and $H1-H2_{\text{EARLY}}$ in each token.
Figure 3. H1-H2 at Early and Late portions of the Vowel in Phrase (2).
The distribution of Early and Late H1-H2 values across all tokens for all speakers.

Figure 4. ΔH1-H2 in Phrase (2). The mean difference between H1-H2$_{\text{LATE}}$ and H1-H2$_{\text{EARLY}}$ in each token.
Figure 5. H1-H2 at Early and Late portions of the Vowel in Phrase (3).
The distribution of Early and Late H1-H2 values across all tokens for all speakers.

Figure 6. ΔH1-H2 in Phrase (3). The mean difference between H1-H2\textsubscript{LATE} and H1-H2\textsubscript{EARLY} in each token.
Figure 7. H1-H2 at Early and Late portions of the Vowel in Phrase (4). The distribution of Early and Late H1-H2 values across all tokens for all speakers.

Figure 8. ΔH1-H2 in Phrase (4). The mean difference between H1-H2\textsubscript{LATE} and H1-H2\textsubscript{EARLY} in each token.
Carrier Phrase (5): [High ___ #] frame

Figure 9. H1-H2 at Early and Late portions of the Vowel in Phrase (5).
The distribution of Early and Late H1-H2 values across all tokens for all speakers.

Figure 10. ΔH1-H2 in Phrase (5). The mean difference between H1-H2_{LATE} and H1-H2_{EARLY} in each token.


**Figure 11.** H1-H2 at Early and Late portions of the Vowel in Phrase (6). The distribution of Early and Late H1-H2 values across all tokens for all speakers.

**Figure 12.** ΔH1-H2 in Phrase (6). The mean difference between H1-H2_{LATE} and H1-H2_{EARLY} in each token.
Carrier Phrase (7): [High ___ Minor syllable] frame

Figure 13. H1-H2 at Early and Late portions of the Vowel in Phrase (7).
The distribution of Early and Late H1-H2 values across all tokens for all speakers.

Figure 14. ΔH1-H2 in Phrase (7). The mean difference between H1-H2_{LATE} and H1-H2_{EARLY} in each token.
Figure 15. H1-H2 at Early and Late portions of the Vowel in Phrase (8). The distribution of Early and Late H1-H2 values across all tokens for all speakers.

Figure 16. ΔH1-H2 in Phrase (8). The mean difference between H1-H2_{LATE} and H1-H2_{EARLY} in each token.
2. H1-A1 Results

Carrier Phrase (1): Isolated Utterances

Figure 17. H1-A1 at Early and Late portions of the Vowel in Phrase (1).
The distribution of Early and Late H1-A1 values across all tokens for all speakers.

Figure 18. ΔH1-A1 in Phrase (1). The mean difference between H1-A1_{LATE} and H1-A1_{EARLY} in each token.
Carrier Phrase (2): [Low __ Low] frame

Figure 19. H1-A1 at Early and Late portions of the Vowel in Phrase (2).
The distribution of Early and Late H1-A1 values across all tokens for all speakers.

Figure 20. ΔH1-A1 in Phrase (2). The mean difference between H1-A1\textsubscript{LATE} and H1-A1\textsubscript{EARLY} in each token.
Carrier Phrase (3): [High __ Low] frame

**Figure 21.** H1-A1 at Early and Late portions of the Vowel in Phrase (3).
The distribution of Early and Late H1-A1 values across all tokens for all speakers.

**Figure 22.** ΔH1-A1 in Phrase (3). The mean difference between H1-A1_{LATE} and H1-A1_{EARLY} in each token.
Carrier Phrase (4): [Low __ High] frame

Figure 23. H1-A1 at Early and Late portions of the Vowel in Phrase (4).
The distribution of Early and Late H1-A1 values across all tokens for all speakers.

Figure 24. ΔH1-A1 in Phrase (4). The mean difference between H1-A1_{LATE} and H1-A1_{EARLY} in each token.
Figure 25. H1-A1 at Early and Late portions of the Vowel in Phrase (5).
The distribution of Early and Late H1-A1 values across all tokens for all speakers.

Figure 26. ΔH1-A1 in Phrase (5). The mean difference between H1-A1$_{\text{LATE}}$ and H1-A1$_{\text{EARLY}}$ in each token.
Carrier Phrase (6): [Low __ #] frame

Figure 27. H1-A1 at Early and Late portions of the Vowel in Phrase (6). The distribution of Early and Late H1-A1 values across all tokens for all speakers.

Figure 28. ΔH1-A1 in Phrase (6). The mean difference between H1-A1_{LATE} and H1-A1_{EARLY} in each token.
Carrier Phrase (7): [High __ Minor syllable] frame

Figure 29. H1-A1 at Early and Late portions of the Vowel in Phrase (7).
The distribution of Early and Late H1-A1 values across all tokens for all speakers.

Figure 30. ΔH1-A1 in Phrase (7). The mean difference between H1-A1\textsubscript{LATE} and H1-A1\textsubscript{EARLY} in each token.
Carrier Phrase (8): [Low __ Minor syllable] frame

Figure 31. H1-A1 at Early and Late portions of the Vowel in Phrase (8). The distribution of Early and Late H1-A1 values across all tokens for all speakers.

Figure 32. ΔH1-A1 in Phrase (8). The mean difference between H1-A1_{LATE} and H1-A1_{EARLY} in each token.
3. H1-A3 Results

Carrier Phrase (1): Isolated Utterances

Figure 33. H1-A3 at Early and Late portions of the Vowel in Phrase (1).
The distribution of Early and Late H1-A3 values across all tokens for all speakers.

Figure 34. ΔH1-A3 in Phrase (1). The mean difference between H1-A3_{LATE} and H1-A3_{EARLY} in each token.
Carrier Phrase (2): [Low __ Low] frame

**Figure 35. H1-A3 at Early and Late portions of the Vowel in Phrase (2).**
The distribution of Early and Late H1-A3 values across all tokens for all speakers.

**Figure 36. ΔH1-A3 in Phrase (2).** The mean difference between H1-A3_{LATE} and H1-A3_{EARLY} in each token.
Carrier Phrase (3): [High __ Low] frame

Figure 37. H1-A3 at Early and Late portions of the Vowel in Phrase (3).
The distribution of Early and Late H1-A3 values across all tokens for all speakers.

Figure 38. ΔH1-A3 in Phrase (3). The mean difference between H1-A3\textsubscript{LATE} and H1-A3\textsubscript{EARLY} in each token.
Carrier Phrase (4): [Low ___ High] frame

**Figure 39.** H1-A3 at Early and Late portions of the Vowel in Phrase (4). The distribution of Early and Late H1-A3 values across all tokens for all speakers.

**Figure 40.** ΔH1-A3 in Phrase (4). The mean difference between H1-A3_{LATE} and H1-A3_{EARLY} in each token.
Carrier Phrase (5): [High __ #] frame

Figure 41. H1-A3 at Early and Late portions of the Vowel in Phrase (5). The distribution of Early and Late H1-A3 values across all tokens for all speakers.

Figure 42. ΔH1-A3 in Phrase (5). The mean difference between H1-A3_{LATE} and H1-A3_{EARLY} in each token.
Figure 43. H1-A3 at Early and Late portions of the Vowel in Phrase (6).
The distribution of Early and Late H1-A3 values across all tokens for all speakers.

Figure 44. ΔH1-A3 in Phrase (6). The mean difference between H1-A3_{LATE} and H1-A3_{EARLY} in each token.
Carrier Phrase (7): [High ___ Minor syllable] frame

Figure 45. H1-A3 at Early and Late portions of the Vowel in Phrase (7). The distribution of Early and Late H1-A3 values across all tokens for all speakers.

Figure 46. ΔH1-A3 in Phrase (7). The mean difference between H1-A3\textsubscript{LATE} and H1-A3\textsubscript{EARLY} in each token.
Carrier Phrase (8): [Low __ Minor syllable] frame

Figure 47. H1-A3 at Early and Late portions of the Vowel in Phrase (8).
The distribution of Early and Late H1-A3 values across all tokens for all speakers.

Figure 48. ΔH1-A3 in Phrase (8). The mean difference between H1-A3_{LATE} and H1-A3_{EARLY} in each token.
4. Open Quotient Results: $\Delta OQ$

Delta ($\Delta$) taken as the mean difference between $OQ_{\text{LATE}}$ and $OQ_{\text{EARLY}}$ in each token. A positive value indicates increasing contact over course of the syllable rhyme.

Figure 49. $\Delta OQ$ in Phrase (1).

Figure 50. $\Delta OQ$ in Phrase (2).

Figure 51. $\Delta OQ$ in Phrase (3).
Figure 52. ΔOQ in Phrase (4).

Figure 53. ΔOQ in Phrase (5).

Figure 54. ΔOQ in Phrase (6).
Figure 55. ΔOQ in Phrase (7).

Figure 56. ΔOQ in Phrase (8).
Appendix II. Consonant Durations Abridged in Chapter Five

1.1 Consonant Durations with an Initial High Tone

<table>
<thead>
<tr>
<th></th>
<th>High Tone CVN vs CV Medial Consonant Durations, Carrier Phrase 2</th>
<th>Sig?</th>
<th>$t$ Test comparing means</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVŋ.gV</td>
<td>CV.kV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaker A: 93.1 (5.0), 4</td>
<td>62.0 (11.3), 4</td>
<td>✓</td>
<td>$t(6) = 5.03$</td>
</tr>
<tr>
<td>Speaker B: 71.8 (10.6), 2</td>
<td>33.4, (7.9), 2</td>
<td>%</td>
<td>$t(2) = 4.11$</td>
</tr>
<tr>
<td>Speaker C: 93.6 (6.1), 4</td>
<td>32.6 (7.7), 4</td>
<td>✓</td>
<td>$t(6) = 12.42$</td>
</tr>
<tr>
<td>Speaker D: 84.7 (19.1), 4</td>
<td>48.7 (28.2), 4</td>
<td>%</td>
<td>$t(6) = 2.11$</td>
</tr>
<tr>
<td>Speaker E: 61.7 (17.5), 4</td>
<td>36.4 (9.0), 4</td>
<td>✓</td>
<td>$t(6) = 2.57$</td>
</tr>
<tr>
<td>Speaker F: 67.3 (2.3), 2</td>
<td>32.0 (2.8), 2</td>
<td>✓</td>
<td>$t(2) = 13.77$</td>
</tr>
<tr>
<td>Speaker G: 79.0 (9.3), 4</td>
<td>38.1 (6.9), 4</td>
<td>✓</td>
<td>$t(6) = 7.06$</td>
</tr>
<tr>
<td>Speaker H: 58.7 (8.4), 2</td>
<td>30.4 (3.6), 2</td>
<td>✓</td>
<td>$t(2) = 4.38$</td>
</tr>
<tr>
<td>Speaker I: 57.6 (13.8), 4</td>
<td>33.4 (17.2), 4</td>
<td>%</td>
<td>$t(6) = 2.19$</td>
</tr>
<tr>
<td>Speaker J: 62.9 (7.5), 4</td>
<td>49.0 (7.5), 4</td>
<td>✓</td>
<td>$t(6) = 2.62$</td>
</tr>
</tbody>
</table>

Expanded results from the summary in (5.18) compare the mean duration in milliseconds (followed by SD in parentheses and n (italicized)) between the C sequences of an embedded High tone coda with the following /k/ onset in Phrase 2. 7/10 speakers produced a significantly longer intervocalic /Nk/ than /k/ sequence. Non-significant differences are shaded.
(2) High Tone CVN vs CV Medial Consonant Durations, Carrier Phrase 3

<table>
<thead>
<tr>
<th>Speaker</th>
<th>CVm.bV</th>
<th>CV.bV</th>
<th>Sig?</th>
<th>t Test comparing means</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>86.8 (4.6), 4</td>
<td>57.2 (7.3), 4</td>
<td>✓</td>
<td>t(6) = 6.86 p &lt; .001</td>
</tr>
<tr>
<td>B</td>
<td>100.6 (6.5), 2</td>
<td>32.6 (6.8), 2</td>
<td>✓</td>
<td>t(2) = 10.22 p = .009</td>
</tr>
<tr>
<td>C</td>
<td>83.3 (2.8), 4</td>
<td>57.4 (12.2), 4</td>
<td>✓</td>
<td>t(6) = 4.14 p = .006</td>
</tr>
<tr>
<td>D</td>
<td>86.1 (3.4), 4</td>
<td>30.1 (8), 4</td>
<td>✓</td>
<td>t(6) = 12.88 p &lt; .001</td>
</tr>
<tr>
<td>E</td>
<td>82.4 (30.3), 4</td>
<td>38.6 (7.7), 4</td>
<td>✓</td>
<td>t(6) = 2.8 p = .031</td>
</tr>
<tr>
<td>F</td>
<td>81.2 (4.0), 2</td>
<td>29.9 (8.3), 2</td>
<td>✓</td>
<td>t(2) = 7.87 p = .016</td>
</tr>
<tr>
<td>G</td>
<td>86.9 (13.7), 4</td>
<td>33.3 (11.0), 4</td>
<td>✓</td>
<td>t(6) = 6.10 p &lt; .001</td>
</tr>
<tr>
<td>H</td>
<td>73.5 (15.5), 4</td>
<td>28.7, 9.9, 4</td>
<td>✓</td>
<td>t(6) = 4.87 p = .003</td>
</tr>
<tr>
<td>I</td>
<td>91.6 (6.8), 4</td>
<td>32.6 (13.0), 4</td>
<td>✓</td>
<td>t(6) = 8.04 p &lt; .001</td>
</tr>
<tr>
<td>J</td>
<td>76.9 (18.3), 4</td>
<td>33.1 (11.3), 4</td>
<td>✓</td>
<td>t(6) = 4.07 p = .007</td>
</tr>
</tbody>
</table>

Expanded results from the summary in (5.18) compares the mean duration between the C sequences of an embedded High tone token’s coda with the following /b/ onset in Phrase 3. All ten speakers produced a significantly longer intervocalic /Nb/ than /b/ sequence. Non-significant differences are shaded.
### 1.2 Consonant Durations with an Initial Creaky Tone

#### 3. Creaky Tone CVN vs CV Medial Consonant Durations, Carrier Phrase 2

<table>
<thead>
<tr>
<th>Speaker</th>
<th>CVŋ.gV</th>
<th>CV.kV</th>
<th>Sig?</th>
<th>t Test comparing means</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>91.1 (25.0), 4</td>
<td>85.6 (20.8), 4</td>
<td>⨯</td>
<td>$t(6) = 0.34 \quad p = .747$</td>
</tr>
<tr>
<td>B</td>
<td>76.7 (24.1), 2</td>
<td>43.0, (0.4), 2</td>
<td>⨯</td>
<td>$t(2) = 1.98 \quad p = .187$</td>
</tr>
<tr>
<td>C</td>
<td>90.8 (7.4), 4</td>
<td>66.2 (10.5), 4</td>
<td>✔</td>
<td>$t(6) = 3.83 \quad p = .009$</td>
</tr>
<tr>
<td>D</td>
<td>67.8 (13.8), 4</td>
<td>35.8 (4.4), 4</td>
<td>✔</td>
<td>$t(6) = 4.42 \quad p = .005$</td>
</tr>
<tr>
<td>E</td>
<td>50.2 (14.8), 4</td>
<td>44.8 (8.1), 4</td>
<td>⨯</td>
<td>$t(6) = 0.64 \quad p = .548$</td>
</tr>
<tr>
<td>F</td>
<td>76.3 (11.6), 4</td>
<td>61.2 (19.0), 4</td>
<td>⨯</td>
<td>$t(6) = 1.36 \quad p = .224$</td>
</tr>
<tr>
<td>G</td>
<td>61.8 (12.4), 4</td>
<td>59.1 (13.6), 4</td>
<td>⨯</td>
<td>$t(6) = 0.29 \quad p = .781$</td>
</tr>
<tr>
<td>H</td>
<td>47.3 (17.0), 4</td>
<td>32.1 (16.1), 4</td>
<td>⨯</td>
<td>$t(6) = 1.30 \quad p = .242$</td>
</tr>
<tr>
<td>I</td>
<td>61.3 (13.5), 4</td>
<td>36.6 (12.5), 4</td>
<td>✔</td>
<td>$t(6) = 2.69 \quad p = .036$</td>
</tr>
<tr>
<td>J</td>
<td>53.6 (6.6), 4</td>
<td>45.6 (11.3), 4</td>
<td>⨯</td>
<td>$t(6) = 5.50 \quad p = .002$</td>
</tr>
</tbody>
</table>

Expanded results from the summary in (5.18) compares the mean duration between the C sequences of an embedded Creaky tone token’s coda with the following /k/ onset in Phrase 2. Three of ten speakers produced a significantly longer intervocalic /Nk/ than /k/ sequence. Non-significant differences are shaded.

#### 4. Creaky Tone CVN vs CV Medial Consonant Durations, Carrier Phrase 3

<table>
<thead>
<tr>
<th>Speaker</th>
<th>CVm.bV</th>
<th>CV.bV</th>
<th>Sig?</th>
<th>t Test comparing means</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>88.7 (17.6), 4</td>
<td>55.2 (7.9), 4</td>
<td>✔</td>
<td>$t(6) = 3.47 \quad p = .013$</td>
</tr>
<tr>
<td>B</td>
<td>77.8 (0.9), 2</td>
<td>31.9 (0.5), 2</td>
<td>✔</td>
<td>$t(2) = 63.05 \quad p &lt; .001$</td>
</tr>
<tr>
<td>C</td>
<td>100.1 (14.9), 4</td>
<td>49.0 (4.2), 4</td>
<td>✔</td>
<td>$t(6) = 6.60 \quad p &lt; .001$</td>
</tr>
<tr>
<td>D</td>
<td>84.9 (7.9), 4</td>
<td>46.8 (11.3), 4</td>
<td>✔</td>
<td>$t(6) = 5.53 \quad p &lt; .001$</td>
</tr>
<tr>
<td>E</td>
<td>89.3 (16.1), 4</td>
<td>65.8 (23.2), 4</td>
<td>⨯</td>
<td>$t(6) = 1.66 \quad p = .147$</td>
</tr>
<tr>
<td>F</td>
<td>90.5 (30.4), 4</td>
<td>48.4 (8.2), 4</td>
<td>✔</td>
<td>$t(6) = 2.67 \quad p = .037$</td>
</tr>
<tr>
<td>G</td>
<td>107.4 (24.8), 4</td>
<td>70.9 (17.7), 4</td>
<td>%</td>
<td>$t(6) = 2.49 \quad p = .054$</td>
</tr>
<tr>
<td>H</td>
<td>60.3 (21.6), 4</td>
<td>36.8, (13.4), 4</td>
<td>⨯</td>
<td>$t(6) = 1.85 \quad p = .114$</td>
</tr>
<tr>
<td>I</td>
<td>97.9 (5.5), 4</td>
<td>45.9 (18.2), 4</td>
<td>✔</td>
<td>$t(6) = 5.47 \quad p = .002$</td>
</tr>
<tr>
<td>J</td>
<td>84.9 (16.6), 4</td>
<td>46.2 (5.5), 4</td>
<td>✔</td>
<td>$t(6) = 4.43 \quad p = .004$</td>
</tr>
</tbody>
</table>

Expanded results from the summary in (5.18) compares the mean duration between the C sequences of an embedded Creaky tone token’s coda and the following /b/ onset in Phrase 3. Seven of ten speakers produced a significantly longer intervocalic /Nb/ than /b/ sequence. Non-significant differences are shaded.
References.


Cornyn, William S. 1944. ‘Outline of Burmese grammar’, Language 20, 5-34.


417


Hollien, Harry and John F. Michel. 1968. ‘Vocal Fry as a Phonational Register’, *Journal of Speech and Hearing Research* 11, 600-4.


Honda, Kiyoshi, Hiroyuki Hirai, Shinobu Masaki, and Yasuhiro Shimada. 1999. ‘Role of vertical larynx movement and cervical lordosis in F0 control’, *Language and Speech* 42.4, 401-411.


420


Kim, Mi-Ryoung, Patrice Speeter Beddor, and Julie Horrocks. 2002. ‘The contribution of consonantal and vocalic information to the perception of Korean initial stops,’ *Journal of Phonetics* 30.1, 77-100.


