Revisiting glottalic production mechanisms: Evidence from phonological and non-phonological ejectives in real-time MRI

Lavinia Price

München 2026

Revisiting glottalic production mechanisms: Evidence from phonological and non-phonological ejectives in real-time MRI

Lavinia Price

Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie
an der Ludwig-Maximilians-Universität
München

vorgelegt von
Lavinia Price

München 2026

Erstgutachterin: Prof. Dr. Marianne Pouplier

Zweitgutachter: Prof. Dr. Phil Hoole

Tag der mündlichen Prüfung: 24.10.2025

Content

Content ..... **VI**
List of Figures ..... **X**
List of Tables ..... **XII**

Chapter 1 General Introduction **1**
1.1. Characterizing Ejectives: Acoustic, Perceptual, and Typological Perspectives **4**
1.1.1. Acoustic Diversity and the Limitations of the Binary Typology 4
1.1.2. Acoustic Cue Integration in the Perception of Ejectives 6
1.2. Theoretical and Empirical Perspectives on the Articulatory Mechanisms of Ejectives **7**
1.2.1. Rethinking the Role of Larynx Raising 8
1.2.2. Timing and Intergestural Coordination in Ejective Production 10
1.2.3. Challenging the Glottalic Model: Pulmonic Ejectives? 10
1.2.4. Integrating Mechanisms: Moving Beyond Canonical Descriptions 12
1.3. Non-Phonological and Historical Perspectives **13**
1.3.1. Non-Phonological Occurrences of Ejectives 13
1.3.2. Diachronic Pathways of Ejectives 14
1.4. The Present Study: A Cross-Linguistic Articulatory Perspective on Ejectives **15**
1.4.1. Acoustic Analysis 16
1.4.2. Articulatory Analysis 16
1.4.3. Structure of the Dissertation and Overview of Chapters 17

Chapter 2 Methods **18**
2.1. Use of Real-Time MRI for Articulatory Analysis **19**
2.2. Imaging **19**
2.3. General Data Collection Procedure **20**
2.4. Processing of Acoustic Signal **21**
2.5. Segmentation **22**
2.5.1. Amharic Segmentation Scheme 23
2.5.2. English Segmentation Scheme 23
2.6. Acoustic Analysis **24**
2.6.1. Closure Duration 25
2.6.2. Burst Duration 25
2.6.3. Voice Onset Time 26
2.6.4. Intensity Rise Ratio of the Following Vowel 26
2.6.5. Relative and Absolute Burst Intensity 27
2.6.6. Post-Burst Lag (PBL) Intensity 28
2.7. Image Analysis **29**
2.7.1. Tracing Vertical Larynx Movement 30

2.7.2. Estimating Intraoral Cavity Size 32
2.7.3. Assessing Supralaryngeal Constriction 33
2.8. Cross-Linguistic Design and Comparative Framework 37
2.9. Statistical Analysis 38

Chapter 3 Phonological Ejectives in Amharic 40
3.1. Introduction 40
3.1.1. The Amharic Language 41
3.1.2. Amharic Phonology 42
3.1.3. Acoustic Studies of Amharic Ejectives 44
3.1.4. Aerodynamic and Articulatory Studies of Amharic Ejectives 45
3.2. Methods 46
3.2.1. Speakers 47
3.2.2. Stimuli 47
3.2.3. Dataset Overview and Exclusions 48
3.2.4. Statistical Modeling 49
3.3. Acoustic Analysis 50
3.3.1. Closure Duration 50
3.3.2. Burst Duration 52
3.3.3. Voice Onset Time 54
3.3.4. Relative Burst Intensity 57
3.3.5. Intensity of the Post-Burst Lag 59
3.3.6. Intensity Rise Ratio of the Stop-Following Vowel 62
3.3.7. Summary of the Acoustics Results 64
3.3.8. Interim Discussion of the Acoustic Findings 64
3.4. Articulatory Analysis 67
3.4.1. Larynx Raising 67
3.4.2. Intraoral Cavity Reduction 71
3.4.3. Constriction of the Hyperpharyngeal Region 74
3.4.4. Summary of the Articulation Results 76
3.4.5. Interim Discussion of Selected Articulatory Findings 76
3.5. Articulatory Relationships and Acoustic Correlates 78
3.5.1. Linking Larynx Raising and Supralaryngeal Cavity Change 78
3.5.2. Linking Hyperpharyngeal Constriction and Supralaryngeal Cavity Change 80
3.5.3. Effect of Larynx Raising on Absolute Burst Intensity 82
3.5.4. Effect of Cavity Size Reduction on Absolute Burst Intensity 84
3.5.5. Effect of Hyperpharyngeal Constriction on Absolute Burst Intensity 85
3.5.6. Summary of the Relationship Analysis Results 87
3.6. Discussion of Amharic Ejectives 88
3.6.1. Acoustic Robustness and Core Articulatory Features of Amharic Ejectives 88

3.6.2. Coordination of Laryngeal and Supralaryngeal Gestures 89
3.6.3. Articulatory Contributions to Burst Intensity 90
3.6.4. Asymmetries Across Place of Articulation 91
3.6.5. Interpretation of the Observed Articulatory Mechanisms 92
3.7. Conclusion 93

Chapter 4 Non-Phonological Ejectives in British English 95
4.1. Introduction 95
4.1.1. Early Accounts and Regional Aspects of English Ejectives 96
4.1.2. Structural Conditioning, Frequency, and Acoustic Properties 96
4.1.3. Phonetic Accounts and Functional Roles of Ejectives in English 98
4.2. Methods 99
4.2.1. Speakers 99
4.2.2. Stimuli 100
4.2.3. Dataset Overview and Exclusions 100
4.2.4. Statistical Modeling 101
4.3. Auditory Categorization Analysis 103
4.3.1. Frequency and Structural Conditioning of Ejectives 103
4.3.2. The Glottal Replacement Category 106
4.3.3. The Unclear Category 106
4.3.4. Interim Discussion of the Auditory Categorization analysis 107
4.4. Acoustic Analysis 108
4.4.1. Closure Duration 108
4.4.2. Relative Burst Intensity 110
4.4.3. Post-Burst Lag Intensity 111
4.4.4. Summary and Interim Discussion of Acoustics 113
4.5. Articulatory Analysis 114
4.5.1. Larynx Raising 114
4.5.2. Intraoral Cavity Reduction 118
4.5.3. Constriction of the Hyperpharyngeal Region 120
4.5.4. Summary and Interim Discussion of Articulatory Results 122
4.6. Articulatory Relationships and Acoustic Correlates 123
4.6.1. Linking Larynx Raising and Supralaryngeal Cavity Change 123
4.6.2. Linking Hyperpharyngeal Constriction and Supralaryngeal Cavity Change 125
4.6.3. Effect of Larynx Raising on Absolute Burst Intensity 127
4.6.4. Effect of Cavity Size Reduction on Absolute Burst Intensity 128
4.6.5. Effect of Hyperpharyngeal Constriction on Absolute Burst Intensity 129
4.6.6. Summary of Relationship Analyses 131
4.7. Discussion of English Ejectives 132
4.7.1. Place of Articulation Effects on Articulatory Patterns 133

4.7.2. Revisiting Pressure Generation in English Ejectives 134
4.7.3. The “Unclear” Category: A Window Into Gradience 135
4.8. Conclusion 136

Chapter 5 Comparing Amharic and British English Ejectives 138
5.1. Introduction 138
5.2. Comparative Analyses 139
5.2.1. Closure Duration 139
5.2.2. Relative Burst Intensity 140
5.2.3. Larynx Raising 141
5.2.4. Supralaryngeal Cavity Reduction 142
5.2.5. Hyperpharyngeal Constriction 143
5.2.6. Summary of Cross-Linguistic Comparison 143
5.3. Contrastive and Emerging Patterns in Ejective Production 143

Chapter 6 General Discussion 145
6.1 Summary of the Main Findings 146
6.2. Revisiting the Glottalic Model 147
6.2.1. A Gradient Model of Ejectiveness 148
6.2.2. Non-Canonical Phonological Ejectives? 149
6.3. English Ejectives as a Sound Change in Progress 150
6.4. Limitations and Future Directions 151
6.5. Conclusion 153

Deutsche Zusammenfassung 153

References 157

Appendices 165
Appendix A Amharic Word List 165
Appendix B Amharic Bilabial Ejectives 168
Appendix C Supplementary Tables for Chapter 3 170
Appendix D Pharynx Length and Larynx Movement Range in Amharic Speakers 171
Appendix E English Word List 175
Appendix F Validation of the Auditory Categorization Procedure for English Ejectives 177
Appendix G Supplementary Tables for Chapter 4 178

Acknowledgements 183

List of Figures

Figure 2.1 Spectrograms of audio recordings from Speaker S04. 22
Figure 2.2 Line plot showing the temporal intensity pattern of stop-following vowels. 27
Figure 2.3 MR image of the vocal tract configuration at stop closure. 30
Figure 2.4 Calculation of larynx displacement (in mm). 31
Figure 2.5 Comparison of the vocal tract configuration at the moment of closure offset for a velar ejective (left panel) and a velar pulmonic voiceless stop (right panel). 32
Figure 2.6 Example of gridline assignments for articulatory regions. 35
Figure 3.7 Raw closure duration values (y-axis, in ms) for ejective and pulmonic voiceless stops. 51
Figure 3.8 Raw burst duration values (y-axis, in ms) for ejective and pulmonic voiceless stops. 52
Figure 3.9 Raw VOT values (y-axis, in ms) for ejective and pulmonic voiceless stops. 55
Figure 3.10 Relative burst intensity values (y-axis, in dB) for ejective and pulmonic voiceless stops. 57
Figure 3.11 Relative burst intensity values (y-axis, in dB) for ejective and pulmonic voiceless stops. 59
Figure 3.12 Intensity values of the post-burst lag (y-axis, in dB) for ejective and pulmonic voiceless stops. 60
Figure 3.13 Raw intensity rise ratio values (y-axis, in dB) for ejective and pulmonic voiceless stops. 62
Figure 3.14 Delta larynx values (y-axis, in mm) of ejective and pulmonic voiceless stops. 68
Figure 3.15 Delta larynx values (y-axis, in mm) by stop type (ejective, pulmonic voiceless) and speaker 70
Figure 3.16 Delta cavity values (y-axis, in mm^2) for ejective and pulmonic voiceless stops. 71
Figure 3.17 Delta cavity values (y-axis, in mm^2) by stop type (ejective, pulmonic voiceless) and speaker. 73
Figure 3.18 Delta hyperph values (y-axis, normalized) for ejective and pulmonic voiceless stops. 74
Figure 3.19 Scatter plot showing the relationship between delta larynx (x-axis, in mm) and delta cavity (y-axis, in mm^2) for ejectives. 79
Figure 3.20 Scatter plot showing the relationship between delta hyperph (x-axis, normalized) and delta cavity (y-axis, in mm^2) for ejectives. 81
Figure 3.21 Scatter plot showing the relationship between delta larynx (x-axis, in mm) and absolute burst intensity (y-axis, in dB) for ejectives. 82
Figure 3.22 Scatter plot showing the relationship between delta cavity (x-axis, in mm^2) and absolute burst intensity (y-axis, in dB) for ejectives. 85
Figure 3.23 Scatter plot showing the relationship between delta hyperph (x-axis, normalized) and absolute burst intensity (y-axis, in dB) for ejectives. 86
Figure 4.24 Distribution of word-final stop realizations (ejective, pulmonic, glottal replacement, and unclear). 104
Figure 4.25 Number of ejective realizations identified per speaker across the dataset. 105
Figure 4.26 Distribution of stop type categories by PoA. 107
Figure 4.27 Raw closure duration values (y-axis, in ms) for four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear). 108

Figure 4.28 Raw relative burst intensity values (y-axis, in dB) for four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear). 110
Figure 4.29 Intensity values of the post-burst lag (y-axis, in dB) for four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear). 112
Figure 4.30 Delta larynx values (y-axis, in mm) for the four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear). 115
Figure 4.31 Delta larynx values (y-axis, in mm) for ejectives by speaker. 117
Figure 4.32 Delta cavity values (y-axis, in mm^2) for the four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear). 118
Figure 4.33 Delta hyperph values (y-axis, normalized) for the four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear). 121
Figure 4.34 Scatter plot showing the relationship between delta larynx (x-axis, in mm) and delta cavity (y-axis, in mm^2) for ejectives and unclear stops. 124
Figure 4.35 Scatter plot showing the relationship between delta hyperph (x-axis, normalized) and delta cavity (y-axis, in mm^2) for ejectives and unclear stops. 126
Figure 4.36 Scatter plot showing the relationship between delta larynx (x-axis, in mm) and absolutely burst intensity (y-axis, in dB) for ejectives and unclear stops. 128
Figure 4.37 Scatter plot showing the relationship between delta cavity (x-axis, in mm^2) and absolutely burst intensity (y-axis, in dB) for ejectives and unclear stops. 129
Figure 4.38 Scatter plot showing the relationship between delta hyperph (x-axis, normalized) and absolute burst intensity (y-axis, in dB) for ejectives and unclear stops. 130
Figure 5.39 Closure duration values (y-axis, in ms) for singleton, word-final ejective stops compared between two languages (Amharic, British English). 139
Figure 5.40 Relative burst intensity values (y-axis, in dB) for singleton, word-final ejective stops compared between two languages (Amharic, British English). 141
Figure 5.41 Delta larynx values (y-axis, in mm) for singleton, word-final ejective stops compared between two languages (Amharic, British English). 141
Figure 5.42 Delta cavity values (y-axis, in mm^2) for singleton, word-final ejective stops compared between two languages (Amharic, British English). 142
Figure 5.43 Delta hyperph values (y-axis) for singleton, word-final ejective stops compared between two languages (Amharic, British English). 142
Figure B.44 Raw closure duration values (y-axis, in ms) for Amharic bilabial ejective and pulmonic voiceless stops. 168
Figure B.45 Relative burst intensity values (y-axis, in dB) for Amharic bilabial ejective and pulmonic voiceless stops. 169
Figure B.46 Delta larynx values (y-axis, in mm) of Amharic bilabial ejective and pulmonic voiceless stops. 169
Figure D.47 Speaker means of *delta larynx* (mm) plotted against pharynx length (mm) for each of the 11 Amharic speakers (S01-S11). 172

List of Tables

Table 2.1 Average gridline assignment per region across speakers and language datasets. 36
Table 3.2 Distribution of tokens by length, stop type, PoA, and word position. 51
Table 3.3 Pairwise comparisons of burst duration (modelled on the log scale) by stop type and position. 54
Table 3.4 Pairwise comparisons of VOT (modelled on the log scale) by stop type and position. 56
Table 3.5 Pairwise comparisons of VOT (modelled on the log scale) by stop type and length. 56
Table 3.6 Pairwise comparisons of relative burst intensity (in dB) by stop type and PoA. 58
Table 3.7 Pairwise comparisons of relative burst intensity (in dB) by stop type and position. 58
Table 3.8 Pairwise comparisons of PBL intensity (in dB) by stop type and PoA. 61
Table 3.9 Pairwise comparisons of PBL intensity (in dB) by stop type and position. 61
Table 3.10 Pairwise comparisons of intensity rise ratio (modelled on the log scale) by stop type and position. 63
Table 3.11 Pairwise comparisons of delta larynx (modelled on the log scale) by stop type and PoA. 69
Table 3.12 Means (*M*) and standard deviations (*SD*, in parentheses) of delta larynx (in mm) values of ejectives for each speaker. 70
Table 3.13 Pairwise comparisons of delta cavity (in mm^2) by stop type and length. 72
Table 3.14 Means (*M*) and standard deviations (*SD*, in parentheses) of delta cavity (in mm^2) values of ejectives for each speaker. 73
Table 3.15 Pairwise comparisons of delta hyperph (normalized) by stop type and PoA. 75
Table 3.16 Pairwise comparisons of delta hyperph (normalized) by stop type and position. 75
Table 3.17 Estimated slopes for the relationship between delta hyperph and delta cavity by PoA and length. 81
Table 3.18 Pairwise comparisons of delta hyperph (normalized) slopes across PoA and length. 82
Table 3.19 Estimated slopes for the relationship between delta larynx (in mm) and absolute burst intensity (in dB) by PoA, position, and length. 83
Table 3.20 Pairwise comparisons of delta larynx (in mm) slopes across PoA, position, and length. 84
Table 3.21 Estimated slopes for the relationship between delta hyperph (normalized value) and absolute burst intensity (in dB) by position. 87
Table 3.22 Pairwise comparisons of delta hyperph (normalized) slopes across position. 87
Table 4.23 Odds ratios from logistic mixed-effects model predicting ejective realization. 104
Table 4.24 Proportion of ejective realizations (%) across structural contexts in the dataset, including PoA, voicing and phrasal accent. 105
Table 4.25 Pairwise comparisons of closure duration (modelled on the log scale) by stop type and PoA. 109
Table 4.26 Pairwise comparisons of relative burst intensity (in dB) by stop type and PoA. 111
Table 4.27 Pairwise comparisons of PBL intensity (in dB) by stop type and PoA. 113
Table 4.28 Estimated marginal means for delta larynx (model-adjusted predictions) by stop type and PoA. 116
Table 4.29 Mean (M) and standard deviation (SD) of delta larynx (in mm) for ejectives by speaker. 117
Table 4.30 Pairwise comparisons of delta cavity (mm^2) by stop type. 119
Table 4.31 Mean (M) and standard deviation (SD) of delta cavity (mm^2) for ejectives by speaker. 119
Table 4.32 Estimated slopes for the relationship between delta larynx (in mm) and delta cavity (in mm^2) by stop type. 124

Table 4.33 Pairwise comparisons of delta larynx (in mm) slopes by stop type. 125
Table 4.34 Estimated slopes for the relationship between delta hyperph (normalized) and delta cavity (in mm^2) by PoA. 126
Table 4.35 Pairwise comparisons of delta hyperph (normalized) slopes across PoA. 127
Table 4.36 Estimated slopes for the relationship between delta hyperph (normalized) and absolute burst intensity (in dB) by PoA. 131
Table 4.37 Pairwise comparisons of delta hyperph (normalized) slopes by PoA. 131
Table C.38 Means (M) and standard deviations (SD, in parentheses) of each acoustic measure by speaker. 170
Table C.39 Means (M) and standard deviations (SD, in parentheses) of each acoustic measure by PoA, word position, and length. 170
Table D.40 Speaker-wise pharynx length (in mm) and mean change in larynx position during ejective closure (*delta larynx*, in mm). 172
Table D.41 Leave-one-out sensitivity analysis for the effect of pharynx length on larynx displacement during ejective stop closure. 173
Table G.42 Pairwise comparisons of delta larynx (modelled on the log scale) by stop type. 178
Table G.43 Pairwise comparisons of delta larynx (modelled on the log scale) by PoA. 179
Table G.44 Pairwise comparisons of delta hyperph (normalized) by stop type. 180
Table G.45 Pairwise comparisons of delta hyperph (normalized) by PoA. 181

Chapter 1
General Introduction

Speech sounds are shaped by languages and the speakers that use them, as well as by the social contexts and speaking styles in which they occur. These factors give rise to multiple "flavors" of a sound – subtle variations that reflect its lived phonetic reality. At the same time, speech sounds are assigned formal definitions and organized into classes that aim to capture their most general characteristics. The International Phonetic Alphabet (IPA), for instance, places each sound within a fixed category based on idealized parameters. But what is lost in this process? What happens to the phonetic nuance – the flavor – when a sound is reduced to a universal categorical definition?

This dissertation poses that question in the context of ejectives, a class of consonants traditionally defined as glottalic egressive sounds. It brings together phonological and non-phonological perspectives by comparing ejective stops in Amharic, an ejective language spoken in Ethiopia, with ejective stops emerging in British English, in order to shed light on the relationship between systemic categorization and lived phonetic realization of these sounds.

Ejectives are attested in roughly 17% of the world's languages. Their geographical distribution is concentrated in the Americas, the Caucasus, and parts of eastern and southern Africa, according to the *World Atlas of Language Structures* (WALS, Maddieson, 2013). Within these regions, they appear across diverse language families. Yet, they are traditionally absent from the languages of Western Europe. The well-documented occurrence of sounds auditorily perceived as ejective in British English (Gordeeva & Scobbie, 2013; McCarthy & Stuart-Smith, 2013) and the curiosity about their emergence in a language originally based on a purely pulmonic phonological system have exposed underlying tensions between abstract categorization and phonetic reality surrounding ejective sounds.

This tension arises from the fact that, despite considerable acoustic research, the articulatory properties of ejectives remain comparatively underexplored. A wide range of acoustic studies has shown that ejectives are highly variable across and within languages, with no single pattern of acoustic features adequately capturing this diversity (Grawunder et al., 2010; Ham, 2007; Ingram & Rigsby, 1987; Warner, 1996; Wright et al., 2002). This variability is echoed in diverse descriptive accounts of how ejectives are produced (Brandt & Simpson, 2021; Demolin, 2002; Halle & Stevens, 1971; International Phonetic Association, 1999; Kingston, 1984; Ladefoged & Maddieson, 1996; Lindau, 1984; Wright et al., 2002).

Yet, only a small number of empirical studies have in fact directly investigated their articulatory implementation (Brandt & Simpson, 2021; Kingston, 1984, 1985). Together, these studies reveal how limited our understanding remains of the actual mechanisms involved in ejective production. This dissertation seeks to address that gap.

Textbook accounts, such as Catford's (1977), describe ejectives as being produced by raising the larynx with a closed glottis while an oral closure is held. This compresses the air trapped between the two closures, building up intraoral pressure that is released as a characteristically intense burst. This places the mechanical action of larynx raising at the core of the airstream mechanism, that is, glottalic initiation. The *IPA Handbook* (1999, pp. 9–10), on the other hand, attributes ejective production to "squeezing" the air confined between glottal and oral closures, notably omitting any reference to a specific articulator performing the action. This omission invites a broader, more holistic interpretation of cavity compression and its role in generating intraoral pressure during ejective production. In fact, findings by Kingston (1985) and Brandt and Simpson (2021) question the role of larynx raising, both its efficacy in building up adequately high pressure and even its necessity in the mechanism at all. This suggests that pressure increase may require assistance from additional articulations above the larynx or, in some cases, may rely on alternative strategies involving pressure generation via a pulmonic airstream.

These findings suggest that ejective mechanisms may vary depending on language and context. Yet, this variation also blurs the boundary between ejective and non-ejective categories. The progression of accounts described above away from the classic glottalic model raises important questions that challenge our current understanding of class distinctions. If a sound is produced without larynx raising, and thus is not glottalically initiated, what, if anything, still qualifies it as an ejective? This question recalls the broader tension introduced at the outset: while ejectives are defined as glottalic consonants, their real-world production may not always align with that definition.

In practice, delineating which phonetic realizations fall within the category of ejective is far from straightforward. One challenging distinction has been that between ejectives and glottalized consonants, which are sounds that involve full or partial closure of the glottis, yet not its vertical movement. It is not always clear how inclusive or restrictive the two categories should be.

Based on the data and interpretations presented in Maddieson et al. (2001), two main criteria emerge as relevant for distinguishing between ejectives and glottalized sounds: the airstream mechanism and the timing of glottal and oral closures.

In this interpretation, ejectives are initiated by a glottalic airstream and exhibit full overlap between glottal and oral closure. Glottalized sounds, by contrast, are at least partially pulmonic and typically involve only partial overlap. Yet, in the *WALS* chapter on glottalized consonants, Maddieson (2013) himself acknowledges the difficulty of drawing clear boundaries. He concedes that the chapter adopts a broader definition of ejectives than the classical model would permit by including sounds in the ejective category even when they appear to lack clear evidence of larynx raising. What ultimately motivates the decision to classify such sounds as ejectives, despite the absence of what is traditionally considered the defining articulatory feature, remains unclear. This categorization issue becomes especially relevant when we enter the non-phonological domain, as in the case of emerging ejectives in English.

Here, we currently call certain sounds ejectives based on their auditory impression. But without articulatory evidence, we cannot know what underlies this perceived ejectiveness. It seems unlikely that these sounds are produced by larynx raising, as there is no phonological foundation within the English sound system for speakers to engage the larynx in this way (Simpson, 2014). The non-glottalic alternative mechanism described in Brandt and Simpson (2021) that foregoes active larynx raising could create the articulatory conditions necessary for such sounds to emerge in an otherwise purely pulmonic language – and, indeed, help shed light on possible diachronic pathways of ejectives.

From a theoretical perspective, however, "non-glottalic ejectives" fail to meet the criteria for ejective classification, at least in the Catfordian sense. Yet, as we have already seen, it remains uncertain to what extent even phonological ejectives consistently adhere to the canonical glottalic mechanism. Precisely because English ejectives challenge these canonical articulatory assumptions and are not anchored in a phonologically contrastive identity, they offer a valuable opportunity to examine the articulatory basis of ejectiveness.

Building on Kingston (1984) and Brandt and Simpson's (2021) work, this dissertation broadens the perspective on ejective production by complementing acoustic insights of ejectives with a detailed investigation of their articulatory underpinnings. This dissertation brings together phonological and non-phonological dimensions, and expands the view on ejective articulation beyond the larynx alone, with the aim of uncovering how existing theoretical accounts are realized in the actual practice of ejective sound production. In doing so, it seeks to contribute to a more nuanced understanding of what qualifies as an ejective and help refine the boundaries of the category itself.

To this end, this dissertation investigates articulation and acoustics of ejective stops in both Amharic and British English, first independently and then through direct comparison. With high-resolution real-time magnetic resonance imaging (rt-MRI), this dissertation makes use of state-of-the-art imaging technologies on a comparably large data sample.

The remainder of this chapter outlines current research on ejectives, to provide a clearer picture of what we know about these sounds and to highlight the key questions that remain unresolved. It is organized into four sections. Section 1.1 examines the acoustic and perceptual properties of ejectives, with particular attention to their cross-linguistic variability. Section 1.2 focuses on articulatory dimensions, confronting theoretical models with empirical findings. Together, these two sections address the phonological aspects of ejectives. Section 1.3 shifts to non-phonological considerations, briefly reviewing what is known about ejective-like sounds that emerge outside of phonological systems. Finally, Section 1.4 introduces the present study, outlining its aims and the specific gaps it seeks to address through a cross-linguistic investigation. More detailed insight into Amharic and English ejectives is reserved for the language-specific chapters: Chapter 3 for Amharic and Chapter 4 for English.

1.1. Characterizing Ejectives: Acoustic, Perceptual, and Typological Perspectives

To situate the present study within the broader phonetic and phonological landscape of ejectives, it is useful to begin with an overview of their acoustic and perceptual properties, along with the typological considerations that have emerged from this body of work. The acoustic domain, in particular, has received the most extensive attention in the literature to date, and thus provides a solid foundation for the articulatory questions addressed in the later section.

1.1.1. Acoustic Diversity and the Limitations of the Binary Typology

Traditionally, ejectives have been categorized as either *stiff* or *slack* types based on cross-linguistic acoustic patterns (Kingston, 1984, 2005; Lindau, 1984). In this binary framework, stiff ejectives (found in languages like Navajo and Tigrinya) are characterized by long voice onset time (VOT), raised fundamental frequency (f0) in the following vowel, sharp intensity rise time yes, and a tense-to-modal voice quality at onset of the following vowel. In contrast, slack ejectives (found in languages such as Hausa and Quiché) show shorter VOT, lowered f0, slower intensity rise time, and are followed by creaky voice quality.

These labels were originally intended to reflect distinct articulatory configurations, particularly differences in vocal fold tension and medial compression, thought to underlie the two ejective types (Kingston, 1984, 2005; Lindau, 1984). Stiff ejectives were hypothesized to involve both high longitudinal tension and strong medial compression, resulting in a tight glottal closure, whereas slack ejectives were thought to rely primarily on medial compression with little or no longitudinal tension. However, as cross-linguistic research on the acoustic properties of ejectives has expanded, it has become clear that this binary classification is no longer adequate.

There are two main approaches to the acoustic study of ejectives. One compares ejectives to their pulmonic counterparts to identify distinguishing acoustic features; the other explores ejective acoustic patterns across languages to document variation and typologize their properties. Studies on a wide range of languages, including Gitksan (Ingram & Rigsby, 1987), Ingush (Warner, 1996), Lushootseed (Kye, 2021), Tsilhqut'in (Ham, 2007), Waima'a (Hajek & Stevens, 2005; Stevens & Hajek, 2004, 2008), and Witsuwit'en (Hargus, 2011; Wright et al., 2002), have demonstrated considerable phonetic variability, both across and within languages, that cannot be captured by a simple dichotomy.

For example, Wright et al. (2002, p. 69) found substantial inter-speaker variability, to the extent that no consistent acoustic profile emerged, leading them to conclude that there is no "average" alveolar ejective in Witsuwit'en. Warner (1996) and Ham (2007) observed that Ingush and Tsilhqut'in ejectives simultaneously displayed features associated with both ejective types. Warner (1996) further notes that no two languages in her cross-linguistic comparison patterned identically across key acoustic features. Evidence from Lushootseed (Kye, 2021) and Waima'a (an Austronesian language, Hajek & Stevens, 2005) showed ejectives that displayed a range of mixed features, occupying an intermediate acoustic space not captured by the binary subdivision.

Taken together, these studies converge on a shared conclusion: the acoustic realization of ejectives is best captured as a multidimensional continuum, shaped by inter-speaker variation as well as patterns conditioned by language and even place of articulation. The traditional binary typology is therefore insufficient for describing the rich acoustic diversity observed in ejective consonants.

This acoustic diversity has even prompted challenges to the physiological assumptions underlying the stiff/slack typology. Interpreting the acoustic patterns observed in their Witsuwit'en data, Wright et al. (2002, pp. 70-72) argue that the traditional assumptions underestimate the interplay of muscular control, pressure dynamics, and the precise timing required to synchronize glottal and oral actions during ejective production.

In response to their acoustic observations, Wright et al. (2002) propose a three-dimensional continuum, defined by larynx height, vocal fold tension, and medial compression, in which ejectives occupy variable positions within the articulatory space. As a preview of articulatory issues addressed in a later section of this chapter, this model is notable for outlining a multidimensional framework in which multiple interacting factors, of which larynx raising is only one, contribute to the observed variability in ejective production. With the acoustic characteristics and typological considerations in view, we proceed by briefly examine what is known about the perception of ejectives.

1.1.2. Acoustic Cue Integration in the Perception of Ejectives

Research on the perception of ejectives shows that listeners rely on a combination of acoustic cues rather than a single feature to distinguish them from pulmonic stops. In a forced-choice identification task with Witsuwit'en listeners, Wright et al. (2002) showed higher identification accuracy of ejectives when multiple acoustic cues were present (92%) than when the contrast relied primarily on intensity rise time (85%), with the measures of closure/VOT ratio and f0 perturbation emerging as the strongest predictors of accurate ejective perception. Cross-linguistic perception experiments by Percival (2023) provide additional evidence that listeners prioritize different acoustic cues depending on their native phonological systems. Post-burst silence emerged as the primary perceptual cue to ejectiveness, followed by burst type (ejective versus non-ejective) and vowel quality.

These studies underscore that no single cue reliably signals ejectiveness; rather, listeners dynamically integrate multiple acoustic dimensions, including the presence of post-burst silence, timing characteristics, voice quality and the nature of the burst itself, guided by the phonological structures of their native language.

Having reviewed the acoustic and perceptual properties of ejectives and painted a picture of their typological variability, we now turn to the central concern of this dissertation: the articulatory basis of ejective production. The next section shifts focus from surface-level acoustic patterns to the underlying articulatory and aerodynamic mechanisms. Drawing on cross-linguistic literature, it outlines what is currently known, and what remains uncertain, about the physiological processes involved in producing ejectives.

1.2. Theoretical and Empirical Perspectives on the Articulatory Mechanisms of Ejectives

To understand the articulatory mechanisms behind ejective production, it is helpful to begin with a broader view of how airflow is generated during speech. At the core of speech production is an airstream that moves through the vocal tract to generate sound. *Airstream mechanisms*, as described in Catford's (1977) model, refer to how this airflow is initiated and form a central basis for phonetic classification. This is reflected in the IPA, where consonants are grouped into pulmonic and non-pulmonic categories.

Most speech sounds are pulmonically initiated, that is, the lungs generate the airflow. However, some languages include sounds produced by other mechanisms. Ejectives, for example, are glottalic consonants in which the larynx, rather than the lungs, initiates the airflow (Catford, 1977). The larynx functions like a piston, raising to compress the air trapped in the intraoral cavity, thus causing a significant increase in intraoral pressure.

Catford's (1977) glottalic model remains one of the most widely cited and foundational accounts of ejective production. Empirical studies that test this model or investigate the details of ejective articulation more generally, remain comparatively rare. This is in part due to the methodological challenges involved in observing laryngeal movement, especially in combination with intraoral pressure dynamics. Because the larynx is obscured by soft tissue, it cannot be directly accessed, making it difficult to track its movement using many of the methods typically employed to study articulatory motion.

Most existing work on laryngeal movement has focused on its role in tone, pitch, and prosody, using tools such as cricothyrometry (Gandour & Maddieson, 1976), X-ray imaging (Andersen & Sonninen, 1960; Lindqvist et al., 1973; Wang & Kong, 2010), ultrasound (Hamlet, 1980), MRI (Honda et al., 1999; Neuschaefer-Rube et al., 1996), and electroglottography (Laukkanen et al., 1999). Studies have used videofluoroscopy or laryngoscopy to examine larynx position during pharyngeal or glottalic sounds (Esling & Moisik, 2011). Overall, however, laryngeal involvement in non-pulmonic sound production has received comparatively little attention.

Research targeting ejectives has relied mostly on indirect methods, such as electromagnetic articulography (EMA) with sensor coils placed externally on the skin just above the larynx (Bückins et al., 2018; Shosted & Rose, 2011). Bückins et al. (2018) use this method to track vertical movement of the larynx in one of the few instrumental studies on ejective production. While their results confirmed larynx raising in Georgian ejectives, the authors caution against overinterpreting the details of the movement patterns due to the indirect nature of the method.

Another indirect method is dual-channel electroglottography (DC-EGG), which infers vertical movement of the larynx from changes in the amplitude of the EGG signal measured across two vertically spaced electrode pairs (Brandt & Simpson, 2021; Simpson & Brandt, 2019). Only more recently has direct imaging been employed with Oh et al. (2024) using real-time MRI and automatic tracking to visualize vertical larynx dynamics in the production of non-pulmonic consonants.

Nevertheless, many observations of vertical larynx movement in ejectives remain impressionistic or anecdotal**,** with limited quantitative analysis (Grawunder et al., 2010; Maddieson, 1998). For instance, Demolin et al. (2021) mention visible larynx raising in video data from Iraqw. Similarly, Dehais-Underdown et al. (2023) note larynx raising along with larynx lowering and pharyngo-laryngeal expansion preceding glottal closure in their laryngoscopic and MRI data of ejective production in beatboxing, without further quantification.

Against this backdrop, and in light of the methodological advances now available, a broader empirical reassessment of the articulatory foundations of ejectives is both timely and necessary. It remains an open question to what extent established accounts of ejective production generalize across different ejective languages. Catford's observations, for example, were based largely on Caucasian languages. Compared to the extensive cross-linguistic literature on the acoustics of ejectives, articulatory and aerodynamic studies, particularly those adopting a cross-linguistic framework, remain limited, with Brandt and Simpson's (2021) study being the one notable exception. We now turn to the set of empirical studies that have investigated the articulatory and aerodynamic properties of ejectives.

1.2.1. Rethinking the Role of Larynx Raising

Kingston (1984) provides one of the earliest and most detailed articulatory investigations of ejective production to date. His study combined physiological data (via thyroumbrometer, a photoelectric technique for tracking larynx movement during speech) from three speakers and aerodynamic modeling. The focus was on ejectives in Tigrinya, a Semitic Ethiopian language closely related to Amharic. While larynx raising was attested in these ejectives, according to Kingston the movement was often slow, limited in magnitude and not synchronized with oral closure. Moreover, considerable inter-speaker variability in larynx movement was found, which Kingston partly attributes to speech style, with more pronounced larynx movement observed in hyperarticulated speech (Kingston, 1984, pp. 52–53).

Bückins et al. (2018) similarly observed variability in the magnitude of movement across speakers. However, they attributed this variability to anatomical differences. Crucially, Kingston's (1984) findings challenge the traditional view that larynx raising alone is responsible for the characteristic high pressure and intense release burst of ejectives. His aerodynamic modeling demonstrated that larynx raising alone cannot generate the pressure levels observed; additional articulatory mechanisms, such as tongue root retraction or contraction of the pharyngeal wall, must also contribute.

A similar conclusion is reached by Demolin (2002), whose deductions from his work on Amharic ejective fricatives mirrors Kingston's claims. Although Demolin does not directly quantify larynx movement, he infers from the pressure levels he measures, and from assumptions about cavity size, that larynx raising alone is insufficient to explain the observed pressure buildup. In line with Kingston (1984), he argues that additional mechanisms, including tongue root retraction and increased supraglottal constriction, must also be involved.

In fact, proposals for additional supralaryngeal involvement in ejective production align with Halle and Stevens (1971, p. 52), who describe ejectives as involving larynx raising *complemented by pharyngeal constriction* to further reduce supraglottal cavity volume. Yet Kingston (1985) even goes so far as to suggest that larynx movement may not be the active driver of pressure changes at all, but rather a passive consequence of other articulatory actions. This directly challenges Catford's (1977) glottalic model, which positions larynx raising at the center of pressure generation in ejective production.

The fact that larynx raising is closely linked to, and even works together with, tongue retraction to reduce pharyngeal space is not, in itself, a novel insight (Esling, 2005; Lindau, 1975; Moisik, 2013; Trigo, 1991). However, the potential interaction of laryngeal and supralaryngeal activity in the specific context of ejective production, particularly in contributing to intraoral pressure buildup, as put forward by Kingston (1984), has received little focused empirical attention to date.

Electropalatographic (EPG) data from Demolin (2002) showed that ejective alveolar fricatives, particularly geminates, exhibited longer and more extensive tongue contact, suggesting a tighter and more sustained constriction. Demolin argued that this may represent a supralaryngeal contribution to cavity compression and the elevation of pressure levels. Similar conclusions were made by Gordon and Applebaum (2006) for Turkish Kabardian, where palatographic data showed greater narrowing and larger contact areas for ejective palato-alveolar fricatives compared to their pulmonic counterparts.

Hermes et al. (2016) present what is, to date, the only study explicitly targeting supralaryngeal correlates of ejectivity. They used rt-MR images from one Tigrinyan speaker to visually compare the supralaryngeal cavity configuration of a geminate velar ejective and corresponding pulmonic stop at the point of maximum cavity constriction. Their qualitative observations support Kingston (1984) in that they showed that ejectives involve tongue root and dorsum retraction and pharyngeal wall advancement, which together reduce the pharyngeal cavity size, articulations absent in the pulmonic stops. They also observed that ejectives are produced with a slightly more posterior place of articulation and appeared to show greater resistance to coarticulation. However, the study provides no quantitative measurements, focuses solely on velar geminate stops, and does not analyze larynx movement, thus leaving open questions about how generalizable these supralaryngeal articulations are across other places of articulation or about how coordinated these additional gestures are with laryngeal movement in ejectives.

1.2.2. Timing and Intergestural Coordination in Ejective Production

In a speculative but significant remark, Kingston (1984, p. 17) proposes that variation in the timing of larynx movement could form the basis of a phonological contrast among ejectives, analogous to VOT in distinguishing aspirated and unaspirated pulmonic stops. Building on the question of timing, the study by Oh et al. (2024) used rt-MRI to track the coordination of laryngeal and oral gestures in Hausa phrase-initial and phrase-medial ejectives (/k', k^w', s'/) produced by three native speakers. Compared to their pulmonic counterparts, ejectives exhibited significantly greater vertical larynx raising. This movement was tightly synchronized with oral closure: larynx raising began just after the onset of oral constriction and peaked near the constriction target. While pulmonic stops showed looser timing between laryngeal and oral gestures, ejectives displayed a remarkably stable coordination pattern across speakers and prosodic contexts. This suggests that precise intergestural timing, not just the presence of larynx raising, is a defining feature of ejective production. Due to Hausa's limited ejective inventory, place-of-articulation effects could not be assessed.

1.2.3. Challenging the Glottalic Model: Pulmonic Ejectives?

By directly testing the assumptions behind Catford's (1977) theoretical model, Kingston's (1984, 1985) work has prompted a reassessment of the larynx's central role in ejective production. His findings challenge the idea that vertical larynx raising alone is sufficient, and have more recently led to speculations that the larynx may not even be essential for ejective

production at all. In his aerodynamic modeling, Kingston (1984) found that one successful configuration for explaining the high intraoral pressure in Tigrinya ejectives involved a slightly open glottis with a contribution from pulmonic airflow. This observation prompted Brandt and Simpson (2021) to further explore the possibility of a pulmonic mechanism for ejective production.

Brandt and Simpson (2021) pursued the idea of pulmonically fueled ejectives in a study comparing phonological ejectives in Georgian with so-called epiphenomenal ejectives in German. While German does not include ejectives in its phonemic inventory, glottalized stop releases can arise as a by-product of glottalized vowel-initial syllables overlapping with preceding word-final stops. These sounds resemble ejectives auditorily, but crucially, there is no reason to assume that such epiphenomenal speech events involve active larynx raising (Simpson, 2007).

To test this, Brandt and Simpson (2021) combined dual-channel EGG with intraoral pressure measurements, allowing them to track vertical larynx movement and glottal contact patterns. Their goal was to uncover the mechanisms underlying these epiphenomenal German ejectives and to explore whether Georgian phonological ejectives might share similar characteristics. Their dataset included 13 Georgian and 14 German speakers, producing ejectives across three places of articulation and three positions: sentence-initial, word-initial, and word-medial.

The results revealed no systematic larynx movement in the German data. In Georgian, larynx raising was observed only in sentence-initial ejectives; no vertical movement occurred in word-initial or medial positions. Interestingly, sentence-initial ejectives in Georgian consistently showed larynx lowering across all places of articulation. This finding echoes results from Bückins et al. (2018), the only other study to investigate larynx movement in Georgian ejectives.

Importantly, the presence of voicing into closure in non-initial Georgian ejectives, paired with the lack of larynx raising in these contexts, supported Brandt and Simpson's hypothesis of a pulmonic mechanism for ejective production. They propose a process in which pulmonic air enters the supraglottal cavity through a still-open glottis during the early part of the oral closure, thereby increasing intraoral pressure. Subsequent glottal closure seals the cavity, and the oral release produces an ejective-like burst, shaped by the resonance of a cavity sealed at the glottis. Crucially, this account explains how an ejective-like sound can be produced without requiring active larynx raising.

More recent evidence from Sulaberidze et al. (2023), based on real-time MRI, corroborates these position-dependent patterns of larynx movement in Georgian. Their findings confirm that larynx raising occurs only in sentence-initial ejectives, and even then, the upward movement is too small to significantly contribute to intraoral pressure, thus aligning with Kingston's (1984) aerodynamic modeling.

Together, these findings present a significant challenge to the canonical glottalic model described by Catford (1977), by demonstrating that active larynx raising may not be a necessary component of ejective production in all contexts, even for phonological ejectives. Moreover, by attributing pressure generation to a pulmonic airstream, this non-glottalic model raises the possibility that active compression of the cavity may not be necessary at all. In doing so, it moves beyond the account offered in the *IPA Handbook* (1999), which still assumes that supraglottal cavity compression – whether through larynx raising or other articulatory maneuvers – is responsible for pressure buildup.

1.2.4. Integrating Mechanisms: Moving Beyond Canonical Descriptions

All the accounts mentioned above share a common goal: explaining how high intraoral pressure is generated, indicating that this, at least, is widely understood to be a central aerodynamic feature of ejective stops across languages (Catford, 1977; Demolin, 2002, 2004; Demolin et al., 2021; Kingston, 1984, 1985; Ladefoged & Maddieson, 1996; Maddieson, 1984; Solé, 2023). The accounts differ, however, in the proposed articulatory strategies to achieve this intraoral pressure. Suggestions range from larynx raising and pharyngeal constriction, including articulations such as tongue root retraction (Catford, 1977; Kingston, 1984), to accounts that frame pressure generation more broadly in terms of overall supralaryngeal cavity compression, without attributing the action to any specific articulator (IPA Handbook, 1999), and even to models that forgo these mechanisms entirely in favor of a pulmonic airstream (Brandt & Simpson, 2021).

This diversity of proposals makes one thing clear: despite general agreement about the aerodynamic goal, we still do not have a complete, empirically well-supported picture of how ejectives are actually produced. Much of our current understanding is based on theoretical descriptions, impressionistic observations, and a limited number of instrumental studies, of only a small set of languages and speakers. Given the complexity of these sounds and the methodological challenges involved in simultaneously analyzing laryngeal and supralaryngeal movement, glottal state and measuring aerodynamic properties, it is perhaps unsurprising that many questions remain.

As demonstrated by Brandt and Simpson (2021), ultimately, to gain a more complete picture of ejective production, it is useful to broaden our perspective and consider cases where ejective sounds arise outside traditional phonological systems, where they may challenge or refine existing models of articulatory mechanisms.

1.3. Non-Phonological and Historical Perspectives

Most of the existing literature is centered on phonologically contrastive ejectives. Next to no attention has been paid to cases where ejective sounds arise outside the phonological system. The following section shifts focus to these cases, setting the stage for the upcoming investigation of English ejectives.

1.3.1. Non-Phonological Occurrences of Ejectives

It is difficult to assess how common non-phonological occurrences of ejectives are across the world's languages, that is, cases where ejective sounds emerge in languages without them being part of the regular phonological inventory. If the number of existing studies is any indication, such occurrences, even when including languages with pronunciation variants that are non-phonological, appear to be relatively rare. Most available research centers on English, which remains by far the most prominently discussed case. A detailed review of studies on English ejectives will be provided in Chapter 4.

Interestingly, the few reports that do exist on non-phonological ejectives point to a recurring structural pattern: these sounds consistently appear in word-final position. This tendency is evident across several unrelated languages. In English, ejectives occur exclusively at the ends of words, typically at phrase boundaries or before pauses, and more frequently in stressed syllables (Ogden, 2024, p. 163). In German, so-called epiphenomenal ejectives arise from gestural overlap when word-final fortis plosives precede glottalized, vowel-initial syllables (Simpson, 2007; Sulaberidze, 2022).

A similar pattern is reported for Eastern Armenian (Toparlak & Dolatian, 2023), a language with a three-way laryngeal contrast but no phonemic ejectives, where voiceless unaspirated stops may be realized as ejectives specifically in word-final position (Ladefoged & Maddieson, 1996). This consistent localization to final position is especially notable given that contrastive ejectives are often disfavored in this environment in languages where they are phonologized (Greenberg, 1970).

In English, ejectives also most frequently occur at the velar pace of articulation (Gordeeva & Scobbie, 2011; McCarthy & Stuart-Smith, 2013), a pattern that aligns with cross-linguistic trends identifying velar ejectives as the most widely attested (Greenberg, 1970; Javkin, 1977; Maddieson, 1984).

Taken together, these observations suggest that the combination of word-final position and velar place of articulation may be especially conducive to the emergence of ejective sounds. Viewed through the lens of sound change in progress, as suggested by their increasing frequency in English (Simpson, 2014), this convergence of factors takes on added significance. Investigating these environments is thus essential for understanding the conditions under which ejectives might arise and potentially stabilize. To explore this further, it is useful to turn to broader work on the historical development of ejectives and the constraints shaping their diachronic emergence.

1.3.2. Diachronic Pathways of Ejectives

When it comes to the diachronic development of ejectives, our understanding remains limited. Easterday and Bybee (2023) highlight how rare and specific the historical pathways are through which ejective consonants are known to emerge. This stands in contrast to the relatively widespread typological presence of ejectives (Disner & Maddieson, 1984; Maddieson, 2013; Moran & McCloy, 2019).

The most well-attested source of ejectives cross-linguistically is the fusion of a voiceless obstruent with a following glottal stop, typically through temporal overlap of oral and glottal articulations (Fallon, 2013). According to Easterday and Bybee (2023), a second pathway, relevant to the English case, is word-final glottalization (Gordeeva & Scobbie, 2013; McCarthy & Stuart-Smith, 2013). While this second source of ejectives appears to be more common than previously thought, it is not known to give rise to phonologically contrastive ejectives. Easterday and Bybee suggest that sound changes involving ejectives may have unusually long trajectories, with phonologization from purely phonetic sources being highly exceptional.

This perspective is consistent with theoretical models of sound change, such as Kiparsky's (2003), which predict that phonetic variants that do not conform to the structural patterns of a language's phonological system are unlikely to be fully integrated. In the case of English, a language whose phonological system is entirely pulmonic, this means glottalic sounds like ejectives are not expected to become contrastive phonemes.

The patterns observed in English call attention to these prevailing assumptions about the limits of sound change. They offer the unique opportunity to study how, when, and under what conditions glottalic initiation can take root in a phonological system where it has no historical precedent. A direct articulatory investigation, combined with a comparative analysis of phonological and emerging ejectives, offers a rare opportunity to observe this process in motion and to better understand the phonetic mechanisms that may underlie such change. With the current state of knowledge on ejectives established, both phonological and non-phonological, the following section introduces the present study.

1.4. The Present Study: A Cross-Linguistic Articulatory Perspective on Ejectives

Despite the centrality of Catford's (1977) glottalic egressive model to our theoretical understanding of ejectives, few studies have directly tested it experimentally, and even fewer have done so cross-linguistically. Notable exceptions, such as Kingston (1984) and Brandt and Simpson (2021), reveal how incomplete our knowledge remains of how ejectives are actually produced. These studies underscore the need for a broader perspective on these sounds, one that recognizes that articulatory strategies may vary across languages, contexts, and speakers. Brandt and Simpson's (2021) comparison of phonological ejectives in Georgian and epiphenomenal ones in German illustrates the value of placing phonological and non-phonological ejectives side by side to reveal alternative articulatory mechanisms and broaden our understanding of ejective production. The present study follows this model but shifts focus to phonological ejectives in Amharic and emerging ejectives in British English.

A broader account also requires moving beyond viewing the larynx in isolation. While supralaryngeal articulation seems likely to play an important role in ejective production, it remains understudied (Hermes et al., 2016), especially in combination with laryngeal movement. Moreover, these articulatory dynamics have not yet been investigated in the context of non-phonological ejective production. Additionally, variables such as place of articulation and word position have been shown to influence ejective realization (Brandt & Simpson, 2021; Kingston, 1984), suggesting further benefit from examining how these contexts additionally shape variation in ejective production within and across languages. Motivated by these open questions around cross-linguistic variability, supralaryngeal involvement, and the status of non-phonological ejectives, this dissertation undertakes a comparative articulatory investigation of Amharic and English ejectives.

Using rt-MRI to analyze laryngeal and supralaryngeal activity, complemented by acoustic analysis, this study aims to illuminate the articulatory strategies that can give rise to ejective sounds by comparing them in both phonological and non-phonological settings. In doing so, it provides an empirical foundation for understanding what the defining characteristics of the sound class are.

1.4.1. Acoustic Analysis

In light of the known cross-linguistic variation, it is essential for any comprehensive investigation into ejective production to first understand how the ejectives in question behave acoustically. In the studies comprising this dissertation, acoustic analysis will therefore precede articulatory investigation, ensuring that any observed articulatory patterns can be meaningfully contextualized.

Establishing key acoustic features serves several foundational purposes: it clarifies how the ejectives differ from their pulmonic counterparts, situates them within broader patterns of cross-linguistic variation, and provides crucial information for interpreting the underlying articulatory mechanisms. The selection of acoustic measures in this dissertation reflects these goals: closure duration, burst duration, VOT, relative burst intensity, intensity of the post-burst lag, and intensity rise ratio of the following vowel were chosen for their interpretative value in cross-linguistic comparison and their relevance to understanding articulatory dynamics.

1.4.2. Articulatory Analysis

The articulatory analyses in this dissertation build on prior instrumental studies (e.g., Brandt & Simpson, 2021; Kingston, 1984) by adopting a broader perspective on articulatory mechanisms. A key obstacle in previous research has been the difficulty of capturing vertical larynx movement. This study employs rt-MRI which enables precise observation of this gesture, allowing the study to assess the presence, magnitude, and variability of larynx raising across contexts. Crucially, rt-MRI provides a full-vocal-tract view that permits detailed analysis of supralaryngeal articulation, including overall intraoral cavity compression as well as more localized pharyngeal constriction, which may arise through tongue root retraction or stiffening of the vocal tract walls. These articulatory components are examined both individually and in coordination to understand how they contribute to cavity compression.

In the absence of direct intraoral pressure measurements, burst intensity is used as an indirect indicator of pressure magnitude. This approach is applied to both phonological ejectives (Amharic) and non-phonological, emerging ejectives (English).

The cross-linguistic comparison enables evaluation of which features are consistent across systems and which are shaped by functional (in this case, phonological contrast vs. non-phonological) or contextual factors. Emergent and phonological systems of ejectives are brought together, allowing us to revisit foundational questions: What defines ejectives?

Since only ejective *plosives* are attested in English, therefore the Amharic analysis is likewise limited to the plosive series and does not include fricatives or affricates. Finally, while this dissertation critically examines whether the label "ejective" is appropriate for these emerging sounds in English, the term English "ejectives" will be used throughout for simplicity, until otherwise noted.

1.4.3. Structure of the Dissertation and Overview of Chapters

The remainder of this dissertation is structured as follows. Chapter 2 outlines the methods of rt-MRI imaging and acoustic data collection and analysis, as used for both language studies. Chapter 3 presents the Amharic study, beginning with a brief overview of the language and prior work on Amharic ejectives. It then analyzes the acoustic properties of Amharic ejectives to provide a comparative reference point within the broader cross-linguistic patterning of ejective acoustics. This is followed by the articulatory investigation, which examines larynx raising, supralaryngeal activity, and their coordination in phonological ejective production as compared to pulmonic stop production. Chapter 4 turns to the analysis of emerging ejectives in British English, starting with a comprehensive review of what is currently known about these sounds. An auditory analysis identifies relevant realizations in the dataset. This is followed by an analysis of the acoustic characteristics to evaluate how they differ from their pulmonic counterparts and whether they exhibit acoustic properties commonly associated with ejective stops. Finally, examination of their laryngeal and supralaryngeal behavior will uncover their articulatory characteristics in an analysis that mirrors the structure of Chapter 3. Chapter 5 brings the findings together in a comparative analysis, directly contrasting the acoustic and articulatory features observed in the two languages. Only through such direct comparison can we begin to disentangle which aspects of ejective production are phonological, which are not, and where the essence of "ejectiveness" truly lies. Chapter 6 discusses the broader implications of the findings for our understanding of ejectives, their typology, and their possible diachronic pathways, ultimately aiming to refine how we define this sound class.

Chapter 2 Methods

This chapter outlines the overarching methodological framework used in this dissertation. It presents the procedures applied consistently across both the Amharic and British English studies, including the experimental setup for rt-MRI data collection, as well as the general approach to image processing and acoustic analysis. Given the central role of rt-MRI in this work, and the fact that both studies were conducted using the same imaging setup, a unified presentation of methods is provided here to avoid unnecessary repetition across chapters. Study-specific methodological details, such as participant numbers and target word selection, are addressed in the respective language chapters (Chapter 3 for the Amharic study and Chapter 4 for the English study).

While the core methodology remained consistent across both studies, two important differences must be noted at the outset. The first stems from the time gap between data collection efforts: data for the Amharic study was collected approximately five years after that for the English study. Although all data was recorded at the same site using the same scanner and research team, several improvements to the post-processing pipeline were introduced in the interim. These technical differences are carefully documented in the relevant sections of this chapter.

The second, and more substantive, difference results from the differing nature of the ejectives in the two languages. Since ejectives in Amharic are phonologically contrastive and occur in predictable environments, the experimental design of the Amharic study, developed specifically for this thesis, could be tailored to systematically probe their production across different contexts. In contrast, ejectives in British English are non-phonological: they occur exclusively in word-final position, but even there, their production is unpredictable, which considerably complicates elicitation. This issue was circumvented by using an existing dataset, originally collected for an independent study on nasalization in British and American English. The decision to examine ejectives in this dataset emerged only after their presence was incidentally observed scattered across the recordings. However, since ejectives were not the original focus of the study, some minor design limitations, particularly in the segmentation scheme, which was not optimized for ejective analysis, had to be accommodated. These limitations had downstream implications for specific acoustic measurements, discussed in the following sections.

The remainder of this chapter provides a comprehensive overview of the data acquisition process, including the rt-MRI imaging procedure and audio recording setup. It also describes the data preparation and processing steps applied to both the acoustic and imaging data, followed by an outline of the analytical methods used in both domains. The chapter begins with a brief overview of rt-MRI and its application in phonetic research.

2.1. Use of Real-Time MRI for Articulatory Analysis

Rt-MRI offers dynamic imaging of the entire vocal tract, making it a powerful tool for articulatory analysis in phonetics. Unlike techniques such as EMA or EPG, rt-MRI provides a comprehensive midsagittal view that simultaneously captures lingual, pharyngeal, and laryngeal structures with relatively high temporal and spatial resolution.

These advantages have made rt-MRI especially valuable for examining articulators that are otherwise difficult to access, including the velum (Byrd et al., 2009; Carignan et al., 2019; Martins et al., 2012; Proctor et al., 2013), the pharynx (Carignan et al., 2015; Shosted et al., 2012; Tiede, 1996), broader laryngeal configurations (Ahmad et al., 2009; Moisik et al., 2015; Moisik et al., 2019), as well as the larynx itself (Demolin et al., 2002; Honda & Tiede, 1998).

In the present case, its ability to image the entire vocal tract, makes rt-MRI particularly well suited for analyzing both laryngeal and supralaryngeal activity during complex speech events such as ejective production, and does so without requiring invasive instrumentation (Lammert et al., 2010; Narayanan et al., 2004; Niebergall et al., 2013; Ramanarayanan et al., 2018).

2.2. Imaging

Rt-MRI data for the studies presented in this thesis was acquired using a 3-Tesla MRI system (Magnetom Prisma Fit, Siemens Healthineers, Erlangen, Germany) at the Max Planck Institute (MPI) for Multidisciplinary Sciences in Göttingen, Germany. Measurements were conducted with a 64-channel head coil secured above the participant's head as they lay supine in the scanner.

The imaging protocol follows established procedures used in similar rt-MRI studies (e.g., Carignan et al., 2020; Cunha et al., 2024; Kunay et al., 2022). It is based on a radiofrequency (RF)-spoiled FLASH sequence, which utilizes highly under-sampled radial gradient echo acquisitions. Regularized non-linear inversion was used for serial image reconstruction (Uecker et al., 2010).

A single set of nine spokes was used to acquire individual images at a repetition time (RT) of 2.22 milliseconds (ms). This results in a reconstructed frame rate of 19.98 ms or 50.05 frames per second (fps). The in-plane pixel size was 1.41 × 1.41 millimeters (mm) and the slice thickness was 8 mm. This returned images of 136 × 136 voxels, which are three-dimensional (3D) volume elements, in a field of view of 192 × 192 mm.

The interval of interest for analyzing the rt-MR images in the studies was the stop closure. The average duration of the stop closure was 104.2 ms in the Amharic dataset and 73 ms in the English dataset. Dividing each interval's duration by the temporal resolution of 19.98 ms yields the number of images analyzed. This means that, on average, approximately 5.2 images were analyzed per stop token in the Amharic study and 3.7 images in the English study.

2.3. General Data Collection Procedure

Before each recording session, the target sentences were randomized and organized into blocks of 13 to 15 sentences, depending on the experiment (English or Amharic). In the Amharic experiment, sentences were presented in Amharic script. Each block was shown as a separate slideshow, with each slide displaying a single sentence and transitioning every 4 seconds (s), resulting in a total block duration of 52-60 s. The Amharic experiment consisted of 21 blocks (excluding an initial practice block of four dummy sentences), while the English experiment comprised 23 blocks, including materials not relevant to the present study. Dummy sentences for English were added at the beginning of each block. During the experiment, slides were displayed on an external screen, and a mirror positioned above the head coil allowed participants to read the sentences while lying supine inside the scanner.[1]

Participants were first screened for MRI compatibility and given time to review and complete informed consent forms from Ludwig-Maximilians-Universität München (LMU) and the MPI. A total of 11 Ethiopian participants, native Amharic speakers, and 27 native speakers of British English were recorded for the respective studies (further details on participants can be found in Chapter 3, Section 3.2.1 for Amharic and Chapter 4, Section 4.2.1 for English).

[1] Gravitational effects on speech in the supine position cannot be entirely ruled out (Kitamura et al., 2005; Stone et al., 2007). However, running speech appears to reduce such effects (Tiede et al., 2000). Also, since both stop types were recorded under identical conditions, any gravitational effects would presumably affect them equally. Ultimately, the benefits of rt-MRI for addressing these research questions clearly outweigh any such concerns.

The participants were then encouraged to familiarize themselves with the speech materials before receiving instructions on their task, emphasizing the importance of speaking naturally, e.g., using a normal speech volume and speed. They had the opportunity to ask questions about the experiment or the scanner before being positioned inside the MRI scanner for testing. Including preparation and scanning, the total procedure lasted about 1.5 hours per participant for the Amharic experiment and approximately 2 hours for the English experiment.

Each speaker typically produced the target items only once to avoid Type I error inflation (Winter, 2015), except in cases of clear reading or pronunciation errors, or technical issues. In such cases, they were given the opportunity to repeat the entire sentence block.

2.4. Processing of Acoustic Signal

Synchronized, noise-suppressed audio was recorded alongside image data during scanning using an Optoacoustics FOMRI III fiber-optic dual-channel microphone (Optoacoustics Ltd.), positioned near the lips once at the start of the session. To further reduce scanner noise, the audio data was processed in MATLAB (The MathWorks Inc., 2022a) a linear predictive coding (LPC) analysis (order 64) applied to a 2 s window to estimate the noise spectrum. The resulting LPC coefficients were used to design an inverse filter, tailored to the specific noise profile, which was then applied to remove scanner noise.

Figure 2.1 presents the spectrograms of the same audio sample before (top panel) and after (bottom panel) noise suppression, clearly illustrating the positive impact of the process. The bottom panel demonstrates a significant reduction in background noise, resulting in a more defined representation of the speech signal. While some scanner noise, such as knocking artifacts, remains visible as regular high-intensity slashes, the overall clarity and intelligibility of the speech are markedly improved.

Though this LPC-based approach is rather conservative, it was nevertheless preferred over more aggressive filtering methods, since it effectively balances noise reduction while maintaining key acoustic features.

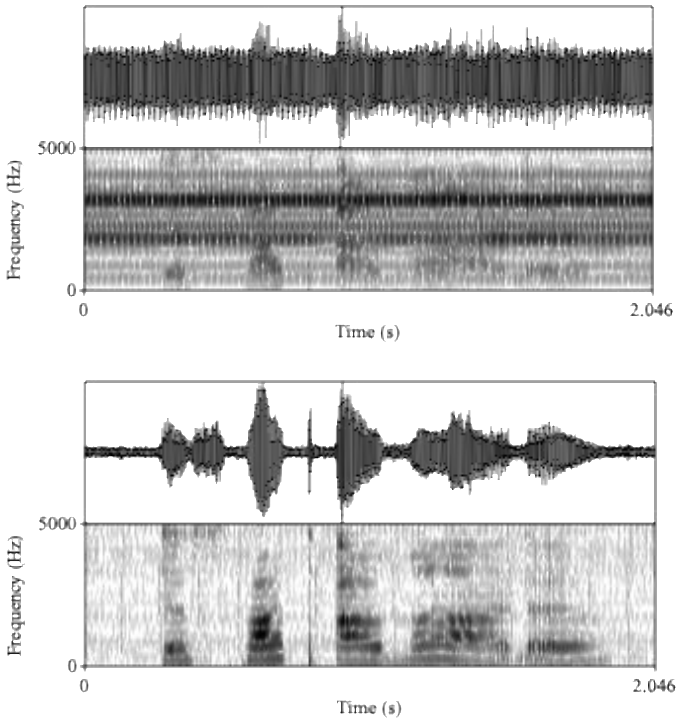

Figure 2.1 Spectrograms of audio recordings from Speaker S04 producing the sentence "Desta t'amä ale," demonstrating the effects of noise suppression. The top panel shows the original recording, while the bottom panel displays the filtered signal after applying LPC-based noise suppression. Although some residual scanner noise artifacts, such as knocking sounds, are still present in the filtered signal, the overall clarity and intelligibility of the speech have significantly improved.

2.5. Segmentation

Prior to analysis, the data was manually segmented through inspection of the acoustic waveform and corresponding broadband spectrogram in PRAAT (Boersma & Weenink, 2025, version 6.4.35). This segmentation provided the foundation for both the acoustic and articulatory analyses. For the former, the intervals of interest included the closure, burst, post-burst lag (PBL), and the vowel following the target stop. For the latter, the closure interval of the target stop served as the primary segment for analysis. Additionally, in the English study, the noise-suppressed audio recorded during MRI scanning was used for auditory stop categorization, necessary to first identify the ejective realizations in the dataset.

The segmentation scheme differed between languages, and some adjustments were necessary for the English data, as the original materials were not specifically designed for ejective analysis. For this reason, the segmentation procedures for each language are described separately below.

2.5.1. Amharic Segmentation Scheme

In segmenting the Amharic data, intervals were defined as follows: the closure of the target stop was segmented from the offset of the preceding vowel, indicated by the cessation of periodicity in the waveform, to the stop's release, marked by a sudden rise in amplitude. This point also served as the onset of the burst segment, which ended at the subsequent sharp drop in amplitude. The burst was further identified in the spectrogram as a band of concentrated energy, clearly distinguishable from the more diffuse energy of any following frication or aspiration. Stops lacking the visible indicators of a release burst described above were treated as unreleased or inaudibly released. These tokens were marked during segmentation and excluded from any analysis involving the burst or its temporal boundaries, such as VOT or post-burst lag intensity.

Immediately following the burst was the post-burst lag (PBL), segmented from burst offset to the onset of voicing in the following vowel. This onset was identified either as the return of regular periodicity in the waveform or as the moment the second formant reached a steady state. Any residual frication (or aspiration, in the case of pulmonic voiceless stops) flanking the burst was included in the PBL interval. The post-burst lag differs from VOT in that VOT includes the burst itself, whereas the post-burst lag measures only the interval following the burst.

Segmentation of the PBL was relatively straightforward for Amharic pulmonic stops, and for ejectives where the glottal release aligned with vowel onset. However, in some ejectives, the glottal release occurred only shortly after the oral release, sometimes visible as a brief, independent burst (like a glottal stop), or as creak that gradually merged into the following vowel. In such cases, it was difficult to distinguish creak resulting from the glottal release from that of a creaky vowel onset. To maintain consistency, any creak in the signal was attributed to the following vowel and excluded from the PBL segment. A similar approach is taken in Derib Ado (2011), though there, the voicing lag is defined from burst onset to vowel onset.

2.5.2. English Segmentation Scheme

In the English dataset, segmentation was adapted to account for the coda position of the target stops and the structure of the stimuli. Because the segmentation of the English dataset was carried out by a different researcher, only the core aspects of the procedure, relevant for the

present analysis, are reported here. The target stop, always a coda plosive, its closure, and the preceding vowel were manually segmented.[2]

For simple codas, closure was segmented from the offset of the preceding vowel to the release of the coda stop. In complex codas, a separate segment marked the first consonant, which was always a nasal. The offset of the nasal was defined as the onset of the following stop's closure. Neither the stop's release burst nor the immediately following lag were segmented individually. However, boundaries were placed to mark the interval between onset of the coda stop's release and the onset of the following word, which always began with the vowel /ʌ/.

2.6. Acoustic Analysis

An acoustic analysis was first conducted to complement the articulatory investigation of ejective and pulmonic stops. This was intended to situate the ejectives examined in this dissertation within the broader cross-linguistic landscape.

As has been established in Chapter 1, the available evidence from a range of acoustic studies suggests that ejectives are too variable to cluster into discrete types (Gordon & Applebaum, 2006; Grawunder et al., 2010; Ham, 2007; Lindau, 1984; Vicenik, 2010; Warner, 1996; Wright et al., 2002; Wysocki, 2004). Thinking in the traditional binary terms (stiff/slack) fails to reflect the gradient nature of ejective production. Accordingly, this dissertation did not attempt to typologize ejectives as stiff or slack. The acoustic measures examined here were intended to provide meaningful cross-linguistic context, both across the two languages studied and across place of articulation (PoA), word position, and consonant length (henceforth simply *length*; i.e., singleton vs. geminate) within the language, and lay the groundwork for subsequent articulatory interpretation. The focus of this dissertation is to uncover the articulatory mechanisms underlying ejective realization, with the understanding that acoustic characteristics emerge through a constellation of interacting factors that remain to be fully identified (Wright et al., 2002). This dissertation helps pave the way in this direction.

Due to the nature of the audio signal recorded during the rt-MRI sessions, it was ultimately decided that factors such as scanner noise and limited bandwidth, despite filtering, compromised the reliability of f0 and voice quality measurements (Kuortti et al., 2018; Lingala et al., 2017; Ramanarayanan et al., 2018). For this reason, while the relevance of measures

[2] Throughout this study, the term *stop* refers exclusively to oral stops. To maintain this distinction, nasal stops are consistently referred to as *nasals*.

involving voice quality and f0 in the analysis of ejectives is generally recognized, for the present dataset they were considered too delicate.

The acoustic studies in this dissertation therefore focused on more robust durational and intensity-based measures to support the articulatory analysis and characterize ejective production acoustically. The following sections describe how each of the six selected features – basic durational properties such as closure duration, burst duration, and VOT, alongside relative burst intensity, the intensity of the post-burst silent gap and the intensity dynamics of the following vowel – was extracted and analyzed using PRAAT.

2.6.1. Closure Duration

Closure duration, while not considered a hallmark feature of ejectives itself (Derib Ado, 2011; Grawunder et al., 2010; Noguchi, 2015; Seid et al., 2009; Stevens & Hajek, 2004; Vicenik, 2010; Warner, 1996), served as an important temporal anchor in this study. This is the interval during which the key articulatory activities (e.g., larynx raising, pharyngeal constriction) were analyzed. The rationale was that, to contribute to the buildup of intraoral pressure, these articulatory activities must occur while both oral and glottal closures are maintained, effectively sealing off the intraoral cavity. For this reason, closure duration was included as a temporal scale against which the timing of pressure-generating gestures can be interpreted.

In the Amharic study, closure duration (in ms) was calculated from the offset of the preceding vowel (V1) to the release of the target consonant. Occasionally, speakers inserted prosodic pauses before the target word, which disrupted the temporal alignment between V1 offset and stop closure onset when the target stop was word-initial. In such cases, the closure onset could not be reliably determined. These tokens were marked during segmentation and excluded from closure duration analysis. In the English study, closure duration was calculated as the difference between coda offset and preceding vowel offset for simple codas and between coda offset and offset of the preceding nasal for complex codas.

2.6.2. Burst Duration

The inclusion of burst duration in the Amharic analysis was motivated by prior findings suggesting that it may be a distinctive cue in the language (Derib Ado, 2021; Seid et al., 2009). In the Amharic study, burst duration (in ms) was calculated from onset to offset of the burst. The burst itself was identified by high-density energy in the spectrogram, which set it apart from any subsequent trailing frication or aspiration in the case of pulmonic voiceless stops. In

the English study, the burst itself was not segmented separately for duration measurement. As a result, burst duration was not included in the acoustic analysis in this dataset.

2.6.3. Voice Onset Time

Voice onset time (VOT) is one of the most commonly used acoustic measures to distinguish between stop types (Lisker & Abramson, 1964). Consistent with the broader conclusions of variability in the acoustic literature on ejectives, VOT in ejectives has been shown to be highly variable both within and across languages (Cho & Ladefoged, 1999; Derib Ado, 2011; Gordon & Applebaum, 2006; Grawunder et al., 2010; Ham, 2007; Hogan, 1976; Kye, 2021; Negesse, 2023; Park & Kim, 2015; Seid et al., 2009; Stevens & Hajek, 2004; Vicenik, 2010; Warner, 1996; Wright et al., 2002; Wysocki, 2004) underscoring that absolute VOT is not a reliable distinguishing feature of ejectives. Rather, its interpretative value and its relevance to this study lay in comparative analyses across languages.

VOT (in ms) was calculated as the difference between the burst onset time and the onset of the following vowel (V2). Creak was considered part of the following vowel, meaning VOT was measured right up to the moment that creaky voice appeared in the acoustic signal. Occasionally, speakers inserted a prosodic break after the target word. When this occurred after a word-final stop, VOT was not analyzed to avoid capturing unnaturally long durations. VOT is included in the Amharic analysis due to its interpretive value for cross-linguistic comparisons of ejective timing.

For English, all target stops were word-final by default, and such prosodic breaks were common, particularly when the target word was in the accented condition. To avoid interfering with measurement accuracy, VOT was excluded from the English acoustic analysis altogether.

2.6.4. Intensity Rise Ratio of the Following Vowel

Across ejective languages, the rate of intensity increase at vowel onset, typically referred to as the *intensity rise time*, is one of the most commonly investigated acoustic features (Derib Ado, 2011; Hajek & Stevens, 2005; Ham, 2007; Kingston, 1984; Kye, 2021; Negesse, 2023; Warner, 1996; Wright et al., 2002). Although prior work shows that this measure does not reliably distinguish ejective types across all languages, it remains useful for capturing the characteristic phonatory influence that ejectives exert on the following vowel relative to pulmonic stops.

In this study, intensity rise is defined as the difference (in dB) between (i) average intensity over a 30 ms window centered at the vowel midpoint and (ii) the intensity measured 30 ms after

vowel onset (cf. Wright et al., 2002). To avoid terminological ambiguity, since *rise time* suggests a temporal measure, this dissertation refers to the acoustic quantity as *intensity rise ratio*. The term denotes the same underlying phenomenon described in earlier work. When discussing previous studies, the original terminology is retained to reflect the authors' usage.

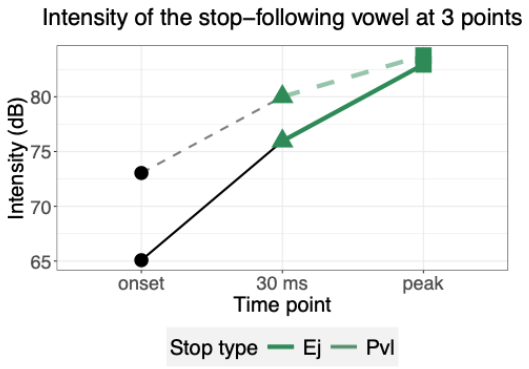

Figure 2.2 Line plot showing the temporal intensity pattern of stop-following vowels, based on mean intensity measurements (in dB) taken at three distinct time points: vowel onset (black circle), 30 ms into the vowel (green triangle), and the vowel's peak intensity (green square). The dashed line represents a pulmonic stop (Pvl), while the solid line represents an ejective stop (Ej). Intensity rise ratio is defined as the difference between the mean intensity in a 30 ms window centered on the vowel's peak and the intensity at 30 ms, illustrated by the green segment of the connecting lines. A flatter dashed green line (Pvl) reflects a smaller rise ratio, indicating relatively little additional increase in intensity after 30 ms. In contrast, the steeper solid green line (Ej) reflects a larger rise ratio, showing that more of the intensity buildup occurs after 30 ms.

Figure 2.2 presents a line plot illustrating the intensity progression of the stop-following vowel over three distinct time points: vowel onset (black circle), 30 ms into the vowel (green triangle), and the vowel's peak intensity (green square). The green segment of the line represents what is captured by the measure of intensity rise ratio. A lower rise ratio, as represented in **Figure 2.2** by the flatter dashed green line for the pulmonic voiceless stop, suggests that the vowel approaches its peak by ~30 ms. A higher rise ratio, as illustrated by the steeper solid green line for the ejective stop in **Figure 2.2**, suggest that the vowel continues to build toward its peak after 30 ms.

2.6.5. Relative and Absolute Burst Intensity

Ejectives are typically described in terms that reference their distinctively forceful burst (Catford, 1977; Ladefoged & Maddieson, 1996). At the same time ejectives reportedly demonstrate considerable variation in burst loudness across languages and contexts (Derib Ado, 2011; Grawunder et al., 2010; Noguchi, 2015; Park & Kim, 2015; Percival, 2019; Vicenik, 2010; Warner, 1996). These facts make burst intensity an important acoustic feature to include

in an analysis of ejectives, both as a feature separating them from their pulmonic counterparts, but also to compare ejectives of different languages.

Relative burst intensity. To facilitate comparisons across speakers and to reduce the influence of recording-level differences, this dissertation analyzed burst intensity using a relative measure (cf. Derib Ado, 2011; Negesse, 2023; Vicenik, 2010). *Relative burst intensity* (in dB) was calculated as max(vowel) - max (burst), with intensity maxima taken over the segmented stop-following vowel interval and burst interval. For the English data, where the release burst was not segmented as a separate interval, burst intensity was approximated over the first 20 ms following the labelled release, and the maximum intensity was taken from this window. Because the measure is vowel minus burst, lower values indicate a stronger burst peak relative to the following vowel.

Absolute burst intensity. To relate acoustic output to articulatory measures relevant to intraoral pressure generation, *absolute burst intensity* (in dB) was additionally computed as the mean intensity over the burst interval (Amharic) or over the first 20 ms following the labeled release (English; cf. Grawunder et al., 2010). Absolute intensity was used as an indirect proxy for pressure at release. However, this interpretation should be treated with some caution, as the inference of pressure is indirect and absolute intensity may reflect recording conditions alongside genuine production differences.

2.6.6. Post-Burst Lag (PBL) Intensity

This measure is widely regarded as an important acoustic cue to ejectives (Demolin, 2004; Gordon & Applebaum, 2006; Grawunder et al., 2010; Ladefoged & Maddieson, 1996; Vicenik, 2010; Warner, 1996; Wright et al., 2002; Wysocki, 2004). It also provides a valuable acoustic window into articulation, as it reflects the strongly reduced or absent airflow caused by a delay in glottal release relative to oral release, an articulatory pattern typically associated with ejective production. This measure can indirectly confirm whether the glottis remains closed at the moment of oral release, making it a crucial feature, particularly when evaluating English ejectives.

In Amharic, post-burst lag (PBL) intensity (in dB) was measured as the mean intensity over the PBL interval. In this study, PBL is defined as the interval between burst offset and the onset of the following vowel. This measure captures the characteristic interval of relative acoustic silence known for ejectives, which results from the delayed glottal release relative to the oral release.

As with VOT, PBL was not analyzed in cases where a prosodic break followed a word-final stop, as this could lead to artificially long intervals. Tokens where vowel onset immediately followed the burst were also excluded. In English, due to the absence of a separate boundary marking the offset of burst release, PBL intensity was calculated as the mean intensity from 40 ms after closure offset to the onset of the following word.

The acoustic measures described above provide a comparative reference point for evaluating the ejectives analyzed in this dissertation, both in terms of whether they exhibit acoustic properties commonly associated with ejective stops, and in how they fit within the broader cross-linguistic patterning of ejective acoustics. Having outlined the acoustic measures, we now turn to the articulatory components of this investigation into ejective production, which draws on rt-MRI data.

2.7. Image Analysis

Rt-MRI allows for direct visualization of vocal tract dynamics and offers a unique opportunity to investigate how specific articulatory strategies, such as larynx raising and supralaryngeal activity, contribute to ejective production. This section begins with an overview of the image processing procedures used to prepare the data for articulatory measurement.

All rt-MR image processing was performed in MATLAB (see also Carignan et al., 2021; Carignan et al., 2020; Kunay et al., 2022). The first step was image registration, which involved aligning each frame to a reference image selected individually for each speaker's dataset. This was achieved by first defining a region of interest (ROI) across the upper section of the head in this reference image, ensuring that only non-speech-related structures were enclosed.

The MATLAB function *imregtform* was then applied to this ROI to compute the rigid transformation, including translation and rotation, between each frame and the reference image. For every speaker, the reference image was chosen from a comparable phonetic context, ideally presenting the vocal tract in a neutral position. Also, special care was taken to avoid selecting an image where the head was in an atypical posture. Ultimately, in the Amharic study, frames from the low-mid vowel /ɜ/ were used as reference images, while in the English study, frames from /ə/, a mid-vowel with prosodic focus, were selected for this purpose.

Registration was essential for compensating minor head movements during recording. However, this method is effective only if the head movement remains within the original mid-sagittal plane. It cannot correct for major positional shifts that may happen throughout the recording process. To verify this was not the case, the correlation between each registered frame

and the reference frame was inspected for each speaker. Across datasets and speakers, correlation values were found to be typically well above 0.9, indicating successful alignment. With image registration complete, the following sections proceed to outline the steps involved in extracting the articulatory measures.

2.7.1. Tracing Vertical Larynx Movement

Ejectives are traditionally defined as glottalic egressive consonants, in which the airstream is generated by upward movement of the larynx, making larynx raising the defining initiatory gesture of the ejective mechanism (Catford, 1977). However, recent evidence from Brandt and Simpson (2021) suggests that larynx raising is not universally present in ejectives and that its implementation may vary depending on contextual factors such as prosodic position. In light of this, the first goal of the articulatory analysis in this dissertation was to determine whether, and to what extent, larynx raising occurs. The following section outlines how vertical larynx displacement was computed from the rt-MRI data.

For this, a semi-manual approach was employed. First, a ROI was placed around the larynx. Within this ROI, a slender vertical rectangle was manually selected, approximately bisecting the center of the larynx, the larynx midline, as indicated by the orange arrow in the left panel of **Figure 2.3**. Next, an edge-detection procedure was applied to the lower end of this midline. This process generated a trace of the larynx's vertical movement over time (represented by the white squiggly horizontal line in the right panel of **Figure 2.3**) by tracking the position of its lower edge at the transition from lower intensity (dark blue) pixels, representing air in the trachea, to higher-intensity (green to yellow) pixels, representing laryngeal tissue. The lower edge was chosen to avoid interference from epiglottal movement.

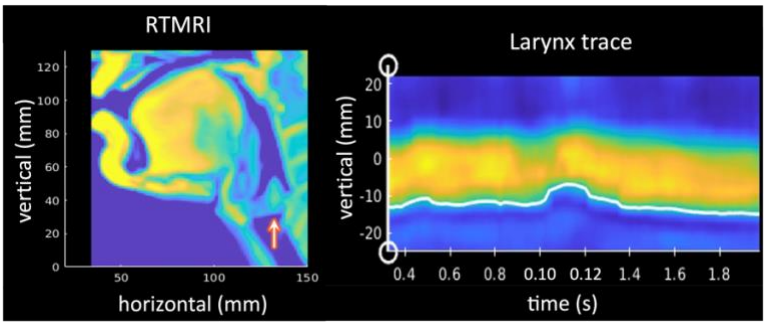

Figure 2.3 MR image of the vocal tract configuration at stop closure (left panel). The orange arrow marks the lower edge of the larynx, used to generate the trace of vertical laryngeal movement over time, represented by the white line in the right panel.

To analyze larynx movement during stop production, larynx displacement was quantified as the difference between two positional values extracted from the larynx trace. These values correspond to the position of the lower larynx's edge at the moment of the stop's closure onset (P1) and offset (P2), as seen in **Figure 2.4**, and were accessed via the previously segmented acoustic signal. Thus, *delta larynx* (in mm) = $\text{larynxPosition}_{\text{closureOffset}}$ - $\text{larynxPosition}_{\text{closureOnset}}$. These timepoints were chosen based on the assumption that active larynx movement can meaningfully contribute to intraoral pressure only when the oral cavity is fully sealed – that is, during the stop's closure.

In interpreting the *delta larynx* results, positive values indicate that the larynx was higher at closure offset than at closure onset of the stop, suggesting upward movement. Negative values indicate that the larynx was lower at closure offset, suggesting larynx lowering.

Outliers or missing values (NaNs) in the positional data sometimes resulted from the edge-detection algorithm failing to identify the lower larynx's boundary. In the larynx's movement trace (see **Figure 2.3**, right panel), such frames would appear as gaps or interruptions in the white line. The difficulties with the edge-detection arose because either the larynx was poorly defined in the image or it was obscured altogether. Such issues are common in rt-MRI data, where rapid articulatory motion can degrade image quality through poor contrast, motion blur, and ghosting, a phenomenon in which moving structures appear smeared or duplicated as faint "shadows" offset from their true position (Lingala et al., 2017; Ramanarayanan et al., 2018).

Although larynx movement was reliably captured in most of the dataset, tracking failures at one of the two key timepoints – closure onset (P1) or closure offset (P2) – occasionally resulted in missing values for *delta larynx* calculations. Tokens affected by this were excluded from further analysis.

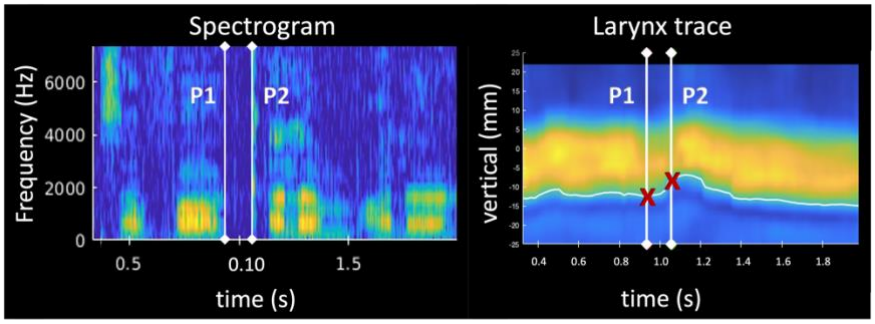

Figure 2.4 Calculation of larynx displacement (in mm) as the difference between two positional values extracted from the larynx trace using the timepoints from the acoustic signal (delta larynx = P2-P1). P1 corresponds to closure onset, and P2 to closure offset of the stop.

2.7.2. Estimating Intraoral Cavity Size

The second goal of the articulatory analysis in this dissertation was to investigate whether ejective production involves an overall compression of the supralaryngeal cavity, a strategy for generating intraoral pressure that does not depend on vertical larynx movement. This articulatory strategy recalls the *IPA Handbook*'s (1999) description of ejective production as simply compression of the intraoral cavity, notably without specification as to the articulators involved.

Support for this possibility emerges from visual comparisons of vocal tract configurations: as illustrated in **Figure 2.5**, the intraoral cavity at closure offset is visibly smaller in an ejective (left panel) than in its pulmonic counterpart (right panel), suggesting that supralaryngeal constriction may play an active role in pressure buildup in ejectives. To assess this possibility systematically, cavity size was measured at both the onset and offset of the stop closure, providing a proxy for overall volume reduction during the pressure-generation phase.

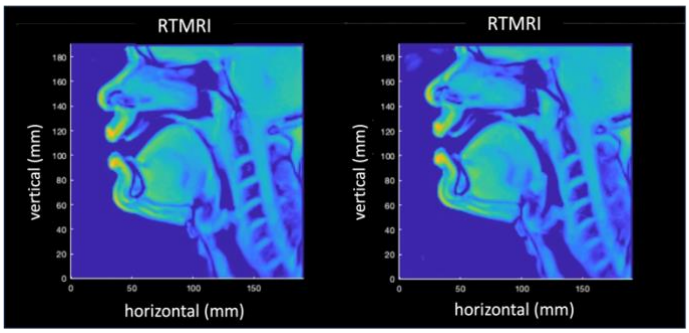

Figure 2.5 Comparison of the vocal tract configuration at the moment of closure offset for a velar ejective (left panel) and a velar pulmonic voiceless stop (right panel) produced by the same Amharic speaker.

The parameter estimating cavity size was calculated using the MATLAB function *bwselect*. This function was used to identify and count the number of directly adjacent low-intensity pixels representing air, starting from a manually selected seed point placed in a region of large aperture at the center of the low-intensity (dark blue) area of the oral cavity. The count expanded outward from the seed point until it encountered high-intensity (green to yellow) pixels marking tissue boundaries, at which point the counting process terminated. This procedure returned a value representing the total area of air within the intraoral cavity, that is, the two-dimensional size of the cavity, hereafter referred to as *cavity size* for simplicity.

A threshold was defined to precisely determine the intensity level above which pixels were classified as tissue rather than air. The challenge lay in calibrating this threshold carefully: if set too low, tissue pixels might be misidentified as air, potentially causing the articulatory

closure in the vocal tract to be missed and resulting in inflated cavity size values; if set too high, articulatory structures or ghosting artifacts might be mistaken as closures, prematurely terminating the count and underestimating the true cavity size. In addition to the cavity size value, a second value was returned indicating the location along the vocal tract where the closure was detected. This location was cross-checked against the expected PoA for the stop token. If the detected closure location aligned with the token's intended PoA, the cavity size measure was considered valid. Otherwise, the token was discarded. Ultimately, different thresholds were used for the two languages, based on the number of valid tokens returned at each threshold level.

Additionally, hard boundaries were set at both ends of the vocal tract, one at the lips and the other at the upper reach of the larynx's raising gesture, as well as a third boundary in the region of the velum. These boundaries prevented the function from "leaking" outside the intraoral cavity, avoiding inflated pixel counts and unrealistic cavity size measures. The lower boundary also ensured that cavity size measurements remained independent of vertical larynx movement.

For some speakers, ghosting artifacts consistently appeared below the velum, leading to artificially reduced cavity size estimates when present. Unlike cases where ghosting confounded the detection of the articulatory closure location, as just described above, these types of distortions are difficult to identify without visually inspecting each individual image. However, since they occurred relatively consistently across all stop types for the affected speakers, their impact on comparisons between stop types is assumed to be minimal.

As with the larynx-raising measure described above, the difference between the two values was calculated as *delta cavity* (in mm^2) = $(cavitySize_{closureOffset} - cavitySize_{closureOnset}) * (-1)$.[3]
A positive *delta cavity* value indicates that the cavity was smaller at closure offset than at closure onset, suggesting progressive cavity reduction. A negative *delta cavity* value indicates expansion of the cavity throughout the closure.

2.7.3. Assessing Supralaryngeal Constriction

In addition to larynx movement and cavity size, a third parameter was computed to further analyze supralaryngeal articulation during ejective production, following methods from Carignan et al. (2020). This parameter quantifies the degree of constriction within individual

[3] Multiplication by (-1) was applied to make the interpretation of the results more intuitive in the figures.

supralaryngeal regions of the vocal tract. The aim of this analysis was to investigate whether targeted articulations above the larynx, such as tongue root retraction, might further contribute to cavity compression and, in turn, support the buildup of intraoral pressure as is hypothesized in the literature (Demolin, 2002; Hermes et al., 2016; Kingston, 1984).

This method involved overlaying a semi-polar grid consisting of 55 lines (or 28 lines in the older version used for the English data) along the length of the vocal tract, spanning from the glottis to the alveolar ridge. The grid was applied semi-manually, allowing flexibility to accommodate individual speaker differences in vocal tract size and shape. The grid was designed as follows: First, key anatomical landmarks – the glottis, velopharyngeal port and alveolar ridge – were manually identified to serve as anchor points for the grid structure.
The midpoint of the diagonal line connecting the alveolar ridge and glottis was designated as the base of the semi-polar grid. An automatic detection process was then used to trim each gridline, ensuring that it terminated at the posterior or superior boundary of the vocal tract.

To measure the degree of constriction within a particular articulatory region of the vocal tract, the mean pixel intensity of each gridline was first calculated. Then, for the English data, gridlines were organized into five clusters representing the hypopharyngeal, hyperpharyngeal, velar, palatal, and alveolar regions. The Amharic data had gridlines divided into nine clusters: the glottal, hypopharyngeal, hyperpharyngeal, post-velar, pre-velar, palatal, post-alveolar, alveolar, and dental regions. This modification of increased number of gridlines enabled a finer partitioning of the vocal tract, allowing for the division of the coronal region into three distinct areas and the separation of the glottal region from the hypopharyngeal region.

The specific number of gridlines assigned to a region varied depending on the speaker's individual anatomy and speaking pattern. To determine a speaker's alveolar region, for example, all images of that speaker producing an alveolar sound were averaged to create a representative image of their typical alveolar articulation. The relevant gridlines for this region were then identified visually from the averaged image. This same process was applied to the post-alveolar, palatal, and pre-velar regions in the Amharic data, and correspondingly to the alveolar, palatal, and velar regions in the English data.

For the remaining regions, gridlines were assigned based on anatomical landmarks.

- Glottal region[4]: Gridlines covering the approximate range of vertical larynx movement.
- Hypopharyngeal region: Gridlines spanning from the highest larynx position to just below the tip of the epiglottis.
- Hyperpharyngeal region: Gridlines spanning from the tip of the epiglottis to the lowest boundary of the open velum. This region is intended to capture articulatory adjustments such as tongue root retraction and pharyngeal wall contraction, both of which are hypothesized to contribute to ejective production (Kingston, 1984).
- Post-velar region: Gridlines covering the soft palate.

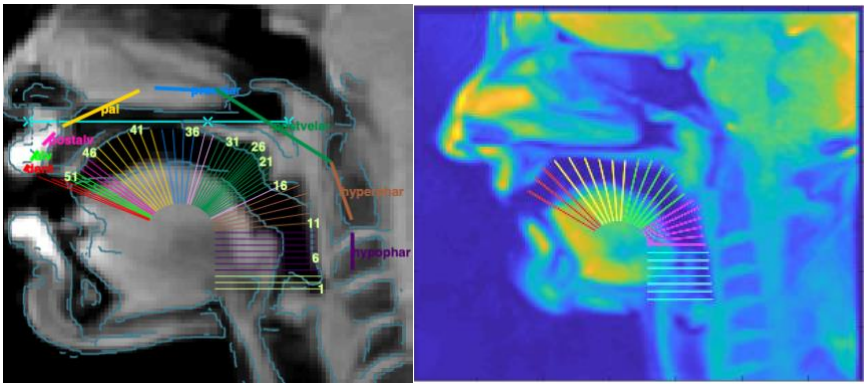

Figure 2.6 Example of gridline assignments for articulatory regions according to the different language study: the left panel displays a speaker from the Amharic dataset, while the right panel features a British English speaker. The grid used in the Amharic study consists of 55 lines divided into nine distinct regions (glottal, hypopharyngeal, hyperpharyngeal, post-velar, pre-velar, palatal, post-alveolar, alveolar, and dental). In contrast, the English grid comprises 28 lines divided into five regions (hypopharyngeal, hyperpharyngeal, velar, palatal, and alveolar). Each region is color-coded for clear differentiation. The Amharic study allows for overlapping gridlines between adjacent regions. These are indicated in pink. This overlap means that the pixel intensities from these gridlines are included in the average signal intensity calculations for both regions.

Figure 2.6 illustrates the gridline assignments for each region in the two studies. The left panel displays an example of the grid subdivision of an Amharic speaker, while the right panel shows an example of the grid subdivision of a British English speaker. Counting from the lowest line above the glottis, the average gridline distribution per region across speakers is given for both languages in **Table 2.1**.

[4] In the English study, this region is not separately assigned, but is included in the hypopharyngeal region.

Table 2.1 Average gridline assignment per region across speakers and language datasets. The Amharic study used a grid that was made up of 55 gridlines, subdivided into nine regions; the English study used a grid-system of 28 gridlines, subdivided into five regions, with the non-used regions left empty accordingly.

		Articulatory region								
		Glottal	Hypo-pharyngeal	Hyper-pharyngeal	Velar		Palatal	Post-alveolar	Alveolar	Dental
					Post-	Pre-				
Language	Amharic	3 - 5	6 - 8	4 - 7	17 - 20	5 - 8	4 - 7	3 - 5	3 - 4	3 - 4
	British English	/	7	5 - 7	5 - 7		5 - 7	/	2 - 3	/

The region-specific constriction values were calculated by averaging the mean intensity of the gridlines within each region. For each speaker, the minimum and maximum intensity values were normalized to a scale from 0 to 1, with the 0.5 percentile set to 0 and the 99.5 percentile set to 1. These percentiles were derived from frames within the VCV segment surrounding the target stop, though normalization was applied to all frames. Higher values indicate a greater degree of constriction in a given region, as they reflect increased presence of tissue (e.g., from the tongue), which raises intensity along the corresponding gridlines. Conversely, lower values suggest a reduced degree of constriction.

A major advantage of this grid-based approach is that it partitions the vocal tract into smaller regions, enabling fine-grained insight into the behavior of specific articulators during ejective stop production, an area that remains largely unexplored. Unlike the overall cavity size measure described above, this method targets local articulatory dynamics and, in principle, allows region-by-region analysis across the entire vocal tract.

As an initial step toward that broader agenda, this dissertation focuses on the hyperpharyngeal region. A comprehensive analysis of all regions is beyond the scope of the present work, and the hyperpharyngeal region is the most theoretically motivated starting point, because this is where the adjustments most commonly hypothesized for ejectives, tongue root retraction and pharyngeal wall constriction, should be observable (Kingston, 1984). By prioritizing this region, the present studies directly test these longstanding proposals, while providing a foundation for future region-by-region analyses of the full vocal tract.

Hyperpharyngeal constriction analysis. The goal of this analysis was to determine whether supralaryngeal adjustments, such as tongue root retraction and constriction of the pharynx, contribute to the buildup of high intraoral pressure associated with ejective stops. If this mechanism is indeed involved, one would expect the degree of constriction in the hyperpharyngeal region to increase over the course of the closure, thereby reducing cavity size and facilitating pressure buildup.

To assess this, two constriction values were extracted from the hyperpharyngeal region's articulatory signal at the same key time points used for the larynx movement and cavity size parameters: closure onset and closure offset of the stop. The closure offset was chosen under the assumption that supralaryngeal articulation contributing to cavity reduction, and thus contributing to increasing intraoral pressure in ejectives, would reach maximum constriction just before the release of the oral closure. As in the previous analyses, the images corresponding to these time points were accessed using the segmented acoustic signal.

The difference between the two values (delta) represents the change in the degree of constriction during the stop's closure. The delta was calculated as $\text{constriction}_{\text{closureOffset}}$ - $\text{constriction}_{\text{closureOnset}}$, meaning that a positive delta indicates a greater degree of hyperpharyngeal constriction at closure offset relative to closure onset, while a negative delta reflects a lesser degree of hyperpharyngeal constriction at closure offset. A delta value of zero indicates no change, meaning the hyperpharyngeal region is constricted to an equal degree at both the onset and offset of the stop's closure. Accordingly, *delta hyperph* is a normalized, unitless measure of change in hyperpharyngeal constriction.

Altogether, three articulatory measures were computed from the rt-MRI data to assess both laryngeal and supralaryngeal contributions to ejective production. Each measure was calculated over the stop's closure. *Delta larynx* captures changes in vertical larynx position. *Delta cavity* measures changes in supralaryngeal cavity size. *Delta hyperph* quantifies changes in the degree of constriction within the hyperpharyngeal region. Having outlined both the acoustic and articulatory measures used in the following ejective studies the next section describes the cross-linguistic study design and comparative framework used to evaluate how ejectives are realized across the two languages and contexts.

2.8. Cross-Linguistic Design and Comparative Framework

The cross-linguistic framework of this thesis enables the investigation of ejective production both in a stable phonological system (Amharic) and in a system where ejectives are not part of the established phonological inventory but are currently emerging (British English). To this end, three empirical studies were conducted: two language-specific analyses and one direct cross-linguistic comparison.

Acoustically, ejectives were examined using measures of stop closure duration, relative burst intensity, post-burst lag intensity, intensity rise ratio of the following vowel, burst

duration, and VOT (the latter two only in the Amharic study). Articulatorily, the analysis focused on three key measures: *delta larynx*, *delta cavity*, and *delta hyperph*. These measures were first analyzed individually. In a final step, relationship analyses were conducted to examine (i) the interaction between larynx raising and supralaryngeal cavity compression, (ii) the contribution of localized hyperharyngeal constriction to overall supralaryngeal cavity compression, and (iii) the extent to which each of these three articulatory measures predicts absolute burst intensity. Because direct intraoral pressure measurements were unavailable, absolute burst intensity was used in these relationship analyses as an indirect proxy for pressure magnitude.

The Amharic study compared ejective stops with voiceless pulmonic stops as a function of place of articulation (alveolar vs. velar), word position (initial, medial, final), and consonant length (singleton vs. geminate). Since the production of voiced stops typically involves larynx lowering to maintain a transglottal pressure differential, they would not serve as appropriate controls. In contrast, voiceless stops generally involve either a neutral larynx position or slight larynx raising, making them a better baseline for comparing laryngeal activity during ejective production. This study aimed both to deepen the understanding of articulatory mechanisms in contrastive ejectives and to provide a reference point for evaluating emerging ejectives in English.

In English, although ejectives usually substitute for voiceless stops, they can occasionally replace originally voiced pulmonic stops (Gordeeva & Scobbie, 2011). Accordingly, the English study included comparisons of ejectives with both voiceless aspirated stops (referred to hereafter as simply “pulmonic voiceless”) and voiced pulmonic stops, exclusively in word-final position, across three places of articulation: bilabial, alveolar, and velar. Including voiced stops enabled a more comprehensive examination of the conditions that give rise to ejective realizations in English.

Together, these investigations asked to what extent being “ejective” is conditioned by glottalic initiation via active larynx raising and to what extent alternative articulatory strategies not traditionally associated with this sound class may also play a role. This line of inquiry challenges, but ultimately aims to refine, our categorical understanding of ejectives.

2.9. Statistical Analysis

All analyses were conducted in R (R Core Team, 2023; version 4.3.2) using lme4 (v1.1-36; Bates et al., 2015). Auditory categorization of English coda stops was analyzed with a logistic mixed-effects model (*glmer*). Continuous articulatory and acoustic outcomes were analyzed

with linear mixed-effects models (*lmer*), with p-values obtained via *lmerTest* (v3.1-3; Kuznetsova et al., 2017).

Fixed effects included stop type and PoA (alveolar, velar; with bilabial additionally in English). Amharic models also included word position (initial, medial, final) and length (singleton, geminate). To test whether the effect of stop type depended on these factors, interactions between stop type and each predictor were included in the initial models. All models included random intercepts for speaker and word; a by-speaker random slope for stop type was added when convergence permitted.

Dependent variables were transformed where needed to improve residual distributions; where applicable, effects were interpreted on the original scale using back-transformed estimates. Model terms were reduced via backward stepwise selection using *step()* in *lmerTest* (starting from the maximal fixed-effects structure).

Post hoc comparisons and simple effects were obtained using *emmeans* (v1.10.1; Lenth et al., 2021) and *emtrends* for continuous moderators. Statistical significance was assessed at $\alpha = .05$. Labels are used consistently throughout: *Ej* = ejectives, *Pvl* = pulmonic voiceless stops, *Pvd* = pulmonic voiced stops. In the tables, abbreviations are: *bil*/*alv*/*vel* = bilabial/alveolar/velar; *i*/*m*/*f* = initial/medial/final; *s*/*g* = singleton/geminate.

This chapter provided an overview of the general methodological approach used in this thesis. Methodological details specific to only one of the two languages will be presented briefly in the respective methods sections of the following chapters: Chapter 3 for Amharic and Chapter 4 for British English. These next two chapters present the empirical analyses of ejectives in each language separately, followed by a comparative analysis in Chapter 5.

Chapter 3
Phonological Ejectives in Amharic

3.1. Introduction

This chapter examines the acoustic and articulatory properties of ejective stops in Amharic, a Semitic language spoken in Ethiopia. The aim is to contribute to a more empirically grounded understanding of how ejectives, traditionally defined as glottalic, egressive consonants (Catford, 1977), are produced. An ongoing debate has questioned the precise role of the larynx in ejective production (Kingston, 1985) and even challenged the assumption that ejectives are inherently glottalic (Brandt & Simpson, 2021). These discussions have given rise to the idea that ejective production may in fact be more variable than previously assumed. In response, this study adopts a broader perspective on the articulatory strategies for producing these sounds by considering not only the role of the larynx but also the potential contribution of supralaryngeal articulators.

To this end, the analysis investigates vertical larynx movement, the defining gesture of glottalic initiation, alongside potential supralaryngeal strategies such as overall cavity compression and pharyngeal constriction. It also explores the laryngeal-supralaryngeal relationship in production and their respective contribution to building up intraoral pressure. The ejective inventory in Amharic makes it well suited for this purpose: ejective stops occur at three places of articulation (bilabial, alveolar, velar), as both singletons and geminates, and in word-initial, medial, and final positions. This range enables a nuanced analysis of how contextual factors variably influence phonological ejective production.

Before turning to articulation, the chapter begins with an acoustic analysis meant to situate the Amharic ejectives within established cross-linguistic patterns and provide essential context for interpreting the articulatory findings that follow. Together, these analyses aim to shed light on the core features of phonological ejective production and establish a foundation for later comparison with non-phonological ejective sounds in English (Chapter 4) and their cross-linguistic interpretation in Chapter 5.

This chapter is structured as follows: Section 3.2 is a short methods section. Section 3.3 presents the acoustic analysis. Section 3.4 details the articulatory investigation based on rt-MRI data, focusing on three components: vertical larynx movement, supralaryngeal cavity compression, and hyperpharyngeal constriction.

Section 3.5 examines the relationship between these articulatory features and absolute burst intensity, with burst intensity serving as an indirect proxy for intraoral pressure. All analyses examine the effects of place of articulation, word position, and consonant length. Before turning to the analyses themselves, the rest of this section gives information on the Amharic language and its phonological system and reviews previous work on Amharic ejectives.

3.1.1. The Amharic Language

Amharic is a South Ethio-Semitic language and one among many languages spoken in Ethiopia (Hetzron & Bender, 1976). It is the Ethiopian government's official working language. It also widely acts as lingua franca within the multilingual state. Nevertheless, within the Ethiopian population it is only the second most frequently spoken primary language (29% of the population as of 2007).

Different theories still exist regarding the time and place of emergence of Amharic. The broadly accepted view today is that Amharic and Geʿez – an ancient South Semitic language, which was the cultural language of the Aksumite Empire and remained the exclusive literary language of Ethiopia until the 19th century – share an ancestral Ethio-Semitic proto-language (Hetzron, 1972), while another view sees Amharic as directly descended from Geʿez (Ullendorff, 1955).

Amharic's writing system originates from the Geʿez script. This is a syllabary script, meaning every grapheme (called fidäl) represents a consonant-vowel unit. Each grapheme corresponds roughly with a consonant and has 7 shapes, depending on the vowel. The basic Amharic word order is SOV and the maximal syllable size is CVCC, where the coda cluster is either a geminate or a cluster of falling sonority. Onset consonant clusters must be of rising sonority (Devens, 1981; Hayward, 1986, p. 304; Leslau, 1997, p. 427ff.).

Regional varieties or dialects exist, which exhibit minor differences, mainly at the level of pronunciation (Leslau, 1995). The Amharic spoken in the capital city, Addis Ababa, is considered the prestige variety (Leslau, 1995). In addition to the Addis Ababa Standard variety, four major varieties are considered, namely, the dialects of Shoa, Begemder, Gojjam, and Menz-Wello (Amare, 1983; Amsalu & Habtemariam, 1973; Baye, 2016; Cowley et al., 1976; Gebre, 1999; Mengistu, 2018).

Ethiopia's is an ancient culture with a rich linguistic history, exemplified by the status of Geʿez, which to date remains the language of church and literature. As a result, Amharic has been extensively studied, particularly in morphology and syntax. In terms of its phonetics, however, less work has been done, leaving many of its distinctive features, such as labialization, gemination, stress, and the obligatory epenthetic vowel, largely underexplored. Among the few phonetic studies that do exist, ejectives have received the most attention. Yet, to my knowledge, no research on Amharic ejectives has systematically examined the glottalic production mechanism to date.

3.1.2. Amharic Phonology

Amharic exhibits several distinctive features that are also characteristic of the broader Ethio-Semitic language branch. Apart from labialization and gemination the presence of glottalic sounds, including even the typologically rare ejective fricative, are of special interest. Well-established as a widespread morphophonemic process, the status and extent of phonemic labialization is, however, still a topic of debate (Derib Ado, 2020). This, as well as the disputed status of some palatal consonants and the glottal stop, means that there currently is no real consensus on the exact size of the Amharic phonological inventory to be found in the literature (Derib Ado, 2020; Leslau, 1995, 1996, 1997; Rose, 1997; Sande & Hedding, 2017; Sumner, 1957).

Gemination is phonemic and can occur at the medial and final positions of a word and in all consonants, except for the glottal stop and fricative. It is sometimes considered non-distinctive in the palatal nasal (Hayward & Hayward, 1992) and is hardly audible in final position (Leslau, 1997). Lexical gemination is rare. Primarily, gemination arises from morphological processes (Ferguson, 1970; Leslau, 1997). Gemination is not marked in Amharic orthography, however, Western scholars will sometimes indicate it using two dots above the symbol in question (Leslau, 1995, pp. 1-3). Amharic has 5 glottalic consonants: 3 plain ejective stops at the bilabial, alveolar and velar places of articulation, as well as a post-alveolar ejective affricate and an alveolar ejective fricative. There is a labio-velar ejective stop, whose phonological status is not definitively established.

Each glottalic consonant has two pulmonic counterparts: a voiceless and a voiced version. However, the exact voicing characteristics of these pulmonic counterparts remain unspecified.

While Hayward and Hayward (1992) characterize pulmonic voiceless stops as moderately aspirated, Derib Ado (2011) describes them as alternating freely between no aspiration and slight aspiration. Similarly, Derib Ado's (2011) data on pre-voicing is inconclusive regarding the presence of lead or lag VOT, or whether both are present. Thus, the voicing characteristics of the pulmonic stops remain unresolved and await targeted investigation.

Interestingly, the pulmonic voiceless bilabial and glottalic bilabial stops, as well as the voiced labiodental fricative, are not original features of Amharic, nor of many other South Ethio-Semitic languages. They appear marginally and are exclusive to loanwords, /p'/ to words originally borrowed from Greek and /p/ to recent words from other European languages (Ferguson, 1970; Hudson, 2000). The exact PoA of /l, r, s, z, n, t, d/ – whether dental or alveolar – remains a matter of debate and requires further investigation (Derib Ado, 2011). The ejective fricative /s'/ originates from Geʿez and is influenced by sociolectal factors (Takkele, 1992). Stress is described as being weak and indeterminate, or rather its place variable, and is apparently marked by the length of the syllable (Cowley et al., 1976; Hayward & Hayward, 1992). However, this is another question which, given the lack of experimental studies, remains unresolved.

Finally, concerning the vowel inventory: in addition to /i, e, o, u, a/, Amharic has two central vowels. They are described as considerably shorter in length and subject to a greater range of variation compared to the other vowels (Ferguson, 1976). In the literature, these vowels are most commonly referred to as /ɨ/ and /ä/, though there is still considerable variation in the phonetic symbols used to represent them. The most recent – and, to my knowledge, only – analysis of the basic properties of Amharic vowels proposes an update to the commonly used vowel chart. Due to an ongoing vowel shift, [ə] for the high-mid vowel and [ɜ] for the low-mid vowel (central, but lower than the classic schwa) are suggested as more accurate representations of the current vowel system (Derib Ado, 2011). Functionally, both high-central vowels play a significant role in the language: the former serves as an obligatory epenthetic vowel to break up consonant clusters, which are largely disallowed in Amharic, while the latter is the most frequently occurring vowel in the language (Ferguson, 1976).

3.1.3. Acoustic Studies of Amharic Ejectives

To date, only a handful of acoustic and aerodynamic studies have examined Amharic stops. An acoustic study by Seid at al. (2009) compared Amharic ejectives and their voiceless pulmonic counterparts, both as singletons and geminates, in different vowel environments within isolated words, spoken by a single native Amharic speaker.

This study found that Amharic ejectives had longer stop closures, shorter VOTs, a slower amplitude rise time of the following vowel, and creaky voice quality compared to voiceless pulmonic stops. Impressionistically, burst amplitude was not judged to differ between the two stop types. According to this study, these properties, together with the notable lack of a silent gap before vowel onset, define the acoustic profile of Amharic ejectives.

Derib Ado's (2011) dissertation is, to my knowledge, the most comprehensive investigation of the acoustic properties of both Amharic vowels and stop consonants to date. It analyzed a string of acoustic properties including durational, burst spectral, and intensity measurements of the stops, as well as voice quality/phonation measurements of the stop-following vowels. The data, gathered from eight speakers (four female) of the Addis Ababa standard variety, included durational, burst spectral, and intensity measurements for /p, p', b, t, t', d, k, k', g/, as well as voice quality measurements of the stop-following vowels. The analysis used real words and focused on stops in word-initial and, except for the bilabials, word-medial positions.

Focusing on the results most relevant to the present study, Derib Ado (2011) found that Amharic ejectives exhibited shorter burst duration and, consistent with Seid at al.'s (2009) observations, slower intensity rise time than pulmonic voiceless stops. Velar stops showed the longest VOTs overall, regardless of stop type. Ejective velars and alveolars had longer VOTs than their pulmonic voiceless counterparts, but for bilabials the opposite VOT distinction was true. Similarly, absolute burst intensity was higher for alveolar and velar ejectives compared to their pulmonic voiceless counterparts, but for bilabials this relationship was reversed. Initial stops had overall higher intensity than intervocalic stops. In terms of relative burst intensity, this was significantly lower for ejectives at all three PoAs compared to voiceless pulmonic stops. Alveolar and velar medial ejectives had higher relative burst intensity than initial ejectives. About half of total ejective cases (46%) in Derib Ado's study (2011) were followed by creak, with the mean duration of the creaky interval lasting some 24 ms.

Stop type had no effect on the duration of the stop's closure. Overall, this study identifies the key characteristic properties of Amharic ejectives as a tendency toward early glottal release, a marked aperiodic vowel onset, and a relatively rapid rise in vowel intensity.

The most recent investigation of Amharic ejectives (Negesse, 2023) is a cross-linguistic study that compared the characteristics of ejective stops of two different Ethiopian languages, Amharic and Oromo. Initial bilabial, alveolar and velar ejective stops of 36 speakers (18 per language) were analyzed within monosyllables (CV). Nine acoustic measures were investigated: VOT, intensity rise time, H1-H2 and f0 of the ejective following vowel, relative burst intensity, and four burst spectral moments.

The results showed no significant acoustic differences between the ejectives of the two languages. Instead, the most notable distinctions emerged across PoA (Hajek & Stevens, 2005; Warner, 1996), with bilabial ejectives standing out due to their unique acoustic profile, particularly in the patterning of spectral moments, as well as their shorter VOT and very low relative burst intensity. While velar ejectives exhibited a relatively stable acoustic profile consistent with a single pattern, alveolar ejectives showed substantially greater internal variability. Notably, the observation of modal phonation in vowels following Amharic ejectives contrasts with earlier reports (Derib Ado, 2011; Seid et al., 2009). Taken together, the three existing acoustic studies of Amharic ejectives present divergent characterizations of their acoustic properties.

3.1.4. Aerodynamic and Articulatory Studies of Amharic Ejectives

Over a series of studies (2002, 2004, 2011), Demolin investigated the aerodynamic and articulatory properties of Amharic ejectives, examining intervocalic stops as well as fricatives and affricates, in both singleton and geminate forms. The dataset, collected from five speakers, included oral airflow (via a silicon mouthpiece), intraoral pressure (via a pressure tube through the nasal cavity), acoustic recordings, and electroglottographic signals (EGG). For one participant only, additional measurements were collected: subglottal pressure (via direct tracheal puncture) and electropalatographic recordings (EPG).

A key finding from these studies was that differences between singleton and geminated ejective and pulmonic fricatives and affricates in Amharic can be explained by changes in the coordination and timing of glottal and oral gestures, which in turn affect pressure and airflow parameters.

Notably, the length contrast in ejectives was realized as a delayed onset of larynx raising, acoustically manifested as a noise increase near release, and aerodynamically as a two-phase rise in pharyngeal pressure. Demolin's observation of timing differences in larynx movement between singleton and geminate ejectives are to date the only available, yet indirect insights into laryngeal activity in Amharic ejectives (Demolin, 2011).

Demolin's (2002) intraoral pressure data showed markedly higher pressure during ejective fricatives than during pulmonic fricatives, with the highest pressure observed in geminate ejectives. Following Kingston (1984), Demolin argues that if this pressure buildup were attributable to larynx raising alone, the observed peak increase (≈42 hPa) would require a sizeable reduction in supralaryngeal volume, on the order of ~4 cm^3 (assuming a starting volume of 100 cm^3 and an initial pressure of ~1030 hPa). Such a reduction is difficult to reconcile with larynx raising alone, given Kingston's claims that vertical larynx movement is too slow and not sufficiently tightly coordinated with oral constriction timing to generate the full pressure rise. The Amharic pressure patterns therefore point to additional cavity-reducing maneuvers (e.g., tongue-root retraction, vocal tract wall tensing, or increased oral contact). Consistent with the latter, Demolin's EPG data showed extended alveolo-palatal contact for /s'/ and /s':/, particularly in geminates, suggesting a smaller effective oral cavity during ejectives compared to their pulmonic counterparts.

Building on this foundation, the present acoustic and articulatory investigation offers a timely and in-depth examination of Amharic ejectives, aiming to fill key gaps in our understanding of what phonological production mechanisms of ejectives entail. The following section first outlines the methodological approach taken in the Amharic study, including participant information and the design of the experimental materials.

3.2. Methods

General methods used in this dissertation related to data preparation, segmentation, and processing are described in Chapter 2. The following sections provide study-specific details, including information about participants and experimental materials.

3.2.1. Speakers

Eleven native Amharic speakers (4 female), aged between 22 and 53 ($\mu = 33.3$, SD = 8.8) were recruited in Germany. This was done mainly through online advertisements and outreach to Ethiopian church communities and expatriate organizations. All speakers originated from Ethiopia, their average length of stay in Germany at the time of recording was 6.2 years (SD = 7.7). Though one speaker was born in the Tigray region, in which Tigrinya is actually the language most commonly spoken natively, all eleven speakers reported Amharic to be their native language. Three speakers reported knowledge of additional Ethiopian languages besides Amharic: one spoke Tigrinya, another spoke Gurage, and the third spoke both Tigrinya and Oromo. All speakers recounted learning English in school – four speakers as early as in elementary school, and most listed German as their third language, acquired after moving to Germany.

Written information about the speakers' language backgrounds, along with written consent confirming their compatibility with the rt-MRI experiment and the absence of any hearing or speaking impairments, was collected prior to the MRI recording sessions. All speakers were remunerated for their participation. The study procedure adhered to the framework detailed in Chapter 2, Section 2.3.

3.2.2. Stimuli

This study intentionally focused on the three plain ejective stops, despite the broader range of glottalic consonants present in the Amharic phonological inventory. This decision was motivated by two factors. First, ejective fricatives and affricates have their own aerodynamic requirements that warrant separate, careful considerations. Second, the focus was restricted to ejective stops to enable a meaningful comparison with non-phonological ejective production in English, where only ejective stops are attested.

The target words comprised 271 lexical items, carefully selected using Amharic dictionaries (Kane, 1990; Leslau, 1976) to fit the criteria for the experiment (see below). The complete list of Amharic words used in this study is provided in Appendix A. Final consultation on the word list was provided by two Amharic native speakers with whom the correct spelling of the words in the Ethiopic script was verified. The target words were elicited via the carrier phrase "*Isana <target-word> ale"* (literally translated as *"Isana <target-word> says")*.

To allow for some degree of variation in the recording process five randomly alternating Ethiopian names (Isana, Amha, Täsäma, Asefa, Desta, Ezana) were used phrase-initially. The target words compare ejective and pulmonic voiceless stops at three places of articulation, bilabial /p’ vs. p/, alveolar /t’ vs. t/, and velar /k’ vs. k/, as singletons and geminates; as well as at three positions in the word, initial (this position does not feature geminates), medial and final. The selection of the target words and the construction of the carrier phrase proceeded in such a way as to permit analysis of the target stops in a strictly VCV context, even across word boundaries. Geminates were elicited via lexical knowledge, given that gemination is not represented orthographically in Amharic.

In order to control for vowel context as much as possible, while still maximizing the number of viable words, the choice of vowels was restricted to the three Amharic mid-vowels /a, ə, ɜ/. To guarantee as many minimal pairs or near minimal pairs as possible, two, three and four syllable words were chosen under the condition that they include only one ejective. The presence of a single ejective within the target word was also controlled for. Overall, the word list comprises 136 ejective and 115 pulmonic voiceless items, which include 20 bilabial, 120 alveolar, and 131 velar tokens.

3.2.3. Dataset Overview and Exclusions

A total of approximately 271 words were recorded per speaker, yielding 2,981 tokens across the dataset. After excluding error tokens, such as incorrectly produced target words, disfluent productions, or technical interruptions, and removing all bilabial tokens, 2,624 tokens remained available for analysis (averaging 239 tokens per speaker).

While data on bilabial stops was recorded, it was ultimately excluded from the final analysis. Given the inherent marginality of bilabial stops in Amharic, the dataset already exhibited a significant imbalance in token distribution across PoA (only 20 bilabial target words were included in the original word list), a disparity exacerbated by the removal of error tokens and outliers (see Chapter 2, Section 2.7). Nevertheless, to offer some insight into their production, Appendix B includes figures illustrating acoustic and articulatory measures for which a sufficient number of bilabial ejective tokens were available.

The remaining 2,624 tokens were used in the acoustic analysis, which examined closure and burst duration, VOT, relative burst intensity, post-burst lag (PBL) intensity, and intensity rise ratio.

As outlined in Chapter 2, additional exclusions were made during articulatory data preparation due to challenges associated with rt-MRI signal quality. After these steps, 2,607 tokens remained for the analysis of larynx movement.

For the cavity size analysis, a reduced subset of 1,643 tokens was available. The number of usable tokens varied considerably across speakers due to ghosting artifacts, imaging distortions below the velum that interfered with accurate cavity estimation. While some speakers' data was unaffected, others showed frequent ghosting, rendering many of their cavity measurements unusable. For supralaryngeal constriction analysis, 2,661 tokens were retained from the hyperpharyngeal region of the vocal tract.

The final set of relationship analyses (e.g., larynx raising vs. cavity size reduction) drew on the relevant subsets from each articulatory measure. As a result, token counts varied across analyses, matching the figures reported above.

3.2.4. Statistical Modeling

The six acoustic measures (closure and burst duration, VOT, relative burst intensity, PBL intensity, and intensity rise ratio) and three articulatory measures (larynx, cavity, hyperph deltas) were analyzed using separate linear mixed models, with each measure serving as the dependent variable. A detailed description of how the measures were computed is given in Chapter 2, Section 2.6 and 2.7. The initial model structure for analyzing Amharic stops in this chapter was as follows:

(1) model = lmer(measure ~ stop type * PoA + stop type * position + stop type * length + (stop type | speaker) + (1 | word)

where stop type included two levels (ejective and pulmonic voiceless), PoA included two levels (alveolar and velar), position included three levels (initial, medial, and final), and length included two levels (singleton and geminate). For the articulatory measures *delta cavity* and *delta hyperph* only, an additional random intercept for vowel context was included – (1 | V1V2) – to account for systematic differences in vocal-tract configuration associated with the surrounding vowels.

(2) model = lmer(deltaCavity/deltaHyperph ~ stop type * PoA + stop type * position + stop type * length + (stop type | speaker) + (1 | word) + (1 | V1V2))

A slightly different model structure was used in the final section of the analysis, which examined the relationship between different measures (e.g., *delta larynx* and absolute burst intensity):

(3) model = lmer(measure1 ~ measure2 * PoA + measure2 * position + measure2 * length + (measure2 | speaker) + (1 | word)

This model was run on the subset of ejective stops. All other factors remained the same. As before, analyses involving *delta cavity* or *delta hyperph* included an additional random intercept for vowel context (1 | V1V2).

Where applicable, models were simplified following the procedure described in Chapter 2, Section 2.9. The outcomes of these models are reported in the following sections, beginning with the results of the acoustic analysis (Section 3.3), followed by the articulatory analysis (Section 3.4), and concluding with the relationship analyses (Section 3.5).

3.3. Acoustic Analysis

The aim of the acoustic analysis was to situate the ejectives examined here within the broader cross-linguistic landscape of ejective acoustics, thereby providing a foundation for the interpretation of the articulatory patterns that follow. The following section presents the acoustic results, first the durational measures (closure duration, burst duration, VOT), followed by the intensity measures (relative burst intensity, PBL intensity and intensity rise ratio).

3.3.1. Closure Duration

Figure 3.7 compares closure duration (in ms) between ejective and pulmonic voiceless stops while making the effect of phonological length explicit. Rather than collapsing across length, the PoA and position panels are faceted by singleton vs. geminate. As shown clearly in the length-only panel, gemination strongly shifts closure duration upward for both stop types. Within each length category, differences between ejectives and pulmonics are generally small and not strongly patterned across PoA or position.

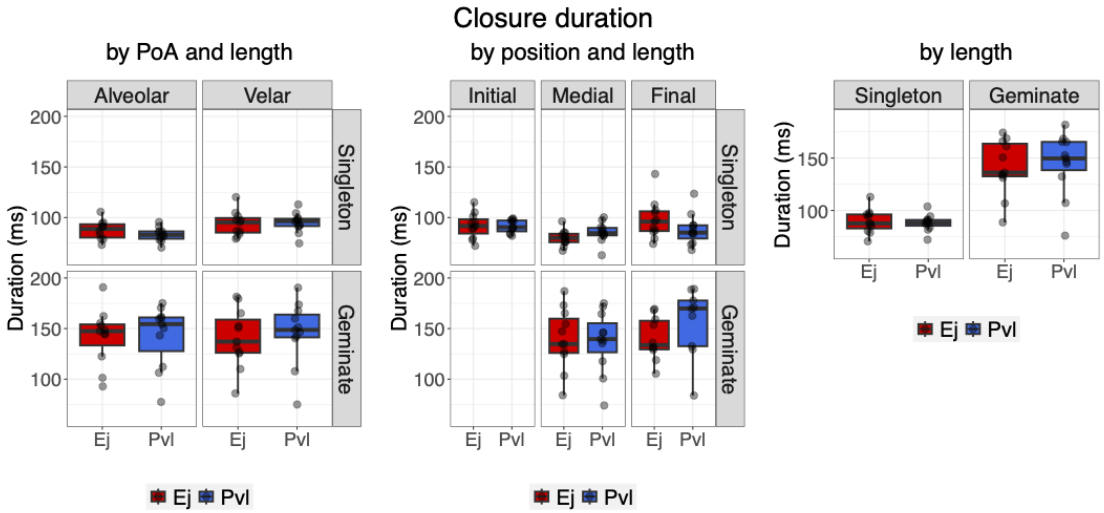

Figure 3.7 Raw closure duration values (y-axis, in ms) for ejective and pulmonic voiceless stops. Speaker-level means (dots) are shown by (i) PoA × length (alveolar vs. velar; singleton vs. geminate), (ii) word position × length (initial/medial/final; singleton vs. geminate), and (iii) length alone. Length is displayed as a facet in the PoA and position panels because gemination substantially increases closure duration, and collapsing across length would obscure this systematic shift and make PoA/position differences harder to interpret.

In addition, because token counts were uneven across PoA, position, and length, the corresponding counts are provided in **Table 3.2** to contextualize the apparent stability of patterns across the facets of **Figure 3.7**.

Table 3.2 Distribution of tokens by length, stop type, PoA, and word position.

Length	Stop type	alv	vel	i	m	f
s	Pvl	448	344	286	211	295
s	Ej	478	529	260	357	390
g	Pvl	143	209	0	281	71
g	Ej	136	215	0	276	75

The lack of observable difference by stop type was confirmed statistically. Closure duration was log-transformed to meet model assumptions due to a positively skewed distribution. Using this transformed measure, neither stop type ($F(1, 31.35) = 0.22$, $p = .64$) nor any of its interactions (stop type × PoA: $F(1, 245.32) = 2.10$, $p = .15$; stop type × position: $F(2, 243.48) = 0.21$, $p = .81$; stop type × length: $F(1, 241.81) = 0.20$, $p = .65$) significantly improved the model fit and were therefore removed during model selection.

The final model retained only the significant main effects of PoA ($F(1, 246.73) = 14.75$, $p < .001$), position ($F(2, 17.81) = 4.96$, $p < .05$), and length ($F(1, 241.36) = 338.89$, $p < .001$).[5]

One outlier speaker (S10) exhibited particularly long closure durations for singleton stops in the final position. However, since this pattern occurred for both stop types, it did not provide insight into how the speaker differentiated ejectives from pulmonic stops. Instead, it likely reflects a general characteristic of their speech. In sum, closure duration did not a differentiate the stop types.

3.3.2. Burst Duration

Of all stop tokens in the dataset, 9% (n=229) had *no* audible release burst (see Chapter 2, Section 2.5.1 for details on how this was determined). Unreleased stops were more common for pulmonic stops (6% of all pulmonic voiceless stops) than for ejectives (3% of all ejectives). Among ejectives, unreleased tokens were more frequent at the alveolar PoA and in final position, with 30% of final alveolar ejectives being unreleased. Notably, one speaker (S01) alone accounted for 34% of all unreleased stops, and this speaker tended to produce unreleased realizations more often with ejectives than with pulmonic stops (20% vs. 14%), contrary to the group trend. Because these tokens lack measurable burst durations, they were excluded from the burst analysis described below.

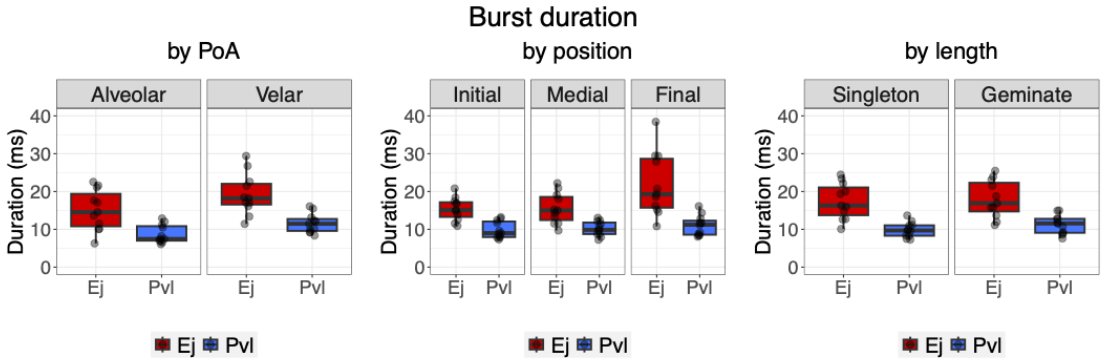

Figure 3.8 Raw burst duration values (y-axis, in ms) for ejective and pulmonic voiceless stops split by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Dots represent speaker-level means; higher values indicate longer burst durations.

[5] Selected model: closureDuration_log ~ PoA + position + length + (position | speaker) + (1 | word).

Figure 3.8 displays the duration (in ms) of all *audible* release bursts for both stop types, shown across PoA, position, and length in the three separate panels. A robust pattern is evident across all panels of this figure, with ejective stops exhibiting longer burst durations than pulmonic stops. The figure further indicates that among ejectives, velars have longer burst durations than alveolars (left panel), and final-position ejectives show both longer and more variable burst durations than initial and medial ones (middle panel), evident in the broader spread of the red boxplots and the wider range of individual speaker means (grey dots).

No substantial difference is visible between singleton and geminate ejectives (right panel). The overall greater spread of individual speaker means (grey dots) in burst durations for ejectives, compared to pulmonic stops, also suggests that speaker variability is more pronounced in the production of ejective bursts.

The results of the statistical analysis align with the visual impression. To meet model assumptions, burst duration was log-transformed prior to analysis. A significant main effect of stop type, as well as of PoA, position, and length (stop type: $F(1, 10.37) = 99.29$, $p < .001$; PoA: $F(1, 253.96) = 250.29$, $p < .001$; position: $F(2, 249.28) = 60.78$, $p < .001$; length: $F(1, 241.58) = 21.46$, $p < .001$) were found. Among the interaction terms, only stop type × position was retained in the final model, indicating a significant interaction ($F(2, 248.19) = 21.49$, $p < .001$). The remaining interactions were dropped during model selection, as they did not significantly improve model fit (stop type × PoA*:* $F(1, 254.32) = 0.09$, $p = .77$; stop type × length*:* $F(1, 243.54) = 3.62$, $p = .06$)[6].

[6] Selected model: burstDuration_log ~ PoA + length + stopType*position + (stopType | speaker) + (1 | word).

Table 3.3 Pairwise comparisons of burst duration (modelled on the log scale) by stop type and position, with back-transformed absolute differences in ms.

Contrast type	Position	Stop type	Contrast	Estimate (log)	SE	DF	t	p	Signif.	Δ (ms)
By position	i		Ej - Pvl	0.45	0.06	16.5	7.70	<.001	***	5.38
By position	m		Ej - Pvl	0.43	0.05	13.1	7.92	<.001	***	4.91
By position	f		Ej - Pvl	0.67	0.06	15.4	11.67	<.001	***	9.42
By stop type		Pvl	m - i	-0.06	0.03	252.6	-1.78	0.1782		-0.51
By stop type		Pvl	f - i	0.04	0.03	269.2	1.32	0.3830		0.41
By stop type		Pvl	f - m	0.10	0.03	263.6	3.27	0.0035	**	0.92
By stop type		Ej	m - i	-0.07	0.03	222.2	-2.42	0.0433	*	-0.99
By stop type		Ej	f - i	0.26	0.03	223.0	9.06	<.001	***	4.44
By stop type		Ej	f - m	0.33	0.02	231.9	13.17	<.001	***	5.43

Pairwise comparisons for the interaction with position (see **Table 3.3**) revealed significantly longer burst durations for ejective stops compared to pulmonic stops across all positions (initial: $\Delta = 5.38$ ms, $p < .001$; medial: $\Delta = 4.91$ ms, $p < .001$; final: $\Delta = 9.42$ ms, $p < .001$), with this effect being particularly strong at the final position. Within the ejectives, burst durations varied significantly by position: final ejectives showed decidedly longer bursts than both medial (final-medial: $\Delta = 5.43$ ms, $p < .001$) and initial ones (final-initial: $\Delta = 4.44$ ms, $p < .001$), confirming the visual observation of **Figure 3.8**. Medial ejectives were minimally shorter than initial ones (medial-initial: $\Delta = -0.99$ ms, $p < .05$). In sum, ejectives showed significantly longer burst durations than pulmonic stops across positions, with the largest effect observed in final position.

3.3.3. Voice Onset Time

Figure 3.9 compares VOT (ms) across stop types, shown by PoA, position and length in the three panels. With the exception of medial position (middle facet of the center panel), VOT tends to be longer for ejective stops than for pulmonic stops across factors. Final-position stops (right-most facet of the center panel) show especially long and variable VOT values, which stand out relative to other positions. These word-final VOT values are not only longest overall but also show the greatest variability across speakers, as reflected in the wide spread of the individual speaker means (grey dots). This mirrors the pattern observed for burst duration in final position described above.

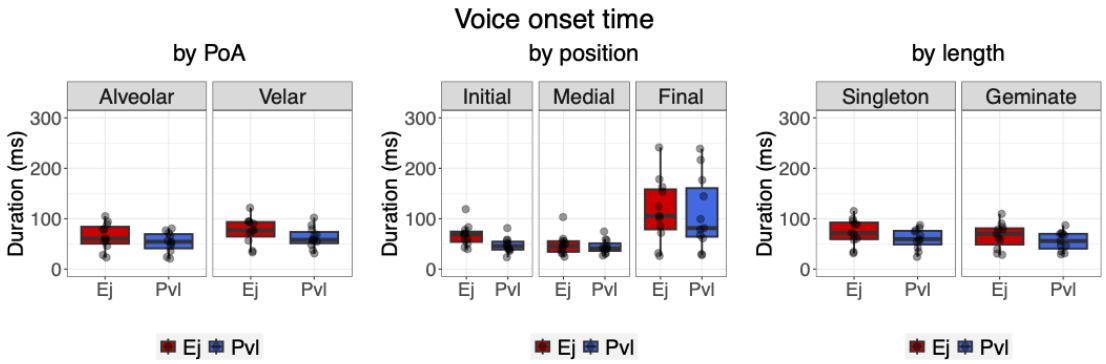

Figure 3.9 Raw VOT values (y-axis, in ms) for ejective and pulmonic voiceless stops split by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Dots represent speaker-level means; higher values indicate longer VOT.

These patterns were confirmed by the statistical analysis. After log-transforming VOT to meet model assumptions, the analysis revealed significant main effects of stop type ($F(1, 13.60) = 16.87$, $p < .01$), PoA ($F(1, 228.01) = 120.58$, $p < .001$), and position ($F(2, 232.19) = 268.03$, $p < .001$). There was also a significant stop type × position ($F(2, 231.52) = 17.10$, $p < .001$) and a stop type × length ($F(1, 220.36) = 4.00$, $p < .05$) interaction. There was no main effect of length ($F(1, 220.47) = 1.14$, $p = .29$) and no significant interaction between stop type and PoA ($F(1, 228.85) = 1.07$, $p = .30$).[7]

As per **Table 3.4**, pairwise comparisons for the interaction with position revealed that VOT was significantly longer for ejectives than pulmonic stops in both initial (Δ = 20.88 ms, $p < .001$) and final position (Δ = 15.79 ms, $p < .01$), but not in medial position (Δ = 1.53 ms, $p = .51$). Within the ejective stop type, significant differences were observed across all comparisons: VOT in medial position was significantly shorter than in both initial (medial-initial: Δ = -21.36 ms, $p < .001$) and final position (medial-final: Δ = -45.20 ms, $p < .001$), while final-position VOT was also significantly longer than initial (final-initial: Δ = 23.83 ms, $p < .001$).

[7] Selected model: VOT_log ~ stopType*position + stopType*length + PoA + (stopType | speaker) + (1 | word).

Table 3.4 Pairwise comparisons of VOT (modelled on the log scale) by stop type and position, with back-transformed absolute differences in ms.

Contrast type	Position	Stop type	Contrast	Estimate (log)	SE	DF	t	p	Signif.	Δ (ms)
By position	i		Ej - Pvl	0.41	0.07	37.7	5.92	<.001	***	20.88
By position	m		Ej - Pvl	0.04	0.06	17.8	0.68	0.5081		1.53
By position	f		Ej - Pvl	0.20	0.06	29.5	3.15	0.0037	**	15.79
By stop type		Pvl	m - i	-0.05	0.05	246.8	-1.04	0.5512		-2.02
By stop type		Pvl	f - i	0.53	0.05	291.2	11.37	<.001	***	28.92
By stop type		Pvl	f - m	0.58	0.04	282.4	12.88	<.001	***	30.94
By stop type		Ej	m - i	-0.42	0.04	214.3	-10.12	<.001	***	-21.36
By stop type		Ej	f - i	0.32	0.04	227.4	7.84	<.001	***	23.83
By stop type		Ej	f - m	0.74	0.04	238.7	20.34	<.001	***	45.20

Regarding the interaction with length, pairwise comparisons given in **Table 3.5** showed that ejectives had significantly longer VOTs than pulmonic stops, both as singletons (Δ = 8.47 ms, $p < .01$) and, more markedly, as geminates (Δ = 15.13 ms, $p < .001$). Within the ejective stop type, VOT was also significantly longer for geminates than for singletons (geminate-singleton: Δ = 5.36 ms, $p < .05$).

Table 3.5 Pairwise comparisons of VOT (modelled on the log scale) by stop type and length, with back-transformed absolute differences in ms.

Contrast type	Length	Stop type	Contrast	Estimate (log)	SE	DF	t	p	Signif.	Δ (ms)
By length	s		Ej - Pvl	0.16	0.05	13.3	3.02	0.0095	**	8.47
By length	g		Ej - Pvl	0.27	0.07	34.1	4.09	<.001	***	15.13
By stop type		Pvl	g - s	-0.03	0.04	250.3	-0.61	0.5414		-1.30
By stop type		Ej	g - s	0.09	0.04	226.7	2.32	0.0211	*	5.36

In summary, ejectives exhibited significantly longer VOT durations than pulmonic stops across all conditions, with the exception of medial position, where no significant difference was observed. Within the ejective stop type, geminates had longer VOTs than singletons, and medial-position ejectives had significantly shorter VOTs than those in both initial and final positions.

3.3.4. Relative Burst Intensity

Figure 3.10 illustrates the effect of stop type on relative burst intensity (in dB) across PoA, position, and length, shown in three separate panels. Across all panels, relative burst intensity values are positive. Because this measure is defined as vowel peak minus burst peak, this indicates that vowel peaks are generally stronger than burst peaks for both Amharic stop types.

Within this overall pattern of positive values, a second, consistent effect emerges: despite the wide dispersion of speaker means (grey dots), ejective stops show lower relative burst intensity than pulmonic stops across all conditions, especially for alveolar and final ejectives.

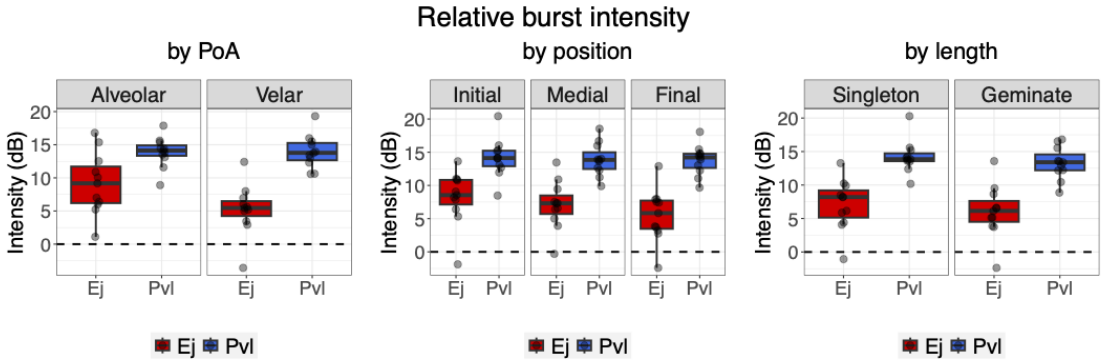

Figure 3.10 Relative burst intensity values (y-axis, in dB) for ejective and pulmonic voiceless stops split by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Dots represent speaker-level means. Values reflect the peak intensity difference between the following vowel and the stop's release burst. Lower values (and especially negative values) indicate a stronger intensity of the burst relative to the vowel.

The results of the statistical analysis revealed a significant main effect of stop type ($F(1, 10.38) = 36.93$, $p < .001$), as well as of PoA, position, and length (PoA: $F(1, 254.03) = 51.13$, $p < .001$; position: $F(2, 249.32) = 8.95$, $p < .001$; length: $F(1, 244.59) = 10.10$, $p < .01$). Furthermore, stop type interacted significantly with both PoA and position (stop type × PoA: $F(1, 254.39) = 60.21$, $p < .001$; stop type × position: $F(2, 249.01) = 6.83$, $p < .001$). The stop type × length interaction was dropped during stepwise selection, and it was not significant in the full model ($F(1, 244.87) = 0.62$, $p = .43$).[8]

[8] Selected model: relBurstIntensity ~ stopType*PoA + stopType*position + length + (stopType | speaker) + (1 | word).

Pairwise comparisons for the interaction with PoA (see **Table 3.6**) confirmed the pattern in **Figure 3.10**: ejective stops showed significantly lower relative burst intensity than pulmonic stops at both the alveolar (Δ = -5.02 dB, $p < .01$) and velar PoA (Δ = -8.65 dB, $p < .001$), with a larger difference at the velar PoA. Within ejectives**,** velars exhibited significantly lower relative burst intensity than alveolar ejectives (velar-alveolar: Δ = -3.50 dB, $p < .001$).

Table 3.6 Pairwise comparisons of relative burst intensity (in dB) by stop type and PoA.

Contrast type	PoA	Stop type	Contrast	Estimate (dB)	SE	DF	t	p	Signif.
By PoA	alv		Ej - Pvl	-5.02	1.15	11.5	-4.36	0.0010	**
By PoA	vel		Ej - Pvl	-8.65	1.15	11.3	-7.55	<.001	***
By stop type		Pvl	vel - alv	0.14	0.35	256.5	0.39	0.6973	
By stop type		Ej	vel - alv	-3.50	0.31	240.9	-11.15	<.001	***

Pairwise comparisons for the interaction with position (see **Table 3.7**) showed that ejective stops had significantly lower relative burst intensity than pulmonic stops in all three word positions: initial (Δ = -5.77 dB, $p < .001$), medial (Δ = -6.70 dB, $p < .001$), and most noticeably in final position (Δ = -8.03 dB, $p < .001$). Within ejectives, final ejectives exhibited the lowest relative burst intensity, differing significantly from both initial (final-initial: Δ = -2.22 dB, $p < .001$) and medial positions (final-medial: Δ = -1.73 dB, $p < .001$).

Table 3.7 Pairwise comparisons of relative burst intensity (in dB) by stop type and position.

Contrast type	Position	Stop type	Contrast	Estimate (dB)	SE	DF	t	p	Signif.
By position	i		Ej - Pvl	-5.77	1.19	13.1	-4.85	<.001	***
By position	m		Ej - Pvl	-6.70	1.16	11.7	-5.80	<.001	***
By position	f		Ej - Pvl	-8.03	1.18	12.5	-6.82	<.001	***
By stop type		Pvl	m - i	0.44	0.45	249.1	0.97	0.5936	
By stop type		Pvl	f - i	0.04	0.46	261.6	0.08	0.9964	
By stop type		Pvl	f - m	-0.40	0.43	257.5	-0.94	0.6173	
By stop type		Ej	m - i	-0.50	0.42	228.8	-1.19	0.4629	
By stop type		Ej	f - i	-2.22	0.42	228.6	-5.24	<.001	***
By stop type		Ej	f - m	-1.73	0.37	235.5	-4.71	<.001	***

Figure 3.11 highlights substantial between-speaker variability in relative burst intensity with speaker mean values ranging from -1.4 dB to 13.3 dB across the dataset. Even so, the stop type contrast is largely consistent across speakers: with the exception of S01 (and a markedly reduced contrast for S09), ejectives show lower relative burst intensity than pulmonic stops.

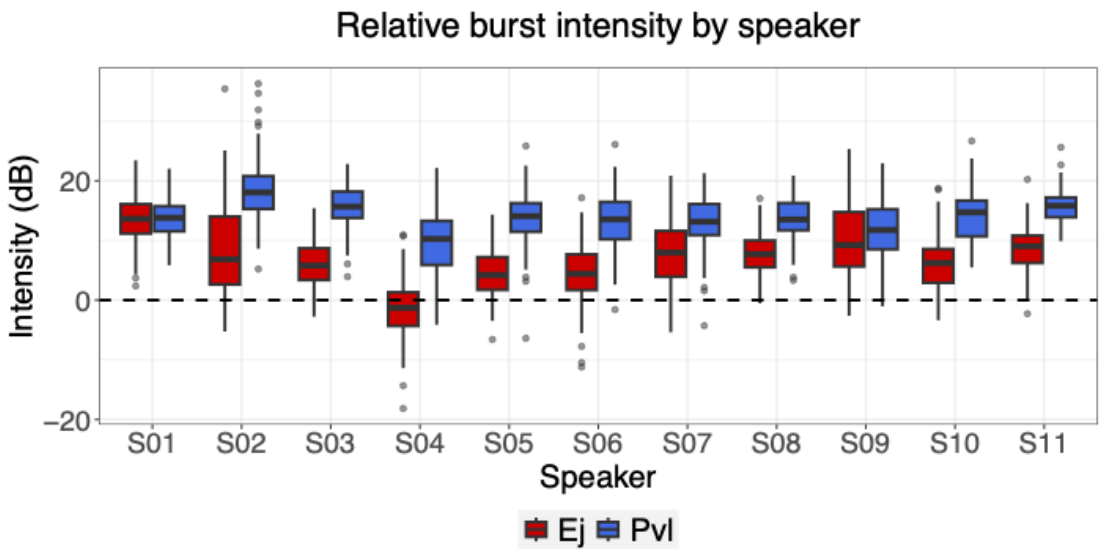

Figure 3.11 Relative burst intensity values (y-axis, in dB) for ejective and pulmonic voiceless stops by speaker. Values reflect the peak intensity difference between the following vowel and the stop's release burst. Lower values (and especially negative values) indicate a stronger intensity of the burst relative to the vowel.

In addition, this figure shows that the pulmonic distributions appear comparatively stable across speakers, whereas ejectives vary more in their central tendency, and in some cases their spread, suggesting greater speaker-specific variation in the relative prominence of the ejective burst. In sum, ejectives patterned as having lower relative burst intensity than pulmonic stops across all conditions, with the largest differences in velars and in final position.

3.3.5. Intensity of the Post-Burst Lag

Figure 3.12 visualizes the intensity of the post-burst lag (PBL), measured in dB from burst offset to vowel onset, for ejective and pulmonic stops, across PoA, word position, and length, each shown in a separate panel. Ejective stops are typically associated with quieter lags than pulmonic stops, as the glottis remains closed briefly after the oral release, delaying the onset of voicing (cf. Kingston, 1984; Wright et al., 2002).

While the PBL for ejectives in this dataset is not completely silent, with average values across panels hovering around 60 dB, they nonetheless consistently exhibit lower intensities than their pulmonic counterparts.

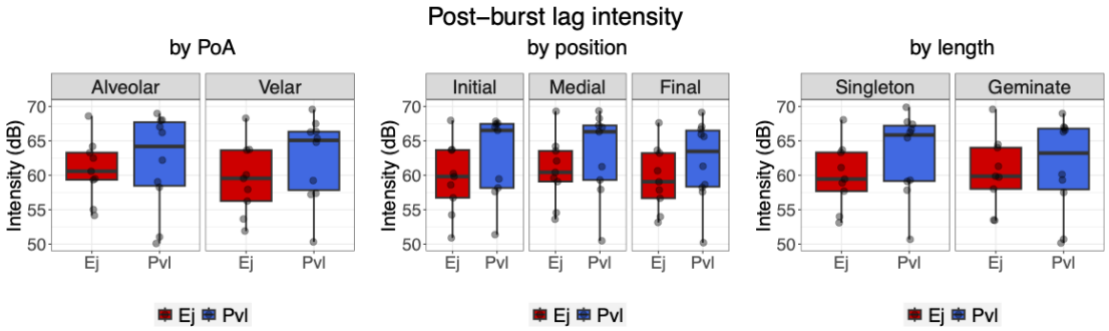

Figure 3.12 Intensity values of the post-burst lag (y-axis, in dB) for ejective and pulmonic voiceless stops split by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Dots represent speaker-level means; lower values indicate a quieter post-burst lag.

The results of the statistical analysis support this visual impression. There was a significant main effect of stop type ($F(1, 10.13) = 22.78$, $p < .001$), along with significant main effects of PoA ($F(1, 246.53) = 147.53$, $p < .001$) and position ($F(2, 237.82) = 57.78$, $p < .001$). Stop type also interacted significantly with both PoA (stop type × PoA: $F(1, 246.48) = 4.50$, $p < .05$) and position (stop type × position: $F(2, 237.82) = 8.45$, $p < .001$). Neither length nor its interaction with stop type significantly improved model fit (length: $F(1, 232.85) = 0.61$, $p = .44$; stop type × length: ($F(1, 233.12) = 0.40$, $p = .53$), and both were removed from the final model.[9]

Pairwise comparisons for the interaction with PoA (see **Table 3.8**) showed that ejective stops had significantly lower PBL intensities than pulmonic stops at both PoAs: alveolar (Δ = -4.25 dB, $p < .001$) and velar (Δ = -3.77 dB, $p < .01$). Within the ejective stop type, PBL intensities were significantly lower following velar stops than alveolars (velar-alveolar: Δ = -1.15 dB, $p < .001$), giving rise to the observed interaction.

[9] Selected model: pblIntensity ~ stopType*PoA + stopType*position + (stopType | speaker) + (1 | word).

Table 3.8 Pairwise comparisons of PBL intensity (in dB) by stop type and PoA.

Contrast type	PoA	Stop type	Contrast	Estimate (dB)	SE	DF	t	p	Signif.
By PoA	alv		Ej - Pvl	-4.25	0.85	10.6	-5.01	<.001	***
By PoA	vel		Ej - Pvl	-3.77	0.85	10.5	-4.45	0.0011	**
By stop type		Pvl	vel - alv	-1.63	0.17	259.5	-9.60	<.001	***
By stop type		Ej	vel - alv	-1.15	0.15	244.0	-7.49	<.001	***

Pairwise comparisons for the interaction with position (see **Table 3.9**) showed that ejective stops had significantly lower PBL intensities than pulmonic stops in all three word positions: initial (Δ = -4.63 dB, $p < .001$), medial (Δ = -4.00 dB, $p < .001$), and final (Δ = -3.39 dB, $p < .01$). Within ejectives, the interaction arises from significantly higher PBL intensities in medial ejectives, compared to both initial (medial-initial: Δ = 1.21 dB, $p < .001$) and final (medial-final: Δ = 1.11 dB, $p < .001$) positions, which did not differ significantly from one another (final-initial: Δ = 0.10, $p = .88$).

Table 3.9 Pairwise comparisons of PBL intensity (in dB) by stop type and position.

Contrast type	Position	Stop type	Contrast	Estimate (dB)	SE	DF	t	p	Signif.
By position	i		Ej - Pvl	-4.63	0.86	11.2	-5.39	<.001	***
By position	m		Ej - Pvl	-4.00	0.85	10.6	-4.70	<.001	***
By position	f		Ej - Pvl	-3.39	0.86	11.0	-3.96	0.0022	**
By stop type		Pvl	m - i	0.57	0.20	246.8	2.79	0.0157	*
By stop type		Pvl	f - i	-1.14	0.22	263.6	-5.14	<.001	***
By stop type		Pvl	f - m	-1.71	0.21	261.8	-8.36	<.001	***
By stop type		Ej	m - i	1.21	0.19	224.2	6.26	<.001	***
By stop type		Ej	f - i	0.10	0.20	223.8	0.48	0.8789	
By stop type		Ej	f - m	-1.11	0.17	236.6	-6.35	<.001	***

In sum, post-burst lag intensity was consistently lower following ejectives than pulmonic voiceless stops. PBL intensity was also shaped by PoA and position, being lowest after velars and higher after medial ejectives.

3.3.6. Intensity Rise Ratio of the Stop-Following Vowel

This measure captures the intensity characteristics of the vowel following the target stop. Intensity rise ratio was calculated as the difference (in dB) between the vowel's mean peak intensity and its intensity at 30 ms into the vowel. Lower values reflect a rapid transition to full intensity, while higher values suggest a more gradual rise.

Figure 3.13 compares intensity rise ratio (in dB) across ejective and pulmonic stops shown by PoA, position, and length in separate panels. Two key observations emerge from the figure: first, there does not appear to be a clear pattern distinguishing the stop types based on the intensity rise ratio of the following vowels. Second, speaker variability is striking across the board, as can be seen by the wide spread of the individual speaker means (grey dots), with values ranging from essentially 0 to 10 across all factors.

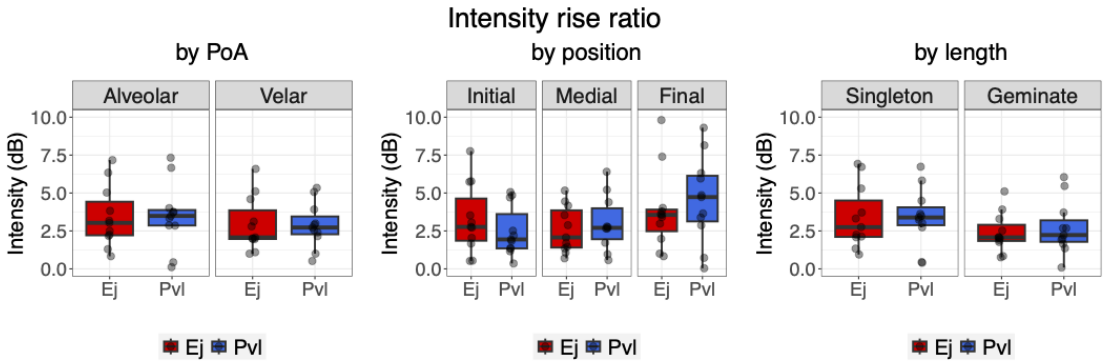

Figure 3.13 Raw intensity rise ratio values (y-axis, in dB) for ejective and pulmonic voiceless stops, split by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Dots represent speaker-level means; higher values indicate a slower rise to full intensity in the vowel following the stop.

In fact, the results of the statistical analysis, using log-transformed intensity rise to meet model assumptions, showed significant main effects of PoA, position, and length (PoA: $F(1, 240.72) = 5.76$, $p < .05$; position: ($F(2, 238.66) = 24.59$, $p < .001$); length: $F(1, 240.87) = 9.51$, $p < .01$) but no main effect of stop type ($F(1, 10.64) = 0.19$, $p = .67$).

There was, however, a significant interaction between stop type and position ($F(2, 239.21) = 6.02$, $p < .01$). No other interactions with stop type were significant (stop type × PoA: ($F(1, 239.64) = 0.10$, $p = .76$; stop type × length: ($F(1, 238.74) = 0.05$, $p = .83$).[10]

Table 3.10 Pairwise comparisons of intensity rise ratio (modelled on the log scale) by stop type and position, with back-transformed absolute differences in dB.

Contrast type	Position	Stop type	Contrast	Estimate (log)	SE	DF	t	p	Signif.	Δ (dB)
By position	i		Ej - Pvl	0.33	0.23	13.8	1.48	0.1608		0.77
By position	m		Ej - Pvl	0.03	0.22	12.0	0.14	0.8903		0.08
By position	f		Ej - Pvl	-0.09	0.22	13.0	-0.39	0.6996		-0.30
By stop type		Pvl	m - i	0.24	0.09	244.2	2.61	0.0259	*	0.52
By stop type		Pvl	f - i	0.61	0.09	243.9	6.74	<.001	***	1.63
By stop type		Pvl	f - m	0.37	0.09	242.2	4.37	<.001	***	1.11
By stop type		Ej	m - i	-0.07	0.09	243.1	-0.78	0.7150		-0.17
By stop type		Ej	f - i	0.19	0.09	241.7	2.17	0.0788	.	0.56
By stop type		Ej	f - m	0.25	0.07	245.3	3.41	0.0022	**	0.73

Pairwise comparisons for the interaction with position revealed no significant differences in intensity rise ratio between ejective and pulmonic stops across any position (initial: Δ = 0.77 dB, $p = .16$; medial: Δ = 0.08 dB, $p = .89$; final: Δ = -0.30 dB, $p = .70$). The interaction arises, because while intensity rise ratio for ejectives remained relatively stable across positions, it increased substantially for pulmonic stops from initial to final position.

Position-related differences in intensity rise ratio were primarily observed among pulmonic stops (see **Table 3.10**). Within ejective stops, results were mixed. There was no significant difference between medial and initial positions (medial-initial: Δ = -0.17 dB, $p = .72$), or the final-initial comparison (final-initial: Δ = 0.56 dB, $p = .08$). The only significant contrast was between final and medial positions (final-medial: Δ = 0.73 dB, $p < .01$), indicating that vowels following final ejectives exhibited a slower rise to peak intensity compared to those following medial ejectives.

[10] Selected model: intensityRiseRatio_log ~ PoA + length + stopType*position + (stopType | speaker) + (1 | word).

In sum, the intensity rise ratio did not significantly differentiate ejectives from pulmonic stops, although within the ejective series, final stops showed a slower rise than medial stops.

3.3.7. Summary of the Acoustics Results

The acoustic analysis revealed several robust patterns distinguishing Amharic ejectives from pulmonic voiceless stops. Ejectives had significantly longer burst durations and VOTs than pulmonic stops, with these effects observed consistently across PoA, word position, and length. Relative burst intensity was the most robust acoustic differentiator: ejectives had markedly lower relative burst intensity than pulmonic voiceless stops across all contexts and especially at the velar place. Post-burst lags in ejectives were characteristically lower in intensity, although not fully silent. However, the stop types did not differ significantly in either their closure duration or the intensity rise ratio of the following vowel.

3.3.8. Interim Discussion of the Acoustic Findings

This section briefly compares the acoustic results to previous work on Amharic and situates the language's ejectives within broader cross-linguistic patterns.

Closure patterns. The finding that closure duration did not differ between stop types aligns with previous work on Amharic (Derib Ado, 2011) and with cross-linguistic studies showing that ejectives often pattern similarly to pulmonic stops in this respect (Gordon & Applebaum, 2006; Lindau, 1984; Maddieson & Ladefoged, 1993). The patterns concerning length, namely, that geminate stops consistently showed longer closures than singletons, confirm universal characteristics of the singleton-geminate contrast in stops (Ladefoged & Maddieson, 1996; Maddieson, 1985).

Release patterns and burst duration. Among all ejective tokens, 3% were unreleased. These were mostly alveolar and word-final, an unusual pattern cross-linguistically, given that alveolars are typically more resistant to reduction (Kirchner, 2013; Lavoie, 2001). The preference for final-position unreleased stops, however, is widely attested across languages (Crystal & House, 1988; Gussmann, 2007; Jun, 2004; Kehrein, 2013).

The particular preference for speaker S01 to weaken ejectives is unexpected, as ejectives are typically considered phonetically robust and resistant to reduction (Disner & Maddieson, 1984; Kingston, 1984; Steriade, 2001; Wright, 2004). Notably, this speaker also exhibited high rates of creak, a common cue of lenition, suggesting a link between ejective weakening and creaky voice.

For released bursts, the strong effect of longer burst durations for ejectives in final position was primarily driven by four speakers, although the trend was evident across the speaker set. One possible explanation is that the experimental setup, characterized by a repetitive task and short, formulaic sentences, may have encouraged prosodic final lengthening, in addition to many participants quickly inferring the study's focus on ejectives. These conditions may have encouraged hyperarticulation in final position, contributing to unusually long burst durations. As such, these patterns may reflect task effects rather than intrinsic phonetic properties of final ejectives, with individual speakers responding to these conditions with more exaggerated articulatory behaviors than others.

Overall, this may help explain the discrepancies with Seid et al. (2009) and Derib Ado (2011), whose findings on burst duration differed from those reported here. A final qualitative observation: several speakers, especially when producing alveolar ejectives, exhibited release bursts with slightly affricated characteristics. This complicated the segmentation of the burst offset and may also have contributed to some of the observed variability in ejective burst durations.

VOT patterns. VOT served as a reliable distinguishing cue for ejectives. Generally speaking, the fact that ejectives here showed longer VOT overall than pulmonic stops does not align with previous findings on Amharic ejectives (Ado, 2011; Seid et al., 2009). Viewed cross-linguistically, the average VOT observed in this study (72 ms) places these ejectives toward the longer end of the range reported for ejective stops, comparable to values described for languages such as Tsilhqut'in, Navajo, and Tigrinya (Ham, 2007; Kingston, 1984; Lindau, 1984).

Burst intensity and PoA differences. The results confirmed that stop types are robustly differentiated by their intensity properties. Although relative values were positive for both stop types, showing stronger vowel than burst peaks, ejective stops had lower relative values than pulmonic stops across conditions, consistent with Derib Ado (2011). Inspection of the components of the measure revealed no systematic difference in vowel peak intensity across stop types (aside from a weak position-specific effect in initial position), but consistently higher absolute burst peak intensities for ejectives. The relative

pattern therefore reflects a direct difference in absolute burst peak intensity between stop types, diverging from Seid et al. (2009), who examined only absolute burst intensity and found no such difference.

The same reasoning applies to PoA-related differences among ejectives: lower relative burst intensity for velars compared to alveolars reflects stronger absolute burst releases, since vowel peaks remain stable across PoA. This pattern accords with Derib Ado (2011) and with cross-linguistic observations that link facilitated pressure buildup to the smaller supraglottal cavity behind velars.

Post-burst lag and intensity rise ratio. Importantly, Amharic ejectives were followed by the characteristic "silent gap" – an interval of low intensity following release burst. This indicates that the glottis was closed at the moment of oral release, confirming a key feature associated with ejective production. The fact that PBL was never fully silent following ejectives in this dataset may be attributed to the nature of the rt-MRI-derived acoustic signal, which remains relatively noisy despite post-processing.

In this study, vowels following stops exhibited no distinct effects of stop type as measured by the intensity rise ratio. The average rise ratio observed was 3.2 dB (SD = 1.9), indicating a rapid rise to full intensity in the following vowels and thus suggesting predominantly modal voice quality. This indication aligns with the findings of Negesse (2023) but contrasts with those of Derib Ado (2011) and Seid et al. (2009). In cross-linguistic terms, these features constitute further similarities with ejectives in Tigrinya (see Ham, 2008 for an overview).

Speaker-level variation and exceptions. Acoustic patterns were generally stable across speakers. As shown in **Table C.38**, provided in Appendix C, only a few speakers diverged notably from the group-level trends (highlighted in bold). Two speakers in particular, S01 and S04, represented the most extreme values for relative burst intensity, though in opposite directions. S04 was the only speaker with a negative average relative intensity, indicating the burst peaks exceeded vowel peaks for this speaker.

By contrast, S01 exhibited unusually high relative burst intensity values, consistent with especially weak burst releases. This speaker further stood out due to short VOT values and a high proportion of unreleased stops. These speaker-specific deviations are best understood as exceptions to otherwise robust and consistent acoustic patterns across speakers.

Context-level variation. Both VOT and intensity rise ratio occupied consistent positions within the cross-linguistic range: VOT values were generally long, and rise ratios indicated a rapid rise to full vowel intensity, in line with predominantly modal phonation. The few contextual effects observed included markedly longer closure durations for geminates, longer burst durations for final-position ejectives, shorter VOTs for medial ejectives, and lower relative burst intensity for velars, which have been discussed in earlier sections. Yet, considering the full set of acoustic measures together, no single contextual factor (PoA, word position, or length) exerted a systematic influence on the overall realization of ejectives.

Including bilabial ejectives in the analysis might have yielded PoA-based distinctions of the kind reported by Negesse (2023), who found bilabials to exhibit a distinct acoustic profile relative to alveolars and velars. In the present dataset, however, the profile of ejectives, summarized in **Table C.39** provided in Appendix C, showed a high degree of uniformity across contexts.

Taken together, the acoustic evidence points to a broadly uniform realization of ejectives across speakers, PoA, word position, and length. Having established this profile, we now turn to the analysis of the underlying articulatory mechanisms involved in their production.

3.4. Articulatory Analysis

The aim of the articulatory analysis was to examine laryngeal and supralaryngeal activity in ejective stop production. The following section presents the results of this analysis, focusing on three key parameters, *delta larynx, delta cavity*, and *delta hyperph*, each of which reflect different articulatory strategies. For a detailed description of the methodology used to attain these measures, see Chapter 2, Section 2.7.

3.4.1. Larynx Raising

Larynx raising is traditionally viewed as the key articulatory component of ejective production (Catford, 1977). Thus, the first articulatory analysis in this section examined larynx movement via the measure *delta larynx,* which quantifies the vertical displacement (in mm) of the larynx during stop closure.

Figure 3.14 compares *delta larynx* between ejective and pulmonic stops across PoA, position and length, each shown in separate panels. Positive *delta larynx* values reflect larynx raising, while negative *delta larynx* values indicate larynx lowering.

The figure reveals a clear pattern: ejectives consistently show positive delta values, indicating systematic larynx raising across all conditions. In contrast, pulmonic stops fluctuate around zero, indicating minimal displacement overall and no consistent direction of larynx movement, unlike the systematic raising observed for ejectives.

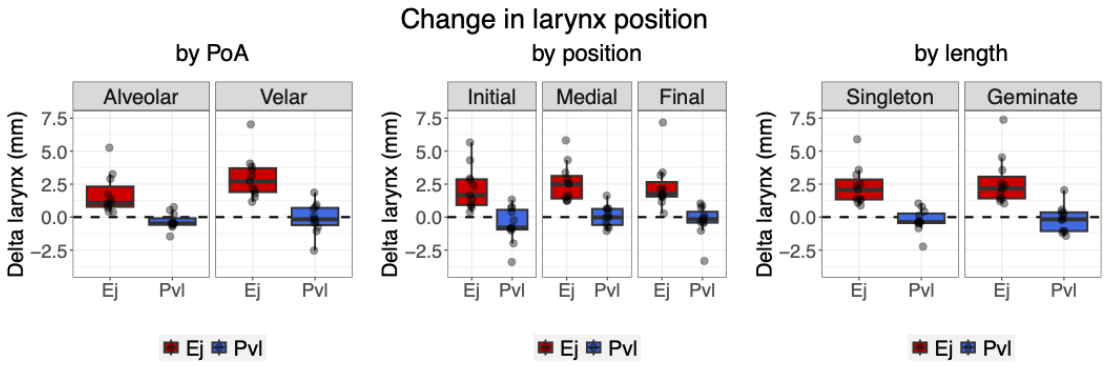

Figure 3.14 Delta larynx values (y-axis, in mm) of ejective and pulmonic voiceless stops, split by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Delta larynx is measured as the difference between the larynx's position at the onset of stop closure and its position at closure offset. Dots represent speaker-level means; positive numbers indicate larynx raising, negative numbers indicate larynx lowering.

The statistical analysis corroborates this visual impression. Using log-transformed *delta larynx* in order to meet model assumptions, significant main effects on *delta larynx* were found for stop type ($F(1, 10.20) = 39.67$, $p < .001$), PoA ($F(1, 246.68) = 84.16$, $p < .001$), position ($F(2, 244.61) = 18.75$, $p < .001$), and length ($F(1, 247.15) = 7.41$, $p < .01$). In addition, a significant stop type × PoA interaction ($F(1, 246.50) = 26.11$, $p < .001$) revealed that the relationship between ejectives and larynx raising varies by PoA. Neither the stop type × position nor the stop type × length interactions significantly improved model fit and were excluded from the final model (stop type × position: $F(2, 240.88) = 2.38$, $p = .10$; stop type × length: $F(1, 246.98) = 1.57$, $p = .21$).[11]

As per **Table 3.11**, pairwise comparisons for the interaction with PoA confirmed that ejective stops exhibited significantly greater larynx raising than pulmonic stops at both PoAs: alveolar ($\Delta = 1.23$ mm, $p < .001$) and velar ($\Delta = 2.27$ mm, $p < .001$).

[11] Selected model: deltaLarynx_log ~ stopType*PoA + position + length + (stopType | speaker) + (1 | word).

Within the ejective stop type, larynx raising was significantly greater at the velar PoA than at the alveolar (velar-alveolar: Δ = 1.16 mm, $p < .001$). This pattern drives the significant stop type × PoA interaction and is illustrated in the leftmost panel of **Figure 3.14**.

Table 3.11 Pairwise comparisons of delta larynx (modelled on the log scale) by stop type and PoA, with back-transformed absolute differences in mm.

Contrast type	PoA	Stop type	Contrast	Estimate (log)	SE	DF	t	p	Signif.	Δ (mm)
By PoA	alv		Ej - Pvl	0.90	0.17	11.0	5.26	<.001	***	1.23
By PoA	vel		Ej - Pvl	1.22	0.17	10.9	7.12	<.001	***	2.27
By stop type		Pvl	vel - alv	0.13	0.05	245.0	2.78	0.0059	**	0.11
By stop type		Ej	vel - alv	0.45	0.04	246.0	10.58	<.001	***	1.16

While the statistical results showed consistent larynx raising for ejectives, a small number of tokens deviated from this pattern. Specifically, 115 out of 1417 ejective tokens exhibited negative *delta larynx* values. These cases are addressed in the discussion below.

Figure 3.15 shows *delta larynx* values for both stop types by speaker. This figure emphasizes that while ejectives are consistently produced with larynx raising, the magnitude of this upward movement varies substantially across speakers. Speaker S06 shows particularly large upward displacements, while other speakers, such as S01, show mean smaller values (incidentally, S01 was also the speaker with the highest proportion of unreleased ejective stops, see Section 3.3.2 above). For almost all speakers, the difference in *delta larynx* between ejective and pulmonic stops is clearly positive. The one exception is S02, for whom the contrast between stop types is reduced due to elevating larynx movement occurring during some pulmonic tokens as well.

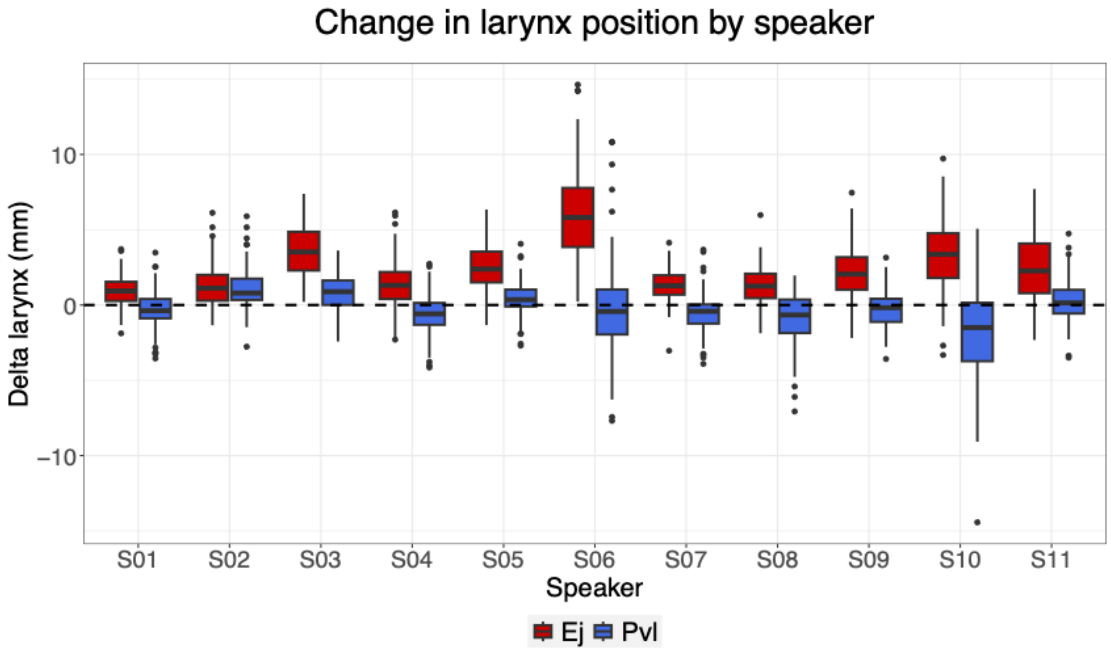

Figure 3.15 Delta larynx values (y-axis, in mm) by stop type (ejective, pulmonic voiceless) and speaker. Delta larynx is measured as the difference between the larynx's position at the onset of stop closure and its position at closure offset. Positive numbers indicate larynx raising negative numbers indicate larynx lowering.

This inter-speaker variability is further illustrated in **Table 3.12**, which summarizes mean *delta larynx* values and standard deviations for each speaker. It highlights the exceptional magnitude of larynx raising in S06 relative to all other speakers.

Table 3.12 Means (*M*) and standard deviations (*SD*, in parentheses) of delta larynx (in mm) values of ejectives for each speaker.

	Speaker										
	S01	**S02**	**S03**	**S04**	**S05**	**S06**	**S07**	**S08**	**S09**	**S10**	**S11**
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Delta larynx (mm)	1.0 (0.9)	1.3 (1.4)	3.6 (1.6)	1.4 (1.6)	2.5 (1.3)	6.2 (3.0)	1.3 (1.0)	1.3 (1.2)	2.1 (1.7)	3.5 (2.3)	2.4 (2.2)

In sum, larynx displacement robustly distinguished stop types, with larynx raising serving as a consistent hallmark of ejective production. Within ejectives, its magnitude varied significantly only by PoA, being greatest for velars, and showed substantial inter-speaker variability.

3.4.2. Intraoral Cavity Reduction

The *IPA Handbook* (1999) describes ejective production simply as compression of the intraoral cavity in order to increase air pressure. Thus, the second articulatory analysis in this section examined the changes in supralaryngeal cavity size (in mm^2) between closure onset and offset using the measure *delta cavity*. Because this measure quantifies only the supralaryngeal portion of the vocal tract from above the larynx (see Chapter 2, Section 2.7.2 for more details), it reflects cavity size changes independently of vertical larynx displacement.

Figure 3.16 displays *delta cavity* values for ejective and pulmonic stops across PoA, position, and length, each in a separate panel. Positive *delta cavity* values indicate a reduction in cavity size over the closure; negative values indicate expansion. The figure shows that both stop types deviate from zero, which suggests that the cavity changes in size during stop closure, but in systematically different directions for ejectives vs. pulmonic stops.

For ejectives, the consistently positive *delta cavity* values across panels indicate that the cavity becomes smaller by closure offset, consistent with cavity size reduction. In contrast, pulmonic stops generally show negative delta values, indicating cavity expansion during closure.

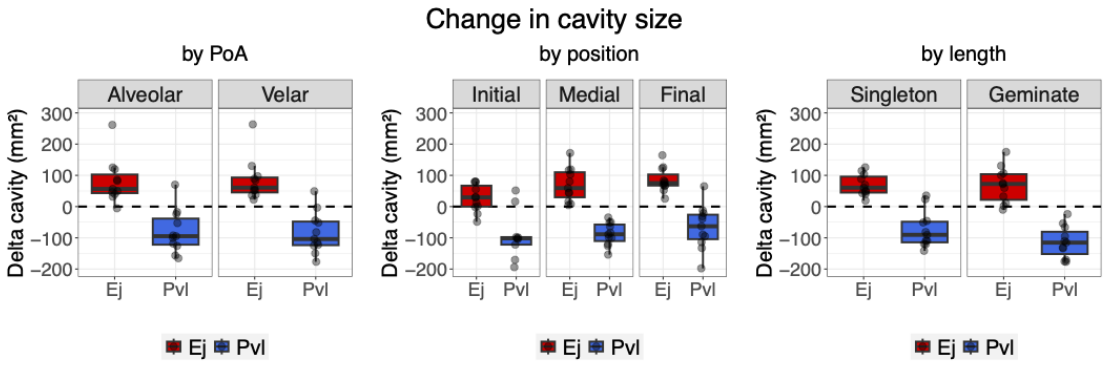

Figure 3.16 Delta cavity values (y-axis, in mm^2) for ejective and pulmonic voiceless stops, split by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Delta cavity is measured as the difference between the cavity's size at the onset of stop closure and its size at closure offset. Dots represent speaker-level means; positive numbers indicate cavity reduction, negative numbers indicate cavity expansion.

The results of the statistical analysis reveal a significant main effect of stop type ($F(1, 13.22) = 91.72$, $p < .001$) and length ($F(1, 244.60) = 17.29$, $p < .001$). There was also a significant stop type × length ($F(1, 244.72) = 4.91$, $p < .05$) interaction. PoA and position did not significantly improve model fit (PoA: $F(1, 265.06) = 1.31$, $p = .25$; position: $F(2, 271.75) = 2.47$, $p = .09$), nor did their interactions with stop type (stop type × PoA: $F(1, 262.59) = 0.51$, $p = .48$; stop type × position: $F(2, 257.81) = 0.82$, $p = .44$). These factors were therefore excluded from the final model.[12]

Table 3.13 Pairwise comparisons of delta cavity (in mm^2) by stop type and length.

Contrast type	length	Stop type	Contrast	Estimate (mm^2)	SE	DF	t	p	Signif.
By length	s		Pvl - Ej	-147.99	17.33	13.7	-8.54	<.001	***
By length	g		Pvl - Ej	-178.84	19.49	21.8	-9.18	<.001	***
By stop type		Ej	g - s	-14.97	9.74	228.3	-1.54	0.1258	
By stop type		Pvl	g - s	-45.81	10.45	225.3	-4.38	<.001	***

Pairwise comparisons for the stop type × length interaction, presented in **Table 3.13**, showed that ejective stops differed significantly from pulmonic stops in cavity size change for both singletons ($\Delta = 147.99$ mm², $p < .001$) and geminates ($\Delta = 178.84$ mm², $p < .001$), indicating a stable ejective-pulmonic contrast across length. Within stop types, no length effect was found for ejectives (geminate-singleton: $\Delta = 14.97$ mm², $p = .13$). The interaction is driven by the singleton-geminate difference in pulmonics.

Figure 3.17 and **Table 3.14** together illustrate the speaker-wise patterns of cavity size change. Across speakers, mean *delta cavity* values are consistently positive, confirming that cavity reduction is a regular feature of ejective production.

[12] Selected model: deltaCavity ~ stopType*length + (stopType | speaker) + (1 | V1V2) + (1 | word).

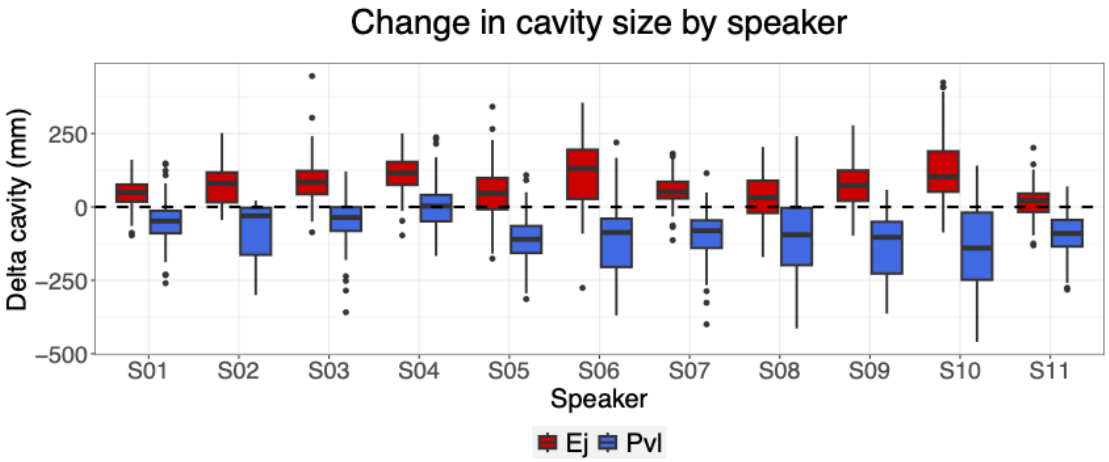

Figure 3.17 Delta cavity values (y-axis, in mm^2) by stop type (ejective, pulmonic voiceless) and speaker. Delta cavity is measured as the difference between the cavity's size at the onset of stop closure and its size at closure offset. Positive numbers indicate cavity reduction; negative numbers indicate cavity expansion.

Speaker variability in this measure appears more constrained than for larynx raising, but two speakers stand out: S06 – the speaker with the greatest range of larynx movement, as seen in Section 3.4.1 above – shows markedly greater *delta cavity* values along with significant variability (*SD*). In contrast, S11 is notable for very low average values, suggesting that any cavity changes for this speaker are very small.

Table 3.14 Means (*M*) and standard deviations (*SD*, in parentheses) of delta cavity (in mm^2) values of ejectives for each speaker.

	Speaker										
	S01	**S02**	**S03**	**S04**	**S05**	**S06**	**S07**	**S08**	**S09**	**S10**	**S11**
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Delta cavity (mm^2)	42.9 (50.7)	78.5 (74.8)	91.4 (73.8)	112.3 (67.8)	46.2 (86.9)	117.1 (138.4)	57.6 (58.3)	31.0 (81.4)	70.4 (73.9)	127.2 (111.3)	14.9 (55.4)

Overall, changes in cavity size significantly distinguished stop types: ejectives were produced with cavity reduction, while pulmonic stops tended to involve cavity expansion. None of the other factors examined exerted a systematic influence on the degree of cavity reduction within the ejectives.

3.4.3. Constriction of the Hyperpharyngeal Region

A number of studies have proposed that additional pharyngeal constriction plays an important role in ejective production (e.g., Demolin, 2002; Hermes et al., 2016; Kingston, 1984). Accordingly, the third analysis in this section examined constriction in the hyperpharyngeal region, defined as the region extending from the tip of the epiglottis to the lowest point of the open velum. Changes in constriction within this region during stop closure were quantified using *delta hyperph*, a normalized measure of the degree of constriction.

Figure 3.18 presents *delta hyperph* values for both stop types across PoA, word position, and length, each displayed in a separate panel. As the figure shows, ejective and pulmonic stops exhibit opposing patterns: ejectives show increased constriction (positive delta values), while pulmonic stops tend to show reduced constriction (negative delta values).

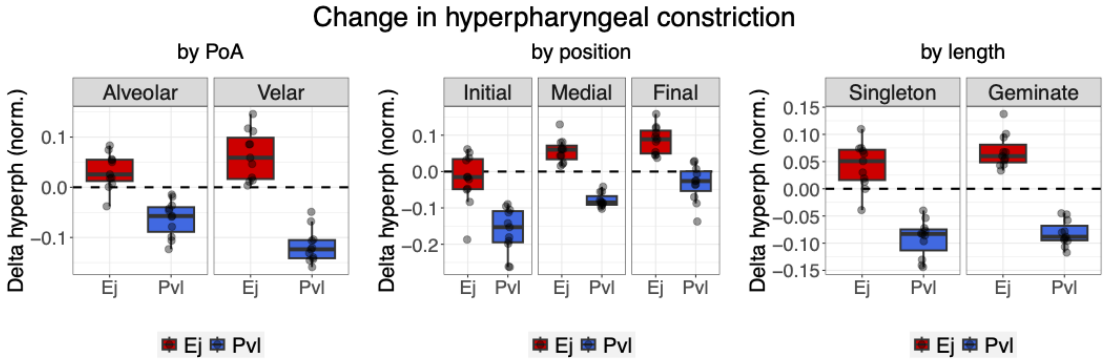

Figure 3.18 Delta hyperph values (y-axis, normalized) for ejective and pulmonic voiceless stops, split by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Delta hyperph is measured as the difference between the degree of hyperpharyngeal constriction at the onset of stop closure and its constriction at closure offset. The black dashed line represents the zero value for the y-axis: observations above it indicate hyperpharyngeal constriction increased; below it indicates hyperpharyngeal constriction relaxed. Dots represent speaker-level means.

This observed pattern is further supported by statistical analysis, which revealed a highly significant main effect of stop type ($F(1, 18.64) = 189.02$, $p < .001$), as well as significant main effects of PoA ($F(1, 241.22) = 6.66$, $p < .05$), position ($F(2, 241.05) = 13.41$, $p < .001$), and length ($F(1, 241.39) = 22.77$, $p < .001$). Stop type also strongly interacted with both PoA (stop type × PoA: $F(1, 241.86) = 54.72$, $p < .001$) and position (stop type × position: $F(2, 239.52) = 7.12$, $p < .001$).

This indicates that the degree of change in hyperpharyngeal constriction varied across both PoA and position depending on stop type. The interaction with length, however, did not significantly improve model fit and was therefore excluded from the final model (stop type × length: $F(1, 240.18) = 0.36$, $p = .55$).[13]

Table 3.15 Pairwise comparisons of delta hyperph (normalized) by stop type and PoA.

Contrast type	PoA	Stop type	Contrast	Estimate	SE	DF	t	p	Signif.
By PoA	alv		Ej - Pvl	0.08	0.01	34.6	7.64	<.001	***
By PoA	vel		Ej - Pvl	0.16	0.01	36.1	15.46	<.001	***
By stop type		Pvl	vel - alv	-0.06	0.01	237.5	-6.78	<.001	***
By stop type		Ej	vel - alv	0.03	0.01	238.3	3.59	<.001	***

Pairwise comparisons for the interaction with PoA (see **Table 3.15**) revealed significant differences in hyperpharyngeal activity between ejective and pulmonic stops at both the alveolar ($\Delta = 0.08$, $p < .001$) and, more markedly, at the velar PoA ($\Delta = 0.16$, $p < .001$). Within ejectives, velars showed a significantly greater increase in constriction than alveolars (velar-alveolar: $\Delta = 0.03$, $p < .001$).

Table 3.16 Pairwise comparisons of delta hyperph (normalized) by stop type and position.

Contrast type	Position	Stop type	Contrast	Estimate	SE	DF	t	p	Signif.
By position	i		Ej - Pvl	0.13	0.013	72.1	10.54	<.001	***
By position	m		Ej - Pvl	0.13	0.011	41.6	12.50	<.001	***
By position	f		Ej - Pvl	0.09	0.012	58.7	7.48	<.001	***
By stop type		Pvl	m - i	0.04	0.011	238.4	3.23	0.0040	**
By stop type		Pvl	f - i	0.02	0.012	237.3	1.93	0.1319	
By stop type		Pvl	f - m	-0.01	0.011	237.2	-1.25	0.4236	
By stop type		Ej	m - i	0.04	0.010	235.9	3.49	0.0017	**
By stop type		Ej	f - i	-0.02	0.012	234.6	-1.92	0.1359	
By stop type		Ej	f - m	-0.06	0.011	236.7	-5.58	<.001	***

[13] Selected model: deltaHyperph ~ stopType*PoA + stopType*position + length + (stopType | speaker) + (1 | V1V2) + (1 | word).

Pairwise comparisons for the interaction with position (see **Table 3.16**) revealed that ejective stops exhibited significantly greater increase in hyperpharyngeal constriction than pulmonic stops across all three positions: initial ($\Delta = 0.13$, $p < .001$), medial ($\Delta = 0.13$, $p < .001$), and final ($\Delta = 0.09$, $p < .001$). Within the ejective stop type, medial ejectives increased their constriction significantly compared to initial ejectives (medial-initial: $\Delta = 0.04$, $p < .01$), and even more so compared to final ejectives (medial-final: $\Delta = 0.06$, $p < .001$). No significant difference was found between ejectives at initial and final positions (initial-final: $\Delta = 0.02$, $p = .14$).

In sum, the results showed systematic differences in hyperpharyngeal constriction between ejective and pulmonic stops: ejectives involved increasing constriction, whereas pulmonic stops showed constriction reduction. Velar ejectives exhibited greater increases in constriction than alveolar ejectives, and medial compared to initial and final ejectives.

3.4.4. Summary of the Articulation Results

Ejective stops in Amharic were consistently produced with upward larynx movement, supralaryngeal cavity reduction, and increased hyperpharyngeal constriction during closure. These patterns held across both PoAs, all word positions, and length categories. Among the examined factors, PoA had the strongest influence, with velar ejectives showing more extreme larynx displacement and hyperpharyngeal constriction than alveolars. Word position also modulated hyperpharyngeal constriction, with medial ejectives showing the greatest degree of constriction. No significant effects were found for the singleton-geminate contrast in ejectives. Cavity-related measures showed less inter-speaker variability than larynx raising, which varied substantially across speakers.

3.4.5. Interim Discussion of Selected Articulatory Findings

A comprehensive interpretation of the articulatory results will follow in the general discussion at the end of this chapter. However, several findings warrant brief commentary at this stage. The points discussed below, such as inter-speaker variability in larynx raising and the presence of negative delta values, raise methodological or interpretive issues that benefit from immediate clarification, as they contribute important nuance to the analysis.

Larynx raising and speaker variability. This study found substantial inter-speaker variability in the extent of larynx raising, consistent with earlier observations by Kingston (1984) and Bückins et al. (2018). While Kingston attributed these differences to variation in speech style, Bückins et al. suggested anatomical factors. To assess the plausibility of anatomical explanations more directly, the present study conducted a supplementary analysis, the details of which are provided in Appendix D, examining the relationship between pharynx length and larynx displacement.

Partially supporting Bückins et al. (2018), the results suggested that anatomical factors, specifically pharynx length, influence the range of larynx movement once a certain threshold is exceeded, as exemplified by speaker S06. Among speakers with more homogeneous pharynx sizes, however, inter-speaker variation is more likely attributable to individual articulatory strategies, potentially including stylistic differences as proposed by Kingston (1984).

Negative delta larynx values. A very small subset of ejective tokens (115 out of 1417) exhibited negative *delta larynx* values in this study. These do not reflect active larynx lowering. Rather, they result from an initial downward movement followed by larynx raising that does not exceed the height at closure onset, yielding a net negative displacement. Visual inspection of the relevant tokens, made possible by their small number, supports this interpretation. The frequency with which such initial lowering gestures occur in this study even in ejectives that ultimately exhibit a net positive displacement cannot yet be determined, although such information is theoretically retrievable. It will be interesting to address this issue in future analyses, since comparable initial downward larynx movement in ejective production has been reported by Bückins et al. (2018) and Brandt and Simpson (2021).

From an aerodynamic perspective, the occurrence of negative delta values presents a challenge: to increase intraoral pressure, the larynx must be higher at closure offset than at onset. A possible explanation is a slight delay in glottal closure relative to oral closure. In this scenario, even limited upward movement, once it takes place and the glottis is shut, could still contribute to intraoral pressure, although this would likely not be very high. The present data does not, however, provide direct insight into glottal timing, so for now, this interpretation cannot be verified. Having examined the individual articulatory properties of Amharic ejectives and discussed some key points in the findings, the next section turns to the analysis of how these components relate to one another.

3.5. Articulatory Relationships and Acoustic Correlates

The purpose of this final set of analyses was twofold: to examine how the articulatory measures are coordinated, and to evaluate how each contributes to absolute burst intensity, used here as a proxy for intraoral pressure. Absolute burst intensity (measured as the mean intensity over the burst interval) was chosen because, although the relative measure facilitates cross-speaker and cross-language comparison, it is not as suitable for understanding how articulatory actions directly influence burst intensity, which itself only indirectly reflects pressure buildup during stop production.

In principle, these questions could be addressed in a single multiple-regression model predicting burst intensity from all articulatory predictors at once. However, once PoA, word position, and length (and their interactions) are included, such a model becomes difficult to interpret and to relate to specific hypotheses. The analyses were therefore kept targeted: coordination between articulatory measures was examined in separate models (larynx raising vs. supralaryngeal cavity reduction; hyperpharyngeal constriction vs. supralaryngeal cavity reduction), and their contributions to absolute burst intensity were evaluated in parallel rather than jointly.

3.5.1. Linking Larynx Raising and Supralaryngeal Cavity Change

The preceding analyses showed that both vertical larynx displacement and supralaryngeal cavity reduction contributed to ejective production. The next analysis examined whether the magnitude of vertical larynx displacement systematically influenced the degree of cavity size reduction.

Figure 3.19 presents scatter plots showing the relationship between *delta larynx* (x-axis, in mm) and *delta cavity* (y-axis, in mm^2) across ejective tokens, separated by PoA, word position, and length. The dashed lines mark zero for both axes: values to the left of the zero line on the x-axis reflect larynx lowering, while values above the zero line on the y-axis indicate a decrease in cavity size.

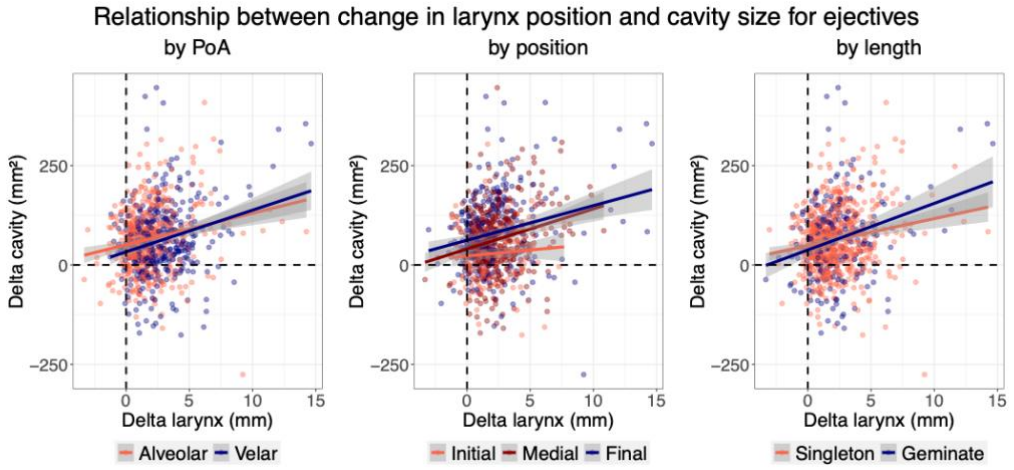

Figure 3.19 Scatter plot showing the relationship between delta larynx (x-axis, in mm) and delta cavity (y-axis, in mm^2) for ejectives, separated by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Superimposed regression lines indicate the strength and direction of the relationship, with shaded regions representing 95% confidence intervals. The black dashed lines represent the zero values for both axes: observations below the zero y-axis indicate an increase in cavity size, while observations to the left of the zero x-axis indicate larynx lowering over the stop closure. Dots represent individual tokens.

Most tokens fall in the upper right quadrant, indicating that larynx raising most frequently co-occurs with cavity size reduction. A secondary cluster appears in the lower right quadrant, showing that larynx raising sometimes occurs with cavity size expansion. Tokens with larynx lowering (left quadrants) are very rare and are more likely to show some degree of cavity reduction (top left quadrant). The superimposed regression lines suggest a weak positive relationship: greater larynx elevation is associated with greater cavity size reduction. The only exception is initial ejectives, which show a flat line, suggesting that cavity reduction plays little role in this position.

Statistical analysis confirmed the visual impression by revealing a modest but significant effect of *delta larynx* ($F(1, 7.07) = 7.84$, $p < .05$) on *delta cavity*, such that greater larynx raising was associated with a greater decrease in cavity size. Model comparison showed that PoA ($F(1, 167.97) = 2.86$, $p = .09$), position ($F(2, 134.54) = 2.35$, $p = .10$), and length ($F(1, 134.82) = 1.63$, $p = .20$), and their interactions with *delta larynx* (delta larynx × PoA: $F(1, 825.02) = 2.01$, $p = .16$; delta larynx × position: $F(2, 839.20) = 0.26$, $p = .77$; delta larynx × length: $F(1, 825.91) = 1.10$, $p = .30$) did not improve model fit. Accordingly, these terms were removed from the final model.[14] In sum, larynx raising

[14] Selected model: deltaCavity ~ deltaLarynx + (deltaLarynx | speaker) + (1 | V1V2) + (1 | word).

was modestly but consistently associated with cavity size reduction, and this relationship was not moderated by PoA, position, or length.

3.5.2. Linking Hyperpharyngeal Constriction and Supralaryngeal Cavity Change

Earlier results showed that ejectives involved both overall cavity reduction and hyperpharyngeal constriction. This next analysis tested whether variation in hyperpharyngeal constriction predicted changes in supralaryngeal cavity size. A lack of association would suggest that cavity compression is driven primarily by more anterior constrictions.

Figure 3.20 plots *delta hyperph* (x-axis, normalized) against *delta cavity* (y-axis, in mm^2) across ejective tokens, separated by PoA, word position, and length. Observations cluster near the horizontal zero line, indicating generally small cavity size changes, whereas hyperpharyngeal constriction varies more widely. The fitted trends show only a weak positive relationship, which is slightly stronger for alveolars than velars. This suggests that greater hyperpharyngeal constriction is associated with modestly greater cavity reduction. However, the substantial overall scatter indicates that cavity compression often occurs independently of it.

Statistical analysis confirmed a significant positive effect of *delta hyperph* on changes in cavity size ($F(1, 10.31) = 27.22$, $p < .001$). Because *delta cavity* was heavy-tailed, the response was modeled using a sign-preserving scaled inverse hyperbolic sine transform. Length showed a significant main effect ($F(1, 171.10) = 5.36$, $p < .05$), while PoA alone did not ($F(1, 168.47) = 1.71$, $p = .19$). Both interactions were significant: delta hyperph × PoA ($F(1, 939.47) = 5.39$, $p < .05$) and delta hyperph × length ($F(1, 807.22) = 6.69$, $p < .01$). Position and its interaction with *delta hyperph* did not improve model fit and were excluded (position: $F(2, 147.60) = 1.40$, $p = .25$; delta hyperph × position: $F(2, 1136.19) = 0.23$, $p = .79$).[15]

[15] Selected model: deltaCavity_asinh_s~ deltaHyperph*PoA + deltaHyperph*length + (deltaHyperph | speaker) + (1 | word).

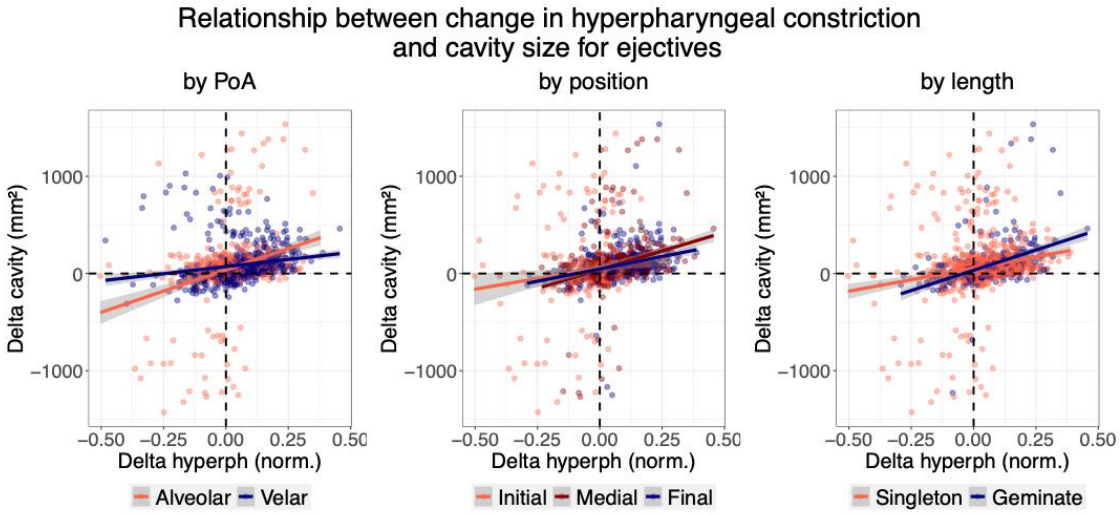

Figure 3.20 Scatter plot showing the relationship between delta hyperph (x-axis, normalized) and delta cavity (y-axis, in mm^2) for ejectives, separated by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Superimposed regression lines indicate the strength and direction of the relationship, with shaded regions representing 95% confidence intervals. The black dashed lines represent the zero values for both axes: observations below the zero y-axis indicate an increase in cavity size, while observations to the right of the zero x-axis indicate an increase in hyperpharyngeal constriction over the stop closure. Dots represent individual tokens.

Estimated slopes for the relationship between hyperpharyngeal constriction and change in cavity size were significantly positive for all groups (see **Table 3.17**). They were larger for alveolars (slope ≈ 441.95 mm^2 per unit) than for velars (slope ≈ 359.60 mm^2 per unit), and larger for geminates (slope ≈ 5.65 mm^2 per unit) than for singletons (slope ≈ 4.23 mm^2 per unit).

Table 3.17 Estimated slopes for the relationship between delta hyperph and delta cavity by PoA and length. Slopes are reported on the transformed (asinh) scale with corresponding raw-scale effect sizes.

Estimate Type	Level	Slope (asinh)	Slope (raw scale)	SE	DF	t	p	Signif.
Slope by PoA	alv	5.51	441.95	0.99	12.2	5.58	<.001	***
Slope by PoA	vel	4.37	359.60	0.97	11.6	4.49	<.001	***
Slope by length	s	4.23	354.28	0.96	10.7	4.43	0.0011	**
Slope by length	g	5.65	447.89	1.02	13.9	5.55	<.001	***

Pairwise comparisons given in **Table 3.18** confirmed that both differences were significant (alveolars-velars: Δ slope = 1.14 on the transformed scale, $p < .05$; geminates-singletons: Δ slope = -1.42 on the transformed scale, $p < .05$).

Table 3.18 Pairwise comparisons of delta hyperph (normalized) slopes across PoA and length.

Contrast Type	Contrast	Δ slope (asinh)	SE	DF	t	p	Signif.
Slope difference by PoA	alv - vel	1.14	0.50	942.2	2.30	0.0215	*
Slope difference by length	s - g	-1.42	0.55	810.7	-2.57	0.0103	*

In sum, hyperpharyngeal constriction was reliably associated with greater cavity reduction in ejectives, with the effect stronger for alveolars than velars and for geminates than singletons.

3.5.3. Effect of Larynx Raising on Absolute Burst Intensity

Taking absolute burst intensity as an indicator of intraoral pressure and building on the conclusion from Section 3.3.8 that high burst intensity peaks characterized Amharic ejectives, this next analysis examined whether the magnitude of vertical larynx displacement predicted the level of absolute burst intensity.

Figure 3.21 presents scatter plots of absolute burst intensity (y-axis, in dB) against *delta larynx* (x-axis, in mm), separated by PoA, word position, and length. Most regression lines show a moderately positive slope. This suggests a general trend: greater larynx raising tends to be associated with higher absolute burst intensity. The trend appears stronger for alveolar ejectives than for velars.

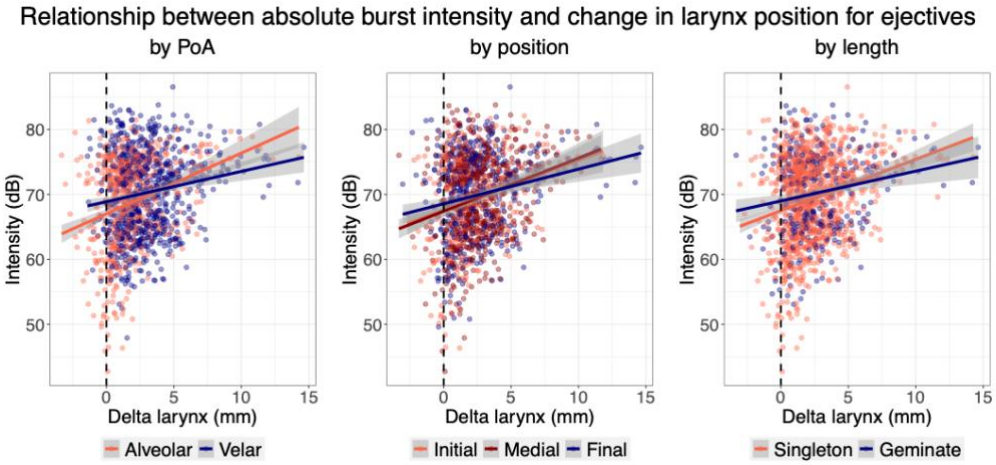

Figure 3.21 Scatter plot showing the relationship between delta larynx (x-axis, in mm) and absolute burst intensity (y-axis, in dB) for ejectives, separated by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Superimposed regression lines indicate the strength and direction of the relationship, with shaded regions representing 95% confidence intervals. The black dashed line represents the zero value for the x-axis: observations to the right of it, indicate larynx raising over the stop closure. Dots represent individual tokens.

Statistical analysis revealed that *delta larynx* alone did not significantly predict absolute burst intensity. In contrast, PoA ($F(1, 513.42) = 73.16$, $p < .001$), position ($F(2, 417.38) = 6.02$, $p < .01$), and length ($F(1, 445.62) = 15.83$, $p < .001$) all showed significant main effects. Several interaction effects were also significant: *delta larynx* × PoA ($F(1, 1303.47) = 22.73$, $p < .001$), *delta larynx* × position ($F(2, 1312.53) = 3.13$, $p < .05$), and *delta larynx* × length ($F(1, 1315.96) = 3.90$, $p < .05$), suggesting that the impact of larynx displacement on absolute burst intensity was modulated by all three contextual factors.[16]

Estimated slopes given in **Table 3.19** showed that the relationship between larynx raising and absolute burst intensity for ejectives was significant for alveolars (slope = 0.48 dB per mm, $p < .05$) but not for velars (slope = 0.12 dB per mm, $p = .49$), and a slope contrast given in **Table 3.20** confirmed that the effect was significantly stronger for alveolars (alveolar-velar: Δ slope = 0.36 dB per mm, $p < .001$).

Table 3.19 Estimated slopes for the relationship between delta larynx (in mm) and absolute burst intensity (in dB) by PoA, position, and length.

Estimate type	Level	Slope (dB per mm)	SE	DF	t	p	Signif.
Slope by PoA	alv	0.48	0.18	12.0	2.75	0.0175	*
Slope by PoA	vel	0.12	0.17	11.3	0.71	0.4913	
Slope by position	i	0.29	0.19	14.9	1.58	0.1343	
Slope by position	m	0.42	0.18	11.8	2.38	0.0352	*
Slope by position	f	0.20	0.18	12.3	1.13	0.2802	
Slope by length	s	0.39	0.17	10.7	2.28	0.0444	*
Slope by length	g	0.22	0.18	13.4	1.21	0.2483	

Across word positions, the relationship reached significance only medially (slope = 0.42 dB per mm, $p < .05$; **Table 3.19**). Pairwise comparisons (**Table 3.20**) revealed that the slope for medial ejectives differed significantly from the final position slope (medial-final: Δ slope = 0.22 dB per mm, $p < .05$), while neither comparison involving the initial position was significant (initial-medial: Δ slope = -0.12 dB per mm, $p = .46$; initial-final: Δ slope = 0.09 dB per mm, $p = .60$).

[16] Selected model: burstIntensity ~ deltaLarynx*PoA + deltaLarynx*position + deltaLarynx*length + (deltaLarynx | speaker) + (1 | word).

Table 3.20 Pairwise comparisons of delta larynx (in mm) slopes across PoA, position, and length.

Contrast type	Contrast	Estimate (Δ slope)	SE	DF	t	p	Signif.
Slope difference by PoA	alv - vel	0.36	0.08	1,304.4	4.76	<.001	***
Slope difference by position	i - m	-0.12	0.10	1,295.8	-1.18	0.4636	
Slope difference by position	i - f	0.09	0.10	1,316.8	0.97	0.5979	
Slope difference by position	m - f	0.22	0.09	1,319.6	2.49	0.0341	*
Slope difference by length	s - g	0.17	0.09	1,316.2	1.97	0.0491	*

In terms of length, the relationship was significant for singleton ejectives (slope = 0.39 dB per mm, $p < .05$) but not geminate ejectives (slope = 0.22 dB per mm, $p = .25$; **Table 3.19**), and the slope contrast between the two was marginally significant (singleton-geminate: Δ slope = 0.17 dB per mm, $p < .05$; **Table 3.20**). In sum, the contribution of larynx raising to absolute burst intensity was consistently positive, but its strength was context-dependent: it was strongest in alveolar, medial, and singleton ejectives.

3.5.4. Effect of Cavity Size Reduction on Absolute Burst Intensity

The analysis in Section 3.4.2 showed that supralaryngeal cavity reduction was a stable feature of Amharic ejectives. This next analysis therefore tested whether its extent predicted absolute burst intensity, taken here as an indicator of intraoral pressure buildup.

Figure 3.22 presents scatter plots showing the relationship between *delta cavity* (x-axis, in mm^2) and absolute burst intensity (y-axis, in dB), separated by PoA, word position, and length. Across all conditions, a weak negative relationship is visible: greater reductions in cavity size are associated with weaker absolute burst intensities. This trend is reflected in the downward-sloping regression lines, although the overall gradient remains shallow.

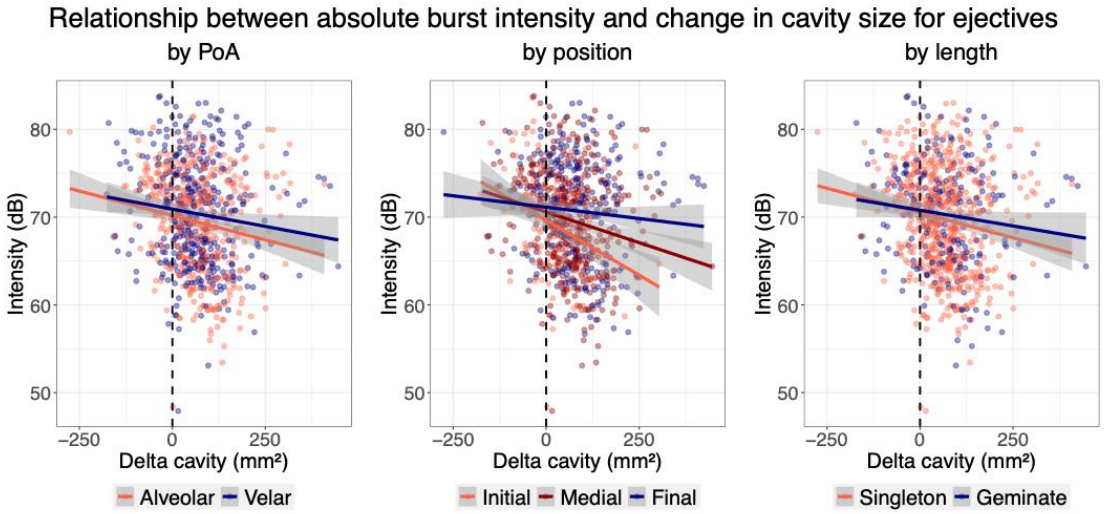

Figure 3.22 Scatter plot showing the relationship between delta cavity (x-axis, in mm^2) and absolute burst intensity (y-axis, in dB) for ejectives, separated by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Superimposed regression lines indicate the strength and direction of the relationship, with shaded regions representing 95% confidence intervals. The black dashed line represents the zero value for the x-axis: observations to the right of it indicate the cavity's size decreased over the stop closure. Dots represent individual tokens.

Statistical analysis showed that PoA ($F(1, 801.86) = 24.14$, $p < .001$), position ($F(2, 801.07) = 3.29$, $p < .01$), and length ($F(1, 801.01) = 5.88$, $p < .01$) had a significant effect on absolute burst intensity. However, *delta cavity* ($F(1, 800.93) = 0.81$, $p = .36$) did not improve model fit, nor did any of the interactions involving *delta cavity* (all $p > .05$). Because these terms did not improve model performance, they were removed from the final model.[17] Thus, changes in cavity size did not predict absolute burst intensity.

3.5.5. Effect of Hyperpharyngeal Constriction on Absolute Burst Intensity

As Section 3.4.3 showed, hyperpharyngeal constriction was a consistent component of Amharic ejectives, and Section 3.5.2 indicated that greater constriction tended to co-occur with increased supralaryngeal cavity reduction in a context-dependent manner. However, because cavity compression itself did not predict absolute burst intensity, there was little theoretical reason to expect the more localized measure to do so. The relationship between changes in hyperpharyngeal constriction and absolute burst intensity was nonetheless examined as a final clarification of how the localized and global measures relate to one another.

[17] Selected model: burstIntensity ~ PoA + position + length + (1 | speaker).

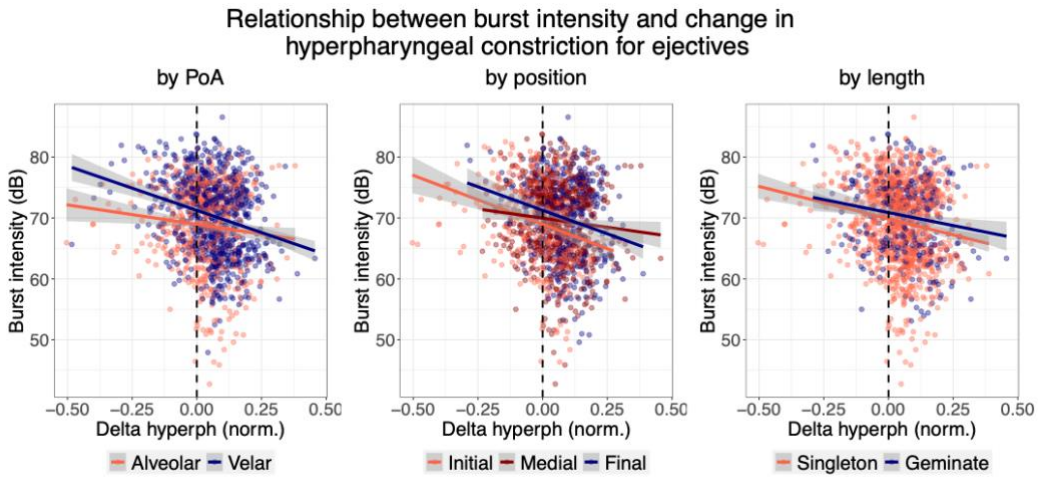

Figure 3.23 Scatter plot showing the relationship between delta hyperph (x-axis, normalized) and absolute burst intensity (y-axis, in dB) for ejectives, separated by PoA (alveolar, velar), word position (initial, medial, final), and length (singleton, geminate). Superimposed regression lines indicate the strength and direction of the relationship, with shaded regions representing 95% confidence intervals. The black dashed line represents the zero value for the x-axis: observations to the right of it indicate an increase in hyperpharyngeal constriction over the stop closure. Dots represent individual tokens.

Figure 3.23 visualizes the relationship between *delta hyperph* (x-axis, normalized) and absolute burst intensity (y-axis, in dB) for ejectives, shown by PoA, word position, and length, with fitted regression lines to illustrate overall trends within each subgroup. The fitted trends are generally slightly negative, and the decline appears steeper for velars than for alveolars. Differences by position and length are comparatively subtle.

Statistical analysis revealed significant main effects of *delta hyperph* ($F(1, 794.28) = 12.03, p < .001$), as well as of position ($F(2, 340.80) = 3.90, p < .05$), PoA ($F(1, 1206.45) = 100.49, p < .001$), and length ($F(1, 1025.86) = 12.23, p < .001$). The delta hyperph × position interaction was significant ($F(2, 1248.59) = 6.72, p < .01$). Neither the delta hyperph × PoA nor the delta hyperph × length interaction was supported, so these terms were not included in the final model (delta hyperph × PoA: $F(1, 1332.8) = 2.92, p = .09$; delta hyperph × length: $F(1, 1310.4) = 1.87, p = .17$).[18]

[18] Selected model: burstIntensity ~ deltaHyperph*position + PoA + length + (1 | V1V2) + (1 | speaker).

Table 3.21 Estimated slopes for the relationship between delta hyperph (normalized value) and absolute burst intensity (in dB) by position.

Estimate Type	Level	Slope (dB per unit)	SE	DF	t	p	Signif.
Slope by position	i	-2.53	1.45	1,333.4	-1.75	0.0805	
Slope by position	m	0.22	1.44	675.0	0.15	0.8788	
Slope by position	f	-7.19	1.61	1,236.8	-4.46	<.001	***

Estimated slopes given in **Table 3.21** showed that *delta hyperph* significantly predicted absolute burst intensity only for ejectives in final position, where the slope was strongly negative (slope = -7.19 dB per unit, $p < .001$). Pairwise slope comparisons (**Table 3.22**) confirmed that the final-position slope differed significantly from the medial slope (final-medial: Δ slope = -7.41 dB per unit, $p < .001$), while the remaining contrasts were not significant (initial-medial: $p = .35$; initial-final: $p = .07$).

Table 3.22 Pairwise comparisons of delta hyperph (normalized) slopes across position.

Contrast Type	Contrast	Estimate (Δ slope)	SE	DF	t	p	Signif.
Slope difference by position	i - m	-2.75	1.98	1,158.4	-1.39	0.3458	
Slope difference by position	i - f	4.65	2.12	1,321.5	2.19	0.0730	
Slope difference by position	m - f	7.41	2.04	1,306.0	3.64	<.001	***

The analysis showed a small effect in word-final position, where greater changes in constriction predicted slightly lower absolute burst intensity, but provided little evidence for a systematic relationship elsewhere. Overall, the analysis suggests that hyperpharyngeal constriction did not play a meaningful role in increasing absolute burst intensity.

3.5.6. Summary of the Relationship Analysis Results

The relationship analyses examined how key articulatory parameters related to each other and to absolute burst intensity in Amharic ejective production. Larynx displacement and changes in supralaryngeal cavity size were positively related, but only weakly: greater larynx raising predicted only a modest additional reduction in cavity size. This relationship was not reliably modulated by other phonetic factors.

Hyperpharyngeal constriction was a consistent predictor of cavity reduction, and its contribution was context-dependent – stronger for alveolars and for geminates. Larynx displacement showed the clearest link to absolute burst intensity: its effect was reliably positive but varied by context, being strongest in alveolar, medial, and singleton ejectives. In comparison, the results provided no evidence that cavity reduction or hyperpharyngeal constriction contributed to increases in absolute burst intensity.

Key acoustic characteristics of Amharic ejectives have now been analyzed, together with their articulatory properties from multiple perspectives: laryngeal activity, overall supralaryngeal cavity compression, and constriction in the hyperpharyngeal region of the vocal tract. The relationship between these articulatory strategies and their influence on absolute burst intensity was also examined. Taken together, these findings pave the way for a nuanced view of the articulatory mechanisms underlying Amharic ejectives and provide a foundation for broader interpretation of phonological ejective production.

3.6. Discussion of Amharic Ejectives

This study set out to investigate the acoustic and articulatory properties of phonological ejectives in Amharic through the use of rt-MRI, in order to address ongoing debates about the mechanisms underlying ejective production and the articulatory strategies involved. The following discussion considers how these empirical patterns support, challenge, or refine existing descriptions of how ejectives are produced.

3.6.1. Acoustic Robustness and Core Articulatory Features of Amharic Ejectives

The acoustic analysis of Amharic ejectives in this study revealed notably robust and consistent patterns across contextual factors and most speakers. In this respect, the present dataset is not representative of the inter-speaker variability or PoA-related variability that other studies report finding even within the same language (e.g., Warner, 1996; Wright et al., 2002).

Closure duration did not differ meaningfully between ejective and pulmonic stops, an expected finding cross-linguistically. From an articulatory perspective, this indicates that the duration of oral closure is not systematically adjusted to accommodate the different initiation mechanisms of glottalic versus pulmonic stops.

However, other acoustic hallmarks of ejective production were observed: all tokens exhibited the characteristic post-release silent gap, confirming glottal closure at the moment of oral release. Also, relative burst intensity values were consistently lower for ejectives than for pulmonic stops, a consequence of their stronger burst amplitudes. This pattern confirms the buildup of intraoral pressure typically associated with ejective sounds.

The rt-MRI analysis provides clear evidence that systematic larynx raising is a robust and consistent feature of Amharic ejectives, across all places of articulation, word positions, geminate and singleton forms, and speakers. This offers empirical confirmation of the classic view of ejectives as sounds initiated by a glottalic airstream driven by active larynx raising (Catford, 1977). However, the results also show that larynx raising alone does not fully exhaust all the details of the mechanism of ejective production on its own. Supralaryngeal activity plays a stable and consistent part. This finding lends empirical support for long-standing claims of supralaryngeal involvement in ejective production, which were previously limited in large part to anecdotal observations (e.g., Demolin, 2002; Hermes et al., 2016; Kingston, 1984).

Included in the supralaryngeal activity are both cavity compression in a global sense, captured in this study by the *delta cavity* measure, as well as localized narrowing of the hyperpharyngeal region, such as that produced by tongue root retraction. The former is evocative of the *IPA Handbook*'s (1999, pp. 9-10) broader description of ejectives as resulting from "squeezing" the air caught within the intraoral cavity without the specification of the individual articulators involved. Against the backdrop of this highly stable and consistent pattern, certain finer-grained details nonetheless merit closer reflection before turning to broader implications.

3.6.2. Coordination of Laryngeal and Supralaryngeal Gestures

The data confirmed that laryngeal activity in the form of systematic upward displacement and supralaryngeal cavity compression not only co-occur in ejective production, but also exhibit a modest, statistically significant positive relationship, suggesting that the two gestures are coordinated components of a unified articulatory mechanism. The data provided additional insight into this coordination. Cavity reduction was less variable across speakers than larynx raising, indicating that observable compression operates within a relatively narrow range.

There may be practical limits to how much speakers compress the cavity, at least within the two-dimensional imaging plane used here. Speakers may implement a relatively fixed "base configuration" of cavity compression, potentially reflecting oral cavity positioning that prevents passive expansion of compliant structures (Kingston, 1984). Such a configuration could involve increased rigidity of the tongue body and tightening of the posterior vocal tract wall to support more effective pressure generation. Within this system, speakers may reduce cavity size only to the extent necessary to complement pressure buildup from larynx raising.

In contrast, the extent of larynx raising showed substantially greater variability across speakers, consistent with findings from Kingston (1984) and Bückins et al. (2018). This variability likely reflects both physiological and habitual factors. Anatomical differences, such as pharynx length, clearly contribute (as illustrated by the notably extended pharynx in speaker S06; see Appendix D), while variation among more average-sized speakers is more plausibly driven by patterns of hyperarticulation (as noted by Kingston, 1984) and by idiosyncratic articulatory habits (Hoole et al., 2024).

Taken together, these findings indicate that the two gestures participate in a unified articulatory mechanism while fulfilling distinct functional roles. Larynx raising directly generates intraoral pressure, whereas supralaryngeal cavity reduction optimizes the vocal tract for efficient pressure buildup. The ejective mechanism can thus be characterized as the coordinated interplay of these complementary components.

3.6.3. Articulatory Contributions to Burst Intensity

More direct support for the functional differentiation of laryngeal and supralaryngeal gestures comes from their relationship to burst intensity. Supralaryngeal cavity reduction showed no significant effect on absolute burst intensity and, in some cases (e.g., alveolars), even trended negatively. This is noteworthy, as it challenges previous claims that cavity reduction is a central contributor to pressure generation (Kingston, 1984, 1985), suggesting instead that though its role is consistent it is limited and, at best, secondary.

Larynx raising, in contrast, was directly associated with increases in absolute burst intensity. Interestingly, this positive relationship was statistically significant only for alveolar ejectives, despite alveolars exhibiting smaller larynx displacements and lower burst intensities than velars.

One possible explanation is that in alveolars, larynx raising remains within a functional range where changes visibly affect pressure, whereas in velars the magnitude of vertical displacement may already approach the upper limit of what larynx raising alone can contribute. Explaining the higher burst intensities of velars likely requires the *combined* action of larynx raising and supralaryngeal compression, aligning with Kingston (1984) and Demolin (2002) who argue that larynx raising alone cannot fully account for high-intensity ejectives.

3.6.4. Asymmetries Across Place of Articulation

Place of articulation further shapes how the ejective mechanism is operationalized. Notably, velar ejectives exhibited more extreme articulatory and acoustic patterns than alveolars: greater larynx raising, higher degrees of hyperpharyngeal constriction, and slightly elevated burst intensities. The only exception was overall supralaryngeal cavity size reduction, which did not differ significantly from that of alveolars, though this finding is based on absolute values. Given that alveolar stops have a more anterior PoA, the cavity behind their constriction is naturally larger than that behind velars. Therefore, from a relative perspective, the same absolute reduction in mm^2 represents a proportionally greater constriction for velars.

Velar ejectives benefit from their posterior place of articulation, which naturally creates smaller post-constriction volumes and thus requires less effort to build intraoral pressure. In contrast, more anterior articulations require more effort for the same effect. Thus, one might expect alveolars, with their larger cavities, to involve the greater articulatory engagement, such as increased larynx raising. However, the opposite pattern emerges.

This discrepancy may be biomechanically explained: the hyoid bone is attached to both the larynx and tongue root. During velar articulation, the tongue body is retracted and elevated toward the soft palate. This posture of the tongue may facilitate co-activation of the laryngeal and pharyngeal musculature (Esling, 2005). Thus, the more extreme laryngeal and pharyngeal gestures observed in velars as opposed to other PoAs do not arise because they are needed for pressure buildup, but rather because they are easily executed.

In contrast, the more anterior constriction for alveolar ejectives requires the tongue to be pulled forward, which restricts tongue root retraction and coordinated larynx raising, limiting the degree of articulatory engagement. Supralaryngeal cavity size reduction alone acting on a larger cavity cannot make up for the more limited laryngeal involvement, resulting in lower burst intensities.

Further PoA-related differences regarding how regional articulatory gestures support overall supralaryngeal cavity reduction also emerge. For velars, hyperpharyngeal constriction is larger in magnitude but does not drive global compression, implying that the main driver of cavity reduction lies elsewhere in the vocal tract. For alveolars, the constriction is smaller but evidently forms a significant component of global compression. In neither case does increased constriction in this region alone affect burst intensity, indicating a limited role in pressure generation.

The idea that pressure changes are easier to generate at more posterior PoAs, due to both proportional and biomechanical factors, is not a new insight (cf. Javkin, 1977). What the present data adds is empirical evidence that this advantage is actively exploited in velar ejectives. In short, velar ejectives are louder not because they have to be, but because they can be. By contrast, the articulatory strategy behind alveolar ejectives appears more conservative, aiming to do just enough to establish a pressure-based contrast with their pulmonic counterparts.

Word position played a limited role in shaping the articulatory patterns, contrary to findings in Georgian (Brandt & Simpson, 2021). Length (singleton vs. geminate) had no significant effect on the articulatory measures, aside from a stronger link between hyperpharyngeal constriction and overall cavity compression. This suggests that these categories differ primarily in temporal coordination, in line with Demolin's (2004) findings.

3.6.5. Interpretation of the Observed Articulatory Mechanisms

Overall, these findings provide evidence that ejective production in Amharic involves an integrated system of laryngeal and supralaryngeal mechanisms aimed at maximizing the conditions under which the buildup of intraoral pressure can efficiently take place. The data suggests that speakers are attuned to articulatory conditions that favor the ejective mechanism, such as those afforded by velar PoA, and readily take advantage of them. Amid the strikingly low variability in the production of Amharic ejectives, this is where

the most systematic variation emerges. In this sense, the data resonates with the *IPA Handbook*'s (1999) characterization of ejectives as targeting compression of the intraoral airspace. This account captures the functional unity of ejective production better than a strictly glottal initiation-centered account (Catford, 1977).

Nevertheless, from the initiation perspective, larynx raising plays the primary role. This can be indirectly inferred from the alveolar results, which show a positive correlation between larynx raising and absolute burst intensity, while supralaryngeal cavity reduction alone does not. It is reasonable to assume that the same principle holds for velar ejectives. However, this relationship is not visually evident in the data, potentially due to a ceiling effect, whereby the already high levels of larynx raising cannot further influence increases in burst intensity.

In omitting reference to larynx raising, the *IPA Handbook*'s (1999) account overlooks the key aspect of ejectives – their *glottalic* initiation. From this perspective, the Amharic data does not support Kingston's (1985) claim that larynx raising is merely a passive consequence of supralaryngeal articulation rather than the primary driver of pressure buildup. The evidence suggests the opposite: supralaryngeal cavity reduction appears to play a supportive role, providing the constricted configuration that optimizes conditions for the pressure-generating function of vertical larynx displacement to take effect, but on its own has limited generative power.

The observed consistency across speakers and conditions supports the view that Amharic ejectives represent a highly cohesive implementation of the glottalic airstream mechanism – one that targets compression of the intraoral airspace through coordinated action from laryngeal and supralaryngeal gestures. At the same time, the wide cross-linguistic range of ejective acoustics does not preclude the possibility that other languages organize these gestures more independently or weight them differently, particularly in systems with either exceptionally intense or comparatively weak ejective bursts.

3.7. Conclusion

This chapter has examined the acoustic and articulatory properties of phonological ejectives in Amharic. The acoustic results reveal robust patterns across speakers and contexts. The articulatory findings align with the classical definition of glottalic initiation but refine this glottis-centered view by showing that ejective production in Amharic involves the integration of both larynx raising and supralaryngeal cavity reduction, a

coordinated mechanism aimed at enhancing intraoral pressure. Velar ejectives, in particular, display more active articulatory engagement than alveolars.

Taken together, these findings provide a detailed rt-MRI-based account of ejective production in Amharic and establish a solid baseline for comparison with the English ejectives examined in the following chapter. This explores to what extent languages without phonological ejectives, such as English, use different articulatory strategies to arrive at similar phonetic outcomes, offering new insights into the variability and adaptability of the ejective mechanism across linguistic systems.

Chapter 4
Non-Phonological Ejectives in British English

4.1. Introduction

This chapter investigates the acoustic and articulatory properties of non-phonological ejective stops in British English. Although ejectives are not part of the original phonological inventory of English, their presence in everyday speech is well established. Yet within the field of phonetics, these English ejectives remain something of a curiosity. Traditionally described as glottalic sounds, ejectives are now emerging in a language whose phonological system is entirely pulmonic. Given the absence of a foundation for glottalic sounds in English, there is little reason to assume that these sounds conform to canonical accounts of ejective production – raising questions about the validity of labeling them as ejectives at all.

English ejectives offer a valuable opportunity to explore the range of articulatory strategies capable of producing ejective-like percepts. In doing so, they shed light on the articulatory sources of ejectiveness independent of any phonologically contrastive function. As such, they provide a foundation for reevaluating the core defining features of this sound class. Furthermore, understanding how such sounds have emerged in English may offer insight into potential diachronic pathways of ejective development, an area that remains largely underexplored. To address these questions, this study evaluates English ejective stops against the traditional typological model of ejective production. Using acoustic analysis and rt-MRI, it first examines whether these English ejectives exhibit typical acoustic features associated with phonological ejectives. Second, it investigates the extent to which these sounds follow the expected glottalic initiation pattern, defined by vertical larynx displacement, or whether alternative articulatory strategies, such as supralaryngeal cavity compression, are instead at play.

This chapter is structured as follows. After a brief methods section (Section 4.2) and a preliminary auditory analysis to identify ejective realizations in the dataset (Section 4.3), the analysis proceeds in parallel with the structure used in the Amharic chapter (Chapter 3). Section 4.4 presents the acoustic analysis. Section 4.5 provides the articulatory analysis, focusing on the three articulatory parameters: vertical larynx movement, supralaryngeal cavity compression, and hyperpharyngeal constriction.

Finally, Section 4.6 examines how these articulatory strategies interact and assesses their respective contributions to absolute burst intensity, used here as an indirect proxy for intraoral pressure. However, before turning to the analysis, the remainder of this section provides an overview of the existing literature on English ejectives, outlining what is currently known about their production and characteristics.

4.1.1. Early Accounts and Regional Aspects of English Ejectives

Ejectives in English are not a new discovery, as evidenced by numerous informal accounts dating back at least 40 years (Catford, 1977; Chirrey, 1999; Fabricius, 2000; Ladefoged & Johnson, 2006; Ladefoged & Maddieson, 1996; Ogden, 2024). Given this long-standing recognition, the limited number of systematic investigations is surprising.

Early observations of ejectives in English largely focus on regional variation, indicating no single dialectal origin (Catford, 1977; Roach, 2002; Wells, 1982). They have been noted in both northern and southern varieties – particularly in the Midlands and the North – and are most often observed in word-final position, especially utterance-finally (Ladefoged & Johnson, 2006; Lass, 1984). In Scottish English, ejectives have been more clearly documented (Gordeeva & Scobbie, 2011, 2013; Macafee, 1983; McCarthy, 2012; McCarthy & Stuart-Smith, 2013).

Overall, their presence in diverse varieties suggests that they likely emerged as a language-internal development (McCarthy, 2012), rather than through borrowing or diffusion from another language (Milroy et al., 1994). The question of where, when, and how these sounds first arose, however, remains unanswered (Simpson, 2014). Despite the lack of systematic documentation, there is a growing impression that ejectives are becoming more geographically widespread and increasingly common in use, though both trends still require empirical substantiation (Simpson, 2014).

4.1.2. Structural Conditioning, Frequency, and Acoustic Properties

Early reports consistently highlight the role of structural context in conditioning ejective production in English. Ejectives occur exclusively word-finally, particularly at phrase boundaries or before pauses. They tend to follow voiced segments, such as vowels, nasals, or laterals, and occur more frequently in stressed syllables (Ogden, 2024).

Building on this framework, McCarthy and Stuart-Smith (2013) analyzed velar ejectives in three consonant clusters: /-ŋk/, /-rk/, and /-sk/. While nasal contexts most strongly favored ejective realizations, 65% of /k/ tokens in the /-sk/ cluster were also realized as ejectives. Their study offered preliminary insights into vowel effects, noting that ejectives were more likely to follow back rounded vowels and less likely after close vowels. In their study of child speech, Gordeeva and Scobbie (2011) also found that 11.7% of ejective realizations occurred in lexemes that phonologically ended in voiced stops (e.g., *pig*, *food*). Their findings further confirmed a place-of-articulation bias, with velar stops more likely than bilabial or alveolar stops to be realized as ejectives.

Having outlined the structural contexts that favor ejective production in English, however, it is important to emphasize that these contexts do not guarantee their occurrence (Simpson, 2014). English ejectives remain unpredictable. This distinguishes them from the epiphenomenal ejectives observed in German, which are highly predictable and tightly context-dependent (Simpson, 2007). The unpredictability of English ejectives represents a major challenge for experimental work, as there is no reliable way to anticipate when, or by whom, ejectives will occur, making consistent data collection difficult (Simpson, 2014; Simpson & Sulaberidze, 2022).

A small number of studies employing systematic analytical methods have begun to quantify the frequency and distribution of ejectives in English. McCarthy's (2012) sociophonetic investigation of word-final /k/ among secondary school girls in Glasgow found that up to 65% of final /k/ tokens were realized as some form of ejective. A follow-up study by McCarthy and Stuart-Smith (2013), analyzing word-final /k/ in the speech of 28 Glasgow high school girls across casual and read speech styles, confirmed this pattern, identifying ejective realizations as the most common variant of /k/. In a related study, Gordeeva and Scobbie (2013) examined word-final stops (/p, t, k/) in the speech of Scottish English pre-school children and found that 10% of all tokens were realized as ejectives. They also observed categorical usage across participants: only five out of seven children produced ejectives, suggesting that ejective production may be both variable and speaker-specific, even from an early age.

To date, only one study has conducted a dedicated acoustic analysis of English ejectives. Gordeeva and Scobbie (2013) found that ejective realizations in Scottish English are characterized by shorter stop burst duration, higher burst intensity, and slightly earlier voicing offset in the preceding vowel compared to their pulmonic counterparts. Based on intensity and duration patterns, they concluded that Scottish

English ejectives share acoustic characteristics with phonological ejectives found in non-European languages (e.g., Grawunder et al., 2010; Lindau, 1984; Vicenik, 2010). Although McCarthy and Stuart-Smith (2013) did not include a full acoustic analysis, they noted a range of ejective realizations, classified as weak, strong, intense, and velar ejective affricates, highlighting the diversity of possible forms.

4.1.3. Phonetic Accounts and Functional Roles of Ejectives in English

Early interpretations suggested a link between ejectives and glottalization, a well-documented phenomenon in English dialects (Chirrey, 1999; Milroy et al., 1994; Ogden, 2024; Roach, 1979; Stuart-Smith, 1999; Wells, 1982). Gordeeva and Scobbie (2013) specifically tested whether ejective production correlated with glottalization of preceding vowels. However, their findings showed no such correlation. They also rejected the idea that ejectives serve a phonological role, such as marking the /±voice/ contrast.

At present, English ejectives appear to hold a quasi-phonological status, serving interactional, pragmatic, and sociolinguistic functions within the language. They frequently occur in phrase-final or turn-final positions, where they may signal discourse boundaries, support floor-holding, or enhance clarity (McCarthy & Stuart-Smith, 2013; Ogden, 2024; Simpson, 2014). While earlier accounts speculated on a paralinguistic role in emphatic or careful speech (Wells, 1982), more recent work also highlights the occurrence of weaker ejectives in sentence-medial positions, suggesting that not all instances have clear discourse functions (Simpson, 2014). In addition, ejectives appear to operate as sociolinguistic markers, with their use influenced by factors such as speech style, age, and ethnicity, particularly in Scottish English, indicating that they form part of socially influenced pronunciation patterns (McCarthy & Stuart-Smith, 2013).

In light of their variation in realization and the range of emerging functions they appear to serve, Simpson (2014) speculates that *multiple* articulatory strategies may be currently in play, possibly linked to these different functions. Some ejective realizations may be entirely epiphenomenal, (e.g., resulting from overlap between glottalized vowel onsets and preceding stops). Others may involve non-glottalic mechanisms (as proposed for Georgian non-initial ejectives; Brandt & Simpson, 2021), while still others may reflect true glottalic initiation or as-yet unidentified processes.

Preliminary observations by Simpson and Sulaberidze (2022) based on multichannel recordings from five young female speakers, support this view. In a reading task designed to elicit both natural and hyperarticulated speech, ejectives occurred in 41% of pulmonic voiceless stop tokens for three of the five speakers, primarily in utterance-final contexts, especially before "*even though....*" The data suggest a continuum ranging from weak, likely epiphenomenal realizations to more forcefully released tokens, possibly involving active laryngeal engagement. These observations align with the idea that production may be shaped by the function these sounds serve.

Building on the issues outlined above, this study undertakes one of the first empirical investigations of the acoustic and articulatory characteristics of English ejectives. Using this combined approach, it aims to shed light on how these sounds are actually being produced. The following section outlines the methods used, including participant information and the design of the experimental materials. It then describes the auditory categorization procedures used to identify ejective productions, setting the stage for the acoustic and articulatory analyses that follow.

4.2. Methods

The following section outlines the participants, stimuli, and analytical procedures used in the English rt-MRI study.

4.2.1. Speakers

This study draws on rt-MRI data from 27 young adult speakers (13 female) of Standard Southern British English, aged between 18 and 46 years ($\mu = 22$, SD = 5.26). The data represents a subset of a larger dataset originally collected for an independent study on nasalization in British and American English. All but two participants were recruited in the UK and flown to Germany for the recordings; the remaining participants were already residing in Germany but had lived there for no more than two years on average. All participants received monetary compensation. Prior to participation, written informed consent was obtained, confirming each participant's compatibility with the rt-MRI procedure and the absence of any hearing or speech impairments. The experimental protocol followed the general procedure outlined in Chapter 2, Section 2.3.

4.2.2. Stimuli

The set of target words selected for this study consisted of 131 lexical items containing voiced and voiceless coda stops, the coda being the critical position where ejective sounds in English are known to occur. The target stops spanned three places of articulation: bilabial, alveolar and velar. Codas were either simplex (e.g., /-p, -t, -k, -b, -d, -g/) or nasal-stop clusters (e.g., /-nt, -nd, -mp, -md, -ŋk, -ŋd/), and occurred with a range of preceding vowels, but were always followed by the vowel /ʌ/. For elicitation, the words were embedded sentence-medially into a variety of carrier phrases (e.g., "Saw *bat* about six"). Two phrasal accent conditions were elicited through orthographic highlighting. Each stimulus was produced once per accent condition, yielding both accented and deaccented versions of each target word. Due to occasional technical issues or mispronunciations during recording, some speakers produced multiple instances of the same item. With one exception, all target words were monosyllabic. A complete list of target words is provided in Appendix E.

4.2.3. Dataset Overview and Exclusions

After excluding erroneous tokens, an average of 197 tokens per speaker were analyzed, yielding a total of 5,324 tokens across all speakers. This included 512 bilabial, 4,117 alveolar, and 695 velar tokens. The distribution of PoA was not balanced, as the corpus was originally compiled for a separate study on nasalization in English.

Due to the unpredictability of ejective occurrence in English, the first analytical step involved the auditory identification of ejective realizations. To this end, all coda stops in the dataset (5,324 tokens from 27 British English speakers) were auditorily classified based on their airstream characteristics into four categories: ejective, pulmonic, glottal replacement (i.e., realized as a glottal stop with no oral closure), and unclear. This classification served as the basis for selecting tokens for subsequent acoustic and articulatory analysis.

The acoustic and articulatory analyses were conducted on the subset of 22 speakers identified as having produced ejective realizations. Only accented tokens were included at this stage, as deaccented ejectives were extremely rare and would have introduced significant imbalance. As a result, accent was not included as a factor in subsequent statistical models.

The acoustic analysis comprised 2,845 tokens. While the auditory analysis grouped tokens broadly by airstream mechanism (ejective vs. pulmonic), the acoustic analysis further distinguished voiceless and voiced pulmonic stops, based on the underlying phonological specification of each lexical item. Relative burst intensity, closure duration, and PBL intensity were analyzed. Intensity rise ratio was initially included but excluded from the final analyses due to lack of informative results.

Known challenges with rt-MRI signal quality (see Chapter 2 for details) influenced token loss for the articulatory analysis. Thus, the number of usable tokens varied by measure. For the analysis of larynx movement, 2,678 tokens (including 462 ejectives) were available. A more limited dataset of 1,098 tokens (including 201 ejectives) from 19 speakers was analyzed for the measure of cavity size. For the analysis of the hyperpharyngeal region, 2,803 tokens (including 473 ejectives) were retained.

The final set of relationship analyses used the respective reduced datasets described above. For example, analyses involving cavity size and larynx height were restricted to the tokens available for the cavity size measure.

4.2.4. Statistical Modeling

Descriptive statistics were used to present the frequencies and distributions resulting from the categorization process in the auditory analysis. A binary logistic mixed-effects model was used to evaluate the phonetic conditioning factors associated with ejective occurrences in English. The model was structured as follows:

(1) model = glmer(realization ~ PoA + accent + voicing + (1 | speaker))

where realization included two levels (ejective and pulmonic), PoA included three levels (bilabial, alveolar, velar), accent included 2 levels (accented and deaccented), and voicing included 2 levels (voiceless and voiced).

Separate linear mixed-effects models were used to analyze three acoustic measures (closure duration, relative burst intensity, and PBL intensity) and three articulatory measures (larynx, cavity, hyperph deltas), with each measure serving as the dependent variable (see Chapter 2, Sections 2.5 and 2.6 for a detailed description of how these measures were computed).

The initial model structure for analyzing English stops was:

(2) model = lmer(measure ~ stop type * PoA + (stop type | speaker) + (1 | word)

where stop type included four levels: ejective, pulmonic voiceless aspirated, pulmonic voiced, and unclear stops; and PoA again included bilabial, alveolar, and velar. For the articulatory measures *delta cavity* and *delta hyperph* only, an additional random intercept for the stop-preceding vowel (V1) was included to account for systematic differences in vocal-tract configuration. The stop-following vowel was not modeled, as it was always /ʌ/ across the study.

(3) model = lmer(deltaClosure/deltaHyperph ~ stop type * PoA + (stop type | speaker) + (1 | word) + (1 | V1))

A slightly different model structure was used in the final section of the analysis, which examined the relationship between different measures (e.g., *delta larynx* and absolute burst intensity):

(4) model = lmer(measure1 ~ measure2 * stop type + measure2 * PoA + (stop type | speaker) + (1 | word)

In this model, stop type included only two levels (ejective and unclear), while PoA remained unchanged (bilabial, alveolar, velar). When these analyses involved *delta cavity* and *delta hyperph,* model (4) also included a random intercept for the preceding vowel (1 | V1).

Where applicable, models were simplified following the procedure described in Chapter 2, Section 2.9. With the modeling framework in place, the first step in the analysis focused on determining which tokens could be classified as ejective based on auditory impression.

4.3. Auditory Categorization Analysis

Ejectives in English are not phonologically contrastive; their occurrences must first be auditorily identified. A preliminary auditory classification conducted for this purpose exposed a broad continuum of auditory ejectiveness, highlighting the challenges of categorical classification. A validation of this auditory categorization is presented in Appendix F. Ejective-like occurrences ranged from highly salient to so subtle as to be nearly indistinguishable from pulmonic stops, yet, somehow still not quite pulmonic. Variability in the ejective realizations also surfaced in features such as affricated releases, as well as the additional presence of creaky phonation and occasional preglottalization. While some speaker-specific tendencies were observed, for instance, a general preference among certain speakers for more intense ejectives, this was not consistent; speakers who often produced strong ejectives still also produced weaker or less distinct variants.

This study did not attempt to further sub-classify all these different types of realizations. Since the auditory analysis was intended as a preliminary step to guide the subsequent articulatory investigation, a more fine-grained classification based on all of these nuanced differences was deliberately set aside in favor of a simplified labeling scheme. The aim at this stage was to establish a robust core category of ejective realization. Stop realizations that lacked a clearly perceptible burst but still carried glottalized cues, rendering them auditorily ambiguous, were categorized as *unclear*. The next section addresses the overall frequency of ejective realizations and examines the distributional characteristics that favored their occurrence.

4.3.1. Frequency and Structural Conditioning of Ejectives

Out of all word-final stops in the dataset, 9.4% (n = 502) were auditorily identified as ejectives, 57.3% (n = 3048) as pulmonic, and 31.1% (n = 1664) as unclear. Glottal replacements accounted for only 2.1% (n = 110) of all coda stops and were produced by just 7 speakers.

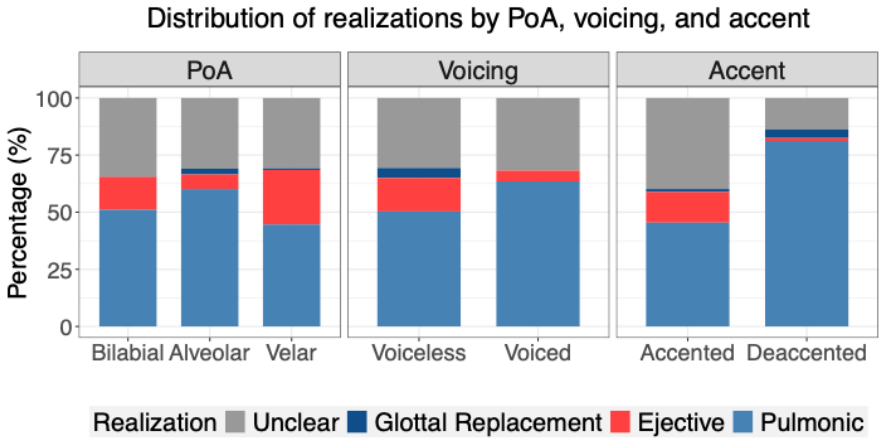

Figure 4.24 Distribution of word-final stop realizations (ejective, pulmonic, glottal replacement, and unclear) as influenced by PoA (bilabial, alveolar, velar), voicing (voiced, voiceless), and phrasal accent (accented, deaccented), aggregated across all speakers. Bars represent distributions (in %) of realization categories within each level of the respective factor.

A binary logistic mixed-effects model was run to evaluate which factors predicted ejective realization, comparing only tokens identified as ejective or pulmonic. Glottal replacements and unclear cases were excluded due to low or ambiguous counts. The model results are summarized in **Table 4.23**.

Table 4.23 Odds ratios from logistic mixed-effects model predicting ejective realization.

Term	Odds ratio
Intercept	0.26
PoA: alveolar vs. bilabial	0.64
PoA: velar vs. bilabial	2.23
Accent: deaccented vs. accented	0.03
Voicing: voiced vs. voiceless	0.20

PoA emerged as a key predictor. Compared to bilabials, velar stops were significantly more likely to be realized as ejectives (OR = 2.23), while alveolars were less likely (OR = 0.64). This pattern aligns with the distributional data (**Table 4.24**, **Figure 4.24**), where velars showed the highest proportion of ejectives (23.9%), followed by bilabials (14.3%) and alveolars (6.4%).

Table 4.24 Proportion of ejective realizations (%) across structural contexts in the dataset, including PoA, voicing and phrasal accent. Values represent the percentage of ejective realizations within each level of the respective factor. These proportions correspond to the red segments in **Figure 4.24**, providing a numeric summary of ejective realization rates by context.

PoA			**Voicing**		**Accent**	
bilabial	alveolar	velar	voiceless	voiced	accented	deaccented
14.3%	6.4%	**23.9%**	**14.7%**	4.7%	**13.4%**	1.3%

Phrasal accent had the strongest effect: deaccented tokens were far less likely to be ejective (OR = 0.03). In fact, only 4.2% of all ejectives occurred in deaccented contexts, and these were overwhelmingly alveolar (n = 20 out of 21). Voicing also significantly influenced ejective realization. Voiceless stops were more likely to be ejective than voiced stops (OR = 0.20), a finding mirrored in the raw data (14.7% vs. 4.7%).

The fact that ejectives replace voiced stops at all, however, is noteworthy, given that glottal constriction typically prevents phonation, making ejectives inherently voiceless. In sum, the statistical findings closely mirror the distributional patterns observed in **Figure 4.24** and **Table 4.24**, confirming that accented, voiceless, velar stops are the most favorable environments for ejective realization in this dataset.

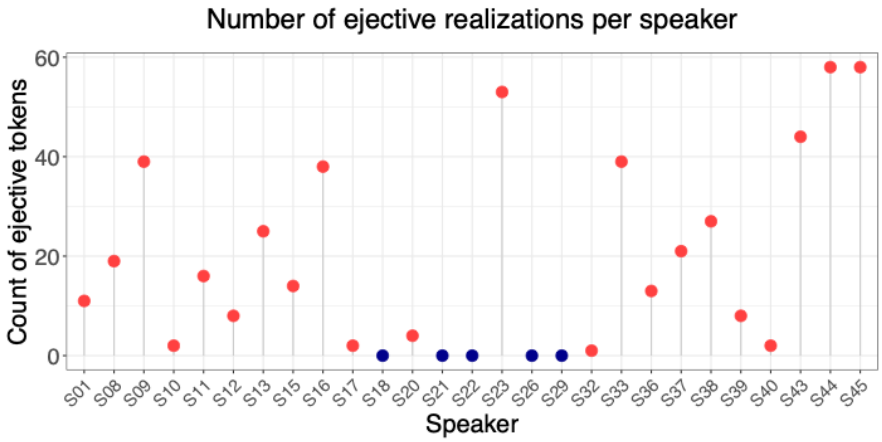

Figure 4.25 Number of ejective realizations identified per speaker across the dataset. Each point represents the total count of ejective tokens produced by an individual speaker (min.: n = 0; max.: n = 58). Blue points indicate speakers who produced no ejectives.

Collectively, the ejectives were produced by 22 of the 27 speakers. Five speakers did not produce any ejectives at all. Inter-speaker variability in ejective use is illustrated in **Figure 4.25**. Among the 22 speakers who did produce ejectives, individual counts ranged from just one occurrence to a maximum of 58 tokens, indicating substantial variation in ejective usage across speakers.

Before proceeding to the acoustic analysis of the identified ejective realizations, several additional aspects of the auditory categorization warrant discussion. These include the role of glottal replacements, the treatment of auditorily unclear tokens, and the reliability of the categorization process itself. These points raise important methodological and interpretive considerations that inform how the following acoustic and articulatory results should be understood.

4.3.2. The Glottal Replacement Category

Although glottal replacements were not the primary focus of this analysis, a brief consideration of their distribution helps contextualize their exclusion from the rest of the study and highlights the distinct conditioning patterns that separate them from ejectives. Glottal stops were found to replace exclusively voiceless, and predominantly unaccented alveolar stops. They accounted for 2.5% of all alveolar stops, 4.4% of deaccented stops, and 3.7% of voiceless stops. These results align with previously reported conditioning factors favoring glottal replacement of stops in English, a well-documented phenomenon.

The overall proportion of glottal replacements in this dataset was substantially lower than that of ejective realizations. It is likely that this low glottal replacement rate reflects characteristics of the experimental context, where the laboratory context encouraged more careful speech and fuller articulation of coda stops. Conversely, this may have promoted more ejective realizations, which have been linked to emphatic speech in English (Wells, 1982). These factors suggest that the current data is not well-suited for investigating glottal replacement patterns in the context of their potential relationship to ejectivization. For these reasons, glottal replacement tokens were excluded from the statistical analyses that follow.

4.3.3. The Unclear Category

The high proportion of tokens categorized as unclear deserves particular attention, as these items sit at the perceptual boundary of what counts as an ejective and may hold important implications for understanding gradient realizations. In the auditory analysis, 31.1% of tokens were categorized as unclear, a result that reflects the conservative approach adopted during the auditory labeling.

As shown in **Figure 4.26**, specifically, 34.4% of all bilabials, 34.4% of all alveolars and 31% of all velar coda stops were judged auditorily ambiguous with respect to how ejective-like they sound.

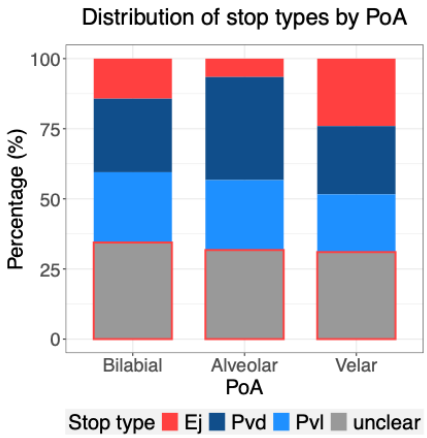

Figure 4.26 Distribution of stop type categories by PoA. Bars represent the proportional distribution (in %) of ejective, pulmonic voiced and voiceless, and unclear realizations across bilabial, alveolar, and velar PoAs. The grey segment (unclear category) is highlighted with a red outline to emphasize the high proportion of tokens that were auditorily ambiguous.

These tokens occupy precisely the zone of perceptual ambiguity – "is it an ejective or not, and why?" – that has yet to be systematically explored. Without necessarily seeking to impose a definitive categorization of these stops in the end, the following analyses aim to further probe these ambiguous realizations to better understand the acoustic and articulatory properties, and their combinations, that give rise to the percept of ejectiveness and underlie variability in its production.

4.3.4. Interim Discussion of the Auditory Categorization analysis

Ejectives were identified in 9.4% of all word-final stops. The results showed that accented, voiceless velars represent the most favorable context for ejective realization in English. These findings align with previous work on the conditions influencing ejective production in English (Gordeeva & Scobbie, 2011; Ogden, 2024).

Notably, 4.7% of ejectives occurred in originally voiced stops. This supports Gordeeva and Scobbie's (2011) proposal that ejective allophones in English reflect phonetic voicelessness rather than functioning as phonological correlates of [-voice]. Although ejectives were occasionally found in unstressed syllables, their very low frequency and uneven distribution precluded further analysis in the present study.

Finally, while some have proposed that ejectives are becoming more common in English (Simpson, 2014), this study provides no diachronic evidence to support such a claim. Without longitudinal data, the question of whether ejective usage is increasing remains unresolved. With the ejective realizations identified and their distributional patterns established, we now turn to an analysis of their acoustic properties.

4.4. Acoustic Analysis

The aim of the acoustic analysis was to determine whether the auditorily identified ejectives differed significantly from pulmonic stops in their acoustic characteristics, specifically, whether they constitute a measurably distinct stop type in English. The results for closure duration, relative burst intensity, and PBL intensity are presented below.

4.4.1. Closure Duration

Figure 4.27 presents closure duration (in ms) across three PoAs for the four stop types. Overall, ejectives and pulmonic voiceless stops exhibit similarly long closure durations, both consistently exceeding those of pulmonic voiced stops, which are shortest across PoAs. At the bilabial PoA (left panel), ejectives stand out with even longer closures than voiceless stops. The unclear category shows intermediate values, between ejectives and voiced stops at the alveolar and velar PoAs, and overlaps largely with ejectives at the bilabial PoA.

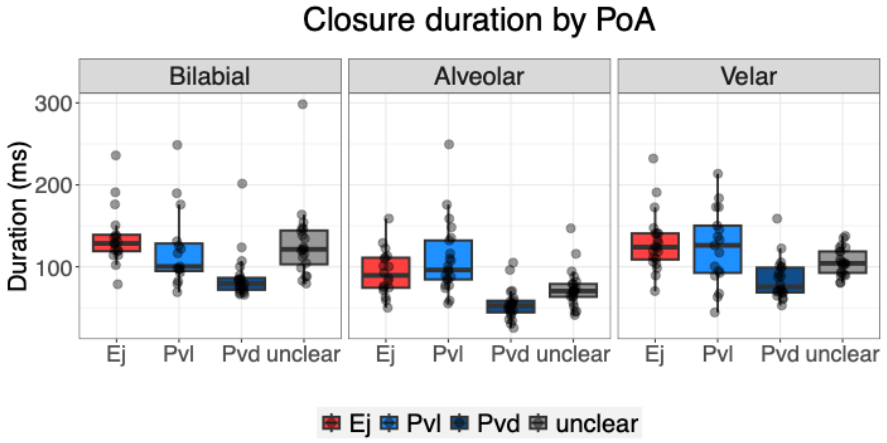

Figure 4.27 Raw closure duration values (y-axis, in ms) for four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear) shown across three PoAs (bilabial, alveolar, velar). Dots represent speaker-level means; higher values indicate longer durations.

Results of the statistical analysis revealed a significant main effect of stop type on log-transformed closure duration ($F(3, 20.12) = 5.98$, $p < .01$), as well as a significant effect of PoA ($F(2, 123.72) = 23.05$, $p < .001$). However, the interaction between stop type and PoA was not supported ($F(6, 2667.65) = 0.25$, $p = .96$) and was not retained in the final model.[19]

Table 4.25 Pairwise comparisons of closure duration (modelled on the log scale) by stop type and PoA, with back-transformed absolute differences in ms.

Contrast type	Contrast	Estimate (log)	SE	DF	t	p	Signif.	Δ (ms)
By stop type	Ej - Pvl	0.14	0.06	20.7	2.49	0.0916	.	12.94
By stop type	Pvd - Pvl	-0.07	0.05	24.9	-1.38	0.5218		-5.62
By stop type	Pvd - Ej	-0.20	0.05	23.5	-3.78	0.0048	**	-18.55
By stop type	unclear - Pvl	0.08	0.05	21.6	1.59	0.4040		7.29
By stop type	unclear - Ej	-0.06	0.03	20.1	-1.67	0.3659		-5.65
By stop type	unclear - Pvd	0.15	0.04	22.4	4.07	0.0026	**	12.90
By PoA	bil - alv	0.51	0.10	128.2	5.11	<.001	***	43.73
By PoA	vel - alv	0.47	0.09	128.8	5.38	<.001	***	39.55
By PoA	vel - bil	-0.04	0.12	128.2	-0.33	0.9428		-4.19

For PoA, closure durations were significantly longer at both the bilabial ($\Delta = 43.73$ ms, $p < .001$) and velar ($\Delta = 39.55$ ms, $p < .001$) place compared to the alveolar place. No significant difference was observed between bilabial and velar stops (bilabial-velar: $\Delta = 4.19$ ms, $p = .94$).

As shown in **Table 4.25**, pairwise comparisons for the main effect of stop type confirm the visual impression from **Figure 4.27:** ejectives had significantly longer closure durations than pulmonic voiced stops (ejective-voiced: $\Delta = 18.55$ ms, $p < .01$), as did unclear stops (unclear-voiced: $\Delta = 12.90$ ms, $p < .01$). No other stop type comparisons reached significance.

In sum, closure duration varied systematically by PoA, with bilabial and velar closures generally longer than alveolar ones. While ejectives were distinguished from pulmonic voiced stops by longer closure durations, they did not differ significantly from voiceless stops.

[19] Selected model: closureDuration_log ~ stopType + PoA + (stopType | speaker) + (1 | word).

4.4.2. Relative Burst Intensity

Figure 4.28 displays relative burst intensity (vowel peak minus burst peak, in dB) for four stop types across three PoAs. Values are not uniformly positive or negative across stops types or PoAs, indicating that burst peaks may either exceed or fall below the following vowel peak. Negative values, most prominent for ejectives at bilabial and alveolar places, indicate that the burst is louder than the vowel.

Nevertheless, a stable ordering emerges across PoA: overall, ejectives show the lowest relative burst intensity, followed by pulmonic voiceless stops, then the unclear category, while pulmonic voiced stops show the highest values. In addition, relative burst intensity generally increases toward velar PoA (right panel), where ejective values approach zero and thus only marginally exceed the following vowel in intensity.

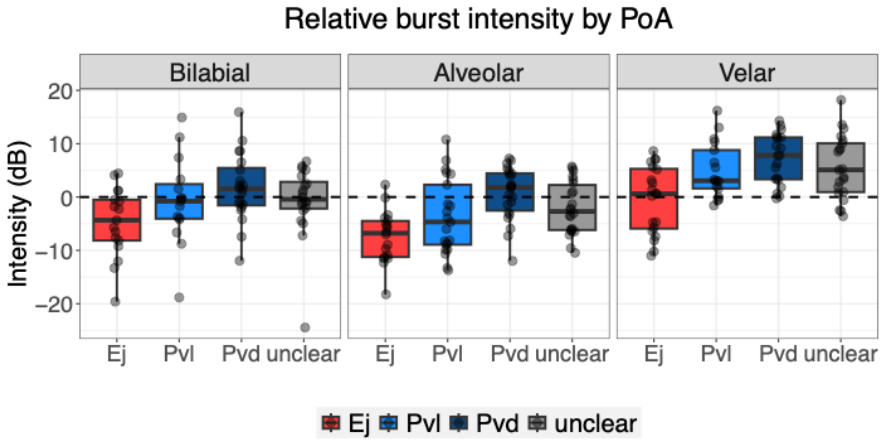

Figure 4.28 Raw relative burst intensity values (y-axis, in dB) for four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear) shown across three PoAs (bilabial, alveolar, velar). Dots represent speaker-level means. Values reflect the peak intensity difference between the following vowel and the stop's release burst. Lower values (and especially negative values) indicate a stronger intensity of the burst relative to the vowel.

The results of the statistical analysis revealed significant main effects of both stop type ($F(3, 19.77) = 44.89$, $p < .001$) and PoA ($F(2, 128.45) = 232.24$, $p < .001$) on relative burst intensity. The interaction term (stop type × PoA) was removed from the final model due to severe collinearity issues, which were resolved upon its exclusion.[20]

[20] Selected model: relBurstIntensity ~ stopType + PoA + (stopType | speaker) + (1 | word).

Table 4.26 Pairwise comparisons of relative burst intensity (in dB) by stop type and PoA.

Contrast type	Contrast	Estimate	SE	DF	t	p	Signif.
By stop type	Ej - Pvl	-3.79	0.69	20.5	-5.53	<.001	***
By stop type	Pvd - Pvl	3.54	0.72	21.3	4.95	<.001	***
By stop type	Pvd - Ej	7.34	0.67	21.3	10.91	<.001	***
By stop type	unclear - Pvl	0.72	0.44	18.2	1.64	0.3831	
By stop type	unclear - Ej	4.52	0.52	18.7	8.65	<.001	***
By stop type	unclear - Pvd	-2.82	0.54	21.0	-5.25	<.001	***
By PoA	bil - alv	2.51	0.38	131.7	6.64	<.001	***
By PoA	vel - alv	7.20	0.34	140.9	21.29	<.001	***
By PoA	vel - bil	4.69	0.45	132.4	10.37	<.001	***

As shown in **Table 4.26,** pairwise comparisons for the significant main effect of stop type confirmed that ejectives exhibited significantly lower relative burst intensity than all other stop types (ejective-voiceless: $\Delta = -3.79$ dB, $p < .001$; ejective-voiced: $\Delta = -7.34$ dB, $p < .001$; ejective-unclear: $\Delta = -4.52$ dB, $p < .001$). Unclear stops had significantly lower relative burst intensity than pulmonic voiced stops (unclear-voiced: $\Delta = -2.82$ dB, $p < .001$), but did not differ significantly from pulmonic voiceless stops (unclear-voiceless: $\Delta = 0.72$ dB, $p = .38$).

Pairwise comparisons for the main effect of PoA confirmed that velar stops had significantly higher relative burst intensity than both alveolar (velar-alveolar: $\Delta = 7.20$ dB, $p < .001$) and bilabial stops (velar-bilabial: $\Delta = 4.69$ dB, $p < .001$). Bilabial stops had significantly higher relative burst intensity than alveolar stops (bilabial-alveolar: $\Delta = 2.51$ dB, $p < .001$).

To summarize, ejectives were categorically distinct from both pulmonic stop types across all three PoAs in relative burst intensity, exhibiting lower values. Within ejectives, velars showed the highest relative burst intensity values.

4.4.3. Post-Burst Lag Intensity

Figure 4.29 illustrates PBL intensity (in dB) for the four stop types, across three PoAs. Overall patterns are less pronounced than those for burst intensity, partly due to greater variability in PBL values across PoAs. Nevertheless, ejective stops show a tendency to cluster at the lower end of the PBL intensity range across all PoAs. A similar, though

weaker, pattern is observed for the unclear category, particularly at the bilabial and alveolar PoAs (left and middle panels).

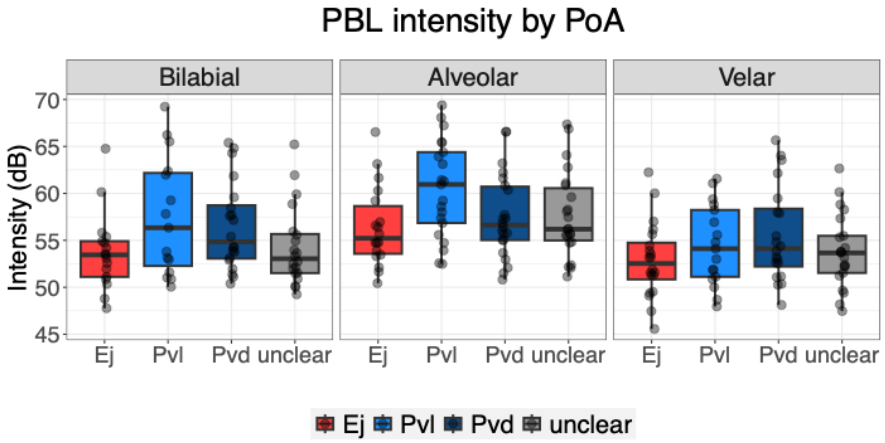

Figure 4.29 Intensity values of the post-burst lag (y-axis, in dB) for four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear) shown across three PoAs (bilabial, alveolar, velar). Dots represent speaker-level means; lower values indicate a quieter post-burst lag.

The statistical analysis of PBL intensity showed significant main effects of both stop type ($F(3, 18.67) = 8.83$, $p < .001$) and PoA ($F(2, 125.68) = 207.73$, $p < .001$). The interaction term (stop type × PoA) was removed from the final model due to severe collinearity issues, which were resolved upon its exclusion.[21]

Pairwise comparisons for the main effect of stop type given in **Table 4.27** revealed that ejectives exhibited significantly lower PBL intensity than pulmonic voiceless stops (ejective-voiceless: Δ = -3.46 dB, $p < .001$), but did not differ significantly from either voiced (ejective-voiced: Δ = -1.43 dB, $p = .13$) or unclear stops (ejective-unclear: Δ = -1.14 dB, $p = .10$).

Unclear stops followed a similar pattern, showing significantly lower PBL intensity than voiceless stops (unclear-voiceless: Δ = -2.32 dB, $p < .01$), but no significant difference from voiced stops (unclear-voiced: Δ = -0.30 dB, $p = .83$). Regarding PoA, bilabial stops showed significantly lower PBL intensity than alveolars (bilabial-alveolar: Δ = -2.84 dB, $p < .001$), while velar stops had significantly lower intensity than both alveolars (velar-alveolar: Δ = -3.76 dB, $p < .001$) and bilabials (velar-bilabial: Δ = -0.92 dB, $p = .0030$).

[21] Selected model: pblIntensity ~ stopType + PoA + (stopType | speaker) + (1 | word).

Table 4.27 Pairwise comparisons of PBL intensity (in dB) by stop type and PoA.

Contrast type	Contrast	Estimate	SE	DF	t	p	Signif.
By stop type	Ej - Pvl	-3.46	0.68	20.4	-5.06	<.001	***
By stop type	Pvd - Pvl	-2.02	0.65	20.7	-3.12	0.0247	*
By stop type	Pvd - Ej	1.43	0.62	21.4	2.32	0.1256	
By stop type	unclear - Pvl	-2.32	0.59	20.0	-3.97	0.0039	**
By stop type	unclear - Ej	1.14	0.47	20.3	2.44	0.1008	
By stop type	unclear - Pvd	-0.30	0.35	21.0	-0.84	0.8338	
By PoA	bil - alv	-2.84	0.23	132.4	-12.36	<.001	***
By PoA	vel - alv	-3.76	0.21	141.1	-18.34	<.001	***
By PoA	vel - bil	-0.92	0.28	133.0	-3.35	0.0030	**

Taken together, English ejectives and unclear stops both showed lower post-burst lag intensity than pulmonic voiceless stops. Neither ejectives nor unclear stops differed significantly from pulmonic voiced stops in the intensity of their post-burst interval.

4.4.4. Summary and Interim Discussion of Acoustics

The results showed that ejectives were robustly distinct from the other stop types in their burst-vowel intensity relationship, in that ejectives had consistently lower relative burst intensity values. PoA also shaped this relationship: while bilabial and alveolar stops generally showed stronger burst than vowel peaks, velars exhibited a reduced contrast, with vowel and burst peaks approaching similar intensity.

Inspection of the components of the measure revealed no systematic difference in vowel peak intensity across stop types. The observed differences therefore reflect variation in burst peak intensity. Ejective burst peaks were stronger than those of pulmonic stops overall, confirming that English ejectives exhibit a key acoustic hallmark of phonological ejectives and pointing to greater pressure buildup during production. Within the ejective series, however, velars displayed comparatively weaker burst releases. This PoA pattern runs counter to the expectation that more posterior stops should produce more forceful bursts due to facilitated pressure buildup, which would predict the lowest relative values for velars.

Low PBL intensity for ejectives – consistent with expectations of glottal closure persisting briefly after oral release – successfully distinguished English ejectives only from pulmonic *voiceless* stops. In contrast, longer closure durations distinguished ejectives from their *voiced* counterparts.

The findings from the acoustic analysis also help clarify the perceptual ambiguity of tokens categorized as unclear. Like ejectives, these stops patterned with voiced stops in having longer closure durations, and with voiceless stops in showing reduced PBL intensity. However, their relative burst intensity was intermediate, falling between ejectives and voiceless stops. This pattern suggests that, although these tokens share some properties with ejectives, their perceptual ambiguity may be due to the absence of acoustic features associated with a forceful release burst.

Having established the acoustic characteristics that distinguish English ejectives from their pulmonic counterparts, we now turn to the articulatory domain. The aim is to investigate how these sounds are produced and the extent to which they align with glottalic or alternative initiation mechanisms.

4.5. Articulatory Analysis

Building on the auditory and acoustic findings, the next step was to examine the articulatory mechanisms underlying ejective production in English. The focus here was on identifying how stops previously classified as ejectives, either clearly or ambiguously, differed articulatorily from their pulmonic counterparts, and on uncovering the strategies that may account for their perceived ejectiveness. This section presents results from three articulatory measures: *delta larynx, delta cavity*, and *delta hyperph* (see Chapter 2 for methodological details).

4.5.1. Larynx Raising

Given that the English phonological system lacks a foundation for glottalic stop production, there is little reason to assume that English ejectives conform to canonical models of ejective production. The first issue examined here, therefore, concerned the extent to which these sounds involved vertical larynx displacement at all. To address this, the first analysis in this section employed the *delta larynx* measure, which quantifies vertical larynx displacement (in mm) during stop closure.

Figure 4.30 compares *delta larynx* values across stop types and PoAs. Overall, English ejectives do not exhibit systematic larynx raising. A gradient is visible across PoAs, with velar ejectives showing the most positive values (i.e., more consistent raising), alveolars clustering around zero, and bilabials showing negative medians (i.e., larynx lowering).

Pulmonic voiced stops cluster around or just below zero, indicating minimal displacement, typically downward. Pulmonic voiceless stops show a clearer trend toward larynx lowering, with delta values falling below zero across all PoAs. The unclear category shows negative deltas at bilabial and alveolar PoAs but a weak positive trend at velars. Speaker variability is high, in that some speakers raise the larynx for velar or alveolar ejectives, while others do not, even for velars.

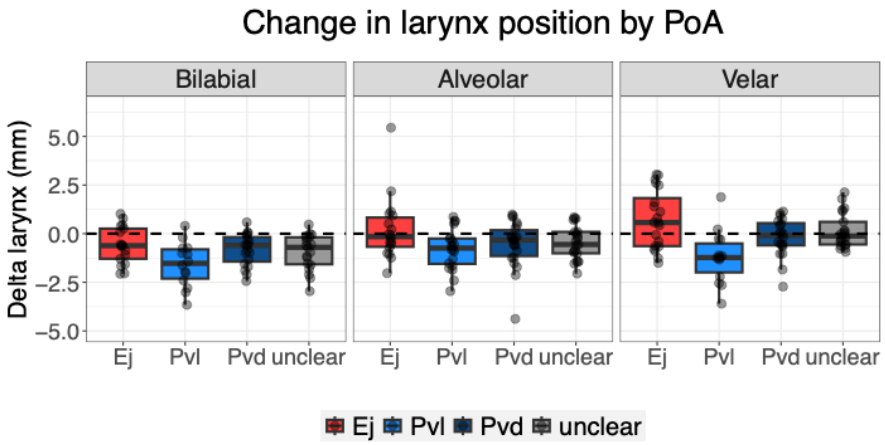

Figure 4.30 Delta larynx values (y-axis, in mm) for the four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear) shown across PoA (bilabial, alveolar, velar). Dots represent speaker-level means; positive numbers indicate larynx raising, negative numbers indicate larynx lowering.

Delta larynx was log-transformed to better meet model assumptions. The analysis showed significant main effects of stop type ($F(3, 34.08) = 9.54$, $p < .001$) and PoA ($F(2, 134.57) = 6.54$, $p < .01$), as well as a significant stop type and PoA interaction ($F(6, 2573.55) = 7.73$, $p < .001$).[22]

Pairwise comparisons shown in **Table G.42** and **Table G.43**, provided in Appendix G, revealed that, at the alveolar PoA, ejectives differed significantly from all other stop types: pulmonic voiceless ($\Delta = 0.42$ mm, $p < .01$), unclear ($\Delta = 0.27$ mm, $p < .05$), and pulmonic voiced stops ($\Delta = 0.26$ mm, $p < .05$).

[22] Selected model: deltaLarynx_log ~ stopType*PoA + (stopType | speaker) + (1 | word).

However, the model-adjusted prediction for alveolar ejectives indicated virtually no change in larynx position (predicted delta = 0.00 mm; see **Table 4.28**), meaning the difference reflects lowering in other types, not raising in ejectives.

Table 4.28 Estimated marginal means for delta larynx (model-adjusted predictions) by stop type and PoA.

Stop type	PoA	Predicted value	SE	95% CI
Pvl	bil	-0.78	0.18	(-1.14, -0.41)
Ej	bil	-0.32	0.17	(-0.65, 0.01)
Pvd	bil	-0.50	0.15	(-0.8, -0.19)
unclear	bil	-0.56	0.14	(-0.84, -0.29)
Pvl	alv	-0.54	0.15	(-0.85, -0.23)
Ej	alv	0.00	0.14	(-0.28, 0.28)
Pvd	alv	-0.30	0.11	(-0.52, -0.08)
unclear	alv	-0.31	0.11	(-0.53, -0.08)
Pvl	vel	-0.86	0.17	(-1.21, -0.5)
Ej	vel	0.33	0.15	(0.03, 0.63)
Pvd	vel	0.00	0.14	(-0.27, 0.28)
unclear	vel	-0.01	0.13	(-0.27, 0.25)

At the velar PoA, ejectives also differed significantly from pulmonic voiceless ($\Delta = 0.97$ mm, $p < .001$) and unclear stops ($\Delta = 0.40$ mm, $p < .05$), but not from pulmonic voiced stops ($\Delta = 0.39$ mm, $p = .08$). Velar ejectives were the only stop type to show predicted larynx *raising*, indicated by the positive model-adjusted delta of 0.33 mm (see **Table 4.28**).

Bilabial ejectives did not differ significantly from other stop types. They were also predicted to involve slight larynx lowering (predicted delta = -0.32 mm; see **Table 4.28**), meaning that in terms of vertical larynx movement, bilabial ejectives are the least differentiated from their pulmonic counterparts at the same PoA.

The unclear stop category showed variable behavior across PoAs. According to **Table 4.28**, alveolar (predicted delta = -0.31 mm) and bilabial (predicted delta = -0.56 mm) unclear stops were not predicted to involve larynx raising and did not differ significantly from any of their pulmonic counterparts (see **Table G.43**).

Velar (predicted delta = -0.01 mm) unclear stops, although not predicted to raise the larynx, nevertheless differed significantly from pulmonic voiceless stops ($\Delta = 0.57$ mm, $p < .001$) in terms of larynx movement.

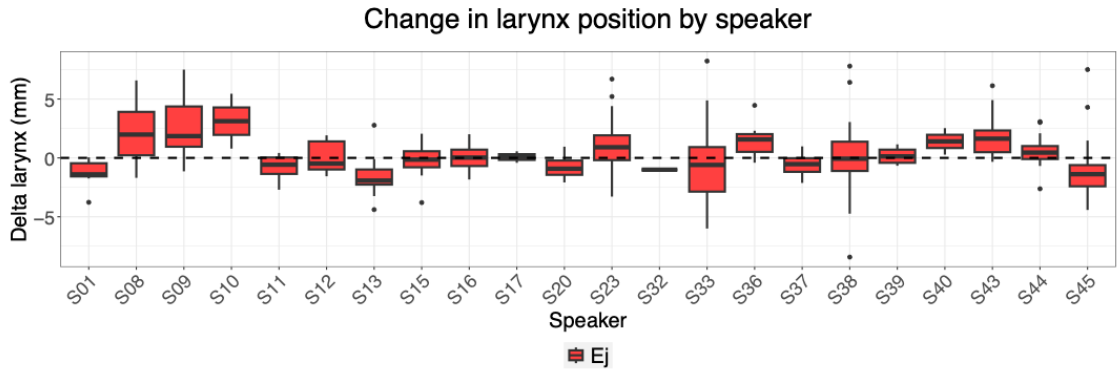

Figure 4.31 Delta larynx values (y-axis, in mm) for ejectives by speaker. Delta larynx is measured as the difference between the larynx's position at the onset of stop closure and its position at closure offset. Positive numbers indicate larynx raising, negative numbers indicate larynx lowering.

Figure 4.31 emphasizes that larynx raising for ejectives in English is not only PoA dependent, but also highly speaker dependent. Inter-speaker variability involves not only differences in the *magnitude* of laryngeal movement, but also differences in whether speakers raise the larynx at all. Some speakers show consistent upward movement and others exhibit minimal or no raising. This pattern of speaker-dependent variation is further illustrated in **Table 4.29**, which provides an overview of mean *delta larynx* values, standard deviations, and the number of ejective realizations for each speaker.

Table 4.29 Mean (M) and standard deviation (SD) of delta larynx (in mm) for ejectives by speaker.

Speaker	**S01**	**S08**	**S09**	**S10**	**S11**	**S12**	**S13**	**S15**	**S16**	**S17**	**S20**
M (SD)	-1.3 (1.1)	1.9 (2.6)	2.4 (2.0)	3.1 (3.3)	-0.8 (1)	0.0 (1.4)	-1.7 (1.4)	-0.3 (1.4)	0.1 (1.0)	0.1 (0.7)	-0.7 (1.3)
Speaker	**S23**	**S32**	**S33**	**S36**	**S37**	**S38**	**S39**	**S40**	**S43**	**S44**	**S45**
M (SD)	1.1 (1.8)	-1 (NA)	-0.6 (3.3)	1.5 (1.6)	-0.6 (0.8)	0.0 (3.4)	0.2 (0.7)	1.4 (1.6)	1.7 (1.5)	0.5 (1.0)	-1.3 (2.1)

Taken together, these findings showed that in English, although larynx displacement distinguished both alveolar and velar ejectives from other stop types, only velar ejectives exhibited larynx raising in a manner consistent with glottalic initiation.

More broadly, larynx behavior was highly speaker-dependent, with some speakers employing it regardless of PoA, and others not at all. Notably, among *unclear* stops, velars alone showed a subtle, speaker-driven upward larynx trend, mirroring that of their ejective counterparts.

4.5.2. Intraoral Cavity Reduction

The preceding analysis demonstrated that vertical larynx displacement was inconsistent in English ejectives. The next step was therefore to examine whether intraoral pressure was instead generated through supralaryngeal cavity reduction. For this, the *delta cavity* measure was used, which quantifies changes in supralaryngeal cavity size (in mm^2) during stop closure independently of larynx movement.

Figure 4.32 compares *delta cavity* values across stop types and PoAs. Compared to the larynx raising patterns described above, ejectives exhibit consistently positive *delta cavity* values at the bilabial and velar PoAs (left and right panels), indicating cavity size reduction. In contrast, pulmonic stops typically show negative values, reflecting expansion. The difference is especially pronounced at the velar PoA. Unclear stops also show cavity reduction at these two PoAs, albeit to a lesser extent.

The alveolar PoA (center panel) diverges from the overall pattern: here, all stop types, including ejectives, show overlapping delta cavity values, with medians clustering around zero, suggesting no reliable cavity size changes. At this PoA, individual speaker values vary widely in both directions across all stop categories.

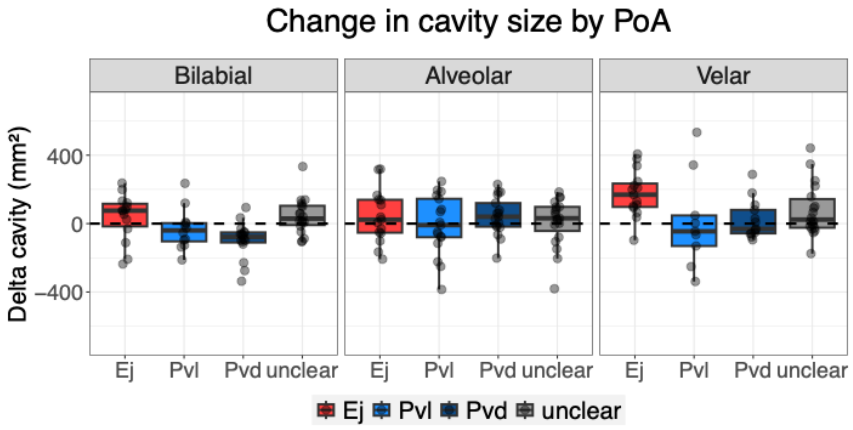

Figure 4.32 Delta cavity values (y-axis, in mm^2) for the four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear) shown across PoA (bilabial, alveolar, velar). Delta cavity is measured as the difference between the cavity's size at the onset of stop closure and its size at closure offset. Dots represent speaker-level means; positive numbers indicate cavity reduction, negative numbers indicate cavity expansion.

The statistical analysis revealed a significant main effect of stop type ($F(3, 1044.91) = 9.38$, $p < .001$) but not of PoA ($F(2, 89.58) = 0.75$, $p = .47$) on *delta cavity*. The stop type × PoA interaction was removed from the final model due to severe collinearity issues, which were resolved upon its exclusion.[23]

Table 4.30 Pairwise comparisons of delta cavity (mm^2) by stop type.

Contrast type	Contrast	Estimate (mm^2)	SE	DF	t	p	Signif.
By stop type	Ej - Pvl	114.61	24.48	1,070.5	4.68	<.001	***
By stop type	Pvd - Pvl	25.60	24.57	902.2	1.04	0.7250	
By stop type	Pvd - Ej	-89.01	22.78	1,029.0	-3.91	<.001	***
By stop type	unclear - Pvl	73.75	21.37	1,066.7	3.45	0.0032	**
By stop type	unclear - Ej	-40.86	17.95	1,077.2	-2.28	0.1044	
By stop type	unclear - Pvd	48.16	18.30	1,054.2	2.63	0.0428	*

The pairwise comparisons for the main effect of stop type summarized in **Table 4.30**, indicated greater cavity size changes for ejectives than for pulmonic voiceless (ejective-voiceless: $\Delta = 114.61$ mm², $p < .001$) and pulmonic voiced (ejective-voiced: $\Delta = 89.01$ mm², $p < .001$), but not unclear stops (ejective-unclear: $\Delta = 40.86$ mm², $p = .10$). Unclear stops showed significantly greater cavity size changes than pulmonic voiced (unclear-voiced: $\Delta = 48.16$ mm², $p < .05$) and pulmonic voiceless stops (unclear-voiceless: $\Delta = 73.75$ mm², $p < .01$).

Speaker-level means and variability in cavity size reduction are summarized in **Table 4.31**. This table illustrates the wide individual differences observed even within the ejective category.

Table 4.31 Mean (M) and standard deviation (SD) of delta cavity (mm^2) for ejectives by speaker. Three speakers are not included in this table, as no ejective tokens remained for them following data exclusion, and they were therefore omitted from the cavity size analysis.

Speaker	S01	S08	S09	S11	S12	S13	S15	S16	S17	S23
M	97.0	126.0	-105.1	73.0	64.8	16.3	-98.5	137.4	209.3	242.2
(SD)	(316.7)	(186.9)	(313.0)	(112.9)	(486.4)	(269.8)	(125.9)	(149.5)	(NA)	(388.5)
Speaker	**S32**	**S33**	**S36**	**S37**	**S38**	**S39**	**S43**	**S44**	**S45**	
M	-165.4	153.1	286.8	41.7	155.8	182.1	192.1	64.4	143.9	
(SD)	(NA)	(123.7)	(126.8)	(121.4)	(183.0)	(194.4)	(210.0)	(100.4)	(372.0)	

[23] Selected model: deltaCavity ~ stopType + PoA + (1 | speaker) + (1 | word) + (1 | V1).

Overall, changes in cavity size significantly distinguished stop types: ejectives consistently involved reduction, whereas pulmonic stops tended toward expansion. This makes cavity reduction a more stable feature of English ejectives than larynx raising. No statistically reliable place differences emerged, but visually, reduction at the alveolar place is not distinctive, as pulmonic stops there also show reduction. Unclear stops patterned similarly to ejectives.

4.5.3. Constriction of the Hyperpharyngeal Region

The third articulatory analysis examined whether hyperpharyngeal constriction, such as that resulting from tongue root retraction, played a role in English ejective production. This was assessed using *delta hyperph* (a normalized value), which captures changes in constriction within the hyperpharyngeal region.

Figure 4.33 shows *delta hyperph* values across stop types and PoAs. The figure highlights a contrast between the more anterior PoAs (bilabial and alveolar) and the velar PoA. Bilabial and alveolar stops (left and center panels) exhibit increased hyperpharyngeal constriction, with ejectives showing higher positive *delta hyperph* values than both pulmonic stop types. This contrast is most pronounced at the bilabial PoA.

At the velar PoA, ejectives show slightly negative values, indicating hyperpharyngeal constriction becomes less over the closure, but with speaker values split between positive and negative ranges. Pulmonic stops show a greater decrease in constriction overall, especially voiceless stops, yielding a clear contrast from ejectives despite the general trend toward negative values. Notably, unclear stops closely mirror the ejectives across all three PoAs.

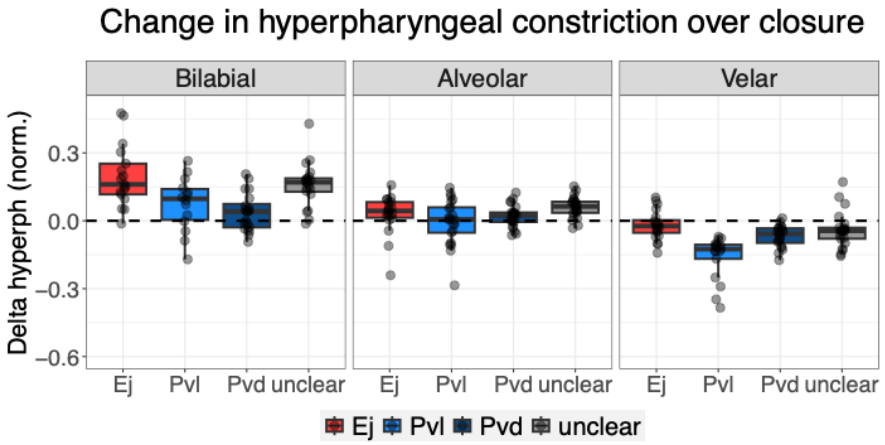

Figure 4.33 Delta hyperph values (y-axis, normalized) for the four stop types (ejective, pulmonic voiced, pulmonic voiceless, and unclear) shown across PoA (bilabial, alveolar, velar). Delta hyperph is measured as the difference between the degree of hyperpharyngeal constriction at the onset of stop closure and its constriction at closure offset. The black dashed line represents the zero value for the y-axis: observations above it indicate hyperpharyngeal constriction increased, below it indicates hyperpharyngeal constriction relaxed. Dots represent speaker-level means.

The statistical analysis confirmed significant main effects of stop type ($F(3, 22.71) = 21.41$, $p < .001$) and PoA ($F(2, 122.81) = 31.88$, $p < .001$), as well as a significant interaction between the two (stop type × PoA: $F(6, 2468.73) = 6.24$, $p < .001$).[24] Pairwise contrasts are presented in **Table G.44** and **Table G.45**, which are provided in Appendix G.

As shown in **Table G.44**, ejectives differed reliably from both pulmonic stop types at all three PoAs. Relative to pulmonic voiceless stops, ejectives showed greater change in hyperpharyngeal constriction at the bilabial ($\Delta = 0.09$, $p < .001$), alveolar ($\Delta = 0.03$, $p < .05$), and velar places ($\Delta = 0.10$, $p < .001$). The same pattern held in comparison with pulmonic voiced stops (bilabial: $\Delta = 0.04$, $p < .05$; alveolar: $\Delta = 0.04$, $p < .001$; velar: $\Delta = 0.05$, $p < .01$).

As suggested by the clear PoA-based trends visible in **Figure 4.33, Table G.45** showed that bilabial ejectives exhibited significantly greater change in hyperpharyngeal constriction than ejectives at the other two PoAs (bilabial-alveolar: $\Delta = 0.11$, $p < .001$; bilabial-velar: $\Delta = 0.12$, $p < .001$), while alveolar and velar ejectives did not differ (velar-alveolar: $\Delta = -0.01$, $p = .88$).

[24] Selected model: deltaHyperph ~ stopType * PoA + (1 | speaker) + (1 | word) + (1 | V1).

A closely parallel pattern was observed for unclear stops: bilabials differed from both alveolars (bilabial-alveolar: $\Delta = 0.10$, $p < .001$) and velars (bilabial-velar: $\Delta = 0.13$, $p < .001$), whereas he velar-alveolar difference did not reach significance (velar-alveolar: $\Delta = -0.03$, $p = .07$).

Changes in hyperpharyngeal constriction varied strongly with PoA but nevertheless distinguished ejectives significantly from other stop types at all PoAs. Ejectives showed increased constriction at bilabial and alveolar places, with the strongest effect at the bilabial place. At the velar place, ejectives remained distinct from pulmonic stops despite exhibiting a slight decrease in constriction.

4.5.4. Summary and Interim Discussion of Articulatory Results

The articulatory analyses showed that larynx raising is not a consistent feature of English ejectives. Only velar ejectives exhibited the larynx raising expected of glottalic initiation, while bilabials showed larynx lowering and alveolars showed no significant change in larynx position. Inter-speaker variability was notable, with some speakers showing larynx raising even for non-velar ejectives.

By contrast, supralaryngeal cavity reduction reliably differentiated ejectives from pulmonic stops, which generally showed expansion; the sole exception was the alveolar PoA, where reduction in ejectives was not distinctive. Hyperpharyngeal activity varied by PoA, with the strongest constriction for bilabial ejectives, minimal ejective-pulmonic differentiation at the alveolar PoA, and net expansion at the velar PoA, where ejectives nevertheless remained distinct from pulmonic stops.

Taken together, these results suggest that while larynx raising is not a reliable articulatory feature of English ejectives, supralaryngeal cavity modulation serves as a more consistent articulatory strategy. Having examined the articulatory components of English ejectives individually, we now turn to their interaction.

4.6. Articulatory Relationships and Acoustic Correlates

The preceding analyses showed that articulatory strategies, most notably larynx raising, were used variably in English ejectives. The next analyses therefore examined whether larynx raising, supralaryngeal cavity reduction, and hyperpharyngeal constriction were systematically coordinated (e.g., in additive or independent patterns) across PoA. Their contributions to absolute burst intensity were also assessed.

To shed light on the articulatory basis of perceptual ambiguity, the same relationships were explored in unclear stops. The following sections focused on ejective and unclear stops only, as the ejective-pulmonic contrast was previously established. As in the Amharic analysis, targeted models were used rather than a single comprehensive regression to maintain interpretability.

4.6.1. Linking Larynx Raising and Supralaryngeal Cavity Change

Figure 4.34 illustrates the relationship between *delta larynx* (x-axis, in mm) and *delta cavity* (y-axis, in mm^2) across three PoAs, based on a reduced dataset (732 tokens, 199 ejectives). Because this subset specifically underrepresents the number of velar ejectives with larynx raising, interpretations should be considered tentative.

Across all PoAs, ejectives and unclear stops display considerable variability in their articulatory characteristics, with tokens distributed across all quadrants of the plot. This includes cases with both larynx raising and cavity size reduction (top right quadrant), as well as cases with larynx lowering and cavity size reduction (top left quadrant), among other patterns.

At the alveolar PoA, both stop types show a small positive relationship, suggesting that increases in cavity size tend to be associated with larger upward shifts in larynx position. For bilabial stops, a similar positive trend is visible for unclear stops, whereas for ejectives, the regression line is relatively flat. In contrast, for velar stops, ejectives exhibit a negative relationship between cavity size and larynx position, suggesting that increases in cavity size are associated with smaller vertical larynx displacements. Unclear stops at the velar PoA show a flatter trend.

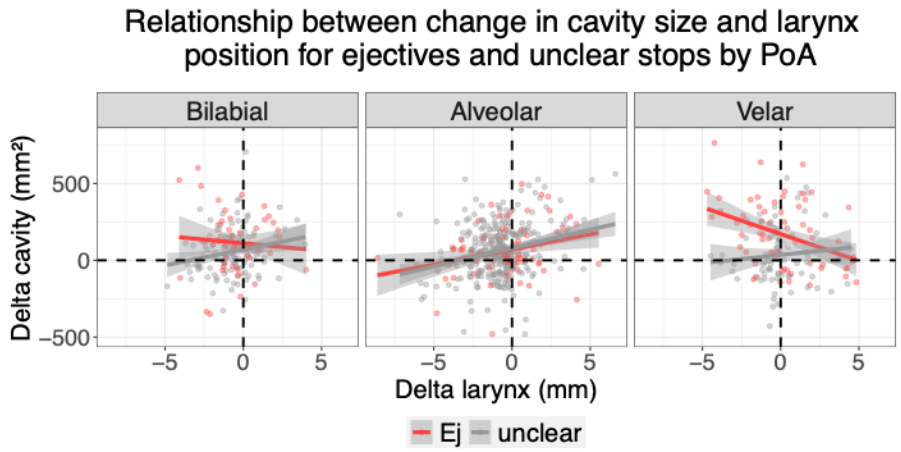

Figure 4.34 Scatter plot showing the relationship between delta larynx (x-axis, in mm) and delta cavity (y-axis, in mm^2) for ejectives and unclear stops by PoA (bilabial, alveolar, velar). Superimposed regression lines indicate the strength and direction of the relationship, with shaded regions representing 95% confidence intervals. The black dashed lines represent the zero values for both axes: observations to the left of the zero x-axis indicate larynx lowering, while observations below the zero y-axis indicate an increase in cavity size over the stop closure. Dots represent individual tokens.

Statistical analysis revealed a significant main effect of *delta larynx* ($F(1, 493.41) = 10.34$, $p < .01$). The main effect of stop type ($F(1, 680.01) = 1.70$, $p = .19$) was not significant, however, stop type did interact significantly with *delta larynx* ($F(1, 698.81) = 7.87$, $p < .01$). Neither PoA ($F(2, 91.33) = 0.09$, $p = .92$) nor its interaction with *delta larynx* ($F(2, 703.53) = 1.89$, $p = .15$) reached significance, suggesting that the divergent patterns observed between PoAs were not strong. Accordingly, PoA and its interaction with *delta larynx* were removed from the final model. [25]

Table 4.32 Estimated slopes for the relationship between delta larynx (in mm) and delta cavity (in mm^2) by stop type.

Estimate type	Level	Slope (mm² per mm)	SE	DF	t	p	Signif.
By stop type	Ej	3.60	6.87	647.9	0.52	0.6004	
By stop type	unclear	26.25	5.53	596.2	4.75	<.001	***

The estimated slopes of the *delta larynx* × stop type interaction (**Table 4.32**) revealed that both stop types displayed a positive association between larynx raising and cavity size reduction. However, this positive association was statistically reliable only for unclear stops (slope = 26.25 mm² per mm, $p < .001$) and not for ejectives (slope = 3.60 mm² per

[25] Selected model: deltaCavity ~ deltaLarynx*stopType + (1 | speaker) + (1 | word) + (1 | V1).

mm, p = .60). The slope difference was significant (ejective-unclear: Δ slope = -22.66 mm² per mm, p < .01; see **Table 4.33**).

Table 4.33 Pairwise comparisons of delta larynx (in mm) slopes by stop type.

Contrast Type	Contrast	Estimate (Δ slope)	SE	DF	t	p	Signif.
Slope difference by Stop type	Ej - unclear	-22.66	8.12	697.5	-2.79	0.0054	**

In sum, larynx displacement significantly predicted changes in supralaryngeal cavity size for unclear stops, but not for ejectives. No reliable effects of PoA were observed. This pattern should be interpreted with caution, as the analysis was based on a reduced dataset and crucially, underrepresented velar ejectives with larynx raising.

4.6.2. Linking Hyperpharyngeal Constriction and Supralaryngeal Cavity Change

This relationship analysis evaluated whether changes in the hyperpharyngeal region varied systematically with changes in cavity size. The aim was to understand the extent to which this local region contributes to global cavity compression.

Figure 4.35 illustrates the relationship between *delta hyperph* (x-axis, normalized) and *delta cavity* (y-axis, in mm^2) for ejectives and unclear stops across the three PoAs. Overall, the data points are distributed across multiple quadrants, particularly for alveolars, indicating that increases in hyperpharyngeal constriction and cavity reduction are not obligatorily coupled.

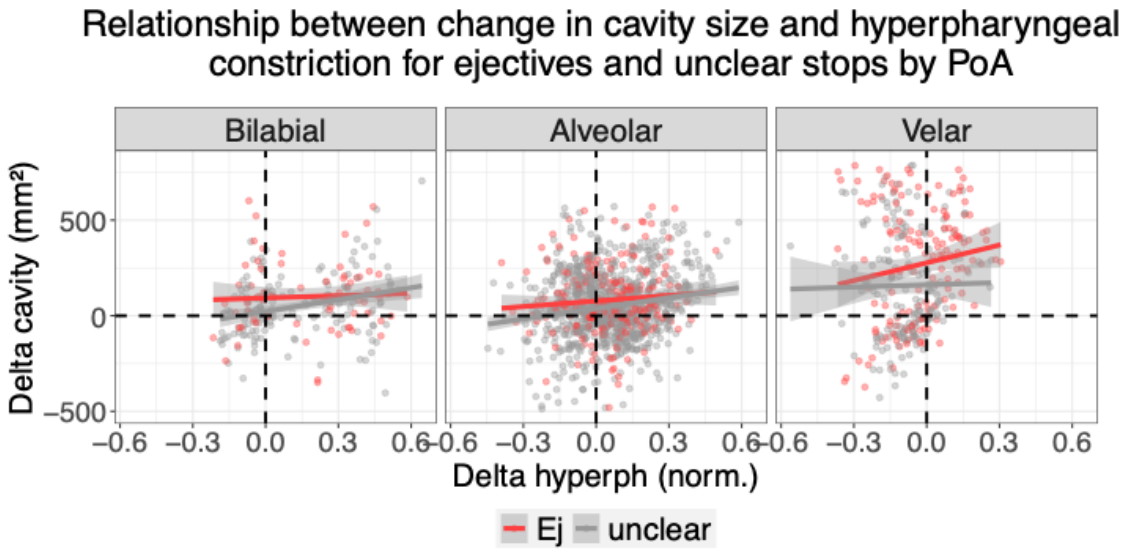

Figure 4.35 Scatter plot showing the relationship between delta hyperph (x-axis, normalized) and delta cavity (y-axis, in mm^2) for ejectives and unclear stops by PoA (bilabial, alveolar, velar). Superimposed regression lines indicate the strength and direction of the relationship, with shaded regions representing 95% confidence intervals. The black dashed lines represent the zero values for both axes: observations to the rigth of the zero x-axis indicate an increase in hyperpharyngeal constriction, while observations below the zero y-axis indicate an increase in cavity size over the stop closure. Dots represent individual tokens.

Statistical analysis revealed significant main effects of *delta hyperph* ($F(1, 1267.83) = 43.20$, $p < .001$) and PoA ($F(2, 141.00) = 19.14$, $p < .001$), as well as a significant delta hyperph × PoA interaction ($F(2, 788.85) = 17.38$, $p < .001$). The stop-type main effect and the delta hyperph × stop type interaction were not supported and were removed from the final model (stop type: $F(1, 18.70) = 2.17$, $p = .15$; delta hyperph × stop type: ($F(2, 1616.82) = 0.07$, $p = .80$).[26]

Table 4.34 Estimated slopes for the relationship between delta hyperph (normalized) and delta cavity (in mm^2) by PoA.

Estimate type	Level	Slope (mm^2 per unit)	SE	DF	t	p	Signif.
Slope by PoA	bil	680.47	121.71	600.0	5.59	<.001	***
Slope by PoA	alv	727.73	72.55	1,161.0	10.03	<.001	***
Slope by PoA	vel	-54.61	124.48	1,561.3	-0.44	0.6610	

[26] Selected model: deltaCavity ~ deltaHyperph*PoA + (stopType | speaker) + (1 | word) + (1 | V1).

Post hoc tests showed that the estimated slopes of *delta hyperph* (see **Table 4.34**) were significantly positive for alveolars (slope = 727.73 mm^2 per unit, $p < .001$) and bilabials (slope = 680.47 mm^2 per unit, $p < .001$), but not for velars (slope = -54.61 mm^2 per unit, $p = .66$).

Pairwise comparisons shown in **Table 4.35** confirmed that the slopes for alveolars and bilabials were each significantly greater than for velars (alveolar-velar: Δ slope = 782.34 mm^2 per unit, $p < .001$; bilabial-velar: Δ slope = 735.08 mm^2 per unit, $p < .001$), whereas the two anterior PoAs did not differ significantly from each other.

Table 4.35 Pairwise comparisons of delta hyperph (normalized) slopes across PoA.

Contrast Type	Contrast	Estimate (Δ slope)	SE	DF	t	p	Signif.
Slope difference by PoA	alv - bil	47.26	130.25	503.0	0.36	0.9300	
Slope difference by PoA	alv - vel	782.34	135.35	1,302.7	5.78	<.001	***
Slope difference by PoA	bil - vel	735.08	165.98	793.5	4.43	<.001	***

In sum, PoA-specific slopes estimated across stop types were significant and positive only for bilabial and alveolar stops.

4.6.3. Effect of Larynx Raising on Absolute Burst Intensity

Figure 4.36 presents scatter plots of the relationship between *delta larynx* (x-axis, in mm) and absolute burst intensity (y-axis, in dB) across three PoAs. Ejectives display a tentative trend toward higher absolute burst intensity with greater larynx raising for bilabials and velars (left and right panels respectively). Unclear stops show a similar weak relationship for bilabials and alveolars.

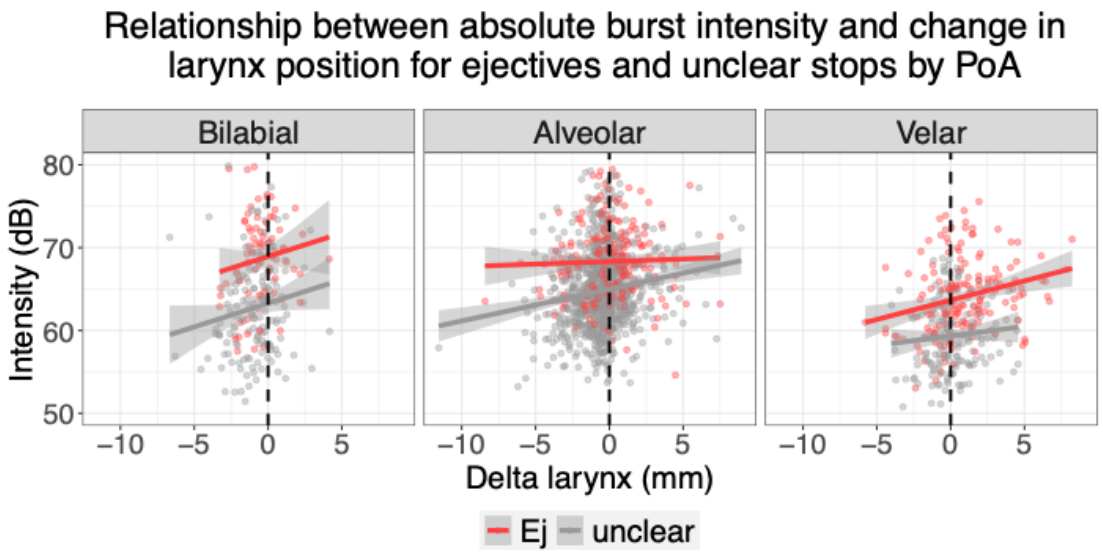

Figure 4.36 Scatter plot showing the relationship between delta larynx (x-axis, in mm) and absolutely burst intensity (y-axis, in dB) for ejectives and unclear stops by PoA (bilabial, alveolar, velar). Superimposed regression lines indicate the strength and direction of the relationship, with shaded regions representing 95% confidence intervals. The black dashed line represents the zero value for the x-axis: observations to the right of the line indicate larynx raising over the stop closure. Dots represent individual tokens.

Statistical analysis confirmed the visual impression of a weak positive relationship between *delta larynx* and absolute burst intensity (delta larynx: $F(1, 1346.38) = 4.84$, $p < .05$). However, *delta larynx* did not interact significantly with either stop type or PoA (delta larynx × stop type: $F(1, 1357.62) = 1.18$, $p = .28$; delta larynx × PoA: $F(2, 1530.93) = 1.49$, $p = .23$). As these terms did not improve model fit, they were removed during model selection.

In the final model[27], both stop type and PoA showed strong main effects on absolute burst intensity (all $p < .001$), consistent with earlier results (Section 4.4.2). In sum, larynx displacement showed only a weak overall association with absolute burst intensity, with no evidence that this relationship varied by stop type or PoA.

4.6.4. Effect of Cavity Size Reduction on Absolute Burst Intensity

Figure 4.37 presents scatter plots showing the relationship between *delta cavity* (x-axis, in mm²) and absolute burst intensity (y-axis, in dB) across three PoAs. Both ejectives and unclear stops exhibit generally flat trends, with no striking visual association between cavity reduction and absolute burst intensity.

[27] Selected model: burstIntensity ~ deltaLarynx + stopType + PoA + (stopType | speaker) + (1 | word).

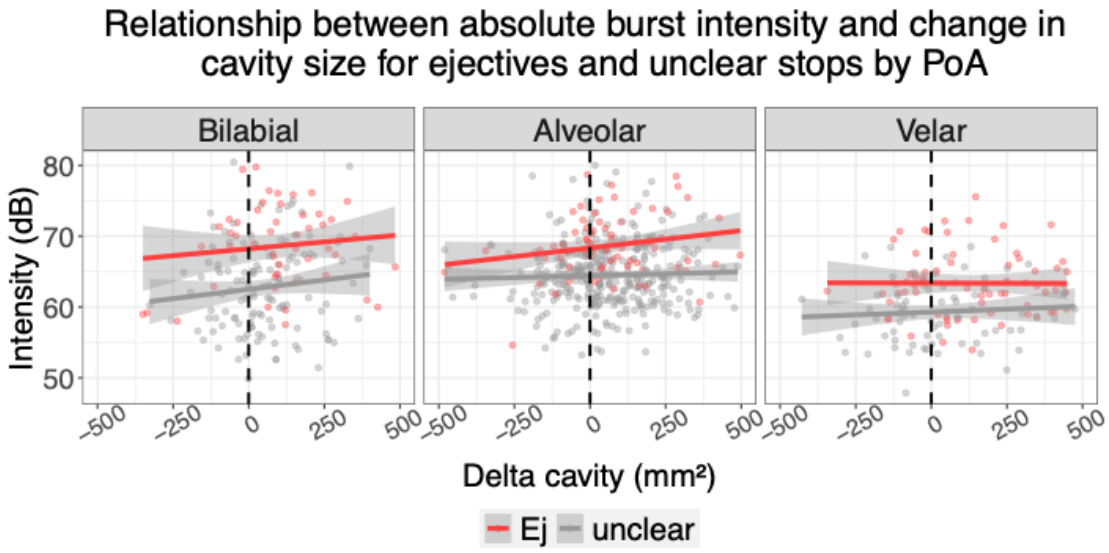

Figure 4.37 Scatter plot showing the relationship between delta cavity (x-axis, in mm^2) and absolutely burst intensity (y-axis, in dB) for ejectives and unclear stops by PoA (bilabial, alveolar, velar). Superimposed regression lines indicate the strength and direction of the relationship, with shaded regions representing 95% confidence intervals. The black dashed line represents the zero value for the x-axis: observations to the right of the line indicate a reduction in cavity size over the stop closure. Dots represent individual tokens.

Statistical analysis confirmed this visual impression. Interactions between *delta cavity* and stop type ($F(1, 715.23) = 1.49$, $p = .22$) and between *delta cavity* and PoA ($F(2, 716.34) = 0.44$, $p = .64$) were not significant and were removed during model selection. The main effect of *delta cavity* was likewise excluded ($F(1, 720.35) = 2.80$, $p = .10$).

The final model therefore retained only stop type and PoA as fixed effects.[28] Both showed strong effects on absolute burst intensity (both $p < .001$), consistent with the previous analysis (Section 4.4.2). Thus, changes in cavity size did not reliably predict absolute burst intensity.

4.6.5. Effect of Hyperpharyngeal Constriction on Absolute Burst Intensity

Hyperpharyngeal constriction was shown to play an important role in the production of bilabial and alveolar ejectives (Section 4.5.3), and Section 4.6.2 further demonstrated that it significantly contributed to overall cavity reduction for these stops. However, overall cavity compression itself did not predict absolute burst intensity.

[28] Selected model: burstIntensity ~ stopType + PoA + (1 | speaker) + (1 | V1).[29] Selected model: burstIntensity ~ deltaHyperph*PoA + stopType + (stopType | speaker) + (1 | word) + (1 | V1).

Against this background, the following analysis tested whether hyperpharyngeal constriction directly predicts absolute burst intensity, particularly for bilabial and alveolar stops. **Figure 4.38** illustrates the relationship between *delta hyperph* (x-axis, normalized) and absolute burst intensity (y-axis, in dB) across three PoAs.

The figure shows that, across all PoAs, ejectives pattern with higher absolute burst intensity overall than unclear stops. Otherwise, PoA-specific tendencies are less pronounced than in previous analyses, and the generally flat-trending lines suggest that hyperpharyngeal constriction does not strongly influence absolute burst intensity.

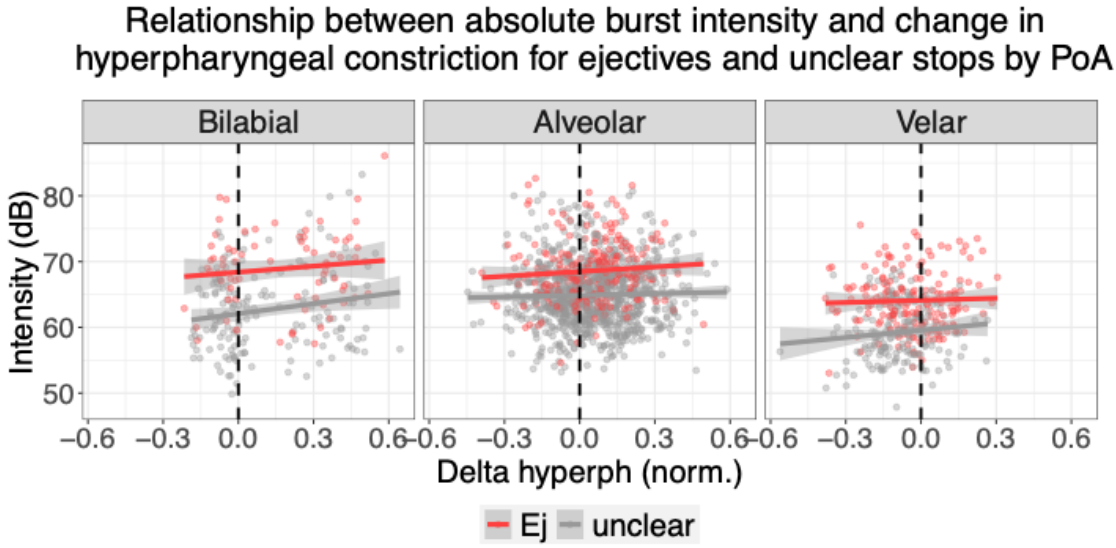

Figure 4.38 Scatter plot showing the relationship between delta hyperph (x-axis, normalized) and absolute burst intensity (y-axis, in dB) for ejectives and unclear stops by PoA (bilabial, alveolar, velar). Superimposed regression lines indicate the strength and direction of the relationship, with shaded regions representing 95% confidence intervals. The black dashed line represents the zero value for the x-axis: observations to the right of the line indicate an increase in hyperpharyngeal constriction over the stop closure. Dots represent individual tokens.

Statistical analysis showed a strong main effect of stop type ($F(1, 15.26) = 72.39$, $p < .001$) and PoA ($F(2, 154.17) = 136.78$, $p < .001$), however, the main effect of *delta hyperph* was not significant ($F(1, 147.67) = 3.69$, $p = .06$). The delta hyperph × stop type interaction was not supported and was removed from the final model ($F(1, 1681.34) = 1.18$, $p = .28$).[29]

[29] Selected model: burstIntensity ~ deltaHyperph*PoA + stopType + (stopType | speaker) + (1 | word) + (1 | V1).

The delta hyperph × PoA interaction was significant ($F(2, 230.84) = 3.97$, $p < .05$), indicating that the relationship between change in hyperpharyngeal constriction and absolute burst intensity varied by PoA.

Table 4.36 Estimated slopes for the relationship between delta hyperph (normalized) and absolute burst intensity (in dB) by PoA.

Estimate Type	Level	Slope (dB per unit)	SE	DF	t	p	Signif.
Slope by PoA	bil	3.93	1.51	199.2	2.60	0.0099	**
Slope by PoA	alv	-0.22	1.11	151.6	-0.20	0.8436	
Slope by PoA	vel	2.35	1.94	457.8	1.21	0.2263	

As **Table 4.36** shows, the estimated slope for alveolars was essentially zero (slope = -0.22 dB per unit, $p = .84$). In contrast, bilabials showed a significant positive slope (slope = 3.93 dB per unit, $p < .01$), which was significantly greater than that of alveolars (bilabial-alveolar: Δ slope = 4.15 dB per unit, $p < .05$; **Table 4.37**). The slope for velars was also positive (slope = 2.35 dB per unit, $p = .23$) but not significant.

Table 4.37 Pairwise comparisons of delta hyperph (normalized) slopes by PoA.

Contrast Type	Contrast	Estimate (Δ slope)	SE	DF	t	p	Signif.
Slope difference by PoA	alv - bil	-4.15	1.53	178.5	-2.70	0.0204	*
Slope difference by PoA	alv - vel	-2.57	1.94	333.4	-1.32	0.3826	
Slope difference by PoA	bil - vel	1.58	2.17	253.7	0.73	0.7478	

Overall, increases in hyperpharyngeal constriction were associated with higher absolute burst intensity only at the bilabial PoA, across ejective and unclear stops.

4.6.6. Summary of Relationship Analyses

A positive association between larynx displacement and cavity size changes was statistically supported only for unclear stops; for ejectives, no reliable coupling between these gestures was observed. Hyperpharyngeal constriction patterned by PoA rather than stop type: more anterior places showed a positive relationship with cavity size changes that was not observed for velars.

Larynx displacement had a modest positive effect on absolute burst intensity, whereas changes in cavity size did not. Hyperpharyngeal constriction affected absolute burst intensity only for bilabials.

4.7. Discussion of English Ejectives

The occurrence of ejectives in English is a well-documented phenomenon, but it has received limited acoustic and articulatory investigation. This study helps fill that gap by combining acoustic and rt-MRI data to examine the extent to which non-phonological ejectives share properties with their phonological counterparts. In doing so, it sheds light on the basis of their perceived ejectiveness.

The preliminary auditory analysis revealed that ejectives occurred in 9.4% of word-final stops, with accented, voiceless velars forming the most favorable environment for their realization. This finding is consistent with prior research linking ejectives to voicelessness, prominence, and a velar PoA (Gordeeva & Scobbie, 2011; Ogden, 2024).

Acoustically, the English ejectives analyzed here exhibited lower relative burst intensity, which reflected their higher absolute burst intensity and a post-burst gap of silence. These patterns are consistent with the only other dedicated acoustic study on English ejectives (Gordeeva & Scobbie, 2013) and confirm that ejectives in English share key acoustic properties with phonological ejectives cross-linguistically (e.g., Grawunder et al., 2010; Lindau, 1984; Vicenik, 2010). Moreover, these are acoustic correlates of two hallmark ejective features: elevated intraoral pressure (reflected in burst intensity) and delayed glottal release (reflected in low post-burst voicing intensity).

The articulatory evidence presents a more complex landscape. It paints a picture of ejective production in English as a flexible system, influenced by multiple articulatory strategies that contribute in different ways. The variability in articulatory strategies appears to depend on several factors. Individual speakers clearly play a considerable role in shaping these strategies. Yet, some consistent trends tied to PoA nevertheless emerge, suggesting that the variability is not entirely random, but structured in ways that warrant closer attention.

4.7.1. Place of Articulation Effects on Articulatory Patterns

For bilabial ejectives, the absence of larynx raising – indeed, the model predicts larynx lowering – precludes any meaningful contribution of laryngeal displacement to absolute burst intensity. Instead, cavity reduction must be achieved supralaryngeally. The results suggest that this strategy is localized: hyperpharyngeal constriction at this PoA is closely linked to overall cavity reduction and significantly predicts burst intensity. Thus, for bilabial ejectives, this region of the vocal tract appears to play a central role in generating the pressure necessary for a strong burst. That this articulatory pattern would be so pronounced in bilabials makes sense, as the lip-based closure affords greater freedom for tongue retraction into the pharynx, facilitating cavity compression. This sets bilabial ejectives apart from both pulmonic bilabials and ejectives at other PoAs, pointing to a distinct articulatory strategy.

Velar ejectives, by contrast, are most likely to involve larynx raising, of which there are even subtle indications in the auditorily ambiguous velar stop category, and they exhibit the most pronounced cavity size reduction. These patterns underscore once again how velar PoA benefits from articulatory conditions clearly favorable for ejective production. And yet, velar ejectives show no evidence of hyperpharyngeal constriction: it is neither statistically related to overall cavity compression nor does it influence absolute burst intensity. This marks a strategic difference from bilabial ejectives. Rather than involving the pharynx, cavity reduction in velar ejectives must therefore be achieved elsewhere in the vocal tract. One plausible explanation is that cavity reduction is regulated directly in the velar region through lengthening of the tongue dorsum's constriction (Demolin, 2002).

Alveolar ejectives are striking in the absence of a consistent group-level articulatory pattern. This makes it particularly challenging to account for their perceived ejectiveness. Across speakers, laryngeal behavior varies, with some exhibiting larynx raising and others larynx lowering. Supralaryngeal constriction likewise varies in both direction and magnitude. When it occurs, its magnitude does not differ from that observed in pulmonic alveolars. The only articulatory feature that robustly differentiates alveolar ejectives is increased hyperpharyngeal constriction. Unlike in velar ejectives, constriction in this region does contribute to the global cavity size reduction; yet, this effect appears modest and, in contrast to bilabial ejectives, does not influence burst intensity.

For alveolar ejectives, then, articulatory patterns do not appear to be tightly regulated by place-specific constraints. Instead, the mixed and inconclusive patterns observed suggest that speaker-specific strategies outweigh systematic PoA-effects.

The central takeaway from this section is that, although the canonical articulatory components of ejective production – larynx raising and cavity compression – are indeed attested in English ejectives, they are neither consistently present nor systematically coordinated. Even when ejectives are considered separately by PoA, it remains unclear whether the articulatory patterns identified here can fully explain the required pressure increase.

4.7.2. Revisiting Pressure Generation in English Ejectives

As a starting point, the present study confirms that pressure buildup occurs during ejective production in English: ejectives consistently show lower relative burst intensity than their pulmonic counterparts, reflecting higher absolute burst peaks. However, the lack of reliable and uniformly deployed articulatory strategies identified in the preceding section raises the question of what mechanisms actually produce this pressure increase.

The inconsistency of laryngeal involvement means that, although larynx raising may contribute to pressure buildup in some ejectives, most notably velars, it cannot serve as a general explanation. Supralaryngeal cavity size reduction emerges as a more consistent strategy across PoA: unlike larynx movement, changes in cavity size are uniform in direction, generally involving reduction and more robust in frequency. The problem with this strategy is firstly, that it is not consistently distinctive for all ejectives, particularly at the alveolar PoA. Secondly, cavity size reduction shows no statistically reliable association with absolute burst intensity. Thus, it cannot serve as general explanation for pressure increase either. More localized supralaryngeal strategies also provide only partial explanatory power. Hyperpharyngeal constriction helps account for how bilabial ejectives compensate for the absence of larynx raising to compress the cavity. However, even bilabial ejectives this strategy alone is unlikely to suffice as a pressure-generating mechanism, given the elevated burst intensities observed.

For velars, the naturally smaller post-constriction cavity compared to more anterior PoAs enables more efficient intraoral pressure buildup: less cavity volume reduction is needed to achieve a significant increase in intraoral pressure (Javkin, 1977). Regardless of the specific strategy considered, this structural advantage makes it is easier to produce

a perceptually salient burst at this PoA. Nevertheless, even with the advantage of a smaller cavity, it remains unclear how effective the previously proposed strategy of extending the dorsal constriction alone can be in velar ejectives that lack larynx raising.

Alveolar ejectives present the greatest challenge. They exhibit no consistent group-level articulatory pattern that could explain the observed differences in burst intensity between alveolar stop types. Moreover, alveolar ejectives demonstrate the most extreme relative burst intensity values across PoA, corresponding to the highest absolute burst peaks. As suggested above, for alveolar ejectives it appears that speaker-specific strategies, rather than PoA effects, primarily determine the mechanisms of intraoral pressure buildup employed.

Taken together, these findings suggest that articulatory mechanisms beyond the three considered in this study must contribute to pressure generation in English ejectives. In this respect, Brandt and Simpson's (2021) proposal of a pulmonic mechanism offers a promising alternative. According to this account, delayed coordination between glottal and oral closures allows pulmonic air to enter the oral cavity, enhancing pressure prior to release. Crucial to this account are timing and coordination of articulatory gestures, factors not directly measured in the present study but which now emerge as central to understanding all possible sources of pressure buildup in English ejectives.

Resolving the question of where and what generates the required pressure ultimately requires more direct inferences of intraoral pressure. But, in the absence of further direct evidence, it seems likely that, alongside the articulatory strategies discussed above, initiation via a pulmonic airstream may also play a role in English ejectives. Overall, the findings from this study support the view that multiple articulatory strategies coexist in English ejectives (Simpson, 2014; Simpson & Sulaberidze, 2022), potentially even within individual speakers.

4.7.3. The "Unclear" Category: A Window Into Gradience

A final word is warranted on the "unclear" stop category, which, under the broad classification used here, serves as a rudimentary catch-all for ejective-like sounds spanning a range of realizations. As such, they occupy a particularly intriguing space: they begin to exhibit ejective-like properties yet diverge enough from canonical expectations to leave their identity ambiguous. While this study does not seek to

determine their identity, it aimed to investigate the sources of this ambiguity, both acoustically and articulatorily.

Acoustically, the unclear stops share several features with clearly defined ejectives. Like them, they showed longer closure durations than voiced pulmonic stops and post-burst intervals of lower intensity than voiceless pulmonic stops, suggesting possible glottal closure during and after oral release – a likely basis for their perceived ejectiveness. However, their relative burst intensity was significantly higher than that of canonical ejectives and statistically indistinguishable from voiceless pulmonic stops. In this respect, they lack the forceful burst that typically "drives the ejective message home."

Articulatorily, the unclear category lives up to its name, showing no consistent speaker-level patterns. Still, visual trends suggest they may approximate the articulatory configurations of their ejective counterparts, hinting that their perceptual ambiguity and partial acoustic similarity have some articulatory basis.

Finally, if we interpret the overall range of realizations, represented here by the binary of *ejective* vs. *unclear*, as gradual variability symptomatic of sound change in progress, then future research must turn to temporal coordination and more direct analysis of pressure to fully understand the emergence and evolution of ejectiveness in English.

4.8. Conclusion

This chapter presents the first systematic articulatory-acoustic study of English ejectives. While their acoustic profile aligns with typological expectations, the articulatory realization of English ejectives is strikingly variable. In contrast to Amharic, no uniform mechanism underlies their production; instead, a range of individualized strategies is exhibited. These strategies consistently draw from the broader ejective framework, with clear, albeit inconsistent, evidence of both glottalic initiation and supralaryngeal cavity reduction.

Yet this very inconsistency highlights a significant gap in our understanding of how intraoral pressure is generated. The findings suggest that other, partially pulmonic strategies (Brandt & Simpson, 2021) that do not rely solely on larynx raising as the initiatory mechanism may be operating alongside or in place of glottalic initiation. Understanding the nature of these alternative strategies, and their role in pressure-generation, emerges as a crucial line of inquiry for future research.

Finally, this chapter has deliberately refrained from addressing the question of whether these English sounds truly merit the label *ejective*, given their variable production and possible non-glottalic initiation. Before taking up this definitional issue, however, the English and Amharic data is brought together and phonological and non-phonological characteristics are directly compared in the following chapter.

Chapter 5
Comparing Amharic and British English Ejectives

5.1. Introduction

This chapter presents a direct cross-linguistic comparison of ejective stops in Amharic and English. While both sets of sounds have been described as ejectives, they occupy fundamentally different roles in their respective phonological systems: in Amharic, ejectives are fully contrastive and phonologized; in English, they are non-contrastive, emerging as free variants that carry social meaning and take on interactional functions (Simpson, 2014). Parallels between the two have largely been drawn on the basis of auditory impression, leaving open the question of how much they truly share in terms of articulatory mechanisms and acoustic signatures.

The preceding chapters established the defining characteristics of ejective production in each language. Chapter 3 demonstrated that Amharic ejectives conform closely to the canonical glottalic model, characterized by consistent larynx raising, with the added support from supralaryngeal strategies of cavity reduction. Chapter 4, in contrast, revealed the variability of English ejectives, showing a range of articulatory realizations underlying their perceived ejectiveness.

This chapter builds on these findings by directly comparing key acoustic and articulatory measures across the ejectives of the two languages. The goal is to evaluate which features are consistent, which strategies differ, and where overlap of articulatory patterns occurs across the two languages. In doing so, the analysis lays the groundwork for later discussions evaluating whether English ejectives truly merit the label *ejective*. More broadly, the comparison aims to provide a firmer basis upon which to reassess the boundaries and internal consistency of the ejective sound class.

The comparison proceeds in parallel across both acoustic and articulatory domains. It begins with closure duration and relative burst intensity, followed by the articulatory features: larynx raising, supralaryngeal cavity size reduction, and hyperpharyngeal constriction. Together, these analyses provide a foundation for the typological and theoretical reflections that follow in Chapter 6.

5.2. Comparative Analyses

For the purposes of direct comparison, the analyses in this chapter focused exclusively on singleton, word-final Amharic ejectives and alveolar and velar English ejectives. It is important to note, however, that the Amharic dataset is based on 11 speakers, each contributing a relatively balanced number of ejective tokens, whereas the English data comes from 22 speakers, with highly variable numbers of ejective realizations. Due to these asymmetries, the comparative results should be interpreted with caution and are presented in a descriptive rather than statistical manner. As the individual results for each language have been presented in previous chapters, the following sections focus on new insights from the direct cross-linguistic comparison of key measures, namely the acoustic measures of closure duration and relative burst intensity, as well as the three articulatory measures, *delta larynx*, *delta cavity*, and *delta hyperph*.

5.2.1. Closure Duration

Closure duration was included in this direct chapter as it provides a useful point of comparison in terms of the durational scales on which ejective production operates in each language.

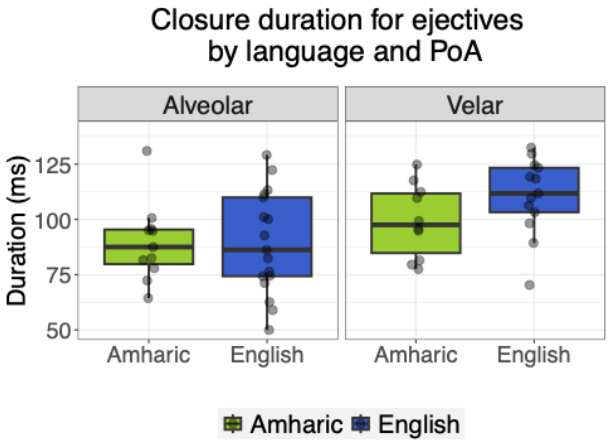

Figure 5.39 Closure duration values (y-axis, in ms) for singleton, word-final ejective stops compared between two languages (Amharic, British English) across two PoAs (alveolar, velar). Dots represent individual speaker means.

Figure 5.39 compares closure duration (in ms) for ejectives in Amharic and English at alveolar and velar PoA.

At the alveolar PoA, median closure durations are equivalent between the two languages at approximately 85 ms. Across both languages, velar ejectives tend to have longer closures than alveolar ones, with English ejectives exhibiting even higher median closure durations (~ 112 ms) than Amharic velar ejectives (~ 95 ms).

The direct comparison shows that English ejectives are markedly more variable in closure duration than their Amharic counterparts, especially at the alveolar PoA. These display a particularly wide range, as reflected in the broader spread of values. At the velar PoA, the overall distribution is more comparable between the two languages.

5.2.2. Relative Burst Intensity

Relative burst intensity was included in the cross-linguistic comparison to capture potential differences in ejective intensity characteristics, operationalized as the relationship between the burst peak and the peak intensity of the following vowel. **Figure 5.40** shows relative burst intensity (in dB) for singleton, word-final ejectives in Amharic and English at alveolar and velar PoA.

Across PoA, Amharic ejectives exhibit higher median relative burst intensity values than English ejectives, with the contrast particularly pronounced at the alveolar PoA. English ejectives additionally show a clearer PoA effect than Amharic, with alveolars displaying lower relative burst intensity than velars. To clarify the source of these differences, the burst peak and the following vowel peak were examined separately. Both components (burst peak and vowel peak) differed between the languages, indicating that cross-linguistic differences in relative burst intensity cannot be attributed to burst strength alone.

In Amharic, relative burst intensity values are positive at both PoAs, indicating that the peak intensity of the ejective-following vowel always exceeds that of the burst. In English, this relationship is attenuated for velars (1.4 dB) and reversed for alveolars, where burst peaks exceed vowel peaks. Thus, the measure captures differences both in the magnitude and, for alveolars, also the direction of the burst-vowel intensity relationship across ejectives of the two languages.

Having compared acoustic characteristics via closure duration, which provides a temporal scale, and via relative burst intensity, which provides an intensity-based relational scale across the ejectives of the two languages, we now turn to the articulatory comparison of larynx raising, supralaryngeal cavity reduction, and hyperpharyngeal constriction.

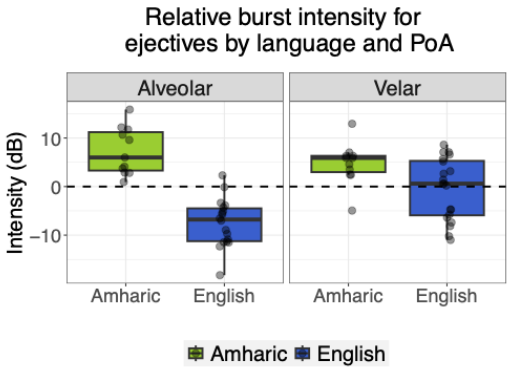

Figure 5.40 Relative burst intensity values (y-axis, in dB) for singleton, word-final ejective stops compared between two languages (Amharic, British English) across two PoAs (alveolar, velar). Dots represent individual speaker means. Values reflect the peak intensity difference between the following vowel and the stop's release burst. Lower values (and especially negative values) indicate a stronger intensity of the burst relative to the vowel.

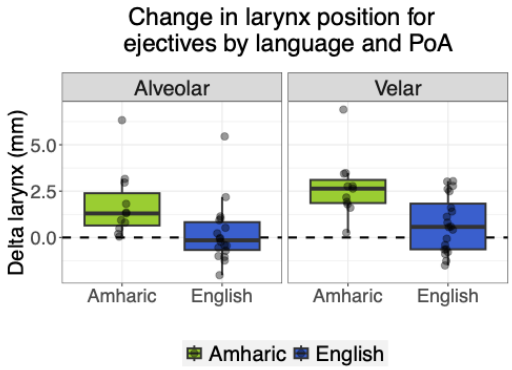

Figure 5.41 Delta larynx values (y-axis, in mm) for singleton, word-final ejective stops compared between two languages (Amharic, British English) across two PoAs (alveolar, velar). Positive numbers indicate larynx raising, negative numbers indicate larynx lowering. Dots represent the mean values for individual speakers.

5.2.3. Larynx Raising

Figure 5.41 compares average larynx displacement (*delta larynx* in mm) during the closure of singleton, word-final ejectives in Amharic and English at alveolar and velar PoAs. Positive values indicate larynx raising, and negative values larynx lowering.

This direct comparison reveals several notable patterns. Although larynx movement is quite variable across speakers, even in Amharic, its direction is not only clearly more systematic but its average magnitude is greater, especially at the velar PoA (Amharic: alveolar ~ 1.25 mm, velar ~ 2.5 mm vs. English: alveolar ~ 0.5 mm, velar ~ 1.25 mm). Two outlier speakers, in the Amharic data, already touched upon in Chapter 3, stand out as reminders of individual variation even within a generally consistent system.

Although larynx raising in English is inconsistent and typically smaller than in Amharic, direct comparison shows that some English speakers do employ raising movements comparable to those of Amharic speakers, particularly in velar ejectives.

5.2.4. Supralaryngeal Cavity Reduction

Figure 5.42 compares changes in supralaryngeal cavity size (*delta cavity* in mm²) during ejective production in Amharic and English, across two PoAs (alveolar and velar). The direct comparison illustrated in this figure is particularly revealing. It highlights that Amharic ejectives involve a highly consistent, yet surprisingly moderate, degree of cavity reduction across PoA, with very narrow inter-speaker variability – indications of a stable and shared articulatory strategy among Amharic speakers.

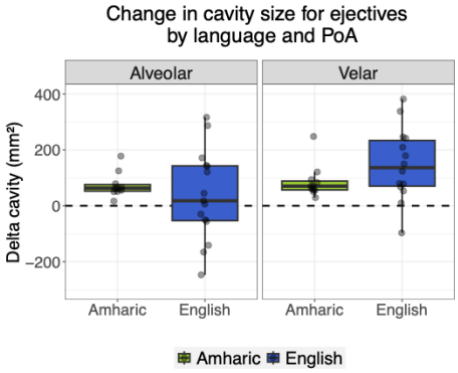

Figure 5.42 Delta cavity values (y-axis, in mm²) for singleton, word-final ejective stops compared between two languages (Amharic, British English) across two PoAs (alveolar, velar). Delta cavity is measured as the difference between the cavity's size at the onset of stop closure and its size at closure offset. Positive numbers indicate cavity reduction, negative numbers indicate cavity expansion. Dots represent the mean values for individual speakers.

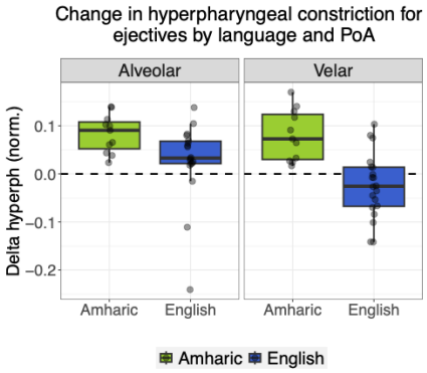

Figure 5.43 Delta hyperph values (y-axis) for singleton, word-final ejective stops compared between two languages (Amharic, British English) across two PoAs (alveolar, velar). Delta hyperph is measured as the difference between the degree of hyperpharyngeal constriction at the onset of stop closure and its constriction at closure offset. The black dashed line represents the zero value for the y-axis: observations above it indicate hyperpharyngeal constriction increased, below it indicates hyperpharyngeal constriction relaxed. Dots represent the mean values for individual speakers.

In stark contrast, English ejectives show high variability, reflected in the broader spread of speaker means and taller boxplots. It is notable, that when English speakers do reduce the cavity, they often show a greater degree of reduction than Amharic speakers, particularly at the velar PoA. However, this strategy is far from reliable, especially at the alveolar PoA, where it is not even contrastive (see Chapter 4).

5.2.5. Hyperpharyngeal Constriction

Figure 5.43 illustrates the change in hyperpharyngeal constriction (*delta hyperph*) during ejective production for Amharic and English, across alveolar and velar PoA. Positive values indicate increased constriction, while negative values indicate expansion.

The direct comparison reveals an interesting pattern: while both languages generally show positive delta values, indicating some degree of increased constriction, English *velars* stand out as an exception, often showing no change or even expansion in the lower pharyngeal region, indicating that this region does not contribute to overall supralaryngeal cavity constriction.

5.2.6. Summary of Cross-Linguistic Comparison

The direct comparisons showed that closure duration is generally longer in English, especially for velars, but on average relative burst intensity across PoA is higher and more stable (both in magnitude and direction) in Amharic. Larynx raising is greater and more consistent in Amharic, though some English speakers show comparable movements for velars. Cavity reduction is moderate and uniform in Amharic, but more extreme and less reliable in English. In English, hyperpharyngeal constriction is reduced and often notably absent in velar ejectives. Overall, the comparison highlights some key differences between structured phonological and variable, non-phonological patterns of ejective production.

5.3. Contrastive and Emerging Patterns in Ejective Production

This chapter brought together acoustic and articulatory evidence to place the ejectives of two languages side by side. The goal was to directly explore similarities in ejective production across systems that differ fundamentally in the status and function of ejectives, while also helping to solidify patterns within each system.

In Amharic, where ejectives are phonological, the direct comparisons brought to light the striking modesty of supralaryngeal cavity constriction. This comparison sharpened our understanding of its role as a consistent, yet moderate and supportive, gesture that complements larynx raising, which remains the primary initiator of the glottalic airstream mechanism (see Chapter 3).

In English, the direct comparisons highlighted how variability extends to the magnitude of the articulatory strategies employed. When larynx raising does occur in English ejectives, it tends to be reduced, seemingly "unpracticed" – it cannot match the highly efficient and routinized use of larynx raising observed in Amharic. This is demonstrated by the fact that Amharic speakers achieve greater larynx elevation in less time than English speakers, as reflected in both larger average larynx displacements and shorter closure durations. Supralaryngeal cavity constriction, by contrast, when employed, can on average be much more extreme in English than in Amharic. Rather than taking on a supportive role, as in Amharic, this may possibly represent a more active attempt to approximate the intraoral pressure buildup crucial for ejective production.

Most English speakers also seem unattuned to the potential benefits of coordinating gestures. The Amharic-specific use of hyperpharyngeal constriction, particularly in velar ejectives, demonstrates an efficient and coordinated use of articulatory resources that is not mirrored in English. In English, while hyperpharyngeal constriction is somewhat more consistent at the alveolar PoA, it is likely not coordinated with larynx raising, which remains unreliable in English alveolar ejectives.

Finally, differences in relative burst intensity suggest that the relationship between burst peak intensity and following vowel peak intensity is weighted differently in the two languages. In Amharic, ejective bursts do not exceed the intensity of the following vowel. This suggests that a high burst-to-vowel intensity ratio is not required to cue ejectivity. Such a pattern is consistent with a phonological system in which ejectives are fully established as a contrastive category: burst intensity functions as one cue among several, and ejectivity does not depend solely on the prominence of the burst peak relative to the following vowel.

In English, by contrast, ejective productions show a stronger tendency for burst peak intensity to exceed that of the following vowel. This pattern may reflect their emergent status. In the absence of a stable phonological category, speakers appear to rely more heavily on highly salient acoustic properties, with burst intensity serving as a primary cue. The exaggerated burst-vowel intensity difference thus enhances the perceptual distinctiveness of the ejective release. This tendency is particularly pronounced for alveolar ejectives. Velar ejectives in English more closely resemble the Amharic pattern, where vowel peak intensity just barely exceeds burst peak intensity. In addition to their greater frequency and higher likelihood of glottalic initiation, this similarity provides

further evidence that velar ejectives in English may be more advanced in the process of phonologization.

The direct comparison sheds light on the robust and highly optimized patterns in Amharic that are reflections of a practiced generational transmission and the learned routines of a phonological category. While individual speakers within this phonological system naturally may vary, a clear categorical contrast is consistently maintained (see Chapter 3).

For English, the direct comparison underscores the emergent and non-phonologized nature of these sounds. They lack the systematic, learned articulatory routines observed in Amharic, with different articulatory features employed variably in more reduced or exaggerated ways, as characteristic of the emergent nature of such sounds. That said, within the highly variable system observed, Chapter 4 showed that, depending on individual speakers and influenced by PoA, strategies may nevertheless approximate surface characteristics of the phonological production mechanism. These comparative insights provide a foundation for a more nuanced understanding of ejectives as a sound class, and for assessing the place of English ejectives within it, feeding directly into the concluding discussion in the next chapter.

Chapter 6
General Discussion

This dissertation set out to explore the articulatory realization of ejectives in both phonological and non-phonological language settings with the goal of providing an empirically grounded account of ejective production mechanisms and thus contributing to a more nuanced understanding of the defining characteristics of this sound class. Building on prior work (notably Brandt & Simpson, 2021; Kingston, 1984), the presented studies used rt-MRI as a central methodological tool to investigate the articulatory details of phonologically contrastive ejectives in Amharic and emerging ejectives in British English, complemented by acoustic analysis.

6.1 Summary of the Main Findings

The findings from the acoustic analysis in the Amharic study (Chapter 3) revealed robust patterns in the acoustic features of Amharic ejectives across speakers and contextual factors. Post-release silence and the relative structure of burst intensity were the clearest acoustic markers of these ejectives. The articulatory data provides strong empirical support for classical accounts of glottalic initiation of ejectives (Catford, 1977), confirming the central role of larynx raising in ejective production. At the same time, the data suggests an extension of this model by revealing that supralaryngeal articulation is a systematic and integrated component of the production mechanism. Rather than acting independently, laryngeal and supralaryngeal gestures seem part of a system that works together to optimize intraoral pressure buildup, a key feature of ejective production.

The auditory analysis of English ejectives (Chapter 4) revealed that they occurred in 9.4% of word-final stops, most frequently in accented, voiceless velars. Acoustically, these tokens exhibited hallmark features including post-burst silence and distinctive relative burst intensity patterns, confirming that they share key properties with phonological ejectives despite their emerging status. In contrast to the robust and consistent patterns observed in Amharic, the articulatory analysis revealed considerable variability across speakers and tokens. Larynx raising was not a consistent feature of production. Instead, supralaryngeal cavity compression emerged as a more robust strategy for generating pressure during stop closure. These patterns point to a heterogeneous class

of productions, suggesting that multiple articulatory mechanisms, ranging from glottalic to non-glottalic, are currently in use.

The cross-linguistic comparison (Chapter 5) further highlighted clear differences in the magnitude and consistency of articulatory gestures, reflecting the distinct status of these sounds in each language. Amharic ejectives were marked by highly stable and integrated articulatory routines, consistent with their phonologized status. In contrast, the flexible and irregular articulatory patterns of English ejectives are tokens of an emergent phenomenon still undergoing stabilization.

By integrating laryngeal and supralaryngeal dimensions and comparing phonological and non-phonological systems of ejective production, the studies included in this dissertation lay the empirical groundwork for the broader discussion that follows in this chapter. This discussion considers how the findings contribute to our understanding of the relationship between the practical realization of ejective stops, the defining characteristics of the ejective sound class, and their relevance to questions of how such sounds are categorized.

6.2. Revisiting the Glottalic Model

A central concern is whether established theoretical accounts of ejective production accurately reflect how these sounds are realized in practice. The findings from Amharic provide robust empirical support for the classical "glottalic egressive" model of ejective production (Catford, 1977). At the same time, by moving beyond an isolated view of the larynx, the findings confirm the presence of a consistent supralaryngeal component to ejective production, one that supports the glottalic mechanism by increasing cavity rigidity and optimizing conditions for pressure buildup.

One of the guiding approaches of this dissertation was the idea that, to better understand the boundaries of this sound class – what "truly" qualifies an ejective and what not – we must examine cases that challenge our assumptions: those that sit at the edges of the category, where the label "ejective" becomes less certain. The analysis of English ejectives showed that larynx raising was not a consistent feature of their production. Nonetheless, precisely because these sounds exhibit certain articulatory, and particularly acoustic, characteristics that give them the *flavor* of ejectiveness, examining a broader spectrum of ejective-related productions, such as those presented here in English, provides valuable insight that can inform categorical deliberations.

6.2.1. A Gradient Model of Ejectiveness

Let us consider the combined range of sounds from Amharic and English that are perceived as ejectives by the listener as a gradient, that is, a continuum spanning from weakly glottalized sounds to canonical ejectives. The continuum begins with sounds that carry the resonance characteristics of a partially or completely sealed glottis, yet do not necessarily involve any active pressure buildup. In this way, even the weakest examples are included. At the other end of the continuum are canonical ejectives – glottalic sounds, where buildup of pressure is achieved through systematic larynx raising, aided by supralaryngeal cavity compression. In between lies a spectrum of articulatory strategies that may converge on a similar auditory percept of ejectiveness. These arise from different combinations of glottal and supralaryngeal components, as well as partially or fully pulmonic airstreams that contribute to pressure building during production.

When considering this continuum of sounds, a first meaningful shift along the gradient of ejectiveness occurs at the point where the articulatory goal transitions to involving active pressure generation, specifically, where simultaneous oral and glottal closure is combined with a deliberate, though not necessarily specified, articulatory effort to compress the oral cavity. The sounds at this point in the gradient may be well serviced by the broader description offered in the *IPA Handbook* (1999), or they may be the result of a pulmonic mechanism such as that proposed by Brandt and Simpson (2021). In any case, it is at this point that sounds move beyond being merely *glottalized*. English ejectives span much of this central range of the gradient. They carry characteristics of sounds produced with a glottal closure at release and exhibit diverse and often uncoordinated articulatory strategies aimed at generating elevated burst intensity. Yet, they do so without reliably employing glottalic initiation.

The second meaningful shift along the gradient of ejectiveness occurs where systematic larynx raising becomes the defining initiatory gesture, marking the transition to sounds that align with classical definitions of ejectives in the IPA (e.g., Catford, 1977). The direct comparison between the Amharic and English data presented in this dissertation thus supports the notion that larynx raising – and, by extension, their glottalic nature – remains the most principled basis for defining ejectives as a sound class.

While surface-level auditory similarities may blur the categorical lines, the underlying articulatory mechanisms are what define category membership. The English sounds, though ejective-like in perception and clearly drawing on the broader articulatory and acoustic framework associated with "true" ejectives, do not consistently meet the criteria for category membership. As such, these sounds should not, at this stage, be classified as "true" ejectives. Having determined this, the term "ejective" is retained for the remainder of this chapter for practical reasons despite the classification caveat.

6.2.2. Non-Canonical Phonological Ejectives?

One question that remains open in this dissertation is how to account for sounds that fall in the middle of the gradient, those that involve some degree of pressure buildup but are not glottalically initiated, and how such sounds should be classified. The presence of these intermediate cases in English was interpreted here as symptomatic of a system in flux, consistent with a sound change in progress.

However, the fact that Amharic does not exhibit this type of non-canonical production does not allow us to conclude whether such variability applies to all languages with phonological ejectives. Detailed articulatory data is still lacking for many of them. It is entirely possible that similar patterns of alternative articulatory strategies exist elsewhere, perhaps in a more systematized form. Indeed, Brandt and Simpson (2021) report such cases in Georgian, where some phonological ejectives do not consistently involve larynx raising. Their findings suggest that non-canonical realizations may occur even within phonologically contrastive systems. To evaluate this possibility, more articulatory studies from a broader and more diverse range of languages are needed.

On the other hand, Georgian may be an exception. As discussed in the introduction, most studies, even when based on indirect evidence, treat larynx raising as a central feature of ejective production. It remains possible that what Brandt and Simpson (2021) observed in Georgian reflects a local innovation or reduction, rather than a broader cross-linguistic trend. Thus, a key question remains: do the non-canonical productions and alternative strategies found in Georgian ejectives reflect true variation within the phonological ejective category, or are they better understood as signs of ongoing sound change and innovation?

6.3. English Ejectives as a Sound Change in Progress

Although the high variability, irregular patterns, and lack of coordinated articulatory routines in English ejectives have been interpreted as signs of a sound change in progress, the basic assumption, that ejectives are increasing in frequency, still requires empirical verification (Simpson, 2014).

Furthermore, when considering the variability of production mechanisms in English ejectives, it is important to acknowledge the potential role of the different functions, including interactional, emphatic, and stylistic, that these sounds are currently taking on in this language. This dissertation did not directly examine how function relates to articulatory strategy, and therefore cannot confirm or refute Simpson's (2014) proposal that functional pressures shape this variability. What the data does show, however, is that even under highly controlled experimental conditions, speakers exhibited considerable divergence in their articulatory patterns. This suggests that while functional pressures may influence ejective production, they do not fully account for it. Instead, the diversity of functions appears to be intertwined with the broader dynamics of sound change.

Brandt and Simpson's (2021) proposal of a pulmonic mechanism for ejectives, based on timing mismatches rather than active larynx raising, provides a system-conforming entry point for such sounds to emerge in a language that otherwise lacks glottalic features. While the present data cannot directly confirm this account, the finding that most English ejective realizations are not glottalically initiated, as anticipated by Simpson (2014), lends it tentative support. At the same time, the data suggests that velars may represent a point of articulatory progression, moving beyond the initial, non-glottalic mechanism. Velar ejectives in English show the highest rates of larynx raising and strongest articulatory engagement, mirroring Amharic patterns and aligning with typological expectations (e.g., Cho & Ladefoged, 1999; Javkin, 1977). Given the biomechanical and aerodynamic advantages of the velar place, such as smaller cavity volume behind the constriction and efficient tongue-larynx coordination, they are well-positioned to support glottalic initiation. In this sense, velars may serve as articulatory footholds for the emergence and stabilization of more canonical ejective production in English.

The further development of English ejectives remains a compelling area of inquiry. Their apparent spread offers a rare opportunity to observe how glottalic sounds may either take root or remain peripheral within a pulmonic language system. Their current status, tied to interactional, emphatic, and stylistic functions, may allow them to stabilize as

independent phonetic variables within those domains. As speakers become more attuned to these functions, articulatory routines may become more predictable, usage more regular, and even anterior places of articulation more likely to involve glottalic initiation. In this scenario, ejectives in English might develop into "true" ejectives in a technical articulatory sense, albeit within a limited functional scope.

Nonetheless, there is little reason to expect that the long-term trajectory of this sound change will diverge from previously observed patterns (Easterday & Bybee, 2023). The positional restriction of these ejective sounds in English, their non-fusion-based origin (Fallon, 2013), and their misalignment with the broader phonological system are precisely the features that, according to Easterday and Bybee's (2023) typological and diachronic analyses and Kiparsky's (2003) theoretical model, suggest a low likelihood of full phonologization. Then again, this may simply mark the beginning of one of the unusually extended trajectories that sound changes involving ejectives are thought to follow (Easterday & Bybee, 2023).

6.4. Limitations and Future Directions

Expanding articulatory research to a broader set of languages with phonological ejectives will be essential for establishing how representative the Amharic patterns observed in this dissertation are. While this dissertation helps to close the gap between acoustic and articulatory approaches to ejectives, much of what is known about ejective production and cross-linguistic variability remains based on acoustic data. The comparison of different languages with phonological ejectives could uncover possible cross-linguistic variation in production strategies. The detailed rt-MRI analysis of one ejective language presented here offers an important starting point, but considerable catch-up work is still needed.

Similarly, integrating intraoral pressure measurements into future research remains a critical step. It is important to remember that the interpretations in this study relied on indirect inferences from burst intensity, which is a significant limitation. A comprehensive account of ejective production will ultimately require intraoral pressure data, with the primary technical challenge arising if such measurements are to be integrated with rt-MRI. Only with direct aerodynamic evidence can we fully model the mechanisms behind ejective production and better understand how these sounds function typologically and phonetically. Future work will need to explore how best to integrate

aerodynamic tools to model pressure dynamics and gesture coordination more precisely. In the meantime, as in the present study, researchers will need to make informed trade-offs about which dimensions of ejective production to prioritize.

Future work must also explore the relationship between acoustic and articulatory variation in a more targeted way. This dissertation prioritized articulatory analysis, and limited its acoustic measures to temporal and intensity-based features. This was done with the articulatory parameters under investigation in mind. Measures capturing effects on the following vowel, where cross-linguistic variability in ejective realization is often observed, were not included. As a result of this prioritization, the study's ability to fully capture key aspects of ejective variability is limited and must be acknowledged as such. This presents a clear opportunity for future research to examine how features like f0 and voice quality map onto articulatory dimensions. Wright et al.'s (2002) continuum, spanning larynx height, vocal fold tension, and medial compression, offers a useful framework for capturing the muscular, aerodynamic, and temporal dynamics that likely shape such acoustic outcomes.

The absence of bilabial ejectives in the Amharic dataset is unfortunate. Given their distinct articulatory environment, compared to alveolar and velar place, which includes larger cavity volume, greater tongue freedom, and a larger area of compliant articulatory structures such as the cheeks and lips, bilabial ejectives could offer valuable insight into articulatory diversity across places of articulation. Future cross-linguistic work that includes this place may help expand our understanding of articulatory variation in ejectives.

This dissertation has shown that emerging ejectives in English are generally not produced with larynx raising. While this is an important first finding, the mechanisms behind these non-glottalic cases remain unclear. Further research is needed to evaluate proposals such as Brandt and Simpson's (2021) pulmonic mechanism or to identify alternative strategies that bypass glottalic initiation. Although supralaryngeal compression appears central, its role in pressure generation is not yet well understood. Addressing this will require moving beyond rt-MRI to incorporate aerodynamic data and dynamic articulatory imaging, allowing for a more detailed account of how pressure is built and how strategies vary across speakers and tokens.

A related priority is to refine our understanding of gesture timing. The current dataset lends itself well to analyzing the dynamic trajectory of articulatory gestures leading up to the release burst – an avenue reserved for future work. Temporally precise modeling of laryngeal and supralaryngeal coordination will offer further valuable insight into how ejective mechanisms are implemented, shedding light not only on production variability but also on the articulatory development of the ongoing sound change in English.

Regarding English, future work could greatly benefit from a more nuanced auditory coding system, aimed at capturing the gradient nature of these events. This could enable clearer differentiation of functional uses, which is important for further considerations of how these sounds are evolving and stabilizing within the English language.

6.5. Conclusion

Through the comparison of phonological ejectives in Amharic and emerging ejectives in British English, this dissertation has shown that while a range of articulatory strategies can give rise to ejective-like sounds, typological classification remains rooted in core defining features, in this case, glottalic initiation. In this light, defining sound classes in terms of formalized articulatory properties is not reductive but necessary: it enables coherence in cross-linguistic classification, even in the face of phonetic variability. The findings support the continued value of traditional categorical definitions and highlight how examining edge cases can clarify, rather than undermine, the structure of the classificatory systems we use to describe and represent sound classes such as ejectives.

.

Deutsche Zusammenfassung

Diese Dissertation untersucht die Produktion von Ejektiven sowohl in phonologischen als auch in nicht-phonologischen Kontexten, anhand von zwei Sprachen: dem Amharischen, einer ethiosemitischen Sprache mit phonologisch kontrastiven Ejektiven, und dem Britisch Englischen, in dem Ejektiv-artige Laute zwar nicht zum regulären phonologischen Lautsystem der Sprache gehören, aber inzwischen im Alltagsgebrauch der Sprache ganz verankert sind. Ziel ist es, ein empirisch fundiertes Verständnis darüber zu gewinnen, wie Ejektivlaute artikulatorisch realisiert werden und welche Eigenschaften sie als Lautklasse definieren. Im Mittelpunkt steht dabei die Frage: Was macht einen Laut zu einem Ejektiv?

Ejektive werden klassischerweise in der phonetischen Literatur durch ein glottalisches Initiationsverfahren beschrieben: Bei geschlossener Glottis wird der Kehlkopf angehoben, wodurch die Luft im supralaryngealen Raum komprimiert wird. Beim Lösen des oralen Verschlusses entweicht diese komprimierte Luft, was zu einem typischerweise intensiven Verschlusslösungsgeräusch führt (Catford, 1977). Diese Beschreibung ist in vielen Lehrwerken zentral, wurde jedoch nur in wenigen empirischen Studien direkt überprüft und kaum im sprachvergleichenden Rahmen (Brandt & Simpson, 2021; Kingston, 1984).

Die vorliegende Arbeit greift diesen Mangel an empirischen Daten auf und verfolgt das Ziel, die artikulatorischen Mechanismen der Ejektivproduktion systematisch und sprachvergleichend zu untersuchen. Mithilfe von Echtzeit-Magnetresonanztomografie (Echtzeit-MRT) werden dabei sowohl vertikale Kehlkopfbewegungen als auch supralaryngeale artikulatorische Gesten analysiert. Ergänzend erfolgt eine detaillierte akustische Analyse, um festzustellen, wie sich Ejektive von pulmonalen Plosiven unterscheiden und wie sie sich in das breitere Sprachumfeld der Ejektivlaute einordnen.

Ein zentrales Anliegen dieser Dissertation ist es, die gängige Annahme zu hinterfragen, dass Kehlkopfanhebung zwingend erforderlich für Ejektive ist. Frühere Studien legen nahe, dass ein erhöhter intraoraler Druck nicht ausschließlich durch Kehlkopfanhebung erzeugt wird, sondern auch durch zusätzliche artikulatorische Strategien wie Zungenwurzelretraktion oder pharyngale Verengung entstehen kann (Kingston, 1984).

Diese Arbeit untersucht zunächst die phonologischen Ejektive im Amharischen (Kapitel 3) und anschließend die nicht-phonologischen, entstehenden Ejektive im Britischen Englisch (Kapitel 4). Im Anschluss werden die Ergebnisse in Kapitel 5 direkt miteinander verglichen, um herauszuarbeiten, welche artikulatorischen und akustischen Eigenschaften konsistent über Sprachsysteme hinweg auftreten, und welche eher durch phonologische oder kontextuelle Bedingungen geprägt sind.

Die Dissertation verfolgt somit nicht nur das Ziel, die konkrete Lautproduktion von Ejektiven differenziert zu beschreiben, sondern auch, bestehende theoretische Modelle zu überprüfen und weiterzuentwickeln. Sie möchte einen Beitrag dazu leisten, die Typologie und Definition von Ejektiven zu schärfen und auf eine empirisch fundierte Grundlage zu stellen.

Der erste Analyseschwerpunkt liegt auf dem Amharischen, einer Sprache, in der Ejektive phonologisch fest etabliert sind und daher als stabiler Referenzpunkt für die Untersuchung kanonischer Produktionsmuster dienen. Die akustischen Ergebnisse zeigen, dass amharische Ejektive zentrale Merkmale dieser Lautklasse aufweisen: Sie zeichnen sich durch charakteristische relative Intensitätsmuster im Verschlusslösungsgeräusch sowie durch eine unmittelbar anschließende Phase akustischer Stille aus. Die Dauer des oralen Verschlusses unterscheidet sich hingegen kaum von derjenigen pulmonischer Plosive.

Auf artikulatorischer Ebene zeigt sich ein konsistentes Muster: eine koordinierte Kehlkopfanhebung in Verbindung mit einer Verengung supralaryngealer Räume, unter anderem im Bereich des Hyperpharynx. Diese Beobachtungen bestätigen klassische Beschreibungen des glottalischen Mechanismus, erweitern sie jedoch um die Erkenntnis, dass laryngeale und supralaryngeale Gesten als integriertes System zusammenwirken, um den intraoralen Druck aufzubauen.

Der Artikulationsort beeinflusst sowohl akustische als auch artikulatorische Eigenschaften: Velare Ejektive sind besonders ausgeprägt und zeigen eine stärkere artikulatorische Beteiligung als alveolare. Dies bestätigt, dass weiter hinten im Vokaltrakt gebildete Konsonanten biomechanische Vorteile für den Druckaufbau bieten. Unterschiede in Wortposition und Konsonantenlänge fallen dagegen nur gering aus. Insgesamt ergibt sich ein detailliertes Referenzbild der artikulatorischen Umsetzung phonologisch kontrastiver Ejektive im Amharischen.

Im Kontrast dazu untersucht der zweite Teil die im britischen Englisch auftretenden, nicht-kontrastiven Ejektivlaute. Diese werden nicht durch das phonologische System der Sprache definiert, sondern auditiv als „ejektivisch“ erkannt. Da sie in einem ansonsten vollständig pulmonischen System auftreten, stellen sie eine besondere Herausforderung für bestehende Beschreibungsmodelle dar.

Die Ergebnisse zeigen ein hohes Maß an inter- und intraindividueller Variabilität. Zwar lässt sich in manchen Fällen eine Kehlkopfanhebung beobachten, vor allem bei velaren Lauten, diese tritt insgesamt jedoch selten und inkonsistent auf. Viele dieser Ejektive entsprechen nicht den klassischen glottalen Beschreibungen. Eine deutlich stabilere Artikulationsstrategie ist hingegen die Komprimierung des supralaryngealen Mundraums, dies gilt selbst für bilabiale und alveolare Laute. Die Ergebnisse legen nahe, dass Ejektive im Englischen nicht durch einen einheitlichen Mechanismus, sondern durch eine Vielzahl artikulatorischer Strategien erzeugt wird, die gemeinsam zu erhöhtem intraoralem Druck führen können.

Der abschließende Vergleich zeigt, dass amharische Ejektive konsistent und systematisch produziert werden: mit regelmäßiger Kehlkopfanhebung und stabiler supralaryngealer Verengung des Mundraums. Im Gegensatz dazu sind englische Ejektive deutlich variabler, sowohl in der Ausführung einzelner Gesten als auch in ihrer Kombination. Während das Amharische ein phonologisiertes, routiniertes Muster zeigt, spiegeln die englischen Daten ein flexibles Produktionssystem wider, Eigenschaften für eine sich im Wandel befindliche Lautklasse.

Diese Gegenüberstellung verdeutlicht die Bandbreite artikulatorischer Konfigurationen, die zu einem ejektivischen Höreindruck führen können, und bildet damit die Grundlage für die theoretische Diskussion zur Definition und Typologie von Ejektiven im abschließenden Kapitel. Dabei wird herausgearbeitet, dass die traditionelle Annahme, Kehlkopfanhebung sei konstitutiv für Ejektive, bestätigt werden kann – und dass diese laryngeale Bewegung sowie die damit verbundene glottale Charakteristik zentrale Merkmale der Lauttypologie von Ejektiven darstellen.

Insgesamt leistet die Arbeit einen Beitrag zur Präzisierung der Ejektivkategorie, indem sie empirisch zeigt, wie sich die Grenze zwischen kanonischen und nicht-kanonischen Realisierungen artikulatorisch manifestiert und unterstreicht somit die Relevanz detaillierter artikulatorischer Analysen für die Lautklassifikation.

References

Ahmad, M., Dargaud, J., Morin, A., & Cotton, F. (2009). Dynamic MRI of larynx and vocal fold vibrations in normal phonation. *Journal of Voice*, *23*(2), 235-239.

Amare, G. (1983). The dialect of Wəllo with a special reference to the Ambassel Awaja. *Unpublished Senior Essay in Linguistics*.

Amsalu, A., & Habtemariam, M. (1973). Regional variation in Amharic: The dialect of Wəllo. *Journal of Ethiopian Studies*, *11*(2), 124-129.

Andersen, K. F., & Sonninen, A. (1960). The function of the extrinsic laryngeal muscles at different pitch. *Acta Oto-Laryngologica*, *51*(1-2), 89-93.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1-48.

Baye, Y. (2016). Phonological features of the Amharic variety of South Wəllo. In B. S. Mendisu & Johannessen, J. B. (Eds.), *Multilingual Ethiopia: Linguistic Challenges and Capacity Building Efforts* (Vol. 8, pp. 9-30). Oslo Studies in Language.

Boersma, P., & Weenink, D. (2025). *Praat: doing phonetics by computer [Computer program]. Version 6.4.35, retrieved 15 June 2025 from http://www.praat.org/.*

Brandt, E., & Simpson, A. P. (2021). The production of ejectives in German and Georgian. *Journal of Phonetics*, *89*, 101111.

Bückins, A., Greisbach, R., & Hermes, A. (2018). Larynx movement in the production of Georgian ejective sounds.

Byrd, D., Tobin, S., Bresch, E., & Narayanan, S. (2009). Timing effects of syllable structure and stress on nasals: A real-time MRI examination. *Journal of Phonetics*, *37*(1), 97-110.

Carignan, C., Coretta, S., Frahm, J., Harrington, J., Hoole, P., Joseph, A., Kunay, E., & Voit, D. (2021). Planting the seed for sound change: Evidence from real-time MRI of velum kinematics in German. *Language*, *97*(2), 333-364.

Carignan, C., Hoole, P., Kunay, E., Joseph, A., Voit, D., Frahm, J., & Harrington, J. (2019). The phonetic basis of phonological vowel nasality: Evidence from real-time MRI velum movement in German. 19th International Congress of Phonetic Sciences, Melbourne, Australia.

Carignan, C., Hoole, P., Kunay, E., Pouplier, M., Joseph, A., Voit, D., Frahm, J., & Harrington, J. (2020). Analyzing speech in both time and space: Generalized additive mixed models can uncover systematic patterns of variation in vocal tract shape in real-time MRI. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *11*.

Carignan, C., Shosted, R. K., Fu, M., Liang, Z.-P., & Sutton, B. P. (2015). A real-time MRI investigation of the role of lingual and pharyngeal articulation in the production of the nasal vowel system of French. *Journal of Phonetics*, *50*, 34-51.

Catford, J. C. (1977). *Fundamental problems in phonetics*. Bloomington, Indiana: Indiana University Press.

Chirrey, D. (1999). Edinburgh: Descriptive material. In P. Foulkes & Docherty, G. (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 223-229).

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, *27*(2), 207-229.

Cowley, R., Bender, M. L., & Ferguson, C. A. (1976). The Amharic language-description. In M. L. Bender, Bowen, J. D., Cooper, R., Ferguson, C. (Eds.), *Language in Ethiopia*. London: Oxford University Press.

Crystal, T. H., & House, A. S. (1988). Segmental durations in connected-speech signals: Current results. *The Journal of the Acoustical Society of America*, *83*(4), 1553-1573.

Cunha, C., Hoole, P., Voit, D., Frahm, J., & Harrington, J. (2024). The physiological basis of the phonologization of vowel nasalization: A real-time MRI analysis of American and Southern British English. *Journal of Phonetics*, *105*, Article 101329.

Dehais-Underdown, A., Vignes, P., Crevier-Buchman, L., Demolin, D., Vuissoz, P.-A., Isaieva, K., Fauvel, M., Laprie, Y., & Felblinger, J. (2023). Non-pulmonic initiation in human beatboxing: A real-time MRI study. 20th International Congress of Phonetic Sciences Prague, Czech Republic.

Demolin, D. (2002). The search for primitives in phonology and the explanation of sound patterns: The contribution of fieldwork studies. *Laboratory Phonology*, *7*, 455-514.

Demolin, D. (2004). Acoustic and aerodynamic characteristics of ejectives in Amharic. *The Journal of the Acoustical Society of America*, *115*(5_Supplement), 2610-2610.

Demolin, D. (2011). The influence of aerodynamic constraints on the shape and the dynamics of phonological systems. 17th International Congress of Phonetic Speech, Hong Kong.

Demolin, D., Ghio, A., & Mous, M. (2021). A phonetic study of Iraqw ejectives consonants. 10th World Congress of African Linguistics, Leiden, Netherlands.

Demolin, D., Hassid, S., Metens, T., & Soquet, A. (2002). Real-time MRI and articulatory coordination in speech. *Comptes Rendus Biologies*, *325*(4), 547-556.

Derib Ado. (2011). *An acoustic analysis of Amharic vowels, plosives and ejectives.* Doctoral dissertation, Addis Ababa University.

Derib Ado. (2020). Revisiting the status of labialised consonants in contemporary Amharic. *Oslo Studies in Language*, *11*(2), 47-58.

Derib Ado. (2021). An acoustic analysis of Amharic fricatives. *Grammatical and Sociolinguistic Aspects of Ethiopian Languages*, *48*, 303.

Devens, M. S. (1981). The Amharic simple imperfect. *Journal of Semitic Studies*, *26*(1), 107-114.

Disner, S. A., & Maddieson, I. (1984). *Patterns of sounds*. Cambridge University Press.

Easterday, S., & Bybee, J. (2023). Diachronic phonological typology: Understanding inventory structure through sound change dynamics. *Linguistic Typology*, *27*(2), 405-427.

Esling, J. H. (2005). There are no back vowels: The larygeal articulator model. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, *50*(1-4), 13-44.

Esling, J. H., & Moisik, S. (2011). Multimodal observation and measurement of larynx height and state during pharyngeal sounds. 20th International Congress of Phonetic Sciences, Hong Kong.

Fabricius, A. H. (2000). *T-glottalling between stigma and prestige: A sociolinguistic study of modern RP.* Doctoral dissertation, Copenhagen Business School.

Fallon, P. D. (2013). *The synchronic and diachronic phonology of ejectives*. Routledge.

Ferguson, C. A. (1970). The Ethiopian language area. *Journal of Ethiopian Studies*, *8*(2), 67-80.

Ferguson, C. A. (1976). The Ethiopian language area. In M. L. Bender, Bowen, J. D., Cooper, R., Ferguson, C. (Eds.), *Language in Ethiopia* (pp. 63–76). London: Oxford University Press.

Gandour, J., & Maddieson, I. (1976). Measuring larynx movement in Standard Thai using the cricothyrometer. *Phonetica*, *33*(4), 241-267.

Gebre, B. (1999). Linguistic survey of the Amhara regional state. *Culture and Tourism Bureau, Bahr Dar. Unpublished*.

Gordeeva, O. B., & Scobbie, J. M. (2011). Laryngeal variation in the Scottish English voice contrast: Glottalisation, ejectivisation and aspiration. In J. M. Scobbie, I. M. Mennen, & Watson, J. (Eds.), *CASL Research Centre Working Paper, WP-19*. Queen Margaret University.

Gordeeva, O. B., & Scobbie, J. M. (2013). A phonetically versatile contrast: Pulmonic and glottalic voicelessness in Scottish English obstruents and voice quality. *Journal of the International Phonetic Association*, *43*(3), 249-271.

Gordon, M., & Applebaum, A. (2006). Phonetic structures of Turkish Kabardian. *Journal of the International Phonetic Association*, *36*(2), 159-186.

Grawunder, S., Simpson, A. P., & Khalilov, M. (2010). Phonetic characteristics of ejectives–samples from Caucasian languages. In S. Fuchs, M. Toda, & Zygis, M. (Eds.), *Turbulent sounds. An interdisciplinary guide* (pp. 209-244). De Gruyter Mouton.

Greenberg, J. H. (1970). Some generalizations concerning glottalic consonants, especially implosives. *International Journal of American Linguistics*, *36*(2), 123-145.

Gussmann, E. (2007). *The phonology of Polish*. Oxford University Press.

Hajek, J., & Stevens, M. (2005). *On the acoustic characterization of ejective stops in Waima'a* Interspeech 2005, Lisbon, Portugal.

Halle, M., & Stevens, K. N. (1971). A note on laryngeal features. *Quarterly progress report, MIT Research Laboratory of Electronics*, *101*, 198-213.

Ham, S. (2007). Tsilhqut'in ejectives. *Annual Conference of the Canadian Linguistic Association* (pp. 1-15).

Ham, S. (2008). *Tsilhqut'in ejectives: a descriptive phonetic study.* Master's thesis, University of Victoria.

Hamlet, S. L. (1980). Ultrasonic measurement of larynx height and vocal fold vibratory pattern. *The Journal of the Acoustical Society of America*, *68*(1), 121-126.

Hargus, S. (2011). *Witsuwit'en grammar: Phonetics, phonology, morphology*. UBC Press.

Hayward, D. (1986). The high central vowel in Amharic: New approaches to an old problem. *The Fergusonian Impact*, *1*, 301-325.

Hayward, K., & Hayward, R. J. (1992). Amharic. *Journal of the International Phonetic Association*, *22*(1-2), 48-52.

Hermes, Z., Fu, M., Rose, S., Shosted, R., & Sutton, B. (2016). Representations of place and airstream mechanism: A real-time MRI study of Tigrinya ejectives. 15th Conference on Laboratory Phonology, Ithaca, USA.

Hetzron, R. (1972). *Ethiopian Semitic: Studies in classification*. Manchester University Press.

Hetzron, R., & Bender, M. L. (1976). The Ethio-Semitic languages. In M. L. Bender, Bowen, J. D., Cooper, R. L., Ferguson, C. A. (Eds.), *Language in Ethiopia* (pp. 2-19). London: Oxford University Press.

Hogan, J. T. (1976). An analysis of the temporal features of ejective consonants. *Phonetica*, *33*(4), 275-284.

Honda, K., Hirai, H., Masaki, S., & Shimada, Y. (1999). Role of vertical larynx movement and cervical lordosis in F0 control. *Language and Speech*, *42*(4), 401-411.

Honda, K., & Tiede, M. (1998). An MRI study on the relationship between oral cavity shape and larynx position. 5th International Conference on Spoken Language Processing, Sydney, Australia.

Hoole, P., Coretta, S., Carignan, C., Kunay, E., Joseph, A., Voit, D., & Frahm, J. (2024). Control of larynx height in vowel production revisited: A real-time MRI study. *Speech Dynamics: Synchronic Variation and Diachronic Change*, *36*, 175.

Hudson, G. (2000). Ethiopian semitic overview. *Journal of Ethiopian Studies*, *33*(2), 75-86.

Ingram, J., & Rigsby, B. (1987). Glottalic stops in Gitksan: An acoustic analysis. 11th International Congress of Phonetic Sciences, Tallinn, Estonia.

International Phonetic Association. (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press.

Javkin, H. (1977). Towards a phonetic explanation for universal preferences in implosives and ejectives. Annual Meeting of the Berkeley Linguistics Society, Berkeley, California.

Jun, J. (2004). Place assimilation. In *Phonetically based phonology* (pp. 58-86). Cambridge University Press.

Kane, T. L. (1990). *Amharic-English dictionary* (Vol. 1). Otto Harrassowitz Verlag.

Kehrein, W. (2013). *Phonological representation and phonetic phasing: Affricates and laryngeals* (Vol. 466). Walter de Gruyter.

Kingston, J. (1984). *The phonetics and phonology of the timing of oral and glottal events.* Doctoral dissertation, University of California, Berkeley.

Kingston, J. (1985). The ineffectiveness of larynx movement. *The Journal of the Acoustical Society of America*, *77*(S1), S86-S87.

Kingston, J. (2005). The phonetics of Athabaskan tonogenesis. In *Amsterdam Studies in the Theory and History of Linguistic Science* (Vol. 269, pp. 137-184). John Benjamins Publishing Company.

Kiparsky, P. (2003). The phonological basis of sound change. *The handbook of historical linguistics*, 311-342.

Kirchner, R. (2013). *An effort based approach to consonant lenition*. Routledge.

Kunay, E., Hoole, P., Gubian, M., Harrington, J., Joseph, A., Voit, D., & Frahm, J. (2022). Vowel height and velum position in German: Insights from a real-time magnetic resonance imaging study. *The Journal of the Acoustical Society of America*, *152*(6), 3483-3501.

Kuortti, J., Malinen, J., & Ojalammi, A. (2018). Post-processing speech recordings during MRI. *Biomedical Signal Processing and Control*, *39*, 11-22.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1-26.

Kye, T. (2021). *Ejective typology: The case of Lushootseed* ICSNL 59, online conference.

Ladefoged, P., & Johnson, K. (2006). *A course in phonetics* (Vol. 3). Thomson Wadsworth Boston.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages* (Vol. 1012). Blackwell Oxford.

Lammert, A. C., Proctor, M. I., & Narayanan, S. S. (2010). Data-driven analysis of realtime vocal tract MRI using correlated image regions. Interspeech 2010, Makuhari, Japan.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174.

Lass, R. (1984). *Phonology: An introduction to basic concepts*. Cambridge University Press.

Laukkanen, A.-M., Takalo, R., Vilkman, E., Nummenranta, J., & Lipponen, T. (1999). Simultaneous videofluorographic and dual-channel electroglottographic registration of the vertical laryngeal position in various phonatory tasks. *Journal of Voice*, *13*(1), 60-71.

Lavoie, L. M. (2001). *Consonant strength: Phonological patterns and phonetic manifestations*. Routledge.

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2021). Emmeans: Estimated marginal means, aka least-squares means (R package version 1.5. 1.)[Computer software]. *The Comprehensive R Archive Network. Available online: https://CRAN. R-project. org/package=emmeans (last viewed 23 June 2025)*.

Leslau, W. (1976). *Concise Amharic dictionary*. University of California Press.

Leslau, W. (1995). *Reference grammar of Amharic*. Wiesbaden: Otto Harrassowitz Verlag.

Leslau, W. (1996). Historical phonetics of Amharic. *The Journal of the American Oriental Society*, *116*(2), 298-302.

Leslau, W. (1997). Amharic phonology. *Phonologies of Asia and Africa*, *1*, 399-430.

Lindau, M. (1975). Larynx height in Kwa. *The Journal of the Acoustical Society of America*, *58*(S1), S12.

Lindau, M. (1984). Phonetic differences in glottalic consonants. *Journal of Phonetics*, *12*(2), 147-155.

Lindqvist, J., Sawashima, M., & Hirose, H. (1973). An investigation of the vertical movement of the larynx in a Swedish speaker. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics*, *7*, 27-34.

Lingala, S. G., Zhu, Y., Kim, Y. C., Toutios, A., Narayanan, S., & Nayak, K. S. (2017). A fast and flexible MRI system for the study of dynamic vocal tract shaping. *Magnetic Resonance in Medicine*, *77*(1), 112-125.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384-422.

Macafee, C. (1983). *Varieties of English around the world* (G. Melchers, Ed. Vol. 198). Benjamins.

Maddieson, I. (1984). Glottalic and laryngealized consonants. In *Patterns of Sounds* (pp. 98-122). Cambridge University Press.

Maddieson, I. (1985). Phonetic cues to syllabification. *Phonetic linguistics: Essays in honor of Peter Ladefoged*, *203*, 221.

Maddieson, I. (1998). Why make life hard? Resolutions to problems of rare and difficult sound types. Annual Meeting of the Berkeley Linguistics Society,

Maddieson, I. (2013). Glottalized consonants. In M. S. Dryer & Haspelmath, M. (Eds.), *WALS Online (last viewed 23 June 2025)* (v2020.4 ed.). Max Planck Institute for Evolutionary Anthropology. https://doi.org/10.5281/zenodo.13950591

Maddieson, I., & Ladefoged, P. (1993). Phonetics of partially nasal consonants. In *Nasals, nasalization, and the velum* (pp. 251-301). Academic Press.

Maddieson, I., Smith, C. L., & Bessell, N. (2001). Aspects of the phonetics of Tlingit. *Anthropological Linguistics*, 135-176.

Martins, P., Oliveira, C., Silva, S., & Teixeira, A. (2012). Velar movement in European Portuguese nasal vowels. Proceedings of IberSpeech—VII Jornadas en Tecnología del Habla and II Iberian SLTech Workshop, Madrid, Spain.

McCarthy, O. (2012). *A sociophonetic study of the realization of word-final velar plosives by female pupils in a Glasgow high school.* Doctoral dissertation, University of Glasgow.

McCarthy, O., & Stuart-Smith, J. (2013). Ejectives in Scottish English: A social perspective. *Journal of the International Phonetic Association*, *43*(3), 273-298.

Mengistu, T. (2018). *Documentation and description of Amharic dialects.* Doctoral dissertation, Addis Ababa University.

Milroy, J., Milroy, L., Hartley, S., & Walshaw, D. (1994). Glottal stops and Tyneside glottalization: Competing patterns of variation and change in British English. *Language Variation & Change*, *6*(3).

Moisik, S. (2013). *The epilarynx in speech.* Doctoral dissertation, University of Victoria.

Moisik, S., Esling, J. H., Crevier-Buchman, L., Amelot, A., & Halimi, P. (2015). Multimodal imaging of glottal stop and creaky voice: Evaluating the role of epilaryngeal constriction. 18th International Congress of Phonetic Sciences, Glasgow, Scotland.

Moisik, S. R., Esling, J. H., Crevier-Buchman, L., & Halimi, P. (2019). Putting the larynx in the vowel space: A pilot study of laryngeal articulatory state across vowel quality in canonical phonetic productions using MRI. 19th International Congress of Phonetic Sciences Melbourne, Australia.

Moran, S., & McCloy, D. (Eds.). (2019). *PHOIBLE 2.0* (Vol. 10). Max Planck Institute for the Science of Human History. http://phoible.org.

Narayanan, S., Nayak, K., Lee, S., Sethy, A., & Byrd, D. (2004). An approach to real-time magnetic resonance imaging for speech production. *The Journal of the Acoustical Society of America*, *115*(4), 1771-1776.

Negesse, F. (2023). An acoustic analysis of Oromo and Amharic ejective stops. *Journal of the International Phonetic Association*, *53*(3), 1033-1048.

Neuschaefer-Rube, C., Wein, B., Angerstein, W., & Klajman, S. (1996). Kernspintomographische Untersuchung der Kehlkopfstellungen beim Singen von Vokalen. *Folia Phoniatrica et Logopaedica*, *48*(4), 201-209.

Niebergall, A., Zhang, S., Kunay, E., Keydana, G., Job, M., Uecker, M., & Frahm, J. (2013). Real-time MRI of speaking at a resolution of 33 ms: Undersampled radial FLASH with nonlinear inverse reconstruction. *Magnetic Resonance in Medicine*, *69*(2), 477-485.

Noguchi, M. (2015). An acoustic study of pulmonic and ejective stops in final coda position in Yucatec. Proceedings of Meetings on Acoustics, Pittsburgh, Pennsylvania.

Ogden, R. (2024). *An introduction to English phonetics*. Edinburgh University Press.

Oh, M., Byrd, D., Goldstein, L., & Narayanan, S. S. (2024). Vertical larynx actions and intergestural timing stability in Hausa ejectives and implosives. *Phonetica*, *81*(6), 559-597.

Park, H., & Kim, H.-J. (2015). Acoustic characteristics of Aymara ejectives: A pilot study. 18th International Congress of Phonetic Sciences, Glasgow, Scotland.

Percival, M. (2019). Contextual variation in the acoustics of Hul'q'umi'num'ejective stops. 19th International Congress of Phonetic Sciences, Melbourne, Australia.

Percival, M. (2023). Perceptual cues to ejectives across languages. 20th International Congress of Phonetic Sciences, Prague, Czech Republic.

Proctor, M. I., Goldstein, L., Lammert, A. C., Byrd, D., Toutios, A., & Narayanan, S. S. (2013). Velic coordination in French nasals: A real-time magnetic resonance imaging study. Interspeech 2013, Lyon, France.

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ramanarayanan, V., Tilsen, S., Proctor, M., Töger, J., Goldstein, L., Nayak, K. S., & Narayanan, S. (2018). Analysis of speech production real-time MRI. *Computer Speech & Language*, *52*, 1-22.

Roach, P. (2002). *A little encyclopaedia of phonetics*. University of Reading.

Roach, P. J. (1979). Laryngeal-oral coarticulation in glottalized English plosives. *Journal of the International Phonetic Association*, *9*(1), 2-6.

Rose, S. (1997). *Theoretical issues in comparative Ethio-Semitic phonology and morphology.* Doctoral dissertation, McGill University.

Sande, H., & Hedding, A. (2017). Syllable weight in Amharic. In P. Newman (Ed.), *Syllable weight in african languages* (Vol. 338, pp. 69-82). John Benjamins.

Seid, H., Rajendran, S., & Yegnanarayana, B. (2009). Acoustic characteristics of ejectives in Amharic. Interspeech 2009, Brighton, United Kingdom.

Shosted, R., Sutton, B. P., & Benmamoun, A. (2012). Using magnetic resonance to image the pharynx during Arabic speech: Static and dynamic aspects. Interspeech 2012, Portland, OR, USA.

Shosted, R. K., & Rose, S. (2011). Affricating ejective fricatives: The case of Tigrinya. *Journal of the International Phonetic Association*, *41*(1), 41-65.

Simpson, A. (2007). Acoustic and auditory correlates of non-pulmonic sound production in German. *Journal of the International Phonetic Association*, *37*(2), 173-182.

Simpson, A. (2014). Ejectives in English and German: Linguistic, sociophonetic, interactional, epiphenomenal? In *Advances in sociophonetics* (pp. 189-204). John Benjamins Publishing Company.

Simpson, A., & Brandt, E. (2019). Detecting larynx movement in non-pulmonic consonants using dual-channel electroglottography. 19th International Congress of Phonetic Sciences, Melbourne, Australia.

Simpson, A., & Sulaberidze, N. (2022). Ejectives in English: Elicitation and analysis. Phonetik und Phonologie im deutschsprachigen Raum, Bielefeld, Germany.

Solé, M.-J. (2023). Aerodynamic parameters of ejectives and pulmonic stops. 20th International Congress of Phonetic Sciences Prague, Czech Republic.

Steriade, D. (2001). The phonology of perceptibility effects: The P-map and its implications for constraint organization. *The Nature of the Word: Essays in Honor of Paul Kiparsky*.

Stevens, M., & Hajek, J. (2004). A preliminary investigation of some acoustic characteristics of ejectives in Waima'a: VOT and closure duration.

Stevens, M., & Hajek, J. (2008). Positional effects on the characterization of ejectives in Waima'a. Interspeech 2008, Brisbane, Australia.

Stuart-Smith, J. (1999). Glottals past and present: A study of T-glottalling in Glaswegian. *Leeds Studies in English*, 181-204.

Sulaberidze, N. (2022). Werden deutsche epiphänomenale Ejektive als 'echte' Ejektive wahrgenommen? Phonetik und Phonologie im deutschsprachigen Raum, Bielefeld, Germany.

Sulaberidze, N., Brandt, E., Hoole, P., Krämer, M., Reichenbach, J. R., & Simpson, A. (2023). Ejectives in Georgian. A real-time MRI analysis of vertical larynx movement. 20th International Congress of Phonetic Sciences, Prague, Czech Republic.

Sumner, C. (1957). *Étude expérimentale de l'amharique moderne*. Addis Ababa: The University College Press.

Takkele, T. (1992). Are s' and t' variants of an Amharic variable? A sociolinguistic analysis. *Journal of Ethiopian Languages and Literature*(2), 103-121.

The MathWorks Inc. (2022a). *Matlab [Computer Program]*.

Tiede, M. K. (1996). An MRI-based study of pharyngeal volume contrasts in Akan and English. *Journal of Phonetics*, *24*(4), 399-421.

Toparlak, T., & Dolatian, H. (2023). Aerodynamics and articulation of word-final ejectives in Eastern Armenian. 20th International Congress of Phonetic Sciences, Prague, Czech Republic.

Trigo, L. (1991). On pharynx—larynx interactions. *Phonology*, *8*(1), 113-136.

Uecker, M., Zhang, S., Voit, D., Karaus, A., Merboldt, K. D., & Frahm, J. (2010). Real-time MRI at a resolution of 20 ms. *NMR in Biomedicine*, *23*(8), 986-994.

Ullendorff, E. (1955). *The Semitic languages of Ethiopia: A comparative phonology*. Taylor's (Foreign) Press.

Vicenik, C. (2010). An acoustic study of Georgian stop consonants. *Journal of the International Phonetic Association*, *40*(1), 59-92.

Wang, G., & Kong, J. (2010). The relation between larynx height and F0 during the four tones of Mandarin in X-ray movie. 2010 7th International Symposium on Chinese Spoken Language Processing, Tainan, Taiwan.

Warner, N. (1996). Acoustic characteristics of ejectives in Ingush. 4th International Conference on Spoken Language Processing, Philadelphia, PA, USA.

Wells, J. C. (1982). *Accents of English* (Vol. 1). Cambridge University Press.

Winter, B. (2015). The other N: The role of repetitions and items in the design of phonetic experiments. 18th International Congress of Phonetic Sciences, Glasgow, Scotland.

Wright, R. (2004). A review of perceptual cues and cue robustness. *Phonetically based phonology*, *34*(57).

Wright, R., Hargus, S., & Davis, K. (2002). On the categorization of ejectives: Data from Witsuwit'en. *Journal of the International Phonetic Association*, *32*(1), 43-77.

Wysocki, T. M. (2004). *Acoustic analysis of Georgian stop consonants and stop clusters.* Doctoral dissertation, University of Chicago.

Appendices

Appendix A
Amharic Word List

The following list presents the lexical items used as stimuli in the Amharic study (Chapter 3). Words are organized by stop type (ejective, pulmonic voiceless) and place of articulation, and further subdivided by word position (initial, medial, final) and consonant length where applicable.

Amharic uses its own script. During the experiment, sentences were presented to participants in the native Amharic writing system. Here, all Amharic words are provided in transliteration following the conventions of Leslau (1976).

Ejective	**Pulmonic voiceless**
Bilabial	
p'awlos	paylät
p'app'as	papaya
p'ag^{w}əme	paket
p'äləqan	pasifik
ityop'əya	tapela
azop'i	səpa
esop'	tep
p'əpp'əsənna	ğappan
lapp'is	filippinə
awropp'a	appartama
Alveolar	
t'afä	tafa
t'ämma	tamma
t'asa	taza
t'änna	täňňa
t'älla	tälla
t'äffa	täffa
t'asä	taššä
t'alä	tägga
t'älat	tammaň
t'äläs	tannaš
t'adä	taffaň
t'ana	täsäffa
t'aba	täballa
t'äbba	täbässa
t'amä	täbazza
t'ata	täsagä
nat'ä	batä
lat'ä	läta
gat'a	gäta
dat'a	data
wat'ä	fata
mat'ä	mata
šät'ä	šätä
gat'ä	gatä
dat'ä	zatä
šat'a	satä

mälat'a	wəläta	att'a	tattämä
tannät'ä	tabbätä	mätt'	fätt
at'äf		nətt'	kətt
at'aň		šəmətt'	säbatt
mät'än		**Velar**	
mat'än		k'älla	källa
lät'äb		k'ana	kanä
zät'äň		k'ata	kadä
gəmmət'	ğəmmat	k'äbba	kabba
səmət'	labät	k'änna	kännä
ləmmət'	ləmat	k'amä	kabä
läwwət'	lamät	k'ämma	käffa
mägat'	wəgat	k'ädda	kädda
mäwat'	fəlat	k'aňňä	kässa
näwwət'	malät	k'ažžä	käzza
mämsät'	lədät	k'äga	kasä
mäblät'	ləfat	k'äläm	kəlləl
mäsmät'	dəfat	k'ällağ	källaš
mäflät'	zəlät	k'äsäm	kətəb
sət'	sat	k'ämaš	kəllet
mət'	bat	zak'ä	zäka
mat'	lat	sak'ä	waka
wät'	lät	lak'ä	lakä
šät'	nät	nak'ä	dakä
mätt'a	mätta	fak'ä	akafa
nätt'a	gätta	mäk'a	akäkä
sätt'ä	šətta	näk'äz	näkač
bätt'a	fätta	mäk'an	mäkan
att'änä	attätä	näk'aš	näkaš
att'äbä	attämä	nək'ät	takač
att'äfä	lättämä	läk'äm	
wätt'a	bättänä	alläk'ä	
tatt'a	tattätä	ammäk'ä	

səllak'	mälʻak	näkk'a	näkka
wälläk'	mäläk	wäkk'a	fäkka
däfäk'	fäläk	däkk'a	täkka
mawäk'	mawäk	bäkk'a	mäkka
dämmak'		läkk'a	läkka
tallak'		akk'äfä	däkka
wäbäk'		akk'ädä	käkka
läbäk'		bäkk'älä	č'akka
bäläk'		säkk'älä	kakka
mädfäk'	məsbak	šäkk'älä	akkämä
mäfläk'	amlak	šakk'älä	akkätä
mädmäk'	məsmak	šäkk'äbä	takkälä
mänfäk'	täwsak	näkk'älä	takkätä
mädälläk'	mälalak	läkk'ämä	takkämä
mädäbäk'	attäkak	makk'aňa	bakkänä
mak'	annäkak	məkk'äňa	bäkkälä
sak'		hakk'	kəkk
mäk'		ləkk'	ləkk
fak'		muləkk'	bäləkk
näk'		fələkk'	mətəkk
wək'	kek		
nək'	nək		

Appendix B
Amharic Bilabial Ejectives

Bilabial stops in Amharic are extremely marginal and occur only in loanwords (see Chapter 3, Section 3.1.2). As a result, the dataset showed a significant imbalance in token distribution, with only 20 bilabial target words included in the original word list. For this reason, bilabial stops were excluded from the main analysis (see Chapter 3, Section 3.2.3).

This appendix presents figures for acoustic and articulatory measures based on the available bilabial ejective tokens, in order to nevertheless provide some insight into their production. However, due to challenges in computing these measures arising from rt-MRI data quality (see Chapter 2, Section 2.7), which further reduced token availability across key factors (e.g., stop type, word position, and consonant length), certain measures, specifically those calculating changes in supralaryngeal cavity size and hyperpharyngeal constriction, are not included here, as they would not yield interpretable patterns even in this supplementary context.

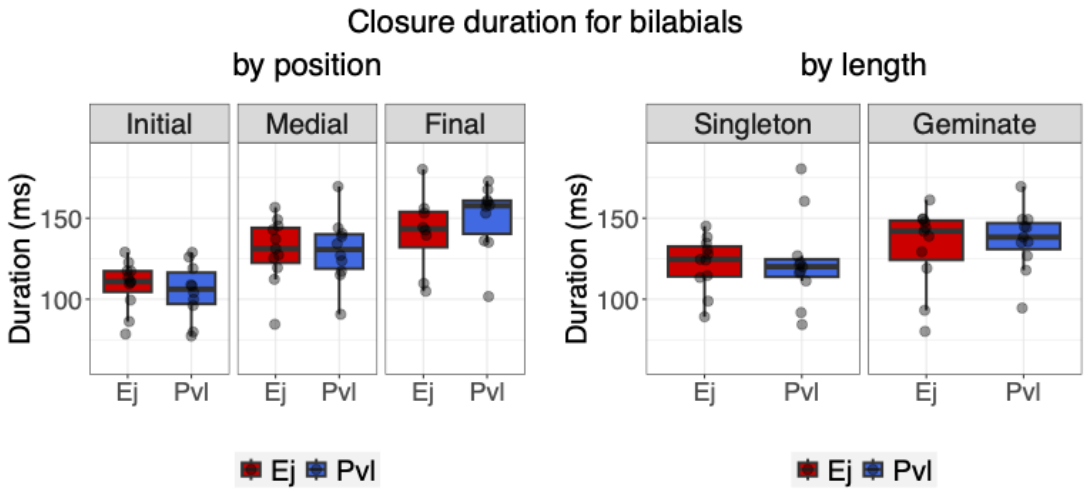

Figure B.44 Raw closure duration values (y-axis, in ms) for Amharic bilabial ejective and pulmonic voiceless stops split by word position (initial, medial, final), and length (singleton, geminate). Dots represent speaker-level means; higher values indicate longer closures.

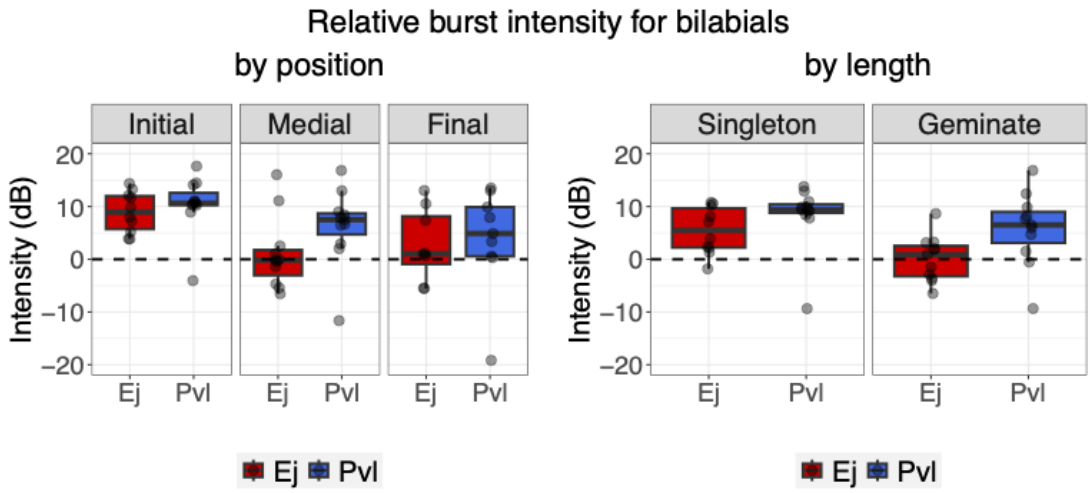

Figure B.45 Relative burst intensity values (y-axis, in dB) for Amharic bilabial ejective and pulmonic voiceless stops split by word position (initial, medial, final), and length (singleton, geminate). Dots represent speaker-level means. Values reflect the peak intensity difference between the following vowel and the stop's release burst. Lower values (and especially negative values) indicate a stronger intensity of the burst relative to the vowel.

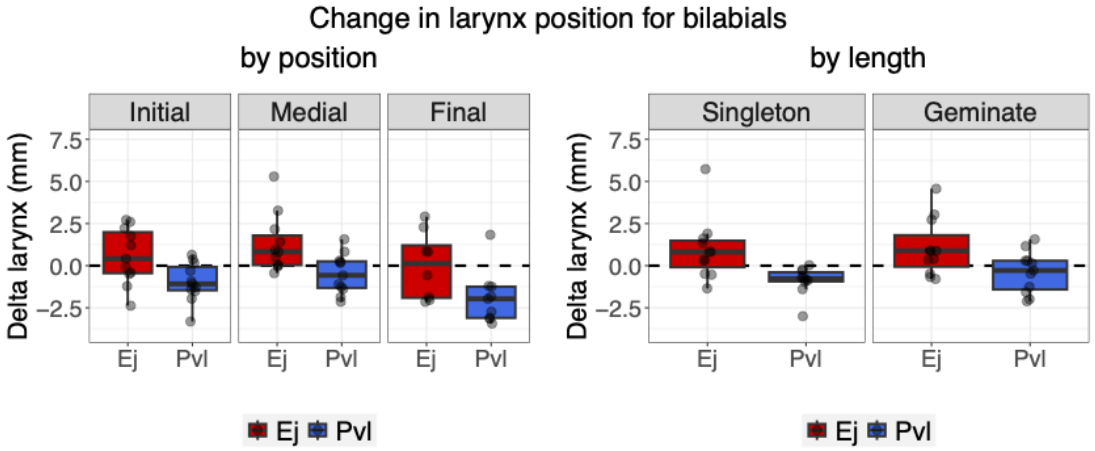

Figure B.46 Delta larynx values (y-axis, in mm) of Amharic bilabial ejective and pulmonic voiceless stops, split by word position (initial, medial, final), and length (singleton, geminate). Delta larynx is measured as the difference between the larynx's position at the onset of stop closure and its position at closure offset. Dots represent speaker-level means; positive numbers indicate larynx raising, negative numbers indicate larynx lowering.

Appendix C

Supplementary Tables for Chapter 3

This appendix presents supplementary tables referenced in Section 3.3.7, containing detailed descriptive statistics for the acoustic analyses reported in Chapter 3, including results by individual speaker, place of articulation (PoA), word position, and consonant length.

Summary of the Acoustics Results by Speaker

Table C.38 Means (M) and standard deviations (SD, in parentheses) of each acoustic measure by speaker. Bold values highlight notable deviations from the overall pattern. The cell outlined in dark grey marks the speaker for whom 20% of ejectives were unreleased.

	Speaker										
	S01	**S02**	**S03**	**S04**	**S05**	**S06**	**S07**	**S08**	**S09**	**S10**	**S11**
Measure	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Intensity rise ratio (dB)	2.5 (2.3)	**6.0 (4.3)**	2.1 (2.2)	0.9 (1.8)	2.2 (3.0)	3.4 (2.8)	2.1 (2.9)	3.0 (2.8)	1.2 (1.6)	**6.5 (3.6)**	**4.8 (2.9)**
VOT (ms)	**33.4 (14.0)**	68.2 (41.2)	**30.8 (14.5)**	113.9 (31.8)	94.6 (52.1)	68.5 (39.9)	67 (41.6)	91.1 (63.2)	52.8 (42.3)	87.1 (63.5)	84.7 (70.4)
Relative burst intensity (dB)	**13.3 (3.9)**	8.4 (7.7)	6.0 (3.8)	**-1.4 (4.9)**	4.3 (3.8)	4.3 (4.6)	7.7 (5.3)	7.8 (3.6)	10.0 (6.0)	5.9 (4.1)	8.5 (3.6)
PBL intensity (dB)	59.5 (1.3)	46.2 (2.0)	53.2 (3.6)	49.6 (1.4)	53.9 (1.5)	59.1 (2.8)	63.5 (1.4)	68.4 (2.2)	63.9 (1.6)	58.0 (3.4)	61.2 (2.5)
Closure duration (ms)	99.8 (40.9)	103.3 (47.6)	81.4 (15.4)	107.1 (52.9)	109.8 (40.7)	103.8 (33.3)	104.6 (40.9)	119.1 (64.3)	82.2 (36.8)	126.2 (55.5)	97.5 (33.4)
Burst duration (ms)	**10.3 (3.8)**	13.7 (5.4)	15.9 (4.1)	20.3 (8.2)	22 (8.6)	23.8 (8.8)	12.5 (4.2)	16.2 (6.3)	14.5 (8.1)	24.3 (11.9)	19.7 (10.8)

Summary of the Acoustics Results by Place of Articulation, Word Position, and Consonant Length

Table C.39 Means (M) and standard deviations (SD, in parentheses) of each acoustic measure by PoA, word position, and length.

	PoA		Word position			Length	
	/t'/	/k'/	**Initial**	**Medial**	**Final**	**Singleton**	**Geminate**
Measure	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Intensity rise ratio (dB)	3.4 (3.4)	2.9 (3.3)	3.3 (3.5)	2.6 (2.8)	3.8 (3.8)	3.4 (3.5)	2.4 (2.7)
VOT (ms)	69.2 (54.3)	76.8 (51.4)	68.2 (29.6)	49.7 (26.7)	113 (70.5)	75.6 (53)	67.3 (51.8)
Relative burst intensity (dB)	8.2 (5.8)	5.2 (5.6)	7.9 (6.2)	6.8 (5.8)	5.0 (5.4)	6.7 (5.9)	5.8 (5.6)
PBL intensity (dB)	58.4 (7.0)	57.4 (6.5)	57.3 (6.7)	58.5 (6.9)	57.3 (6.5)	57.7 (6.7)	58.2 (7.0)
Closure duration (ms)	98.4 (44.0)	107.4 (45.7)	92.3 (19.1)	105.7 (49.2)	106.2 (48.7)	90.3 (35.3)	140.8 (49.1)
Burst duration (ms)	16.1 (8.8)	19.4 (8.9)	15.5 (5.6)	15.8 (6.5)	22.7 (11.7)	17.9 (8.8)	18.3 (9.5)

Appendix D
Pharynx Length and Larynx Movement Range in Amharic Speakers

1. Purpose of Analysis

This supplementary analysis investigates whether pharynx length influences the range of vertical larynx movement during ejective stop closure. The main analysis in Chapter 3, Section 3.4.1 showed that, across speakers, ejective stops in Amharic are consistently produced with larynx raising, but that the magnitude of larynx raising varies widely across individuals. Similar variability has been noted in prior work (Bückins et al., 2018; Kingston, 1984), where it has been attributed to both stylistic and anatomical factors. This analysis evaluates the anatomical account by testing whether inter-speaker differences in pharynx length predict the extent of larynx raising during ejective stop closure. Establishing this relationship is necessary in order to determine whether the variability observed in Chapter 3 reflects anatomical constraints.

2. Methods

Data. This analysis uses the subset of the dataset containing only ejective tokens, since ejectives are the stop type that actively relies on vertical larynx displacement for airstream generation.

Measure of larynx movement. As in the main analyses (Chapter 3), larynx displacement was quantified as *delta larynx* (mm), defined as the difference between larynx position at onset and offset of stop closure (see Chapter 2, Section 2.7.1). Speaker-wise means and standard deviations for *delta larynx* of ejectives are given in **Table D.40.**

Pharynx length. Pharynx length was measured per speaker from the rt-MRI data (**Table D.40**). One speaker (S06) stands out for exhibiting both an unusually long pharynx (92.7 mm) and the greatest mean larynx displacement during ejective closure (6.2 mm).

Table D.40 Speaker-wise pharynx length (in mm) and mean change in larynx position during ejective closure (*delta larynx*, in mm). Data are from Amharic speakers.

Speaker	Pharynx length (mm)	Delta larynx (mm) *M(SD)*
S01	57.7	0.96 (0.95)
S02	58.0	1.3 (1.4)
S03	67.4	3.6 (1.6)
S04	62.9	1.4 (1.6)
S05	76.9	2.5 (1.3)
S06	92.7	6.2 (3.0)
S07	68.3	1.3 (1.0)
S08	68.2	1.3 (1.2)
S09	74.7	2.1 (1.7)
S10	75.6	3.5 (2.3)
S11	64.7	2.4 (2.2)

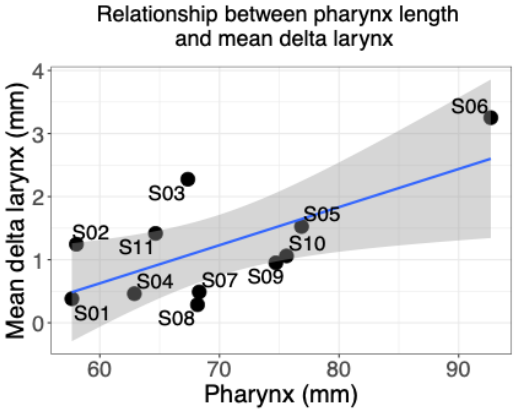

Figure D.47 Speaker means of *delta larynx* (mm) representing the average change in larynx position over stop closure, plotted against pharynx length (mm) for each of the 11 Amharic speakers (S01-S11) in this dataset. The blue line represents the fitted linear trend and the shaded region the 95% confidence interval.

Statistical model. A linear mixed-effects model tested whether pharynx length predicts larynx displacement during ejective stop closure. Pharynx length was mean-centered (Pharynx_c), meaning that the intercept corresponds to the expected *delta larynx* for a speaker with average pharynx length. Random intercepts for speaker and word were included to account for repeated measures and baseline differences across speakers not explained by pharynx length, and lexical items:

model = lmer(deltaLarynx ~ Pharynx_c + (1 | Speaker) + (1 | Word)).

3. Results

The model revealed a significant positive effect of pharynx length on *delta larynx* ($F(1, 9.01) = 22.23$, $p < .01$). The estimated slope indicates an increase of approximately 0.13 mm in larynx raising for each additional 1 mm of pharynx length (estimate = 0.13, SE = 0.03). Random-effects estimates indicated non-trivial baseline variability across speakers (SD = 0.87 mm) and words (SD = 0.69 mm), with substantial residual variation at the token level (SD = 1.63 mm). This suggests that anatomical differences contribute to, but do not fully account for inter-speaker variability in the extent of larynx raising.

Because pharynx length included a clear outlier speaker (see S06 in **Table D.40**), this point may exert substantial leverage on the fitted slope in a dataset with only 11 speakers (as is hinted at in **Figure D.47**). Therefore, a leave-one-out analysis was conducted to evaluate the robustness of the pharynx-length effect.

Table D.41 Leave-one-out sensitivity analysis for the effect of pharynx length on larynx displacement during ejective stop closure. Each row reports the fixed-effect slope for pharynx_c from the model refit with the indicated speaker excluded. The final column shows the change in this slope relative to the full model. Boldface marks the most influential case (S06).

Left out speaker	Estimate (mm per mm)	SE	DF	t	p	Signif.	Diff. from full model estimate (mm per mm)
S01	0.13	0.03	8.01	4.14	0.0032	**	0.002
S02	0.14	0.03	8.00	4.36	0.0024	**	0.006
S03	0.14	0.02	8.00	5.63	<.001	***	0.004
S04	0.13	0.03	8.01	4.30	0.0026	**	-0.001
S05	0.14	0.03	8.01	4.88	0.0012	**	0.007
S06	**0.08**	**0.04**	**8.02**	**2.09**	**0.0697**		**-0.049**
S07	0.13	0.03	8.01	4.70	0.0015	**	-0.001
S08	0.13	0.03	8.01	4.71	0.0015	**	-0.002
S09	0.14	0.03	8.01	5.00	0.0011	**	0.006
S10	0.13	0.03	8.01	4.32	0.0025	**	-0.002
S11	0.14	0.03	8.01	4.69	0.0016	**	0.004

As shown in **Table D.41**, the pharynx-length effect was highly stable when excluding any speaker other than S06: the slope estimates clustered around 0.13–0.14 mm/mm and remained statistically reliable (all $p \leq .003$). In contrast, excluding S06 reduced the estimated effect to 0.08 mm/mm (SE = 0.04), and the effect no longer reached significance ($p = .07$). Notably, the estimate remained positive, indicating that the direction of the relationship is consistent, but that statistical support for the effect depends in part on the inclusion of the speaker with the most extreme pharynx length.

4. Interpretation of Pharynx Length Effects

Overall, the results are consistent with a positive association between pharynx length and the magnitude of larynx raising during ejective stop closure. This supports the idea that anatomical variation contributes to the inter-speaker differences observed in Chapter 3.

This is broadly in line with Bückins et al. (2018), who proposed anatomical differences as a source of speaker variability. At the same time, the leave-one-out analysis shows that the effect is sensitive to an outlier speaker (S06). This underscores that conclusions drawn from a small speaker sample ($n = 11$) should be interpreted cautiously, as individual cases can disproportionately affect results.

This sensitivity also leaves room for the possibility raised by Kingston (1984) that some variability reflects stylistic or speaker-specific articulatory strategies in addition to anatomical differences.

Note on Gender. Gender was not included as a predictor because the sample is unbalanced (female $n = 4$, male $n = 7$), which limits the interpretability of any gender-effects. Moreover, Bückins et al. (2018) reported no systematic male–female differences that would explain variability in larynx movement magnitude.

With this supplementary check in mind, the main analysis can remain focused on linguistic factors such as PoA, word position, and the geminate–singleton contrast while more targeted evaluation of anatomical and gender-related effects is left to future work with larger, more balanced samples. Finally, because the English dataset was preprocessed differently, an equivalent pharynx-length measure could not be derived, and no parallel analysis was conducted for English speakers.

Appendix E
English Word List

The following list presents the lexical items used as stimuli in the English ejective study (Chapter 4). The words represent a subset drawn from a larger corpus originally designed for a study on nasalization. For the present investigation on English ejectives, items were selected to include both voiced and voiceless stops in coda position – the environment in which ejective productions in English are observed. Stimuli span three places of articulation (bilabial, alveolar, and velar).

Bilabial
pap
sap
beep
seep
Pip
pimp
sip
boop
bump
hump
sump
sup
fab
beeb
boob
hub
nub
snub

Alveolar
bat
fat
pat
sat
bet
pet
set
bit
fit
pit
sit
beet
peet
seet
bait
fate
paight
sate
boot
but
hut
put
fought
pot
saught
pant
bent
sent
bint
faint
paint
saint
bunt
hunt
punt
shunt
bad
fad
pad
bed
fed
said
bid
Sid
beed
peed
seed
bade
fade
made
paid
smade
spade
shade
booed
bud
fud
pawed
sawed
beamed
seemed
boomed
bummed
band
fanned
panned
sand
bend
fend
penned
send
binned
finned
pinned
sinned
beaned
scened
feigned
pained
fund
punned
shunned
fond
pawned
banged
fanged
pinged
bunged

Velar
back
pack
sack
sank
pick
sick
pink
sink
buck
huck
suck

bunk	fag	mug	Sig
hunk	bug	pig	smug
sunk	hug	pug	
bag	lug	plug	

Appendix F

Validation of the Auditory Categorization Procedure for English Ejectives

This appendix presents an interrater reliability assessment for the auditory classification of coda stops described in Section 4.3 of the English chapter (Chapter 4). Given the subjectivity involved in auditory labeling, it is important to assess the reliability of the categorization used to identify the ejective realizations in the English dataset. This section presents interrater agreement data to evaluate the consistency of judgments across two listeners – a first rater (the author) and a second, independent researcher – and to contextualize the strength of the auditory results.

Interrater reliability for the auditory categorization of stop types was assessed using a randomly selected subset of the data. A total of 492 tokens were included, comprising one third of the ejective tokens initially identified by the first rater, supplemented by an equal number of tokens from the pulmonic and unclear categories. A second rater, who was informed of the three available label options but received no further guidance, independently classified the tokens.

The overall observer agreement was 55.3%. Cohen's Kappa was calculated to assess interrater reliability and yielded a value of $\kappa = 0.329$, indicating fair agreement according to the Landis and Koch (1977) scale. When analyzed by category, agreement was highest for pulmonic tokens (90.9%), followed by ejectives (68.9%). In contrast, agreement for the unclear category was notably low (6.1%), likely reflecting the second rater's greater willingness to assign a categorical label, in contrast to the more conservative labeling approach taken by the first rater.

Appendix G
Supplementary Tables for Chapter 4

This appendix includes supplementary tables showing post hoc pairwise contrasts referenced in Chapter 4. The tables below show the full post hoc pairwise comparisons from the linear mixed-effects model described in Section 4.5.1, for *delta larynx* (log scale). They include back-transformed absolute differences in mm, grouped by stop type (**Table G.42**) and place of articulation (**Table G.43**).

Table G.42 Pairwise comparisons of delta larynx (modelled on the log scale) by stop type, with back-transformed absolute differences in mm.

Contrast type	Stop type	Contrast	Estimate (log)	SE	DF	t	p	Signif.	Δ (mm)
By stop type	Pvl	bil - alv	-0.24	0.13	395.0	-1.86	0.1515		-0.12
By stop type	Pvl	vel - alv	-0.32	0.12	442.6	-2.70	0.0197	*	-0.16
By stop type	Pvl	vel - bil	-0.08	0.15	393.1	-0.50	0.8723		-0.03
By stop type	Ej	bil - alv	-0.32	0.13	391.8	-2.55	0.0300	*	-0.28
By stop type	Ej	vel - alv	0.33	0.11	332.1	3.09	0.0061	**	0.39
By stop type	Ej	vel - bil	0.65	0.14	305.4	4.61	<.001	***	0.67
By stop type	Pvd	bil - alv	-0.20	0.13	373.0	-1.48	0.3009		-0.13
By stop type	Pvd	vel - alv	0.30	0.11	369.5	2.69	0.0204	*	0.26
By stop type	Pvd	vel - bil	0.50	0.16	385.1	3.14	0.0051	**	0.39
By stop type	unclear	bil - alv	-0.26	0.10	186.9	-2.48	0.0368	*	-0.17
By stop type	unclear	vel - alv	0.30	0.09	211.7	3.17	0.0050	**	0.25
By stop type	unclear	vel - bil	0.55	0.13	202.5	4.39	<.001	***	0.42

Table G.43 Pairwise comparisons of delta larynx (modelled on the log scale) by PoA, with back-transformed absolute differences in mm.

Contrast type	PoA	Contrast	Estimate (log)	SE	DF	t	p	Signif.	Δ (mm)
By PoA	bil	Ej - Pvl	0.46	0.18	45.8	2.56	0.0646	.	0.27
By PoA	bil	Pvd - Pvl	0.28	0.16	198.6	1.77	0.2912		0.15
By PoA	bil	Pvd - Ej	-0.18	0.16	129.8	-1.10	0.6898		-0.12
By PoA	bil	unclear - Pvl	0.22	0.12	91.6	1.80	0.2821		0.11
By PoA	bil	unclear - Ej	-0.24	0.13	60.2	-1.80	0.2820		-0.16
By PoA	bil	unclear - Pvd	-0.07	0.12	847.5	-0.57	0.9411		-0.04
By PoA	alv	Ej - Pvl	0.54	0.16	24.8	3.51	0.0089	**	0.42
By PoA	alv	Pvd - Pvl	0.24	0.10	31.4	2.40	0.0981	.	0.16
By PoA	alv	Pvd - Ej	-0.30	0.11	29.4	-2.73	0.0496	*	-0.26
By PoA	alv	unclear - Pvl	0.24	0.09	27.1	2.62	0.0636	.	0.15
By PoA	alv	unclear - Ej	-0.31	0.11	26.4	-2.81	0.0433	*	-0.27
By PoA	alv	unclear - Pvd	-0.01	0.05	34.3	-0.12	0.9994		0.00
By PoA	vel	Ej - Pvl	1.19	0.17	34.2	7.09	<.001	***	0.97
By PoA	vel	Pvd - Pvl	0.86	0.14	123.3	6.01	<.001	***	0.58
By PoA	vel	Pvd - Ej	-0.33	0.14	66.6	-2.44	0.0798	.	-0.39
By PoA	vel	unclear - Pvl	0.85	0.12	85.5	7.05	<.001	***	0.57
By PoA	vel	unclear - Ej	-0.34	0.12	39.5	-2.82	0.0360	*	-0.40
By PoA	vel	unclear - Pvd	-0.01	0.09	394.7	-0.13	0.9992		-0.01

The following tables show the full post hoc pairwise comparisons from the linear mixed-effects model described in Section 4.5.3, for *delta hyperph* (Table G.36 and Table G.37).

Table G.44 Pairwise comparisons of delta hyperph (normalized) by stop type.

Contrast type	Stop type	Contrast	Estimate (norm.)	SE	DF	t	p	Signif.
By stop type	Pvl	bil - alv	0.06	0.02	266.1	3.39	0.0023	**
By stop type	Pvl	vel - alv	-0.07	0.02	270.9	-4.68	<.001	***
By stop type	Pvl	vel - bil	-0.13	0.02	273.7	-6.64	<.001	***
By stop type	Ej	bil - alv	0.11	0.02	261.3	6.41	<.001	***
By stop type	Ej	vel - alv	-0.01	0.01	206.2	-0.48	0.8802	
By stop type	Ej	vel - bil	-0.12	0.02	227.0	-6.15	<.001	***
By stop type	Pvd	bil - alv	0.11	0.02	289.4	5.85	<.001	***
By stop type	Pvd	vel - alv	-0.01	0.02	253.4	-0.78	0.7151	
By stop type	Pvd	vel - bil	-0.12	0.02	303.1	-5.59	<.001	***
By stop type	unclear	bil - alv	0.10	0.01	154.6	6.60	<.001	***
By stop type	unclear	vel - alv	-0.03	0.01	164.0	-2.22	0.0706	
By stop type	unclear	vel - bil	-0.13	0.02	166.0	-7.37	<.001	***

Table G.45 Pairwise comparisons of delta hyperph (normalized) by PoA.

Contrast type	PoA	Contrast	Estimate (norm.)	SE	DF	t	p	Signif.
By PoA	bil	Ej - Pvl	0.09	0.02	131.8	5.62	<.001	***
By PoA	bil	Pvd - Pvl	0.04	0.02	412.3	2.33	0.0939	
By PoA	bil	Pvd - Ej	-0.04	0.02	492.2	-2.69	0.0374	*
By PoA	bil	unclear - Pvl	0.06	0.01	91.0	4.38	<.001	***
By PoA	bil	unclear - Ej	-0.02	0.01	305.6	-2.06	0.1686	
By PoA	bil	unclear - Pvd	0.02	0.02	279.3	1.36	0.5234	
By PoA	alv	Ej - Pvl	0.03	0.01	30.3	3.21	0.0158	*
By PoA	alv	Pvd - Pvl	-0.01	0.01	40.8	-0.59	0.9357	
By PoA	alv	Pvd - Ej	-0.04	0.01	41.8	-4.47	<.001	***
By PoA	alv	unclear - Pvl	0.02	0.01	28.1	2.15	0.1632	
By PoA	alv	unclear - Ej	-0.01	0.01	36.8	-1.65	0.3628	
By PoA	alv	unclear - Pvd	0.03	0.01	26.7	3.39	0.0110	*
By PoA	vel	Ej - Pvl	0.10	0.01	75.2	7.78	<.001	***
By PoA	vel	Pvd - Pvl	0.06	0.02	229.9	3.72	0.0014	**
By PoA	vel	Pvd - Ej	-0.05	0.01	221.3	-3.54	0.0027	**
By PoA	vel	unclear - Pvl	0.07	0.01	86.3	4.77	<.001	***
By PoA	vel	unclear - Ej	-0.04	0.01	144.0	-3.75	0.0014	**
By PoA	vel	unclear - Pvd	0.01	0.01	129.2	0.79	0.8600	

Acknowledgements

Thank you:

To Bo, Cecilia, Eric, Julia and Alma. No words are truer – home is where family is.

To Miri. For every “You got this!”

To Phil. For the many contributions that brought this dissertation into being, starting with a P3 lecture on non-pulmonic sounds that I still remember, some eight years later. You were introducing ejectives – to me at the time, impressively exotic sounds – when you mentioned with an almost roguish smile their presence in British English. My interest was piqued!!

Finally, to Marianne. For eight years of reasoned guidance and steady molding of my often naïve, instinct-driven approaches to “sciencing”. The things I have and still can learn from you are endless. Thank you for getting me to this point.