The Development of Superheavy Syllables in Jordanian Child Speech: Acoustic and Computational Analyses

By

Bara’ah Al-Duneibat

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Central Lancashire

September 2024

RESEARCH STUDENT DECLARATION FORM

Type of Award	**Doctor of Philosophy**
School	**School of Psychology and Humanities**

1. **Concurrent registration for two or more academic awards**

   I declare that while registered as a candidate for the research degree, I have not been a registered candidate or enrolled student for another award of the University or other academic or professional institution

2. **Material submitted for another award**

   I declare that no material contained in the thesis has been used in any other submission for an academic award and is solely my own work

3. **Collaboration**

   There is no collaboration with any other parties for the completion of my PhD thesis.

4. **Use of a Proof-reader**

No proof-reading service was used in the compilation of this thesis.

Signature of Candidate

Print name: Bara’ah Al-Duneibat

Abstract

The study investigates syllable structure development in Jordanian Arabic (JA) child speech, focusing on superheavy syllables, which are often avoided in the literature due to their complexity. JA provides unique typological features for phonological analysis but remains under-researched compared to other Arabic dialects and West Germanic languages. The effects of age group, lexical stress, syllable structure, and syllable position on syllable and vowel durations were analysed. JA durational patterns contribute to the understanding of normal phonological development, typological variation, and stress assignment in Arabic dialectical phonology.

Twenty-one children aged between 24 and 72 months old and four adults were recruited. Adults carried out semi-spontaneous speech and repetition tasks whereas children completed an additional picture-elicitation task. For data analysis, a quantitative approach was employed, which involved acoustic analysis and Bayesian modelling of raw durations. An analysis of type and frequency of phonological processes was reported demonstrating an aspect of the normal developmental trajectory.

Results revealed that durations decreased with maturation, reflecting improved articulatory control, although the oldest child group did not match adult-like patterns. Lexical stress was not a strong predictor for duration as stress is determined by syllable weight and position. However, consistent word-final lengthening was observed. Syllable structure was a key predictor for durations, with superheavy syllables being longer than heavy and light syllables. Vowel shortening in non-final superheavy syllables was evident, rendering them bimoraic and not trimoraic. With maturation, superheavy syllables increased in frequency and complexity, demonstrating increased prosodic variability. Phonological processes were more frequent in younger age groups, particularly in stressed, final, and superheavy syllables.

The findings emphasize the role of syllable development in informing practical implications for speech-language pathology. Superheavy syllable durational patterns may serve as early indicators of speech production delays and deficits, aiding in the diagnosis and intervention of speech disorders in JA children.

Table of Contents

Abstract ii

Table of Contents iv

Acknowledgments ix

List of Figures x

List of Tables xix

List of Abbreviations and Symbols xxii

1. Chapter One: Introduction 1

    1.1. Overview 2

        1.1.1. Basic Units of Child Language: The Syllable 2

        1.1.2. Rhythm Typology and Domain-Related Durational Effects 5

    1.2. Scope of the Study 9

    1.3. Aims of the Study 11

    1.4. Research Questions 11

    1.5. Significance of the Study 12

    1.6. Structure of the Thesis 13

2. Chapter Two: Literature Review 15

    2.1. The Arabic Language in Jordan 16

        2.1.1. Arabic: An Overview 16

        2.1.2. Modern Standard Arabic 18

        2.1.3. Jordanian Arabic 20

            2.1.3.1. Background on JA 20

            2.1.3.2. Phonological Properties of JA 23

2.1.3.2.1.Syllable Structure 23

2.1.3.2.2.Stress and Metrical Structure 27

2.1.3.2.3.Geminates 33

2.1.3.2.4.Epenthesis 35

2.1.3.2.5.Syncope 38

2.1.4. Superheavy Syllables 40

2.2. Language Development 47

2.2.1. Phonological and Phonetic Development in Infancy 47

2.2.1.1. Perceptual Development in Infants 47

2.2.1.2. From Prelinguistic Vocalizations to Early Words 51

2.2.2. Language Universals and Articulatory Motor Learning Perspectives 54

2.2.2.1. Language Universals 54

2.2.2.2. Articulatory Learning 58

2.2.3. Phonological Development in Arabic 60

2.2.3.1. Phonological Development in Arabic Dialects 61

2.2.3.2. Phonological Development in JA 65

2.2.4. Lexical Stress in Child Speech 72

2.3. Summary and Predictions 81

2.3.1. Jordanian Arabic: A summary 81

2.3.2. Language Development: A summary 82

2.3.3. Lexical Stress: A summary 84

2.3.4. Predictions 85

3. Chapter Three: Methodology 89

3.1. Approach 90

3.1.1. Participants 92

3.2. Child Group 98
  3.2.1. Semi-Spontaneous Speech Task 98
  3.2.2. Picture Elicitation Task 99
  3.2.3. Repetition Task 100
3.3. Adult Group 101
  3.3.1. Semi-Spontaneous Speech Task 101
  3.3.2. Repetition Task 101
3.4. Data Analysis 101
  3.4.1. Acoustic Analysis 101
    3.4.1.1. Word Identification and Text-Grids 101
    3.4.2. Acoustic Markers for Phoneme-Level Segmentation 105
      3.4.2.1. Fricatives 105
      3.4.2.2. Stops 106
      3.4.2.3. Nasals 107
      3.4.2.4. Affricates 108
      3.4.2.5. Approximants 109
      3.4.3.5. Vowels 110
    3.4.4. Segmentation complexities in JA productions 111
  3.4.2. Phonological Processes Analysis 114
  3.4.3. Statistical Analysis 123
    4.6.1. Bayesian Multi-Level Modelling 124
    4.6.2. Monotonic Predictors 125
    4.6.3. Priors 126
    4.6.4. Efficiency and Convergence 127
3.5. Limitations 127

4. Chapter Four: Results 129

4.1. Semi-Spontaneous Speech Task 131

4.1.1. Syllable Duration 132

4.1.2. Vowel Duration 139

4.1.3. Interaction between Syllable Structure and Syllable Position 145

4.1.4. Summary 150

4.2. Repetition Task 152

4.2.1. Syllable Duration 152

4.2.2. Vowel Duration 161

4.2.3. Summary 169

4.3. Picture Elicitation Task 170

4.3.1. Syllable Duration 170

4.3.2. Vowel Duration 177

4.3.3. Summary and Interim Discussion 185

4.4. Superheavy Syllables 188

4.4.1. Semi-Spontaneous Speech Task 188

4.4.1.1. Superheavy Syllable Frequency 189

4.4.1.2. Bayesian Models and Descriptive Statistics 207

4.4.2. Repetition Task 228

4.4.2.1. Superheavy Syllable Frequency 228

4.4.2.2. Bayesian Models and Descriptive Statistics 242

4.4.3. Picture Task 257

4.4.3.1. Superheavy Syllable Frequency 257

4.4.3.2. Bayesian Models and Descriptive Statistics 267

4.4.4. Summary 288

4.5. Phonological Processes 293

5. Chapter Five: Discussion 299

5.1. Age Group 300

5.2. Lexical Stress 307

5.3. Syllable Position 312

5.4. Syllable Structure and Moraicity 318

5.5. Superheavy Syllables 323

5.6 Summary and Theoretical Implications 332

5.7. Clinical Implications 336

5.8. Limitations and Directions for Future Research 336

References 338

Appendix: Repetition Task Word List 383

Acknowledgments

I owe the successful completion of this thesis to my exceptional supervisor, Dr. Hae-Sung Jeon, for your unwavering support, invaluable guidance, and dedication to my growth as a researcher. I am deeply grateful for the countless hours you spent providing feedback and pushing me beyond my expectations. I also extend my heartfelt thanks to my supervisory team, Dr. Daniel Bürkle and Dr. Nick Palfreyman for their insights and commitment, which made this journey truly enriching.

To my wonderful parents, Khalid and Manal, and my amazing sisters, Farah, Rama, and Hala, your love and support have been the driving force behind this achievement. Your unwavering belief in me, patience, and encouragement helped me overcome challenges and persevere through difficult times. I am deeply grateful for every sacrifice you made to help me succeed on this academic journey.

To my sister Aiswarya and her family, I am deeply grateful for your words of wisdom and unwavering support, especially during challenging times. The countless cups of coffee and your belief in me made this journey much easier. To my best friend Jehan, thank you for your encouragement, heartfelt conversations, and shared laughter, which provided much-needed inspiration, and for celebrating every milestone as if it were your own.

List of Figures

Figure 2.1: Map of Arabic varieties 17

Figure 2.2: MSA vowels 20

Figure 2.3: A coda clusters with rising sonority 25

Figure 2.4: A word-final superheavy syllable /malaak/ ‘angel’ 27

Figure 2.5:A word non-final superheavy syllable /daarsiin/ ‘they studied’ 27

Figure 2.6: Examples of a S-w trochaic pattern 30

Figure 2.7: Gemination of a monomoraic /sad*/ becoming bimoraic /sadd/ 31

Figure 2.8: CVC syllable in final and non-final positions 33

Figure 2.9: Skeletal tiers for CVCVC and CVCCVC templates 34

Figure 2.10: Moraic representations of /fataħ/ and /fattaħ/ 34

Figure 2.11: Vowel epenthesis in a geminate coda 37

Figure 2.12: Coda consonant in a superheavy syllable 43

Figure 2.13: Stress assignment to the visible foot by ERR 44

Figure 2.14: Stress assignment in a suffixation example 45

Figure 2.15: Phonological components of SPE 56

Figure 3.1: Textgrid demonstrating the tiers used for data acoustic analysis 103

Figure 3.2: Textgrid demonstrating the difference in syllable and production tiers 104

Figure 3.3: Segmentation of a fricative in /farawlah/ 106

Figure 3.4: Segmentation of a stop in /samat/ 107

Figure 3.5: Segmentation of a nasal in /muuz/ 108

Figure 3.6: Segmentation of an affricate in /nad͡ʒmah/ 109

Figure 3.7: Segmentation of an approximant in /murabbaʕ/ 110

Figure 3.8: Segmentation of a vowel in /ʔaħmar/ 111

Figure 3.9: (A) A detailed waveform (b) Spectrogram of a liquid-vowel segmentation 112

Figure 3.10: (A) A detailed waveform (b) Spectrogram of a glide-vowel segmentation 113

Figure 3.11: (A) A detailed waveform (b) Spectrogram of a trill-vowel segmentation 114

Figure 3.12: Segment modification- metathesis 115

Figure 3.13: Segment modification- lateralization 116

Figure 3.14: Segment modification- stopping 116

Figure 3.15: Segment modification- assimilation 117

Figure 3.16: Vowel epenthesis 118

Figure 3.17: Assimilation and gemination 119

Figure 3.18: Syllable deletion 120

Figure 3.19: Cluster reduction 120

Figure 3.20: Weak syllable deletion 121

Figure 3.21: Coda deletion 122

Figure 3.22: Syncope 123

Figure 4.1: The Distribution and mean of the syllable duration (ms) by age group 134

Figure 4.2: (A) The distribution and the mean syllable duration (ms) for stress (b) Posterior predictive plot for syllable duration for stress 135

Figure 4.3:The Distribution and the mean syllable duration (ms) for syllable structure 136

Figure 4.4: The distribution and the mean syllable duration (ms) for syllable position 137

Figure 4.5: Posterior predictive plot for syllable duration for the interaction between age group, stress assignment, and syllable position 138

Figure 4.6: (A) The distribution and the mean vowel duration (ms) for age group (b) posterior predictive plot for vowel duration for age group 141

Figure 4.7: The distribution and the mean vowel duration (ms) for stress assignment 142

Figure 4.8: The distribution and the mean vowel duration (ms) for syllable structure 143

Figure 4.9: The distribution and the mean vowel duration (ms) for syllable position 143

Figure 4.10: Posterior predictive plot for vowel duration for the interaction between age group, stress assignment, and syllable position 145
Figure 4.11: Posterior predictive plots for syllable duration for syllable structure and syllable position and the three-way interaction 147
Figure 4.12: Posterior predictive plots for vowel duration for syllable structure and syllable position and the three-way interaction 149
Figure 4.13: The distribution and the mean syllable duration (ms) by age group 155
Figure 4.14: (A) The distribution and the mean syllable duration (ms) for stress (B) Posterior predictive plot for syllable duration for stress 156
Figure 4.15: The distribution and the mean syllable duration (ms) for syllable position 157
Figure 4.16: The distribution and the mean syllable duration (ms) for syllable structure 158
Figure 4.17: (A) Posterior predictive plot for syllable duration for the interaction between age group, stress assignment, and syllable position (B) Posterior predictive plot for syllable duration for the interaction between age group and syllable position 160
Figure 4.18: The distribution and the mean vowel duration (ms) by age group 163
Figure 4.19: (A) The distribution and the mean vowel duration (ms) for stress (B) Posterior predictive plot for vowel duration for stress 164
Figure 4.20: The distribution and the mean vowel duration (ms) for syllable position 165
Figure 4.21: The distribution and the mean vowel duration (ms) for syllable structure 166
Figure 4.22: (A) Posterior Predictive plot for vowel duration for the interaction between age group, stress assignment, and syllable position (B) Posterior predictive plot for vowel duration for the interaction between age group and syllable position ... 168
Figure 4.23: The distribution and the mean syllable duration (ms) for age group 173
Figure 4.24: (A) The distribution and the mean syllable duration (ms) for stress (B) Posterior predictive plot for syllable duration for stress 174

Figure 4.25: The distribution and the mean syllable duration (ms) for syllable position......175

Figure 4.26: The distribution and the mean syllable duration (ms) for syllable structure.....176

Figure 4.27: Posterior predictive plot for syllable duration for the interaction between age group, stress assignment, and syllable position...177

Figure 4.28: The distribution and the mean vowel duration (ms) by age group...................180

Figure 4.29: (A) The distribution and the mean vowel duration (ms) for stress (B) Posterior predictive plot for vowel duration for stress 181

Figure 4.30: The distribution and the mean vowel duration (ms) for syllable position.........182

Figure 4.31: The distribution and the mean vowel duration (ms) for syllable structure........183

Figure 4.32: Posterior predictive plot for vowel duration for the interaction between age group, stress assignment, and syllable position...184

Figure 4.33: The frequency and the percentage of superheavy syllables in st by age group.191

Figure 4.34: Spectrograms of superheavy syllables in age group 24–30..............................194

Figure 4.35: Spectrograms of superheavy syllables in age group 31–37..............................196

Figure 4.36: Spectograms of superheavy syllables in age group 38–44...............................198

Figure 4.37: Spectrograms of superheavy syllables in age group 45–51..............................199

Figure 4.38: Spectrogram of a superheavy syllable in age group 52–58...............................200

Figure 4.39: Spectrograms of superheavy syllables in age group 59–65..............................202

Figure 4.40: Spectograms of superheavy syllables in age group 66–72...............................204

Figure 4.41: Spectrograms of superheavy syllables in the adult group 206

Figure 4.42: Frequency and percentage of superheavy syllables in monosyllabic words 208

Figure 4.43: The distribution and the mean syllable duration (ms) for age group in monosyllabic words 209

Figure 4.44: The distribution and the mean syllable duration (ms) for sub-structure in monosyllabic words...210

Figure 4.45: Posterior predictive plot for syllable duration of the interaction between sub-structure age group in monosyllabic words..211
Figure 4.46: The distribution and the mean vowel duration (ms) for age group in monosyllabic words 212
Figure 4.47: The distribution and the mean vowel duration (ms) for sub-structure in monosyllabic words..213
Figure 4.48: Posterior predictive plot for vowel duration of the interaction between sub-structure age group in monosyllabic words..214
Figure 4.49: Frequency and percentage of superheavy syllables in disyllabic words 215
Figure 4.50: Frequency and percentage of superheavy syllables in multisyllabic words......216
Figure 4.51: The distributions and the mean syllable durations (ms) for age group in disyllabic and multisyllabic words 218
Figure 4.52: The distributions and the mean syllable durations (ms) for sub-structure in disyllabic and multisyllabic words...219
Figure 4.53: The distributions and the mean syllable durations (ms) for stress in disyllabic and multisyllabic words 221
Figure 4.54: The distributions and the mean syllable durations (ms) for syllable position in disyllabic and multisyllabic words...222
Figure 4.55: The distributions and the mean vowel durations (ms) for age group in disyllabic and multisyllabic words 224
Figure 4.56: The distributions and the mean vowel durations (ms) for sub-structure in disyllabic and multisyllabic words 225
Figure 4.57: The distributions and the mean vowel durations (ms) for stress in disyllabic and multisyllabic words 226

Figure 4.58: The distributions and the mean vowel durations (ms) for syllable position in disyllabic and multisyllabic words 228
Figure 4.59: Frequency and percentage of superheavy syllables in rt 229
Figure 4.60: Spectrogram of a superheavy syllable age group 24–30 232
Figure 4.61:Spectrogram of a superheavy syllable in age group 31–37 233
Figure 4.62: Spectrograms of superheavy syllables in age group 38–44 234
Figure 4.63: Spectrogram of A Superheavy Syllable In Age Group 45–51 235
figure 4.64: Spectrograms of superheavy syllables in age group 52–58 237
Figure 4.65: Spectrograms of superheavy syllables in age group 59–65 239
Figure 4.66: Spectrogram of a superheavy syllable in age group 66–72 240
Figure 4.67: Spectrogram of a superheavy syllable in the adult group 241
Figure 4.68: Frequency and percentage of superheavy syllables in monosyllabic words 243
Figure 4.69: The distribution and the mean syllable duration (ms) for age group in monosyllabic words 244
Figure 4.70: The distributions and the mean syllable and vowel durations (ms) for sub-structure in monosyllabic words 245
Figure 4.71: Posterior predictive plot for vowel durations across the age groups in monosyllabic words 246
Figure 4.72: Posterior predictive plot for vowel duration for sub-structure in monosyllabic words 247
Figure 4.73: Frequency and percentage of superheavy syllables in disyllabic words 248
Figure 4.74: Frequency of superheavy syllables in multisyllabic words 249
Figure 4.75: The distributions and the mean syllable durations (ms) for age group in disyllabic and multisyllabic words 251

Figure 4.76: The distributions and the mean syllable duration (ms) for sub-structure in disyllabic words 252

Figure 4.77: The distributions and mean syllable duration (ms) for stress in disyllabic words 253

Figure 4.78: The distributions and the mean vowel durations (ms) for age group in disyllabic and multisyllabic words 255

Figure 4.79: The distributions and the mean vowel duration (ms) for sub-structure in disyllabic words 255

Figure 4.80: The distributions and the mean vowel duration (ms) for stress in disyllabic words 256

Figure 4.81: Frequency and percentage of superheavy syllables in PT 258

Figure 4.82: Spectrogram of a superheavy syllable in age group 24–30 261

Figure 4.83: Spectrograms of superheavy syllables in age group 31–37 262

Figure 4.84: Spectrogram of a superheavy syllable in age group 38–44 263

Figure 4.85: Spectrogram of a superheavy syllable in age group 52–58 265

Figure 4.86: Spectrogram of a superheavy syllable in age group 66–72 266

Figure 4.87: Frequency and percentage of superheavy syllables in monosyllabic words 268

Figure 4.88: The distribution and the mean syllable duration (ms) for age group in monosyllabic words 270

Figure 4.89: The distributions and the mean syllable durations (ms) for sub-structure in monosyllabic words 271

Figure 4.90: Posterior predictive plot for syllable duration for the interaction between age group and sub-structure in monosyllabic words 272

Figure 4.91: The distribution and the mean vowel duration (ms) for age group in monosyllabic words 273

Figure 4.92:The distribution and the mean vowel duration (ms) for sub-structure in monosyllabic words......274

Figure 4.93: Posterior predictive plot for vowel durations for the interaction between age group and sub-structure monosyllabic words......275

Figure 4.94: Frequency and percentage of superheavy syllables in disyllabic words......276

Figure 4.95: Frequency of superheavy syllables in multisyllabic words......277

Figure 4.96: The distributions and the mean syllable durations (ms) for age group in disyllabic and multisyllabic words......279

Figure 4.97: The distribution and the mean syllable duration (ms) for sub-structure in disyllabic words......280

Figure 4.98: The distributions and the mean syllable durations (ms) for stress in disyllabic and multisyllabic words......281

Figure 4.99: The distributions and the mean syllable durations (ms) for syllable position in disyllabic and multisyllabic words......282

Figure 4.100: The distributions and the mean vowel durations (ms) for age group in disyllabic and multisyllabic words......284

Figure 4.101: The distributions and the mean vowel duration (ms) for sub-structure in disyllabic words......285

Figure 4.102: The distributions and the mean vowel durations (ms) for stress in disyllabic and multisyllabic words......286

Figure 4.103: The distributions and the mean vowel durations (ms) for syllable position in disyllabic and multisyllabic words......287

Figure 4.104: Number of processes – Syllable Structure......295

Figure 4.105: Number of processes – Lexical Stress......296

Figure 4.106: Number of processes – Syllable Position......297

Figure 4.107: Number of processes – Interaction between stress, position and age group ...297

Figure 5.1: Spectrograms of the production form /taab/ and the target form /ktaab/ 304

Figure 5.2: Spectrograms of the production form /ʔiɣziir/ and the target form /zɣiir/ 306

Figure 5.3: Posterior predictive plot for syllable duration of the interaction between stress and age group in ST 311

Figure 5.4: Posterior predictive plot for syllable duration for syllable position across the age groups in ST 314

Figure 5.5: Posterior predictive plot for syllable duration for the three-way interaction 317

Figure 5.6: Metrical trees for superheavy syllable representation in word non-final positions 322

Figure 5.7: Spectrograms of the words /kik.kiin/ and /sˤub.buun/ 328

List of Tables

Table 2.1: MSA consonants 19

Table 2.2: AA consonantal inventory 22

Table 2.3: Superheavy syllables emerging due to syncopation 40

Table 3.1: Participants 96

Table 4.1: Frequency of syllable structures across the age groups 131

Table 4.2: Frequency of syllable structures in ST 131

Table 4.3: Summary of the mean and standard deviation for syllable durations for syllable structure lexical stress and syllable position for each age group in ST (N=1440)132

Table 4.4: Bayesian model output summary for syllable duration/ST 133

Table 4.5: Summary of the mean and standard deviation for vowel durations for syllable structure lexical stress and syllable position for each age group in ST (N =1440)139

Table 4.6: Bayesian model output summary for vowel duration/ST 140

Table 4.7: Bayesian model output summary for syllable duration in ST interaction 146

Table 4.8: Bayesian model output summary for vowel duration in ST interaction 148

Table 4.9: Frequency of syllable structures in RT 152

Table 4.10: Summary of mean and standard deviation for syllable durations for syllable structure lexical stress and syllable position for each age group in RT (N = 613).152

Table 4.11: Bayesian model output summary for syllable durations/RT 154

Table 4.12: Summary of mean and standard deviation for vowel durations for syllable structure lexical stress and syllable position for each age group in RT (N=613) 161

Table 4.13: Bayesian model output summary for vowel duration/RT 162

Table 4.14: Frequency of syllable structures in PT 170

Table 4.15: Summary of the mean and standard deviation for syllable durations for syllable structure lexical stress and syllable position for each age group in PT (N=1282). 170

Table 4.16: Bayesian Model Output Summary For Syllable Duration-PT............................172

Table 4.17: Summary of the mean and standard deviation for vowel durations for syllable structure lexical stress and syllable position for each age group in PT (N=1282). 177

Table 4.18: Bayesian model output summary for vowel duration/PT..................................178

Table 4.19: Summary of bayesian model outputs across tasks (S: Strong, NS: Not Strong) 186

Table 4.20:Superheavy syllables ST...191

Table 4.21: Bayesian model output summary for syllable duration in monosyllabic words 209

Table 4.22: Bayesian model output summary for vowel duration in monosyllabic words ...211

Table 4.23: Bayesian model output summary for syllable duration in di/multi-syllabic words 217

Table 4.24: Bayesian model output summary for vowel duration in di/multisyllabic words 223

Table 4.25: Superheavy syllables RT 230

Table 4.26: Bayesian model output for superheavy syllables in monosyllabic words/ syllable durations..244

Table 4.27: Bayesian model output summary vowel duration in monosyllabic words.........246

Table 4.28: Bayesian model output summary for syllable duration in di/multisyllabic words 249

Table 4.29: Bayesian model output summary for vowel duration in di/multisyllabic words 253

Table 4.30: Superheavy syllables PT...258

Table 4.31: Bayesian model output summary for syllable duration in monosyllabic words.269

Table 4.32: Bayesian model output summary for vowel duration in monosyllabic words ...272

Table 4.33: Bayesian model output summary for syllable duration in di/multisyllabic words 277

Table 4.34: Bayesian model output summary for vowel duration in di/multisyllabic words 283

Table 4.35: Summary of superheavy syllable productions across tasks (Mono=Monosyllabic; Di=Disyllabic; Multi=Multisyllabic). 288

Table 4.36: Summary of bayesian model outputs in superheavy syllable productions (S: Strong, NS: Not Strong) 292

Table 4.37: Phonological processes in JA child speech 293

List of Abbreviations and Symbols

σ	Syllable
μ	Mora
AA	Ammani Arabic
Bulk-ESS	Bulk Effective Sample Size
CA	Classical Arabic
CI	Credible Interval
ERR	End Right Rule
ST	Semi-Spontaneous Speech Task
IPA	International Phonetic Alphabet
JA	Jordanian Arabic
l-95% CI	Lower limit of the 95% credible interval
MCMC	Markov Chain Monte Carlo
MSA	Modern Standard Arabic
OT	Optimality Theory
PT	Picture Elicitation Task
R-hat	Potential Scale Reduction Factor
RT	Repetition Task
SD	Standard Deviation
SSP	Sonority Sequencing principle
Tail-ESS	Tail Effective Sample Size
u-95% CI	Upper limit of the 95% credible interval
WPR	Weight by Position Rule

1. Chapter One: Introduction

Chapter One: Introduction

The current study investigates the development of syllable structure in JA child speech, focusing on the development of superheavy syllables. The study highlights durational patterns influenced by predictors, such as age group, lexical stress, syllable position within a word, and syllable structure. This chapter provides an overview of syllable role in child development and analytical approaches to speech timing. The study's scope, aims, questions, significance, and overall structure are also provided.

1.1. Overview

1.1.1. Basic Units of Child Language: The Syllable

In language development research, what constitutes the basic unit of child speech production has received much debate. The debate centers on whether children primarily acquire language through larger or holistic units such as words (Ferguson & Slobin 1973, Ferguson & Farewell 1975, Menn 1978), or smaller units such as moras, syllables, and phonemes (Zharkova 2004, Rojczyk & Porzuczek 2012). The current study highlights the syllable as the primary domain of analysis due to it being shaped by the phonotactic properties of each language, serving as a domain of prosodic phenomena such as stress assignment.

The syllable has been widely recognized as a fundamental speech constituent for multiple reasons. Research on perception and auditory processing extensively explored the role of the syllable in the acquisition and organization of phonological knowledge, providing strong evidence for its primacy (Savin & Bever 1970, Warren 1971, Massaro 1972, Studdert-Kennedy 1976). For example, Studdert-Kennedy (1976) suggested that speech is segmented into syllables, serving as the basis of phonetic recognition and providing the information necessary for phoneme recognition. Massaro (1972) showed that perceptual processing, where an examination of the physical features of the stored sequential pattern to identify input, cannot be at the phoneme level and must be more complicated. However, it also cannot be as large as two or three words since larger units would be complex to process efficiently. Savin & Bever (1970) found that a speaker's recognition was faster and more efficient when a target was a complete syllable compared to a target that was a phoneme from a syllable. Such a finding was further supported by Warren (1971), whose results showed the identification of monosyllabic words and nonsense syllables was faster than the identification of phoneme clusters. This suggests that syllables are processed as whole units, facilitating faster recognition compared to phonemes, which require extraction or inference from within syllables.

The syllable role in children's speech perception has been documented in the literature (Bertoncini and Mehler 1981, Mehler et al., 1988, Cutler and Butterfield 1992, Jusczyk et al., 1993, Kuhl 1994, and Saffran et al., 1996). Children demonstrated the tendency to rely on the syllable for speech perception, aiding in the acquisition of their ambient language(s). For example, Jusczyk et al. (1993) found that infants as young as seven months can segment continuous speech into syllabic units, indicating an early sensitivity to the syllable structure. Additionally, Mehler et al. (1988) suggested that infants can distinguish between different languages based on rhythmic properties, parsing speech input in syllabic patterns. The early ability to discriminate languages by their syllable-based rhythm further evidences that syllables play a crucial role in auditory processing and language acquisition. Kuhl's (1994) analysis demonstrated that infants have a heightened sensitivity to the phonetic properties of syllables, aiding in the categorization of speech sounds that fosters the development of phonetic awareness and language acquisition.

Syllable primacy is evident in the developmental progression of syllable complexity in child speech (Oller 1980, Fikkert 1994, Demuth 1996, Lleó & Prinz 1996, Vihman & Velleman 2000). The progression from simple to complex syllables reflects a developmental trajectory in phonological acquisition (2.2.2.1). Demuth (1996) emphasized that in early language acquisition, children’s initial word productions are organized around basic syllable structures such as CV and CVC. Moreover, the mastery of syllable structures serves as a critical milestone in the phonological development of a child's speech. Vihman & Velleman (2000) observed that children’s early lexical development is closely related to their mastery of syllable structures. Children who produce more structures tend to produce a larger set of vocabulary with more advanced phonological skills.

Furthermore, the production of syllables is influenced by the phonotactic constraints of the target language (Ingram 1989, Levelt, Schiller & Levelt 2000, Goad & Brannen 2003, Kehoe & Lleó 2003). Phonotactic rules govern permissible combinations of sounds within a language, shaping the structure and complexity of syllables that children produce. Children acquiring different languages develop syllable structures that conform to the phonotactic rules of each language (Goad & Brannen 2003). In examining bilingual children acquiring German and Spanish, Kehoe & Lleó (2003) showed that the child’s syllable adhered to the specific phonotactic rules of each language. German children produced complex consonant clusters following the language’s phonotactic norms whereas Spanish children produced simpler syllable structures. Similarly, Ingram (1989) discussed the differences in syllable structure preferences based on phonotactic rules in English and French-speaking children. English-speaking children tended to produce a higher number of complex syllables containing consonant clusters while French-speaking children predominantly produced simpler syllables. Levelt, Schiller & Levelt (2000) further explored the influence of language-specific constraints on the acquisition of Dutch syllable structure. Children’s early productions seemed to mirror the permissible syllabic structures in Dutch, demonstrating a phonological awareness of the phonotactic rules from an early age (e.g., Dutch children did not produce complex marginal syllables such as VCC, CCVC, or CCVCC before producing simpler structures such as CVCC or CCV).

Finally, the syllable is considered the basic unit of the prosodic hierarchy (Selkirk 1982, Nespor & Vogel 1986, Goldsmith 1990, Demuth 1996, Kehoe 2001). The main levels of the prosodic hierarchy include the prosodic word (PW), foot (Ft), syllable (σ), and mora (μ), with each higher level being composed of units from the level directly below it (Fikkert 1994). Selkirk (1982) and Goldsmith (1990) suggested that syllables have a central role in organizing phonological representations, serving as the unit that encodes and influences the phonological information and rule application (Ito 1986, Nespor & Vogel 1986, Goldsmith 1990). For example, Demuth (1996) suggested that syllable structure influences the application of phonological processes, with certain rules applied more frequently to specific syllable types (e.g., weak syllable deletion). Kehoe (2001) emphasized that the interaction between the syllable role and the prosodic hierarchy is reflected in the phonological and prosodic developmental domains. Children employ syllable structures to demonstrate early sensitivity to syllable boundaries in parsing and producing speech (Inkelas & Zec 1995). Prieto (2006) highlighted that the acquisition of stress patterns and intonation contours is linked to their mastery of syllable structure, further emphasizing the foundational role of syllables in prosodic hierarchical development. Therefore, the syllable's role in the prosodic hierarchy is central to the organization and structure of speech, influencing phonological representations, application of phonological rules, and development of prosodic patterns.

1.1.2. Rhythm Typology and Domain-Related Durational Effects

The traditional dichotomy between stress-timed and syllable-timed languages has long been a topic of discussion among linguists. Stress-timed languages, such as English, Russian, and Arabic are characterized by having relatively equal intervals of time between stressed syllables regardless of the number of intervening syllables (Pike 1946, Abercrombie 1967, Roach 1982, Heliel 1982). Syllable-timed languages, such as French, Telugu, and Yoruba have syllables that occur at equal intervals regardless of stress (Pike 1946, Abercrombie 1967, Roach 1982). A third classification of languages is the mora-timed languages, such as Japanese, where timing is based on moras rather than syllables or stress (Port, Dalby & O'Dell 1987, Ota 1999).

However, this classification system has increasingly been questioned for its capacity to accurately represent the rhythmic patterns of languages. Mitchell (1969) argued that no language is purely syllable-timed or stress-timed; instead, languages exhibit varying degrees of both types of timing, with one typically predominating. Roach (1982) further critiqued this dichotomy, particularly challenging the idea that syllable-timed languages maintain equal syllable lengths. Roach noted that syllable-timed languages with phonemically long and short vowels would still show syllable length variation. This variability, found across both stress-timed and syllable-timed languages, suggests that the traditional classification may oversimplify the complexity of speech rhythms.

Some studies provided support for the rhythm typology using 'rhythm metrics' to categorize languages such as %V (the proportion of vocalic intervals), ΔC (the standard deviation of consonantal intervals), and PVIs (Pairwise Variability Index, measuring the variation in duration between successive vowels or consonants) (Nespor & Vogel 1989, Ramus et al., 1999, Grabe & Low 2002, Lee & McAngus Todd 2004, Rouas et al., 2005). Nevertheless, the 'rhythm metrics' reliability and their interpretations have been questioned (Arvaniti 2009). However, these metrics have shown inconsistencies and limited success, particularly for languages that do not fit into prototypical categories (i.e., languages that do not exhibit clear characteristics of either stress-timed or syllable-timed rhythms). Arvaniti suggested that rhythm in all languages should be understood through universal principles of grouping (how speech segments are organized into larger units) and prominence (how elements are highlighted relative to others). This perspective denotes that different languages employ various cues to indicate prominence, resulting in different perceptions of rhythm. For example, English listeners rely on stress patterns whereas Japanese listeners may focus more on pitch and moraic timing (Beckman

1986, de Jong 1994, Ota 2001). Phonetic studies have shown that isochrony (i.e., regularity in timing) is not consistently observed in actual speech across these categories (Nespor, Shukla & Mehler 2011). For instance, the duration of interstress intervals in stress-timed languages varies according to the number of syllables they contain while syllable-timed languages exhibit variation in syllable duration depending on the number of segments (Peterson 1962, O'Connor 1965, Lea 1974, Borzone de Manrique & Signorini 1983).

In Arabic, which is often classified as a stress-timed language (Miller 1984), the rhythmic pattern does not exhibit the same level of regularity as English, largely due to the complexity of its syllable structure (Milelr 1984, Bertinetto 1989). Arabic features a range of syllable structures which include simple syllables, such as light CV or CCV, and more complex structures, such as superheavy CCVCC and CVVC syllables (Section 2.1.3.2.1). This aligns with Abercrombie's (1967) observation that stress-timed languages tend to have a greater variety of syllable structures, which results in variability in syllable length. Syllable structure plays a significant role in stress-timed languages as it interacts with lexical stress, reinforcing each other; for instance, heavy syllables are typically stressed, and light ones are unstressed (Dauer 1983). Nonetheless, experimental studies argued that although Arabic is classified as a stress-timed language, it shows a lesser degree of durational contrast between stressed and unstressed syllables (i.e., particularly due to stress assignment being determined by syllable weight and position within a word) compared to languages such as English (Heliel 1982, Roach 1982, Bertinetto 1989, Zawaydeh, Tajima & Kitahara 2002, Watson 2002).

Another approach for analysing speech timing concerns how durational variation occurs in relation to prosodic structure (Wightman et al., 1992, Arvaniti 2009, Nolan & Jeon 2014). Lengthening effects are often observed at both the heads and edges of prosodic domains, expanding the understanding of how timing is influenced by linguistic organization rather than arbitrary rhythmic categories (Arvaniti 2009). Domain-head lengthening typically serves as a cue to prominence while domain-edge lengthening may signal constituent boundaries (Wightman et al., 1992, Turk & Shattuck-Hufnagel 2000). These prosodic mechanisms facilitate effective communication by structuring speech rhythmically, thereby enhancing the listener's ability to parse and interpret language (Nolan & Jeon 2014, White 2014).

Domain-edge effects refer to phonological processes' sensitivity to the boundaries of prosodic domains, such as syllables, feet, prosodic words, and intonational phrases (White and Turk 2010). An example of domain-edge effects is final lengthening, which involves the lengthening of vowels and consonants towards the end of prosodic domains (Nolan & Jeon 2014, Turk & Shattuck-Hufnagel 2000). This lengthening provides listeners with cues about the end of prosodic units, facilitating easier subsequent prosodic boundary predictions. The increased duration at the edges of words, phrases, or sentences signals to the listeners the finality of a prosodic unit, thus aiding in speech segmentation and comprehension (White and Turk 2010). The biomechanical constraints framework interprets speech timing as a product of the physical limitations and mechanical properties of the human articulatory system (Berkovits 1994, Cummins 1999, Tabain 2003). Accordingly, features such as final lengthening arise naturally from the speech apparatus behaviour (Berkovits 1994, Cummins 1999). Final lengthening is suggested to reflect a generalized and diffused deceleration of the articulatory system as the speaker approaches the end of a prosodic unit (Fowler 1990, and Tabain 2003). The physical limitations and mechanical properties of the human articulatory system naturally slow down articulatory movements at prosodic boundaries, resulting in longer durations for syllables and vowels (Berkovits 1994, Arvaniti 2009). However, the final lengthening is not solely determined by biomechanical constraints but also differs across languages, suggesting that it is bound to linguistic structure (Fletcher 2010).

Domain-head effects emphasize the importance of head positions within prosodic domains. The head typically refers to the most prominent or stressed syllable within a domain. Domain-head processes include the lengthening of lexically stressed syllables and words under sentence stress (Oller 1973, White 2014). These phonological processes enhance the prominence and perceptual distinctiveness of syllable onset, crucial for conveying prosodic information (Fry 1955, Saffran et al., 1996). In stress-timed languages such as English, the head of the prosodic word, the stressed syllable, exhibits distinct phonetic characteristics (Oller 1973, Price et al., 1991, White 2014). Stressed syllables are produced with longer durations and greater amplitudes compared to unstressed syllables, rendering the stressed as more prominent (Fry 1955, Klatt 1976, Beckman 1986). Nevertheless, quantifying the magnitude of lengthening due to lexical stress is often confounded by other factors, such as the influence of phrasal accent in natural discourse (Oller 1973, Ladd 1996).

1.2. Scope of the Study

This study investigates the development of syllable structure in JA child speech, particularly superheavy syllables. It analyses durational patterns based on predictors, such as age group, lexical stress, syllable structure, syllable position within a word, and their interactions. Durations can represent children's sensitivity to phonological aspects such as syllable weight and its interaction with stress assignment rules (Section 2.1.3.2.2). The emergence of syllables is crucial in child language acquisition and development, requiring both general phonetic mastery and adaptation to language-specific properties. These properties incorporate constraints on metrical structure and phonotactics, which vary typologically across languages and dialects (Section 2.2.3).

JA has distinctive phonological, morphological, and syntactic features from well-documented West Germanic languages (e.g., English, Dutch, and Spanish). Additionally, JA has been understudied compared to other Arabic dialects, such as Lebanese, Palestinian, Kuwaiti, Najdi, and Egyptian. JA's structural composition as a VC dialect is influenced by multiple socio-economic, geographical, and historical factors (2.1.3). Moreover, the exploration of Ammani

Arabic (AA), as the JA dialect under investigation, facilitates the analysis of Jordanian dialects mapping, which is poorly understood and not well classified.

JA exhibits a range of syllable structures (e.g., CV, CVV, CVVCC, CCVC), syllable weight categories (light, heavy, and superheavy), a lexical stress system that is rule-governed by a trochaic moraic system, phonological constraints on syllable formation (e.g., clusters adhering to the sonority sequencing principle), and phonological processes (e.g., epenthesis and syncope). The study examines how these language-specific constraints shape the development of syllable structure and superheavy syllable production in JA child speech.

The temporal aspects of JA children's speech are expected to be influenced by a combination of language-specific constraints and universal principles of acquisition and motor learning (Section 2.2.2). To explore this, the study examines key predictors influencing syllable and vowel durations in JA speech. First, analysing age group effect provides insights into durational trends determining when a child's productions begin to approximate adult patterns, which serves as a reference for typical development. Second, the study investigates prosodic factors such as lexical stress and final lengthening in children's phonological development. The study seeks to investigate how JA children assign stress, and whether they employ duration as an acoustic cue to mark stressed elements, which is a point of debate in the literature. Third, the study focuses on syllable structure, mainly superheavy syllables, as the complexity of syllable structure is often linked to the languages' classification and their rhythmic patterns. Superheavy syllables have been relatively overlooked in the literature due to their linguistic complexity, focusing mainly on their theoretical phonological accounts (e.g., interaction with stress, morphological markers and phonological processes) (Hayes 1995). Durational patterns associated with syllable structures may provide insights into the development of syllable complexity and phonological processing skills as children transition from language universals to more language-specific patterns. Thus, the study fills gaps in the literature by linking theoretical phonological accounts to empirical data and provides insights into the prosodic patterning and developmental strategies of JA child speech.

1.3. Aims of the Study

The study aims to explore syllable structure development, emphasizing superheavy syllables in JA child speech. It focuses on the progression of syllabic productions as a key aspect of child language, which requires both, general phonetic mastery and the mapping of language-specific constructs, including constraints on metrical structure and phonotactics. The study examines how predictors and their interactions, such as age group, lexical stress, syllable structure, and syllable position within a word influence syllable and vowel durations. The specific aims are as follows:

1. To trace the extent to which predictors, such as age group, lexical stress, syllable structure, and syllable position affect durations
2. To examine the developmental trajectory of the potential three-way interaction between age group, lexical stress, and syllable position on durational patterns, highlighting when durations start approximating adult-like patterns
3. To examine the frequency distribution and durational patterns of superheavy syllables across the age groups
4. To report the type and frequency of phonological processes evident in JA child speech

1.4. Research Questions

The overarching question is how the development of syllable production, mainly superheavy syllables, is shaped by predictors, such as age group, lexical stress, syllable structure, and syllable position within a word. Varieties of Arabic provide a rich context for phonological acquisition and development due to the typological variation in their metrical structure and phonotactic constraints. Given the limited assessment of the emergence and development of superheavy syllables in Arabic child speech, the study seeks to determine the frequency distribution patterns and durational properties of these syllables, and how predictors such as sub-structure and word length (in addition to the aforementioned predictors) influence durations.

Specific Questions

1. How do predictors, such as age group, lexical stress, syllable structure, and syllable position within a word affect durational patterns of JA child productions?
2. How is the developmental trajectory of the three-way interaction between age group, lexical stress, and syllable position depicted in JA child productions? Particularly, when do durations start approximating adult-like patterns?
3. What are the phonological and temporal properties of superheavy syllables in JA child speech?
4. What aspects of syllabic development, mainly superheavy syllables, are influenced by language-universal and language-specific properties in JA child speech?
5. What types and frequencies of phonological processes are evident in JA child productions?

1.5. Significance of the Study

This study examines the durational patterns of syllable structure production in JA child speech, particularly superheavy syllables, which have received limited consideration in the literature. Through addressing a gap in Arabic child speech research, the study seeks to bridge theoretical phonological frameworks with empirical data of durational properties specific to JA child speech. The study enhances the understanding of phonological development by emphasizing the fundamental role of the syllable in child language development. This allows for the exploration of the typological variation in syllable structure and its role in stress assignment between Arabic dialects, which is of broader theoretical significance within Arabic linguistics. The contribution to the deeper understanding of how syllable-level constraints and stress patterns differ across languages and dialects draws the connection between universal phonological patterns and language-specific properties.

Beyond theoretical contributions, the study has significant practical implications for speech-language pathology. The current data may serve as a reference for what constitutes normal production in JA children, facilitating the identification and detection of abnormal speech patterns. The inclusion of superheavy syllables highlights their potential as early indicators of speech production delays or deficits. Through identifying durational patterns and areas of difficulty, speech-language pathologists can develop more targeted interventions to enhance phonological skills in children with speech abnormalities. Insights gained from analysing JA child speech may assist in developing linguistically appropriate assessment tools and intervention strategies, offering practical benefits for speech-language pathologists.

1.6. Structure of the Thesis

The remainder of this thesis is structured as follows: Chapter Two highlights (1) Arabic language and its varieties, phonological aspects of Modern Standard Arabic (MSA), background on JA, phonological properties of JA, such as syllable structure, stress and metrical structure, geminates, epenthesis and syncope, and analysis of superheavy syllables; (2) language development including phonological and phonetic development, comprising the perceptual development of infants and the production patterns in prelinguistic vocalizations to early words, language universals and articulatory motor learning approaches, phonological development in Arabic dialects and JA, and acquisition of word stress. Chapter Three details the methodology, describing the approach, participants, tasks, and data analysis including acoustic (word identification and text-grids, acoustic markers for phoneme level segmentation, and segmental complexities in JA), statistical analysis (Bayesian multi-level modelling, monotonic predictors, priors, efficiency, and convergence), and limitations. Chapter Four presents the results divided into three sections: (1) tasks results, including semi-spontaneous speech task (ST), repetition task (RT), and picture elicitation task (PT) durations of predictors and their interactions, such as age group, stress, syllable position, and syllable structure; (2) the production of superheavy syllables and the Bayesian model outputs for predictors effects on durations across word lengths; and (3) type and frequency of phonological processes in JA child speech. Chapter Five discusses the effect of age group, lexical stress, syllable structure and moraicity, and syllable position within a word on durations in addition to a detailed analysis of superheavy syllable results. Furthermore, the chapter highlights the theoretical and clinical implications, limitations, and directions for future research.

2. Chapter Two: Literature Review

Chapter 2. Literature Review

This chapter first discusses the Arabic language in Jordan, highlighting the background and classification of JA, the phonological properties of JA, and superheavy syllables. Second, the chapter provides a survey on children's language development. Aspects such as perceptual development in infants, prelinguistic vocalizations, early words, phonological developmental theories, phonological development in Arabic, and lexical stress development in child speech are discussed. The study's predictions are also provided.

2.1.The Arabic Language in Jordan

2.1.1. Arabic: An Overview

The Arabic language belongs to the Semitic language family which descended from a broader language family, the Afro-Asiatic or the Hamito-Semitic (Al-Zabibi 1990, Ryding 2005, Huehnergard & Pat-El 2019). Arabic is the official language for members of the Arab League, covering geographical regions from the Arabian Gulf to Northern Africa (Newman 2002, Ryding 2005, Abdoh 2011, Altakhaineh & Zibin 2014). The varieties of Arabic include Classical Arabic (CA), MSA, and many regional dialects. There is a further stylistic variation often called colloquial Arabic (Ryding 2005, Chiang et al., 2006, Diab, Ghoneim & Habash 2007, Biadsy, Hirschberg & Habash 2009, Al-Saidat & Al-Momani 2010).

CA is dated to the pre-Islamic period and is the language of the holy Qur'an. Although it is the mother tongue of nobody, it is not considered dead for its religious significance for more than 400 million Muslims around the globe (Diab, Ghoneim & Habash 2007, Biadsy, Hirschberg & Habash 2009, Al-Saidat & Al-Momani 2010). MSA, derived from CA, is the standard language for written and spoken media, culture, education, and literature (Ryding 2005, Diab, Ghoneim & Habash 2007, Al-Saidat & Al-Momani 2010, Altakhaineh & Zibin 2014). MSA has a similar phonological, morphological, and syntactic structure to CA, but the MSA lexicon is considered more modern (Diab, Ghoneim & Habash 2007). MSA is used in official settings, but it is not acquired as the first language in any Arab country. MSA and colloquial Arabic coexist, serving different communicative purposes (Diab, Ghoneim & Habash 2007).

Children acquire the regional and colloquial varieties as their first language(s) (Chiang et al., 2006, Diab, Ghoneim & Habash 2007, Altakhaineh & Zibin 2014). Colloquial/dialectical Arabic is used for daily conversations, popular media, and culture such as songs, movies, folksongs, tales, and TV shows (Diab, Ghoneim & Habash 2007, p.5). However, it is not taught at schools, standardized, or used in formal settings (Ryding 2005, Biadsy, Hirschberg & Habash 2009). Figure 2.1 shows Arabic varieties that are broadly classified into five dialectical zones, namely, Arabian, Mesopotamian, Levantine, Egyptian, and Maghrebi[1] (Al-Zabibi 1990, Embarki et al., 2007, Biadsy, Hirschberg & Habash 2009, Altakhaineh & Zibin 2014). This classification is not only based on linguistic forms but also extra-linguistic (i.e., social, ethnic, geographic, and historical) criteria (Embarki et al., 2007, Biadsy, Hirschberg & Habash 2009).

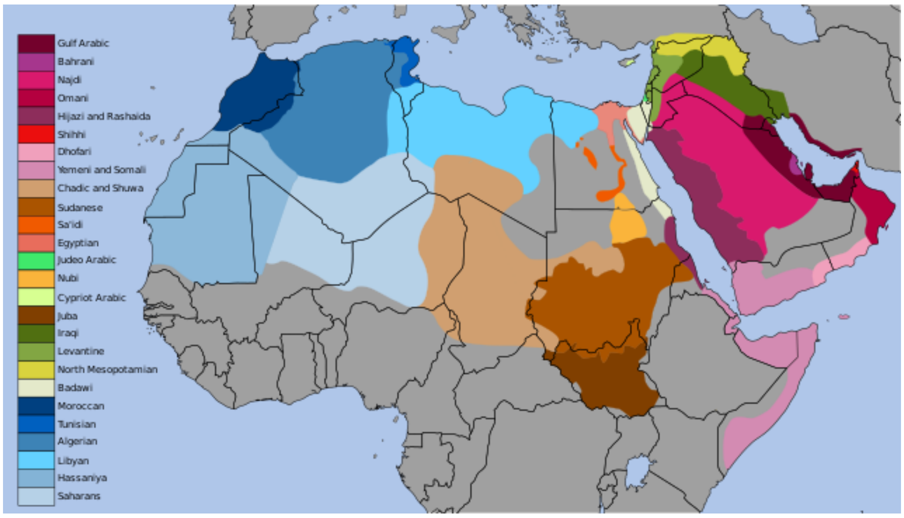


Figure 2.1: Map of Arabic varieties

---

[1] Map of Varieties of Arabic. Map distributed under a CC-BY 3.0 license from Wikipedia.

The segmental inventory of regional varieties is broadly similar to that in MSA; however, they differ in phonological processes. Kiparsky (2003) classified Arabic dialects into CV, VC, and C dialects; the main distinctions are drawn from their semi-syllable licensing and syllabification patterns (Watson 2002, Alqattan 2015, Alhammad 2018). Semi-syllables are defined as un-syllabified moras not affiliated with the syllable node or foot level in the prosodic hierarchy (Kiparsky 2003). Thus, they do not count in syllable weight or foot size; instead, they are directly adjoined to higher prosodic constituents (Selkirk 1981, Ito 1986, Watson 2007). Generally, semi-syllables are typically unstressed, less sonorous than syllable nuclei, restricted to peripheral positions, and prosodically invisible (Kiparsky 2003). Kiparsky's classification is based on the licensing of semi-syllables (at the word level or post-lexically): CV dialects do not permit semi-syllables at any level; VC dialects permit semi-syllables at the word level only, and C dialects permit semi-syllables in word and post-lexical (sentence) levels (Watson 2007, Alhammad 2018).

CV dialects, known as onset dialects, comprise Egyptian dialects, spoken in Cairo, Delta, parts of the Libyan desert, and Middle Egypt (Kiparky 2003). VC dialects are referred to as coda dialects (cf. Irshied & Kenstowicz 1984, Kiparsky 2003); they include dialects spoken in Syria, Lebanon, Jordan, Palestine, Iraq, Turkey, Bedouin areas (Bani-Hassan), Central Arabic (Hijazi), Eastern Libya, the easternmost part of Delta, and Asyut in Upper Egypt. Finally, C dialects, which are characterized by long consonant sequences comprising complex structures and consonantal nuclei, are found in North African regions, namely, Morocco, Tunis, Mauretania, and some Bedouin dialects (Watson 2002, Kiparsky 2003, Alhammad 2018).

2.1.2. Modern Standard Arabic

MSA has a rich consonantal system with 28 consonants including emphatic and guttural consonants in addition to plosives, fricatives, nasals, liquids, and glides (Abdoh 2011, Alotaibi

& Meftah 2013, Embarki 2013). Table 2.1 shows the consonantal inventory of MSA. MSA consonants are characterized by length contrast; singleton consonants are lexically contrasted with their geminate counterparts (/bakaa/ ‘he cried’ vs /bakkaa/ ‘made someone cry’, Newman 2002, Jackson 2000, Embarki 2013, Mustafawi 2017).

Table 2.1: MSA consonants

Place Manner	Bilabial	Labio-dental	Dental	Alveo-dental	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	b			d t dˤ tˤ		k	q		ʔ
Fricative		f	ð θ ðˤ	z s sˤ	ʃ		ɣ x	ʕ ħ	h
Affricate					d͡ʒ				
Nasal	m			n					
Liquid				l					
Tap/Trill				r					
Glide	w				j				

MSA is considered to have six vowel phonemes (Maddieson 1980) (Figure 2.2). The short low central unrounded /a/, short high back rounded /u/, and short high front unrounded /i/ have long counterparts /a:/, /u:/, and /i:/ (Abdoh 2011, Alotaibi & Meftah 2013). In addition, MSA has two diphthongs including /aw/ and /aj/[2]. In MSA, syllables cannot be onset-less (Amayreh & Dyson 2000, Abou-Elsaad, Baz, & El-Banna 2009). MSA vowels occur in syllable medial and final positions only; at the syllable-initial position, vowels are always preceded by the glottal stop /ʔ/.

---

[2] This transcription is used by most Arab linguistics and is adopted in the *Handbook of IPA* (2005) as it reflects Arabic orthography. An alternative transcription could be /ai:/ for /aj/, and /au:/ for /aw/ (Kalaldeh 2018).

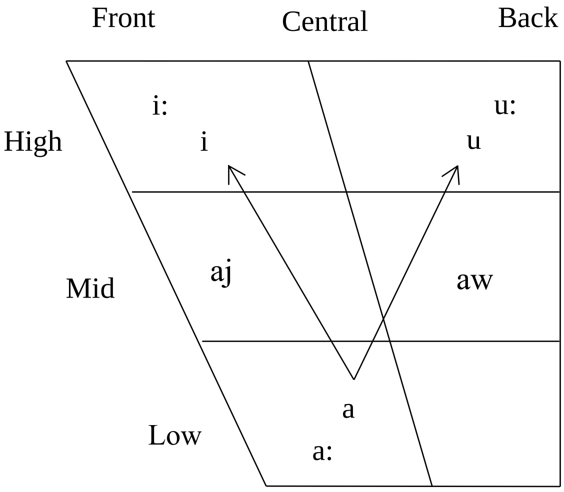


Figure 2.2: MSA vowels

2.1.3. Jordanian Arabic

2.1.3.1.Background on JA

In Jordan, the sociolinguistic phenomenon of diglossia is apparent, with Jordanians utilizing both colloquial JA and MSA (Amer, Adaileh & Rakieh 2011). MSA is predominately used in official settings, education, media, religious sermons, and literature (Laks & Berman 2014). JA is used in informal communication and casual social settings. JA children do not encounter MSA to any large extent before starting school, although they may hear it in some television programs and religious prayers, and it has limited use in literacy-related activities (Amayreh 2000). Therefore, little to no attempts are required from JA children to use MSA before school (Elgibali 1996, Abu-Rabia 2000, Khamis-Dakwar & Froud 2007, Khamis-Dakwar, Froud & Gordon 2012, Laks & Berman 2014).

JA is a linguistically and culturally rich dialect within the Levant region, and it is defined by socio-economic, geographical, and historical factors (Cleveland 1963, Palva 1984, Al-Wer 2002, Al-Wer 2007, Al Mashaqba 2015). Despite limited documentation, JA exhibits diverse variations reflecting Jordan’s complex landscape. JA’s geographical and linguistic boundaries are complex, as multiple linguistic varieties within a single area coexist (Al Mashaqba 2015). Additionally, historical and political influences, particularly with Palestine, have significantly impacted JA. The influx of Palestinians who received Jordanian citizenship after 1949 due to the Palestinian-Israeli conflict has shaped JA's dialectal variations and features (Ferguson 1962, Cleveland 1963, Palva 1984).

Geographically, Jordanian dialects are classified into three types: urban- madanī (Abd-el-Jawad 1981), rural-fallāḥī (Al-Sughayer 1990), and Bedouin (Sakarna 2005) which are distinguished by different phonological, morphological, and lexical characteristics (Suleiman 1993, Sawaie 2007). Socio-economically, Jordanian dialects are further divided into city vs. village dialects, and sedentary vs. Bedouin dialects (Al-Wer 2002, Al-Wer 2007, Al Mashaqba 2015). Cleveland (1963)[3] established a categorization of JA dialects dividing them into four groups corresponding to social and economic stratification in the country, with major reference to voiceless uvular stop /q/ realization:

1. Group One (/yigūl/ 'he says'[4]): Spoken by Bedouins in the eastern and southern deserts of Jordan including Karak in the south of Jordan. This group lacks the indicative marker b- (a prefix used to indicate mood verb forms in Arabic) as /yigūl/ is produced instead of /bigūl/. The uvular stop /q/ in /yiqūl/ is realized as the velar /g/ as in /yigūl/.
2. Group Two (/bəgūl/ 'he says'): Spoken in southern Palestine, the Jordan Valley, and by nomads. The production of /q/ in /baqūl/ is in the form of velar /g/ as in /bagūl/.

---

[3] This classification reflects the realization of the imperfect form of the verb /gāl/ ‘to say’ in different dialects by Cleveland (1963).

[4] Cleveland (1963) suggested that the form of the common expression for ‘he says’ has been chosen as a characterizing feature, as it indicates both an important phonetic and an important morphological characteristic.

3. Group Three (/bəkūl/ 'he says'): Spoken by villagers around Jerusalem and in the northern part of Central Palestine and Jordan. The realization of /q/ in /bəqūl/ is in the form of a prevelar (or postpalatal) /k/ as in /bəkūl/.
4. Group Four (/bəʔūl/ 'he says'): Represents the urban variety found in Jerusalem, Jenin, Hebron, and the capital city of Jordan, Amman. Instead of producing /bəqūl/, this group has /q/ realized as the glottal stop /ʔ/ as in /bəʔūl/.

The current study examines AA in Group Four. The phonetic inventory of AA appears to be very similar to MSA. Table 2.2 shows AA phonemes adopted by Abu Guba (2016, p.13). AA has five vowels including: /i/, /u/, /e/, /o/, /a/ and their long counterparts /i:/, /u:/, /e:/, /o:/, /a:/, respectively. The presence of /e:/ and /o:/ is a result of the diachronic monophthongization process; /aj/ and /aw/ changed to /e:/ and /o:/ (e.g. /sajf/ > /seef/ ‘sword‘, /sawt/ > /soot/ ‘voice’, Abu Guba 2016, p.14). The major phonological differences distinguishing AA and MSA were documented by Amer, Adaileh & Rakhieh (2011, p.34). The differences are as follows:

1. Mid-front vowel /e:/ and mid-back long vowel /o:/ are evident in AA compared to only /aj/ and /aw/ in MSA.
2. The MSA uvular voiceless stop /q/ is realized as a glottal voiced stop /ʔ/ in AA.
3. Although interdental fricatives /θ/ and /ð/ are sometimes produced in AA (e.g., /θaaʔir/ ‘Thaer’ and /muʕaað/ ‘Moath’), these phonemes are commonly realized as /t/ and /d/ or /z/, respectively (e.g., /ðiʔb/ ‘wolf’ being produced as /diib/).

Table 2.2: AA Consonantal inventory

Place **Manner**	**Labial**	**Labio-dental**	**Inter-dental**	**Dento-alveolar**	**Palato-alveolar**	**Palatal**	**Velar**	**Uvular**	**Pharyngeal**	**Laryngeal**
Stop	b			t d			k g	q		ʔ

				tˤ dˤ						
Fricative		f	θ ð	s z zˤ sˤ	ʃ		x ɣ		ħ ʕ	h
Affricate					dʒ					
Nasal	m			n						
Lateral				l						
Trill				r						
Glide	(w)					j	w			

2.1.3.2. Phonological Properties of JA

This section highlights the phonological properties of JA as a VC dialect. It focuses on the role of metrical theory in shaping the understanding of the phonological properties in JA using metrical constraints. Additionally, a moraic approach to analyse syllable structure is employed.

2.1.3.2.1. Syllable Structure

JA has eight syllable types: CV, CVC, CVCC, CVV, CVVC, CVVCC, CCVC, and CCVVC (AbuAbbas 2003, Crossley 2023). Each syllable comprises an obligatory onset, an obligatory vowel, and an optional coda (Amayreh & Dyson 2000). The obligatory onset can either be a non-branching (C-) or branching (CC-) as illustrated in examples (1) and (2). Vowels can be short (V) or long (VV) as in (3) and (4), respectively.

1. /**t**uf.faaħ/ ‘apple’ CVC
2. /**sl**aaħ/ ‘weapon’ CCVVC
3. /k**a**s.sar/ ‘he broke’ CVC
4. /f**aa**r/ ‘mouse’ CVVC

A coda is not obligatory (Example 5). However, when present, it can comprise a single segment (6) or a coda cluster (7) (Watson 2002, Abdoh 2011).

5. /**daa**.ris/ ‘educated’ CVV
6. /maf.**tuuħ**/ ‘opened’ CVVC
7. /ka**lb**/ ‘dog’ CVCC

Examples (8 & 9) show how consonant clusters in JA can occur in word-initial and final positions while word medial clusters are syllabified as a coda-onset sequence (i.e., C.C as in 10) (Abdoh 2011).

8. CCVC /slaħ/ ‘weapon’
9. CVCC /dars/ ‘lesson’
10. CV.CV**C.C**V /ʔa.kalt.ha/ ‘I ate it’

Cluster production is restricted by the Sonority Sequencing Principle (SSP)[5] (Al-Ani 1970, Kiparsky 2003). While onset clusters can have a falling or rising sonority, coda clusters only have a falling sonority in JA (Al-Ani 1970, Kiparsky 2003, Daana 2009, Al Tamimi & Shboul 2013). For example, the production of onset cluster in /zlaam/ ‘men’ and /dmuuʕ/ ‘tears’ obeys SSP, but the production of /nħa.raʔ/ ‘burned’ and /rfuuf/ ‘shelves’ flouts SSP. JA coda clusters appear in two forms, heterogeneous (i.e., containing different consonants /kalb/ ‘dog’ and /fard/ ‘pistol’) and homogeneous clusters (i.e., containing the same consonant creating a geminate /dubb/ ‘bear’ and /sadd/ ‘dam’). Figure 2.3 demonstrates the production of /ʔakl*/ ‘food’ which is not permissible due to the rising sonority; thus, it surfaces as /ʔakil/ which abides by SSP through epenthesis.

---

[5] SSP stipulates that the nucleus has the peak sonority of the syllable while surrounding segments’ sonority levels depend on the segment's distance from the vowel. The sonority hierarchy in Arabic, based on the manner of articulation, is as follows: Vowels > Glides > Liquids > Nasals > Fricatives > Stops (Carlisle 2001, Watson 2002, Daana & Khrais 2018).

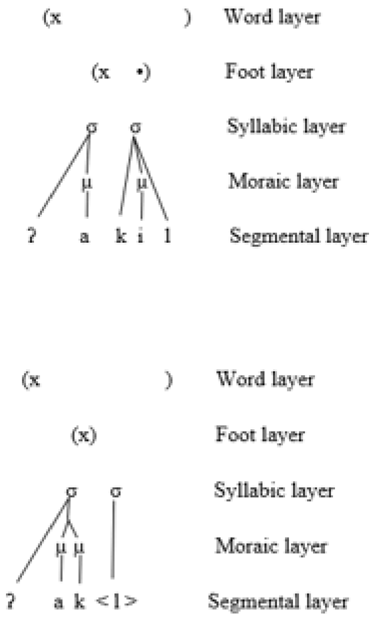


Figure 2.3: A coda clusters with rising sonority

JA syllables are classified into three categories: light, heavy, and superheavy (Watson 2002, Abdoh 2011, Alotaibi & Meftah 2013, Alqattan 2015). Based on the Moraic Theory (Hayes 1995), a light CV syllable (as in 11) and a word-final CVC (as in 12) are monomoraic (Watson 2002, 2011, Davis & Ragheb 2014). A heavy syllable is bimoraic, where the two vowels in a CVV pattern receive two moras (as in 13). Similarly, a non-final CVC pattern is assigned two moras (Al-Thamery & Ibrahim 2005, Saiegh-Haddad & Henkin-Roitfarb 2014). Any syllable that exceeds two moras is classified as a superheavy syllable (as in 15-17).

11. CV /**ʔa**.lam/ ‘pen’
12. CVC /daf.**tar/** ‘notebook’
13. CVV /**kaa**.tib/ ‘writer’
14. CVC /**mal**. ʕab/ ‘field’
15. CVVC /baasˤ/ ‘bus’
16. CVCC /kalb/ ‘dog’

17. CCVVC /ħsˤaan/ ‘horse’

Superheavy syllables are linguistically complex structures that seem to be avoided cross-linguistically (Hayes 1995, Davis 2011, Alahmari 2021). In JA, superheavy syllables include CVVC, CVVCC, CVCC, and CCVVC (AbuAbbas 2003). Superheavy syllables contain a larger number of segments compared to other structures consisting of either a long vowel (e.g., /baab/ ‘door’ CVVC), a coda cluster (e.g., /bint/ ‘girl’ CVCC), or a combination of both (e.g., /saadd/ ‘clogging up’ CVVCC). These syllables often arise from morphologically complex words that include plural forms (e.g., /sajjaar**aat**/ ‘cars’ CVC.CVV.**CVVC**) and derivatives (e.g., prefixes and suffixes as in /ʔaaʕ**diin**/ ‘they are sitting’ CVVC.**CVVC**).

Stipulated by the bimoraicity constraint that does not permit trimoraic syllables (Broselow et al., 1995), superheavy syllables in JA are suggested to be bimoraic. Then, it is expected that JA superheavy syllable productions undergo processes such as epenthesis and vowel shortening to abide by this constraint (Watson 2002). McCarthy & Prince (1986) suggested that superheavy syllables in the word-final position receive two moras only (Figure 2.4). The final consonant of a superheavy syllable is rendered as extrasyllabic rather than extrametrical (Watson 2002, p.58). Extrasyllabicity accounts for the final C that falls outside the target syllable's domain, comprising a canonical syllable that is left un-syllabified until a later stage in derivation (Watson 2002, Watson 2011). In conformity with the End Right Rule (ERR), the canonical syllable receives the main stress as the extrasyllabic final consonant prevents the canonical syllable from being peripheral (Hayes 1981, Watson 2002).

18. CVVC /ma.**laak/** ‘angel’

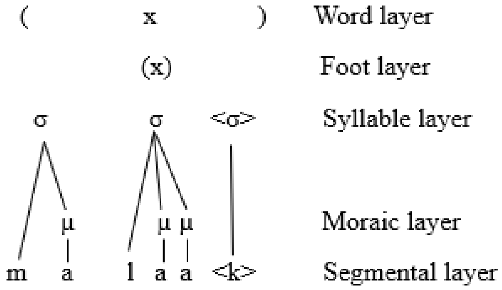


Figure 2.4: A word-final superheavy syllable /malaak/ 'angel'

However, Figure 2.5 demonstrates how a superheavy syllable appearing in a non-final position (as in 13) poses a challenge. A mora is assigned to a non-final coda by the Weight by Position Rule (WPR) which is essential for determining syllable weight distribution (Hayes 1995). As the long vowel receives two moras and a third mora is assigned to the non-final coda by WPR, the syllable is supposed to be trimoraic and not bimoraic.

19. CVVC /**daar**.siin/ 'they studied'

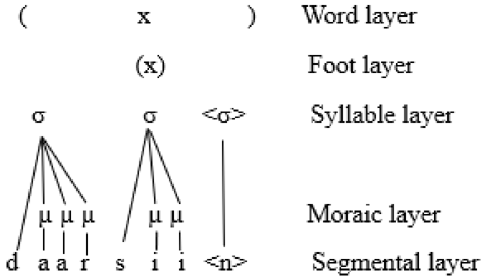


Figure 2.5:A word non-final superheavy syllable /daarsiin/ 'they studied'

2.1.3.2.2. Stress and Metrical Structure

It is widely accepted that a stressed syllable within a word is metrically strong, and stress is an abstract property of a word inherent to its phonological structure (Ladd 2008). For example, in English, noun and verb pairs such as '**per**mit' (noun) and 'per**mit**' (verb) are differentiated by stress patterns. Acoustic cues, such as lengthening and an increase in loudness, are associated with stressed syllables, while vowels are reduced in unstressed syllables (Fry 1958, Van Heuven 1987, Gussenhoven 2004, Ladd 2008). However, the acoustic cues to lexical stress are not always clear. While Bolinger (1958) suggested that pitch is a crucial acoustic correlate of lexical stress, subsequent studies revealed that pitch is not a reliable cue to stress in English. While the noun and verb 'permit' contrasts exhibit clear pitch distinctions in isolation, with a pitch rise associated with the stressed syllable, these distinctions might not be notable in different intonational contexts (Ladd 2008). For instance, when these words are spoken as questions (i.e., 'permit?'), then there would be an overall pitch rise. Thus, pitch accents, which involve changes in pitch that highlight specific syllables or words within utterances play an important role in determining utterance-level intonation (Liberman & Prince 1977, Ladd 1996). Stress and pitch accents are interlinked but not synonymous (cf. Liberman 1975, Bruce 1977, Pierrehumbert 1980, Ladd 1996, Gussenhoven 2004).

In Arabic, stress is defined as a phonological property of a word, determined by stress placement rules (Hayes 1995). Syllable weight, the internal structure of a syllable, and word length play an important role in determining lexical stress location (Hayes 1995, de Jong & Zawaydeh 1999). Therefore, the location of the stressed syllable in a word is predictable. An early study on Standard Arabic as produced in Iraq showed that stressed syllables tend to be longer, louder, and higher in pitch compared to unstressed syllables (Al-Ani 1970). De Jong and Zawaydeh (1999) examined acoustic correlates of lexical stress in AA, such as duration, spectral properties, and pitch in the speech of four female adult speakers. The data consisted of ten words placed in prosodic conditions comprising final and non-final positions in utterances forming statements and questions. Results showed that vowels in stressed syllables were significantly longer and had higher pitch than those in unstressed syllables. Moreover, Zuraiq (2005) analysed the production of lexical stress by native speakers of JA, English, and Arab learners of English. Eight participants, including four males and four females, were recruited from the northern region of Jordan. The stimuli consisted of six minimal pairs of disyllabic words. These words represented differences in stress with six nouns being stressed on the first syllable, and their minimal pair counterparts (comparative adjectives) being stressed on the second syllable ('a.sad - a.'sad ‘lion’). Results showed that longer duration, higher amplitude, and higher pitch were correlates of stressed syllables compared to unstressed ones. Nevertheless, JA speakers did not employ vowel reduction or F2 as a correlate of stress

Building on these phonetic observations, the phonological structure of stress in JA follows specific rules and patterns. Stress in JA does not contribute to differentiating lexical meaning (Al-Ani 1970, Watson 2002, De Jong & Zawaydeh 1999, Abdoh 2011). Analyses of Jordanian dialects, such as Bani Ḥasan Arabic (Irshied 1984), Wadi Mousa Arabic (Huneety 2015), and Wadi Ramm Arabic (Al Mashaqba 2015) demonstrated that similar to MSA, JA follows a fixed stress system that is rule-governed and follows a trochaic moraic system. This system is associated with intensity contrasts that organize syllables into trochees of (S-w) pairs with the first segment of the pair being more prominent than the second (Hayes 1995, Watson 2002). The following examples in Figure 2.6 show the (S-w pair) trochaic system of JA.

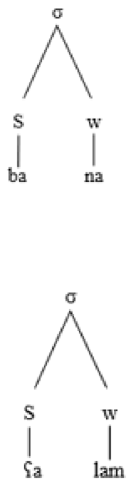


Figure 2.6: Examples of a S-w trochaic pattern

JA words must meet the bimoraicity condition to surface. A word must contain at least one foot, which is maximally assigned two moras (cf. McCarthy and Prince 1990, Prince & Smolensky 1993, Hayes 1995, Watson 2002). The metrification directionality of JA has a left to right foot parsing with an absolute ban on degenerate feet (i.e., stranded moras at word edges that are left un-footed and are ineligible to construct a foot) (Hayes 1995, Kager 1995, Watson 2011). Therefore, a mono-moraic word will not be eligible to construct a foot or surface and the coda consonant undergoes gemination CVC-CVCC (Hayes 1995, Davis 2011). Gemination results in the monomoraic syllable becoming bimoraic, allowing it to construct a visible foot that receives the main stress. Figure 2.7 demonstrates the moraic configuration of the words /sad*/ and /sadd/ 'dam'.

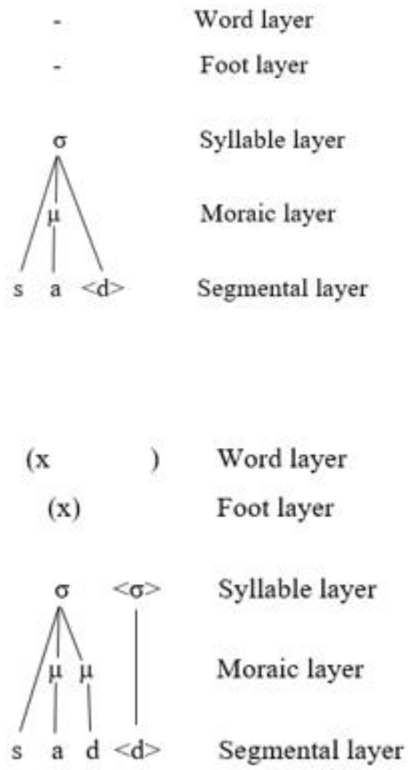


Figure 2.7: Gemination of a monomoraic /sad/ becoming bimoraic /sadd/*

In monosyllabic words, JA word stress falls on the only bimoraic syllable (Huneety & Mashaqba 2016). Examples of monosyllabic words comprised of heavy or superheavy syllables are as follows:

20. CVV /ˈlaa/ ‘no’
21. CVCC /ˈbint/ ‘girl’
22. CCVVC /ˈktaab/ ‘book’
23. CVCC /ˈsadd/ ‘dam’

In disyllabic words, stress is assigned to a word-final superheavy syllable; otherwise, the penultimate heavy syllable receives stress (Example 16-18). In the case of two light syllables in a disyllabic word (Example 19-20), stress is assigned to the penultimate syllable as JA has a trochaic foot system (Watson 2002, Ayyad 2011, Watson 2011).

24. CVC. **CVVC** /mak.ˈtuub/ ‘written’

25. CV.**CVCC** /ma.ˈħall/ ‘shop’

26. **CVC**.CVC /ˈmak.tab/ ‘desk’

27. **CV**.CVC /ˈdˤa.rab/ ‘he hit’

28. **CV**.CV /ˈsa.ma/ ‘sky’

In multisyllabic words, stress is assigned to one of the word-final three syllables (Davis & Ragheb 2014, Mashaqba & Huneety 2018). Thus, a heavy CVC syllable in a pre-antepenultimate position fails to attract and receive the main stress. To illustrate, al Huneety et al. (2023) indicated that in /maʕ.ka.roo.na/ ‘pasta’ a foot over the bimoraic preantepenultimate syllable /maʕ/ is constructed, whereas the antepenultimate monomoraic syllable /ka/ fails to construct a foot. Another foot is constructed over /roo/ which receives the main stress. The heavy syllable /maʕ/ failed to receive the main stress due to ERR that assigns stress to the rightmost heavy syllable /roo/. The rules of stress assignment can be summarized as a series of conditional statements as follows:

1. Stress falls on the ultimate superheavy syllable (CVVC or CVCC), if not;
2. Stress falls on the penultimate heavy syllable (CVC or CVV), if not;
3. Stress falls on the antepenultimate syllable.

A CVC syllable is rendered as heavy in non-final word positions (e.g., /**mak**.tab/ CVC.CVC ‘desk’) by WPR. However, a CVC is light in the word-final position, where the coda is deemed as extrametrical by the peripherality condition[6] (e.g., /ka.**tab**/ CV.CVC ‘he wrote’). Figure 2.8 shows how a CVC syllable can be light and heavy, depending on syllable position.

---

[6] The peripherality condition suggests that a word-final C is rendered as extrametrical as it is located at a domain final position or edge (Liberman & Prince 1977, Hayes 1995).

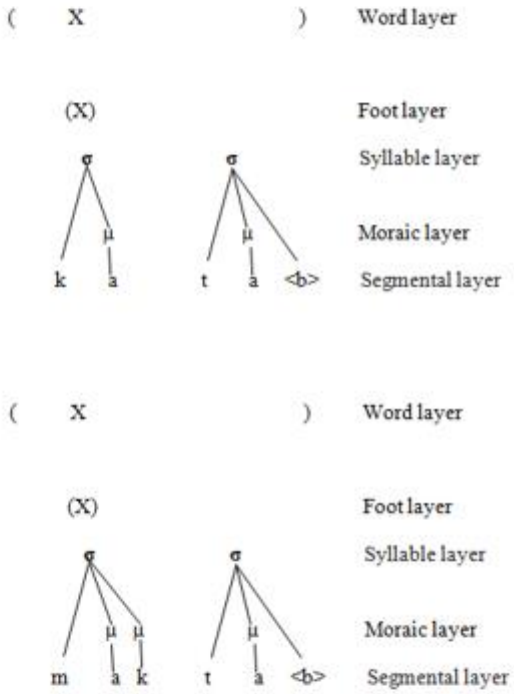


Figure 2.8: CVC syllable in final and non-final positions

2.1.3.2.3. Geminates

Geminates play a significant role in shaping the phonological patterns and the temporal properties of JA influencing stress assignment, surrounding vowel durations, and the moraic organization of productions (Watson 2007, Davis 2011, Davis & Ragheb 2014). However, a geminate can be defined in different ways.

Phonetically, geminates are long sounds that require greater muscular tension in the articulators for an extended period compared to their singleton counterparts (Trubetzkoy 1969, Catford 1977). Geminates are produced with greater energy and stronger articulation of a sound segment (Ridouane 2007, Khattab & Al-Tamimi 2013). Phonologically, a geminate is defined as a long or a double consonant that is phonemically contrastive to its short singleton counterpart (Davis 2011, Al-Deaibes 2016). There are two main views to account for geminates: prosodic length and moraic weight representations. The prosodic length representation views geminates as two elements linked to two C-slots on a skeletal tier while singletons are linked to one C-slot (Hayes 1989, Watson 2002). In the skeletal tier of a CVCVC template in the word /fataħ/ 'opened', the segment /t/ is linked to a C slot (Figure 2.9). Alternatively, the CV**CC**VC template of the word /fattaħ/ 'blossomed' has a geminate /tt/ which is linked to two C slots. Thus, the segments of a geminate are assigned two C slots in the syllable boundary, occupying the coda position of the first syllable and the onset position of the following one (i.e., CV**C**.**C**VC).

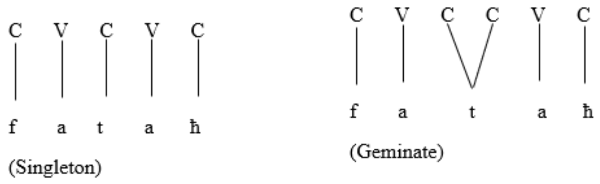


Figure 2.9: Skeletal tiers for CVCVC and CVCCVC templates

In the moraic weight representation, geminates have an inherent weight. In the case of a word non-final geminate occurring at the syllable boundary, the first segment of the geminate occupying the coda position is assigned a mora by WPR (cf. Watson 2002, Davis and Ragheb 2014). Alternatively, the singleton appearing in an onset position is not assigned a mora due to onsets being weightless in Arabic (Mashaqba et al., 2019). Figure 2.10 demonstrates the moraic representation difference between singleton and geminate environments, /fataħ/ and /fattaħ/, respectively.

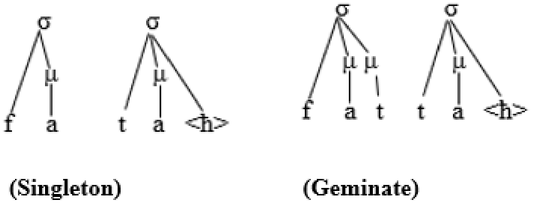


Figure 2.10: Moraic representations of /fataħ/ and /fattaħ/

Acoustically, previous studies had conflicting results on the extent to which temporal compensation (i.e., the adjustment of duration in one segment of speech to balance changes in another) cues are evident in geminates (Al-Deaibes 2016). Temporal compensation patterns shed light on the correspondence between geminate consonant lengthening and shortening in adjacent vowels. For example, Ferrat & Guerti (2017) suggested that Algerian MSA geminate productions exhibited decreased duration of the preceding vowel and increased duration of the following vowel. In Lebanese Arabic, Khattab and Al-Tamimi (2014) discussed that vowel shortening preceding intervocalic geminates affected long vowels rather than short ones. In JA, Al-Deaibes's (2016) study confirmed that short vowels in geminates are significantly shorter than those in singleton contexts, while long vowels in geminates are significantly longer than in singletons. Al-Tamimi, Abu-abbas & Tarawneh (2010) argued that temporal compensation between word-final geminates and their preceding vowels was evident in JA adult productions. Nonetheless, Thnaibat (2019) suggested that in JA child speech, no temporal compensation cues were evident in the preceding or following vowels in intervocalic geminate environments.

2.1.3.2.4. Epenthesis

Epenthesis, or the post-lexical insertion of a prothetic vowel, is used to avoid consecutive consonantal sequences (Kager 1999). Kiparsky's (2003) classification of dialect typology is constructed on the position of the epenthetic vowel in CCC clusters. Epenthesis in VC dialects, such as Iraqi Arabic, forms a C(V)CC pattern where the vowel appears to the left side of the first C (e.g., /gilt.la/ > /gilit.la/ 'I told him', Watson 2007). However, in CV dialects, such as Cairene Arabic, the epenthetic vowel in a CC(V)C structure appears to the left side of the last C in a CCC cluster (e.g., /ʔult-lu/ > /ʔultilu/ 'I told him', Watson 2007). Epenthesis is a crucial phonological phenomenon to consider in addressing syllable structure and stress assignment in Arabic. This is due to the interaction between epenthesis and stress being opaque wherein VC-dialects, such as Baghdad Jewish and Christian Arabic, consider the epenthetic vowel invisible to lexical stress (Abu-Salim 1980, Abu-Haidar 1991). Kiparsky (2003) explains that not only is the epenthetic vowel invisible to stress but the whole syllable it appears in is not considered in any lexical processes. For example, in the word /fihm.na/ ‘our understanding’, the stress is assigned to the antepenultimate syllable when produced as /ˈfi.him.na/ and not the penultimate based on ERR.

Coda cluster formation and vowel epenthesis in Arabic superheavy syllables have been investigated in Qassimi Arabic (Alhoody & Aljutaily 2020), JA (Daana 2019), Bedouin JA (Kenstowics 1986), Karaki Arabic (Btoosh 2006), Lebanese Arabic (Abdul-Karim 1980, Haddad 1984), Maani (Rakhieh 2009), Najdi (Alhammad 2018, McCarthy 2008), Palestinian (Abu-Salim 1982), Hadrami (Bamakhramah 2010), Urban JA (Na'eem, Abudalbuh & Jaber 2020), San’ani Arabic (Watson 2002), Cariene Arabic (Ragheb and Davis 2014, Abu-Mansour 1987), Hijazi Arabic (Bokhari 2020), and Iraqi/Egyptian/Omani/ Sudanese (Elashhab 2018).

Epenthesis in JA productions is expected to avoid trimoraic syllables (Abu-Salim 1980, Benhallam 1980, Broselow 1992, Abu-abbas 2003, Mashaqba et al., 2019). Daana (2018) discussed the strategies used by JA speakers to handle coda clusters formed in superheavy syllables, such as CVCC and CVVCC. Recalling that JA coda clusters are not permitted unless they have a falling sonority, her OT analysis justifies that elements of the coda cluster (geminates) have equal sonorities or plateaus. This renders the emergence of CVCC syllables tolerable. Nonetheless, CVVCC is regarded as a controversial syllable structure in the literature as it violates the *3µ constraint (Kager 1999). This is due to the two vowels being assigned two moras inherently, and the first C of the geminate comprising the coda cluster receives a third mora by WPR (Hayes 1995). Daana suggested that the insertion of a vowel between the geminate segment entails a violation of the geminate constraint, yet it seems that triggering this constraint is ranked below the *3µ constraint. Figure 2.11 shows vowel epenthesis in a geminate coda cluster in the word /maarr/ ‘passer-by’ to yield /maa.rir/ that confirms the *3μ constraint.

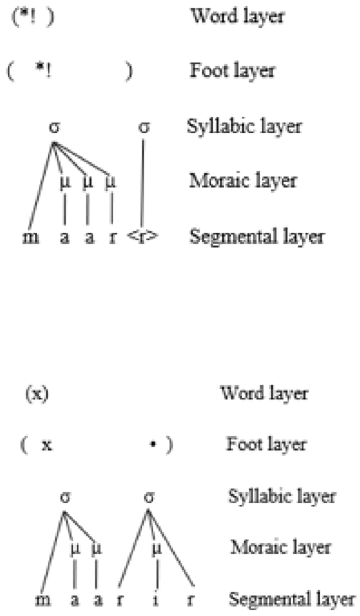


Figure 2.11: Vowel epenthesis in a geminate coda

In his analysis of CVCC clusters attested in Bedouin JA lexical items, Kenstowics (1986) noticed that SSP is not functional in some coda cluster examples. SSP operates in words such as /bint/ ‘girl’ and /dars/ ‘lesson’ but not in /himl/ ‘load’ as it surfaces post-lexically as /himil/. Kenstowics mentioned that due to /himl/ not demonstrating a falling sonority profile, an epenthetic vowel is inserted to alter the syllable sonority scale. This process has been reported in other Arabic dialects such as Lebanese Arabic (Abdul-Karim 1980, Haddad 1984a), Palestinian Arabic (Abu-Salim 1982), and Hijazi Arabic (Bokhari 2020). Btoosh (2006) supported this notion arguing that Karaki Arabic clusters with rising sonorities or tri-consonantal clusters behave differently than other syllables. To illustrate, the tetra-consonantal medial cluster in the word /qult.lha/ ‘I told her’ is broken up by epenthesis forming /qul.til.ha/. The epenthetic vowel serves as the nucleus of the newly established syllable to regulate the number of assigned moras. Cases of irregular sonority patterns such as /habs/ ‘prison’ or /ʔamr/ are prohibited from surfacing unless epenthesis occurs resulting in /habis/ and /ʔamir/.

2.1.3.2.5. Syncope

Syncope is the deletion or loss of an unstressed high short vowel in an open syllable CV (Rakhieh 2009). It is considered an essential phonological process triggering modifying syllable structure in most Arabic dialects including JA (Sakarna, 2006), San’ani Arabic (Watson 2002, McCarthy 2008), Ammani Arabi (Daana 2018), Karaki Arabic (Btoosh 2006), Hijazi Arabi (Al-Mozainy 1981), and Hadrami Arabic (Bamakhramah 2010). Kiparsky (2003) shed light on how syncope is linguistically tolerated to create initial clusters if the cluster's first segment is deemed as semi-syllable. For example, the word /klaab/ ‘dogs’ CCVVC has an underlying representation of /kilaab/ CVCVVC. Due to the syncopation process, the unstressed short vowel in the open syllable /ki/, the /kl/ onset cluster emerges with /k/ being assigned an affiliated mora that is directly parsed to the word level. In VC dialects, syncope is permitted to create word-initial clusters (McCarthy 2008). Alhoody & Aljutaily (2020) reported that CCVVC syllables emerge in Qassimi Arabic due to syncopating **CV**CVVC syllables, such as /ħimaar/-/ħmaar/ ‘donkey’, /turaab/ - /traab/ ‘sand’, and /firaaʃ/-/fraaʃ/ ‘bed’.

There are two categorizations of Arabic dialects based on the syncopation pattern including differential and non-differential (Ito 1989, McCarthy 2008, Alhoody & Aljutaily 2020). Differential dialects refer to those that only delete weak high short vowels in open syllables, and these dialects are also divided into two sub-groups. The first sub-group contains dialects, such as JA and Egyptian Arabic, that delete any weak high short vowels in open syllables, whether it is a high front short vowel /i/or a high back short vowel /u/. JA deletes all high short vowels in non-final open syllables (Rakhieh 2009, Mashaqba & Huneety 2018). However, when the low short vowel /a/ is followed by a non-final open syllable, the short vowel /a/ is deleted in Bedouin JA (e.g., /bagara/ ‘cow’ is syncopated into /bgara/) (Irshied 1984, Sakarna 1999, Rakhieh 2009). In Egyptian Arabic, the unstressed high vowel /i/ in the word /wi.ħi.ʃa/ ‘ugly’ undergoes syncopation to become /wiħ.ʃa/, and the word /xulusit/ ‘she finished’ becomes /xulsit/ after syncope (Watson 2002, p.71). The second sub-group is comprised of dialects deleting only high front short vowels, such as Meccan Arabic, which has more restrictions on syncopation (e.g., /misiku/ ‘they held’ is produced as /misku/, Kabra 2004). In contrast, non-differential dialects delete weak short vowels in open syllables, including high, mid, and low short vowels. For example, in Syrian Arabic, there are no rules governing syncopation that make the distinction between high and low vowels that are weak and short. The high short vowel /i/ in the underlying representation /nizilna/ ‘we went down’ undergoes syncope as in /nzilna/, and the low short vowel /a/ in the underlying representation /laħamt/ ‘I welded’ becomes /lħamt/ (Adra 1999, p.38).

In superheavy syllables, syncope triggers the resyllabification of the final consonant to become the onset of the newly formed syllable, which happens in concurrence with the deletion of the unstressed short vowel (Kiparsky 2003, Watson 2007, McCarthy 2008, Alamro 2016). Consequently, superheavy syllables appearing in non-final word positions risk being syncopated to avoid provoking a violation in the syllabic-moraic structure (Alamro 2016). Previous studies reported that syncope is allowed more frequently in geminates when compared to other CC clusters (Kiparsky 2003, Watson 2007, McCarthy 2008). VC dialects tend to delete high vowels after a geminate consonant to result in a superheavy syllable that is prosodically licensed as it does not violate the bimoraic constraint (e.g., /y-kallim-u/ > /y(i)kal(l)mu/ ‘they talk to someone’ and /tixiil-na/ > /tixiinna/ ‘you confuse us’, Kiparsky 2003, Watson 2007 Rakhieh 2009).

However, the characteristics of syncope in JA and its association with stress have not received much attention. Scholars have neglected the synchronization of how superheavy syllables and clusters are formulated based on the syncopation process. As a VC dialect, Table 2.3 demonstrates examples of JA superheavy syllables and clusters being by-products of syncopation.

Table 2.3: Superheavy syllables emerging due to syncopation

Word	**Syllabification**	**Word (Syncope)**	**Syllabification**	**Gloss**
kitaab	CV.CVVC	ktaab	CCVVC	Book
silaaħ	CV.CVVC	slaaħ	CCVVC	Weapon
fa.him.tu	CV.CVC.CV	fhim.tu	CCVC.CV	You understood
ħa.fizt	CV.CVCC	Ћfizt	CCVCC	Memorized
ka.biir	CV.CVVC	Kbiir	CCVVC	Big
sa.ɣiir	CV.CVVC	Sɣiir	CCVVC	Small
ʔin.kabb	CVC.CVCC	Nkabb	CCVCC	Got spilled

2.1.4. Superheavy Syllables

Analysing superheavy syllables provides insights into Arabic phonology features, including prosodic structure, stress patterns, and morphology (Hyman 1985, McCarthy & Prince 1986, Hayes 1989, Owens 2013). However, the existing literature only offers phonological accounts for the behaviour and representation of superheavy syllables. Acoustic data supporting or defying the theoretical frames remains scarce. The previous studies tackled the production of superheavy syllables in adult speech, neglecting child speech. Thus, superheavy syllable production in child speech remains an intriguing yet under-researched area despite its potential implications on prosody, language learning, and speech development.

Superheavy syllables are problematic to the theoretical accounts for multiple reasons. First, they have more segments compared to other structures affecting their weight distribution (Bamakhramah 2014, Crossley 2023). Second, they appear in critical positions, such as word-final positions, where phonological processes (e.g., lengthening or vowel reduction) and morphological markers (e.g., tense or case) are evident (Hayes 1989, Broselow et al., 1995). Third, they interact with lexical stress assignment which adds to their complexity. Fourth, they have a two-way interaction between syllable weight and position leading to debates about whether they should be classified as bimoraic (Broselow et al., 1995, Watkins 2001, Zec 2007, Bamakhramah 2010) or trimoraic (Fery 1998, Crossley 2023).

Hubbard (1994) suggested that languages with superheavy syllables tend to neutralize them. Thus, rather than maintaining the complex structure of these syllables, languages modify them to confirm with phonological constraints. This is indicative of superheavy syllables behaving differently from other syllable structures posing unique challenges and triggering phonological processes to simplify their productions. Hubbard (1994) summarized two reasons for positing a universal constraint prohibiting these trimoraic syllables (McCarthy & Prince 1986, Steriade 2001, Broselow et al., 1995, Bamakhramah 2014). First, the lack of evidence for a three-way syllable weight distinction: differences between monomoraic (CV) and bimoraic (CVV/ CVC) syllables are commonly observed; however, there is limited evidence supporting the existence of trimoraic syllables. Second, the restructuring of CVVC syllables wherein CVVC syllables are restructured to bimoraic CVC syllables or are limited to the edges of target constituents rendering them as extraprosodic (cf. McCarthy & Prince 1990). In Cairene Arabic, CVVC syllables in the word-final position are produced as /kitaab/ ‘book’ CV.CVVC but when it appears in word non-final positions, it is produced as /kitabna/ ‘our book’ CV.CVC.CV.

Phonetically, few studies have demonstrated that superheavy syllables undergo vowel reduction and mora sharing to prevent violating the bimoraicity constraint. Mora sharing refers to the process where a single mora is linked to multiple segments i.e., a vowel and a coda share a mora to avoid contributing for a third mora in languages that have bimoraic syllables only such as Arabic (Broselow et al., 1995, Hayes 1995, Watson 2002). To illustrate, Broselow et al.

(1995) traced the durational properties of word-internal CVVC syllables with a particular focus on Lebanese and Syrian (permit CVVC) and Egyptian dialects (do not permit CVVC. They argued that in languages with non-final codas being moraic by WPR, CVVC should be trimoraic. However, Broselow and colleagues suggested that durational and phonological data supported the bimoraic representation of CVVC syllables in Arabic with mora sharing between the final consonant and the long vowel. The Lebanese data demonstrated the vowel in /ki.**taab**.na/ (Mean = 97.8 ms) being shorter than the long vowel in an open syllable in /ki.**taa**.bi/ (Mean = 115.4 ms). Similar results were attested for the Syrian data where the vowel in /ki.**taab**.na/ (Mean = 112.2 ms) was shorter than the vowel in /ki.**taa**.bi/ (Mean = 123.9 ms). Egyptian dialects exhibited bimoraic non-final CVVC syllable productions as vowel shortening, rather than mora sharing, was attested. The long vowel in a CVVC syllable was produced as a short vowel with no durational difference between CVVC and CVC syllables where the vowel and the coda consonant are assigned separate moras.

Similarly, Khattab and Al-Tamimi's (2014) analysis of superheavy syllables in Lebanese Arabic highlighted mora sharing. Their analysis of medial geminates showed the correspondence between phonetic timing and moraic weight according to syllable structure. Their predictions implied if there was no language-specific constraint on superheavy syllables, there would be no motivation for restructuring the target syllable (McCarthy and Prince 1990). The analysis focused on geminate productions, and data was comprised of comparing the following environments: CV **ma**.lak 'name', CVG **mal**.lak 'owner', CVV **maa**.lak 'what's wrong', and CVVG **maal**.la 'bored'. The statistical analysis confirmed that phonological length was reflected in duration, wherein environments with short vowels and consonants had a mean duration of 300 ms. When the vowel or the consonant was long, the mean duration was around 400 ms whereas the duration was approximately 500 ms when both the vowel and the consonant were long. In VVC.CV sequences (Mean = 149 ms, SD = 32), mora sharing was evident in shorter VV segments compared to VV.CV sequences (Mean = 166 ms, SD = 36). Nevertheless, the CC segments were not shortened in VCCV sequences (Mean = 182 ms, SD = 41) compared to VVCCV ones (Mean = 181 ms, SD = 38) suggesting that mora sharing only influenced the preceding vowel. Such results were attributed to the inherent weight of geminate consonants requiring shortening to avoid trimoraic syllables in Lebanese Arabic. The question remains whether the attested durational results were due to geminate effects or syllabic structure.

Watson (2007) described that superheavy syllables are restricted to word-final positions. CVCC syllables are restricted to the word-final position in San’ani and to utterance final in Cairene, whereas CVCCC syllables only appear in San’ani in the word-final position. Watson explained that a superheavy syllable comprises a canonical syllable and a degenerate syllable (Figure 2.12). The degenerate syllable blocks foot extrametricality, preventing the canonical syllable from being peripheral, as the degenerate syllable appears between the constructed foot over the canonical syllable and the rightmost edge of the target word. Thus, the superheavy syllable constructs a foot that receives the main stress by ERR and an additional degenerate syllable.

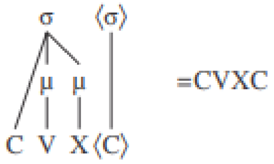


Figure 2.12: Coda consonant in a superheavy syllable

Watson further investigated foot parsing in the presence of a superheavy syllable. She suggested that in San’ani Arabic, a CVCC superheavy syllable fails to receive the main stress if a non-final CVV or CVG (geminate) syllable appears. For example, in the words /daw.wart/ ‘looked for’ and /saa.fart/ ‘travelled’, the stress is assigned to the penultimate syllables. She argued that there is a sonority disparity between the syllables, where the former syllables are assigned two moras in the upper moraic level but the latter syllables are only assigned one. Then, a foot is constructed over the initial syllables only. ERR assigns the main stress to the only visible foot, as shown in Figure 2.13 (cited in Watson 2007, p.104).

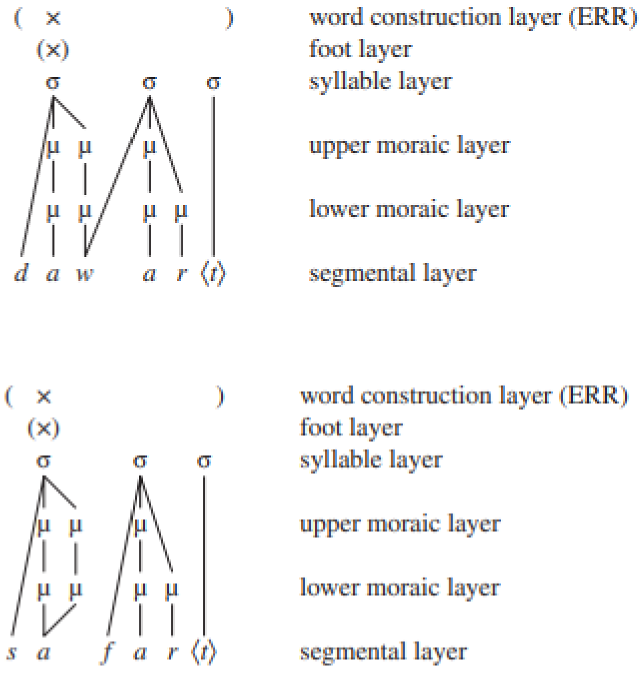


Figure 2.13: Stress assignment to the visible foot by ERR

Nevertheless, in cases of words with suffixations, such as /daw.war.naa/ ‘we looked for’ and /jaa.rat.naa/ ‘our neighbor’, two feet are constructed (i.e., one over /daw/ and /jaa/ and another over /naa/). Although two feet are present, foot extrametricality is applied, rendering the suffixed foot peripheral and thus invisible to stress. Stress is then assigned to the rightmost visible foot over the initial syllable by ERR (Figure 2.14, cited in Watson 2007, p.105).

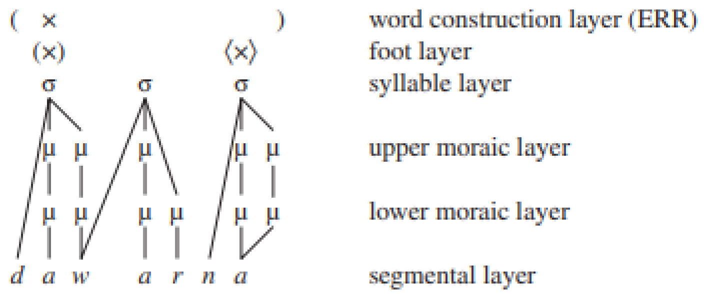


Figure 2.14: Stress assignment in a suffixation example

Watson proposed that sub-structural varieties of superheavy syllables, such as CVVC and CVCC syllables, are not prosodically similar. Parallel observations were present in Aoun (1979) for Lebanese Arabic, Selkirk (1981) for Cairene Arabic, Broselow (1992) for dialectical Arabic, and Farwaneh (1992) for JA. Superheavy syllables across dialects in the forms of CVVC and CVCC structures are argued not to behave similarly and are not theoretically accounted for comparably. They demonstrated that a CVVC syllable is recognized as a single superheavy syllable whereas a CVCC syllable is a canonical syllable plus a degenerate syllable (Broselow 1992). Then, CVVC syllables are expected to occur more frequently in non-final positions than CVCC syllables.

Broselow (1992) examined Sudanese, Syrian, Iraqi, Lebanese, and Gulf dialects. The data showed that CVVC syllables could be found in derived word internal environments, including subject suffixes (CVVCVC, /kitaab+na/ ‘our book’) and possessive suffixes (CVCVVC, /maask+iin/ ‘we are holding’). CVVC syllables in MSA are formed in active-participle environments forming geminates, such as /maarruun/ ‘passer-by’ and /maadduun/ ‘stretching’. Alternatively, CVCC syllables do not appear in word non-final positions because they are usually restructured by vowel epenthesis (e.g., /kalb/ > /kalib/ ‘dog’).

Crossley (2023) examined superheavy syllables in JA adult speech. His analysis employed the contrast between bimoraic and trimoraic syllables using affixation as a variable in productions.

The objective was to provide compelling evidence for trimoraic syllables in JA, challenging the trimoraic ban active in Levantine dialects (Broselow et al., 1995, Khattab & Al-Tamimi 2014). Vowel and coda consonant durations in superheavy syllables were reported in two environments: trimoraic CVVG and non-final CVCC syllables. The study examined variables such as emphasis, gemination, and stress in superheavy syllables to record the durational variations based on moraic and non-moral-related effects. For emphasis, results showed no evidence for the durational difference between plain /sɑ:d/ “ruled” and emphatic geminates /sˤɑ:d/ “hunted”. Vowel durations were not statistically significantly different in stressed and unstressed syllables in the pair /ga.sˤad/ and /ga.sˤad.ha/, where the stress falls on /ga/ in the first word and /sˤad/ in the second. Crossley further investigated the difference in vowel and coda consonants in bimoraic and trimoraic CVCC syllables in word medial and final positions. The study revealed that position affected durations as shorter non-final CVCC syllables were produced when the syllable had an affix. The results showed that coda shortening in CVCC and CVVG syllables occurred, but the consistent coda shortening regardless of consonantal length and lack of long vowel shortening led to the conclusion that phonetic effects of a trimoraic ban were manifested.

2.2. Language Development

This section explores the literature on phonetic and phonological development in child speech perception and production. For the production patterns, language universals, articulatory learning, the stages of linguistic development, and lexical stress development are discussed. A thorough review of Arabic studies examining the phonological patterns and processes is conducted.

2.2.1. Phonological and Phonetic Development in Infancy

2.2.1.1.Perceptual Development in Infants

Infants exhibit early perceptual attunement to aspects such as consonant and vowel contrasts, rhythmic patterns, word boundary detection, intonational properties, and prosody (Eimas et al., 1971, Kuhl 1991, Best 2017). In the first year of life, infants' perceptual sensitivity shifts from language universals to a narrowing focus on language-specific categorizations (Vihman 2015, Best 2017). Infants are described as 'citizens of the world', where they exhibit sensitivity to discriminations between sounds across languages by six months of age (Kuhl 2004, Werker & Tees 1984). However, by twelve months, this language-universal discrimination sensitivity decreases as infants' language-specific phonetic ability improves, turning them into 'culture-bound listeners' (Eimas 1975, Werker & Tees 1984, Kuhl 1979, 2001, 2004).

Perceptual sensitivity refers to the ability to detect and discriminate acoustic differences in speech sounds during infancy and early childhood (Eimas et al., 1971, Kuhl 1996). By six months, infants are perceptually attuned to the vowels in their ambient languages, and by twelve months, they show an increased sensitivity to consonant contrasts (e.g., English /r/ and /l/ and English voiced /b/ and voiceless /p/, Vihman 2015, Best 2017). By six months, infants' attunement to the acoustic properties of their ambient languages is apparent (Eimas et al., 1971, Kuhl 1991). For example, English native infants exhibited the ability to discriminate between voiced and voiceless stop consonants through an acoustic difference of 20 ms in voice onset time VOT (Eimas et al., 1971, Tsao, Liu & Kuhl 2004).

To account for infants' ability to discriminate and categorize language-specific phonemes, multiple models were proposed, such as Native Language Magnet (Kuhl 1991, 2008), Perceptual Reorganization (Werker et al., 1981), Perceptual Assimilation Model (Best 1994, 1995), and Natural Referent Vowels (Polka and Bohm 2003, 2011). Highlighting a few, Kuhl (1991, 2008) proposed in her Native Language Magnet Approach that infants go through a critical period. This period entails the exposure to language-specific patterns resulting in the "brain's initial state of universal sensorineural responsiveness to the acoustic features" to be narrowed down from language universals to language specifics (Best 2017, p.476). Thus, processing language-specific distinctions is only achieved when "the brain's auditory phonetic neutral circuitry becomes neurally committed" to language-specific prototypes (Best 2017, p.476). Therefore, prototypes (central representations of linguistic categories, Kuhl 2004) function as magnets that attract surrounding sounds, mainly vowels and prosodic features (stress, pitch, and intonation), initiating the intonational and prosodic structures. Lively (1993, p. 2423) expanded on the prototype proposition by clarifying that prototypes tend to attract other stimuli more strongly than non-prototypes. Thus, vowels that surround a prototypical stimulus are more difficult to discriminate than vowels that surround a non-prototype". Then, infants show a prototypical sound development where distinctive properties of language-specific units are mapped neuro-psychologically, reflecting prototype learning in cognitive psychology (Guenther & Gjaja 1996, Kuhl 2004).

Werker et al. (1981) argued that 10- to 12-month-old infants show signs of perceptual reorganization. Perceptual reorganization refers to the cognitive process through which the infant's perception of speech sounds becomes increasingly tuned to the language-specific phonetic distinctions of the ambient language (Werker et al., 1981). For example, Mattock & Burnham (2006) investigated perceptual reorganization for tones in 6 to 9-month-old infants from tonal, Chinese, and non-tonal English environments for speech and non-speech discrimination. Their results suggested that Chinese infants performed better in both speech and non-speech tone discriminations. Nonetheless, English infants' discrimination of lexical tone declined between 6–9 months of age. They stated that the reorganization of tone perception is a function of the ambient language environment. Additionally, Werker and Tees (1984) demonstrated that between 10–12 months of age, English-speaking infants could initially discriminate non-native sounds, such as Hindi and Salish consonant contrasts. However, by the same age, Hindi and Salish-speaking infants showed higher sensitivity to their respective native contrasts, indicating a tuning of their perceptual abilities to their ambient language. Werker and Tees (1984) postulated that the declining ability of English-speaking infants to perceive contrasts in other languages is due to their lack of exposure to those non-native sounds rather than a permanent loss of their capacity for sensorineural discrimination of linguistic phenomena (Best 2017, Kuhl 2004).

Multiple studies suggested that prosodic and phonotactic properties of speech are evident in infants' linguistic preferences (Bion, Benavides-Varela & Nespor 2011). During the first year, infants can recognize prosodic patterns (i.e., stress patterns, pitch contour, and phrasing), and they gradually develop their perceptual sensitivity to supra-segmental features to adapt to language-specific structures (Abdoh 2010). This turning point from speech perception universality to specificity is depicted as evidence for infants' attunement to language-specifics motivated by the narrowing preference for "the prosodically enhanced registers" (Fernald 1985) and "for the stress patterns of the ambient language" (Kuhl 1979, Jusczyk et al., 1992, Kuhl 2001, Vihman 2015).

There is an early discrepancy in stress pattern perception, by which infants begin to develop language-specific perceptual patterns within the first year of life (Jusczyk, Houston & Newsome 1993, Skoruppa et al., 2009). Distinguishing stress patterns varies among infants learning fixed and free-stress languages (Skoruppa et al., 2009). Studies demonstrated that by nine months of age, infants learning languages with fixed stress patterns, such as Spanish, exhibit the ability to discriminate stress patterns in segmentally varied stimuli[7] (Pons and Bosch 2010). In contrast, infants learning languages with free stress patterns, such as French, may show difficulty in this discrimination task yet differentiate stress patterns in repeated nonword task (Skoruppa et al., 2011).

Infants' preferences for stress patterns are assumed to be based on language-specific exposure. For instance, nine-month-old English infants tend to prefer stress-initial disyllabic words (e.g., ‘pliant’, ‘falter’) over stress-final ones (e.g., ‘comply’, ‘befall’), a preference not observed in Spanish-learning infants (Jusczyk, Houston & Newsome 1993, Pons & Bosch 2010, Skoruppa et al., 2011). This distinction is attributed to the higher overall frequency of stress-initial disyllables in English compared to Spanish (Jusczyk, Houston & Newsome 1999). Moreover, it is suggested that infants exhibit a bias for the predominant stress patterns of words they hear frequently (Jusczyk, Houston & Newsome 1999). This was evident in English learning infants demonstrating a preference for words with strong/weak stress patterns over weak/strong ones (Jusczyk, Cutler & Redanz, 1993, Morgan 1996). This sensitivity emerges between 6 and 9 months of age, suggesting a developmental trajectory in the perception of stress patterns and word boundaries in English (Jusczyk, Houston & Newsome 1999).

Furthermore, prosodic sensitivity is assumed to include the distributional phonotactic properties of the sounds in the ambient language (Kuhl 1993, Jusczyk, Luce & Charles-Luce

---

[7] Segmentally varied stimuli refer to linguistic inputs where consonants and vowels change while stress patterns remain consistent (Pons & Bosch 2010)

1994). This sensitivity is evident in the infant’s ability to distinguish between more and less frequency phonetic and phonological sequences by nine months of age (e.g., CV syllables are more frequent than CCV ones) (Kuhl 1996, Vihman 2018). Exposure to related linguistic stimuli assists in ‘categorical language learning’ which involves acquiring the phonotactic patterns and the rules governing them (e.g., Arabic-learning infants can discriminate between different syllabic patterns such as /baa/ (CVV) and /baba/ (CVCV) 'father') (Eimas 1975, Kuhl 1979, Lively 1993, Maye, Werker & Gerken 2002, Vihman 2015).

2.2.1.2. From Prelinguistic Vocalizations to Early Words

Typically developing infants naturally progress through several stages of speech development (Eimas 1975). At the cooing stage, until three months of age, infants produce non-verbal sounds and by four months, they produce vowel-like sounds (Maye, Werker & Gerken 2002, Tsao, Liu & Kuhl 2004, Stoel-Gammon 2011). With the development of the vocal control mechanism, infants’ linguistic productions become more diverse (Jusczyk, Luce & Charles-Luce 1994). They demonstrate evidence of sensory-motor learning, exploring various vocalizations and experimenting different sounds (Liu & Kuhl 2004). As they engage in vocal play and produce a wide range of non-verbal sounds (i.e., cooing, gurgling, and laughter), their vocalizations do not yet represent meaningful speech before the canonical babbling stage from 5 to 10 months of age (Ferguson, Menn & Stoel-Gammon 1992, Jusczyk, Luce & Charles-Luce 1994, Abdoh 2010).

Observational studies examined how the emergence and progression of babbling coincide with the broader development of the infant’s speech production and language skills (Stoel-Gammon 1985, McCune & Vihman 2001). Studies suggest that babbling constitutes a key step for linguistic production for two main reasons. First, babbling enables the child to master articulatory gestures by constructing a sensorimotor representation aided by an ‘auditory feedback loop' (Fry 1966, Davis et al., 2000, Thelen 1991, Vihman et al., 2013). This means that as infants babble, they hear the sounds they produce and use this feedback to refine their articulatory movements, thus developing a mental representation of these specific sounds (McCune & Vihman 2001, Stoel-Gammon 2011). Then, the feedback process is associated with perceptual representations, where the child's ability to categorize and perceive speech sounds contributes to their learning and production (Tsao, Liu & Kuhl 2004, Vihman 2015).

Second, the infant's vocal behaviour is assumed to increase the attention of adults interacting with the child (McCune 1992, Majorano, Vihman & Depaolis 2014). Babbling often triggers an immediate interactive parent/caregiver response, described as 'proto-conversations' (Veneziano 1981, Tsao, Liu & Kuhl 2004). With the increased attention and interactive response from the parent/caregiver, infants seem to modify their productions according to perceived phonological patterns and vocalizations (Vihman 2014). The temporal patterns observed in their canonical babbling initially mirror the patterns found in the speech input they perceive (Tsao, Liu & Kuhl 2004). Consequently, the production of temporal patterns resembling adult-like ones is argued to precede the production of phoneme and syllable-sized units (Majorano, Vihman & Depaolis 2014).

On the other hand, there are debates on the extent to which babbling is indicative of early word production (McCune & Vihman 2001). Some studies indicated that infants who show advanced levels of preverbal productions (i.e., producing more babbling vocalizations) demonstrate greater stability, accuracy, and syllabic complexity in their productions (Vihman & McCune, 1994, Majorano & Vihman 2009, Majorano, Vihman & Depaolis 2014). Increased babbling vocalization is associated with increased vocal-gestural practice that results in mastery of the ambient language (Stoel-Gammon 1985). Some scholars extend this argument to discuss how babbling contributes to the emergence of consonantal preferences at 10–12 months of age

(Werker & Tees 1984, Maye, Werker & Gerken 2002). Furthermore, heightened sensitivity to distinctions and preferences between sounds is linked to an enhanced ability to detect acoustic cues such as pitch, intensity, and duration (Maye, Werker & Gerken 2002, Kuhl 2004). This allows infants to differentiate between phonetic distinctions and understand distributional patterns more effectively (Vihman & McCune, 1994, Majorano, Vihman & Depaolis 2014).

As infants become increasingly attuned to the features of their ambient language (Tobin 1997, Lust 2006), their vocalizations and babbling seem to lead to significant developmental milestones. Normally developing infants, aged 12–18 months, begin to produce single-word utterances (Vihman & McCune 1994). The progression towards language-specific features establishes the basis of lexical development, a phase referred to as the holophrastic phase[8] (Vihman 1993, Vihman & McCune 1994, Abdoh 2010, Menggo 2017). During the holophrastic stage, children resort to phonetic patterns recognized during the first postnatal year and the babbling stage to produce 25+ spoken words (Jusczyk, Houston & Newsome 1993, Vihman 1996, Lust 2006). By 18–24 months, children are expected to have acquired a total of 50-100+ items of expressive vocabulary displaying ambient language phonological rules and systematic contrasts (Tobin 1997).

At this stage, the child’s ability to produce various syllabic structures shapes the prosodic framework of their lexicon (Stoel-Gammon & Williams 2013). By the end of the first year and into the second year, children shift from holophrastic components to more stable canonical forms influenced by adult patterns and their developing motor skills (Kuhl 1991, Thelen 1991, Vihman & McCune 1994). This progression is marked by an increase in the production of canonical word shapes, such as CVC, CV/CV, and CVC/CV, reflecting improved accuracy in

---

[8] Anwar, a 13-old-month child, produces the word /ʔimbuu/ ‘water’ serving a communicative function and used in a holophrastic sense where the child meant ‘I want water’ or ‘give me water’ (Salim & Mehawesh 2014)

generating adult-like word forms (Vihman 1993, Abdoh 2010). Initially, children's early word forms are constrained by simpler syllabic patterns such as CV and CVC, contrasting with the more complex structures in adult speech, such as consonant clusters (Ferguson 1977, Tobin 1997). As they progress in their development, children gradually transition to more complex syllabic patterns with emerging structures, such as CV, CVC, CVV, and CVVC, becoming more frequent, thereby showing increased phonological correspondence with adult forms (Vihman & Croft 2007).

2.2.2. Language Universals and Articulatory Motor Learning Perspectives

Two major approaches are highlighted in this section: innateness and motor learning. They offer contrasting yet complimentary explanations for language acquisition and development. The first suggests the universality of language development as children are hypothesized to be born with innate linguistic knowledge that is activated by the input (van den Berg 2012). The latter posits that acquisition occurs through exposure to linguistic input and subsequent learning mechanisms (i.e., cognitive processes extracting and internalizing linguistic patterns and structures over time) (Kent 1976, 1992, Clark & Casillas 2015).

2.2.2.1. Language Universals

Phonological acquisition and development theories aim to provide universally valid explanatory concepts for child speech (Alqattan 2015). Similarities across languages are emphasized by these theories; however, there is much debate on which aspects of speech development are universal. Jakobson's (1968) structuralist approach to language acquisition postulated predictions based on a universal hierarchy of structural laws governing phonetic systems. The analysis proposed that sounds would be acquired early when they are more distributed among the world's languages. Jakobson (p.35) claimed the following developmental patterns: (1) babbling is unrestricted and does not predict a child's later acquisition of adult phonology; (2) phonological development is best described in the mastery of distinctive features; (3) the child does not approximate the adult's phonemes one by one; instead, they establish their system of phonemic contrasts, not always using the same features as adults to distinguish between words; (4) phonological development in children is systematic and universal. Jakobson (1968) suggested a universal order of syllable acquisition, following a predictable pattern across languages. This order progresses from simple to more complex linguistic structures with the first syllable structure to develop being a consonant-vowel structure, CV and CVCV followed by CVC. Moreover, Jakobson highlighted that children acquire back consonants only after they acquire front consonants; fricatives only after they acquire stops; affricates only after acquiring stops and fricatives; and nasal vowels after corresponding oral vowels (Tobin 1997).

Another prominent approach is the nativist approach to language acquisition associated with Chomsky (1965). It proposes that humans are inherently equipped to learn a language with the language acquisition device which is an innate mechanism enabling children to grasp the grammar of their ambient language. Accordingly, children have knowledge of their language as they naturally discriminate linguistic patterns from the language they are exposed to (Abdoh 2010). Chomsky argued that linguistic knowledge is acquired based on certain innate principles called 'universal grammar' defined as "a certain fixed language-independent schematism that determines what counts as linguistic experience and what knowledge is acquired, what grammars are constructed, on the basis of this experience" (p.319). Thus, this concept implies that all languages share fundamental structural principles and the children's exposure to their ambient language triggers the activation and refinement of these universal principles.

In their work 'The Sound Pattern of English' (SPE, 1968), Chomsky & Halle introduced the Generative Phonology Approach, proposing that phonological systems adhere to universal principles governed by innate cognitive mechanisms. Their basic principle is that children's speech acquisition is achieved by applying a set of phonological rules to abstract underlying forms similar to those of adults. These phonological rules map underlying representations to surface forms accounting for the systematic variations observed in speech sounds. Chomsky & Halle speculated that children accurately perceive and store lexical representations to guide their linguistic output. As development progresses, children gradually assimilate additional features and unlearn earlier patterns leading to the formation of the SPE feature system. The SPE system classifies sounds primarily on articulatory aspects of speech, explaining phonological alternations and patterns across languages. The mapping of phonological components of SPE is presented in Figure 2.15 (Dresher 2004, p.6).

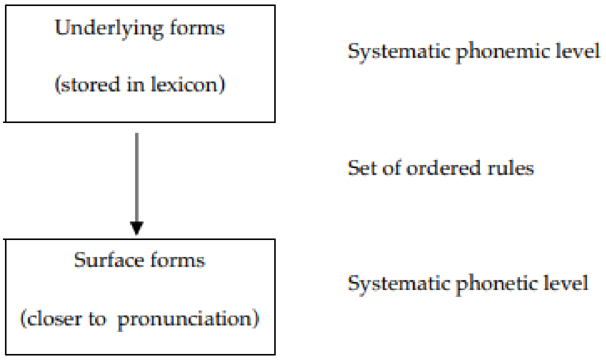


Figure 2.15: Phonological components of SPE

Moreover, multiple studies focused on the universal development of syllable structure and early word shapes. Two main models that discuss early speech development emerged during the 1990's including Fikkert's (1994) parametric theory of syllable structure and Demuth & Fee's (1995) theory of children's early word shapes based on the minimal word concept. Both models predict the rate of developmental aspects of the rhyme and syllable size in early acquisition.

Focusing on Dutch, Fikkert hypothesized syllable type acquisition has a syllable parameter that develops from unmarked to marked[9] units of input language (van den Berg 2012). Fikkert

---

[9] The universal linguistic patterns are classified into marked and unmarked structures. Unmarked structures are common, frequently used across the languages, and appear in the early stages of

suggested that children tend to acquire final consonants before they differentiate between vowel lengths, where the relation of vowel length to weight is constrained by length contrasts being only evident in closed syllables. She proposed specific syllable parameters comprising four stages of children's rhyme development as follows:

1. Stage I: children have set no syllable parameters where they only produce core syllables comprising a consonant and a vowel. Vowel length is not contrastive at this stage although short and long vowels are produced, thus vowels are represented with a non-branching nucleus.
2. Stage II: children set the branching rhyme parameter to include a nucleus and a coda, and obstruents appear before sonorants in the syllable final position.
3. Stage III: children establish the branching nucleus parameter, allowing for the representation of short and long vowels. Acquiring vowel length representation is attested before sonorants. This was evident in children shortening long vowels appearing before sonorants and lengthening short vowels when sonorants were deleted. However, no length relationship was found in the case of obstruents.
4. Stage IV: a bipositional rhyme is allowed accommodating an additional consonant. This highlights the extrathymal parameter achieved by children at this stage. Vowel length representation is acquired before obstruents.

Demuth & Fee’s (1995) model proposes that core syllables are the first prosodic forms produced by children in the first stage of development, where neither coda consonants nor vowel length contrasts are implemented. This is followed by the minimal word stage, where productions are minimally and maximally bimoraic. This stage comprises three sub-stages, with the first encompassing the tendency to reduplicate monosyllables or epenthesize syllables,

---

development (e.g. CV- syllable, Clements & Keyser 1983). Marked structures are less frequent and tend to appear in later stages (e.g. onset clusters CCVC, Fikkert 1994, Zamuner 2003).

wherein the minimal word is satisfied by a bisyllabic CVCV form (Demuth & Fee 1995). The second sub-stage includes the acquisition of coda consonants (CVC) followed by the third sub-stage comprising the acquisition of vowel length contrasts (CVV). Demuth and Fee observed that the minimal word is the simplest and most common prosodic form found across languages. This form provides a constrained learning environment for children, allowing them to learn the specific syllable structures of their language while also producing well-formed prosodic words (Kehoe & Stoel-Gammon 2001, p. 401).

2.2.2.2. Articulatory Learning

The question of when and to what extent child productions become more adult-like has received much attention (Tingley & Allen 1975, Sharkey & Folkins 1985). The development of temporal control in child speech determines the improvement in phonological and phonetic consistency occurring with age (Tingley & Allen 1975). Previous studies showed two major factors attributed to motor learning and exposure which contribute to the development of child productions.

First, the developmental patterns evident in children's speech productions are influenced by their articulatory coordination and motor skills. With maturation, children maintain a higher vocal tract articulatory movement control compared to younger children reflected mainly in their durations (Kent 1976, 1992, Green & Wilson 2006, Nip & Green 2013). To illustrate, Smith (1978) investigated the developmental aspects of temporal parameters in the speech of English-speaking children ranging from 2;6–4;6 years of age. Smith’s analysis addressed that child productions are significantly different and longer than adults as younger children exhibit less refined neuromotor capabilities. Thus, children were not able to exhibit adult-like control of their speech mechanisms. Smith added that durational patterns are associated with language-specific constraints, such as syllable structure, syllable position, and stress rules. For instance, durational patterns in French infants during the babbling stage (age range of 8–14 months) were influenced by phonological parameters such as syllable position and structure. Canault et al. (2020) argued that although French infants produced longer word-final syllables compared to their counterparts, syllable structure did not exhibit a significant effect on durations. They suggested that the the developmental trajectory of temporal control in speech was not linear, as the advancement of motor abilities is accompanied by the integration of more complex linguistic forms with age. This finding indicates that while motor skills develop, the interaction with linguistic complexity becomes more prominent, highlighting the multifaceted nature of language acquisition.

Within this line, Robb & Saxman (1990) investigated the continuity in the development of syllable duration patterns in seven English-speaking children as they progressed from pre-word to multi-word periods of vocalization development. They emphasized how final syllable lengthening is a manifestation of syllable structure and stress assignment (Canault et al., 2020). Robb & Saxman suggested children exhibit a temporal pattern of word-final syllable lengthening associated with age-related articulatory development, with younger children producing longer word-final syllables compared to older children (i.e., production performance advances due to the elevation of motor control[10]) yielding a progressive declining pattern in durations.

Second, language learning and experience affect the patterning of linguistic productions (Kent & Forner 1980, Smith, Sugarman & Long 1983). Children develop their production patterns

---

[10] Motor learning includes: more controlled coordination (Nip and Green 2013), increased rate and amplitude of jaw and lop movements (Nip, Green & Marx 2011), and variability rate in orofacial movements i.e. lips, tongue and jaw (Green, Moore & Reilly 2002, Smith & Zelaznik 2004, Canault et al., 2020).

and processes not only through maturation but also through experience, resulting in shorter durations (Smith, Kenny & Hussain 1996, Payne et al., 2012). Correspondingly, Nip and Green (2013, p. 1) stated that learned skills such as "motor learning, semantic and lexical maturation, phonological access, and motor planning and programming" contribute to the development of target units. Thus, with maturation, linguistic learning and experience are expected to contribute to phonological development and by extension to phonetic development.

For example, cross-linguistic studies documented that the emergence of word-final lengthening and prosodic marking are reflective of language-specific experience and age-related motor skills (Delattre 1966, Oller & Smith 1977, Smith 1978, Robb & Saxman 1990). While this lengthening is consistent with adult productions in some languages, such as French (Pollock Brammer & Hageman 1993, Snow 1998b), discrepancies may arise in others, such as Japanese and Mandarin Chinese (Ota 2001, Ota et al., 2016). Younger children may exhibit greater word-final lengthening reflecting their awareness of prosodic features of their ambient language (s) (Robb & Saxman 1990). Nonetheless, their phonological and motor skills do not demonstrate motor control, creating greater discrepancies between word-final and non-final syllables that tend to be reduced with age (Kent 1976, Zec 2007). Word-final lengthening then provides insights into the interaction between language-specific experience/motor advancement in phonological development.

2.2.3. Phonological Development in Arabic

The literature focuses on early developmental stages in European languages, such as English, Dutch, and Spanish, while less attention is devoted to Arabic or children of two years of age and above (Stoel-Gammon 1992, Dodd & Gillon 1997, Stoel-Gammon & Williams 2013). Since the linguistic repertoire advances with maturation, the assessment of child phonology expands drastically to include consonant accuracy, syllable structure, stress patterns, and phonological processes (Menn 1971, Ferguson & Farewell 1975, Amayreh and Dyson 1998, Fikkert 1994, Dyson and Amayreh 2000, Ammar 2002). Aspects such as articulatory and perceptual development, articulatory challenges, details on early acoustic sensitivity, and categorical perception capabilities have not been well documented for Arabic-speaking children. So far, some scholars such as Omar (1973), Amyareh (1996), Amayreh & Dyson (1998), Dyson and Amayreh (2000), Amayreh and Dyson (2000), Saleh et al. (2007), Abdoh (2010), Ayyad (2011), Alqattan (2015), and AlAjroush (2019) studied normal phonological development.

2.2.3.1. Phonological Development in Arabic Dialects

Saleh et al. (2007) investigated the phonemic inventory and phonological processes in Egyptian Arabic children with an age range of 12–30 months. Thirty Cairene Egyptian-speaking children were recruited and divided into three age groups with a 6-month interval. One-hour tape recording for each child was collected and data were analysed impressionistically. One of the major aims of Arabic studies is to document at which stage phonemes are acquired to assess when the child productions start approximating adult-like ones. The results revealed that in Cairene child speech, and before the age of 30 months, six stops (/b/, /t/, /d/, /k/, /g/, /ʔ/), nine fricatives (/s/, /z/, /ʃ/, /f/, /x/, /ɣ/, /h/, /ħ/, /ʕ/), two liquids (/l/, /r/), two nasals (/m/, /n/), two glides (/w/, /j/), non-Egyptian phonemes (/v/, /θ/), and one emphatic /dˤ/ appeared. The authors suggested that the phonemes acquired early in Egyptian child speech (e.g., /b/, /t/, /d/, /k/) almost matched those attested in English (Smit et al., 1990), and in JA (Amayreh & Dyson 1998). However, some phonemes, such as /l/, /s/, /j/, and /ʔ/, were acquired earlier in Egyptian Arabic compared to English and JA due to their prevalence in adult Egyptian language.

They documented the main and sub-phonological processes to compare the child and adult productions. These included syllable structure processes (e.g., weak syllable deletion), substitution (e.g., glottal replacement), and assimilation (e.g., regressive contiguous).

Employing syllable structure and substitution phonological processes were more frequent than assimilation processes, with a notable trend of a gradual decrease in the use of syllable structure processes with age. The authors concluded that this trend was indicative of children relying mainly on syllable structure processes during the first six months. These included attempts to reproduce CV syllables used in reduplicated babbling or to reduce the number of syllables in target words to match the articulatory capacity.

Studies of Arabic phonological development lack adequate information on the extent to which factors such as feature frequency and phonological salience[11] influence the early stages of speech acquisition. This motivated Alqattan (2015) to trace how Kuwaiti Arabic child acquisition is influenced by salient aspects of the ambient language. A total of 70 typically developing Kuwaiti children were recruited with an age range of 1;4 to 3;7 years. Spontaneous speech recordings of utterances were recruited as participants were playing with toys or while viewing picture books with parents/caregivers. Data was manually transcribed and evaluated based on frequency calculations, production accuracy, and phonological error patterns (i.e., child productions deviating from adult productions). Consonant production accuracy was traced with an emphasis on the influence of type and token frequencies[12] on the order of acquisition and development of error patterns. The consonant production accuracy was calculated based on the following criteria: (1) Mastery production: when a sound was produced accurately in at least 90% of attempted targets; (2) Acquisition production: when a sound was produced accurately in at least 75% of attempted targets; (3) Customary production: when a sound was produced accurately in at least 50% of attempted targets.

---

[11] Phonological saliency is based on the diagnostic measure, sonority index that accounts for the oral cavity opening and voicing propensity (Rice 1999, Hume 2006), with more sonorous sounds being easier to be perceived and therefore, acquired (Pye, Ingram & List 1987, Mowrer & Burger 1991).

[12] Input frequency is divided into type frequency: relating to the segment's occurrence in unique words, and token frequency: reflecting the raw number of exposures to the segments (Ferguson & Farwell 1975, Mehler et al., 1988).

Kuwaiti children were able to produce more than half of the consonants with 75% accuracy before the age of 3;7 years including /p/, /b/, /t/, /d/, /k/, /g/, /ʔ/, /m/, /n/, /f/, /s/, /w/, /l/, /ɫ/, /r/, /z/, /ʃ/, /x/, /ħ/, /ʕ/, /h/, /j/, /ʤ/, /ʧ/, /tˤ/, /sˤ/, /q/, /ɾ/, /ɣ/, and /ðˤ/. However, /ŋ/, /v/, /θ/, /ð/, /ʒ/, /dˤ/, and /zˤ/ were not acquired by the oldest group, 3;4–3;7 years (Based on the study's criteria: these consonants were produced accurately in less than 50% of the total number of occurrences). Similar to Egyptian and Jordanian-speaking children, Kuwaiti children produced stops, fricatives, and nasals more frequently than other sounds. Notably, there was a discrepancy between the acquisition of Kuwaiti consonants and their frequency due to reasons such as word length and structural differences (i.e., consonants in prevocalic positions may be more salient than those in postvocalic positions), and the child's linguistic preferences (i.e., children avoid difficult sounds limiting the range of words produced). For example, Alqattan reported that since the voiceless alveolar stop /t/ was produced with a higher frequency (6%) than /k/ (4%), then /t/ is expected to be acquired before /k/. Nevertheless, the production accuracy analysis showed that /k/ was mastered (90% accuracy between the ages of 2;0–2;3) earlier than /t/ (93% accuracy between the ages of 3;4–3;7). Alqattan highlighted that syllable shape and complexity produced by children increased with age. For syllable shapes, CV, CVC, CVVC, and CVCC syllables were the most frequently produced. Onset clusters in CCV and CCVC syllables and coda ones in CVCC syllables appeared as early as 1;4–1;7 years. However, CCVCC, CVVCC, and CCVV were the least frequent and they did not appear before 3;0–3;3 years. In terms of word length, disyllabic words were the most produced followed by monosyllabic and trisyllabic words.

Alqattan divided phonological errors into two categories, segmental (affecting place, manner, and voicing) and prosodic (affecting syllables and word shapes). According to the number of error patterns divided by the total target words, the author suggested that de-emphasis, stopping and de-affrication were the most frequent segmental errors with 36%, 13%, and 7%, respectively. Nonetheless, coda deletion and cluster reduction were the most frequent prosodic errors with 8% and 6%, respectively. Alqattan claimed that existing phonological concepts, such as markedness, functional load[13] , and feature hierarchies did not account for some of the patterns attested in her study reflecting due to language-specific characteristics and cross-linguistic differences of reported findings in the literature.

Ayyad (2011) documented the development of the phonological skills of 80 typically developing Kuwaiti Arabic children in the age range of 3;10–5;2 years. The study adopted the non-linear phonological[14] framework for data analysis. A standard single-word picture- and object-based elicitation was used to evaluate consonants and vowels across word positions within a variety of word lengths and structures. Ayyad presented her findings descriptively, focusing on the segmental inventory and error tokens, without employing statistical analysis. The mastery criterion of 90% was set by Ayyad, unlike other Egyptian and Jordanian studies that set it at 75%, while 75% was set for the consonant acquisition level. The analysis showed that the younger group mastered (/b/, /t/, /t:ˤ/, /d/, /k/, /g/, /qː/, /ʔ/, /m/, /n/, /ðˤ/, /ħ/, /h/, /x:/, /tʃ/, /rː/, /w/, /j/), and acquired (/tˤ/, /q/, /sːˤ/, /ð/, /ʃ/, /ʁ/, /χ/, /ʕ/, /l/). Nonetheless, consonants /s/, /sˤ/, /θ/, /z/, /dʒ/, and /r/ were not acquired to a 75% level, and the features not acquired were Coronal [+grooved], [-grooved], [-cont]-[+cont] (affricate), and [+trilled]. The features acquired to a level of 75%-89% included [Pharyngeal], [Dorsal, -high, -low] (uvular), and [+lateral]. The older group mastered (/s/, /b/, /bː/, /t/, /d/, /tˤ/, /tːˤʰ/, /k/, /g/, /q/, /qː/, /ʔ/, /m/, /n/, /f/, /ðˤ/, /ʃ/, /χ/, /ʁ/, /ħ/, /h/, /tʃ/, /rː/, /l/, /w/, /j/, /jː/) and acquired (/θ/, /ð/, /dʒ/, /ʕ/). The consonants with Coronal [+ grooved] (and [Pharyngeal]) and [+trilled] features (/s/, /sˤ/, /r/) were not yet acquired by 75% of children.

---

[13] Functional load refers to the importance of phonemes in a specific phonological system and it influences the order of phonological acquisition (Hua & Dodd 2006).

[14] Describing the phonological form of words in terms of a hierarchy of phonological elements (Ayyad 2011)

Furthermore, Ayyad highlighted that monosyllabic words included CVV, CVC, and CVCC shapes, disyllabic words included CVCV, CVVCV, and CVCCV shapes, and multisyllabic words had CVCVCV, CVCCVCVC, CVCVCVCV, and CVCCVCVCV shapes. The youngest age group exhibited the acquisition of codas, diphthongs, and coda clusters. However, onset clusters, as in the CCVCV word shape, did not appear in the younger age groups reflecting a high-ranked constraint on complexity. Based on these results, Ayyad suggested that Kuwaiti children showed a more advanced consonant inventory and syllabic patterning at an earlier age compared to Egyptian and Jordanian-speaking children. Nonetheless, the acquisition of the majority of the consonant inventory by Kuwaiti children before the age of five may suggest potential influences stemming from differences in study design, acquisition rate criteria, or dialectical variations compared to other studies.

2.2.3.2. Phonological Development in JA

In the context of Jordan, Amayreh & Dyson (1998) conducted a normative study for the acquisition of consonants in MSA as spoken by Jordanian children. They recruited 180 monolingual Arabic-speaking children with an age range of 2;0 to 6;4 years of age, divided into nine gender-balanced groups. The aim was to find at which age MSA consonants are acquired in word-initial, medial, and final positions. A picture naming test was used to elicit 58 target words. Following Sander (1972), Amyreh & Dyson defined three levels of consonant acquisition: (1) Mastery level (90% correct production in all positions); (2) Acquisition level (75% correct production in all positions); (3) Customary production level (50% correct production in at least two-word positions). Production correctness was measured using standard and acceptable[15] forms caused by dialectical variance. Amayreh & Dyson analysed

---

[15] Acceptable forms are variants of standard consonants such as (Amayreh & Dyson 1998, p.643):
/dˤ/→ /d/, /ð/
/q/→ /k/, /g/, /G/, /ʔ/
/ʔ/→ ø (when word-medial)
/ðˤ/→ /dˤ/

the overall accuracy percentage in all positions combined; the difference in accuracy by position and in sound classes; and gender-based accuracy.

The results showed that there was a clear developmental trend with older children producing more accurate productions, medial consonants were reported to be produced more accurately than initial and final ones, and gender was not reported to affect the acquisition of consonants. Th analysis demonstrated that /b/, /t/, /d/, /k/, /q/, /ʔ/, /f/, /ħ/, /m/, /n/, /l/, and /w/ were acquired by 2;0–3;10 years; /s/, /ʃ/, /x/, /ɣ/, /h/, /r/, and /j/ were acquired by 4;0–6;4 years; and /q/, /tˤ/, /dˤ/, /θ/, /ð/, /ðˤ/, /z/, /sˤ/, /ʕ/, and /d͡ʒ/ were not acquired by the age of 6;4. The authors reported that in the early period (2;0 to 3;10), children acquired at least ten out of 26 consonants in Arabic, including the acceptable forms. In the intermediate period (4;0 to 6;4), most of the fricatives, the affricate, and liquid /r/ were acquired. As for the late period, consonants such as emphatics, which were not acquired by this stage, were expected to be acquired by children older than 6;4 years. Such results supported Jakobson's (1968) universal theory, where Jordanian children demonstrated the acquisition of stops before fricatives and front consonants before back ones. Jordanian children also exhibited the acquisition of voiceless cognates before voiced ones demonstrated in the late acquisition of emphatic sounds.

The relative early acquisition of /ħ/ and /k/ was attributed to their high frequency in Arabic. This was linked to the articulatory learning theory suggesting that recurring sounds are acquired fast (Ingram 1992, 1989b, Pye, Ingram & List 1987). The early acquisition of /l/ in Arabic further supported the notion of consonants appearing quite frequently in a small set of very commonly used words by children. This was evident by Arabic /l/ being acquired a year earlier than English /l/ (Omar 1973). The relative early acquisition of fricative /x/ was attributed to

---

/sˤ/→ /z/ (when word-final)
/θ/→ /t/
/ð/→ /d/
/d͡ʒ/→ /ʒ/

Ingrarm’s (1989a) functional load hypothesis as the significance of a phoneme is determined by the number of contrastive oppositions it has. Arabic /x/ contrasts with fricatives, stops, and sonorants in all positions. Moreover, it was noted that auditory saliency might have played a role in the early acquisition of /x/. Amayreh & Dyson explained that /x/ appears to be an easy phoneme to recognize as it is characterized by a low-frequency concentration of acoustic energy compared to alveolar and palatal fricatives.

In a follow-up of Amayreh & Dyson’s (1998) study, Amyreh (2000) conducted further analysis of the acquisition of MSA consonants that had not been acquired by the age of 6;4 years. The study investigated what age JA children acquire these consonants in relation to their accuracy and error patterns according to the position within a syllable. A total of 60 Arabic-speaking children (age range of 6;6–8;4 years) from Jordan were recruited. The study focused on emphatic consonants produced with a secondary articulation (i.e., the root of the tongue is retracted into the pharynx, Mitchell 1990) including /q/, /tˤ/, /dˤ/, /θ/, /ð/, /ðˤ/, /z/, /sˤ/, /ʕ/, and /d͡ʒ/. Amayreh highlighted that emphatic consonants affect adjacent front vowels resulting in a farther back production. For example, the long vowel /a:/ is produced as [ɑ:] after an emphatic consonant (e.g., [ta:b] repented [tɑ:b] recovered, p.518). These consonants were examined due to their relatively “low functional load”[16] in Arabic (Pye, Ingram, & List 1987), and they are commonly replaced by more frequently occurring consonants or dialectical variants in child productions. In markedness terms, Amayreh suggested that these marked consonants are acquired later than their less marked substitutions as summarized below (p.519):

1. A consonant with secondary articulation is replaced by a similar consonant with no secondary articulation: /tˤ, dˤ/ → [t, d], /q/ → [k, g, ʔ], /ðˤ/ → [ð], /sˤ/ → [s].
2. A fricative is replaced with a stop: /ð, θ, ðˤ/ → [t, d, dˤ].

---

[16] Low functional load is where the phoneme in this case has a minimal impact or a limited significance as a linguistic feature (King 1967, Ingram 1989).

3. An affricate is replaced with a fricative: /d͡ʒ/ → [ʒ].
4. A voiced consonant is replaced with a voiceless cognate: /ʕ/ → [ħ]

Results showed that /dˤ/, /q/, /z/ and /ʕ/ were acquired by age group 6;6–7;4 years. By 7;8–8;4, /tˤ/ met the criterion for the acquisition (produced with at least 75% accuracy). However, /ð/, /θ/, /ðˤ/, /sˤ/, and /d͡ʒ/ did not meet the criterion by even the oldest children (>8;4), although they were produced in an acceptable form by all children. The author suggested that these consonants simply require a degree of articulatory precision not yet available in children before 7 or 8 years of age. Amayreh emphasized that the stop and affricate consonants demonstrated a higher accuracy in the onset positions, whereas fricatives were either more accurately produced in coda positions or approximately equal in accuracy in the two positions. The study illustrated that although these sounds have not been acquired yet, a child in first or second grade should not be suspected of articulation or phonological disorder. A clear distinction between acquisition and mastering should be considered, by which multiple productions had acceptable forms of target consonants constituting dialectical variations, such as the production of /q/ as /ʔ/ and /d͡ʒ/ as /ʒ/. Amayreh suggested that this notion is particularly evident in female participants due to females adapting to the urban linguistic variants more frequently than males.

Furthermore, Dyson & Amayreh (2000) highlighted the articulation/phonological processes and sound changes in 50 normally developing Arabic-speaking children located in Amman with an age range of 2;0–4;4 years. A 58-word picture-naming articulation test was designed to elicit single-word responses representing the initial, medial, and final consonants of Educated Spoken Arabic (ESA). ESA refers to the variety used by educated Arabic speakers and it is commonly used by the media (Zughoul 1980). The results demonstrated that the youngest children produced more than 40% of consonants with modifications including errors (43.7%)

or changes from ESA[17] (46.6%). By 4;0–4;4 years, the aforementioned percentages decreased by about one-half. In the youngest age group, the emphatic stops /q/, /dˤ/, and /tˤ/, the emphatic fricatives /ðˤ/, and /sˤ/, the dental non-emphatic fricatives /ð/ and /θ/, and the /r/ met the most difficult criterion. The most difficult criterion represented sounds whose productions involved changes from ESA or errors more than 75% of the time. The stops /b/, /t/ and /ʔ/, the nasals /m/ and /n/, and the back fricatives /ħ/ and /x/, and /l/ were produced with changes from ESA constituting less than 25% of the time. As for the oldest age group, only /q/ met 'the most difficult' criterion, whereas half of the consonants matched the ESA targets with little difficulty (< 25%). Nasals and non-emphatic stops were reported to be the most accurate manners produced, followed by fricatives and glides, while emphatics were the least accurate. For the attested phonological errors, the authors reported ten processes including de-emphasis (occurred 50% or more of the time); stridency deletion, and lateralization of /r/ (occurred 25-50%); syllable reduction, final consonant deletion, consonant sequence reduction, fronting, final devoicing, initial voicing, and stopping (occurred 1-24%).

Mashaqba et al. (2019) investigated early word syllable structure and phonological processes appearing in JA children. They recruited 20 participants aged 1;0 to 3;0 years, divided into four age groups: 1;0–1;6, 1;7–2;0, 2;1–2;6, and 2;7–3;0. The study documented the acquisition of phonological aspects of JA child speech to fill the gap in the literature which only highlighted adult productions and other dialects such as Cairene Arabic (Ammar & Morsi 2006, Ragheb & Davis 2010, Saleh et al., 2007), and Kuwaiti Arabic (Ayyad 2011, Ayyad & Bernhardt 2009).

---

[17] The authors differentiated between errors and changes (Dyson & Amayreh 2000, p.82):
1) Errors are productions that are neither ESA nor acceptable (would not be used by normal-speaking adults even in casual speech).
2) Changes from ESA refer to all productions that do not match ESA forms and combine productions that are acceptable as well as others that are true errors.

Data analysis included manually tracing syllable structure frequency, syllable shape, word syllable structure, and notable phonological processes across age groups.

JA children adhered to the universal development with CV developing into CVC, with older children producing a greater variety of syllabic structures. Results demonstrated that JA children produced thirteen syllables: CV, CVV, CVC, CVVC, CCV, CCVV, CCVVC, CCVC, CCCVV, CCVCC, CVCC, CCVC, and CVVCC. The most occurring syllables were CVC, CV, CVVC, and CVV with 38%, 30%, 14.7%, and 12%, respectively. The least occurring syllables were CCCVV and CVVCC. In the youngest age group, 1;0–1;6, CV, CVV, CVC, and CVVC syllables were produced comprising open and closed syllables, short and long vowels, codas, one superheavy syllable, and no onset or coda clusters. Notably, the two age groups, 1;7–2;0 and 2;1–2;6, had similar patterns to those attested in the youngest age group, with no syllable production development. This was attributed to a phonological plateau, where children might have resorted to developing their existing skills before advancing to more complex structures (Vihman 2004). This observation was also attributed to individual variation contributing to overall age group development. An expansion in the syllabic development was only evident in the oldest age group, 2;7–3;0, where children produced onset clusters in CCV, CCVVC, CCVC, CCVCC and CCVC, coda clusters in CCVCC, and CVCC. However, these clusters were reported to be limited to monosyllabic words only due to their complexity (e.g., /burʒ/>[burd] ‘bridge’).

JA children maintained the bimoraic weight parameter in their productions, as the moraic weight of a prosodic word in JA is minimally bimoraic. Results showed that the youngest age group tended to produce more monosyllabic words than other word lengths in CVV, CVG (G-geminate), and CVVC shapes with 20%, 36%, and 21%, respectively. Disyllabic words followed comprising CVV/CV, CVC/CV, and CV/CV shapes with 26%, 24%, and 21%, respectively. Trisyllabic words were rare and comprised only 3% of productions with two-word shapes CVC/CVC/CV and CVC/CVV/CV. In the second age group, a significant decrease in the production of monosyllabic words was reported, where 17% of the productions were monosyllabic compared to 50% in the previous group. Disyllabic words were more frequently produced with 82% including CVC/CVC, CVC/CVVC, CVC/CV, CVVC/CV, and CVV/CV shapes, with 28%, 20%, 25%, 5%, and 22%, respectively. The increase in disyllabic words was attributed to increased vocabulary, improved motor skills, and language exposure (Kuhl 1991, Jusczyk, Houston & Newsome 1993, Vihman 2018). Trisyllabic word shapes expanded to include CV/CV/CVVC, CVC/CV/CV, and CVC/CV/CVVC shapes. By 2;1–2;6 years, monosyllabic word production was further reduced and limited to CVG (11%) and CVVC (14%) superheavy forms. Disyllabic words were the most frequent with 61%, while an increase in trisyllabic words was attested with 14%. In the oldest age group, monosyllabic words had a slight increase of 29%, attributed to consonant cluster acquisition. Disyllabic words were the most frequent (39%), followed by trisyllabic words (32%). No quadrisyllabic words were attested up to 2;7–3;0 years.

The notable phonological processes involving syllable structure were cluster reduction, onset deletion, weak syllable deletion, coda deletion/closed syllable to open syllable (CVVC to CVV), vowel lengthening, syllable epenthesis, assimilation, gemination and assimilation, and metathesis. Although the study reported that word-initial and final clusters appeared in the oldest age group, cluster reduction was evident across the age groups referring to word medial positions (e.g., /kalb.na/ ‘our dog’ > [kab.na]). Weak syllable deletion was attested across the age groups, with the youngest age group having the most frequent occurrences (21.4%) and the oldest having the least (1.5%) (e.g., /laj.mu:n/ > [mu:n]‘lemon’ and /til.fiz.jo:n/ > [ʔiz.jo:n] ‘television’). Gemination and assimilation was another frequently used phonological process that persisted even in the oldest age group (e.g., /ʕaṣ.fu:r/ > [ʔaf.fu:l]‘bird’ and /mas.ba.ħah/ >

[sab.ba.ħah] ‘rosary’). Metathesis was regularly attested and persisted until age group 2;7–3;0 years (e.g., /ta.la.fo:n/ > [ta.fa.lo:n]‘phone’ and /nas.ka.fe:h/ > [san.ka.fe:h] ‘Nescafe’).

To conclude, speech development in Arabic has been under-studies, where most of the studies focused on the segmental level of early phonological acquisition, the order and rate of acquisition of phonemes, the developmental phonological processes, and the acquisition rate of syllable structure. The methods used were often qualitative/observational, where the researchers made subjective judgments or interpretations based on their perceptual impressions. However, acoustic and computational methods of analysis were not reported in the literature. To the best of the researcher’s knowledge, no study has traced the durations of syllable development patterns in normally developing JA-speaking children.

2.2.4. Lexical Stress in Child Speech

Children’s capacities to produce and recognize stress patterns serve as a significant indicator of their phonological development. This reflects their ability to internalize and apply language universal and specific prosodic rules (Hochberg 1988, Fikkert 1994, Gerken 1994, Kehoe & Stoel-Gammon 1997). The section explores debates on stress development stages and metrical constraints, trochaic versus neutral bias in stress learning, and lexical learning versus rule-based stress acquisition (Allen & Hawkins 1980, Hochberg 1988, Klein 1984). The section also considers the role of acoustic cues in stress production and how children’s mastery of these cues indicates the complexity and variability in stress pattern application (Pollock et al., 1993, Kehoe et al., 1995, Arias & Lleó 2009,).

First, whether children have a bias towards a particular foot type has been debated. Findings suggest that the preference for a trochaic stress pattern may be a feature of early language development, regardless of the ambient language. Allen & Hawkins (1980) argued for a universal bias towards trochaic patterns in English-speaking children. This was evident in the frequent deletion of weak syllables, particularly in word initial unstressed positions. Further supporting the trochaic bias argument, Allen & Hawkins (1980) observed that children exhibited higher accuracy in producing trochaic words compared to iambic words. Even when children produced iambic words, they often emphasized the initial syllable, indicating an underlying preference for trochaic structures. While the trochaic preference is considered universal, Allen & Hawkins discussed that this preference becomes more prominent when the target language has a high frequency of trochaic words input. This bias was also evident in children organizing their productions in a Sw trochaic stress pattern (Gerken 1991, 1994, Kehoe & Stoel-Gammon 1997, Roy & Chiat 2004). For example, Bolinger (1986) suggested that weak syllables preceding strong ones (i.e., wS or wSw) are more fragile and thus more prone to omission. Gerken's (1994) study on English 2-year-olds further supported this claim. In producing nonsense words, children showed a tendency to preserve the second weak syllable more than the first weak syllable in a weak-Strong-weak-Strong stress pattern (e.g., /pazamkasis/), resulting in a Strong-weak-Strong pattern (e.g., /zamkasis/). Levey and Schwarts (2002) discussed the ability of two-year-old children to perceive and produce minimal pairs of trisyllabic words with primary stress on either the first or second syllable. The results showed that English monolingual children exhibited a Strong-weak-weak bias (e.g., elephant), and instances of syllable omission were less frequent in Strong-weak pairs compared to other patterns. However, children tended to omit the first syllable of trochaic weak-Strong-weak pairs to retain a Strong-weak pattern, reflecting a trochaic bias. In Dutch, Fikkert (1994) highlighted that Dutch children exhibit a bias towards trochaic feet. This was supported by the high production frequency of disyllables with initial stress (e.g., [ˈbo:mi:] 'ballon') and children tending to delete initial unstressed syllables in their productions (/ˈte:kana/ produced as [ˈka:ka] 'tekenen'). Furthermore, while Dutch-speaking children seemed to segment and extract both trochaic and iambic words by one year of age (Jusczyk, Houston & Newsome

1999), they showed more sensitivity to mispronunciations and misplacements of trochaic stress words compared to iambic ones since Dutch has a trochaic stress pattern. Similar patterns have been observed in children’s early productions in Spanish (Macken 1993), Czech, Slovenian, and Estonian (Vihman 1980).

While the debate regarding the trochaic bias is well supported in the literature, there has been a contrasting argument supporting a neutral bias in child productions. Hochberg (1988) proposed that Spanish-speaking children approach stress learning without a predisposition towards any particular foot type. According to Hochberg, stress placement was not consistently on the same syllable across different words in Spanish. She noted an increasing percentage of correct placements for penultimate stressed words over time. Hochberg interpreted this gradual increase in accuracy as evidence against an innate trochaic bias. If there were a trochaic bias, children would exhibit consistent stress patterns from the initial stages of development, rather than showing an increase in accuracy over time. This suggests that children's stress acquisition is influenced by language-specific characteristics rather than a universal trochaic predisposition. In languages such as Spanish, where stress patterns vary and do not conform to a predominant trochaic pattern, children adapt their stress production in alignment with the specific prosodic rules of their language. This finding supports the notion that children's phonological development is significantly shaped by the linguistic environment and the specific stress patterns to which they are exposed.

Second, whether stress is achieved by lexical learning or by rule has also been debated. Hochberg (1988) argued for the systematic acquisition of stress rules. She suggested that if stress acquisition was purely lexical, child productions would closely match adult patterns with minimal form deviations. Her data consisted of spontaneous and imitated productions of 3,4 and 5-year-old Spanish learners. The study predicted that if children do learn stress rules, then they should find words with regular stress easier to imitate than words with irregular stress and they should tend to regularize stress in words with irregular stress, but not the opposite. The results confirmed these predictions as Spanish children had difficulties imitating words with irregular stress patterns, tending to regularize stress in such words, indicating an active process of learning and applying stress rules. On the other hand, Klein (1984) analysed the polysyllabic utterances of an English-speaking 2-year-old. Results demonstrated that the child's stress placement was the same as in the adult target word (62%). This emphasized lexical primacy, suggesting that children prioritize learning stress patterns at the lexical level, with early productions closely resembling adult forms. Nonetheless, Klein's approach was critiqued (Fikkert 1994). The limited scope and reliance on imitated utterances, including overlooking aspects, such as monosyllable reductions of polysyllabic words and target words being disyllables with initial stress only, may not reflect a child's understanding of stress rules.

Third, children seem to integrate inputs and outputs constrained by the parameters of the structural organization proposed in metrical theory. Fikkert (1994) suggested four stages in the development of stress patterns based on longitudinal data collected from twelve Dutch-speaking children with an age range of 1;0 to 1;11 years. The model suggests that in the initial stage, children produce disyllabic words with final stress as monosyllabic words maintaining the final stress (e.g., the disyllabic word banana /ba:ˈna:n/ being produced as [ˈba:n], p.202). Children select the segmental material from the most salient syllable, in particular the foot in which this syllable is contained. The child circumscribes a trochaic foot from the right side of the adult target word and maps it onto his trochaic template. For instance, in cases of disyllabic target forms with initial stress, these words are never truncated, and if the circumscribed foot contains two syllables, they are both realised (e.g., the word /ˈbe:bi:/ baby is never produced as *[be:], but is produced as [ˈbe:bi:], p.210). In the second stage, children realize both syllables of the target word, but contrary to the adult stress pattern, the stress falls on the first syllable. Fikkert discussed that since what is circumscribed is not clear (i.e., it cannot be a foot, a two-syllable window, or the entire word, as none of which align with the prosodic structure of the adult target) trisyllabic words need to be examined. The analysis showed that trisyllabic words with stress on the first and third syllables (e.g., ooievaar /ˈo:jaˌfa:r/) are truncated to disyllables. The first and third syllables are produced, with the segmental material from the stressed elements being selected (e.g., /ˈo:ˌfa:r/). At the third stage, children realize both syllables of the adult target word but both syllables receive an equal amount of stress, resulting in productions with level stress (e.g., banana /ba:ˈba:n/ being produced as [ˈma:ˈna:n]). Although children’s prosodic templates have expanded to two feet, they have not acquired the main stress rule which assigns a greater prominence to one of the feet in the template. At the final stage, words are produced with an adult-like stress pattern (e.g., muziek /my:ˈsi:k/ being produced as [ˌmy:ˈsi:k]).

The acoustic cues for stress seem to be complex, and there are indications that children may not be able to control them very well (Bernhardt & Stemberger 2020). Many studies attempt to bring empirical evidence for or against theoretical claims related to the early representations of word stress (Correia 2009). Allen & Hawkins (1980) showed that duration and pitch were employed by children to derive target-like accented phrases. Three English-speaking children, with an age range of 2;8–3;4 years were recruited. The authors suggested that the child productions resembled the adult ones as final and post-nuclear stressed syllables tended to be longer than the initial syllables in addition to falling contours being evident in most of the syllable types.

Pollock, Brammer & Hageman (1993) analysed the productions of eighteen English-speaking children ranging from 2–4 years of age. The study traced the use of pitch, intensity, and duration in nonsense disyllabic words. Results showed that two-year-olds produced longer stressed syllables than 3- and 4-year-olds, and they did not correctly control pitch and intensity measures to mark stressed syllables in ˈCVCV and CVˈCV targets. The acoustic cues were employed correctly by only 3- and 4-year-olds to mark stressed syllables. All groups produced significantly longer stressed syllables compared to unstressed ones.

Additionally, acoustic correlates of stress in spontaneous speech samples of English-speaking children with an age range of 1;6–2;6 were examined by Kehoe, Stoel-Gammon & Buder (1995). Results showed that higher pitch, greater amplitude, and longer durations were employed by children to create stress contrast in an adult-like manner. Despite this, 30% of the tokens that were marked as incorrectly stressed had different acoustic properties compared to those that were correctly stressed. This suggests that by the age of 2:6 years, children had not yet fully mastered the use of acoustic parameters to control stress accurately in producing new words. This was attributed to the insufficient control of phonetic parameters, rather than the incorrect use of lexical stress patterns.

An analysis of word and phrasal stress acquisition in Spanish-speaking child speech was conducted by Arias & Lleó (2009). The spontaneous speech data of two Spanish monolinguals with an age range of 1;0–2;6 years showed that children placed stress correctly in their early productions. The authors highlighted that although children do not necessarily produce the acoustic correlates of stress in an adult-like manner, Spanish children tend to mirror the stress patterns of their ambient language in their productions. However, Arias & Lleó noted that in single trochees, the effect of final lengthening is evident highlighting that durations in trochees are not yet controlled (amplitude and pitch were controlled). In trisyllabic words with penultimate stress, the penultimate vowel had longer durations and higher pitch values compared to unstressed ones; while in quadrisyllabic words, control of durations was evident as penultimate syllables were lengthened compared to other syllables. Results demonstrated that by 1;8–1;9 years, some words with an iambic stress pattern were produced as trochees.

As for Arabic, only a few studies have addressed the acquisition of stress (al Huneety et al., 2023). Research on child speech suggests that syllable structure plays a significant role in stress placement with children gradually mastering stress patterns and exhibiting more complex syllable structures with maturation (Ammar 2002, Daana 2009, Alqattan 2015, al Huneety et al., 2023). For example, Ammar (2002) conducted a study on the acquisition of syllable structure and stress patterns in colloquial Egyptian Arabic among two to three-year-old children. Ammar emphasized the significant role of syllable structure in Arabic phonology highlighting that word stress is entirely predictable from syllable structure. An audio-visual analysis was conducted and a wave editing program was used to locate the stress position from the acoustic records. Results showed that heavy syllables such as CVCC are always stressed and there is only one stressed syllable per word. When children altered syllable structures, they tended to preserve the prosodic weight of the syllable maintaining the original stress pattern (e.g., short syllables were lengthened in cases of syllable deletion in polysyllabic words). Ammar reported that children employed pitch accents in cases of stressed syllables but no assessment of employing duration was conducted.

In Alqattan's (2015) study, the acquisition of stress patterns was examined as part of a larger analysis of Kuwaiti Arabic early phonological development. The study involved 70 children, aged 1;4 to 3;7 years, and spontaneous speech samples were recorded. Only perceptual analysis was conducted, stress was determined by Kuwaiti Arabic stress rules with no reference to the acoustic parameters of stress production. Results showed a clear developmental variability in stress patterns with age. Younger children (1;4–1;7) produced only five stress patterns including Sw, wS, wSw, wwS, wwSw, whereas the oldest group (3;4–3;7) produced three additional patterns including wwwS, wwwSw, and wwwwS. Alqattan observed that the increase in word shape complexity and syllable structures correlated with the development of stress patterns, exhibiting a marked development in the use of longer and more complex syllables after the age of 2;4.

As for JA child speech, Daana (2009) conducted a study on the development of consonant clusters, stress, and plural nouns. A sample of 30 monolingual children with an age range of two to five years were recruited from Amman. Spontaneous speech samples were collected during naturalistic interactions including conversations and play sessions. Samples were transcribed and analysed using perceptual and acoustic methods, wherein stress patterns were identified and measured in terms of pitch, duration, and intensity. Results showed that in the early stages of development, JA children tend to simplify stress patterns as they often stress the penultimate syllable, regardless of the target word's stress pattern. With age, children begin to produce more complex stress patterns, and only by the age of five, do children exhibit adult-like stress patterns. JA children marked stressed syllables by a higher pitch, longer duration, and greater intensity compared to unstressed syllables. Daana indicated that the acquisition of stress patterns was closely linked to development in other phonological aspects including the mastery of consonant clusters and syllable structures (i.e., children who exhibited advanced stress patterns also tended to show more developed consonant clusters and syllable structure productions). The study emphasized the significant role of syllable structure in stress placement, as the correct application of stress was often dependent on the child's ability to produce syllables accurately, particularly those with long vowels or consonant clusters.

Al Huneety et al. (2023) examined the stress patterns of Ammani-speaking children. A total of 48 typically developing children were recruited and divided into four age stages: 1;0–1;6, 1; 7–2;0, 2;1–2;6, and 2;7–3;0 years old. Data was collected through spontaneous speech samples and a picture-naming task. Data was analysed perceptually, and stress assignment was assessed based on the stress rules of JA. Results demonstrated that Ammani-speaking children go through four developmental stages of acquiring stress until they become adult-like by the age of three years as follows (p.8):

1. Stage One (1;0–1;6): children correctly produced stress on 63% of disyllabic words; overall, stress shift occurs when children violate the stress rules of adults and place it on the rightmost syllable in CV.CV words (creating an iambic pattern) and when they place stress on a syllable with a medial geminate irrespective of the weight of the final syllable.
2. Stage Two (1;7–2;0): children up to 24 months do not fully acquire stress rules; accuracy in placing stress increases reaching 83% for disyllabic words; children started to produce trisyllabic words with an accuracy rate of 60%.
3. Stage Three (2;1–2;6): children produced stress in disyllabic words similar to adults; however, they misplaced stress in 25% of trisyllabic words.
4. Stage Four (2;7–3;0): stress is produced in an adult-like fashion.

Ammani children exhibited no bias for any stress type as they employed both trochaic and iambic feet in their productions, in line with Hochberg' (1988) neutral bias analysis. JA children in the youngest two age groups exhibited stress misplacements as they produced iambic foot forms instead of trochaic ones. Al Huneety and colleagues indicated that JA children placed stress on the syllable with a geminate (G), irrespective of the weight of the other syllables (a process described as the process of assimilation and gemination), but this process stopped by 30 months (e.g., production /ˈbak.ka/- target /baˈqara/ CVGVV- CVCVCV 'cow' and production /ˈtun.nan/- target /ˈʃukran/ CVGVC- CVC.CVC 'thank you').

2.3.Summary and Predictions

2.3.1. Jordanian Arabic: A summary

JA, a VC dialect, embodies a rich linguistic and cultural significance shaped by socio-economic, geographical, and historical influences. JA syllable structures are CV, CVC, CVCC, CVV, CVVC, CVVCC, CCVC, and CCVVC (AbuAbbas 2003, Crossley 2023). JA syllables are classified as light, heavy, or superheavy (Watson 2002, Abdoh 2011, Alotaibi & Meftah 2013, Alqattan 2015). Stress patterns in JA are rule-governed and predictable, based on syllable weight, position, and word length (Al-Ani 1970, Watson 2002, Holes 2004). Constraints such as bimoraicity condition, ERR, extrasyllabicity, extrametricality, and WPR contribute to syllable weight and stress placement (Watson 2002). Onset clusters in JA appear regardless of sonority level while coda clusters adhere to SSP, (Al-Ani 1970, Kiparsky 2003, Daana 2009, Mashaqba et al., 2021). Epenthesis is expected in JA productions (e.g., /fikr/ CVCC ‘thought’, /fikir/ CVC(V)C). As for syncope, JA permits the creation of word-initial clusters (e.g., /klaab/ 'dogs' that has an underlying pattern of /kilaab/).

Analysing Arabic superheavy syllables remains limited despite its potential implications for phonology, prosody, and language acquisition. Previous research has primarily focused on theoretical phonological accounts, highlighting the complexity of superheavy syllables and their interaction with stress, morphological markers, and phonological processes. Some studies suggest that superheavy syllables are bimoraic (Broselow et al., 1995, Zec 2007), although some scholars argue for a trimoraic analysis (Fery 1997, Crossley 2023). Evidence from Arabic dialects (e.g., Meccan and Lebanese) supports the bimoraic analysis with vowel shortening, mora sharing, and epenthesis being evident to avoid trimoraicity. However, there are not sufficient acoustic data supporting or rejecting these theoretical proposals, and there is a lack of research on superheavy syllables in child speech.

2.3.2. Language Development: A summary

Infants demonstrate early perceptual attunement to speech elements (Eimas et al., 1971, Kuhl 1991). Infants' perceptual journey involves transitioning from 'citizens of the world' to 'culture-bound listeners', becoming attuned to language-specific phonetic distinctions (Werker & Tees 1984). By six months, infants show perceptual attunement to vowels, followed by increased sensitivity to consonant contrasts by twelve months (Best 2017). In production, infants progress through stages such as cooing, canonical babbling, and single-word utterances (Eimas 1975). Babbling plays a crucial role in articulatory gesture mastery and interactive responses from parents/caregivers (Jusczyk, Luce & Charles-Luce 1994). Infants become increasingly attuned to the features of their ambient language leading to the emergence of early word forms (Vihman 1996). Between 12–18 months, normally developing infants enter the holophrastic stage, producing single-word utterances (McCune & Vihman 2001, Menggo 2017). By 18–24 months, their expressive vocabulary grows to 50-100+ words. From the end of the first year through the middle of the second year, children's linguistic production shifts from holophrastic components to more stable canonical forms, such as CVC and CV/CV, reflecting improved accuracy and maturation (Vihman 1996).

The ‘innateness’ perspective proposes that children are born with innate linguistic knowledge with children exhibiting a universal and predictable order of syllable acquisition (Jakobson 1968, Chomsky 1965, Chomsky & Halle 1968). The ‘motor learning’ approach suggests that language acquisition occurs through exposure to input and subsequent learning mechanisms, such as word-final lengthening, that reflect language-specific experience and motor skills (Robb & Saxman 1990, Clark & Casillas 2015). Younger children may exhibit greater final lengthening but lack motor control, creating greater discrepancies between word-final and non-final syllables. This highlights the interaction between language experience, motor advancement, and phonological development (Zec 2007).

Existing literature predominantly focuses on early linguistic development in children acquiring West Germanic languages, such as English, Dutch, and Spanish. Research on phonological acquisition and development in Arabic is scarce and mainly evaluated accuracy in producing consonants, syllable structure, stress patterns, and phonological processes (Saleh et al., 2007, Ayyad 2011, Alqattan 2015). Although some studies have examined normal phonological development, the literature lacks an in-depth exploration of aspects such as articulatory and perceptual development, articulatory challenges, and early acoustic sensitivity in Arabic-speaking children. Research on phonological development in JA has explored aspects, such as the acquisition of MSA consonants, identifying developmental trends, the influence of phoneme frequency and functional load on the acquisition, phonological processes, and early word syllable structure (Amayreh & Dyson 1998, Amyreh 2000, Dyson & Amayreh 2000, Mashaqba et al., 2019). However, the methods used were often qualitative and/or observational, and acoustic or computational methods of analysis were not reported in the literature.

2.3.3. Lexical Stress: A summary

Stress is an abstract property of a word inherent to its phonological structure (Ladd 2008). However, the acoustic correlates of stress are not always clear. In Arabic, stress is predictable, and influenced by syllable weight and structure. Previous studies showed that compared to unstressed syllables, stressed syllables have longer durations, and higher amplitudes (Al-Ani 1970, de Jong & Zawaydeh 1999, Zuraiq 2005). Research on the developmental process of lexical stress perception and production highlights the significance of recognizing stress patterns and sensitivity to acoustic differences (Cutler & Carter 1987, Demuth 1996). Stages of stress development correspond to the progression of metrical constraints reflecting language-specific input (Allen & Hawkins 1980, Fikkert 1994). The debate on whether children have a bias towards a particular foot type remains controversial, with some arguing for a universal trochaic bias (Allen & Hawkins, 1980), while others, such as Hochberg (1988), suggest a neutral bias. The acquisition of stress rules versus lexical learning is significant, with Hochberg (1988) arguing for systematic rule acquisition emphasizing lexical primacy.

Acoustic studies show that while children can produce stress contrasts, mastery of acoustic parameters, such as duration, pitch, and intensity develops gradually (Bernhardt & Stemberger, 2020). In Arabic, few studies addressed the acquisition of stress and its acoustic parameters. Ammar & Morsi (2006) and Alqattan (2015) emphasized the role of syllable structure in stress patterns. Daana (2009) found that JA-speaking children tended to simplify stress patterns, with adult-like production achieved by age five. Al Huneety et al. (2023) demonstrated four developmental stages in AA-speaking children, achieving adult-like stress patterns by age three, with no bias towards any stress pattern.

2.3.4. Predictions

General predictions

Drawing from language universals and language-specific features of JA, this section presents predictions on syllabic development and its durational properties. The statistical analysis is conducted using Bayesian Multi-level Models as they allow estimate incorporation of parameters posterior distributions based on probabilistic ‘predictions’ rather than testing specific ‘hypotheses’ (Winter 2013, Franke & Roettger 2019, Gelman et al., 2021). This flexibility enables the exploration of multiple predictions simultaneously providing a more adaptable approach to data analysis (Bürkner & Charpentier 2020). The models utilizing predictors seem to handle uncertainty by estimating a range of probability values instead of a single outcome (Gelman et al., 2021) This is particularly useful for modelling the continuous nature of linguistic data, which may be oversimplified by binary hypotheses (Winter 2013). The specific predictors and their interactions under investigation are age group, lexical stress, syllable position within a word, and syllable structure.

Syllable structure

In line with language universals, it is anticipated that core syllable structures will be produced more frequently in JA child speech compared to marginal ones (Mashaqba et al., 2019). Younger children are expected to exhibit less variability in their syllabic structures, gradually exhibiting adult-like patterns as children age (Alqattan 2015). This developmental trajectory suggests a gradual increase in the production of superheavy syllables over time, accompanied by a refinement in syllabic patterning towards adult norms. Despite their complexity, superheavy syllables are anticipated to emerge as early as two years of age. This early appearance can be attributed to the phonological saliency and frequency of such structures in Arabic due to its rich consonantal inventory and morphological complexity (Broselow et al., 1995). Within JA typology, cluster formation poses a notable challenge for children, influenced by constraints such as SSP (al Huneety et al., 2023). Consequently, clusters are expected to appear more frequently in older age groups, with younger children resorting to epenthesis as a strategy to reduce the complexity of clusters. Structures such as CVCC are expected to emerge to satisfy the bimoraicity, with geminates playing a significant role in the production of superheavy syllables (Davis & Raghib 2014). JA productions would adhere to multiple properties governing syllabic structures including the bimoraicity condition, ERR, WPR, extrametricality, and extrasyllabicity (Mashaqba et al., 2019). Even in the presence of expected phonological processes (e.g. weak syllable deletion, assimilation and gemination, cluster reduction), child productions will abide by these constraints.

Superheavy syllables in JA child speech are predicted to be bimoraic and not trimoraic. Processes such as vowel shortening and epenthesis are anticipated to occur in superheavy syllable productions. This prediction is informed by theoretical considerations and empirical observations, highlighting the challenges associated with trimoraic syllable structures in Arabic phonology (Broselow et al., 1995, Watson 2007, Owens 2013). JA child productions are expected to exhibit phonological processes, where deviations from the adult forms will be evident. The processes are predicted to fall under two categories: segmental and prosodic (Section 3.4.2). Younger children are expected to produce more phonological processes deviating from adult forms compared to older ones. Phonological processes will be influenced by syllable structure, lexical stress, and syllable position. The phonological processes will be reflective of the durational and developmental trajectories of JA child productions.

Duration

Younger children are predicted to produce longer durations compared to older children (Bunta & Ingram 2007). The transition towards adult-like durations will exhibit different patterns for syllables and vowels. Specifically, the durations of syllables will become more adult-like at a later stage of development compared to vowels. This discrepancy arises from the inherent durational variability of consonants (Payne et al., 2012), with some acquired later in the developmental stages of JA (Amayreh & Dyson 1998, Amayreh 2000). This delayed consonant acquisition is expected to manifest in longer durations, whereas vowel durations are expected to be more consistent across age groups (Lee, Potamianos & Narayanan 1997). In JA speech, duration is predicted to serve as a robust cue for marking word-final syllables, stressed syllables, and superheavy syllables. Below are specific predictions for the influence of the predictors on syllable and vowel durations in JA productions, in addition to superheavy syllables in particular.

Specific predictions

The following predictions apply to syllable and vowel durations:

1. Syllable Development

    a) A decreasing trend in durations will be evident, with younger age groups producing longer syllables compared to older age groups.
    b) The oldest age group will exhibit durations that are close to the adult ones but the durations of the two groups will not intersect.
    c) Lexical stress will influence syllables resulting in longer durations for stressed vowels/syllables compared to unstressed counterparts.
    d) Word final lengthening will be evident in longer durations for word final syllables and vowels compared to non-final ones.
    e) Syllable structure complexity will influence durations with syllables containing more constituents displaying longer durations.

f) Younger children will produce more phonological processes deviating from adult forms compared to older children, and these processes will be more evident in superheavy, stressed, and final syllables compared to their counterparts.

2. Superheavy Syllable Production

a) Younger children are expected to produce longer superheavy syllables compared to older age groups.
b) Stress will affect superheavy syllables with stressed syllables exhibiting longer durations than unstressed ones.
c) Word-final lengthening will be observed in superheavy syllable productions with final syllables being longer than non-final ones.
d) Superheavy syllables are bimoraic and not trimoraic as vowel shortening is anticipated to occur in non-final superheavy syllables.

3. Chapter Three: Methodology

Chapter 3. Methodology

This chapter presents the research methodology including participant recruitment procedure, details on tasks including experimental materials and procedure (ST, PT, and RT), the analysis of phonological processes, details on the acoustic and statistical analyses, and limitations of the methodology.

3.1.Approach

The present study adopts a quantitative approach. The focus is on analysing the development of syllable structure with an emphasis on superheavy syllables in JA child speech. The literature review revealed a scarcity of acoustic data attributed to JA child speech. Previous studies have predominantly relied on auditory analysis (Dyson & Amayreh 1999, Amayreh & Dyson 2000, Daana 2009, Abdoh 2011, Ayyad 2011, Alqattan 2015, Mashaqba et al., 2019). The review of literature highlighted superheavy syllables posing challenges to children due to their complex structure and the additional articulatory efforts they require. However, despite their linguistic significance (Section 2.1.4), there has been limited investigation into superheavy syllable development in Arabic.

The first part of the analysis concerns raw syllable and vowel durations. The predictors were age group (children ranging from 24–30 months to 66–72 months and adults aged 20–30 years), syllable structure (light, heavy, and superheavy), lexical stress (stressed, unstressed), and syllable position within a word (word non-final and word-final). The second part presents an analysis of phonological processes occurring in JA child speech. The third part concerns raw syllable and vowel durations in superheavy syllables. The predictors were age group, lexical stress, syllable position within a word, sub-structure (the composition of vowel length: short/long; clusters: onset/coda of the superheavy syllable including CVCC, CVVC, CCVVC, and CCVCC syllables), and word length (monosyllabic, disyllabic, multisyllabic).

Duration of child production forms was chosen as the dependent variable for the following reasons. First, analysing durational properties provides insights into linguistic properties, such as syllable structures (e.g., CV, CCV, or CVVC) and stress patterns (Khattab & Al-Tamimi 2014). Second, temporal properties of speech vary across the developmental stages, reflecting changes in speech production skills (Nip & Green 2013). Third, the duration can be reliably measured compared to other cues. As the data collection was conducted online, intensity could be significantly influenced by uncontrollable factors such as the placement and type of recording device used by the parent/caregiver or adult participant. These factors can lead to inconsistent intensity measurements (Kent & Read 2002). Pitch was excluded as it includes both linguistic (e.g., intonation and stress) and paralinguistic variation (e.g., emotional tone and speaker-specific characteristics) in spontaneous speech (Ladd 2008, Banse & Scherer 1996). While children's development in the use of intonation is worth investigating, it is outside the scope of this study. Finally, providing acoustic and statistical analyses of durational patterns in JA productions presents a methodological advancement diverting away from impressionistic subjective methods in previous studies (Alqattan 2015).

Raw durations of child production forms, rather than proportional durations, were analysed in the current study. Raw syllable and vowel durations provide clear insights into developmental patterns of child speech, particularly for their motor control abilities. They capture the actual timing without normalizing for speech rate variation (Smith & Kenney 1994). This is valuable in tracking the developmental milestones of child productions to investigate how they gradually approximate temporal properties of adult speech (Green, Moore & Reilly 2002). Moreover, analysing raw durations allows for an in-depth understanding of how children acquire and develop language-specific features. For example, JA includes emphatic consonants that pose a challenge to children up to 6–7 years of age, a phonemic distinction between short and long vowels, and constraints for producing clusters such as SSP (Amayreh & Dyson 2000, Amayreh

2003, Mashaqba et al., 2019). This focus on raw durations provides an overview of the natural progression and difficulties that JA children encounter in mastering such complex aspects of their ambient language. Furthermore, as this study aims to construct a reliable reference for what constitutes normal/typical JA speech, raw durations can be an asset to the clinical considerations in diagnosing and treating atypical productions. Specific temporal patterns in raw durations (e.g., consistent prolonged vowels or irregular syllable timing) may indicate the presence of language delays or disorders that could be obscured by normalization techniques (Shriberg, Tomblin & McSweeny 1999). Thus, presenting raw durations can be more useful for diagnosis, clinical monitoring, and intervention.

3.1.1. Participants

In the first stage, participants were recruited from the experimenter's immediate social network. Her acquaintances, who were parents/caregivers of children aged 24 to 72 months, were contacted via telephone. The rationale for this recruitment strategy was to guarantee that the collected data would not contain dialectical variation, which could be a confounding factor. The experimenter provided potential participants' parents/caregivers with information about the study, and they were asked whether they agreed to their children's participation. Parents/caregivers were provided with an electronic copy of the participant information sheet which included study aims, questions, methods, data protection and analysis measures, potential risks, and the voluntariness of their decision to take part. Parents/caregivers were given 1–2 weeks to decide. If they agreed for their children to participate in the research, an electronic copy of the consent form was provided for a signature. For the child group, a total of 21 participants were recruited in seven age groups ranging from 24 to 72 months old (24–30, 31–37, 38–44, 45–51, 52–58, 59–65, and 66–72 months). Four adults (two males and two females, aged 20–30 years) were recruited for the control group, and the same recruitment procedure was used. Based on the ethical procedures, participants in the child and adult groups received customized materials such as the consent form and the information sheet.

All experiments were conducted online using Microsoft Teams provided by the University of Central Lancashire (UCLan) owing to COVID-19 regulations that restricted travel and in-person activities. The experimenter joined a video call with each participant in the adult group and each child in the child group, where children were accompanied by their parents/caregivers. Participants in the adult group and parents/caregivers in the child group were instructed to record in a quiet room to limit background noise. Participants and parents/caregivers were instructed to have access to two devices, a mobile phone, and a laptop, intended to be used for recording and conducting the meeting, respectively. During the video call, participants were asked to audio-record their voices with applications such as Recorder and Awesome Voice Recorder. To ensure that all audio recordings had a sound quality suitable for acoustic analysis, participants were directed to use mono-channel files at 44.1khz, 256 bps, in lossless formats (.wav for the Voice Recorder and .ogg for the Recorder). Participants were requested not to use any headphones or microphones to ensure a consistent recording quality across the participants. They were advised not to put the mobile phone near the laptop, as the laptop's fan noise might affect the recording quality. Subsequently, when all experiments were completed, participants uploaded the audio recordings to a shared OneDrive destination linked to the experimenter's Office 365 account provided by UCLan. Each participant had a separate file destination to upload the audio recordings.

All participants had Arabic as their first language and JA (Section 2.1.3) as their main dialect, and they lived in Jordan. Parents/caregivers were asked to identify any cognitive, behavioral, or hereditary disorders that might alter the perception or production abilities of their child before signing the consent form to ensure that the collected data was linguistically normal. Consequently, one child participant (from the age group 24–30 months) was excluded from the experiment before the recording session, as the parent disclosed that the family has a history of a speech disorder. This study was approved by the BAHSS Ethics Review Panel at UCLan (BAHSS2 0280).

The justification for selecting and dividing age groups relates to the following criteria. Firstly, participants under 24 months were unsuitable for the current study as they exhibit simple forms of production ranging from babbling to a vocabulary of ≤ 50 words (Feldman 2019). Children aged 1–2 years can understand approximately five times the number of words that they can express verbally (NHS Foundation Trust 2016). Secondly, beyond 72 months, there is a drastic expansion of children's linguistic repertoire by education, as MSA is implemented in schools around the age of 7 years. Finally, some sounds in Arabic, such as emphatics/gutturals, are not entirely acquired by 6–7 years of age. Amayreh & Dyson (1998) and Amayreh (2000) reported that emphatic sounds such as /θ/, /ð/, /ʒ/, /dˤ/, /zˤ/ are acquired at a later stage of development. This is due to these consonants simply requiring a degree of articulatory precision not yet available in children before 6 or 7 years of age. Based on F1 and F2 values of vowels (taken at onset, midpoint, and offset) neighboring the emphatic consonants, Mashaqba et al. (2022) suggested that the production of emphatic consonants resembles adult-like values by the age of 6 years in word-initial and median positions. They added that the child forms exhibit significant errors and do not match the adult forms acoustically, and only by the age of seven years, do the consonants in word-final positions exhibit adult-like patterns. Therefore, selecting participants for up to 72 months ensures that most of the phonetic repertoire is acquired, reducing syllabic errors and phonological processes that influence syllable duration patterns. If the phonetic repertoire is fully acquired, target segments within a syllable will be produced accurately, reflecting the typical timing needed to articulate the syllable. However, if one or more segments are not acquired, the child may resort to phonological processes such as deletion or epenthesis. These compensatory strategies alter the syllable's structure, thereby affecting its duration. Table 3.1 provides details on the recruited participants including name tag, age, gender, and parent/caregiver name tag (for the child group).

Table 3.1: Participants

Name	**Age**	**Gender**	**Parent/ Caregiver**
1IG	24 months	Male	BAY
YRO	28 months	Male	SAL
M0N	28months	Female	DEN
DDK	30 months	Male	SON
2GS	32 months	Female	TAH
M37	32 months	Male	HAD
6VI	39 months	Male	SAB
DLE	42 months	Female	LAY
5PB	44 months	Female	ISL
G9I	51 months	Female	ZEI
3VO	51 months	Female	ZEI
P0X	49 months	Male	HAN
MD5	53 months	Male	AMN
I8T	53 months	Female	HAN
N8E	53 months	Male	HAL
6V9	60 months	Female	BAY
FDW	61 months	Female	RAH
JII	65 months	Male	HAN
MMA	66 months	Female	AHE
QR6	66 months	Female	AYA
3Q1	71 months	Male	HAL
6ED	23 years	Male	Adult
80X	24 years	Male	Adult
BMG	25 years	Female	Adult
MS2	27 years	Female	Adult

Three experiments were conducted: ST, PT, and RT. The child group completed all three tasks, while the adult group completed ST and RT only. The rationale for conducting three tasks was to account for speech sample diversity: ST has a natural speech flow and is less controlled compared to other tasks, with a focus on narration skills (Levelt & Van de Vijver 2004); PT has memory and visual stimulation effects (Ellis 2003); RT is more controlled and has a phonological proficiency assessment, with a focus on the ability to reproduce sounds, syllables, and words (Gathercole & Baddeley 1990). Performing multiple tasks is crucial in capturing variations in speech samples/patterns, vocabulary usage, and linguistic ability through different contexts (Dickinson & Tabors 2002, Hoff 2003). Thus, it allows for a thorough understanding of speech development, minimizing the risk of drawing conclusions based on a limited set of linguistic scenarios (Bornstein et al., 2014).

Task-related effects are expected to be evident in the results. To illustrate, ST is expected to reflect the children's ability to produce syllables in an interactive setting. PT is expected to show potentially less spontaneous productions, due to the emphasis on specific phonetic and phonological features predetermined by the chosen pictures. RT is expected to offer clearer insights into the development of superheavy syllables. Yet, RT may present complexities in analysing durational patterns, as the task involves the experimenter producing the target words first, and participants repeating them, potentially introducing a confounding effect due to imitations (Snow 1998). During the data analysis phase, recruited data points from the three tasks were organized according to the designated task. Data separation was considered to increase reliability measures, minimizing the impact of task-specific effects contributing to the data validity (Fletcher 1991). One sound recording, including the three experiments, was created for each session. Then, recordings were saved according to their designated groups (adult or child) and tasks (ST- semi-spontaneous speech task, PT-picture elicitation task, or RT-repetition task). The naming convention of the recorded sound files follows a three-letter/number name tag, age, gender, and task (e.g. AB9_30_M_rt).

3.2.Child Group

3.2.1. Semi-Spontaneous Speech Task

Materials

The experimenter used randomly selected questions from a list with child-oriented topics and interests, such as: “*Tell me about your favorite cartoon show. What is your favorite color? What is your favorite animal?”*. School-oriented topics were not used to ensure the use of a colloquial JA dialect instead of a standardized one.

Experimental Procedure

A semi-spontaneous speech was recorded for each participant for five minutes. To build rapport with the children, the experimenter, a female native speaker of JA, engaged in a closed context dialogue with each child. When the recording session started, the experimenter first spoke directly to the child to introduce her name and ask the child's name. Each child was assured that their recording session was only a regular chat and that all answers were correct to ensure they were not pressured. The experimenter asked the questions consecutively, and she gave the child sufficient time to think and answer the questions. In some cases, when children produced one-word answers, the experimenter asked further related questions. For example, if the child answered the question “Do you prefer playing inside or outside the house?” with only “inside” or “outside”, the experimenter asked, “Why?” to elicit further productions. Moreover, the experimenter asked questions about the participant’s surroundings, and parents/caregivers were encouraged to use toys and nearby objects to encourage responses. Parents/caregivers were instructed not to correct or interfere when the participant was speaking to guarantee spontaneous syllable productions. In cases where child participants repeated inaccurate but consistent productions during the recording session, parents were asked, if possible, to clarify them when the experimenter could not.

3.2.2. Picture Elicitation Task

Materials

For the second task, the experimental materials were a set of twenty pictures, including colorful pictures of animals, fruits and vegetables, numbers, shapes, and colors. The pictures were chosen from *My First Illustrated Dictionary (Arabic-English)*, which offers colorful pictures for the children to learn basic vocabulary in Arabic and English. Participants were asked to identify objects whose names include various syllable structures such as CVCC (/kalb/ ‘dog’), CVC.CVC (/dˤuf.dˤaʕ/ ‘frog’), and CV.CV.CVC (/sa.ma.kih/ ‘fish’); shapes: open (CV.CVC /ʔa.sad/ ‘lion’) and closed syllables (CVC.CVC /bis.sih/ ‘cat’); vowel length: short (CVC.CVC /ʔar.nab/ ‘rabbit’) and long vowels (CVVC /fiil/ ‘elephant’); clusters: onset (CCVVC /ħmaar/ ‘donkey’) and coda clusters (CVCC /kalb/ ‘dog’); and geminates (CVCC /xass/ ‘lettuce’). Phonemic sub-classes of JA’s phonetic inventory, such as plosives, fricatives, affricates, nasals, liquids, glides, and emphatics were included. Furthermore, the experimenter and the supervisors carefully selected and reviewed the pictures to confirm they were age-appropriate for JA children. The chosen pictures were intended to be familiar to children from 24–72 months of age as they avoid abstract representations (e.g., seasons or feelings), and only fall within the Jordanian context and not from any foreign contexts (e.g., igloos or polar bears).

Experimental Procedure

The experimenter used the ‘Sharing Screen’ option on Microsoft Teams to present one picture at a time, and participants were asked to identify each picture. If the participant had difficulty identifying the picture, s/he was given clues about it. For example, if the participant could not identify an animal, the experimenter described the animal (e.g., this animal is big), used an onomatopoeic sound representing the animal (e.g., this animal meows), or used a context that has the animal to make the identification process easier (e.g., this animal lives on the farm). On five occasions, parents/caregivers reminded their children of situations and memories linked to the objects in the pictures to help the identification process. If participants could not identify the picture, the experimenter uttered the name of the object and asked the participant to repeat the word. However, the experimenter and caregivers did not attempt to correct the child's responses.

3.2.3. Repetition Task

Materials

A repetition task focusing on superheavy syllables was used. A total of 46 superheavy syllables occurred in 36 monosyllabic words, 9 disyllabic words, and one multisyllabic word comprised the word list (Appendix: Repetition Task Word List). Target words with complex constituents, such as geminates, clusters, and long and short vowels were considered. The following words are examples of syllable structures included: CCVVC (/ktaab/ 'book'), CV. CVVC (/ma.laak/ 'angel'), CV.CVCC (/ma.ħall/ 'shop'), CV.CV.CVVC (/sa.ka.kiin/ 'knives'), CVVC (/baab/ 'door'), CVCC (/bint/ 'girl'), and CVCC (/dubb/ 'bear').

Experimental Procedure

Participants were asked to repeat at least ten words chosen randomly by the experimenter from the list. Participants in the child group showed different levels of willingness to repeat words, particularly in the two youngest groups ranging from 24–30 months and 31–37 months. Therefore, the number of repeated words across the age groups varied. Participants were asked to repeat the words as they were uttered sequentially by the experimenter. If they could not perceive a word, the experimenter would first attempt to repeat it to elicit a clearer production. If a child participant failed to respond satisfactorily after the second attempt, parents were instructed to repeat the word for the child for more clarity. The experimenter did not correct any productions, and parents/caregivers were instructed not to interfere.

3.3.Adult Group

3.3.1. Semi-Spontaneous Speech Task

The same method of data recording used in the child group was applied to the adult group. The experimenter asked random questions from a list with various topics, such as: "*Tell me about your favorite movie. What is your favorite memory with your friends? How did the COVID-19 pandemic affect your community?*" to each participant for five minutes.

3.3.2. Repetition Task

The same word list for the child group was used for the adult group. The experimenter used the 'Share Screen' option during the Microsoft Teams meeting to show the table of target words. Participants were asked to read aloud the word list and were instructed to produce the words in their colloquial dialect, not a standardized one.

3.4.Data Analysis

The data analysis section is two-fold: the first includes the acoustic analysis conducted using Praat Software; the second concerns the statistical analysis carried out using R Software and RStudio employing Bayesian Multi-level modelling.

3.4.1. Acoustic Analysis

This section presents the strategies for identifying words and phoneme-level segments, Praat textgrid annotation, and the segmentation criteria. Some annotation examples of complex segmentation cases are presented.

3.4.1.1.Word identification and text-grids

Praat Software (version 6.1.51, Boersma & Weenick 2020) was used for all audiovisual and acoustic analyses of the speech data. Word identification was conducted using context-based and shape-based criteria following Vihman & McCune (1994) and Alqattan (2015). The context-based criteria involved identifying sequences of syllables as words if their plausible meaning could be inferred from the context. For example, Alqattan (2015) suggested that onomatopoeic productions (phonetic imitations that are linked to meaningful productions) are considered identifiable words. For instance, a participant identified a picture of a donkey as /maaʕ/ instead of /ħmaar/, and another participant used the onomatopoeic sound /ʕaww/ instead of /kalb/ in identifying a picture of a dog. Shape-based criteria relied on assessing the similarity between child and adult forms. In this approach, words were analysed if they were audible and identifiable by the experimenter even when they did not match the adult form. One child repeatedly but consistently produced the word /gur.gaʕ/ instead of /dˤuf.dˤaʕ/, wherein the child's production was recognized by his parent as an identification of a picture of a frog. The aforementioned tokens were identified as words in the analysis.

For data annotation, the experimenter carried out audiovisual analysis using spectrograms and textgrids in Praat. The spectrogram settings were set to view the frequency range 0-8000 Hz, with a window length of 0.005s and a dynamic range of 50 dB. A custom segmentation criterion was generated for annotations. This was necessitated by the need to avoid special International Phonetic Alphabet (IPA) characters, such as /zˤ/, /ħ/, and /ɣ/, that might not be recognized across the chosen programs for data analysis. As shown in Figure 3.1, the textgrid included eight tiers: Utterance, Word, Syllable, Production, Transcription, Position, Vowel, and Comment. The word, syllable, and vowel boundaries were annotated to measure their duration.

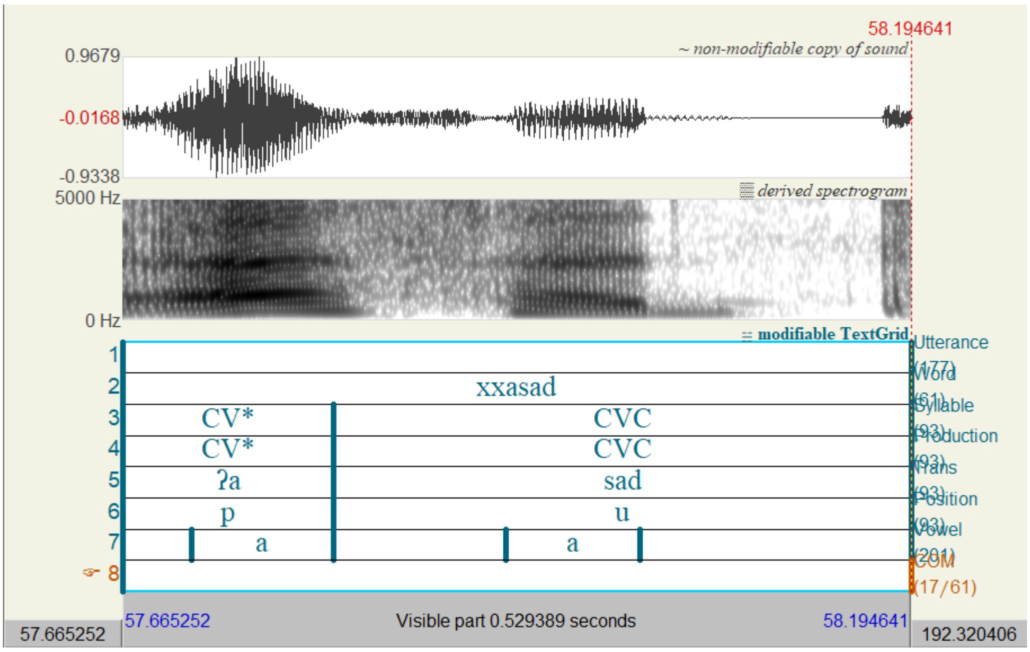


Figure 3.1: Textgrid demonstrating the tiers used for data acoustic analysis

Utterance tier: While the tier ‘utterance’ was created for annotating utterances or intonational phrases, its utilization for data analysis was deemed impractical. Children mainly produced one-word utterances, and therefore the ‘utterance’ coding was redundant. In ST, the majority of the answers for children, specifically the younger age groups, were one-word responses (Section 3.4.3).

Word tier: Used to annotate the identified words. The phoneme-level segmentation was used to set the onset and offset boundaries of each word.

Syllable tier: Used to annotate the adult form of the word’s syllabic structure. the asterisk was used to denote lexical stress. Stressed syllables were identified by JA metrical rules based on parameters such as syllable structure, weight, and position in addition to constraints such as ERR and WPR (Section 2.1.3.2.2). For example, Figure 3.1 demonstrates the production of the word /ʔa.sad/ ‘lion’ CV.CVC, where the adult and child forms match. Since JA has a trochaic foot parsing that goes from left to right, and both syllables are light, stress falls on the penultimate syllable /ʔa/, and the asterisk (CV*) was added to both production and syllable tiers. Nonetheless, Figure 3.2 shows the child production of the monosyllabic word /tiix/ ‘watermelon’ CVVC. The adult target form is disyllabic /batˤtˤiix/ CVC.CVVC. In the child form, the participant deleted the penultimate syllable and the stress was assigned to the only superheavy syllable. However, the adult form has stress assigned to the penultimate syllable as it contains a geminate, which attracts and receives the main stress based on JA stress rules. The asterisk placement in the ‘Syllable’ and ‘Production’ tiers is different due to the process of syllable deletion.

__Production tier:__ Used to annotate the child’s syllable production form and the marked syllable duration boundaries were used as the dependent variable. As in the Syllable tier, the asterisk marked a stressed syllable. The ‘Syllable’ and ‘Production’ tiers were used to compare the child production and the adult target forms to determine the occurrence of any phonological processes.

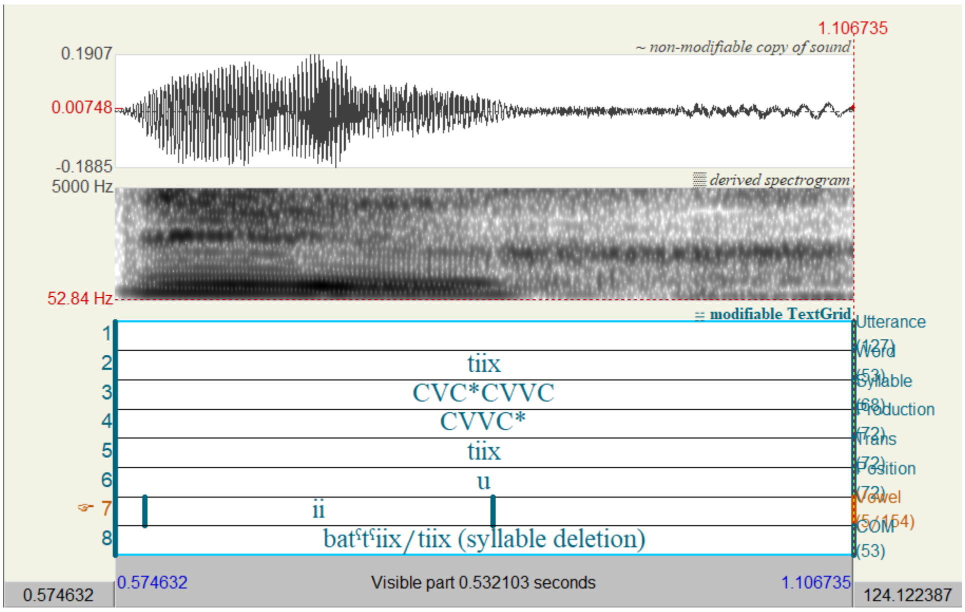


Figure 3.2: Textgrid demonstrating the difference in syllable and production tiers

__Transcription tier (Trans):__ Used to record the IPA transcriptions of syllables.

__Position tier:__ Used to mark the syllable’s position within a word (u = ultimate; p = penultimate; a = antepenultimate; pre = preantepenultimate; pro = propreantepenultimate).

Vowel tier: Used to mark the boundaries for vowels based on their acoustic markers and the surrounding segments.

Comment tier (COM): Used to document any comments on the production form, including the occurrence of phonological processes. In such cases, a summary of both target and production forms in addition to the process occurring were annotated (e.g., batˤtˤiix/tiix: syllable deletion). Three Praat scripts were used to process and extract the annotated data. The first script (save_selection_to_sound_and_textgrid) was used to save selected productions into WAV and textgrid files to organize the productions according to the tasks. Second, to extract data from the annotations on Praat, the script (read_text_from_interveal_tiers) was used. The script interprets the following: file name, word, syllable structure, production, transcription, and comments. In addition, the (v_syll_duration) script reads the file name, word, word duration, syllable structure, production, transcription, syllable duration, vowel, vowel duration, syllable position, and comments. Excel sheets were used to document the readings from the scripts for further data analysis on R Software.

3.4.2. Acoustic markers for phoneme-level segmentation

The segmentation process included transcribing speech using the IPA and annotating syllable and vowel boundaries. Below are the main acoustic cues for the manners of articulation in JA including fricatives, stops, nasals, affricates, approximants (liquids and glides), and vowels. In cases of word medial geminates (CC), the onset and offset points were selected on the textgrid, and the duration of each segment was determined by moving the cursor to the mid-point, using the ‘Set cursor to’ option in the Time-Set selection option.

3.4.2.1.Fricatives

The friction noise caused by the turbulent airflow through a narrowed vocal tract (Harrington 2010) was considered a marker for the presence of a fricative. To distinguish sibilant and non-sibilant fricatives, the amplitude characteristics of the target fricative were considered. In

Figure 3.3 /farawlah/ ‘strawberry’, the static noise-like shape was evident in the spectrogram; the low average frequency, compared to what is expected to be exhibited in /s/ or /ʃ/; and the aperiodic signal attributed to the lack of voicing, were the main cues to identify /f/.

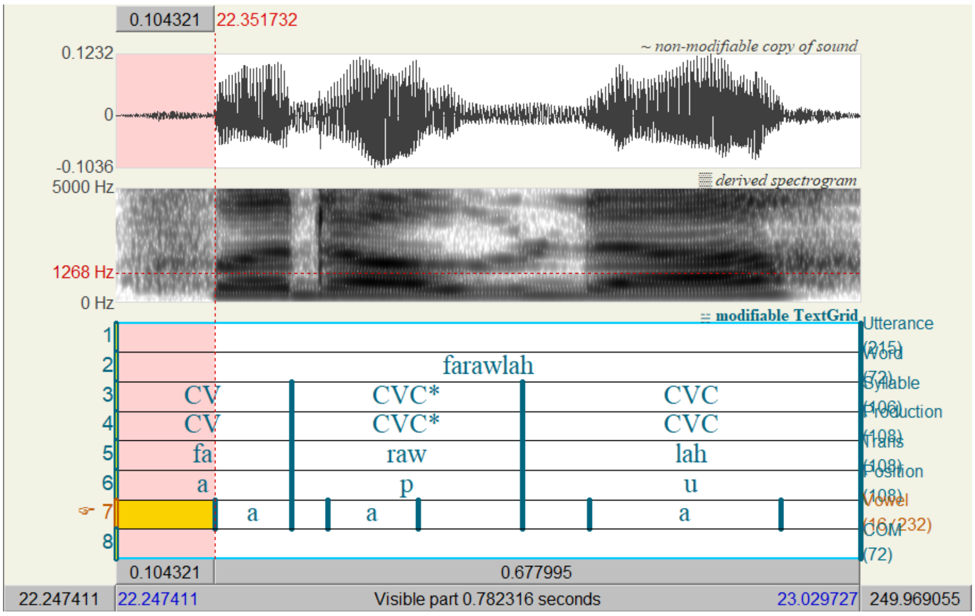


Figure 3.3: Segmentation of a fricative in /farawlah/

3.4.2.2.Stops

The major cue for voiceless stops was the stop burst (Harrington 2010). The significant portion of silence preceding a burst of noise marked the onset of the stop. The offset of the stop was at the release of the stop closure. To identify stop closures, a lowering in F2 frequency compared to adjacent vowels or consonants was considered. In Figure 3.4 showing /samat/ ‘fish’, the spectral characteristics of plosive /t/ included the drop in the periodic energy following the vowel /a/; the short release burst of noise; and the silent closure phase, during which the airflow is completely blocked by the tongue at the alveolar ridge.

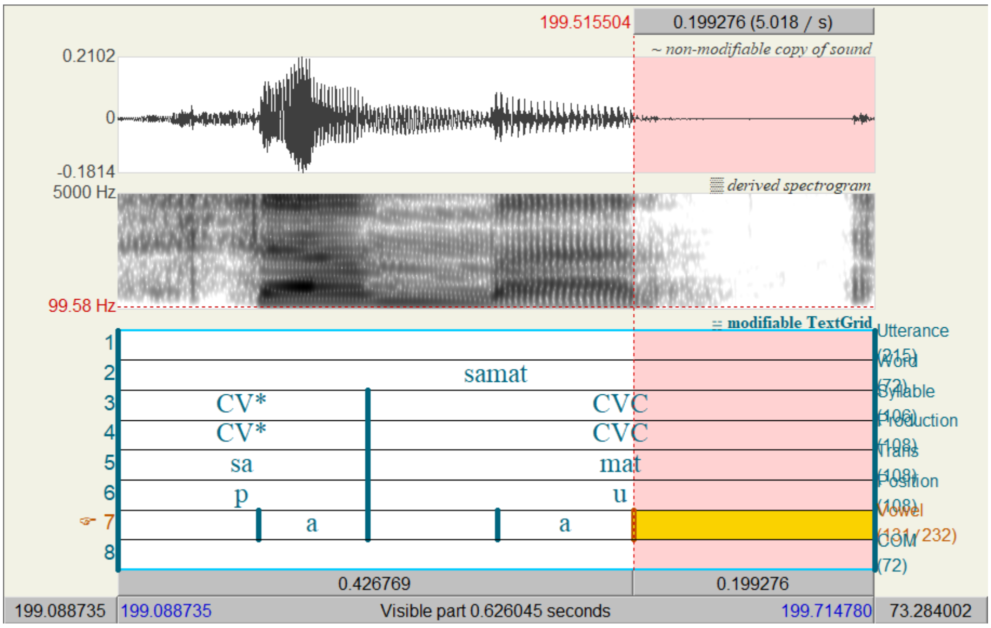


Figure 3.4: Segmentation of a stop in /samat/

3.4.2.3.Nasals

Nasals exhibited visible formant patterns that were not as clear as those of vowels. Nasals were identified by the lack of upper formant energy in the spectrograms (Harrington 2010). The onset of nasals often had a sharp discontinuity in the waveform, reflecting the closure of the nasal cavity by other articulators (e.g., tongue or lips), resulting in a lower amplitude (Duanmu 1994). F1 is typically lower compared to vowels, around 200Hz (Al-Zabibi 1990). The smooth periodic patterns of the acoustic waveform were a clear marker for nasals as they differed from vowels which have high-frequency fluctuations corresponding to high-frequency energy. Moreover, nasals have abrupt spectral changes at closure onset and release coinciding with a brief v-like dip followed by a rise in the waveform (Sudhoff et al., 2006). In Figure 3.5 for /muuz/ ‘banana’, segmenting nasal /m/ using the spectrogram was challenging, whereas using waveforms provided a clearer distinction. The ‘Show Formant’ option aided in identifying the lower F1 regions for nasal /m/ compared to the following long vowel /uu/, with F1 frequency being around 239.8Hz.

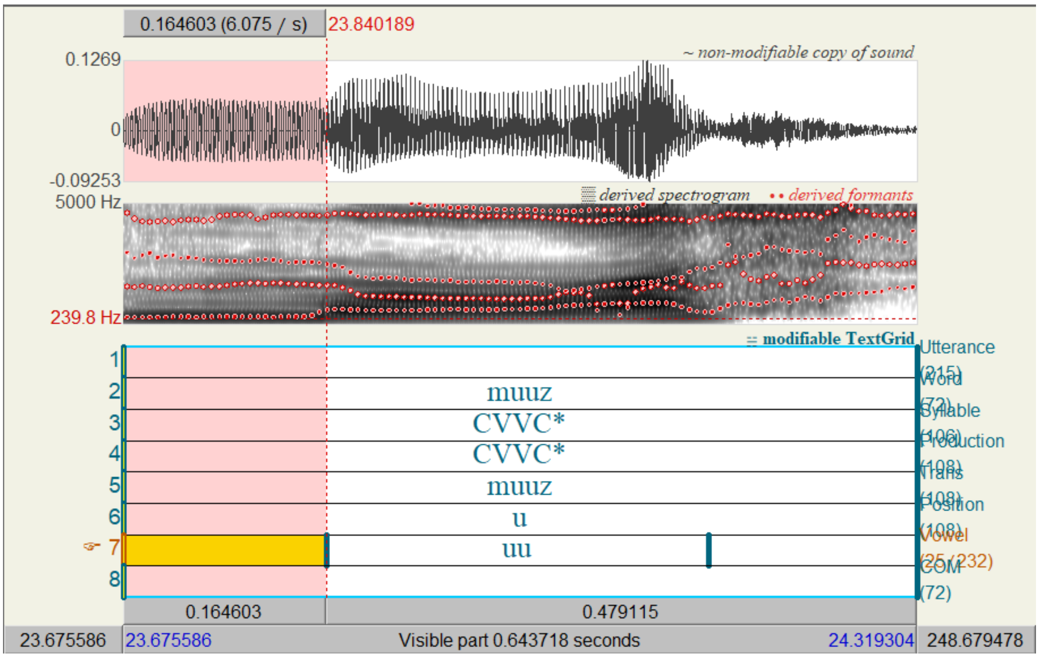


Figure 3.5: Segmentation of a nasal in /muuz/

3.4.2.4.Affricates

Affricates show a rapid rise in the amplitude at the onset phase, with random aperiodic noise for a prolonged period (Harrington 2010). In the example /nad͡ʒmah/ ‘star’, the energy appeared to be concentrated in the mid-to-high frequency range, around 3000Hz; the spectral shape resembled a band of noise representing the turbulent airflow; vertical striations in the spectrogram (Harrington 2010); and the absence of F1 frequency in /d͡ʒ/ region, but the presence of F1 for the surrounding sounds /i/ and nasal /m/ were the main cues to mark the onset and offset of the segment.

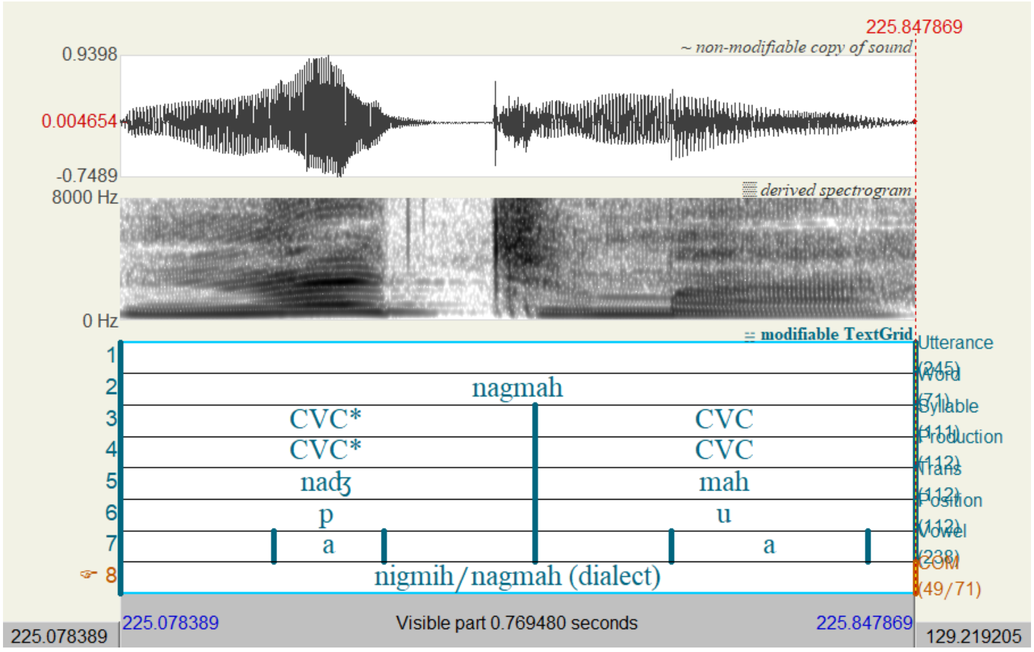


Figure 3.6: Segmentation of an affricate in /nad͡ʒmah/

3.4.2.5.Approximants

Although approximants have a similar formant structure to vowels, F1 tends to be mid-to-high, F2 is low, and F3 is low in frequency (Lehiste 1964). In the example /murabbaʕ/ ‘square’, the major cues for the trill /r/ were the relative silence phase with low intensity; in the trill phase, more periodic energy bands appeared in the spectrogram (Espy-Wilson 1994); and the shift in the F2 region to a higher frequency transitioning between /r/ and /a/ is apparent, rendering the increased F2 values as a marker for the onset of the vowel.

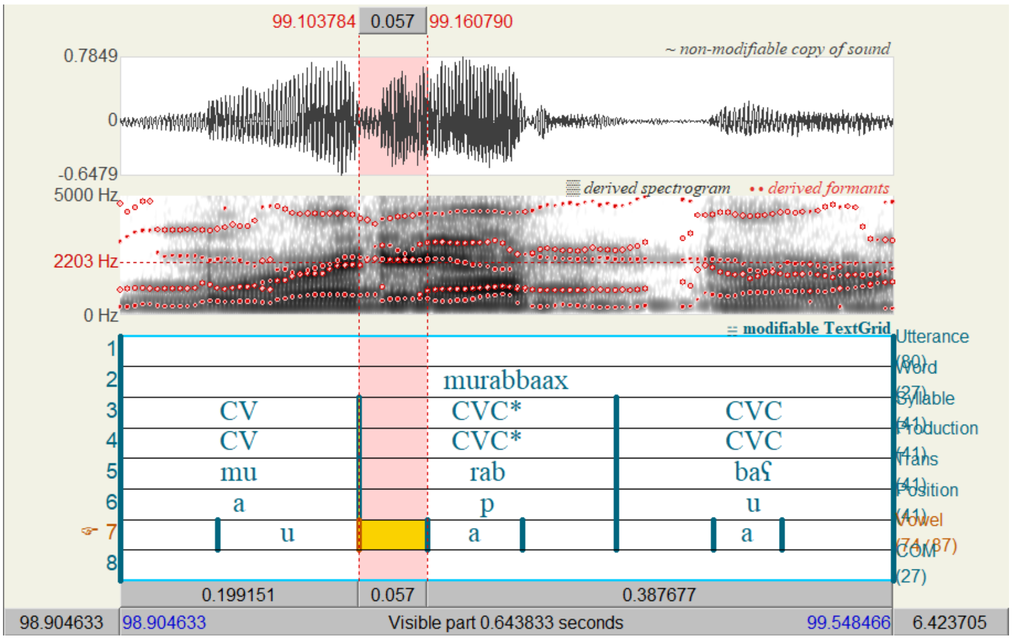


Figure 3.7: Segmentation of an approximant in /murabbaʕ/

3.4.3.5.Vowels

A visual examination of the spectrogram including the visible F1 and F2 in high-frequency areas and the waveform was conducted to detect the vowel interval. If the formants were not clear enough, formant listing was resorted to as formant values aid in marking the onset and the offset of the vowel compared to the surrounding sounds. The endpoint of the vowel was detected by visually examining the last glottal pulse in the spectrogram. In cases of vowels following voiceless stops, the onset of the vowel was marked at the consonantal release (Sudhoff et al., 2006). In the example /ʔaħmar/ ‘red’ (Figure 3.8), the cues identifying the vowel’s onset and offset were the high F1 frequency around 800.6Hz compared to the surrounding segments /m/ and /r/, and the periodic vocal fold vibrations.

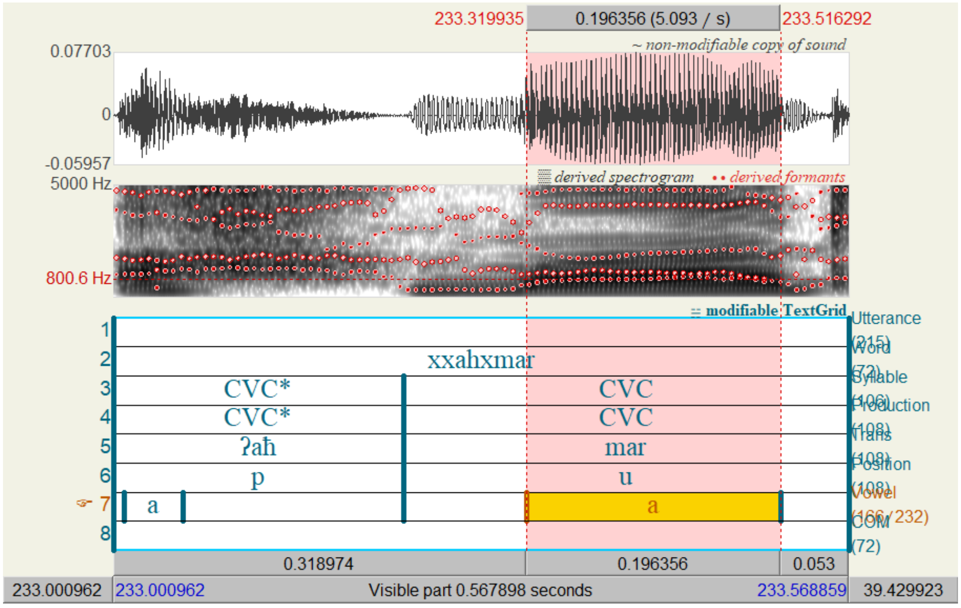


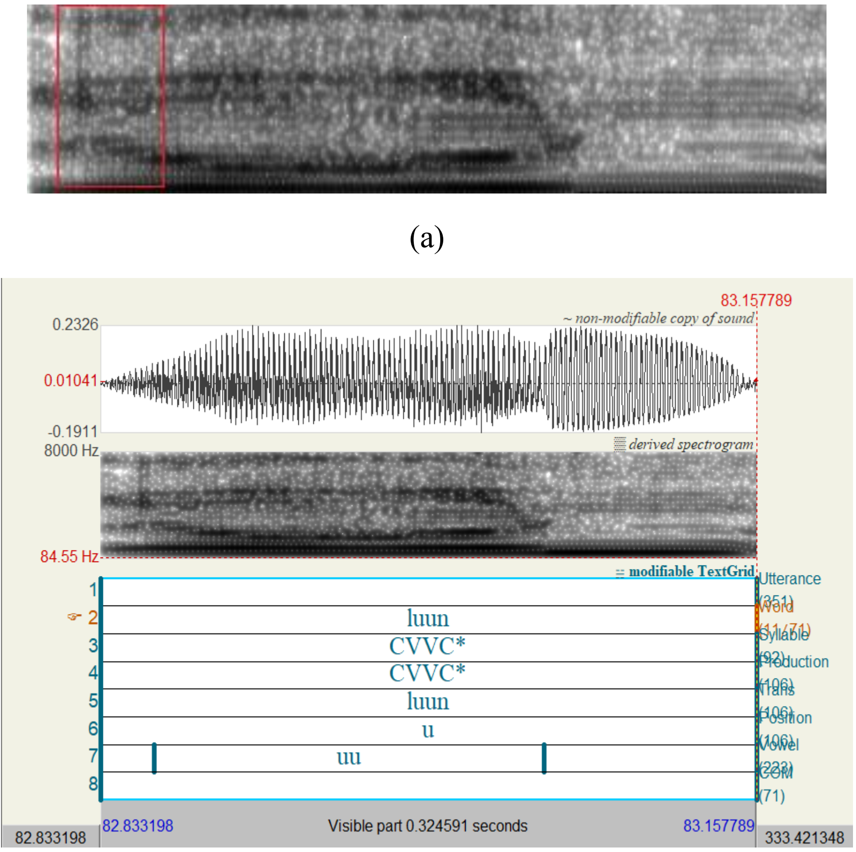
Figure 3.8: Segmentation of a vowel in /ʔaħmar/

3.4.4. Segmentation complexities in JA productions

The examples below demonstrate how segmentation was carried out for a sequence of segments with acoustic similarities such as vowel-liquid, vowel-glide, and vowel-trill sequences.

Liquids and vowels

Liquids and vowels (e.g., in /luun/ ‘color’, the liquid /l/ and /u, Figure 3.9) were challenging to segment when they occurred consecutively. They showed formant frequencies and gradual formant transitions (Ladefoged 2001, Kent & Read 2002). A closer examination of the spectrogram demonstrated the presence of the horizontal band, indicative of the resonance associated with /l/. In addition, the formant frequency transition between /l/ to /u/ aided in the visual analysis of the segmental boundary.



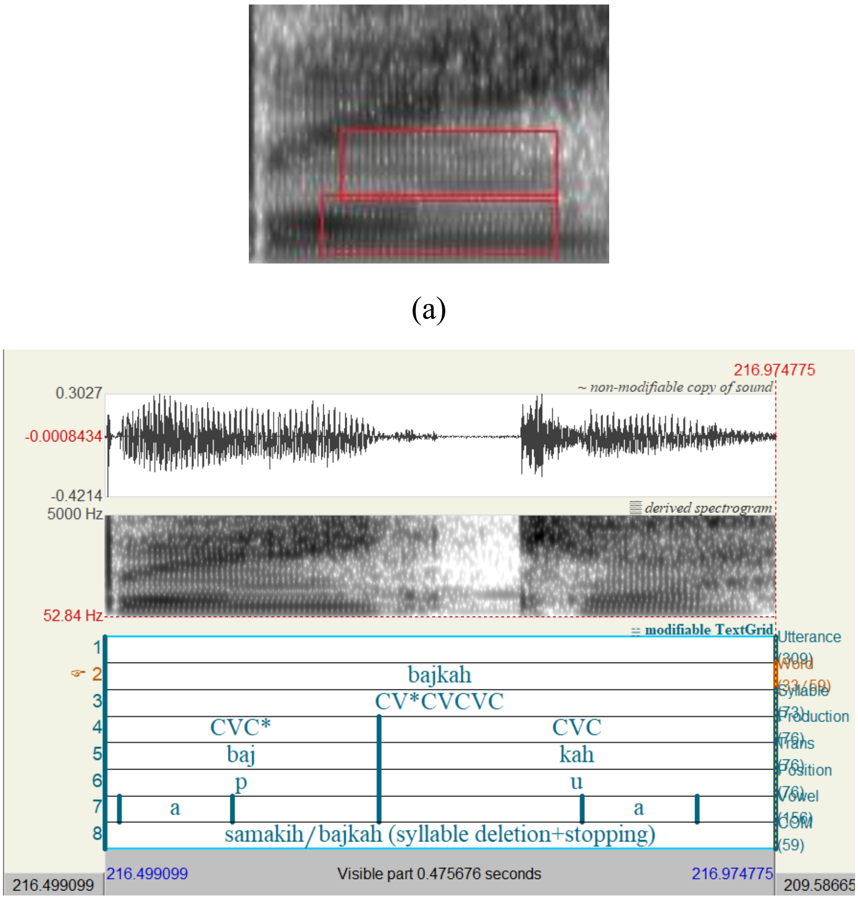
(a)

(b)

Figure 3.9: (a) A detailed waveform (b) spectrogram of a liquid-vowel segmentation

Glides and vowels

In the case of the word /bajkah/ 'fish' (Figure 3.10), the boundary between the vowel /a/ and the glide /j/ was not clear. The main cue for marking the vowel offset and the onset of /j/ was the change in the formant frequency, where lower formant frequencies are evident for /j/ compared to /a/.



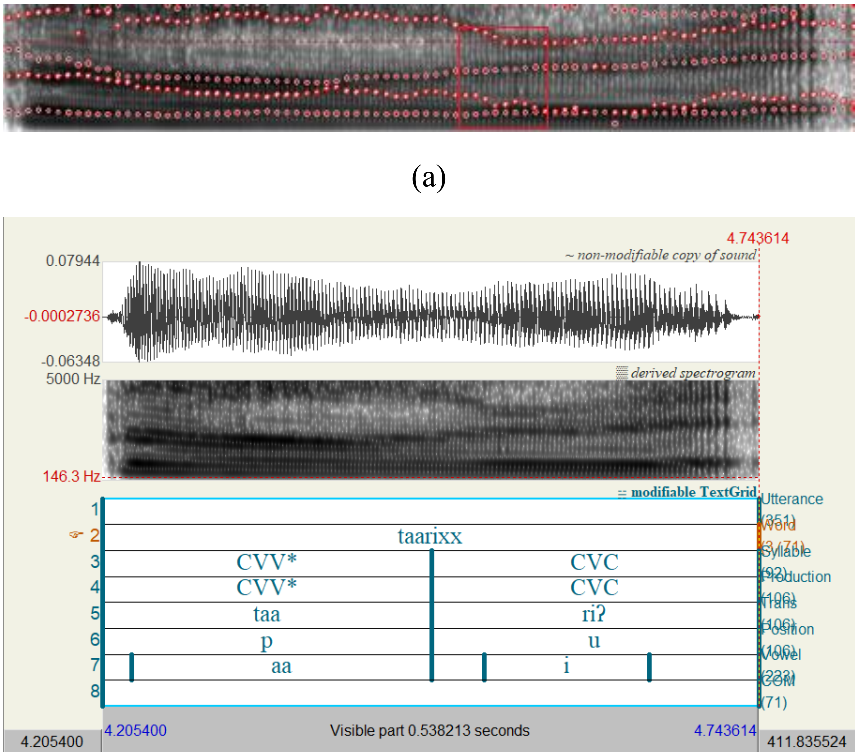
(a)

(b)

Figure 3.10: (a) A detailed waveform (b) spectrogram of a glide-vowel segmentation

Trills and vowels

The expected regular patterns of vocal fold vibrations and the accompanying small closures between them for /r/ were not visually evident in Figure 3.11. In the production of /taariʔ/ ‘Tariq-Name’, marking /r/ which had /a/ and /i/ as surrounding vowels was challenging. The lower formant frequencies compared to the surrounding vowels were the main cue to locate /r/.



(a)

(b)

Figure 3.11: (a) A detailed waveform (b) spectrogram of a trill-vowel segmentation

3.4.2. Phonological Processes Analysis

The phonological processes were analysed to explore linguistic development and the phonological characteristics specific to JA typology. A phonological process refers to a consistent pattern of sound modification or simplification resulting in a deviation from the standard target form (Alqattan 2015, Mashaqba et al., 2019). Following the classification proposed by al-Huneety (2023) and Alqattan (2015), the phonological processes were classified as segmental and prosodic. Segmental processes do not alter the syllable structure or the word shape of the target form (al Huneety et al., 2023). These include processes such as metathesis, assimilation, gliding, stopping, backing, fronting, and de-emphasis. On the other hand, prosodic processes alter the syllable structure or word shape of the target form (Alqattan 2015). These include processes such as vowel epenthesis, assimilation and gemination, syllable deletion, cluster reduction, weak syllable deletion, coda deletion, and syncope. The identification and definition of the phonological process were based on previous studies' descriptions by Amayreh and Dyson (1998), Dyson and Amayreh (2000), Amayreh (2000), Alqattan (2015), and Mashaqba et al. (2019). The age group, syllable position, lexical stress, and syllable structure that the processes occurred in were documented for further descriptive analysis in Chapter 4. The following spectrograms demonstrate examples of phonological processes observed.

Segmental processes

First, for segmental modification, Figure 3.12 shows the production of /sulfaħah/ 'turtle' instead of /sulħafah/, which was analysed as metathesis, the reordering of the consonantal segments (Alqattan 2015).

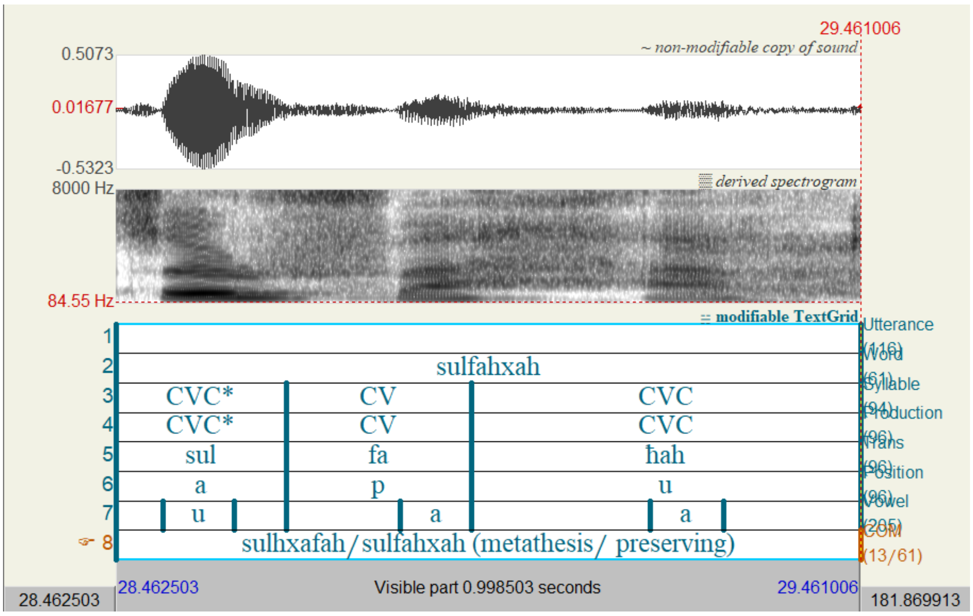


Figure 3.12: Segment modification- metathesis

Another example in Figure 3.13 is the production of /banduurah/ 'tomato' as /banduulah/, showing lateralization, the substitution of /l/ for /r/ (Alqattan 2015).

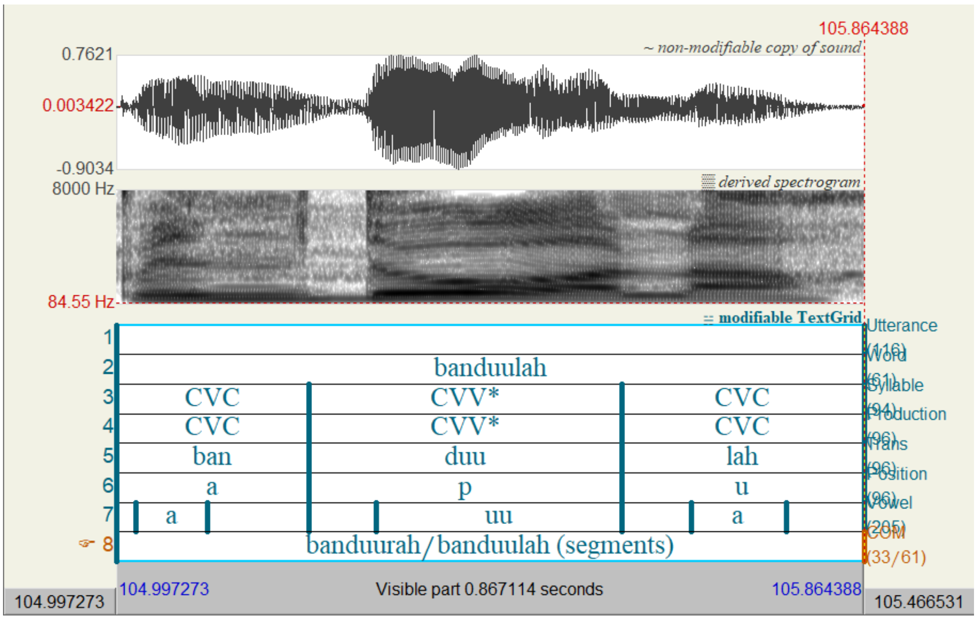


Figure 3.13: Segment modification- lateralization

Figure 3.14 shows the production of /muuzih/ 'banana' as /buuzih/, where stopping is evident as nasal /m/ was produced as the stop /b/.

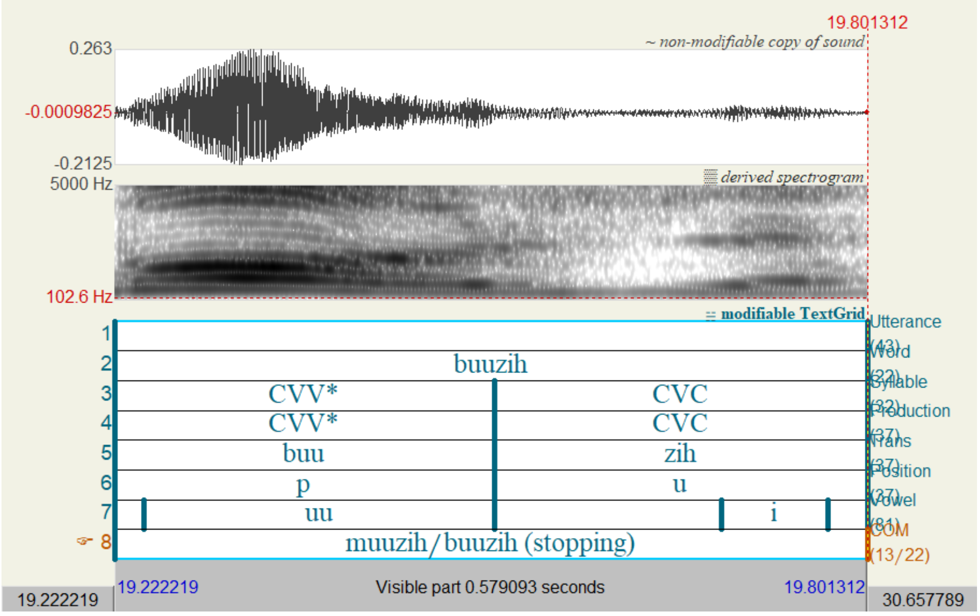


Figure 3.14: Segment modification- stopping

Figure 3.15 shows the process of assimilation, where the child produced /farawlah/ 'strawberry' as /fawawlah/, assimilating the trill /r/ as approximant /w/.

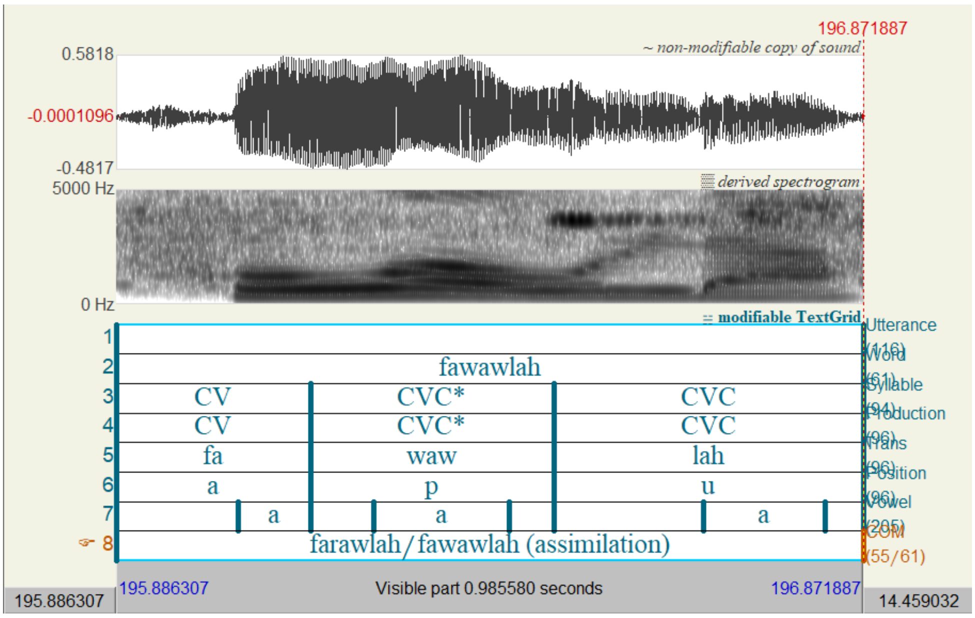


Figure 3.15: Segment modification- assimilation

Prosodic processes

Figure 3.16 presents an example of vowel epenthesis, where a vowel is inserted to break up a consonantal cluster (McCarthy 2008). The target form /ħsˤaan/ ‘horse’ CCVVC was produced as /ʔisˤaan/ CV.CVVC. The cluster /ħsˤ/ is broken up by the epenthetic vowel /ħ(i)sˤ/, and the fricative /ħ/ is produced as glottal /ʔ/ creating a disyllabic word including a light syllable, /ʔi/, and a superheavy syllable, /sˤaan/, instead of a monosyllabic word.

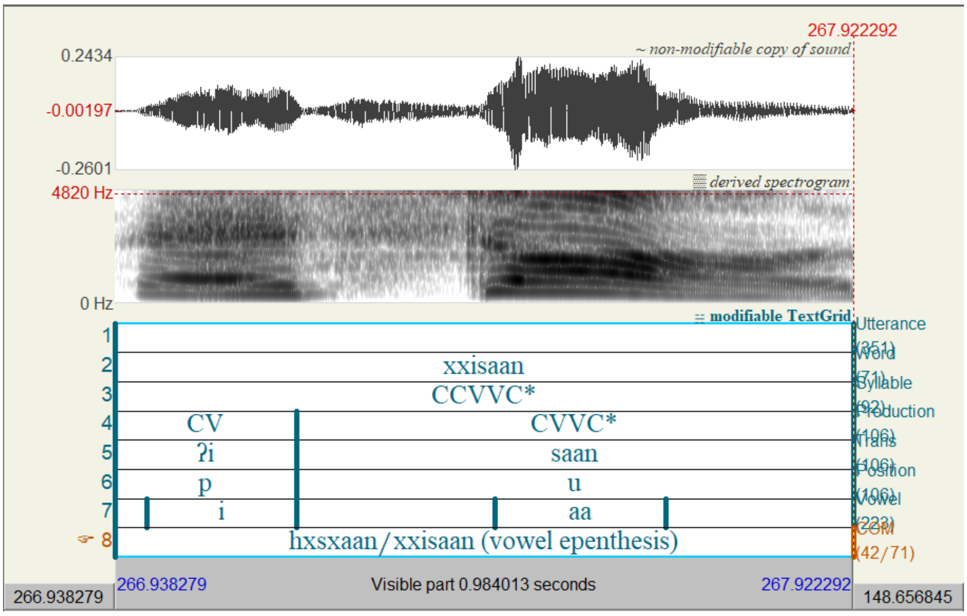


Figure 3.16: Vowel epenthesis

The process of assimilation and gemination refers to the tendency of JA children to assimilate the coda of the first syllable with the onset of the proceeding syllable creating a geminate at the syllable boundary (Mashaqba et al., 2021). Figure 3.17 shows the production of /ʔannab/ ‘rabbit’ instead of /ʔarnab/. The trill /r/, occupying a non-final coda position, was assimilated as nasal /n/, which occupies the onset position of the following syllable. This process results in stress being assigned to the geminate, regardless of the syllable structure of the other syllables.

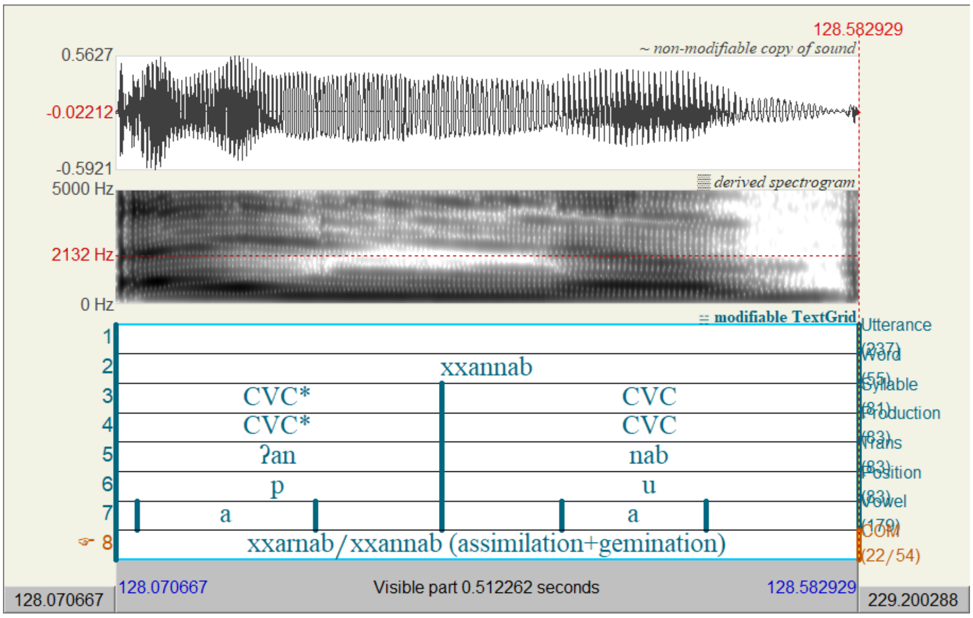


Figure 3.17: Assimilation and gemination

Syllable deletion refers to the modification of the word’s syllabic shape, where a syllable or more are deleted (e.g., stressed and heavy/superheavy syllables may be deleted, Dodd et al., 2003). Figure 3.18 demonstrates the production of /tˤiix/ ‘watermelon’, which has /batˤtˤiix/ as a target form. The target word shape is CVC.CVVC, but the deletion of the heavy and stressed syllable /batˤ/, preserving the superheavy syllable /tˤiix/, resulted in a monosyllabic (CVVC) word.

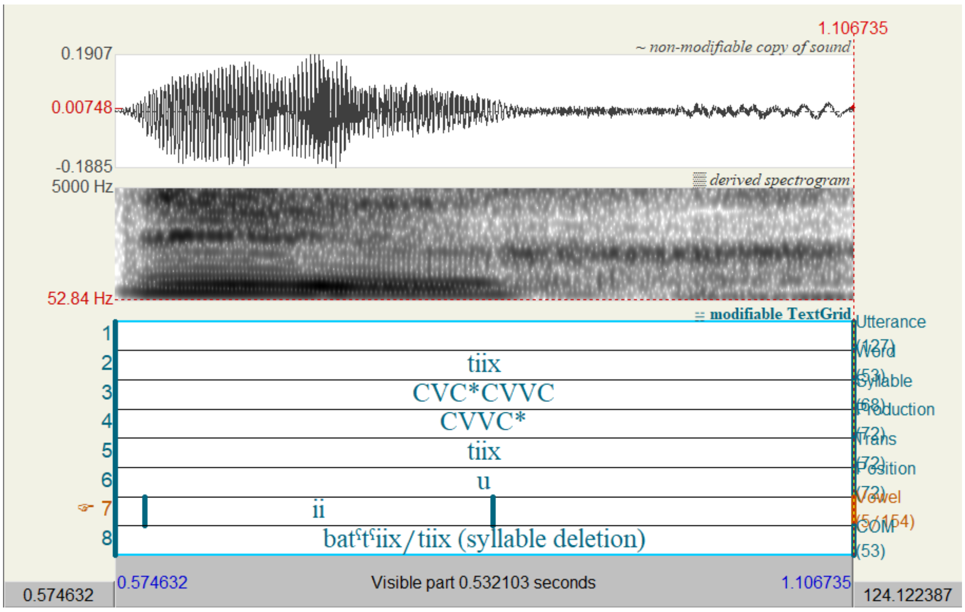


Figure 3.18: Syllable deletion

Cluster reduction involves the deletion of one or more segments, resulting in a simplified structure of that cluster (Alqattan 2015). Figure 3.19 demonstrates the production of /sˤaan/ CVVC ‘horse’, which has a target form of /ħsˤaan/ CCVVC. The cluster is reduced by deleting the first segment, /ħ/, and reserving the second, /sˤ/.

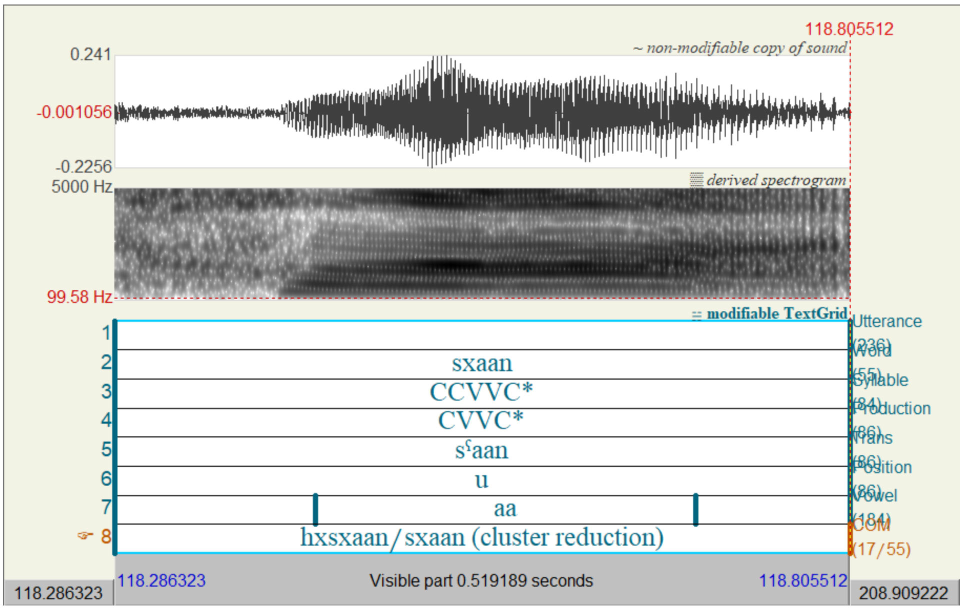


Figure 3.19: Cluster reduction

Weak syllable deletion refers to the deletion of unstressed or light syllables from the target words (Dodd et al., 2003, Alqattan 2015). Figure 3.20 shows /maamih/ ‘dove’ CVV.CVC, which has the target adult form /ħamaamih/ CV.CVV.CVC. The light non-final unstressed syllable /ħa/ was deleted, preserving the stressed syllable /maa/ and the final syllable /mih/.

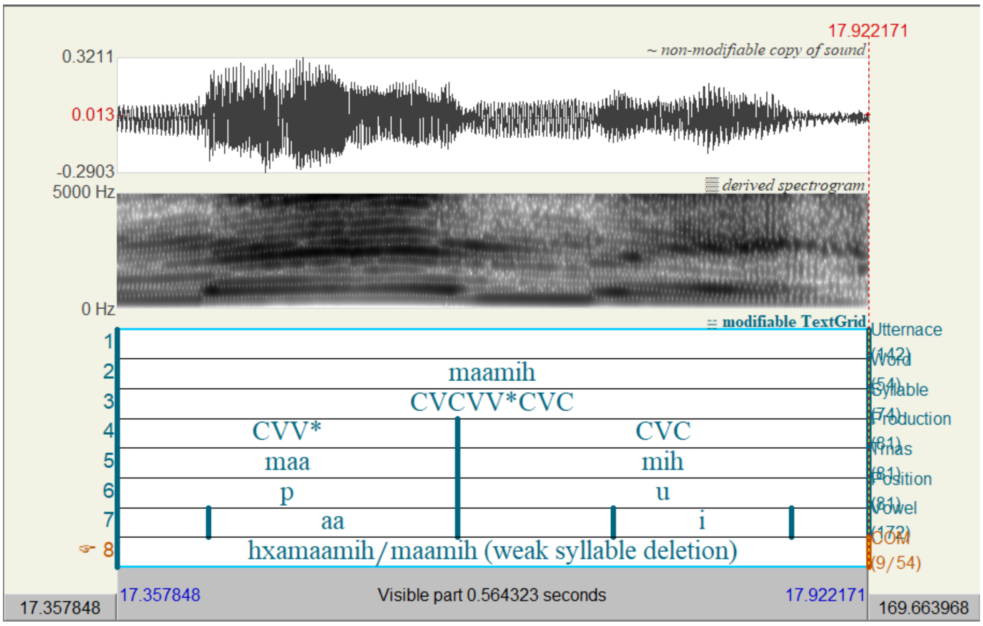


Figure 3.20: Weak syllable deletion

Coda deletion refers to the deletion of the final consonant of a syllable (Alqattan 2015). Figure 3.21 demonstrates the production of /ʔadˤdˤa/ ‘green’ CVC.CV, which has the target form of /ʔaxdˤar/ CVC.CVC. The target syllable /dˤar/ CVC was produced as /dˤa/ CV, deleting the coda /r/.

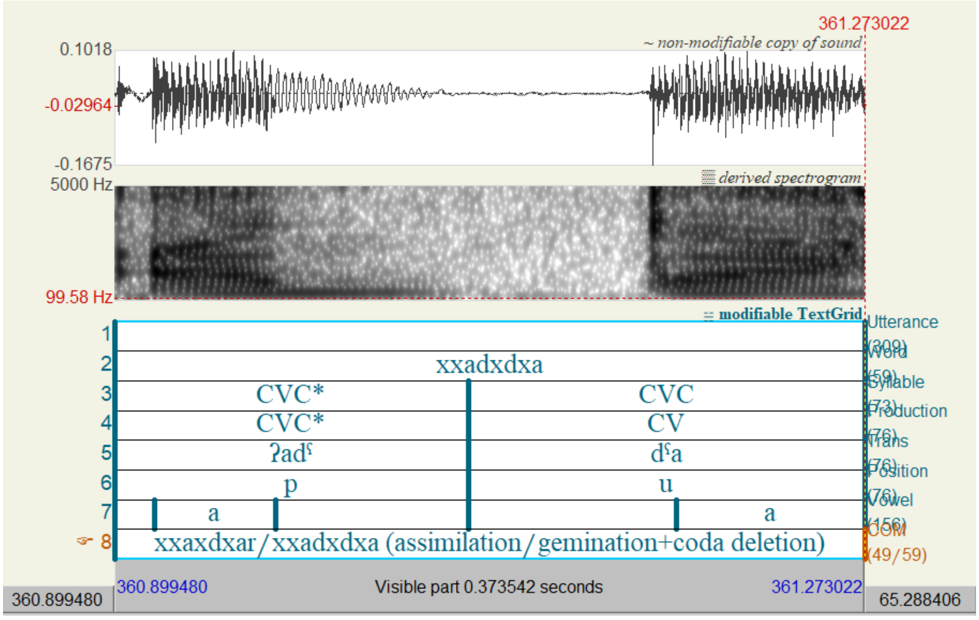


Figure 3.21: Coda deletion

Syncope is the deletion of a vowel typically in short unstressed syllables (Ferguson, Menn & Stoel-Gammon 1992). For example, Figure 3.22 shows the word /ʃlaatah/ ‘chocolate’ CCVV.CVC, which has a target form of /ʃukalaatah/ CV.CV.CVV.CVC. The production form has the deletion of the weak syllable /ka/ in addition to syncopating the vowel from the preantepenultimate syllable /ʃu/. The weak syllable deletion and vowel syncope resulted in the formation of the onset cluster /ʃl/.

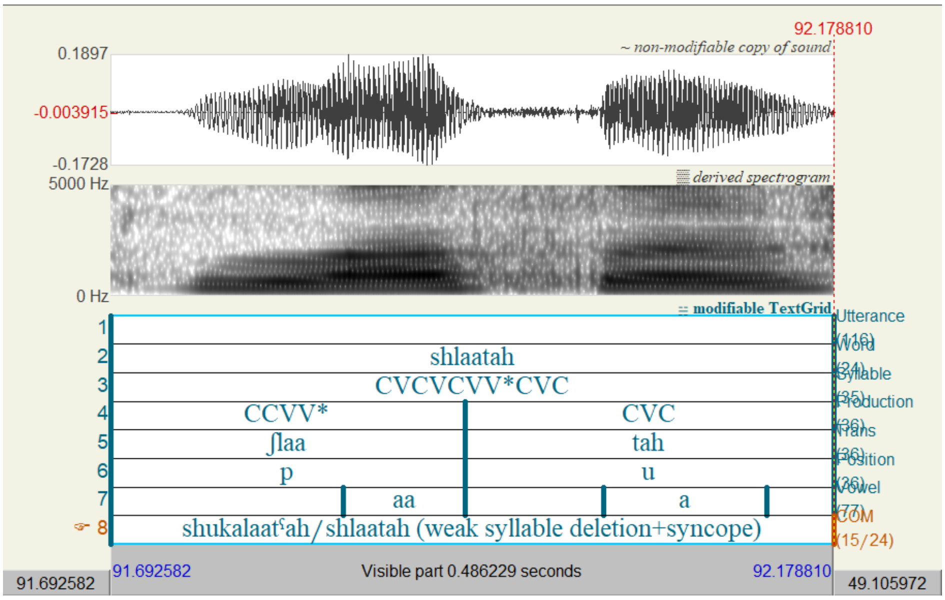


Figure 3.22: Syncope

3.4.3. Statistical Analysis

Using R Software (R Development Team 2013) and RStudio (2022.07.2, Build 576), Bayesian multi-level models were employed for several reasons. First, they are suitable for analysing hierarchical data structures commonly encountered in language development research (Gelman, Hill & Vehtari 2021). The data often involves multiple observations nested within individual participants, as well as participants within various groups (e.g., age groups, genders). Second, Bayesian models are robust at handling the inherent variability in data (resulting from children being divided into seven age groups, and data comprising three tasks) by incorporating random effects at multiple levels of analysis (Winter 2013, Gelman & Shalizi 2013). This capability allows for group-level pattern analysis. Third, Bayesian modeling provides a robust framework for dealing with small sample sizes. Unlike traditional frequentist statistical tests, Bayesian methods do not encounter issues related to small sample sizes and offer more reliable estimates of parameters (Hoijtink, Klugkist & Boelen 2008). Fourth, Bayesian models allow for prior information incorporation about parameters of interest, enhancing the stability and efficiency of estimates (Winter 2013). Finally, Bayesian models estimate the probability of predictors' uncertainty yielding credible intervals and posterior predictive distributions, thereby offering more interpretable results compared to traditional statistical approaches (Gelman & Shalizi 2013).

4.6.1. Bayesian Multi-Level Modelling

Bayesian multi-level models were fitted to the data to determine the effect of predictors affecting syllable and vowel durations using the R program (R Core Team 2013), Stan modelling language (Carpenter et al., 2017), and the brms package (Buerkner 2017). The models estimated the relationship between syllable/vowel duration as the dependent variable and four main independent variables. The independent variables were age group (a categorical monotonic predictor with eight levels coded as 24, 31, 38, 45, 52, 59, 66 months, and adult) (Section 4.6.2); stress (a binary categorical predictor with stressed and unstressed levels); syllable position within a word (a categorical predictor with two levels, word-final and non-final)[18]; and syllable structure (a categorical predictor with three levels light, heavy and superheavy syllables). For the superheavy syllable analysis, additional predictors were considered including sub-structure (a categorical predictor coded as CVCC, CVVC, CCVVC, and CCVCC syllables) and word length (a categorical predictor with three levels coded as monosyllabic, disyllabic, multisyllabic). The models included two-way interactions (e.g., age and stress, stress and position) and three-way interactions (e.g., age, stress, and position). The participant's name and the word item were incorporated as random effects, allowing the predictors and their interaction to vary by participants and words (Franke & Roettger 2019).

---

[18] In the analysis of syllable position using R software, 'final' was used for ultimate syllables, and 'non-final' was used for penultimate, antepenultimate, preantepenultimate, and propreantepenultimate syllables. This step was deemed necessary to enhance the model's performance in the Bayesian analysis. Monosyllabic words were annotated as 'ultimate' while other word lengths were annotated according to the number of syllables present.

The intercept is the expected value of the dependent variable when all independent variables are at their reference levels (Winter 2013). The reference level is chosen for each categorical variable in the model, against which all other levels are compared. The reference level associated with the intercept and variables of the models was: ‘light’ for syllable structure, ‘stressed’ for stress assignment, ‘non-final’ for syllable position, ‘'CVCC’ for sub-structure, and ‘monosyllabic’ for word length.

Using Markov Chain Monte Carlo (MCMC) to obtain samples from the posterior distribution, four sampling chains ran for 30,000 iterations, with a warmup period of 5000 iterations, yielding 100,000 total post-warmup draws. The expected values were reported under the posterior distribution and their 95% credible intervals (CIs) (Franke & Roettger 2019). For differences between levels within each predictor, the posterior probability of the variable was represented as a strong predictor when the difference between the lower limit of the 95% CI (l-95% CI) and the upper limit of the 95% CI (u-95% CI) did not include zero. The estimates of the fixed effects coefficients were reported.

4.6.2. Monotonic Predictors

The current study assigned the age group as a monotonic predictor for the models. Monotonic predictors are ordinal categorical variables that have a consistent directional relationship with the response variable. As the value of the predictor variable increases, the value of the response variable either decreases or increases accordingly (Bürkner & Charpentier 2018, Bürkner & Charpentier 2020). The population-level estimate of a monotonic effect shows the response difference between the minimum and maximum category, where categories are not assumed to be equally impactful with respect to their effect on the response variable (Leitenstorfer & Tutz 2007). If there are theoretical reasons to assume a monotonic relationship between variables, introducing monotonicity assumptions into the model's design can enhance the model's interpretation. That is, the monotonicity can increase the accuracy of parameter estimates and predictors, ensuring improved model calibration and predictive performance, regularizing the model's predictions to mitigate overfitting risks, and assessing the validity of the assumed monotonic relationship (Leitenstorfer & Tutz 2007, Bürkner 2021).

4.6.3. Priors

A prior is the probability distribution representing the degree of uncertainty or knowledge about a set of parameters before observing the data (Gelman & Shalizi 2013, Piironen & Vehtari 2017, Bürkner & Charpentier 2020). Priors can be either informative, wherein subject-specific knowledge is integrated into the model, or non-informative when the preference is to minimize the posterior inference impact (Gelman & Shalizi 2013, Stan Development Team 2013, Stan Reference Manual 2016: Version 2.32). Specifying informative priors can bias the results as prior assumptions can influence the output more than the data itself, while weakly informative or non-informative priors are less likely to bias the results (Bürkner & Charpentier 2020). The present study is exploratory, and the main goal was to avoid biasing the results. Thus, default priors were used for the models in the current study for three reasons: the data had a large variance as participants were recruited from eight different age groups; each participant produced a different set of words/syllables; and three tasks were used to collect data.

For predictors, the default priors were non-informative flat priors. Student_t (df = 3, mean = 453, sd = 220.5) was used as a default prior including the degrees of freedom, intercept, and standard deviation, respectively. Student_t refers to a probability distribution representing the t-distribution with specified degrees of freedom (Stan Development Team, Stan Reference Manual 2016, Version 2.32). The prior distribution for the standard deviation (sigma) is an inverse gamma distribution with two parameters that are equal to .01. These parameters are shape (determining the shape of the distribution) and scale (determining the spread of the distribution), both equal to .01.

4.6.4. Efficiency and Convergence

Three main diagnostics were considered to assess the sampling efficiency and convergence of the posterior distribution including Bulk Effective Sample Size (Bulk-ESS), Tail Effective Sample Size (Tail-ESS), and Potential Scale Reduction Factor (R-hat) (Winter 2013). Any fitted models that encountered fatal warning messages were not reported, as they do not suffice as reliable statistical measures. First, Bulk-ESS is a sampling efficiency diagnostic tool within the bulk of the posterior. Using split chains, it quantifies the effective sample size for rank-normalized values (Stan Reference Manual 2016, Vehtari et al., 2021). Second, Tail-ESS is utilized to assess the sampling efficiency and quantile estimates in the tail of the posterior. It represents the minimum effective size for the 5% and 95% percentile (Stan Reference Manual 2016, Vehtari et al., 2021). Third, R-hat is a diagnostic tool to assess the convergence of the MCMC algorithm by comparing estimates of model parameters and other univariate quantities of interest between and within chains. A value close to 1 indicates convergence of the MCMC chains, whereas if the chains have not mixed well, and the between and within chain estimates do not agree, the R-hat is larger than 1 (Vehtari et al., 2021).

3.5. Limitations

One of the key challenges encountered during data collection and analysis was the variability in timing, responses, and accuracy rates among participants, particularly in the child group, despite carefully planning each task’s time frame. Another challenge was the nature of child speech which has non-linguistic inconsistencies due to their mood-driven behaviour. Factors such as boredom, impatience, refusal to answer or cooperate, demands for treats, shyness, irregular speech rate such as screaming, and selective responding had to be considered.

Additionally, due to COVID-19, the online recording setup served as a communication barrier between the experimenter and the child participants. An obstacle encountered is the quality of recordings, despite providing detailed instructions to parents/caregivers regarding optimizing the recording setup. Although some productions were discernable audibly, they lacked the clarity necessary for precise annotations and, therefore, were discarded. Finally, the order of tasks has affected the general response acceptability in child groups. Participants performed relatively well in the initial two tasks, especially in the visually stimulating PT. However, RT, which was the last task, had notable behavioral differences. Participants demonstrated signs of fatigue, refusal to repeat words, impatience, crying, making noises in disapproval, and the desire to stop the experiment. Varying the order of the tasks across the participants could have mitigated the order effect. However, the order of the tasks was not randomized as one of the aims was to focus on less controlled tasks that potentially exhibited more accurate insights into the natural developmental patterns of JA child speech and also to maintain consistency across participants.

4. Chapter Four: Results

Chapter 4. Results

This chapter presents descriptive statistics and Bayesian model outputs for syllable and vowel durations for the three tasks: ST, RT, and PT (Section 4.1, 4.2, and 4.34.3, respectively). Section 4.4 presents a detailed examination of superheavy syllable production in JA speech, and Section 4.5 offers descriptive statistics of the occurrence of phonological processes across the child groups. The model's posterior predictive distribution is also provided when necessary. The posterior predictive probability is represented using the median values and not mean values for two reasons. First, means are used when the data distribution is primarily symmetrical, while medians are used when the distribution is skewed (Ali et al., 2015). Second, means are heavily influenced by outliers, where smaller or larger values pull the mean away from the central tendency of data, whereas medians are more robust as they are not affected by outliers. Therefore, the more skewed the distribution, the greater the difference between the mean and the median, and the greater the emphasis should be placed on using the median instead of the mean (Pham-Gia & Hung 2001).

The following table summarizes the frequency distribution of syllable structures across the child and adult groups in the three tasks. CVC, CV, CVVC, and CVVC syllables were the most frequent with 39.4%, 17.2%, 16.9%, and 11.3%, respectively. However, CCVC, CCVV, and CCV syllables were the least frequent with .3%, .2%, and .1%, respectively. An overall increasing trend of the number of attested syllables is depicted across the child groups, with a slight decrease in age group 45–51 months. It can be noted that the variability and complexity of structures increases with age, as long vowels, coda clusters, and onset clusters emerge. Nonetheless, even the oldest age group, ranging from 66–72, did not produce all structures evident in the adult group.

Table 4.1: Frequency of syllable structures across the age groups

Syllable Structure	24–30	31–37	38–44	45–51	52–58	59–65	66–72	adult	Total	Frequency
CVC	126	136	159	137	170	213	217	113	1271	39.4%
CV	43	59	69	63	87	83	85	65	554	17.2%
CVVC	58	56	73	61	87	58	85	66	544	16.9%
CVV	43	45	37	33	48	67	36	54	363	11.3%
CVCC	24	22	28	36	44	33	48	46	281	8.7%
CCVVC	7	7	20	12	25	20	31	55	177	5.5%
CCVC	0	0	1	1	0	2	1	8	13	0.4%
CCVV	0	0	0	1	1	3	4	2	11	0.3%
CCV	0	0	0	0	0	0	0	5	5	0.2%
CCVCC	0	0	0	0	0	0	0	3	3	0.1%
Total	301	325	387	344	462	479	507	417	3222	100.0%

4.1.Semi-Spontaneous Speech Task

This section explores the effects of predictors, including age group, stress, syllable structure, syllable position, and the three-way interaction between age group, stress, and syllable position on syllable and vowel durations in JA child and adult speech for the semi-spontaneous speech task. The following table summarizes the attested syllable structures and their frequencies across the age groups in ST. CVC, CVV, CV, and CVVC syllables are the most frequent with 42.5%, 16.6%, 16.3%, and 14.2%, respectively. As for the least frequent structures, CCVV, CCVC, and CCV had the following percentages respectively .7%, .5%, and .3%. CCVCC syllables were not produced in this task.

Table 4.2: Frequency of syllable structures in ST

Syllable Structure	Frequency in ST	Percentage
CVC	558	42.5%
CVV	218	16.6%
CV	214	16.3%
CVVC	186	14.2%
CVCC	102	7.8%
CCVVC	16	1.2%
CCVV	9	0.7%

CCVC	6	0.5%
CCV	4	0.3%
CCVCC	0	0.0%
Total	1313	100%

4.1.1. Syllable Duration

Table 4.3 summarizes the mean (in ms) and standard deviation for syllable durations for the independent variables, syllable structure (light, heavy, superheavy); lexical stress (stressed, unstressed); and syllable position within a word (final, non-final) for each age group.

Table 4.3: Summary of the mean and standard deviation for syllable durations for syllable structure lexical stress and syllable position for each age group in ST (n=1440).

Age Group	Syllable Structure		Stress		Syllable Position	
24–30	Light	208.2 (90.4)	Unstressed	236.2 (89.7)	Non-Final	237.9 (101.7)
	Heavy	262.3 (125.7)	Stressed	375.1 (195.9)	Final	374.0 (192.8)
	Superheavy	473.5 (179.2)				
31–37	Light	168.0 (81.5)	Unstressed	291.0 (153.0)	Non-Final	244.3 (122.2)
	Heavy	328.5 (147.2)	Stressed	349.5 (170.0)	Final	388.9 (167.1)
	Superheavy	444.6 (157.5)				
38–44	Light	200.6 (93.6)	Unstressed	312.7 (138.0)	Non-Final	263.0 (107.2)
	Heavy	329.3 (113.2)	Stressed	421.7 (202.0)	Final	472.0 (179.6)
	Superheavy	581.2 (171.9)				
45–51	Light	154.5 (64.3)	Unstressed	231.9 (86.3)	Non-Final	210.3 (74. 6)
	Heavy	250.7 (70.8)	Stressed	303.9 (156.0)	Final	322.4 (149.9)
	Superheavy	405.9 (180.8)				
52–58	Light	138.2 (75.2)	Unstressed	223.5 (130.3)	Non-Final	187.1 (80.7)
	Heavy	227.7 (105.6)	Stressed	285.8 (153.8)	Final	317.2 (162.9)
	Superheavy	370.2 (163.6)				

59–65	Light	135.7 (52.2)	Unstressed	236.2 (114.3)	Non-Final	217.2 (87.1)
	Heavy	263.6 (92.5)	Stressed	283.3 (106.2)	Final	303.4 (119.9)
	Superheavy	355.6 (121.3)				
66–72	Light	138.5 (50.3)	Unstressed	261.2 (126.3)	Non-Final	221.71 (93.64)
	Heavy	273.1 (89.6)	Stressed	315.8 (130.2)	Final	349.0 (129.7)
	Superheavy	413.6 (134.9)				
Adult	Light	120.8 (74.9)	Unstressed	175.8 (89.3)	Non-Final	171.1 (67.7)
	Heavy	192.3 (111.2)	Stressed	226.6 (94.7)	Final	231.9 (110.8)
	Superheavy	286.9 (171.8)				

For ST, 1440 tokens were annotated and analysed for the syllable duration (Table 4.3). The Bayesian analysis was performed using four chains running for 30,000 iterations, and the warmup period consisted of 5,000 iterations to mitigate the potential effect of the initial phase of the sampling process. The total post-warmup draws available for analysis was 100,000. Table 4.4 shows the Bayesian model output for syllable duration.

Table 4.4: Bayesian model output summary for syllable duration/ST

Data: STC data (Number of observations: 1440)
Draws: 4 chains, each with iter = 30000; warmup = 5000; thin = 1; total post-warmup draws = 1e+05

Population-level effect	Estimate	Est. Error	l-95% CI	u-95% CI	Bulk_ESS	Tail_ESS
Intercept	183.15	21.07	141.42	224.60	38476	55496
Stress (Unstressed)	2.20	18.01	-32.74	38.62	47697	59685
Position (Final)	116.33	17.67	82.08	151.55	55434	72316
Structure (Heavy)	92.25	9.16	74.32	110.16	112628	84461
Structure (Superheavy)	223.16	14.39	195.09	251.40	96744	83104
Stress (Unstressed): Position (Final)	-88.93	35.29	-157.68	-19.18	39867	49717
Age group	-6.63	4.86	-16.35	2.81	30140	47737
Age group: Stress (Unstressed)	-3.03	3.51	-10.02	3.83	47110	57094
Age group: Position (Final)	-18.90	3.15	-25.18	-12.83	54793	72118

Age group: Stress (Unstressed): Position (Final)	20.52	5.70	9.13	31.56	38869	49412

Age Group

Syllable durations demonstrate a general decreasing trend across the age groups. Figure 4.1 shows that the mean syllable durations exhibit a slight increase in the youngest three age groups (24–30 months, Mean = 320.2 ms, SD = 175.7; 31–37 months, Mean = 322.7 ms, SD = 164.4; 38–44 months, Mean = 366.8 ms, SD = 180.7). At 45–51 months, the durations demonstrate a decrease in the mean durations, by which the four older age groups exhibit comparable values (45–51 months, Mean = 270.7 ms, SD = 132.9; 52–58 months, Mean = 257 ms, SD =146.3; 59–65 months, Mean = 259.1 ms, SD =112.7; 66–72 months, Mean = 290.8 ms, SD =130.7). Although the durations decreased after 45–51 months of age, children's data showed longer durations than the adult group (Mean = 199.4 ms, SD = 95.1). However, the model output shows that age group is not a strong predictor for syllable duration (β = -6.63, CI [-16.35, 2.81]).

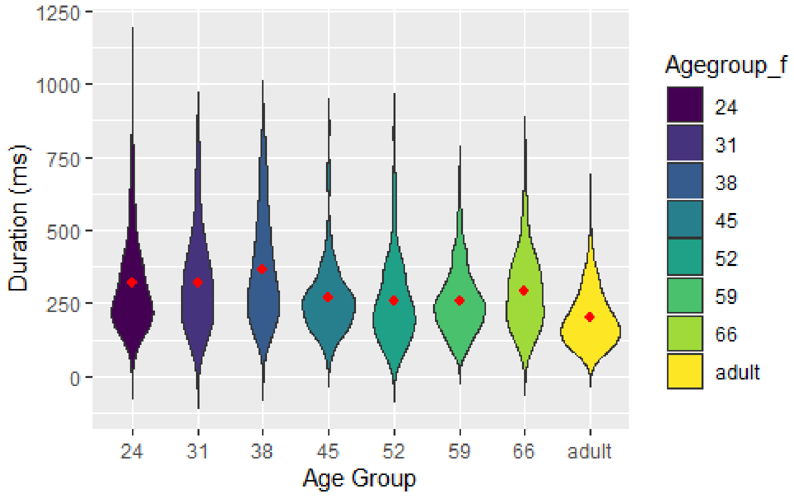
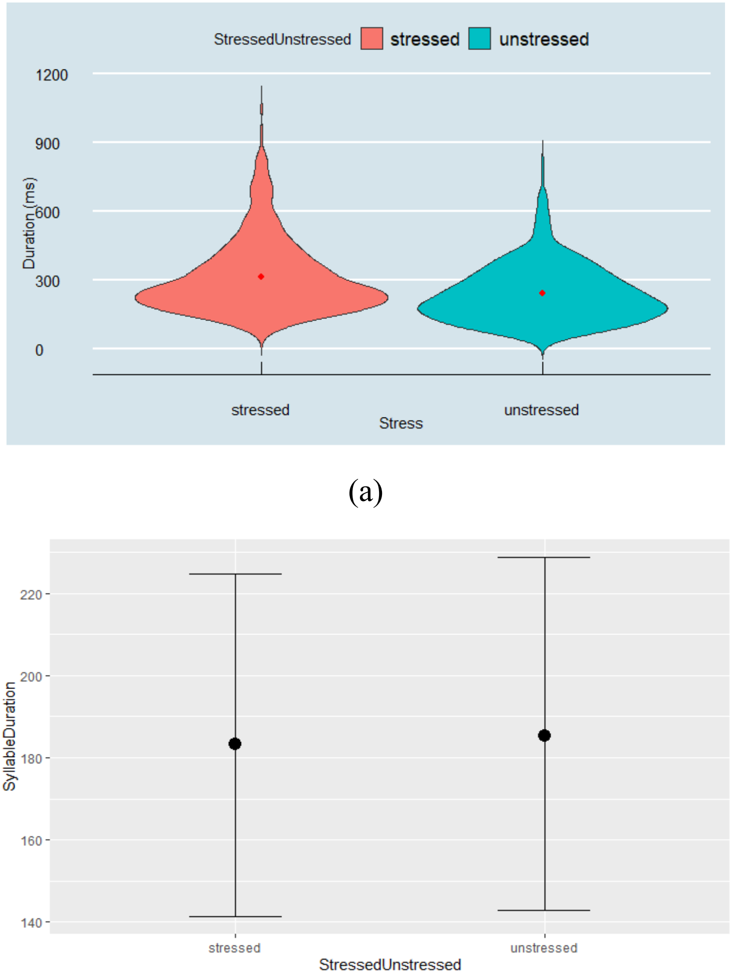


Figure 4.1: The distribution and mean of the syllable duration (ms) by age group

Stress

Figure 4.2/a shows that stressed syllables (Mean = 311.6 ms, SD = 159.2) were longer than unstressed syllables (Mean = 239.0 ms, SD = 124.0). Stressed syllables are 1.3 times longer than their unstressed counterparts on average. However, the model output indicates that stress assignment is not a strong predictor for syllable duration (β = 2.68, CI [-32.74, 38.62]). Figure 4.2/b demonstrates that the median duration of unstressed syllables appears slightly higher than that of stressed syllables.



(a)

(b)

Figure 4.2: (a) The distribution and the mean syllable duration (ms) for stress (b) Posterior predictive plot for syllable duration for stress

Syllable Structure

Figure 4.3 shows that light syllables are the shortest (Mean = 153.7 ms, SD = 74.9), followed by heavy syllables (Mean = 259.9 ms, SD = 111.2), while superheavy syllables are the longest (Mean = 407.2 ms, SD = 171.8). On average, superheavy syllables are 2.6 times longer than short syllables and 1.6 times longer than heavy syllables. Similarly, the model output shows that syllable structure is a strong predictor for syllable duration (heavy, β = 92.25, CI [74.32, 110.16]; superheavy, β = 223.16, CI [195.09, 251.40]).

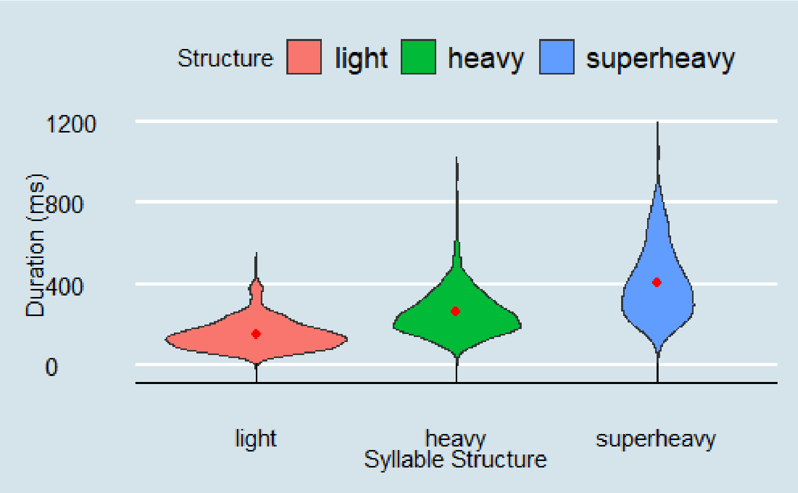


Figure 4.3:The distribution and the mean syllable duration (ms) for syllable structure

Syllable Position

Figure 4.4 shows that syllables produced in the word final position (Mean = 335.2 ms, SD = 94.9) are longer than syllables in word non-final positions (Mean = 214.1 ms, SD = 163.7). On average, word final syllables are approximately 1.6 times longer than non-final syllables. The model output shows that syllable position is a strong predictor for syllable duration (β = 116.33, CI [82.08, 151.55]).

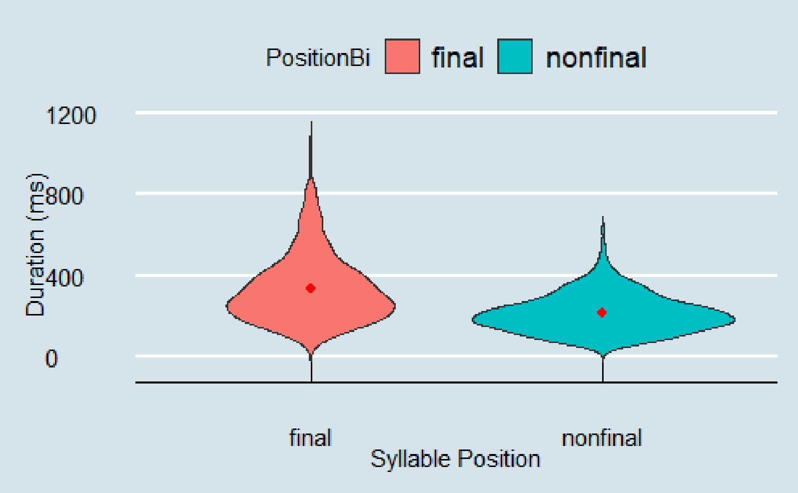


Figure 4.4: The distribution and the mean syllable duration (ms) for syllable position

Age-Stress-Position Interaction

Across the age groups, stressed final syllables (Mean = 378.2 ms, SD = 178.6) are longer than their stressed non-final counterparts (Mean = 240.2 ms, SD = 92.2). Also, unstressed final syllables (Mean = 288.9 ms, SD = 131.4) are longer than unstressed final syllables (Mean = 186.7 ms, SD = 89.9). On average, stressed final syllables are 1.6 times longer than stressed non-final syllables, and unstressed final syllables are 1.5 times longer than unstressed non-final syllables. The model output in Table *4.4* shows that the three-way interaction between age group, stress, and position is a strong predictor for syllable duration (β = 20.52, CI [9.13, 31.56]). Figure 4.5 shows that for the non-final positions, stressed syllables tend to be longer than unstressed syllables for all age groups. The figure indicates that the difference between stressed and unstressed non-final syllables becomes more evident with age. The large overlap of CIs in stressed and unstressed shows the marginal difference between the estimated probabilities. For the word-final position, Figure 4.5 also shows that stressed syllables are longer than unstressed syllables only in the first three age groups, aged 24–30 to 38–44 months. At age 45–51 months, for the word-final position, durations of stressed and unstressed syllables were estimated to be similar. A change follows this overlap in the trend where stressed syllables become shorter than unstressed syllables in the word-final position. In the youngest three age groups, stressed and unstressed syllables overlap less, indicating a larger difference between the probability estimates. However, by 45–51 months, the overlap between attested CIs becomes more evident, indicating a smaller durational difference between stressed and unstressed syllables.

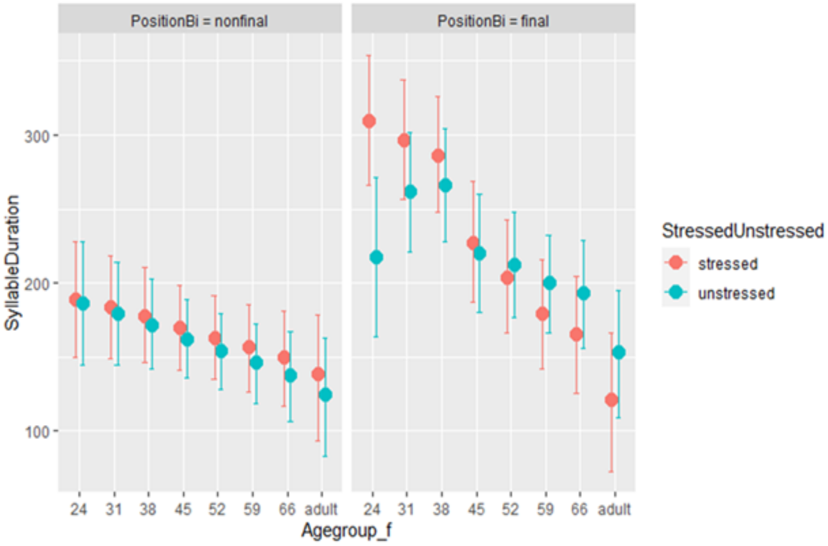


Figure 4.5: Posterior predictive plot for syllable duration for the interaction between age group, stress assignment, and syllable position

4.1.2. Vowel Duration

Table 4.5 summarizes the mean and standard deviation for vowel durations for the independent variables, syllable structure (light, heavy, superheavy); lexical stress (stressed, unstressed); and syllable position within a word (final, non-final) for each age group.

Table 4.5: Summary of the mean and standard deviation for vowel durations for syllable structure lexical stress and syllable position for each age group in ST (n=1440).

Age Group	**Syllable Structure**		**Stress**		**Syllable position**	
24–30	Light	110.8 (44.2)	Unstressed	97.7 (45.4)	Non-Final	113.92 (54.3)
	Heavy	109.6 (89.0)	Stressed	149.8 (101.2)	Final	139.21 (102.5)
	Superheavy	170.3 (90.0)				
31–37	Light	107.3 (36.5)	Unstressed	130.4 (65.7)	Non-Final	126.72 (57.1)
	Heavy	150.9 (92.6)	Stressed	172.2 (104.7)	Final	175.24 (107.5)
	Superheavy	199.8 (100.5)				
38–44	Light	111.7 (54.8)	Unstressed	132.4 (63.1)	Non-Final	119.35 (57.9)
	Heavy	138.8 (64.1)	Stressed	190.3 (107.6)	Final	203.46 (101.7)
	Superheavy	252.3 (111.8)				
45–51	Light	88.8 (25.7)	Unstressed	95.4 (36.9)	Non-Final	101.09 (38.6)
	Heavy	110.7 (52.0)	Stressed	154.8 (90.4)	Final	149.9 (92.8)
	Superheavy	198.6 (106.9)				
52–58	Light	76.5 (37.7)	Unstressed	107.5 (58.5)	Non-Final	86.73 (45.5)
	Heavy	107.1 (54.2)	Stressed	127.7 (86.4)	Final	145.60 (84.8)
	Superheavy	159.9 (100.3)				
59–65	Light	83.3 (29.0)	Unstressed	98.7 (42.0)	Non-Final	93.43 (38.6)
	Heavy	107.3 (45.3)	Stressed	120.6 (58.2)	Final	126.2 (58.0)
	Superheavy	141.2 (71.6)				

66–72	Light	84.4 (28.8)	Unstressed	114.5 (57.6)	Non-Final	89.8 (38.6)
	Heavy	107.1 (46.8)	Stressed	138.8 (85.9)	Final	159.6 (83.2)
	Superheavy	198.7 (97.0)				
Adult	Light	54.9 (24.8)	Unstressed	79.9 (60.8)	Non-Final	66.7 (35.4)
	Heavy	83.8 (59.5)	Stressed	89.1 (56.2)	Final	104.4 (72.4)
	Superheavy	109.9 (65.4)				

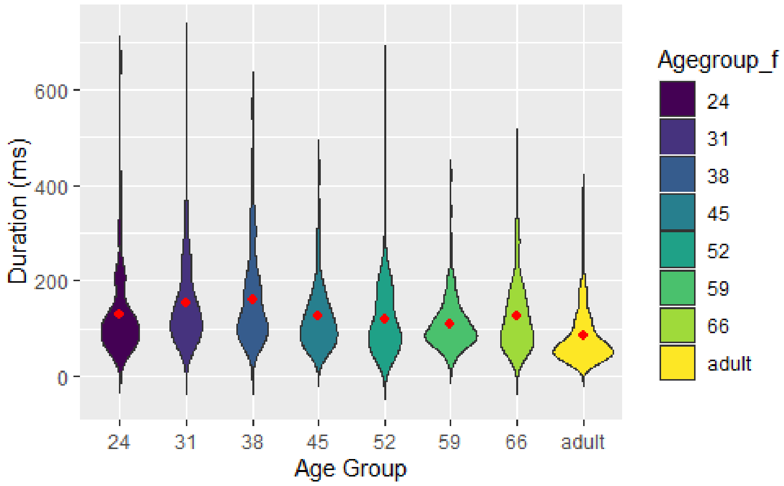
A total of 1440 vowels were annotated and analysed for ST. The Bayesian analysis was performed using four chains running for 30,000 iterations. The warmup period consisted of 5000 iterations to mitigate the potential effect of the initial phase of the sampling process. The total post-warmup draws available for analysis was 100,000. Table 4.6 summarizes the Bayesian model output for vowel duration.

Table 4.6:Bayesian model output summary for vowel duration/ST

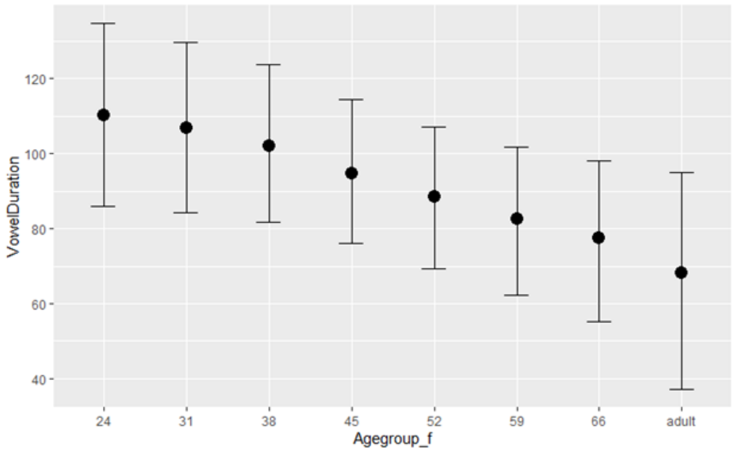
Data: STC data (Number of observations: 1440)						
Draws: 4 chains, each with iter = 30000; warmup = 5000; thin = 1; total post-warmup draws = 1e+05						
Population-level effect	Estimate	Est. Error	l-95% CI	u-95% CI	Bulk_ESS	Tail_ESS
Intercept	110.16	12.42	85.85	134.74	41727	58301
Stress (Unstressed)	-12.49	11.51	-35.76	9.80	43749	57611
Position (Final)	72.25	11.32	49.49	93.73	44561	46864
Structure (Heavy)	15.80	5.62	4.74	26.88	125303	84505
Structure (Superheavy)	40.08	8.78	22.92	57.39	109790	84426
Stress (Unstressed): Position (Final)	-67.12	19.55	-106.32	-29.37	36382	47913
Age group	-6.08	2.87	-11.84	-0.49	33012	51560
Age group: Stress (Unstressed)	2.10	2.28	-2.40	6.65	40776	54816
Age group: Position (Final)	-4.86	2.23	-9.00	-0.19	39595	41958
Age group: Stress (Unstressed): Position (Final)	7.80	3.60	0.60	14.81	33960	42257

Age Group

Figure 4.6/a demonstrates that, broadly, the mean vowel durations decrease across the age groups. In the first three age groups, vowel durations increase as the age increases (24–30 months, Mean = 129.2 ms, SD = 87.3) to reach a maximum of Mean = 161.1 ms (SD = 91.1) by 38–44 months. At 45–51 months (Mean = 127.4 ms, SD = 76.6), the durations decrease steadily, reaching the mean (Mean = 84.2 ms, SD = 58.8) in the adult group. The model output shows that age group is a strong predictor for vowel duration (β = -6.08, CI [-11.84, -0.49]). Vowel durations decrease with age, becoming closer to the median duration of the adult group Figure 4.6/b.



(a)



(b)

Figure 4.6 (a) The distribution and the mean vowel duration (ms) for age group (b) Posterior predictive plot for vowel duration for age group

Stress

Figure 4.7 shows that vowels in stressed syllables (Mean = 137.1 ms, SD = 89.1) are 1.3 times longer than in unstressed syllables (Mean = 104.2 ms, SD = 57.2). Nevertheless. The model output shows that stress assignment is not a strong predictor for vowel duration (β = -12.49, CI [-35.76, 9.80]).

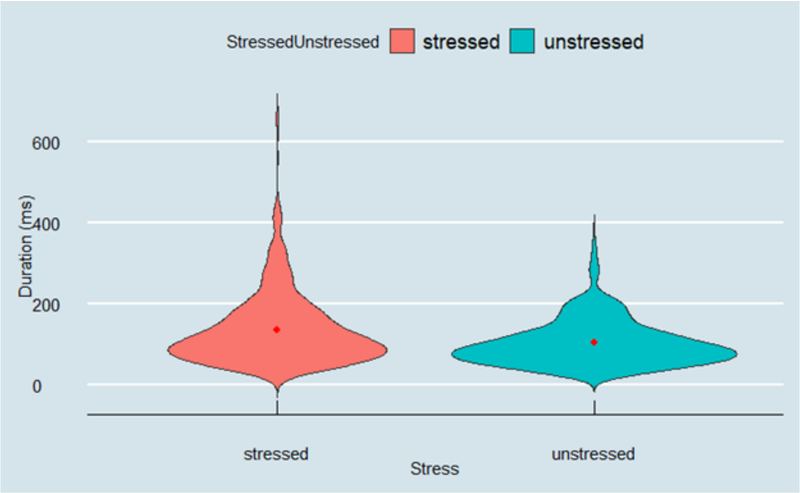


Figure 4.7:The distribution and the mean vowel duration (ms) for stress assignment

Syllable Structure

Figure 4.8 shows that vowels in superheavy syllables are the longest (Mean = 171.7 ms, SD = 99.9), followed by vowels in heavy syllables (Mean = 111.2 ms, SD = 64.6), while vowels in light syllables are the shortest (Mean = 86.7 ms, SD = 40.5). Approximately, vowels in superheavy syllables are two times longer than those in light syllables, while vowels in heavy syllables are 1.3 times longer than in light syllables. The model output shows that syllable structure is a strong predictor for vowel duration (heavy, β = 15.80, CI [4.74, 26.88]); superheavy, β = 40.08, CI [22.92, 57.39]).

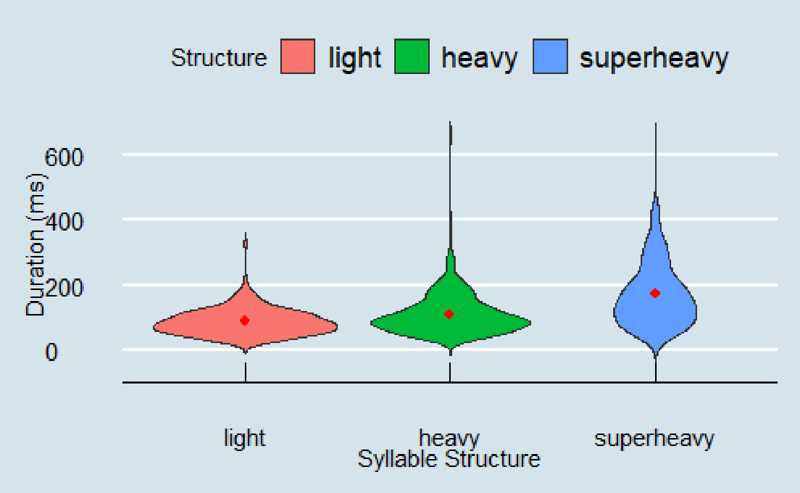


Figure 4.8:The distribution and the mean vowel duration (ms) for syllable structure

Syllable Position

Word final lengthening is observed in vowel duration as demonstrated in Figure 4.9. Vowels appearing in word final syllables (Mean = 145.5 ms, SD = 90.0) are longer by 1.5 times on average than vowels in word non-final syllables (Mean = 95.3 ms, SD = 48.6). The model output shows that syllable position is a strong predictor for vowel duration (β = 72.25, CI [49.49, 93.73]).

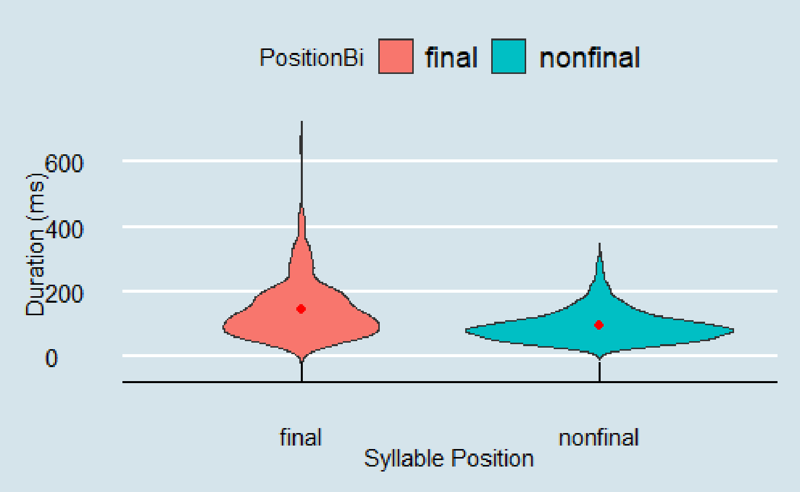


Figure 4.9:The distribution and the mean vowel duration (ms) for syllable position

Age-Position-Stress Interaction

Across the age groups, vowels in stressed final syllables (Mean = 172.2 ms, SD = 102.4) are longer than vowels in stressed non-final syllables (Mean = 99.5 ms, SD = 49.9). Similarly, vowels in unstressed final syllables (Mean = 116.8 ms, SD = 63.1) are longer than vowels in unstressed non-final syllables (Mean = 90.9 ms, SD = 46.9). On average, vowels in stressed final syllables are 1.7 times longer than vowels in stressed non-final syllables, and vowels in unstressed final syllables are 1.3 times longer than vowels in unstressed non-final syllables. Correspondingly, the model output shows that the three-way interaction between age group, stress, and position is a strong predictor for vowel duration ($\beta = 7.80$, CI [.60, 14.81]). Figure 4.10 suggests that in the younger age groups, vowels in stressed word non-final syllables are longer than vowels in unstressed non-final syllables. However, this trend becomes less evident with age, where the durational difference between vowels in stressed and unstressed non-final syllables becomes smaller as age increases, reaching an overlap in the adult group. As for vowels in the word-final positions, Figure 4.10 also indicates that vowels appearing in stressed final positions are longer than vowels appearing in unstressed final syllables across age groups. The durational difference between vowels in stressed and unstressed final syllables decreases with age. For the final position, from 24–30 to 45–51 months, vowels in stressed and unstressed final syllables do not overlap, indicating a larger difference in the probability estimates. At the age of 52–58 months, the overlap between vowels in stressed and unstressed final syllables becomes more evident with age, suggesting a smaller difference in the probability estimates.

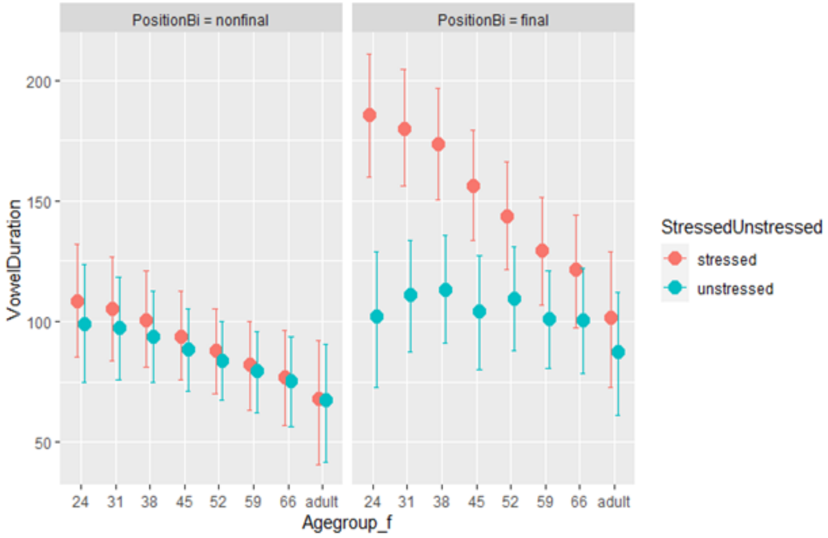


Figure 4.10: Posterior predictive plot for vowel duration for the interaction between age group, stress assignment, and syllable position

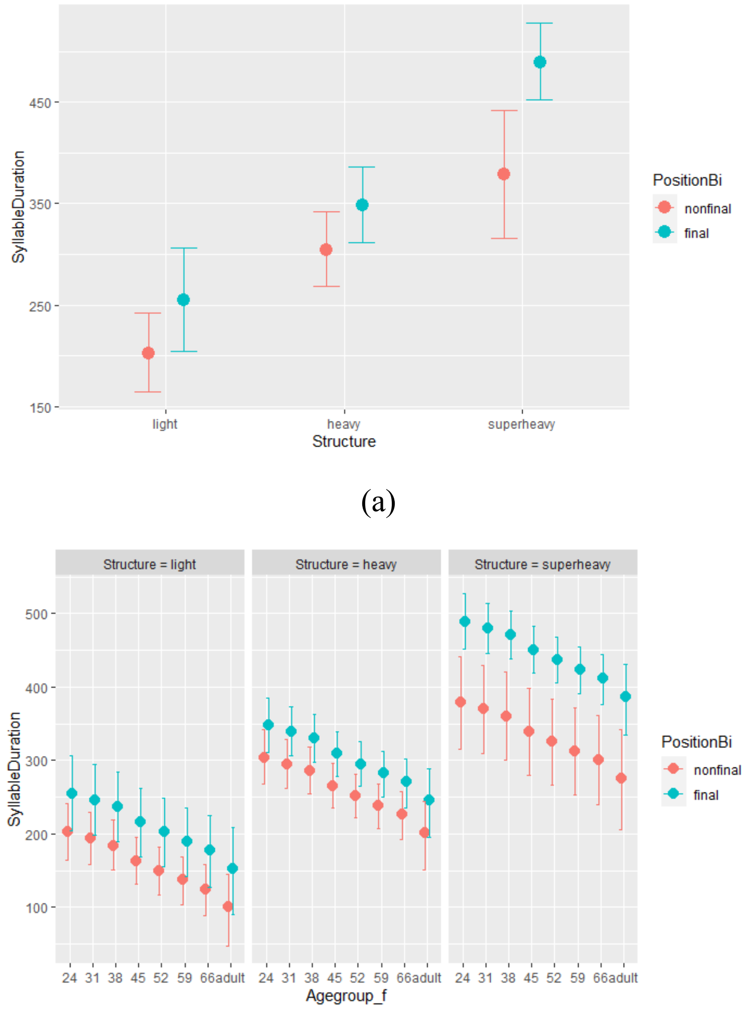
4.1.3. Interaction between Syllable Structure and Syllable Position

Another Bayesian model was constructed to examine the interaction between syllable structure and position. The aim was to examine whether JA productions show evidence of abiding by the trimoraic ban and the bimoraic constraint (Section 2.1.4). As syllable position affects syllable structure by contributing to its weight based on WPR, light and heavy syllables do not pose a problem for the analysis, as they are monomoraic and bimoraic, respectively. However, recalling that syllables are maximally bimoraic in JA, superheavy syllables are problematic to this account as a third mora would presumably be assigned to the non-final superheavy syllable (i.e., the third mora would be assigned to the word non-final coda by WPR). If superheavy syllables in JA are bimoraic, then the durations of word non-final superheavy and heavy syllables would be comparable, as both heavy and superheavy syllables are bimoraic. On the other hand, if superheavy syllables are trimoraic, then it is expected that the duration of a non-final superheavy syllable would be longer than a heavy one, as an extra mora would be present. Thus, Bayesian models were fitted to determine the effect of age group, syllable position, syllable structure, and the interaction between syllable position and structure on syllable and vowel durations. For each model, four sampling chains ran for 10,000 iterations, with a warmup period of 2000 iterations, yielding 32,000 total post-warmup draws. Table 4.7 demonstrates the model output for the effect of syllable position, syllable structure, and age group on syllable duration.

Table 4.7: Bayesian model output summary for syllable duration in ST interaction

Population-Level Effects	**Estimate**	**Est. Error**	**l-95% CI**	**u-95% CI**	**Bulk_ESS**	**Tail_ESS**
Intercept	202.75	19.54	164.89	242.07	15950	21044
Position (Final)	52.21	20.68	12.03	92.87	28790	24158
Structure (Heavy)	101.61	9.41	83.09	119.98	40549	27612
Structure (Superheavy)	175.91	27.68	122.12	229.72	42522	25970
Position (Final): Structure (Heavy)	-8.28	22.20	-51.84	35.01	28553	23147
Position (Final): Structure (Superheavy)	58.50	34.36	-8.54	125.83	30210	25514
Age group	-14.82	4.65	-24.13	-5.73	13694	18442

The model output indicates that syllable position (β = 52.21, CI [12.03, 92.87]), syllable structure (heavy, β = 101.61, CI [83.09, 119.98]) (superheavy, β = 175.91, CI [122.2, 229.72]), and age group (β =-14.82, CI [-24.13, 5.73]) are strong predictors for syllable duration. However, the model output indicates that the interaction between syllable structure and syllable position (heavy, β =101.61, CI [-51.84, 35.01]) (superheavy, β =58.50, CI [-8.54, 125.83]) is not a strong predictor for syllable duration. A closer examination of the posterior predictive probability plots in Figure 4.11 indicates that syllable durations in light and heavy syllables in final and non-final positions overlap, indicating a weaker difference between the probability estimates. Nevertheless, superheavy syllable durations in word final and non-final positions do not show an overlap. This indicates that superheavy syllables in word non-final positions are shorter than their counterparts.



(a)

(b)

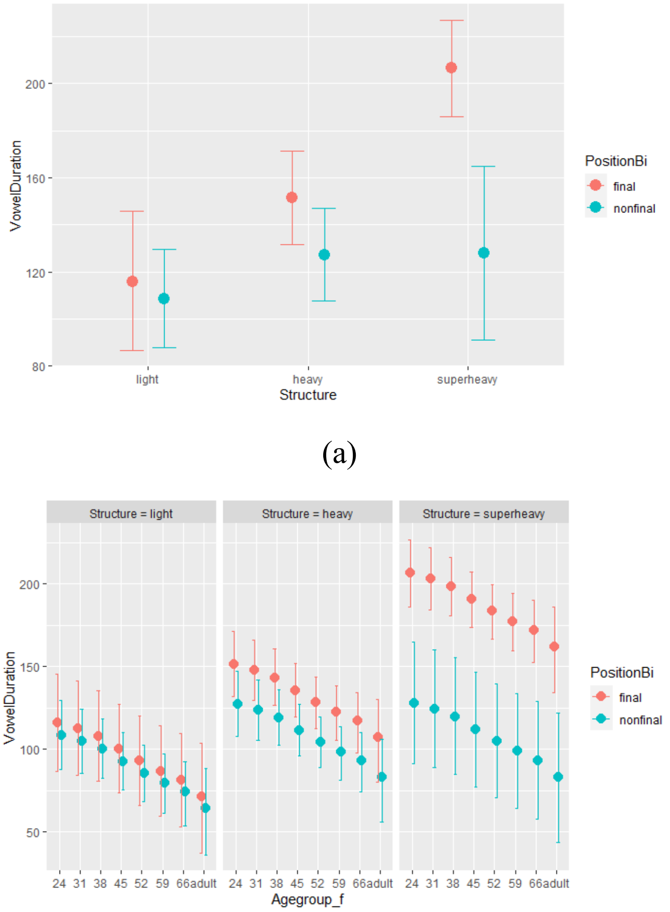
Figure 4.11: Posterior predictive plots for syllable duration for syllable structure and syllable position and the three-way interaction

As for vowel durations, Table 4.8 demonstrates the model output for the effect of syllable position, syllable structure, age group, and the interaction between syllable position and syllable structure on vowel durations. A closer examination of the median vowel durations across the structures specifies that non-final CVVC syllables (Median = 79.4 ms, SD = 54.0) are shorter than short vowels appearing in CVCC syllables in the same position (Median = 85.7 ms, SD = 16.1). CVVC syllables have vowel durations comparable to CVC syllables in word non-final positions (Median = 79.9 ms, SD = 33.6). On the other hand, CVV syllables have longer vocalic durations (Median = 156.5 ms, SD = 60.6).

Table 4.8: Bayesian model output summary for vowel duration in ST interaction

Population-Level Effects	**Estimate**	**Est. Error**	**l-95% CI**	**u-95% CI**	**Bulk_ESS**	**Tail_ESS**
Intercept	115.82	15.04	86.75	145.69	10560	15243
Position (Final)	-7.41	12.77	-32.41	17.66	11674	16931
Structure (Heavy)	35.62	12.27	11.60	59.46	12012	17233
Structure (Superheavy)	90.70	12.67	65.82	115.54	11815	17778
Position (Final): Structure (Heavy)	-16.81	13.69	-43.50	10.08	11670	17275
Position (Final): Structure (Superheavy)	-71.16	21.07	-112.49	-30.06	17779	23105
Age group	-6.41	2.47	-11.33	-1.47	8730	13900

The model output indicates that syllable position (β = -7.41, CI [12.03, 92.87]), syllable structure (heavy, β = 101.61, CI [83.09, 119.98]) (superheavy, β = 175.91, CI [122.2, 229.72]), and age group (β =-14.82, CI [-24.13, 5.73]) are strong predictors for syllable duration. However, the model output indicates that the interaction between syllable structure and syllable position (heavy, β =101.61, CI [-51.84, 35.01]) (superheavy, β =58.50, CI [-8.54, 125.83]) is not a strong predictor for vowel duration. Results suggest that the interaction between syllable position and syllable structure is only a strong predictor for superheavy syllables but not for light and heavy syllables. The probability plot in Figure 4.12/b demonstrates that vowel durations in word final and non-final positions for light and heavy syllables overlap, with less overlap for heavy syllables, suggesting a weaker difference in the probability estimates. Vowels in light and heavy syllables are longer in the word final position. Nevertheless, superheavy syllables do not show a similar pattern. Word non-final superheavy syllables have shorter vowels than word final superheavy syllables. The median vowel durations of word final and non-final superheavy syllables do not overlap, indicating a larger difference in the probability estimates. This indicates that vowel shortening occurs, suggesting that superheavy syllables are rather bimoraic and not trimoraic. Figure 4.12/a displays the median range of vowels in non-final superheavy syllables closely resembling the vowel durations in heavy syllables.



(a)

(b)

Figure 4.12: Posterior predictive plots for vowel duration for syllable structure and syllable position and the three-way interaction

4.1.4. Summary

Core syllables (CVC, CVV, CV, and CVVC) are the most frequent compared to marginal ones (CCVC, CCV, and CCVCC). The variability and complexity of syllable structures increased with age as clusters appeared at later stages of development. For syllable duration, children in the youngest three age groups, ranging from 24–30 to 38–44 months, produce longer syllables than older age groups, ranging from 45–51 months to adults, and even the oldest child group did not produce adult-like durations. Additionally, the mean durations show that stressed syllables are 1.3 times longer than unstressed syllables. The results suggest that syllable structure affects syllable durations where superheavy syllables are 2.6 times longer than short syllables and 1.6 times longer than heavy ones. Word final lengthening is observed across the age groups, with word final syllables being 1.6 times longer than their counterparts. The Bayesian model output indicates that only syllable structure, syllable position, and the three-way interaction are strong predictors for syllable duration. Stressed syllables are longer than unstressed syllables in word non-final positions only in the youngest three age groups while the difference between stressed and unstressed syllables becomes more evident with age in the word final position.

Moreover, the data distribution indicates that vowel durations have a general decreasing pattern across the age groups. Vowels in stressed syllables are 1.3 times longer than vowels in unstressed syllables. Vowels in superheavy syllables are two times longer than light syllables and 1.3 times longer than heavy syllables. Word final lengthening is observed in vocalic durations, where vowels in the word-final position are 1.5 times longer than their counterparts. The Bayesian model output suggests that only stress assignment is not a strong predictor for vowel duration in JA productions. However, the interaction between stress and syllable position, and the three-way interaction are strong predictors for vowel durations. Vowels in stressed non-final syllables become longer than their counterparts with age, while vowels in stressed final syllables decrease in duration with age to become similar to their unstressed counterparts. The interaction between syllable structure and syllable position demonstrated that vowel shortening is evident in word non-final superheavy syllables, indicating that superheavy syllables are bimoraic and not trimoraic.

4.2.Repetition Task

This section explores the effects of predictors, including age group, stress, syllable structure, syllable position, and the three-way interaction between age group, stress, and syllable position on syllable and vowel durations in JA child and adult speech in RT. The following table summarizes the attested syllable structures and their frequencies across the age groups in RT. CVVC, CCVVC, CVCC, and CV syllables are the most frequent with 27.8%, 21.5%, 19.6%, and 16.4%, respectively. On the other hand, CCVCC, CCVV, and CCV were the least frequent with .5%, .2%, and .2%, respectively.

Table 4.9: Frequency of syllable structures in RT

Syllable Structure	Frequency in RT	Percentage
CVVC	174	27.8%
CCVVC	135	21.5%
CVCC	123	19.6%
CV	103	16.4%
CVC	57	9.1%
CVV	26	4.1%
CCVC	4	0.6%
CCVCC	3	0.5%
CCVV	1	0.2%
CCV	1	0.2%
Total	627	100%

4.2.1. Syllable Duration

Table 4.10 summarizes the mean and standard deviation for syllable durations for the independent variables, syllable structure (light, heavy, superheavy); lexical stress (stressed, unstressed); and syllable position within the word (final, non-final) for each age group.

Table 4.10: Summary of the mean and standard deviation for syllable durations for syllable structure lexical stress and syllable position for each age group in RT (n=613).

Age Group	Syllable Structure	Stress	Syllable position

24–30	Light	154.9 (NA)	Unstressed	244.6 (92.7)	Non-Final	289.4 (71.6)
	Heavy	289.4 (71.6)	Stressed	499.9 (210.0)	Final	483.5 (218.1)
	Superheavy	500.0 (210.0)				
31–37	Light	194.1 (NA)	Unstressed	272.3 (26.4)	Non-Final	223.9 (42.0)
	Heavy	272.3 (26.4)	Stressed	627.8 (174.9)	Final	633.9 (159.9)
	Superheavy	656.8 (135.7)				
38–44	Light	148.2 (66.3)	Unstressed	246.6 (144.8)	Non-Final	199.5 (86.1)
	Heavy	319.9 (122.3)	Stressed	532.7 (186.3)	Final	547.5 (170.9)
	Superheavy	579.0 (174.0)				
45–51	Light	161.4 (66.1)	Unstressed	211.9 (119.8)	Non-Final	189.7 (94.9)
	Heavy	255.2 (107.4)	Stressed	509.2 (162.9)	Final	514.5 (154.4)
	Superheavy	534.1 (142.4)				
52–58	Light	142.9 (63.0)	Unstressed	180.9 (80.3)	Non-Final	169.3 (69.1)
	Heavy	254.2 (87.8)	Stressed	532.6 (133.0)	Final	538.2 (119.8)
	Superheavy	555.7 (107.5)				
59–65	Light	147.7 (56.1)	Unstressed	210.3 (113.8)	Non-Final	181.3 (76.9)
	Heavy	319.6 (104.0)	Stressed	513.9 (129.1)	Final	522.1 (114.4)
	Superheavy	523.5 (115.6)				
66–72	Light	144.4 (62.8)	Unstressed	193.9 (130.5)	Non-Final	183.4 (85.3)
	Heavy	278.5 (80.2)	Stressed	499.9 (90.8)	Final	503.7 (122.3)
	Superheavy	512.6 (117.7)				
Adult	Light	155.3 (57.4)	Unstressed	199.9 (103.0)	Non-Final	202.7 (98.5)
	Heavy	251.0 (83.5)	Stressed	495.9 (135.3)	Final	492.8 (140.5)
	Superheavy	520.2 (112.7)				

To analyse the effect of the predictors, as mentioned earlier, on syllable duration in JA child and adult speech, a total of 613 tokens were analysed for RT. The Bayesian analysis was performed using four chains running for 30,000 iterations. The warmup period consisted of 5000 iterations to mitigate the potential effect of the initial phase of the sampling process. The total post-warmup draws available for analysis was 100,000. Table 4.11 reports the Bayesian model output for syllable duration.

Table 4.11: Bayesian model output summary for syllable durations/RT

Data: RTC data (Number of observations: 613) **Draws: 4 chains, each with iter = 30000; warmup = 5000; thin = 1; total post-warmup draws = 1e+05**						
Population-level effect	Estimate	Est. Error	l-95% CI	u-95% CI	Bulk_ESS	Tail_ESS
Intercept	127.59	62.68	-7.61	242.58	18672	31816
Stress (Unstressed)	45.26	64.33	-76.17	179.27	19253	29112
Position (Final)	143.28	59.14	29.96	262.75	22036	32233
Structure (Heavy)	102.00	18.13	66.53	137.49	71095	74247
Structure (Superheavy)	299.58	35.27	230.26	368.83	68220	70823
Stress (Unstressed): Position (Final)	-130.94	95.31	-328.42	47.76	21744	32704
Age group	8.80	11.24	-12.45	32.21	16272	25854
Age group: Stress (Unstressed)	-8.96	11.04	-31.20	12.89	17859	24942
Age group: Position (Final)	-18.37	8.85	-36.08	-0.80	17590	23353
Age group: Stress (Unstressed): Position (Final)	20.04	15.63	-10.66	51.60	21013	31315

Age Group

Figure 4.13 shows the distribution of syllable duration across the age groups. The mean syllable duration in the youngest age group is 466.6 ms (SD = 215.9), which increases to 588.3 ms (SD = 200.6) by age group 31–38 months. The mean durations then decline by age group 38–44 months, reaching 464.4 ms (SD = 214.9), and the durations across the older groups exhibit comparable values, with adults having a mean duration of 420.3 ms (SD = 181.7). The model output indicates that age group is not a strong predictor for syllable duration (β = 8.80, CI [-12.45, 32.21]).

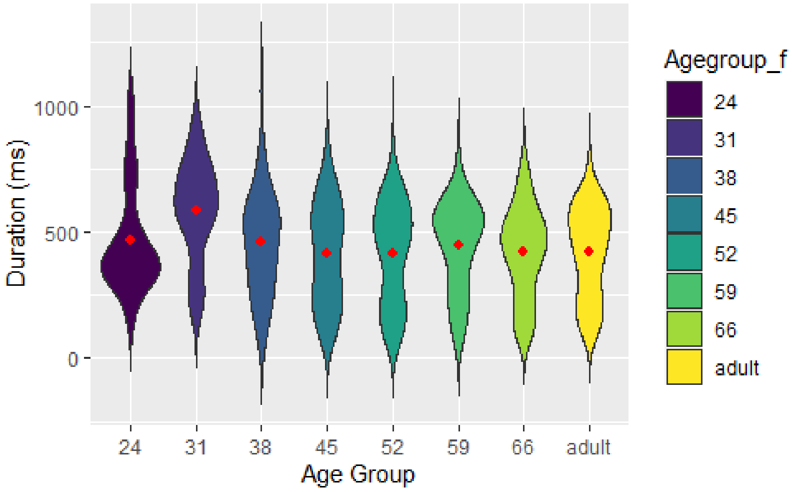
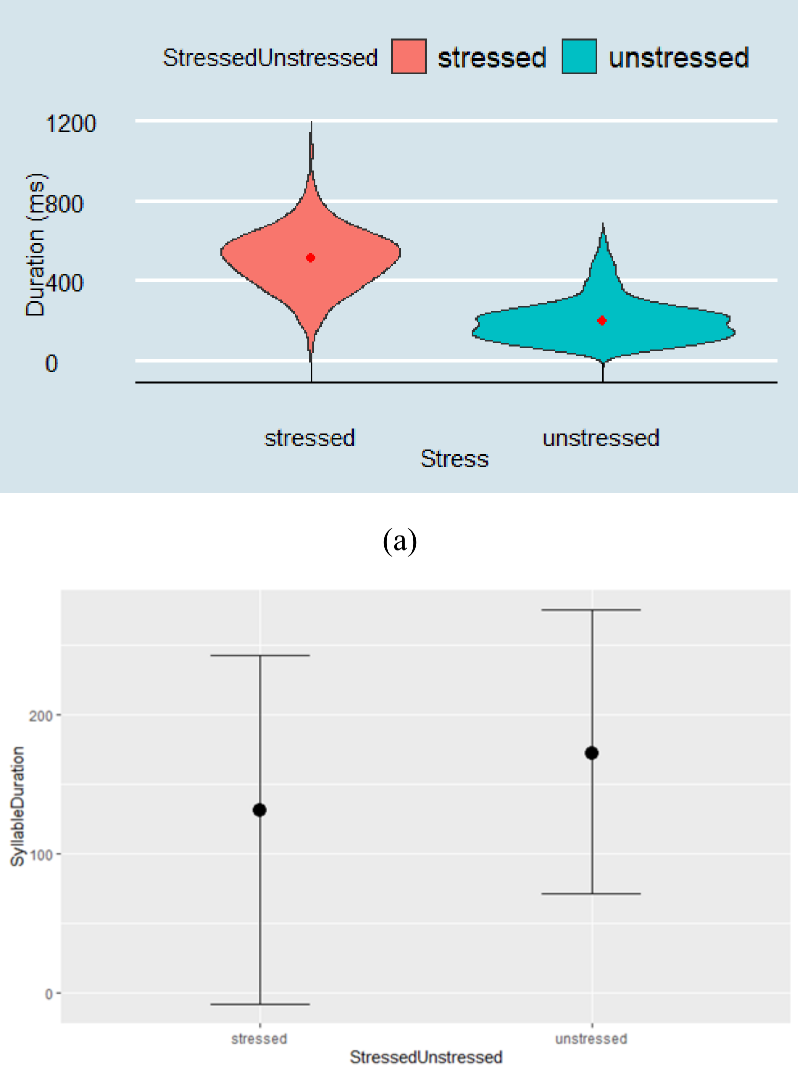


Figure 4.13: The distribution and the mean syllable duration (ms) by age group

Stress

Figure 4.14/a shows that stressed syllables (Mean = 513.5 ms, SD = 150.4) are longer by 2.5 times on average than unstressed syllables (Mean = 204.6 ms, SD = 105.4). Nonetheless, the model output shows that stress assignment is not a strong predictor for syllable duration (β = 45.26, CI [-76.17, 179.27]). Figure 4.14/b demonstrates that the median durations of unstressed syllables are higher than those of stressed syllables. There is greater uncertainty of the estimated probability of syllable durations in stressed syllables indicated by the wider CI bar than in unstressed syllables. This discrepancy in results may be attributed to variance related to random factors such as the speaker and the word item.



(a)

(b)

Figure 4.14: (a) The distribution and the mean syllable duration (ms) for stress (b) Posterior predictive plot for syllable duration for stress

Syllable Position

Figure 4.15 syllables in the word final position (Mean = 516.5 ms, SD = 145.7) are longer by 2.7 times on average than syllables in word non-final positions (Mean = 190.5 ms, SD = 87.4). Moreover, the model output indicates that syllable position is a strong predictor for syllable duration ($\beta = 143.28$, CI [29.96, 262.75]).

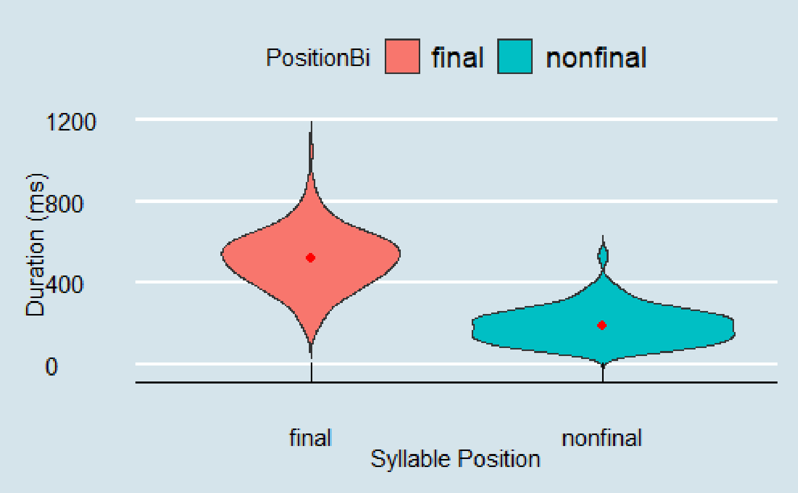


Figure 4.15: The distribution and the mean syllable duration (ms) for syllable position

Syllable Structure

Figure 4.16 demonstrates that syllable durations are affected by syllable structure. Light syllables are the shortest (Mean = 151.4 ms, SD = 60.2), followed by heavy syllables (Mean = 268.2 ms, SD = 92.4), while superheavy syllables are the longest (Mean = 532.6 ms, SD = 133.9). Superheavy syllables are 3.5 times longer than light syllables and two times longer than heavy syllables. The model output shows that syllable structure is a strong predictor for syllable duration (heavy, β = 102.00, CI [66.53, 137.49]), (superheavy, β = 299.58, CI [230.26, 368.83]).

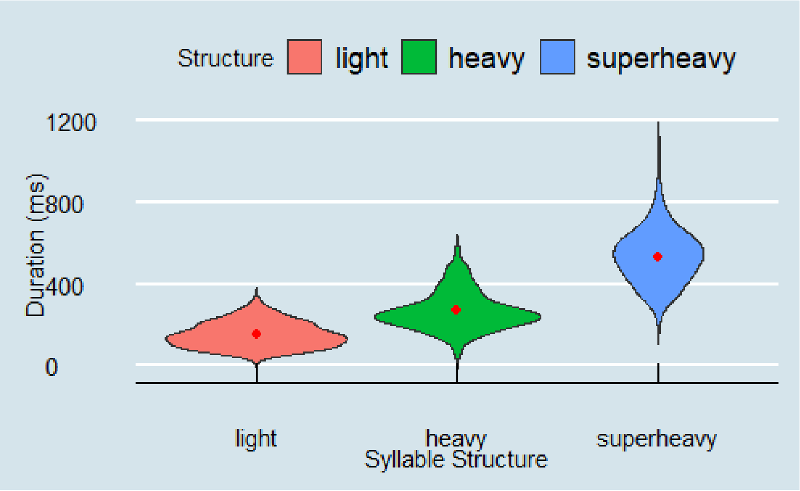


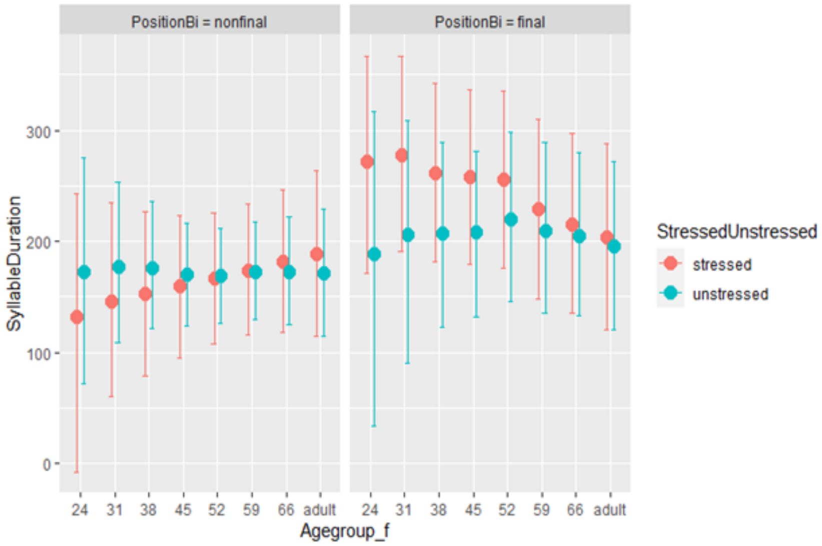
Figure 4.16: The distribution and the mean syllable duration (ms) for syllable structure

Age-Stress-Position Interaction

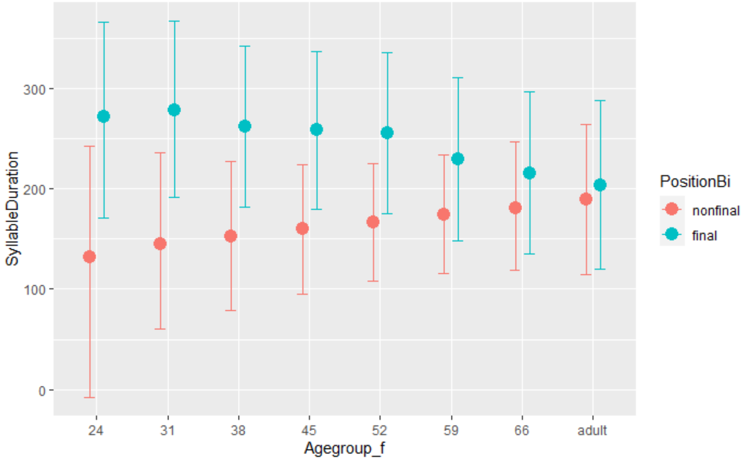
Across the age groups, stressed final syllables (Mean = 530.9 ms, SD = 135.7) are longer than stressed non-final syllables (Mean = 243.5 ms, SD = 103.2). Similarly, unstressed final syllables (Mean = 313.7 ms, SD = 131.7) are longer than unstressed non-final syllables (Mean = 179.4 ms, SD = 79.8). On average, stressed final syllables are 2.2 times longer than stressed non-final syllables, and unstressed final syllables are 1.7 times longer than unstressed non-final syllables. Notwithstanding, the model output indicates that the three-way interaction between age group, stress, and position is not a strong predictor for syllable duration (β = 20.04, CI [-10.66, 51.60]), while the two-way interaction between age group and syllable position is a strong predictor for syllable duration (β = -18.37, CI [-36.08, -0.80]).

Figure 4.17/a shows that syllable duration in stressed non-final environments is shorter than unstressed ones in the groups ranging from 24–30 to 45–51 months. However, this trend changes by 52–58 months, where syllable duration medians of stressed and unstressed non-final syllables intersect. This is followed by stressed non-final syllables having longer durations than unstressed non-final syllables by the age of 66–72 months. Furthermore, the model output suggests that stressed syllables are longer than unstressed syllables in the word-final position. Nevertheless, the durational difference between stressed and unstressed syllables for the word final position decreases with age, where the data distribution for the oldest child group, 66–72 months, and the adult group overlap.

Nevertheless, in the youngest three age groups, ranging from 24–30 to 38–44, stressed non-final syllables show greater uncertainty of probability estimates compared to older groups. The CIs become shorter with age, which suggests that the parameters could be estimated more precisely. As for the final position, less overlap is present in the younger age groups, indicating a greater difference between stressed and unstressed syllables. On the other hand, more overlap is attested in the older age groups, ranging from 52–58 months to the adult group, suggesting a weaker durational difference between stressed and unstressed syllables in the word-final position. On the other hand, Figure 4.17/b shows that as age increases, the difference in median syllable durations between word final and non-final syllables remarkably decreases. The younger age groups show a greater difference between syllables in word final and non-final positions, while older age groups exhibit a smaller difference.



(a)



(b)

Figure 4.17: (a) Posterior predictive plot for syllable duration for the interaction between age group, stress assignment, and syllable position (b) Posterior predictive plot for syllable duration for the interaction between age group and syllable position

4.2.2. Vowel Duration

Table 4.12 summarizes the mean and standard deviation for vowel durations for the independent variables, syllable structure (light, heavy, superheavy); lexical stress (stressed, unstressed); and syllable position within a word (final, non-final) for each age group.

Table 4.12: Summary of the mean and standard deviation for vowel durations for syllable structure lexical stress and syllable position for each age group in RT (n=613).

Age Group	**Syllable Structure**		**Stress**		**Syllable position**	
24–30	Light	80.2 (NA)	Unstressed	81.1 (42.0)	Non-Final	81.5 (59.4)
	Heavy	81.5 (59.4)	Stressed	228.0 (75.2)	Final	220.9 (80.1)
	Superheavy	228.0 (75.2)				
31–37	Light	108.0 (NA)	Unstressed	93.3 (46.4)	Non-Final	84.2 (33.6)
	Heavy	93.3 (46.4)	Stressed	301.7 (112.3)	Final	302.8 (110.2)
	Superheavy	314.6 (103.2)				
38–44	Light	72.9 (26.6)	Unstressed	102.1 (57.3)	Non-Final	87.1 (44.0)
	Heavy	127.2 (65.7)	Stressed	213.5 (114.1)	Final	218.2 (110.7)
	Superheavy	218.8 (112.9)				
45–51	Light	90.8 (32.8)	Unstressed	104.3 (63.4)	Non-Final	97.4 (56.0)
	Heavy	112.1 (69.2)	Stressed	224.4 (79.6)	Final	225.7 (78.3)
	Superheavy	234.8 (74.5)				
52–58	Light	74.3 (25.4)	Unstressed	84.6 (35.9)	Non-Final	77.9 (30.0)
	Heavy	124.1 (85.0)	Stressed	225.2 (65.3)	Final	228.5 (59.1)
	Superheavy	229.9 (55.5)				
59–65	Light	85.7 (26.5)	Unstressed	91.7 (28.4)	Non-Final	88.7 (28.6)
	Heavy	98.7 (31.9)	Stressed	194.1 (52.6)	Final	195.0 (50.8)

	Superheavy	197.4 (49.2)				
66–72	Light	76.5 (26.3)	Unstressed	93.1 (41.4)	Non-Final	91.0 (44.8)
	Heavy	119.3 (55.6)	Stressed	195.8 (65.6)	Final	196.6 (63.5)
	Superheavy	200.5 (61.9)				
Adult	Light	73.1 (27.5)	Unstressed	77.6 (35.1)	Non-Final	82.3 (48.8)
	Heavy	88.8 (46.3)	Stressed	182.7 (75.1)	Final	180.3 (75.1)
	Superheavy	190.0 (72.2)				

To examine the effect of the predictors on vowel duration in JA child and adult speech, vowels from 613 tokens were analysed for RT. The Bayesian analysis was performed using four chains running for 30,000 iterations. The warmup period consisted of 5000 iterations to mitigate the potential effect of the initial phase of the sampling process. The total post-warmup draws available for analysis was 100,000. Table 4.13 presents the Bayesian model output for vowel duration.

Table 4.13: Bayesian model output summary for vowel duration/RT

Data: RTC data (Number of observations: 613)
Draws: 4 chains, each with iter = 30000; warmup = 5000; thin = 1; total post-warmup draws = 1e+05

Population level effect	Estimate	Est. Error	l-95% CI	u-95% CI	Bulk_ESS	Tail_ESS
Intercept	92.54	32.65	22.20	152.21	28307	47703
Stress (Unstressed)	5.23	32.29	-55.82	72.34	29632	44766
Position (Final)	97.07	31.78	38.41	163.65	32663	51416
Structure (Heavy)	6.15	8.95	-11.38	23.75	122260	80785
Structure (Superheavy)	74.99	17.35	40.96	108.90	122162	78163
Stress (Unstressed): Position (Final)	-60.22	49.41	-161.51	33.60	32756	45823
Age group	3.61	5.76	-7.20	15.67	25065	41294
Age group: Stress (Unstressed)	-4.98	5.32	-15.63	5.45	28327	40187
Age group: Position (Final)	-12.56	4.63	-22.24	-4.05	26493	39123

Age group: Stress (Unstressed): Position (Final)	10.44	7.76	-4.58	26.22	32117	46639

Age Group

Figure 4.18 shows that vowel duration increases from 24–30 months (Mean = 208.8 ms, SD = 87.2), reaching 278.5 ms (SD = 125.6) by 31–38 months of age. At 38–44 months (Mean = 186.9 ms, SD = 113.6), a decrease in the durations is evident, and the age groups exhibit comparable values, reaching 155.8 ms (SD = 81.3) in the adult group. Nevertheless, the model output shows that age group is not a strong predictor for vowel duration (β = 3.61, CI [-7.20,15.67]).

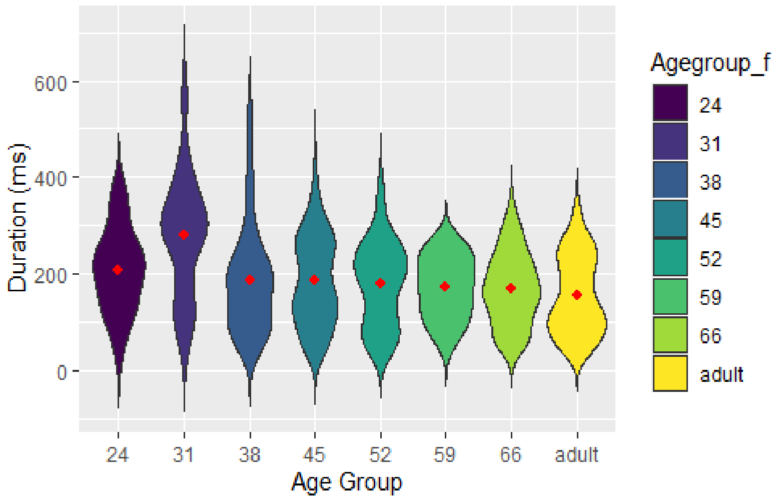
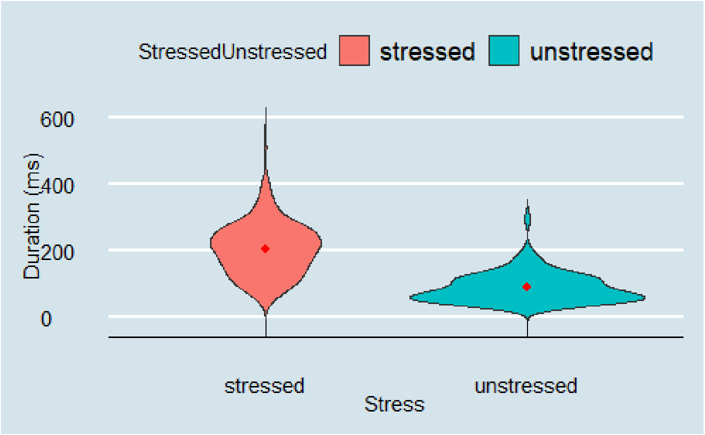


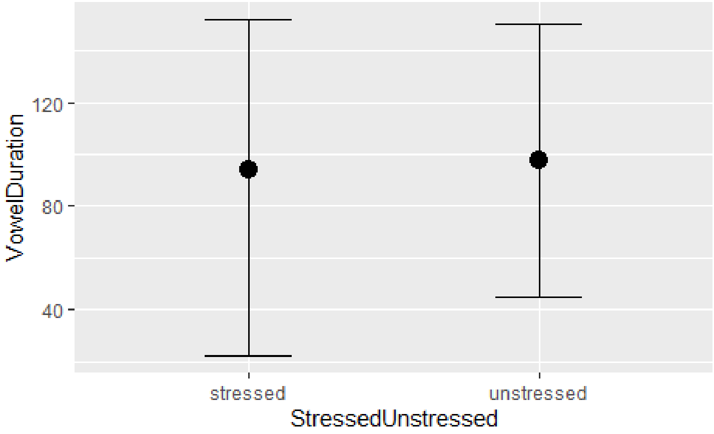
Figure 4.18: The distribution and the mean vowel duration (ms) by age group

Stress

Figure 4.19/a indicates that vowels in stressed syllables (Mean = 206.7 ms, SD = 81.6) are, on average, 2.3 times longer than vowels in unstressed syllables (Mean = 90 ms, SD = 45.2). However, the model output suggests that stress is not a strong predictor of vowel duration (β = 5.23, CI [-55.82, 72.34]). The posterior probability distribution in Figure 4.19/b shows that vowels in unstressed syllables are slightly longer than those in stressed ones. However, the wide CI, particularly for the stressed syllables, indicates uncertainty in the probability estimate of vocalic durations. This discrepancy in results may be attributed to variance related to random factors such as the speaker and the word item.



(a)



(b)

Figure 4.19: (a) The distribution and the mean vowel duration (ms) for stress (b) Posterior predictive plot for vowel duration for stress

Syllable Position

Figure 4.20 shows that vowels in the word final position (Mean = 207.1 ms, SD = 80.4) are longer than their counterparts (Mean = 86.7 ms, SD = 44.6). Approximately, vowels in word final syllables are 2.4 times longer than those in word non-final syllables. Thus, word final lengthening is observed in JA vocalic productions. The model output indicates that syllable position is a strong predictor for vowel duration (β = 97.07, CI [38.41, 163.65]).

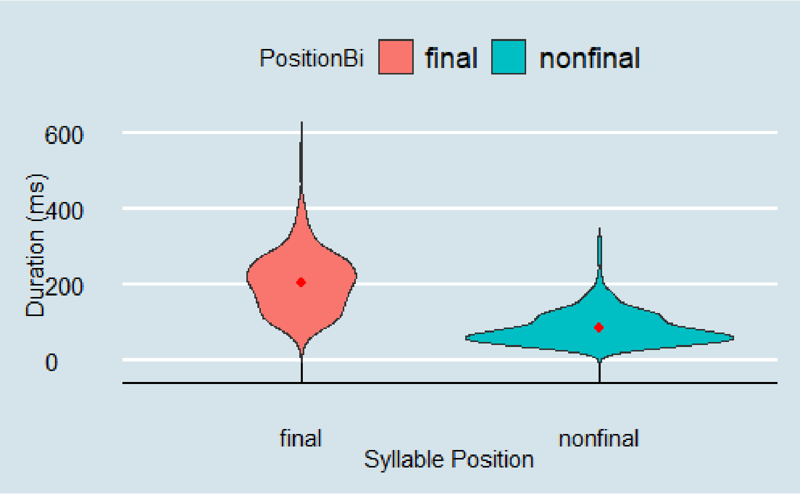


Figure 4.20: The distribution and the mean vowel duration (ms) for syllable position

Syllable Structure

Figure 4.21 demonstrated that vowels in superheavy syllables are the longest (Mean = 213 ms, SD = 27.9), followed by vowels in heavy syllables (Mean = 108.3 ms, SD = 61.9), while vowels in light syllables are the shortest (Mean = 78.5 ms, SD = 78.3). On average, vowels in superheavy syllables are two times longer than vowels in heavy syllables and 2.7 times longer than those in light syllables. The model output shows that syllable structure is a strong predictor for vowel duration in (superheavy, β = 74.99, CI [40.96, 108.90]), but not in (heavy, β = 6.15, CI [-11.38, 23.75]).

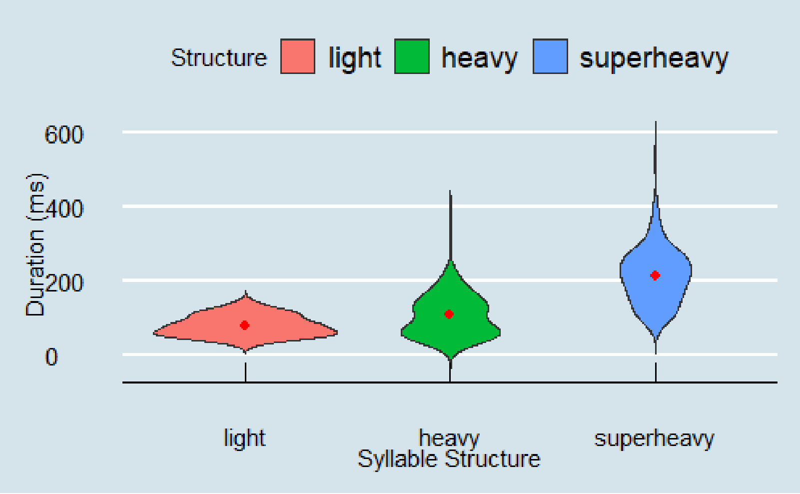
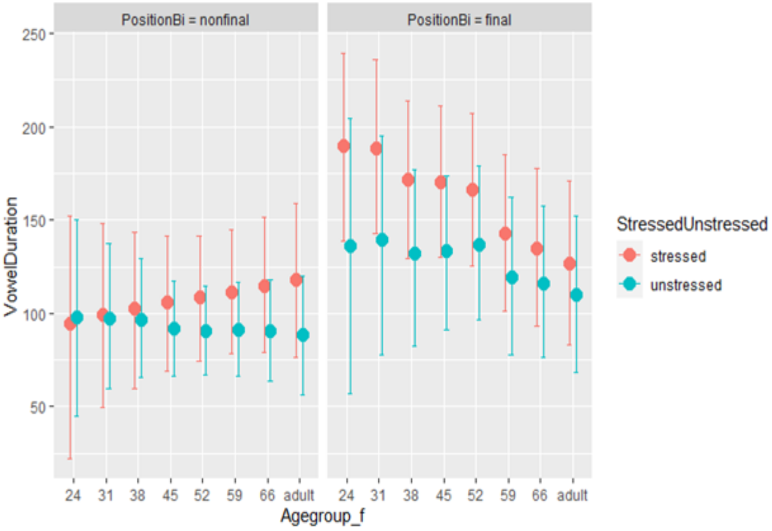


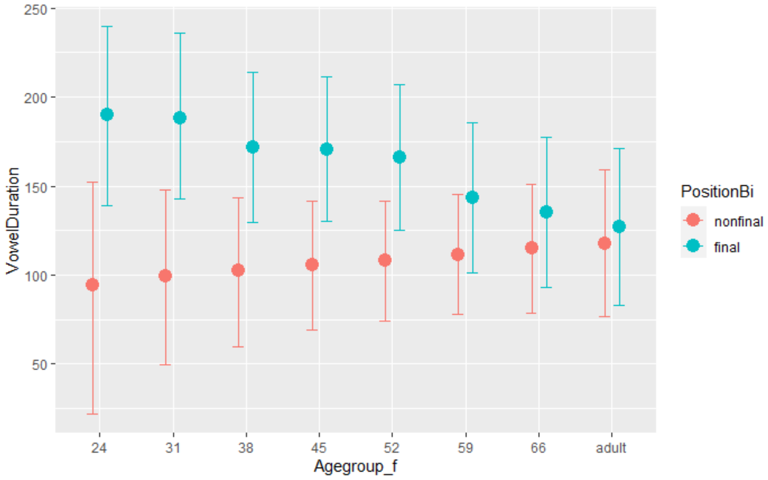
Figure 4.21: The distribution and the mean vowel duration (ms) for syllable structure

Age-Stress-Position Interaction

Across the age groups, vowels in final stressed syllables (Mean = 213.3 ms, SD = 78.5) are longer than vowels in stressed non-final syllables (Mean = 104.6 ms, SD = 58.7), and vowels in unstressed final syllables (Mean = 120.4 ms, SD = 52.8) are longer than their counterparts (Mean = 82.9 ms, SD = 40.3). On average, vowels in stressed final syllables are two times longer than their counterparts, and vowels in unstressed final syllables are 1.5 times longer than vowels in unstressed non-final syllables. Nevertheless, the model output indicates that the interaction between age group, stress, and position is not a strong predictor for vowel duration ($\beta = 10.44$, CI [-4.58, 26.22]), while the two-way interaction between age group and syllable position is a strong predictor ($\beta = -12.56$, CI [-22.24, -4.05]). Figure 4.22/a demonstrates the three-way interaction between syllable position, age group, and stress, where vowels in non-final stressed syllables are longer than vowels in unstressed syllables, except for the youngest age group. Children aged 24–30 months produce slightly longer vowels in unstressed syllables than in stressed syllables. The durational difference between vowels in non-final stressed and unstressed syllables increases with age. Additionally, the model output shows that vowels in word-final stressed syllables are longer than their counterparts. The durational difference between vowels appearing in word-final stressed and unstressed environments decreases with age. In the youngest three age groups, ranging from 24–30 to 38–44 months, the durations show a small difference between probability estimates of stressed and unstressed word non-final syllables due to the overlap attested. However, fewer CI bar overlap levels exist in the older age groups, from 52–58 months to adults, indicating a large difference between the estimates. As for the word-final position, vowels in stressed and unstressed syllables overlap more with age, indicating that the difference between probability estimates decreases. On the other hand, Figure 4.22/b demonstrates that the difference between vowel durations in word final and non-final positions becomes less evident with age. Vowels in the word final position are remarkably longer than their counterparts in the younger age groups, but by 59–65 months, vowels in word final and non-final positions do not exhibit notable durational differences.



(a)



(b)

Figure 4.22: (a) Posterior predictive plot for vowel duration for the interaction between age group, stress assignment, and syllable position (b) Posterior predictive plot for vowel duration for the interaction between age group and syllable position

4.2.3. Summary

For syllable duration, the data distribution shows that children in the youngest two age groups, ranging from 24–30 to 31–37 months, produce longer syllables than the older age groups. Stressed syllables are 2.5 times longer than unstressed syllables. Word final lengthening is observed as word final syllables are 2.7 times longer than their counterparts. Superheavy syllables are 3.5 times longer than light syllables and two times longer than heavy syllables. The Bayesian model output suggests that only syllable position and syllable structure are strong predictors for syllable duration in JA productions. The three-way interaction between age group, stress, and syllable position is not a strong predictor for syllable duration, while the two-way interaction between age group and syllable position is a strong predictor. The difference in syllable durations for word final and non-final syllables decreases with age. As for vowels, the durations decrease by 38–44 months. Vowels in stressed syllables are 2.3 times longer than their counterparts. Vowels in word final syllables are 2.4 times longer than vowels in word non-final syllables. Vowels in superheavy syllables are 2.7 times longer than vowels in light syllables and two times longer than vowels in heavy syllables. However, the Bayesian model output suggests that syllable position and syllable structure are the only strong predictors for vowel duration in JA productions. The three-way interaction between age group, stress, and syllable position is not a strong predictor for vowel duration, while the two-way interaction between age group and syllable position is a strong predictor. The difference in vowel durations for word final and non-final syllables becomes smaller with age.

4.3.Picture Elicitation Task

This section explores the effects of predictors, including age group, stress, syllable structure, syllable position, and the three-way interaction between age group, stress, and syllable position on syllable and vowel durations in JA child speech in PT. Only the child group carried out this task. The following table summarizes the attested syllable structures and their frequencies across the age groups in PT. CVC, CV, and CVVC syllables were the most frequent with 51.2%, 18.5%, and 14.4%, respectively. CCVC and CCVV were the least frequent with .2%, .1%, respectively. CCV and CCVCC syllables were not produced in this task.

Table 4.14: Frequency of syllable structures in PT

Syllable Structure	Frequency in PT	Percentage
CVC	656	51.2%
CV	237	18.5%
CVVC	184	14.4%
CVV	119	9.3%
CVCC	56	4.4%
CCVVC	26	2.0%
CCVC	3	0.2%
CCVV	1	0.1%
CCV	0	0.0%
CCVCC	0	0.0%
Total	1282	100%

4.3.1. Syllable Duration

Table 4.15 summarizes the mean and standard deviation for syllable durations for the independent variables, syllable structure (light, heavy, superheavy); lexical stress (stressed, unstressed); and syllable position within a word (final, non-final) for each age group.

Table 4.15: Summary of the mean and standard deviation for syllable durations for syllable structure lexical stress and syllable position for each age group in PT (n=1282).

Age Group	Syllable Structure		Stress		Syllable position	
24–30	Light	197.9 (91.8)	Unstressed	330.6 (129.9)	Non-Final	260.4 (71.6)

	Heavy	324.4 (112.9)	Stressed	358.2 (210.0)	Final	429.2 (218.1)
	Superheavy	512.1 (174.2)				
31–37	Light	228.0 (166.6)	Unstressed	339.5 (26.4)	Non-Final	259.8 (42.0)
	Heavy	329.1 (103.4)	Stressed	424.6 (174.9)	Final	504.4 (159.9)
	Superheavy	630.6 (189.3)				
38–44	Light	150.7(53.9)	Unstressed	280.5 (144.8)	Non-Final	231.6 (86.1)
	Heavy	299.9 (124.9)	Stressed	341.9 (186.3)	Final	392.6 (170.9)
	Superheavy	491.1 (205.6)				
45–51	Light	173.0 (66.0)	Unstressed	268.1 (119.8)	Non-Final	214.4 (94.9)
	Heavy	269.1 (91.9)	Stressed	292.2 (162.9)	Final	349.3 (154.4)
	Superheavy	428.9 (113.2)				
52–58	Light	146.0 (51.2)	Unstressed	286.5 (80.3)	Non-Final	200.9 (69.1)
	Heavy	290.5 (91.5)	Stressed	338.3 (133.0)	Final	421.5 (119.8)
	Superheavy	503.1 (119.5)				
59–65	Light	185.4 (79.9)	Unstressed	266.6 (113.8)	Non-Final	222.6 (76.9)
	Heavy	277.0 (75.7)	Stressed	314.5 (129.1)	Final	364.1 (114.4)
	Superheavy	471.6 (92.9)				
66–72	Light	156.7 (62.8)	Unstressed	297.1 (130.5)	Non-Final	204.6 (85.3)
	Heavy	285.3 (54.2)	Stressed	318.1 (90.8)	Final	411.3 (122.3)
	Superheavy	507.8 (89.8)				

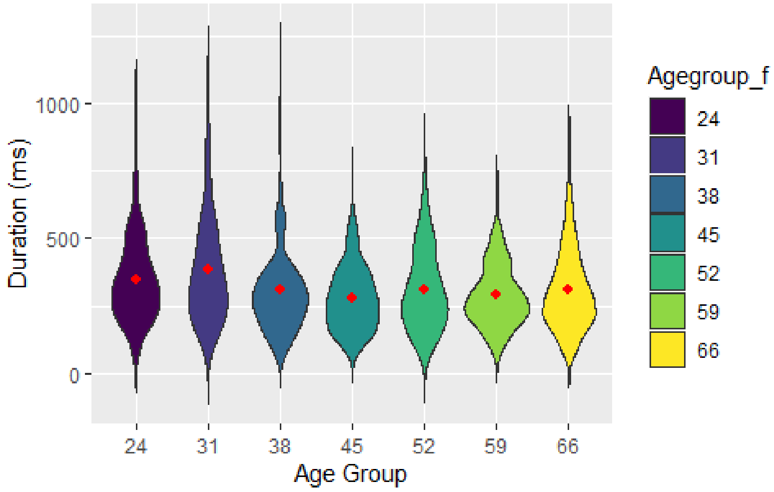
To examine the degree to which the predictors, age group, stress, syllable position, and syllable structure affect syllable durations, a total of 1282 syllables were analysed for PT. The Bayesian analysis was performed using four chains running for 30,000 iterations. The warmup period consisted of 2000 iterations to mitigate the potential effect of the initial phase of the sampling process. The total post-warmup draws available for analysis was 112,000. Table 4.16 summarizes the Bayesian model output for syllable duration in JA child productions.

Table 4.16: Bayesian model output summary for syllable duration-PT

Data: PTdata (Number of observations: 1282) Draws: 4 chains, each with iter = 30000; warmup = 2000; thin = 1; total post-warmup draws = 112000						
Population level effect	Estimate	Est. Error	l-95% CI	u-95% CI	Bulk_ESS	Tail_ESS
Intercept	168.23	19.23	130.19	206.11	47546	66269
Stress (Unstressed)	20.33	15.63	-9.07	52.37	69918	77907
Position (Final)	124.58	23.00	79.39	169.56	68572	75583
Structure (Heavy)	93.60	9.60	74.78	112.54	108078	88545
Structure (Superheavy)	264.63	19.31	226.88	302.59	89525	86191
Stress (Unstressed): Position (Final)	-35.05	27.02	-89.00	16.87	61374	75527
Age group	-3.48	4.35	-11.87	5.33	39755	56829
Age group: Stress (Unstressed)	-3.99	3.63	-11.31	2.98	71009	77661
Age group: Position (Final)	-13.82	3.37	-20.43	-7.20	70175	77186
Age group: Stress (Unstressed): Position (Final)	14.84	4.92	5.45	24.75	59922	74733

Age Group

Figure 4.23 shows that the youngest age group's mean syllable duration is 344.3 ms (SD = 154.9). Then, the mean duration increases and reaches a maximum of 382.1 ms (SD = 201.4) at 31–38 months. However, by 38–44 months (Mean = 310.6 ms, SD = 174.8), the durations decrease, and the age groups exhibit comparable values ranging from 280 ms to 313 ms. The oldest age group, 66–72 months, has a mean duration of 307.6 ms (SD = 147.3). However, the model output suggests that age group is not a strong predictor for syllable duration (β = -3.48, CI [-11.87, 5.33]).

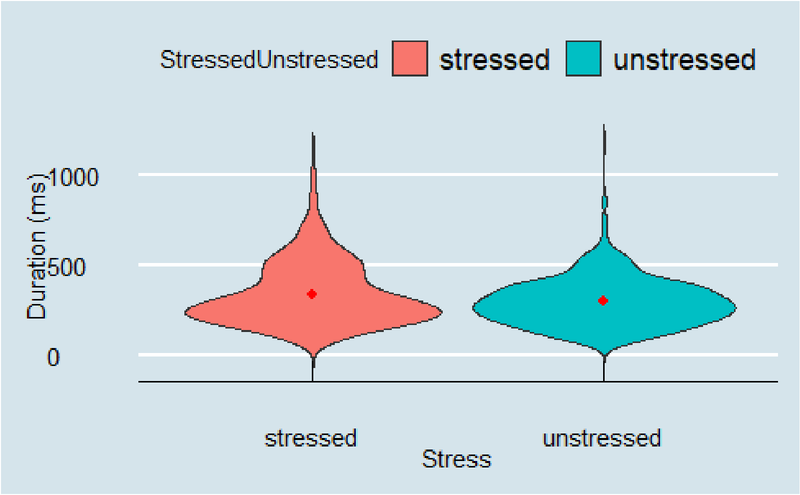


(a)

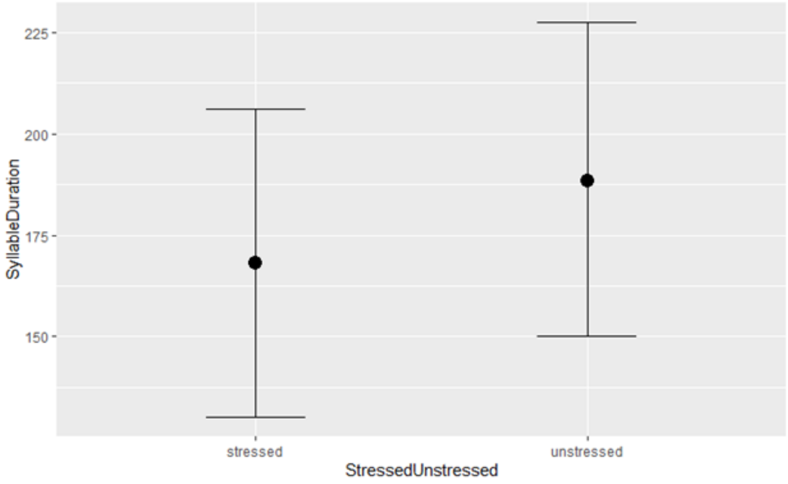
Figure 4.23: The distribution and the mean syllable duration (ms) for age group

Stress

Figure 4.24/a suggests that stressed syllables (Mean = 340.7 ms, SD = 177.4) are, on average, 1.2 times longer than unstressed syllables (Mean = 296.2 ms, SD = 134.1). However, the posterior predictive distribution results in Figure 4.24/b show that unstressed syllables exhibit longer durations than their counterparts. The model output indicates that stress assignment is not a strong predictor for syllable duration (β = 20.33, CI [-9.07, 52.37]).



(a)



(b)

Figure 4.24: (a) The distribution and the mean syllable duration (ms) for stress (b) Posterior predictive plot for syllable duration for stress

Syllable Position

Figure 4.25 shows that word final syllables (Mean = 413.5 ms, SD = 154.6) are approximately 1.8 times longer than word non-final syllables (Mean = 225 ms, SD = 93.7). The model output suggests that syllable position is a strong predictor for syllable duration (β = 124.58, CI [79.39, 169.56]).

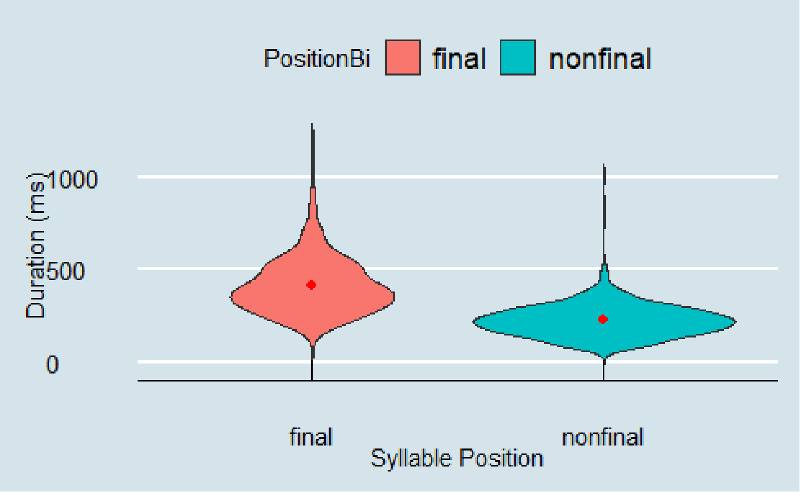


Figure 4.25: The distribution and the mean syllable duration (ms) for syllable position

Syllable Structure

Figure 4.26 demonstrates that JA child productions are affected by syllable structure, where light syllables are the shortest (Mean = 172.9 ms, SD = 89.2), followed by heavy syllables (Mean = 295.9 ms, SD = 100.8), while superheavy syllables are the longest (Mean = 513.3 ms, SD = 161.4). On average, superheavy syllables are three times longer than light syllables and 1.7 times longer than heavy syllables. The model output indicates that syllable structure is a strong predictor for syllable duration for (superheavy, β = 264.63, CI [226.88, 302.59]), (heavy, β = 93.60, CI [74.78, 112.54]).

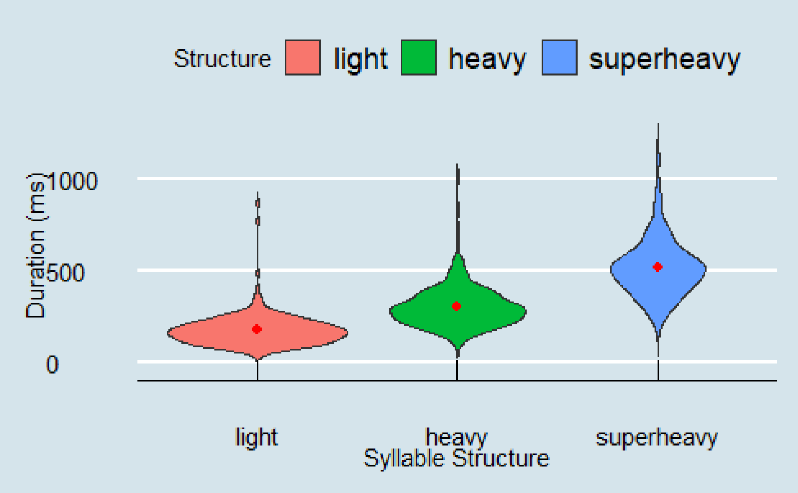


Figure 4.26:The distribution and the mean syllable duration (ms) for syllable structure

Age-Stress-Position Interaction

Across the age groups, final stressed syllables (Mean = 511.3 ms, SD = 157.7) are longer than stressed non-final syllables (Mean = 239.1 ms, SD = 88.1), and final unstressed syllables (Mean = 355.2 ms, SD = 119.5) are longer than unstressed non-final syllables (Mean = 202.6 ms, SD = 98.2). On average, stressed final syllables are 2.1 times longer than stressed non-final ones, while unstressed final syllables are 1.8 times longer than their counterparts. The model output indicates that the three-way interaction between age group, stress, and position is a strong predictor for syllable duration (β = 14.84, CI [5.45, 24.75]). Figure 4.27 demonstrates the three-way interaction between age group, stress, and position. JA children, in age groups 24–30 to 52–58 months, produce shorter syllables in stressed environments than unstressed ones in word non-final positions. However, this durational trend changes at 59–65 months, where there is an overlap between the medians of stressed and unstressed non-final syllables. Subsequently, children in the oldest age produce slightly longer stressed syllables in non-final positions than unstressed syllables. As for the word-final position, children in the first two age groups, 24–30 to 31–37 months, produce longer stressed syllables than unstressed ones. However, by age 38–

44 months, JA children start producing longer unstressed syllables when compared to stressed ones.

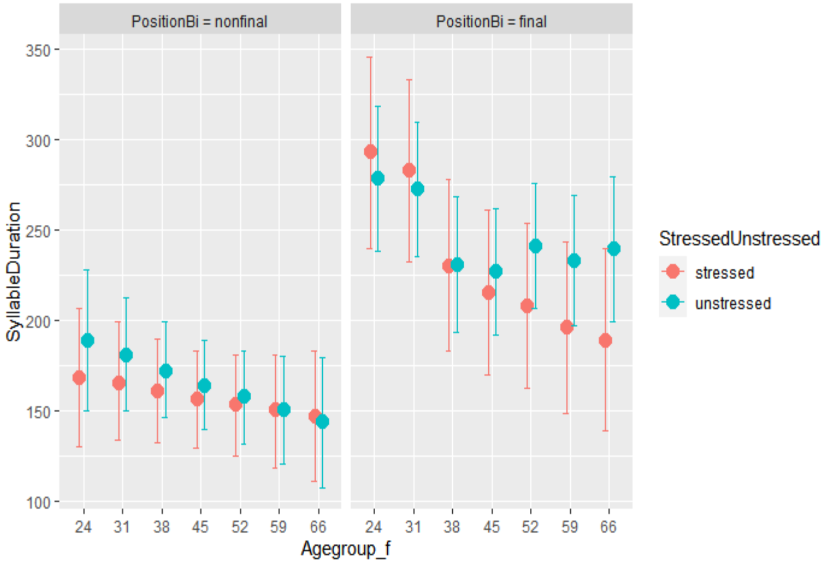


Figure 4.27: Posterior predictive plot for syllable duration for the interaction between age group, stress assignment, and syllable position

4.3.2. Vowel Duration

Table 4.17 summarizes the mean and standard deviation for vowel durations for the independent variables, syllable structure (light, heavy, superheavy); lexical stress (stressed, unstressed); and syllable position within a word (final, non-final) for each age group.

Table 4.17: Summary of the mean and standard deviation for vowel durations for syllable structure lexical stress and syllable position for each age group in PT (n=1282).

Age Group	**Syllable Structure**		**Stress**		**Syllable position**	
24–30	Light	103.1 (41.7)	Unstressed	129.0 (59.4)	Non-Final	125.2 (69.3)
	Heavy	129.6 (65.4)	Stressed	161.8 (112.6)	Final	165.7 (105.3)
	Superheavy	227.9 (135.0)				
31–37	Light	98.7 (34.6)	Unstressed	132.9 (72.9)	Non-Final	121.0 (62.5)
	Heavy	137.5 (65.9)	Stressed	191.3 (112.7)	Final	203.3 (111.4)

	Superheavy	271.2 (115.8)				
38–44	Light	81.3 (25.8)	Unstressed	103.4 (72.2)	Non-Final	103.6 (49.9)
	Heavy	110.5 (52.1)	Stressed	154.4 (84.2)	Final	154.2 (99.7)
	Superheavy	222.9 (109.9)				
45–51	Light	81.6 (34.1)	Unstressed	105.6 (50.7)	Non-Final	91.3 (48.0)
	Heavy	105.7 (52.6)	Stressed	117.7 (64.9)	Final	133.0 (60.5)
	Superheavy	163.7 (65.8)				
52–58	Light	84.0 (30.1)	Unstressed	122.7 (63.6)	Non-Final	91.7 (50.0)
	Heavy	118.9 (58.1)	Stressed	145.8 (82.8)	Final	175.9 (71.3)
	Superheavy	211.4 (76.6)				
59–65	Light	92.5 (33.1)	Unstressed	112.1 (47.3)	Non-Final	102.4 (50.5)
	Heavy	115.8 (49.7)	Stressed	141.7 (75.0)	Final	152.9 (66.5)
	Superheavy	210.2 (72.4)				
66–72	Light	86.0 (26.8)	Unstressed	126.4 (56.2)	Non-Final	97.8 (47.2)
	Heavy	123.2 (55.6)	Stressed	147.8 (85.8)	Final	176.6 (73.4)
	Superheavy	224.2 (76.8)				

A total of 1282 tokens were analysed for PT to investigate the effects of the predictors on JA child vowel durations. The Bayesian analysis was performed using four chains running for 30,000 iterations. The warmup period consisted of 2000 iterations, and the total post-warmup draws available for analysis was 112,000. Table 4.18 shows the Bayesian model output for vowel durations.

Table 4.18: Bayesian model output summary for vowel duration/PT

Data: PTdata (Number of observations: 1282) **Draws: 4 chains, each with iter = 30000; warmup = 2000; thin = 1; total post-warmup draws = 112000**

Population level effect	Estimate	Est. Error	l-95% CI	u-95% CI	Bulk_ESS	Tail_ESS
Intercept	98.10	11.70	75.31	121.24	62875	79451
Stress (Unstressed)	-4.43	10.29	-23.94	16.95	75572	77855
Position (Final)	27.16	15.16	-1.96	57.28	70470	84163
Structure (Heavy)	22.06	5.89	10.46	33.57	165707	94401
Structure (Superheavy)	110.23	11.84	87.01	133.45	130699	92700
Stress (Unstressed): Position (Final)	-22.84	18.86	-61.02	13.26	64654	78343
Age group	-2.57	2.61	-7.53	2.75	49650	69095
Age group: Stress (Unstressed)	-1.91	2.50	-6.96	2.88	68829	76505
Age group: Position (Final)	-1.21	2.98	-6.89	4.61	55117	84733
Age group: Stress (Unstressed): Position (Final)	7.79	3.86	0.25	15.48	52469	72919

Age Group

In age group 24–30 months, the mean vowel duration is 145.3 ms (SD = 91.1), followed by an increase at 31–37 months, reaching the mean 162.1 ms (SD = 99). Then the durations start exhibiting a decreasing trend by 45–51 months (Mean = 111.5 ms, SD = 58.1). The durations increase slightly in the 52–58 months group (Mean = 134.4 ms, SD = 74.7), while the oldest age group has a mean duration of 137.1 ms (SD = 73.1). The model output shows that age group is not a strong predictor for vowel duration (β = -2.57, CI [-7.53, 2.75]).

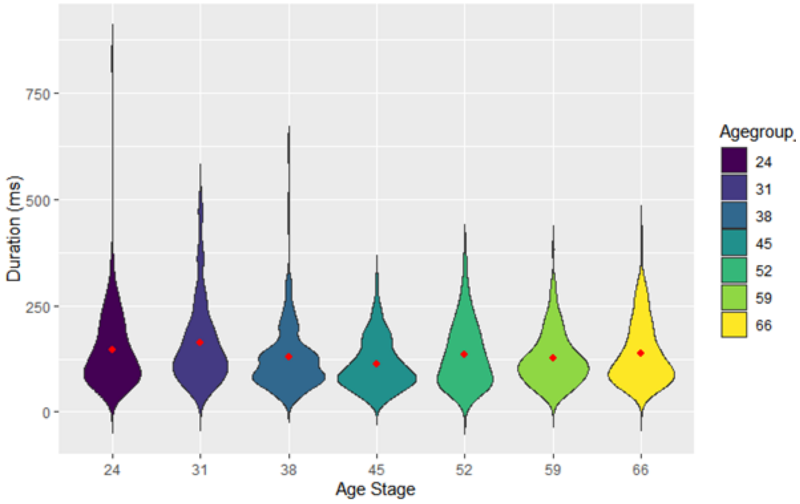
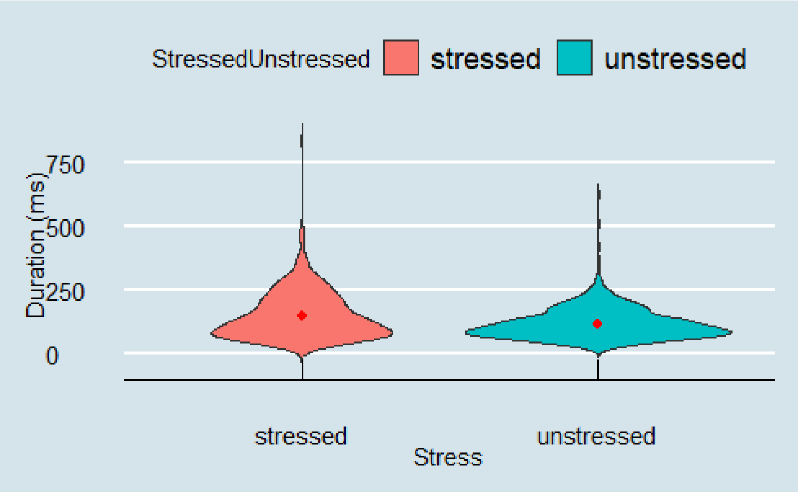


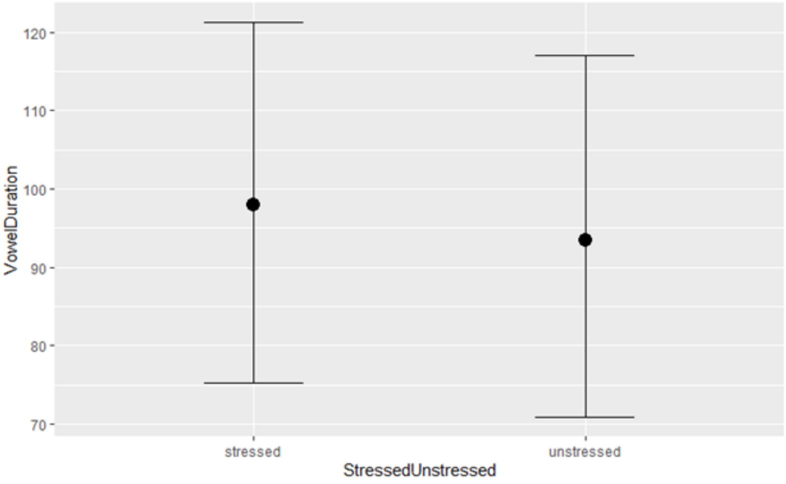
Figure 4.28:The distribution and the mean vowel duration (ms) by age group

Stress

Figure 4.29/a shows that vowels in stressed syllables (Mean = 151.8 ms, SD = 91.6) are marginally longer than vowels in unstressed syllables (Mean = 119.9 ms, SD = 61.6). On average, vowels in stressed syllables are 1.3 times longer than vowels in unstressed syllables. The model output in Figure 4.29/b suggests that stress assignment is not a strong predictor for vowel duration (β = -4.43, CI [-23.94, 16.95]). JA children produce marginally longer vowels in stressed environments than in unstressed ones.



(a)



(b)

Figure 4.29: (a) The distribution and the mean vowel duration (ms) for stress (b) Posterior predictive plot for vowel duration for stress

Syllable Position

Vowels in word final syllables (Mean = 168.4 ms, SD = 87.2) are approximately 1.6 times longer than vowels in word non-final syllables (Mean = 103.7 ms, SD = 54.9), as demonstrated in Figure 4.30. Nonetheless, the model output indicates that syllable position is not a strong predictor for vowel duration (β = 27.16, CI [-1.96, 57.28]).

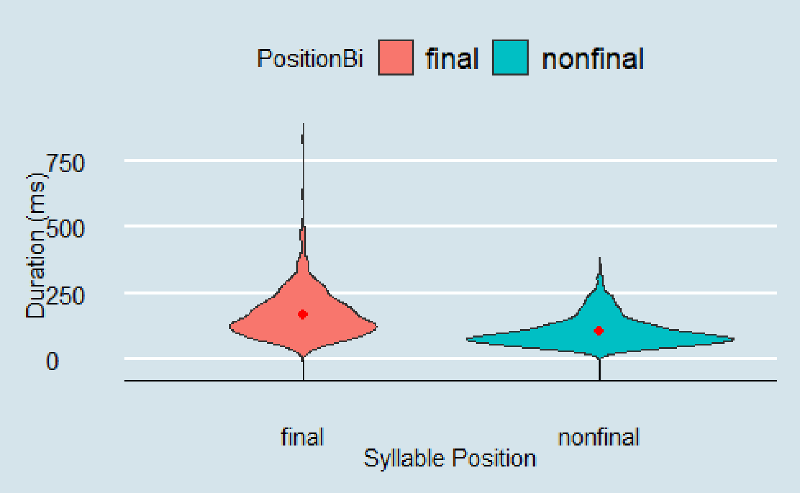


Figure 4.30:The distribution and the mean vowel duration (ms) for syllable position

Syllable Structure

Figure 4.31 demonstrates that JA vocalic productions are affected by syllable structure. Vowels in light syllables are the shortest (Mean = 88.5 ms, SD = 32.1), followed by vowels in heavy syllables (Mean = 120.6 ms, SD = 58.1), then vowels in superheavy syllables are the longest (Mean = 221.9 ms, SD = 98.3). On average, vowels in superheavy syllables are 2.5 times longer than vowels in light syllables and 1.8 times longer than vowels in heavy syllables. The model output indicates that syllable structure is a strong predictor for vowel duration (heavy, β = 22.06, CI [10.46, 33.57]; superheavy, β = 110.23, CI [87.01, 133.45]).

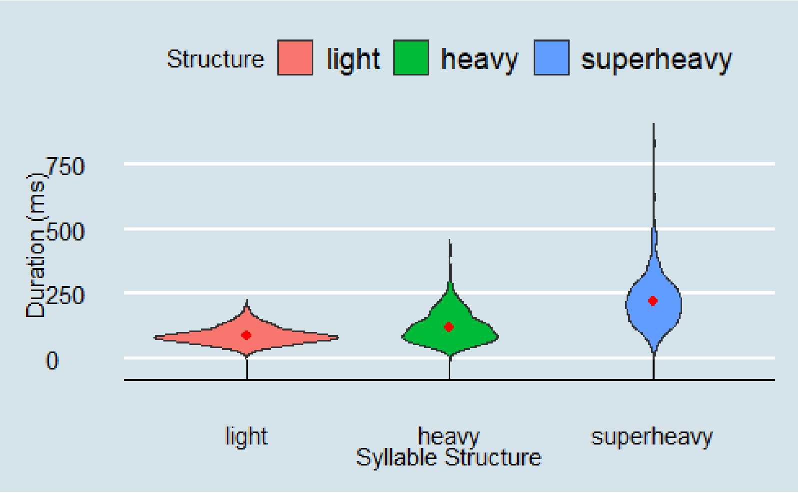


Figure 4.31: The distribution and the mean vowel duration (ms) for syllable structure

Age-Stress-Position Interaction

Across the age groups, vowels in final stressed syllables (Mean = 222.4 ms, SD = 157.7) are longer than vowels in stressed non-final syllables (Mean = 109.7 ms, SD = 88.1), while vowels in final unstressed syllables (Mean = 136.1 ms, SD = 119.5) are longer than vowels in unstressed non-final syllables (Mean = 94.2 ms, SD = 98.2). On average, vowels in stressed final syllables are two times longer than their counterparts, while vowels in unstressed final syllables are 1.4 times longer than their counterparts. The model output indicates that the three-way interaction between age, stress, and position is a strong predictor for vowel duration (β = 7.79, CI [.25, 15.48]). Figure 4.32 shows that JA children produce longer vowels in stressed non-final positioned syllables than their counterparts. The durational difference between vowels in stressed and unstressed non-final syllables increases with age. Moreover, results suggest that children from the 24–30 to 45–51 months groups produce longer vowels in stressed final syllables. However, this durational trend changes by 52–58 months of age, where JA children produce longer vowels in unstressed final syllables compared to their counterparts. Vowel durations in stressed final syllables decrease with age, while the vowel durations in unstressed final syllables increase.

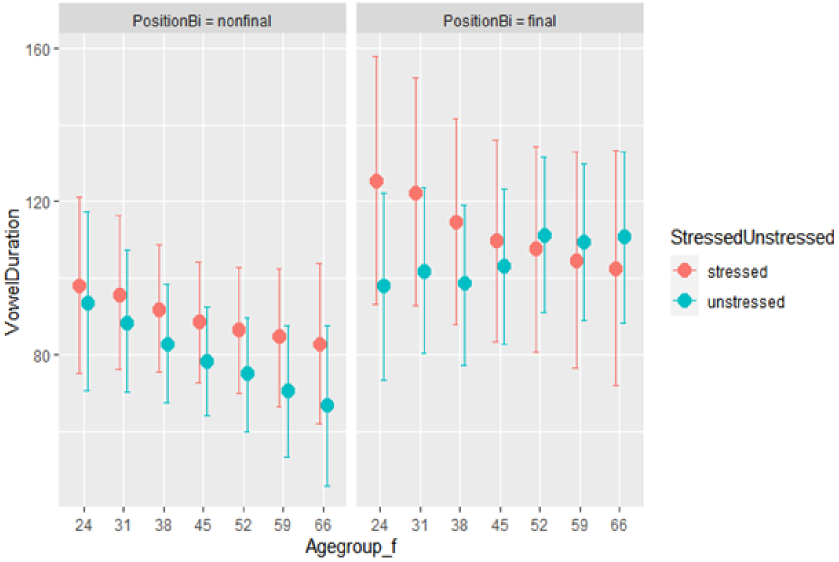


Figure 4.32: Posterior predictive plot for vowel duration for the interaction between age group, stress assignment, and syllable position

4.3.3. Summary and Interim Discussion

The youngest two age groups, ranging from 24–30 to 31–37 months, produced the longest syllables, while the durations decreased by age group 45–51 months, showing comparable values across the older groups. On average, stressed syllables are only 1.2 times longer than unstressed syllables. Word final lengthening is observed as word final syllables are 1.8 times longer than word non-final syllables. Superheavy syllables are three times longer than light syllables and 1.7 times longer than heavy syllables. The Bayesian model output indicates that all predictors played a role in determining the syllable and vowel durations; they significantly affected the durations or were involved in the interaction effects. Syllable position and syllable structure in addition to the three-way interaction between age, stress, and position are strong predictors for syllable duration in JA child productions. Children aged groups 24–30 to 52–58 months produce longer unstressed non-final syllables than stressed non-final ones. By 59–65 months, the durational difference between stressed and unstressed syllables becomes less evident. Stressed word final syllables are longer than unstressed final syllables only in the youngest two age groups. Older age groups demonstrated the production of longer unstressed final syllables than stressed ones.

As for vowels, younger age groups exhibit longer durations than older groups, and a decrease is evident by 38–44 months. Vowels in stressed syllables are 1.3 times longer than unstressed syllables. Vowels in the word final position are 1.6 times longer than vowels in word non-final positions. Vowels in superheavy syllables are the longest, followed by vowels in heavy syllables, while vowels in light syllables are the shortest. Nonetheless, the Bayesian model output indicates that syllable structure and the three-way interaction between age, stress, and position are the only strong predictors for vowel duration in JA child productions. Vowels in stressed non-final syllables are longer than vowels in unstressed non-final syllables. On the other hand, age groups 24–30 to 45–51 months produce longer vowels in stressed final syllables, while age groups from 52–58 months exhibit longer vowels in unstressed final syllables compared to their stressed final ones. The following table summarizes the Bayesian model outputs across the tasks.

Table 4.19: Summary of Bayesian model outputs across the tasks showing the effects of predictors (S: Strong, NS: Not Strong)

Task	**ST**		**RT**		**PT**	
Duration	**Syllable**	**Vowel**	**Syllable**	**Vowel**	**Syllable**	**Vowel**
Age group	NS	S	NS	NS	NS	NS
Stress	NS	NS	NS	NS	NS	NS
Syllable position	S	S	S	S	S	NS
Syllable structure	S	S	S	S-Superheavy Only	S	S
Stress x position	S	S	NS	NS	NS	NS
Age x stress	NS	NS	NS	NS	NS	NS
Age x position	S	S	S	S	S	NS
Age x stress x position	S	S	NS	NS	S	S

Moreover, two findings need to be highlighted regarding the three-way interaction effects. First, the three-way interaction effects were evident across ST and PT but not for RT. This could be attributed to RT's nature focusing on superheavy syllables (Section 4.2). These syllables appear in stressed word final environments, which explains the three-way interaction not being a strong predictor for syllable or vowel durations. Alternatively, the two-way interaction between age group and syllable position was a strong predictor for syllable and vowel durations probably due to word final lengthening effects. Second, although the three-way interaction was a strong predictor for syllable and vowel durations in ST and PT, the interaction showed different developmental patterns. The younger age groups produced larger stressed vs. unstressed differences for the final syllables in ST (Section 4.1.1, Figure 4.5), whereas the difference between stressed and unstressed final syllables was remarkably smaller in PT

(Section 4.3.1, Figure 4.27). In ST, the older age groups produced longer unstressed final syllables than stressed final ones, exhibiting a decreasing trend. However, the older age groups in PT produced longer unstressed final syllables than stressed ones, with only unstressed syllables exhibiting an increasing trend. The source of this interaction is not yet clear, as it may be attributed to the word properties evident in the tasks. That is, PT had more monosyllabic words, while ST had higher frequencies of disyllabic and multisyllabic words, which might have contributed to the interaction effects across the age groups. Thus, the second part of the analysis, which focuses on superheavy syllables, has the word length incorporated as a variable. The target superheavy syllables were divided into three groups according to the word length, including monosyllabic, disyllabic, and multisyllabic words.

The following is a summary of predictions and whether they were met or not based on the results.

a) A decreasing trend in durations will be evident, with younger age groups producing longer syllables compared to older age groups	✓
b) The oldest age group will exhibit durations that are close to the adult ones, but the durations of the two groups will not intersect	✓
c) Lexical stress will influence syllables, resulting in longer durations for stressed vowels/syllables compared to unstressed counterparts	Descriptive statistics ✓ Bayesian model output **X**
d) Word final lengthening will be evident in longer durations for word final syllables and vowels compared to non-final ones	✓
e) Syllable structure complexity will influence durations, with syllables containing more constituents displaying longer durations	✓

4.4.Superheavy Syllables

This section consists of three parts examining superheavy syllable production: ST, RT, and PT, across JA groups. For each task, the frequency of superheavy syllable productions is reported including factors, such as word lengths and sub-structure (the composition of vowel length: short/long; clusters: onset/coda of the superheavy syllable including CVCC, CVVC, CCVVC, and CCVCC syllables). Furthermore, the analysis incorporates Bayesian models and the mean durational distribution results to investigate the effects of predictors and their interactions, such as age group, sub-structure, syllable position within a word, and lexical stress on superheavy syllables across the word lengths.

4.4.1. Semi-Spontaneous Speech Task

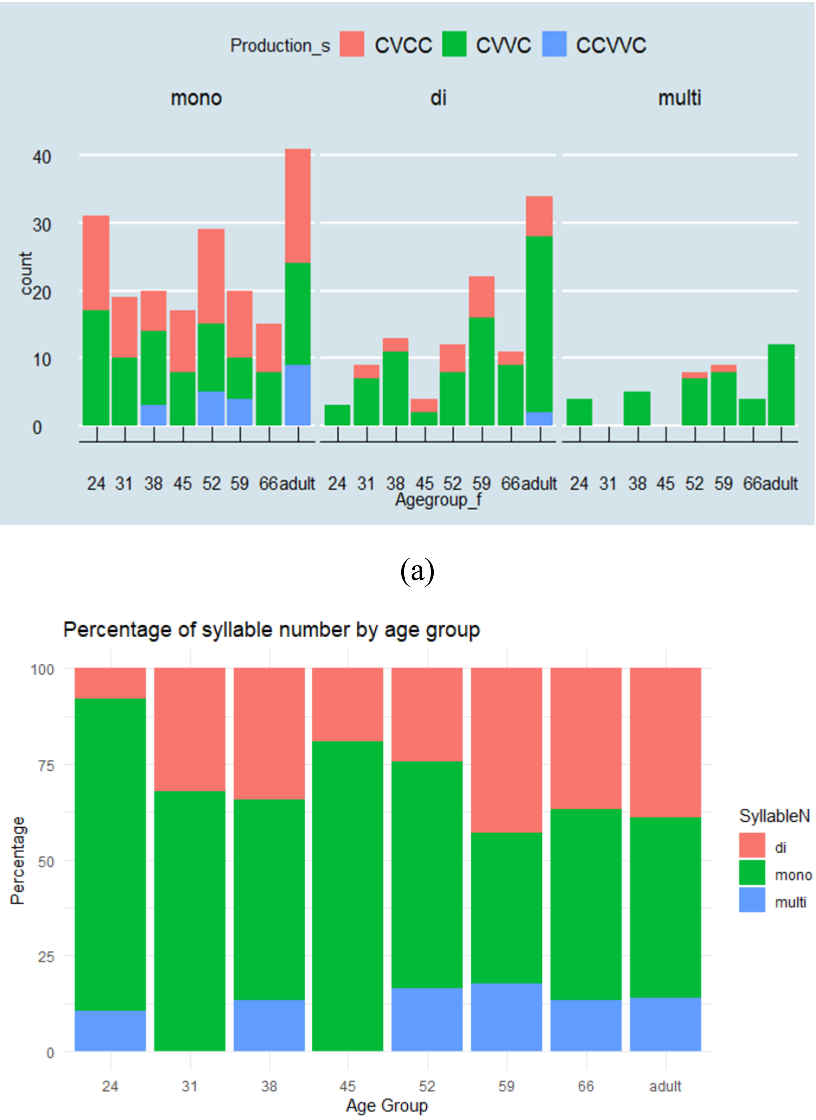
Lexical Items

The child's ability to produce superheavy syllables is mediated by lexical diversity and complexity (Vihman 1996). Children in the younger age groups tended to use a limited lexical set compared to older children. In age groups 24–30, 31–37, and 38–44 months, children produced words comprised of superheavy syllables in three main categories, names, onomatopoeic words, and yes/no answers. For example, Speaker 1IG (age group 24–30 months) produced the word /ʕabiir/ ('Abeer', the mother's name, CV.CVVC). Speaker M37 (age group 31–37 months) produced /d͡ʒawaad/ ('Jawad', the child's name, CV.CVVC) as he was answering the question, "What is your name?". Speaker 1IG (24–30 months), M0N (24–30 months), and DDK (31–37 months), all produced the word /ʕaww/ ('for dogs', CVCC) when they were answering the question "What is your favorite animal?". The word /maaʕ/ ('for sheep', CVVC) was produced by Speaker 2GS (31–37 months) to answer the aforementioned question. Speaker 1IG produced the word /ʕann/ ('for cars', CVCC) as he described his favorite outdoor activity with his parents. Additionally, multiple examples of the words /ʔaah/ ('yes', CVVC) and /laʔʔ/ (‘no’, CVCC), as a yes/no answer, were produced by speakers in the younger age groups, including Speaker 1IG, Speaker DDK, Speaker YRO, and Speaker 2GS.

Older children in age groups 45–51 months to 66–72 months produced superheavy syllables in words that entail morphological patterns (morphological boundary marked by the dash -) such as plural forms and derivatives, including prefixes and suffixes. Examples of plural forms include /marad͡ʒiiħ/ (‘swings’, CV.CV.CVVC, Speaker 3Q1, 66–72 months); /sanaw-aat/ (‘years’, CV.CV.CVVC, Speaker 6V9, 59–65 months); /sajjaar-aat/ (‘cars’, CVC.CVV.CVVC, Speaker 3Q1, 66–72 months); /ħajwaa-naat/ (‘animals’, CVC.CVV.CVVC, Speaker MD5, 52–58 months); and /ʔal-ʕaab/ (‘toys’, CVC.CVVC, Speaker 3Q1, 66–72 months). As for derivative forms, children in the older age groups produced superheavy syllables in words with prefixes such as the word /ʔil-ʔalwaan/ (‘the colors’, CVC.CVC.CVVC, Speaker 6V9, 59–65 months); /ʔil-fiil/ (‘the elephant, CVC.CVVC, Speaker MD5, 52–58 months); /ʔil-bajt/ (‘the house’, CVC.CVCC, Speaker POX, 45–51 months). JA children also produced superheavy syllables in words with suffixes such as the word /baħibb-hum/ (‘I love them’, CV.CVCC.CVC, Speaker 6V9, 59–65 months); /bitˤalʕuu-naaʃ/ (‘they do not take us out’, CV.CVC.CVV.CVVC, Speaker 6V9, 59–65 months); /ħilw-iin/ (‘they are pretty’, CVC.CVVC, Speaker QR6, 66–72 months); /maʕ-aah/ (‘with him’, CV.CVVC, Speaker MMA, 66–72 months); and /maask-ih/ (‘holding, Female’, CVVC.CVC Speaker 6V9, 59–65 months).

4.4.1.1.Superheavy Syllable Frequency

Figure 4.33/a shows the frequency of superheavy syllables in monosyllabic, disyllabic, and multisyllabic words across the age groups, ranging from 24–30 months to the adult group. Superheavy syllables are the most frequent in monosyllabic words, followed by disyllabic words, and they are the least frequent in multisyllabic words. Additionally, more sub-structural varieties of superheavy syllables are attested in monosyllabic words compared to disyllabic and multisyllabic words. Furthermore, younger age groups produced fewer sub-structures of superheavy syllables compared to older age groups. Figure 4.33/b presents an overview of the distribution of superheavy syllable production percentages based on the number of syllables in the target word by the age group predictor. From age group 24–30 months to 45–51 months, children mainly produced superheavy syllables as monosyllabic words, as it has the highest production percentage. However, there is a decline in monosyllabic productions as age increases, with age groups ranging from 52–58 months to adults displaying an increase in the production of disyllabic and multisyllabic words.



(a)

(b)

Figure 4.33: The frequency and the percentage of superheavy syllables in ST by age group

Table 4.20 shows the frequency of superheavy syllables in JA child and adult productions. The table shows the following variables: age group (children: 24–30, 31–37, 38–44, 45–51, 52–58, 59–65, 66–72 months, and adults), word length (monosyllabic, disyllabic, and multisyllabic), sub–structure (CVCC, CVVC, and CCVVC), stress assignment (stressed, unstressed), and syllable position (final, nonfinal).

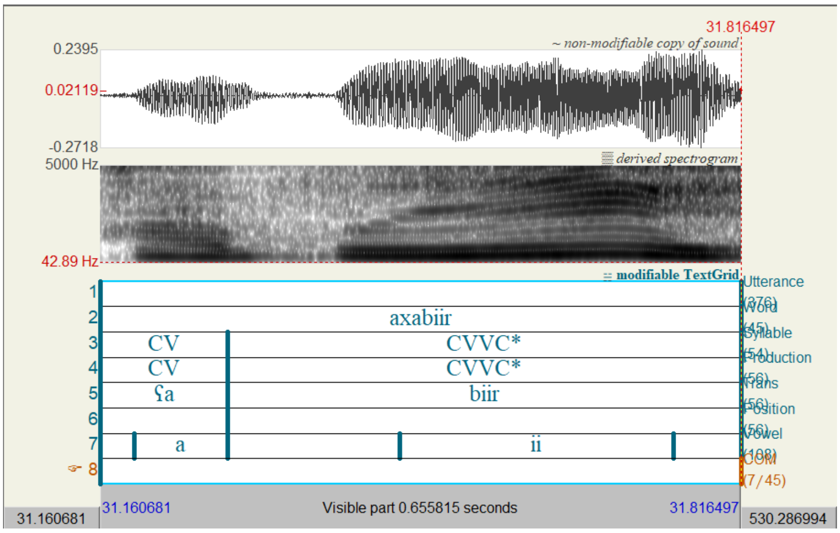
Table 4.20:Superheavy syllables ST

Age Group	Word Length	Sub–Structure	Stress	Syllable Position	Frequency (n)
24–30	Monosyllabic	CVCC	Stressed	Final	14
		CVVC	Stressed	Final	17
	Disyllabic	CVVC	Stressed	Final	3
	Multisyllabic	CVVC	Stressed	Final	4
31–37	Monosyllabic	CVCC	Stressed	Final	9
		CVVC	Stressed	Final	10
	Disyllabic	CVCC	Stressed	Final	2
		CVVC	Unstressed	Final	1
		CVVC	Stressed	Final	6
38–44	Monosyllabic	CVCC	Stressed	Final	6
		CVVC	Stressed	Final	11
		CCVVC	Stressed	Final	3
	Disyllabic	CVCC	Unstressed	Nonfinal	2
		CVVC	Unstressed	Final	1
		CVVC	Stressed	Final	10
	Multisyllabic	CVVC	Stressed	Nonfinal	2
		CVVC	Stressed	Final	3
45–51	Monosyllabic	CVCC	Stressed	Final	9
		CVVC	Stressed	Final	8
	Disyllabic	CVCC	Stressed	Final	2
		CVVC	Stressed	Final	2
52–58	Monosyllabic	CVCC	Stressed	Final	14
		CVVC	Stressed	Final	10
		CCVVC	Stressed	Final	5
	Disyllabic	CVCC	Stressed	Final	4
		CVVC	Unstressed	Final	1
		CVVC	Stressed	Final	7
	Multisyllabic	CVCC	Stressed	Nonfinal	1
		CVVC	Unstressed	Final	6
		CVVC	Stressed	Final	1
59–65	Monosyllabic	CVCC	Stressed	Nonfinal	10
		CVVC	Stressed	Final	6
		CCVVC	Stressed	Nonfinal	4
	Disyllabic	CVCC	Stressed	Final	6
		CVVC	Stressed	Nonfinal	1
		CVVC	Stressed	Final	15
	Multisyllabic	CVVC	Stressed	Nonfinal	1
		CVVC	Unstressed	Final	2
66–72	Monosyllabic	CVVC	Stressed	Final	6
		CVCC	Stressed	Final	7
	Disyllabic	CVVC	Stressed	Final	8

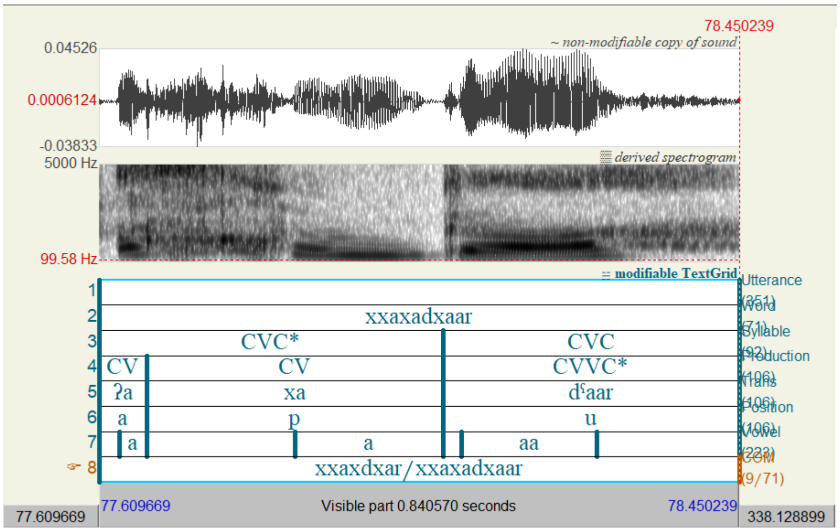
		CVCC	Stressed	Final	2
		CVVC	Unstressed	Final	1
	Multisyllabic	CVVC	Stressed	Final	8
		CVVC	Unstressed	Final	2
		CVVC	Stressed	Nonfinal	1
Adult	Monosyllabic	CVVC	Stressed	Final	1
		CVCC	Stressed	Final	17
		CVVC	Stressed	Final	15
	Disyllabic	CCVVC	Stressed	Final	9
		CVCC	Stressed	Final	6
		CVVC	Unstressed	Nonfinal	2
		CVVC	Unstressed	Final	3
		CVVC	Stressed	Nonfinal	4
		CVVC	Stressed	Final	17
	Multisyllabic	CCVVC	Stressed	Nonfinal	2
		CVVC	Unstressed	Final	5
		CVVC	Stressed	Nonfinal	2
		CVVC	Stressed	Final	5

Age group 24–30

Superheavy syllables in monosyllabic words had the highest frequency (n = 31, 81.6%), followed by multisyllabic words (n = 4, 10%), then disyllabic words (n = 3, 7.9%). In monosyllabic words, two sub-structures were produced, CVCC (n = 14, 45.2%) and CVVC (n = 17, 54.8%), with CVVC being more frequent. In disyllabic words, only CVVC syllables appeared in the stressed final environment. Figure 4.34/a demonstrates the production of a stressed final CVVC syllable in the word /ʕabiir/ CV.CVVC (‘Abeer’ a name, Speaker 1IG). As for superheavy syllables in multisyllabic words, only stressed final CVVC syllables were attested. No productions of CCVVC were observed in this age group. Figure 4.34/b demonstrates the production of a stressed final CVVC syllable in the word /ʔaxadˤaar/ (‘green’, CV.CV.CVVC, Speaker YRO).



(a) /ʕabiir/ Speaker 1IG

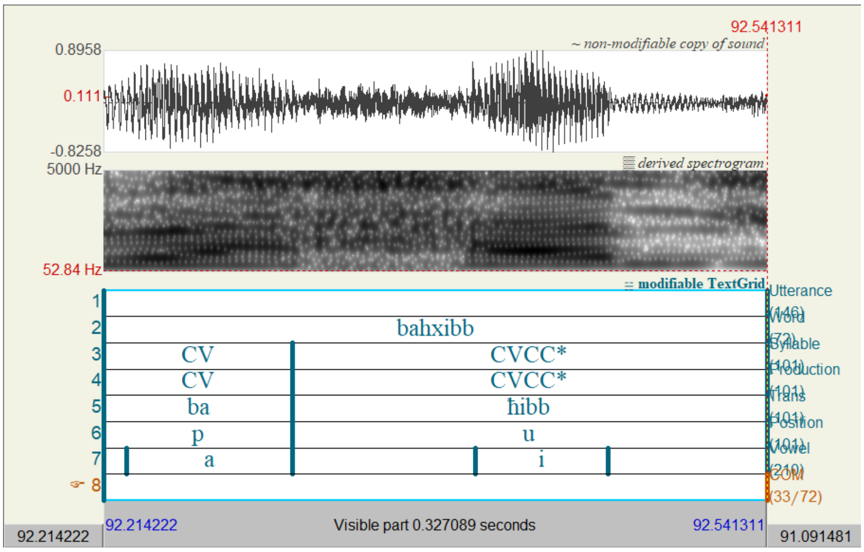


(b) /ʔaxadˤaar/ Speaker YRO

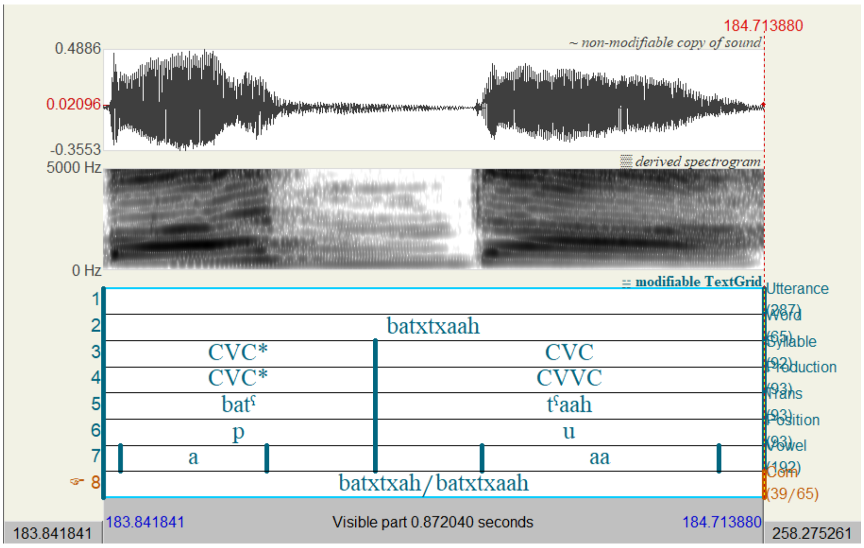
Figure 4.34: Spectrograms of superheavy syllables in age group 24–30

Age group 31–37

For 31–37 months, participants produced more superheavy syllables in monosyllabic words (n = 19, 67.9%) compared to disyllabic words (n = 9, 32.1%). For monosyllabic words, two sub-structures, CVVC (n = 10, 52.6%) and CVCC (n = 9, 47.4%), were observed at a similar frequency. As for disyllabic words, two sub-structures were observed, CVVC (n = 7, 77.8%), which was maintained from the youngest age group, and CVCC (n = 2, 22.2%), which emerged in this group. Figure 4.35/a demonstrates the production of a stressed final CVCC syllable in the word /baħibb/ (‘I love’, CV.CVCC, Speaker DDK). No productions of the CCVVC structure were attested. An increase in the frequency of superheavy syllables in disyllabic words occurred compared to the younger age group, where the frequency increased from 7.9% to 32.1%. In addition, unstressed final CVVC syllables appeared; for example, Figure 4.35/b demonstrates the production of the word /batˤtˤaah/ (‘duck’, CVC.CVVC, Speaker 2GS). Superheavy syllables in multisyllabic words were not attested in this age group.



(a) /baħibb/ Speaker DDK



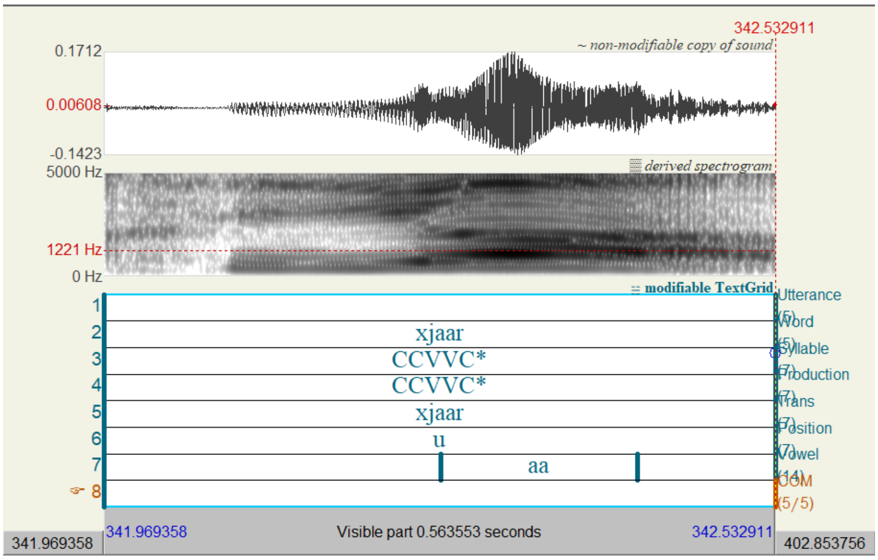
(b) /batˤtˤaah/ Speaker 2GS

Figure 4.35: Spectrograms of superheavy syllables in age group 31–37

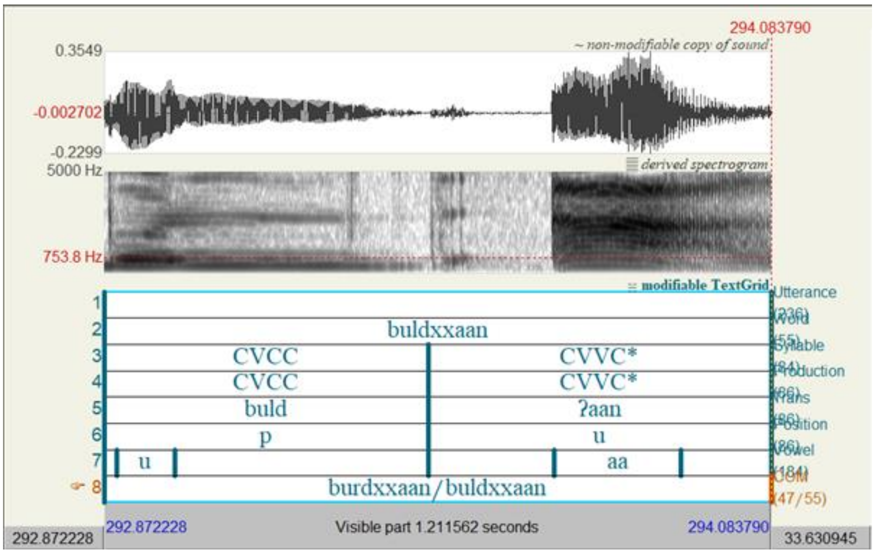
Age group 38–44

For this group, the majority of superheavy syllables were produced for monosyllabic words (n = 20, 52.6%). In addition to CVCC (n = 6, 30%) and CVVC (n = 11, 55%) that were observed in monosyllabic words for the younger age groups, the CCVVC (n = 3, 15%) structure emerged. Figure 4.36/a shows the CCVVC syllable in /xjaar/ (‘cucumber’, Speaker 6VI). Moreover, the number of superheavy syllables appearing in disyllabic (n = 13, 34.2%) and multisyllabic (n =

5, 13.2%) words has slightly increased compared to age group 24–30 months (no superheavy syllables in multisyllabic words in age group 31–37 months were observed). For disyllabic words, CVCC (n = 2, 15.4%) and CVVC (n = 11, 84.6) syllables were produced, while for multisyllabic words, only the CVVC structure was produced. Similar to age group 31–37 months, one token of the CVVC syllable was produced in an unstressed final environment for disyllabic words. However, an expansion in the superheavy syllable production was noted. First, stressed non-final CVVC syllables in multisyllabic words were observed. Second, non-final unstressed CVCC syllables in disyllabic words were observed. Figure 4.36/b demonstrates the production of an unstressed non-final CVCC syllable in the word /buldʔaan/ (‘orange’, CVCC.CVVC, Speaker DLE).



(a) /xjaar/ Speaker 6VI

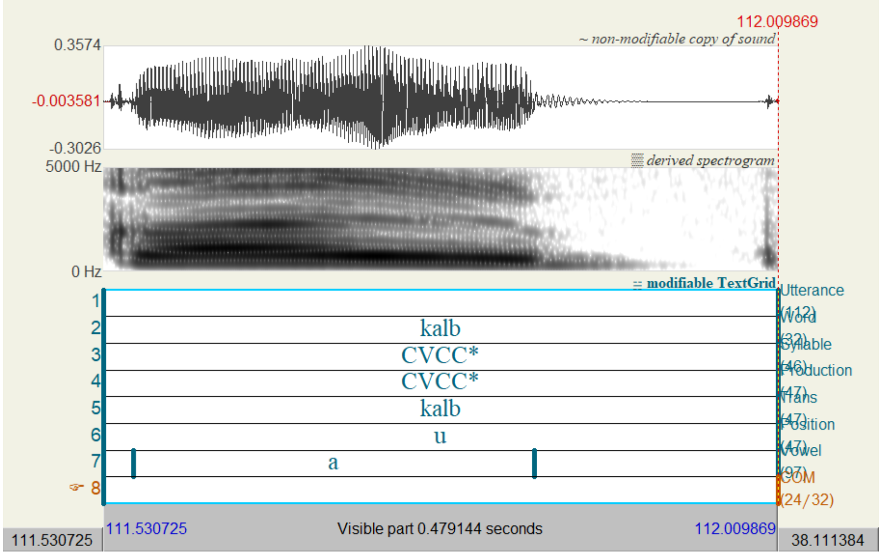


(b) /buldʔaan/ Speaker DLE

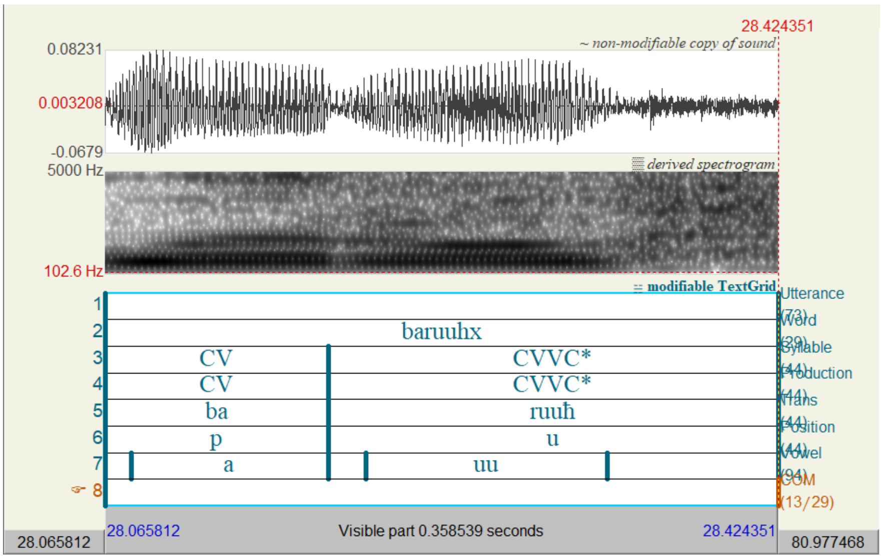
Figure 4.36:Spectograms of superheavy syllables in age group 38–44

Age group 45–51

For the age group 45–51 months, the frequency of superheavy syllables in monosyllabic words (n = 17, 81%) was higher than in disyllabic words (n = 4, 19%). Two sub-structures appeared in monosyllabic words including CVCC (n = 9, 52.9%) and CVVC (n = 8, 47.1%) syllables. Figure 4.37/a shows a stressed final CVCC syllable in the word /kalb/ (‘dog’, CVCC, Speaker G9I). As for disyllabic words, only CVCC (n = 2, 50%) and CVVC (n = 2, 50%) structures emerged in stressed final environments. Figure 4.37/b demonstrates a stressed final CVVC syllable in the word /baruuħ/ (‘I am going’, CV.CVVC, Speaker P0X). CCVVC syllables were not attested in this group, and no superheavy syllables in multisyllabic words were recorded.



(a) /kalb/ Speaker G9I

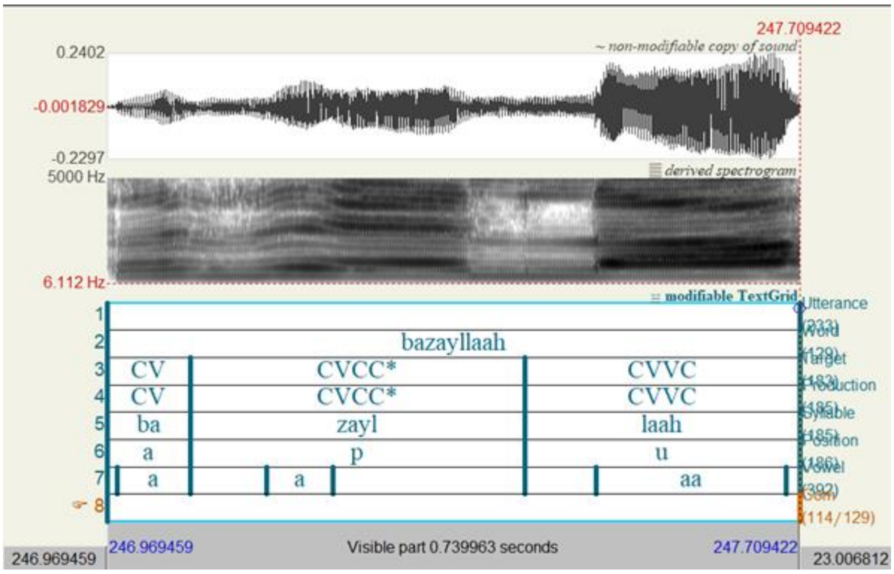


(b) /baruuħ/ Speaker P0X

Figure 4.37:Spectrograms of superheavy syllables in age group 45–51

Age group 52–58

At 52–58 months, an increase in the frequency of multisyllabic tokens (n = 8, 16.3%) was observed. Superheavy syllables in monosyllabic words were produced in three sub-structures, CVCC (n = 14, 48.3%), CVVC (n = 10, 34.5%), and CCVVC (n = 5, 17.2%). An expansion in the number of superheavy productions in the CCVVC structure appearing in monosyllabic words was observed in this age group. In disyllabic words, superheavy syllables were attested in two sub-structures, CVCC (n = 4, 33.3%) and CVVC (n = 8, 66.7%), with no production of CCVVC syllables. Unstressed final CVVC structures appeared in disyllabic words. Progressively, the disyllabic occurrences of superheavy syllables in the non–core CVCC structures increased for this age group. As for multisyllabic words, only CVCC (n = 1, 12.5%) and CVVC (n = 7, 87.5%) sub-structures were observed. Stressed non–final CVCC structures in multisyllabic words emerged in this age group. For example, Figure 4.38*Figure 4.38* demonstrates the production of a stressed non–final CVCC syllable in the word /bazajllaah/ ('beans', CV.CVCC.CVVC, Speaker MD5). There was a notable increase in the number of unstressed final CVVC syllables in multisyllabic words.

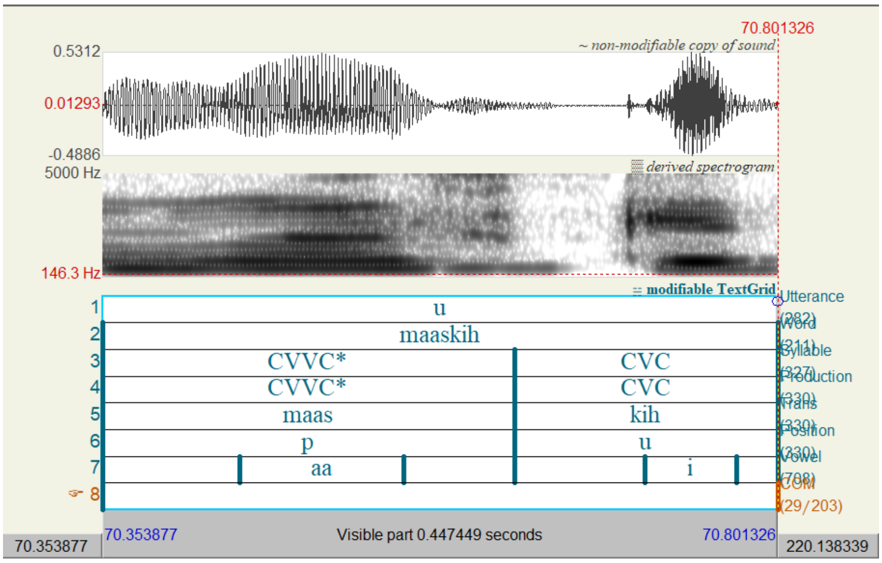


/bazajllaah/ Speaker MD5

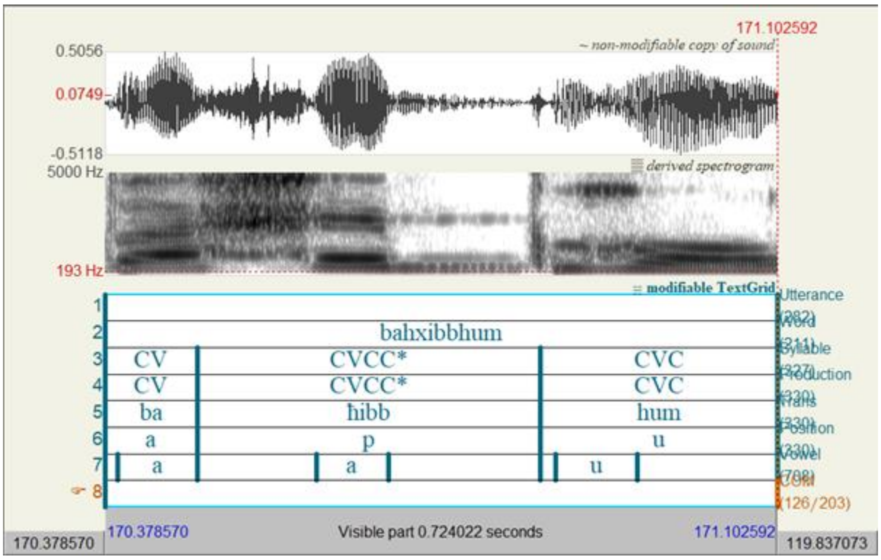
Figure 4.38:Spectrogram of a superheavy syllable in age group 52–58

Age group 59–65

At 59–65 months, three sub-structures were observed of superheavy syllables in monosyllabic words, CVCC (n = 10, 50%), CVVC (n = 6, 30%), and CCVVC (n = 4, 20%). As for disyllabic words, two sub-structures were recorded, CVCC (n = 6, 27.3%) and CVVC (n = 16, 72.7%) syllables, with no production of CCVVC. An expansion in the number of superheavy syllables produced in disyllabic words was attested (n = 22, 43.1%) compared to the younger age groups that produced more superheavy syllables in monosyllabic words. Stressed non–final CVVC syllables in disyllabic words were reported. Figure 4.39/a demonstrates the production of a stressed non–final CVVC syllable in the word /maaskih/ (‘holding’, CVVC.CVC, Speaker 6V9). In multisyllabic words, CVCC (n = 1, 11.1%) and CVVC (n = 8, 88.9%) syllables were observed, with CVVC being more frequent. Figure 4.39/b demonstrates the production of a stressed non–final CVCC syllable in the word /baħibbhum/ (‘I love them’, CV.CVCC.CVC, Speaker 6V9). Stressed non–final CVCC syllables and unstressed final CVVC syllables were observed.



(a) /maaskih/ Speaker 6V9



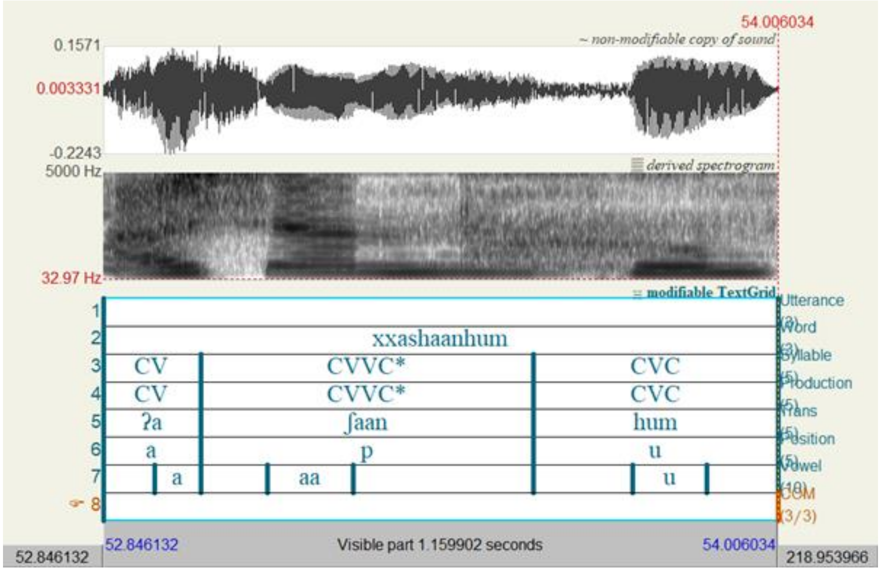
(b) /baħibbhum/ Speaker 6V9

Figure 4.39: Spectrograms of superheavy syllables in age group 59–65

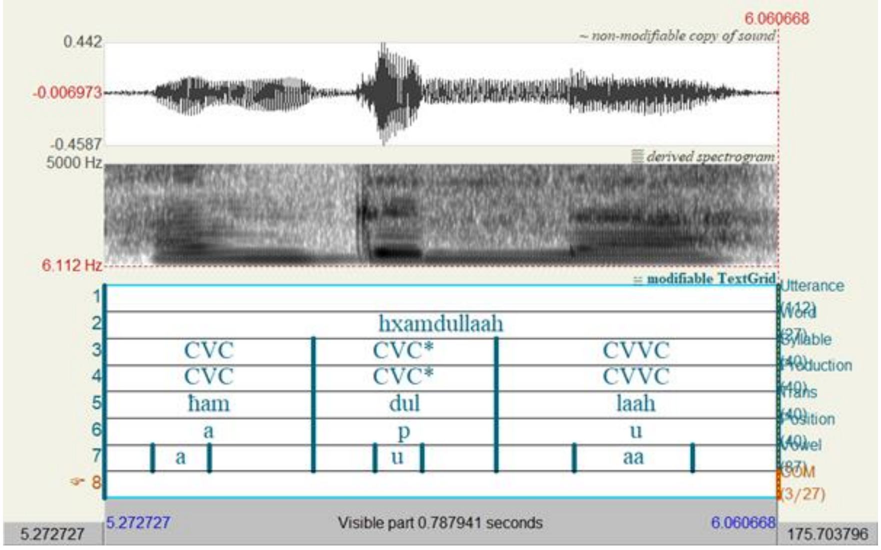
Age group 66–72

By 66–72 months, superheavy syllables appeared across all word lengths, with monosyllabic tokens being the most frequent (n = 15, 50%), followed by disyllabic (n = 11, 36.7%), then multisyllabic (n = 4, 13.3%). For monosyllabic words, superheavy syllables appeared in two sub-structures, CVCC (n = 7, 46.7%) and CVVC (n = 8, 53.3%). As for disyllabic words, superheavy syllables were attested in CVCC (n = 2, 18.2%) and CVVC (n = 9, 81.8%)

structures. Participants produced unstressed final CVVC syllables in disyllabic words. Moreover, all superheavy syllables in multisyllabic words were CVVC syllables. Varieties of superheavy syllable productions in multisyllabic words, such as stressed non–final CVVC syllables and unstressed final CVVC syllables, were observed. Figure 4.40*Figure 4.40*/a demonstrates the production of a stressed non–final CVVC syllable in the word /ʕaʃaanhum/ (‘for them’, CV.CVVC.CVC, Speaker QR6). Figure 4.40/b demonstrates the production of the word /ħamdullaah/ (‘thank God’, CVC.CVC.CVVC, Speaker 3Q1).



(a) /ʕaʃaanhum/ Speaker QR6

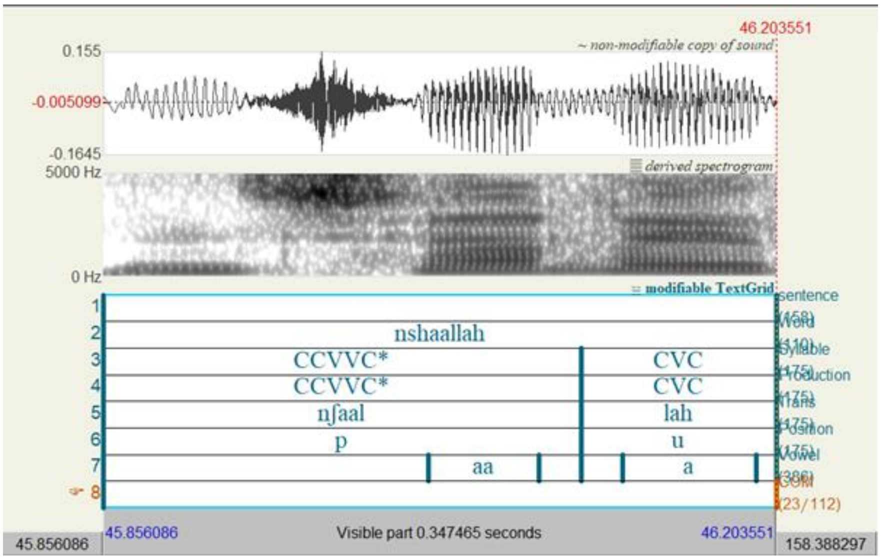


(b) /ħamdullaah/ Speaker 3Q1

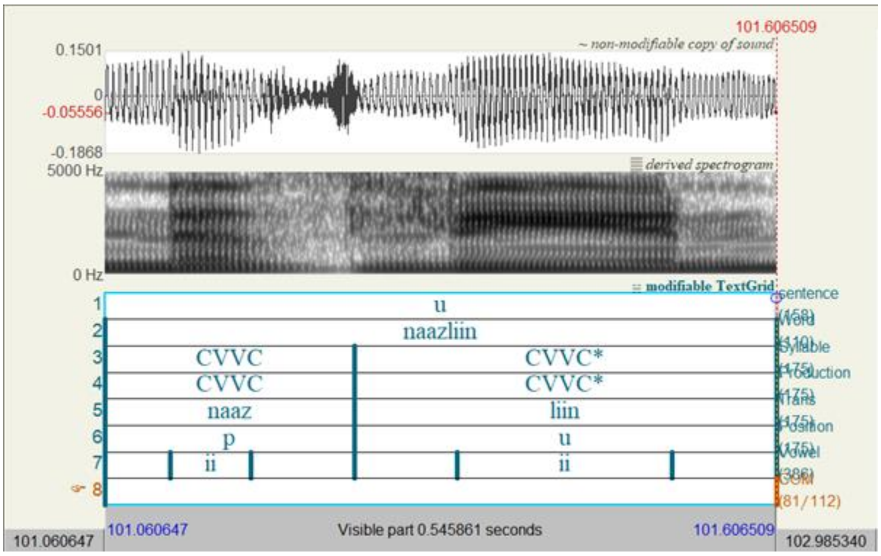
Figure 4.40:Spectograms of superheavy syllables in age group 66–72

Adults

In adult speech, superheavy syllables were produced more frequently in monosyllabic words (n = 41, 47.1%) than in disyllabic words (n = 34, 39.1%) or multisyllabic words (n = 12, 13.8%). Three sub-structures occurred in monosyllabic words, including CVCC, CVVC, and CCVVC, with CVCC (n = 17, 41.5%) and CVVC (n = 15, 36.6%) syllables being more frequent. Superheavy syllables were attested in three sub-structures in disyllabic words, CVCC, CVVC, and CCVVC, with CVVC (n = 26, 76.5%) syllables being the most frequent. Multiple linguistic structures were observed in superheavy disyllabic productions, such as stressed non–final CVVC syllables, unstressed final CVVC syllables, unstressed non–final CVVC syllables, and stressed non–final CCVVC syllables. Figure 4.41/a demonstrates a stressed non–final CCVVC syllable in the word /nʃaallah/ (‘if God wills’, CCVVC.CVC, Speaker BMG). Figure 4.41/b demonstrates an unstressed non–final CVVC syllable in the word /naazliin/ (‘heading downstairs’, CVVC.CVVC, Speaker BMG). Additionally, only CVVC syllables were observed in multisyllabic words. Stressed non–final CVVC syllables and unstressed final CVVC syllables in multisyllabic words were reported. Adults produced more varieties of superheavy syllables compared to the child groups, i.e., the number of productions for sub-structures and word lengths, indicating a linguistic complexity that advances with age.



(a) /nʃaallah/ Speaker BMG



(b) /naazliin/ Speaker BMG

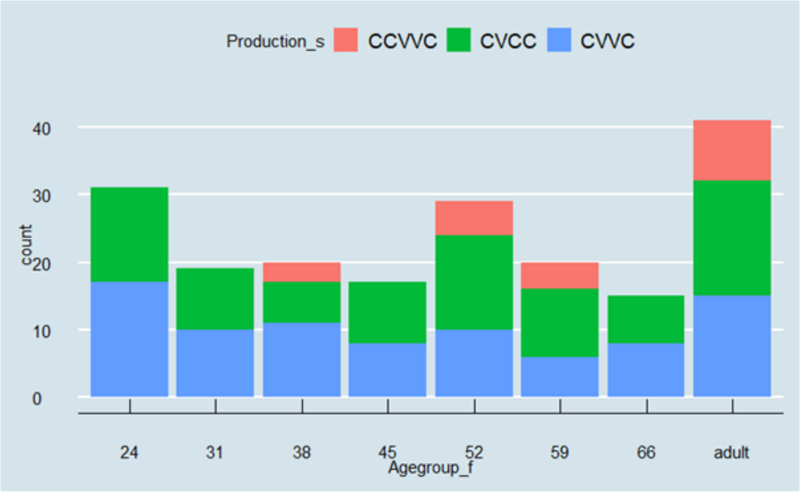
Figure 4.41: Spectrograms of superheavy syllables in the adult group

4.4.1.2.Bayesian Models and Descriptive Statistics

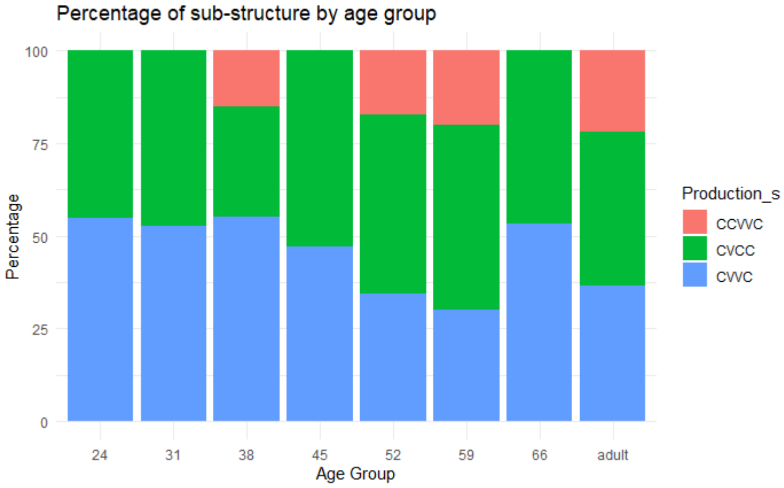
First, this section summarizes the frequency distributions, Bayesian model outputs, and data distribution means for syllable and vowel durations for superheavy syllable productions across the word lengths.

Superheavy Syllables in Monosyllabic Words

Figure 4.42/a shows the frequency distribution of superheavy sub-structures in monosyllabic words across the age groups, ranging from 24–30 months to the adult group. Figure 4.42/b shows the percentages of sub-structure production across the age groups, with CVVC and CVCC syllables being the most frequent.



(a) Frequency



(b) Percentage

Figure 4.42: Frequency and percentage of superheavy syllables in monosyllabic words

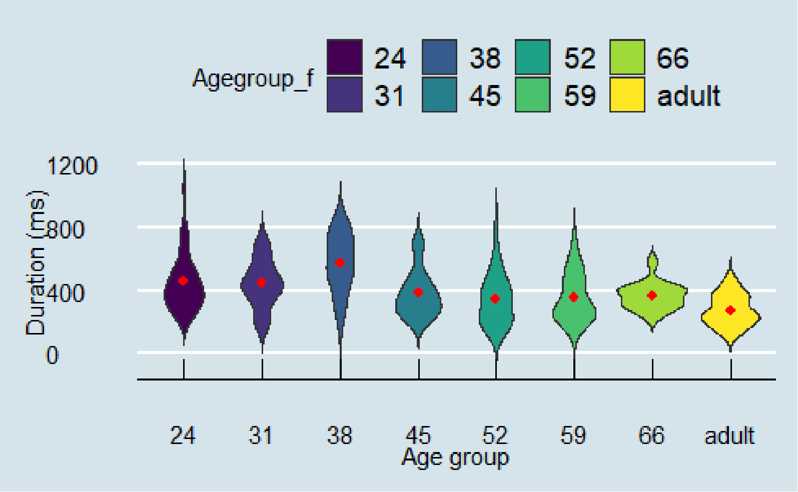
For superheavy syllables in monosyllabic words, a total of 192 syllables were analysed. For the syllable duration model, the Bayesian analysis was performed using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Since superheavy syllables in monosyllabic words only appear in stressed final environments, the model was fitted only to determine the effect of age group and sub-structure on durations. Table 4.21 summarizes the Bayesian model output for superheavy syllables in monosyllabic words.

Table 4.21: Bayesian model output summary for syllable duration in monosyllabic words

Population–Level Effects	**Estimate**	**Est. Error**	**l–95% CI**	**u–95% CI**	**Bulk_ESS**	**Tail_ESS**
Intercept	506.55	41.22	427.44	590.42	12705	17479
Sub-structure (CVVC)	–35.77	46.20	–129.75	52.92	13330	17063
Sub-structure (CCVVC)	231.38	148.04	–8.11	574.52	12203	10392
Age group	–28.94	8.70	–46.34	–11.88	11892	16569
Age group: Sub-structure (CVVC)	12.46	9.07	–5.10	30.34	15093	20938
Age group: Sub-structure (CCVVC)	–24.84	24.36	–79.53	16.97	12484	10421

Age Group

The model output indicates that age group is a strong predictor for superheavy syllable duration (β = –28.94, CI [–46.34, –11.88]). Figure 4.43 shows that mean syllable durations are the longest in the three youngest age groups, including 24–30 months (Mean = 454.3 ms, SD = 184.4), 31–37 months (Mean = 445.4 ms, SD = 152.2), and 38–44 months (Mean = 576.6 ms, SD = 184.3). Durations show a decreasing trend by 45–51 months of age (Mean = 381.3 ms, SD = 146.9), and adults produced the shortest mean duration (Mean = 277 ms, SD = 97.9).

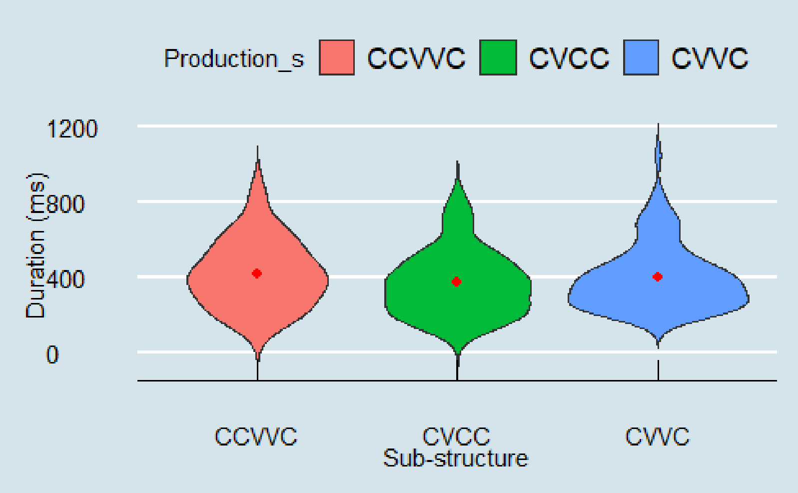


Syllable duration

Figure 4.43: The distribution and the mean syllable duration (ms) for age group in monosyllabic words

Sub-Structure

Figure 4.44 demonstrates that CCVVC syllables are the longest (Mean = 419.4 ms, SD = 174.0), followed by CVVC syllables (Mean = 399.7 ms, SD = 171.9), and CVCC syllables are the shortest (Mean = 370.7 ms, SD = 169.5). However, the model output indicates that sub-structure is not a strong predictor for superheavy syllable duration (CVVC, β = –35.77, CI [–129.75, 52.92]), (CCVVC, β = 231.38, CI [–8.11, 574.52]).



Syllable duration

Figure 4.44: The distribution and the mean syllable duration (ms) for sub-structure in monosyllabic words

Interaction between Age Group and Sub-Structure

The model output indicates that the two-way interaction between age group and sub-structure is not a strong predictor for superheavy syllable duration (Age-CVVC, β = 12.46, CI [–5.10, 30.34]), (Age-CCVVC, β = –24.84, CI [–79.53, 16.97]). CVCC and CVVC exhibit a negligible difference in the median syllabic durations, whereas the duration of CCVVC syllables becomes shorter with age. With age, the difference between durations across the sub-structures becomes less evident.

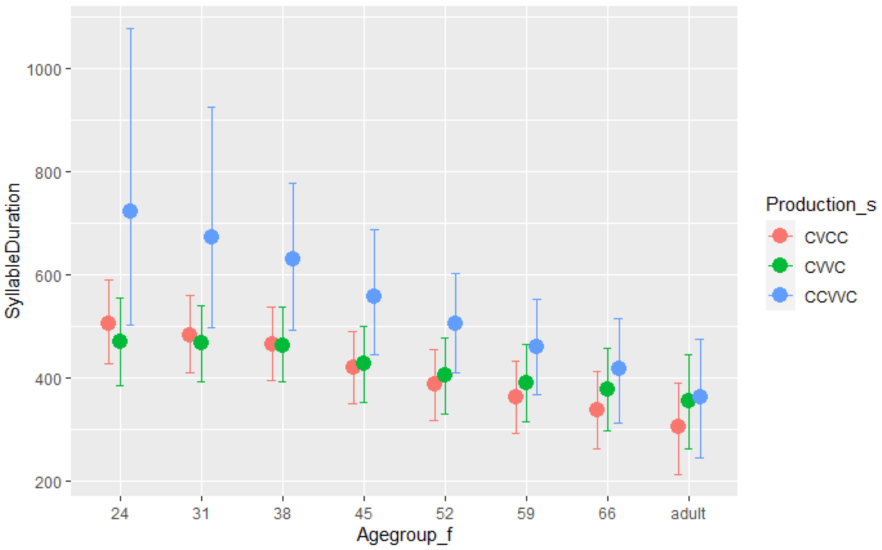


Figure 4.45: Posterior predictive plot for syllable duration of the interaction between sub-structure age group in monosyllabic words

Vowel Duration

As for the vowel duration model, the Bayesian analysis was performed on a total of 192 vowels using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Since superheavy syllables in monosyllabic words only appear in stressed final environments, the model was fitted only to determine the effect of age group, sub-structure, and the interaction between age group and sub-structure on syllable duration. Table 4.22 summarizes the Bayesian model output for superheavy syllables in monosyllabic words.

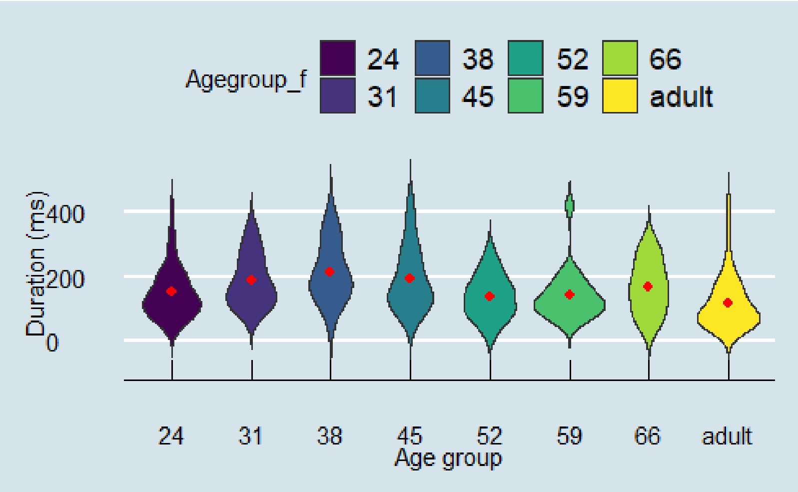
Table 4.22: Bayesian model output summary for vowel duration in monosyllabic words

Population–Level Effects	Estimate	Est. Error	l–95% CI	u–95% CI	Bulk_ESS	Tail_ESS
Intercept	137.74	20.11	97.83	177.48	9095	13798
Sub-structure (CVVC)	87.80	17.74	51.48	121.81	14144	19028
Sub-structure (CCVVC)	140.68	77.30	34.95	327.13	10571	6877
Age group	–6.07	4.78	–15.51	3.35	8203	12903
Age group: Sub-structure (CVVC)	0.75	3.90	–6.86	8.45	16493	21401

Age group: Sub-structure (CCVVC)	–16.83	12.44	–45.65	1.91	10845	6864

Age Group

Figure 4.46 shows that mean vowel durations increase from 152.6 ms (SD = 81.5) in age group 24–30 months to 216.4 ms (SD=92.9) in age group 38–44 months. At 45–51 months of age (Mean = 195.5 ms, SD = 98.2), a decline in the mean durations is observed, reaching the adult group mean (Mean = 119.6 ms, SD = 83.4). However, the model output indicates that age group is not a strong predictor for superheavy syllable vowel duration (β = –6.07, CI [–15.51, 3.35]). Although the durations decrease with age, approaching the adult group mean, they do not intersect, suggesting that children as old as 66–72 months do not produce adult-like vocalic durations in superheavy syllable productions.

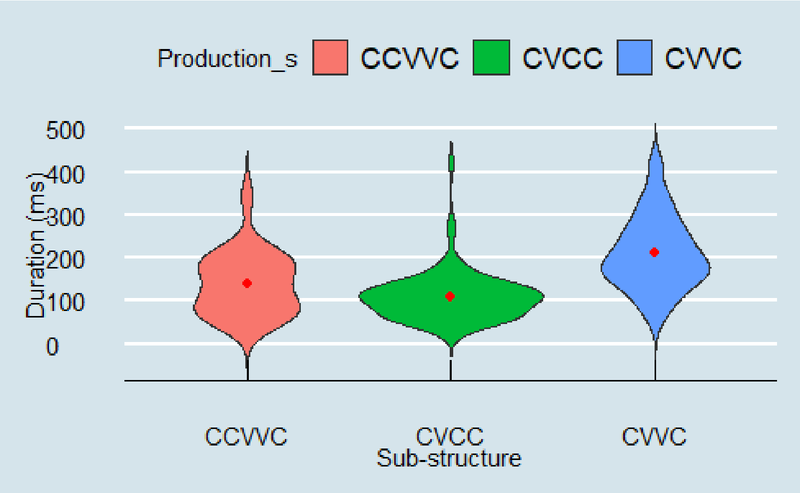


Vowel duration

Figure 4.46: The distribution and the mean vowel duration (ms) for age group in monosyllabic words

Sub-Structure

Figure 4.47 shows that vowels in CVVC syllables are the longest (Mean = 211.1 ms, SD = 84.6), followed by vowels in CCVVC syllables (Mean = 139.9 ms, SD = 72.5), and vowels in CVCC syllables are the shortest (Mean = 109.6 ms, SD = 57.5). On average, vowels in CVVC syllables are 1.5 times longer than vowels in CCVVC syllables and 1.9 times longer than vowels in CVCC syllables. Sub-structure is a strong predictor for vowel duration in superheavy syllables (CVVC, β = 87.80, CI [51.48, 121.81]), (CCVVC, β = 140.68, CI [34.95, 327.13]).



Vowel duration

Figure 4.47: The distribution and the mean vowel duration (ms) for sub-structure in monosyllabic words

Interaction between Age Group and Sub-Structure

The model output indicates that the interaction between age group and sub-structure is not a strong predictor for vowel duration (Age-CVVC, β = .75, CI [–6.86, 8.45), (Age-CCVVC, β = –16.83, CI [–45.65, 1.91]). In the youngest three age groups, 24–30, 31–37, and 38–44 months, vowels in CCVVC syllables are the longest, followed by vowels in CVVC syllables, while vowels in CVCC syllables are the shortest. By age group 45–51 months, vowels in CVVC syllables appear to be the longest, followed by vowels in CCVVC syllables, and vowels in CVCC remain the shortest. A noticeable decrease in vowel durations in CCVVC syllables is observed across the age groups.

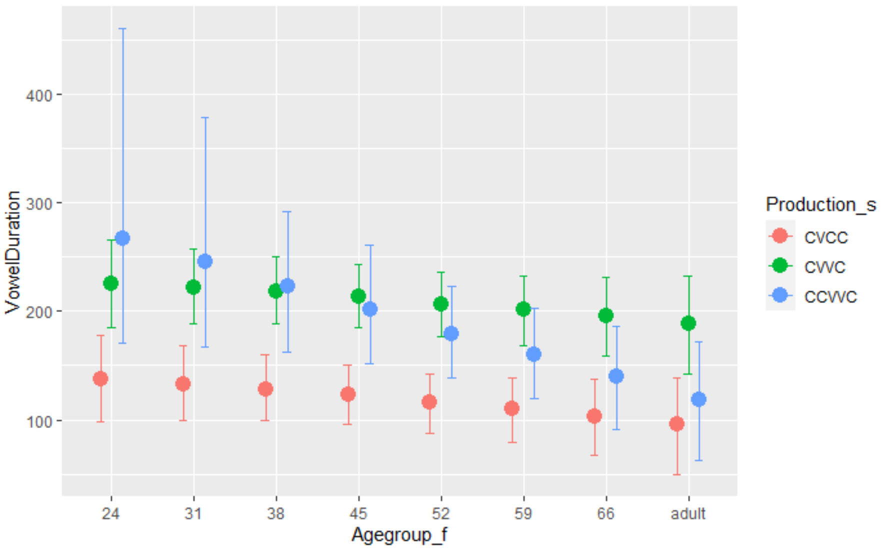
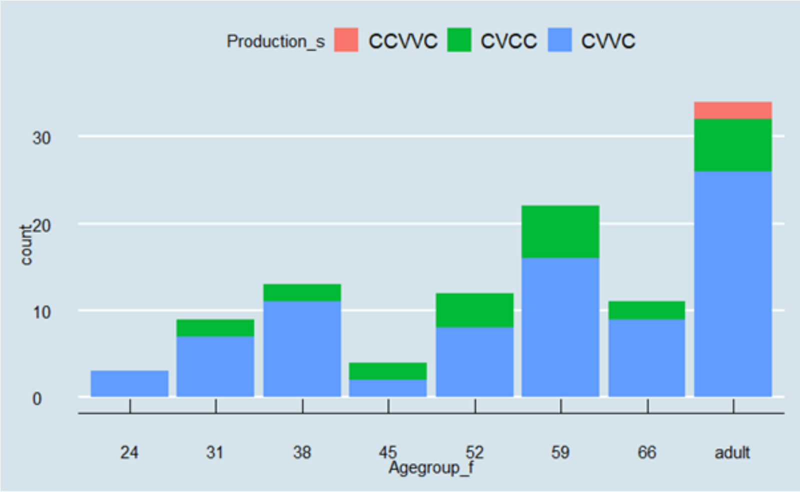


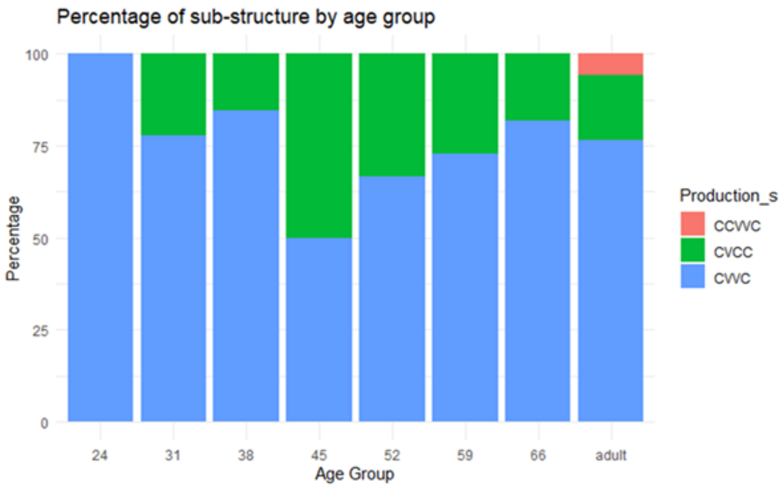
Figure 4.48: Posterior predictive plot for vowel duration of the interaction between sub-structure age group in monosyllabic words

Superheavy Syllables in Disyllabic and Multisyllabic Words

Figure 4.49 shows the frequency and percentage distributions of superheavy sub-structures in disyllabic words across the age groups, ranging from 24–30 months to the adult group. CVVC syllables are the most frequently produced syllables. The youngest age group, 24–30 months, only produced two sub-structures, whereas adults were the only group to produce all three sub-structures.



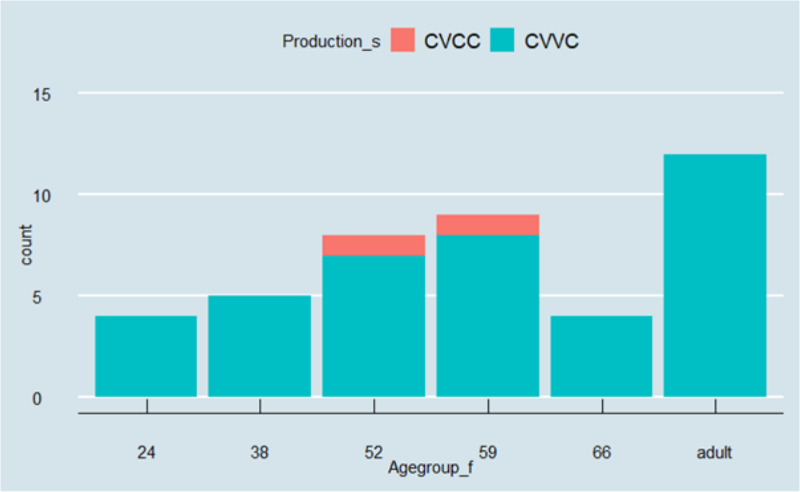
(a) Frequency



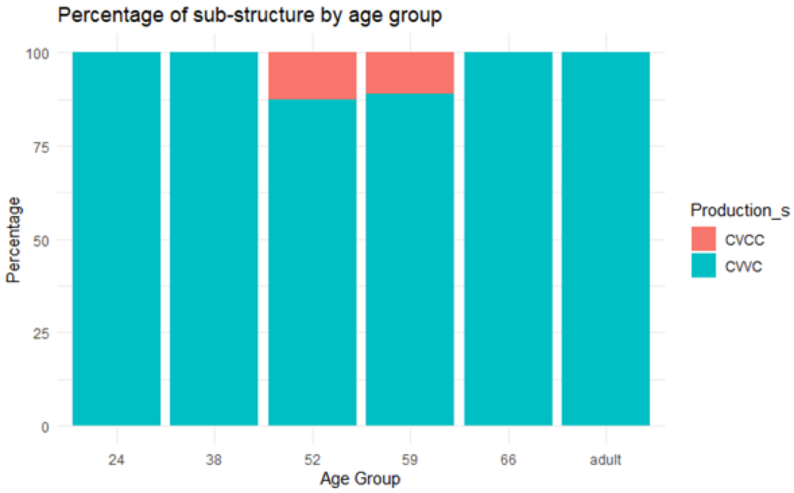
(b) Percentage

Figure 4.49: Frequency and percentage of superheavy syllables in disyllabic words

Figure 4.50 shows the frequency and percentage distributions of superheavy sub-structures in multisyllabic words across the age groups. Age groups 31–37 and 45–51 did not produce superheavy syllables in multisyllabic words. Also, only CVCC and CVVC syllables were produced, with CVVC syllables being the most frequent.



(a) Frequency



(b) Percentage

Figure 4.50: Frequency and percentage of superheavy syllables in multisyllabic words

Syllable Duration

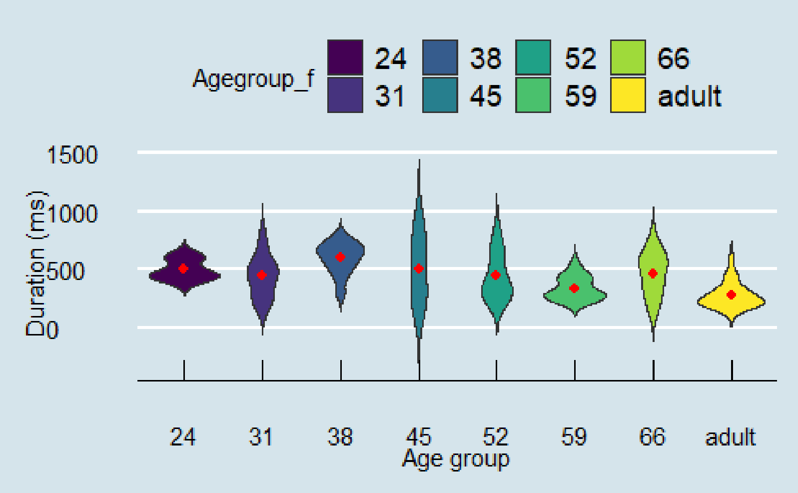
Given the small sample size (disyllabic, n = 108; multisyllabic, n = 42), attempts to fit separate models for superheavy syllables in disyllabic and multisyllabic words led to the emergence of multiple unresolved warning messages. Therefore, disyllabic and multisyllabic word data were combined, with a total of 150 syllables analysed. Age group, syllable position, stress, and sub-structure were assigned as independent variables. For the syllable duration model, the Bayesian analysis was performed using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Table 4.23 summarizes the Bayesian model output for superheavy syllables in disyllabic and multisyllabic words. On the other hand, to account for the syllable and vowel durations for superheavy syllables in disyllabic and multisyllabic words, the data distribution violin plots were created separately according to the word length.

Table 4.23: Bayesian model output summary for syllable duration in di/multi-syllabic words

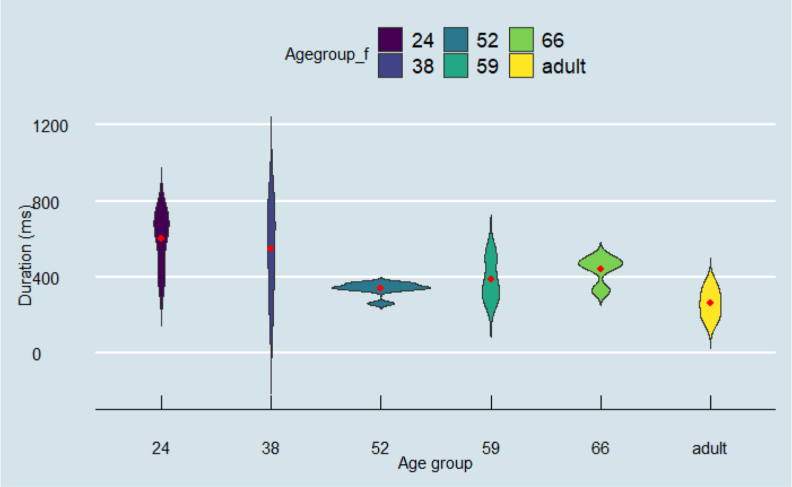
Population–Level Effects	Estimate	Est. Error	l–95% CI	u–95% CI	Bulk_ESS	Tail_ESS
Intercept	461.62	67.56	329.76	596.20	13472	18066
Sub-structure (CVVC)	–23.02	41.34	–105.73	56.70	24478	22811
Sub-structure (CCVVC)	–4.24	120.98	–244.95	231.24	26744	25214
Stress (Stressed)	25.30	32.37	–37.85	88.89	26472	24137
Position (Final)	104.95	37.79	31.26	179.19	10580	13988
Age group	–31.26	9.12	–48.79	–12.71	13472	18066

Age Group

The model output suggests that age group is a strong predictor for superheavy syllable duration in disyllabic and multisyllabic words ($\beta = -31.26$, CI [–48.79, –12.71]). For disyllabic words, Figure 4.51/a shows that mean syllable durations increase from age group 24–30 months to age group 45–51 months, then durations show an overall declining trend. The youngest age group has a mean duration of 501.8 ms (SD = 97), followed by a mean duration of 442.9 ms (SD = 177.7) in age group 31–37 months, while age group 38–44 months has a mean duration of 600.4 ms (SD = 135.3). By 45–51 months (Mean = 510.2 ms, SD = 291.2), durations fluctuate in a descending pattern reaching the adult group mean (Mean = 277.9 ms, SD = 109.2). For multisyllabic words, Figure 4.51/b shows that mean syllable durations are the longest in the youngest two age groups, but the durations tend to decrease with age. The youngest age group has a mean duration of 601.2 ms (SD = 153.1), followed by 549.8 ms (SD = 233.4) in age group 38–44 months. The adult group's mean duration is 262.3 ms (SD = 73.9).



(a) Superheavy syllables in disyllabic words

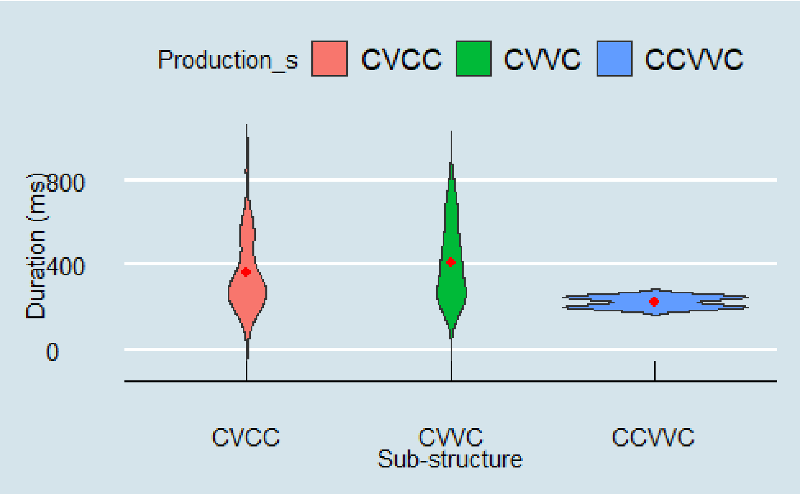


(b) Superheavy syllables in multisyllabic words

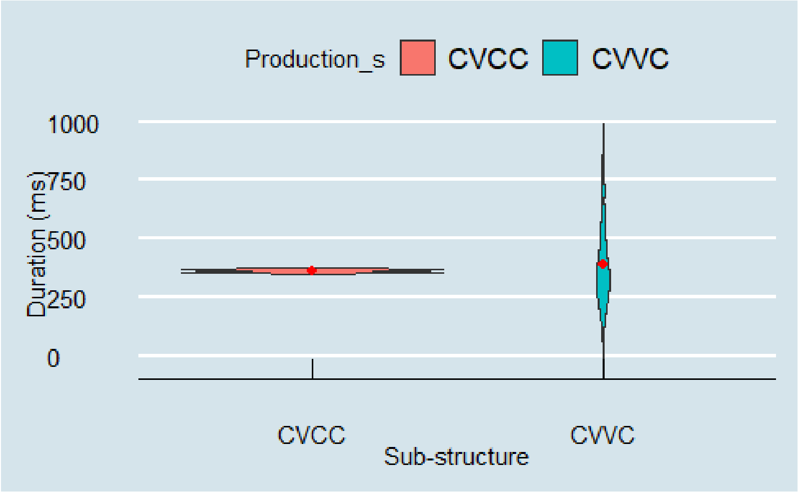
Figure 4.51: The distributions and the mean syllable durations (ms) for age group in disyllabic and multisyllabic words

Sub-Structure

Sub-structure is not a strong predictor for superheavy syllable duration in disyllabic and multisyllabic words (CVVC, β =–23.02, CI [–105.73, 56.70]), (CCVVC, β = –4.24, CI [–244.95, 231.24]). For disyllabic words, Figure 4.52/a demonstrates that CVVC syllables are the longest (Mean = 410.7 ms, SD = 178.5), followed by CVCC syllables (Mean = 223.1 ms, SD = 174.3), while CCVVC syllables are the shortest (Mean = 139.9 ms, SD = 33.3). Approximately, CVVC syllables are 1.8 times longer than CVCC syllables and 2.9 times longer than CCVVC syllables. For superheavy syllables in multisyllabic words, CVVC syllables (Mean = 389.3 ms, SD = 159.6) and CVCC syllables (Mean = 363.5 ms, SD = 8.5) have similar mean duration, as demonstrated in Figure 4.52/b. However, the distributions of CVVC and CVCC durations are different, where CVCC syllables have a relatively homogeneous distribution, indicated by the flat-shaped violin. On the other hand, CVVC syllables exhibit a broader range of values, indicating a higher variation in the data.



(a) Superheavy syllables in disyllabic words

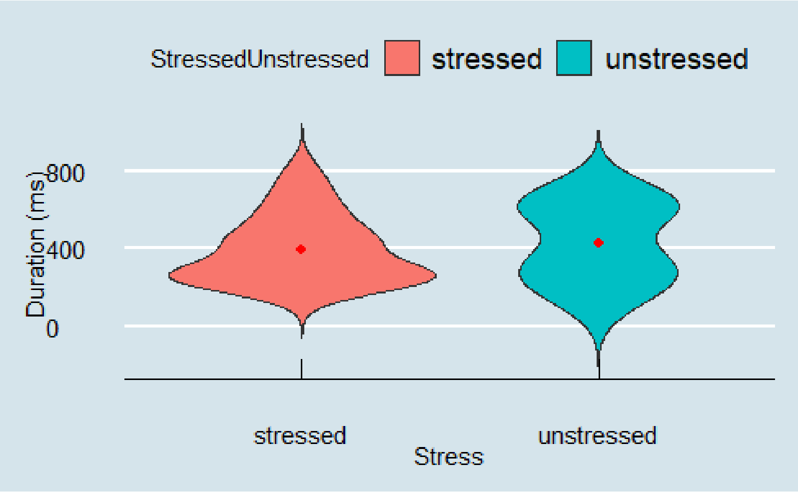


(b) Superheavy syllables in multisyllabic words

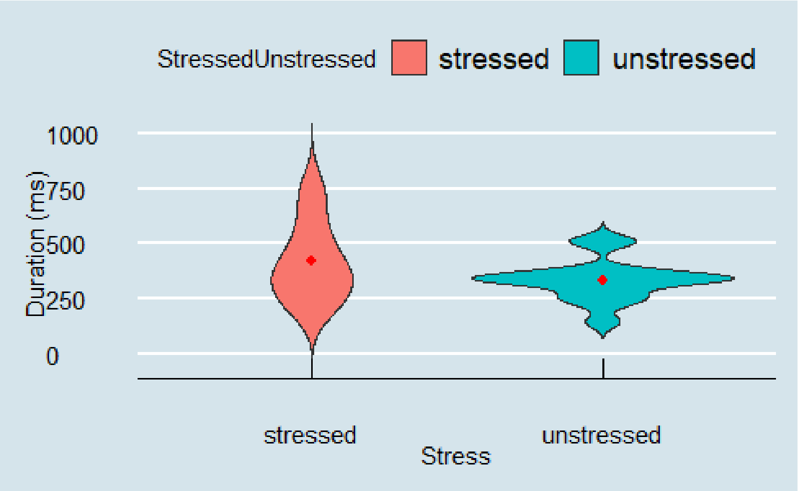
Figure 4.52: The distributions and the mean syllable durations (ms) for sub-structure in disyllabic and multisyllabic words

Stress

Moreover, stress assignment is not a strong predictor for superheavy syllable duration (β = 25.30, CI [–37.85, 88.89]). Figure 4.53/a demonstrates that stressed superheavy syllables in disyllabic words (Mean = 392.1 ms, SD = 176.0) are shorter than unstressed syllables (Mean = 426.8 ms, SD = 201.1). Approximately, unstressed syllables are 1.1 times longer than their counterparts. Alternatively, for superheavy syllables in multisyllabic words, Figure 4.53/b shows that stressed superheavy syllables (Mean = 175.8 ms, SD = 176.0) are longer than their counterparts (Mean = 93.9 ms, SD = 201.1). Approximately, stressed syllables are 1.9 times longer than unstressed ones.



(a) Superheavy syllables in disyllabic words



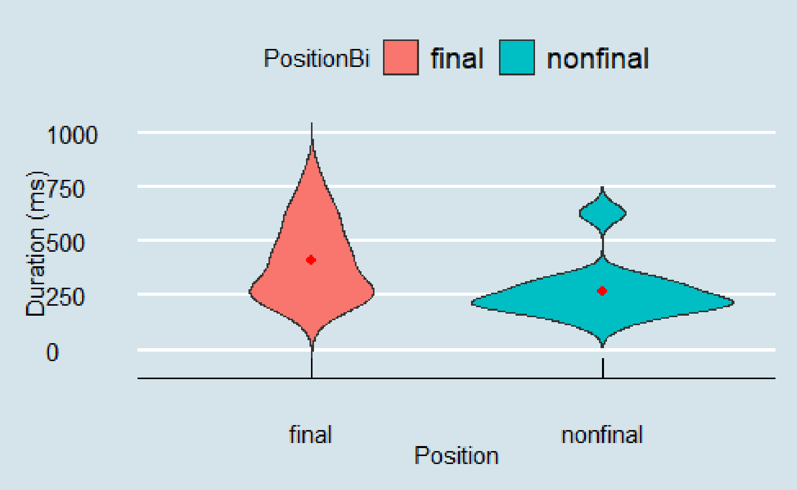
(b) Superheavy syllables in multisyllabic words

Figure 4.53: The distributions and the mean syllable durations (ms) for stress in disyllabic and multisyllabic words

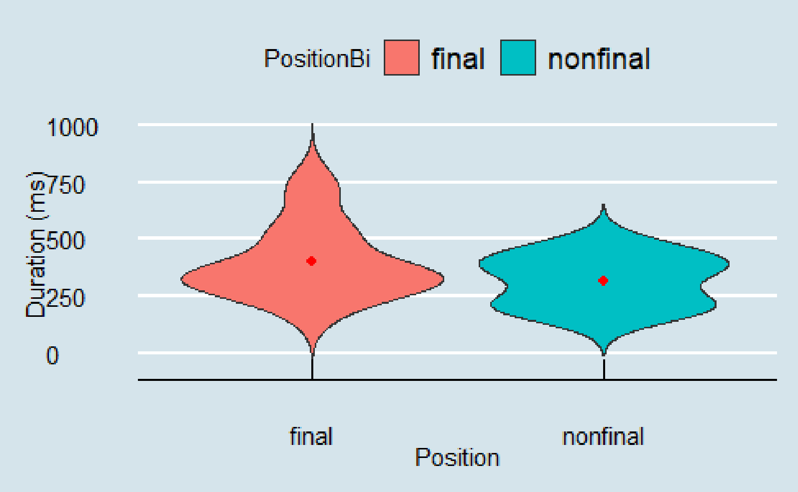
Syllable Position

The model output indicates that syllable position is a strong predictor for superheavy syllable duration (β = 104.95, CI [31.26, 179.19]). In disyllabic words, word-final lengthening is observed. Figure 4.54/a shows that syllables in the word-final position (Mean = 409.7 ms, SD = 177.5) are longer than their counterparts (Mean = 271.2 ms, SD = 132.5). On average, superheavy word-final syllables are 1.5 times longer than word non–final ones. Moreover, word final lengthening is also observed in superheavy syllables in multisyllabic words. Figure

4.54/b suggests that syllables in the word-final position (Mean = 402.2 ms, SD = 177.5) are longer than their counterparts (Mean = 317.2 ms, SD = 132.5). On average, word-final syllables are 1.3 times longer than non–final syllables.



(a) Superheavy syllables in disyllabic words



(b) Superheavy syllables in multisyllabic words

Figure 4.54: The distributions and the mean syllable durations (ms) for syllable position in disyllabic and multisyllabic words

Vowel Duration

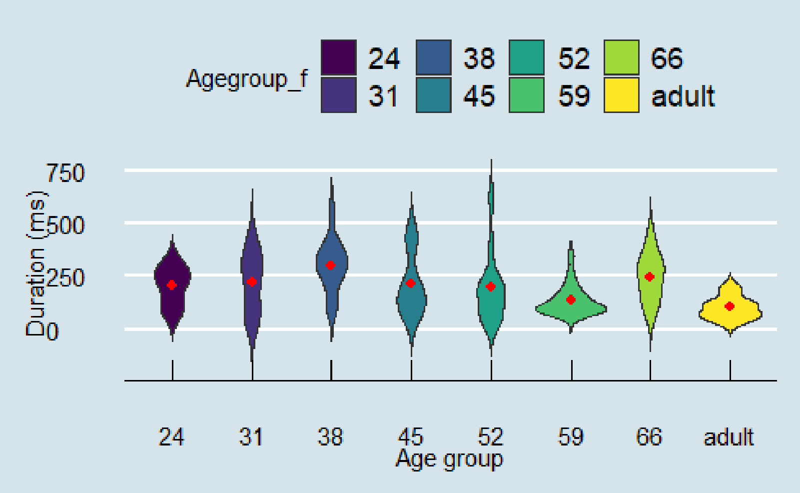
As for vowel durations, a total of 150 vowels were analysed. The Bayesian analysis was performed using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Table 4.24 summarizes the Bayesian model output for superheavy syllables in disyllabic and multisyllabic words.

Table 4.24: Bayesian model output summary for vowel duration in di/multisyllabic words

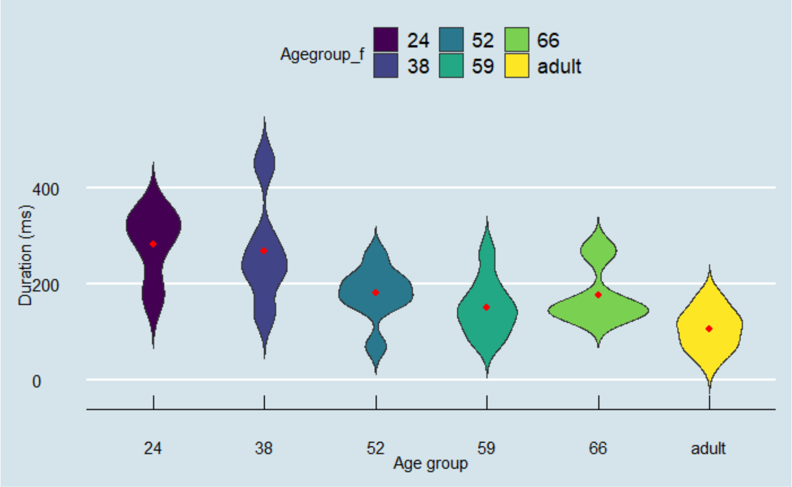
Population–Level Effects	Estimate	Est. Error	L–95% CI	U–95% CI	Bulk_ESS	Tail_ESS
Intercept	105.16	41.84	22.42	186.12	19320	20562
Sub-structure (CVVC)	88.69	22.46	43.09	132.00	23671	14718
Sub-structure (CCVVC)	85.15	69.47	–53.12	220.93	35039	23016
Stressed (Stressed)	20.00	20.44	–19.68	59.88	35371	24600
Position (Final)	71.19	23.89	24.43	117.97	38356	24417
Age group	–17.15	5.87	–28.23	–5.06	12497	15593

Age Group

The model output indicates that age group is a strong predictor for superheavy vowel duration in disyllabic and multisyllabic words ($\beta = -17.15$, CI [–28.23, –5.06]). For superheavy syllables in disyllabic words, Figure 4.55/a shows that mean vowel durations increase from 202.7 ms (SD = 99.1) in age group 24–30 months to 301.1 ms (SD = 124.9) in age group 38–44 months. At 45–51 months of age (Mean = 211.9 ms, SD = 156.6), a decline in the mean durations is observed, reaching the adult group mean (Mean = 102 ms, SD = 52.7). For multisyllabic words, Figure 4.55/b displays that mean vowel durations decrease with age from 283.4 ms (SD = 74.3) in age group 24–30 months to 106.9 ms (SD = 41.8) in the adult group. Although children produce shorter vowels with age, children as old as 66–72 months did not produce durations that intersect with adult productions.



(a) Superheavy syllables in disyllabic words

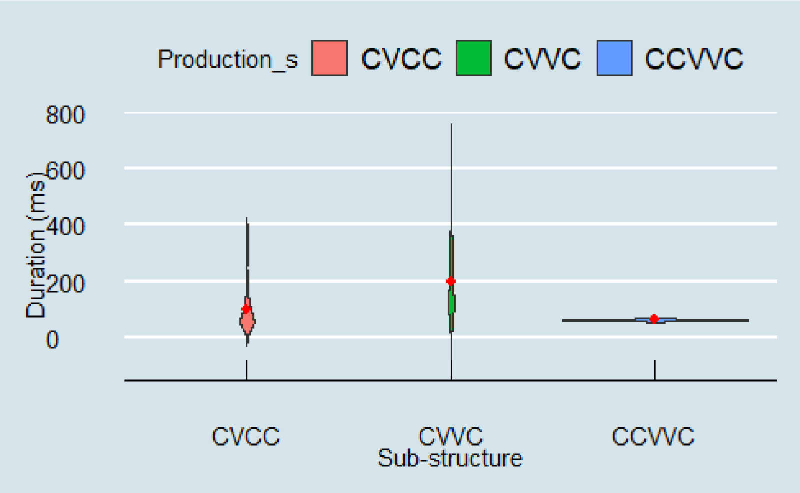


(b) Superheavy syllables in multisyllabic words

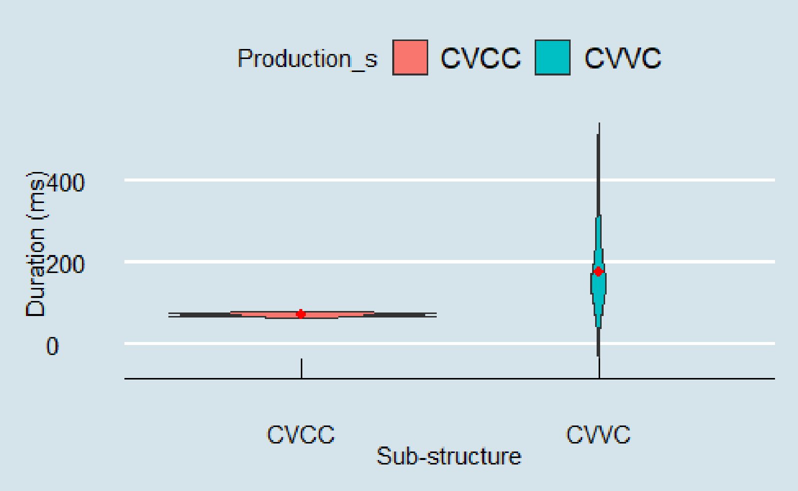
Figure 4.55: The distributions and the mean vowel durations (ms) for age group in disyllabic and multisyllabic words

Sub-Structure

Sub-structure is a strong predictor for superheavy vowel duration in disyllabic and multisyllabic words (CVVC, β = 88.69, CI [43.09, 132.00]), but not for (CCVVC, β = 85.15, CI [–53.12, 220.93]). For disyllabic words, Figure 4.56/a shows that vowels in CVVC syllables are the longest (Mean = 199.9 ms, SD = 121.9), then vowels in CVCC syllables (Mean = 99.8 ms, SD = 82.9), and vowels in CCVVC syllables are the shortest (Mean = 59.8 ms, SD = 3.9). Vowels in CVVC syllables are two times longer than vowels in CVCC syllables and 3.3 times longer than those in CCVVC syllables. For multisyllabic words, the effect of the sub-structure is observed in vowel durations. Figure 4.56/b shows that vowels in CVVC syllables (Mean = 178.4 ms, SD = 84.8) are longer than those in CVCC syllables (Mean = 74.2 ms, SD=4.7). On average, vowels in CVVC syllables are 2.4 times longer than vowels in CVCC syllables.



(a) Superheavy syllables in disyllabic words

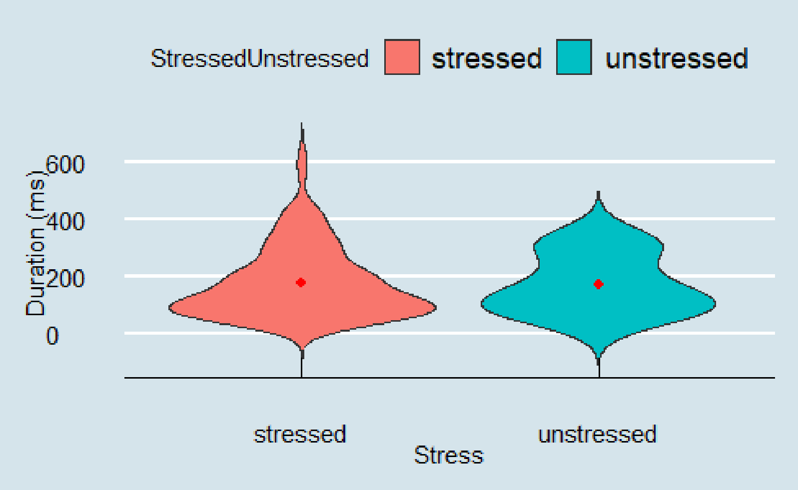


(b) Superheavy syllables in multisyllabic words

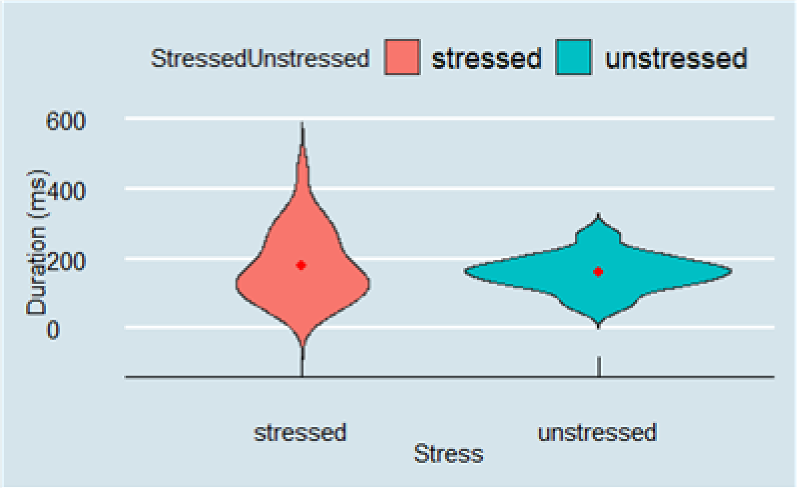
Figure 4.56: The distributions and the mean vowel durations (ms) for sub-structure in disyllabic and multisyllabic words

Stress

The model output indicates that stress assignment is not a strong predictor for vowel duration (β = 20.00, CI [–19.68, 59.88]). For disyllabic words in Figure 4.57/a, stress assignment does not seem to affect vowel duration of superheavy syllables. On average, vowels in stressed superheavy syllables (Mean = 175.3 ms, SD = 123.8) are marginally longer than their counterparts (Mean = 171.2 ms, SD = 102.6). In multisyllabic words, Figure 4.57/b shows vowels in stressed superheavy syllables (Mean = 180.5 ms, SD = 99.9) appear slightly longer than in unstressed syllables (Mean = 160.6 ms, SD = 52.1). Vowels in stressed syllables are only 1.1 times longer than vowels in unstressed syllables.



(a) Superheavy syllables in disyllabic words

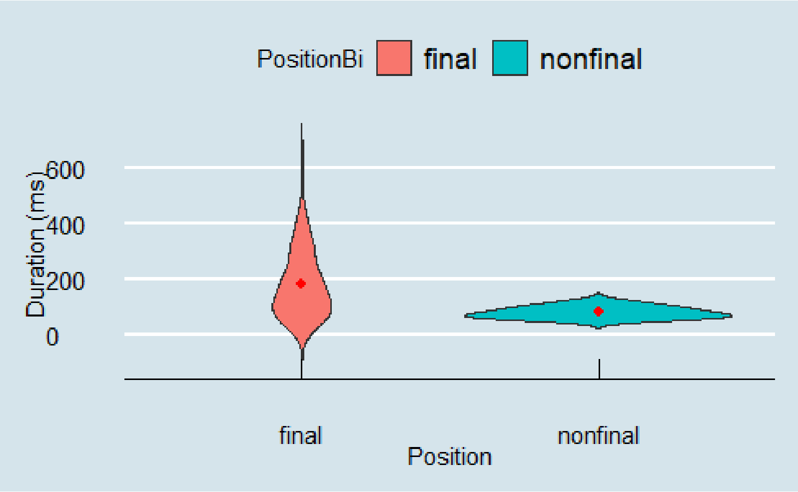


(b) Superheavy syllables in disyllabic words

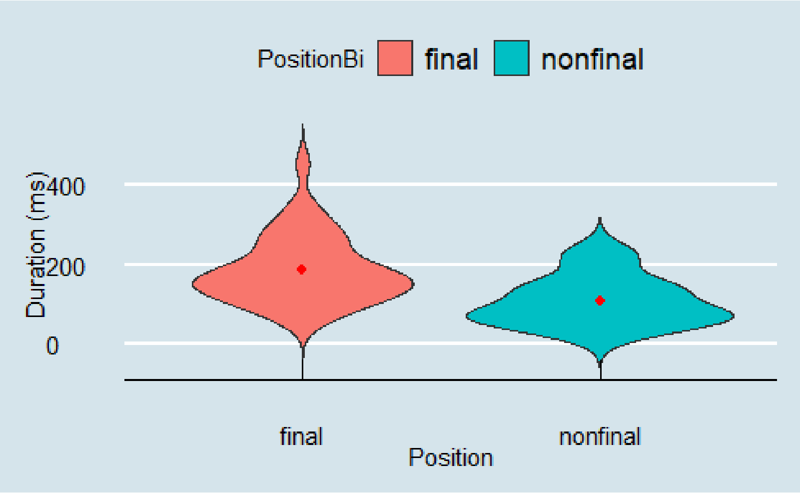
Figure 4.57: The distributions and the mean vowel durations (ms) for stress in disyllabic and multisyllabic words

Syllable Position

The model output suggests that syllable position is a strong predictor for vowel duration (β = 71.19, CI [–24.43, 117.97]). In Figure 4.58/a, word-final lengthening is observed for vowel durations for disyllabic words. Vowels in word-final syllables (Mean = 185.5 ms, SD = 123.6) are longer than vowels in word non–final syllables (Mean = 82.0 ms, SD = 21.0). On average, vowels in word-final syllables are 2.3 times longer than vowels in word non–final syllables. For multisyllabic words, Figure 4.58/b demonstrates that vowels in word-final syllables (Mean = 186.3 ms, SD = 84.5) are longer than vowels in word non–final syllables (Mean = 109.0 ms, SD = 62.9). On average, vowels in word-final syllables are 1.7 times longer than vowels in word non–final syllables.



(a) Superheavy syllables in disyllabic words



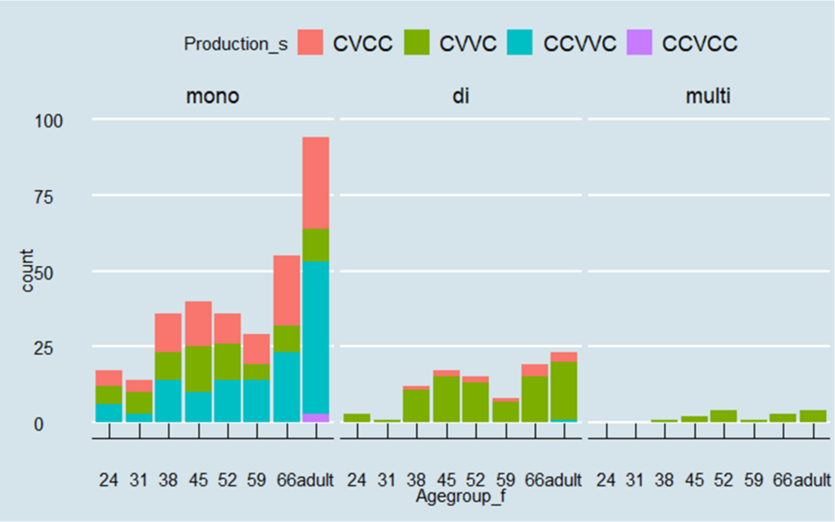
(b) Superheavy syllables in multisyllabic words

Figure 4.58: The distributions and the mean vowel durations (ms) for syllable position in disyllabic and multisyllabic words

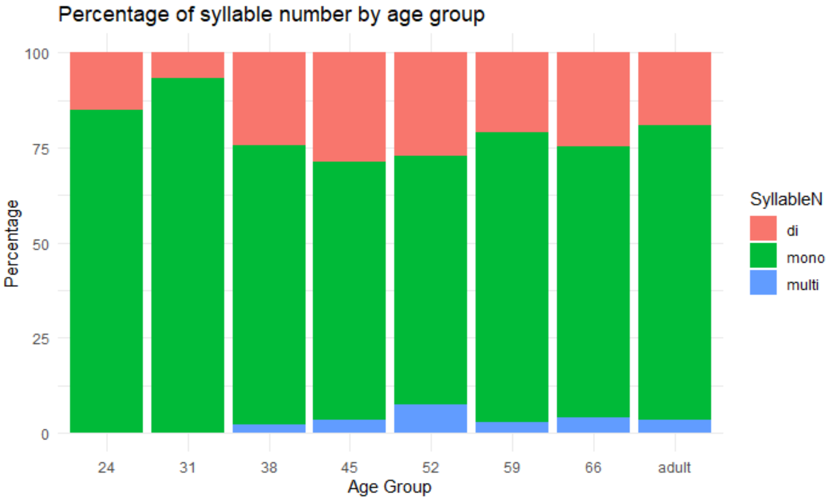
4.4.2. Repetition Task

4.4.2.1.Superheavy Syllable Frequency

Figure 4.59/a demonstrates the frequency of superheavy syllables in monosyllabic, disyllabic, and multisyllabic words across the age groups ranging from 24–30 months to the adult group. Superheavy syllables are the most frequent in monosyllabic words, followed by disyllabic words, and the least frequent in multisyllabic words. Furthermore, monosyllabic words exhibit a greater range of sub-structure varieties than disyllabic and multisyllabic word lengths. There is a general increasing trend of the count of superheavy syllables produced as the age group increases. Additionally, Figure 4.59/b shows that the syllabic complexity of superheavy syllables in monosyllabic words increases with age. The youngest two age groups, 24–30 months to 31–38 months, produce superheavy syllables only for monosyllabic and disyllabic words. However, by age group 38–44 months, children start producing superheavy syllables in multisyllabic words.



(a) Frequency



(b) Percentage

Figure 4.59: Frequency and percentage of superheavy syllables in RT

Also, Table 4.25 shows the frequency of superheavy syllables in JA child and adult productions. The table shows the age group (children: 24–72 months, adults), word length (monosyllabic, disyllabic, and multisyllabic), sub-structure (CVCC, CVVC, and CCVVC), stress assignment (stressed, unstressed), and syllable position (final, nonfinal).

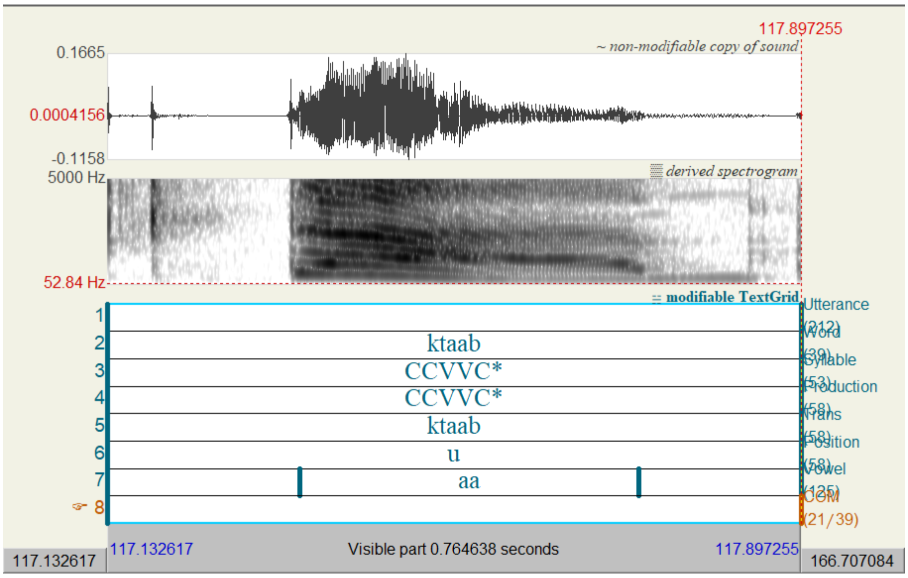
Table 4.25: Superheavy syllables RT

Age Group	**Word Length**	**Sub-Structure**	**Stress**	**Syllable Position**	**Frequency (n)**
24–30	Monosyllabic	CVCC	Stressed	Final	5
		CVVC	Stressed	Final	6
		CCVVC	Stressed	Final	6
	Disyllabic	CVVC	Stressed	Final	3
31–37	Monosyllabic	CVCC	Stressed	Final	4
		CVVC	Stressed	Final	7
		CCVVC	Stressed	Final	3
	Disyllabic	CVVC	Stressed	Final	1
38–44	Monosyllabic	CVCC	Stressed	Final	13
		CVVC	Stressed	Final	9
		CCVVC	Stressed	Final	14
	Disyllabic	CVCC	Stressed	Final	1
		CVVC	Stressed	Final	9
	Multisyllabic	CVVC	Unstressed	Final	1
	Multisyllabic	CVVC	Stressed	Final	1
45–51	Monosyllabic	CVCC	Stressed	Final	15
		CVVC	Stressed	Final	15
		CCVVC	Stressed	Final	10
	Disyllabic	CVCC	Stressed	Final	2
		CVVC	Stressed	Final	14
	Multisyllabic	CVVC	Stressed	Final	2
52–58	Monosyllabic	CVCC	Stressed	Final	10
		CVVC	Stressed	Final	12
		CCVVC	Stressed	Final	14
	Disyllabic	CVCC	Stressed	Final	2
		CVVC	Stressed	Final	13
	Multisyllabic	CVVC	Stressed	Final	4
59–65	Monosyllabic	CVCC	Stressed	Final	10
		CVVC	Stressed	Final	5
		CCVVC	Stressed	Final	14
	Disyllabic	CVCC	Stressed	Final	1

		CVVC	Stressed	Final	7
	Multisyllabic	CVVC	Stressed	Final	1
66–72	Monosyllabic	CVCC	Stressed	Final	23
		CVVC	Stressed	Final	9
		CCVVC	Stressed	Final	23
	Disyllabic	CVCC	Stressed	Final	4
		CVVC	Stressed	Final	15
	Multisyllabic	CVVC	Stressed	Final	3
Adult	Monosyllabic	CVCC	Stressed	Final	30
		CVVC	Stressed	Final	11
		CCVVC	Stressed	Final	50
		CCVCC	Stressed	Final	3
	Disyllabic	CVCC	Stressed	Final	3
		CVVC	Stressed	Final	19
		CCVVC	Stressed	Final	1
	Multisyllabic	CVVC	Stressed	Final	4

Age group 24–30

At 24–30 months, participants produced superheavy syllables in monosyllabic and disyllabic words only. Superheavy syllables were more frequent in monosyllabic words (n = 17, 85%) than in disyllabic words (n = 3, 15%). For monosyllabic words, a total of three sub-structures were attested, including CVCC (n = 5, 29.4%), CVVC (n = 6, 35.3%), and CCVVC (n = 6, 35.3%). On the other hand, only CVVC syllables in the stressed final environments were observed in disyllabic words. Figure 4.60 demonstrates the production of a stressed final CCVVC syllable in the word /ktaab/ (‘book’, Speaker YRO). Superheavy syllables in multisyllabic words were not attested in this age group.

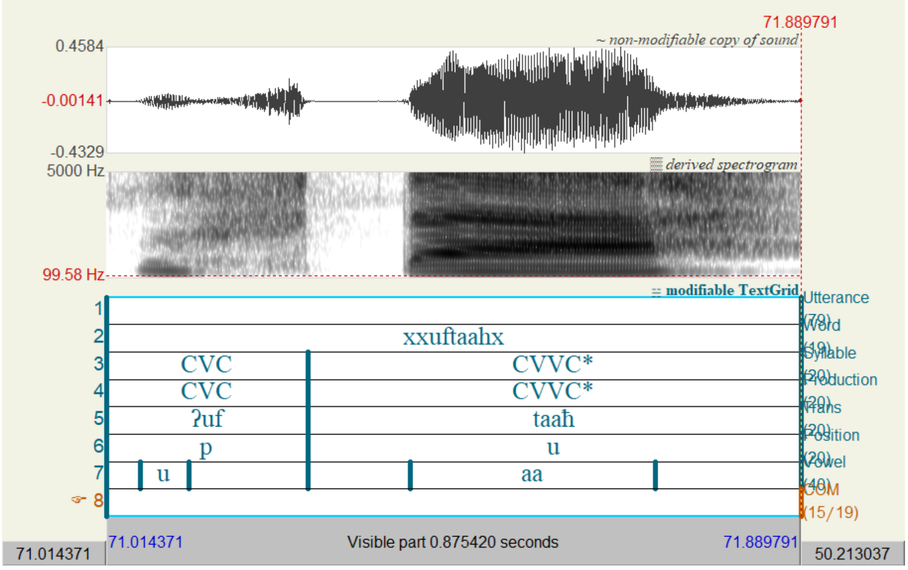


/ktaab/ Speaker YRO

Figure 4.60: Spectrogram of a superheavy syllable age group 24–30

Age group 31–37

By 31–37 months, participants maintained producing superheavy syllables in monosyllabic and disyllabic words only. Superheavy syllables in monosyllabic words (n = 14, 93.3%) are more frequent than those in disyllabic words (n = 1, 6.7%). Three sub-structures, including CVCC (n = 4, 28.6%), CVVC (n = 7, 50%), and CCVVC (n = 3, 21.4%), appeared in monosyllabic words. However, only CVVC syllables were attested in disyllabic words. Figure 4.61 demonstrates the production of a stressed final CVVC syllable in the word /ʔuftaaħ/ (‘key’, Speaker DDK).

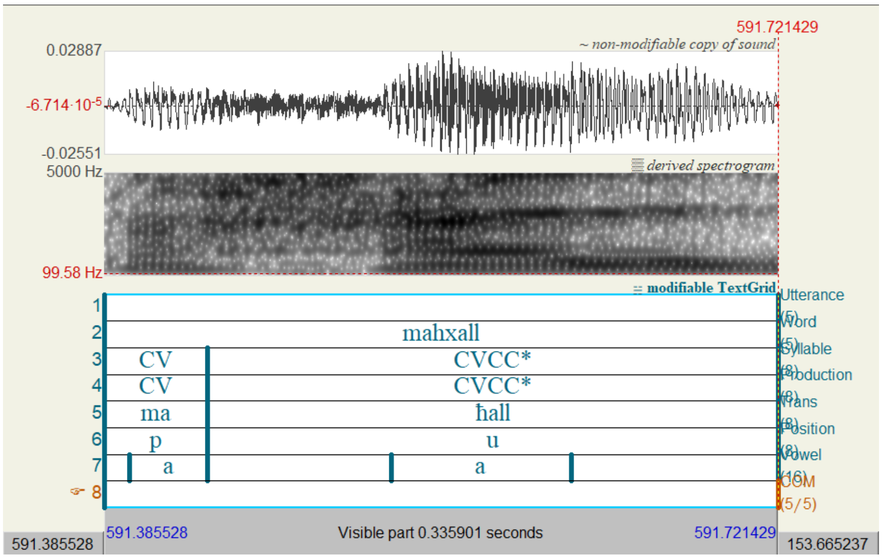


/ʔuftaaħ/ Speaker DDK

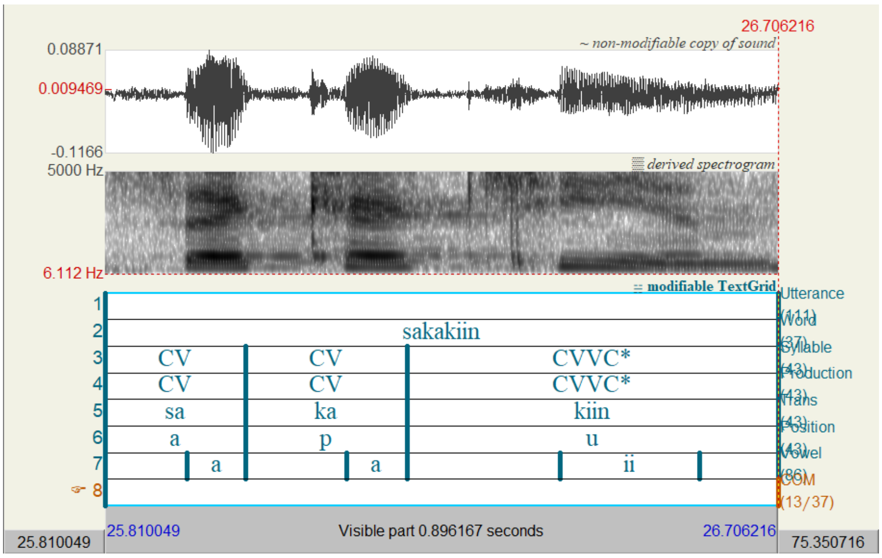
Figure 4.61:Spectrogram of a superheavy syllable in age group 31–37

Age goup 38–44

At 38–44 months, superheavy syllables were produced in all word length categories. Superheavy syllables were the most frequently produced in monosyllabic words (n = 36, 73.5%). All three sub-structures appeared in monosyllabic words, including CVCC (n = 13, 36.1%), CVVC (n = 6, 25%), and CCVVC (n = 14, 38.9%) syllables. A slight expansion in the observed sub-structures in disyllabic words was reported as a CVCC (n = 1, 8.3%) token emerged in this group, in addition to CVVC syllables (n = 11, 91.7%). Figure 4.62/a demonstrates a stressed final CVCC syllable in the word /maħall/ (‘shop’, CV.CVCC, Speaker MDD). Only CVVC syllables were produced in multisyllabic words. Figure 4.62/b demonstrates a stressed final CVVC syllable in the word /sakakiin/ (‘knives’, CV.CV.CVVC, Speaker 5PB). The superheavy syllable CVVC is the first to appear in multisyllabic words due to the high frequency of this sub-structure in JA speech. All superheavy productions in disyllabic and multisyllabic words were reported in the final stressed environment.



(a) /maħall/ Speaker 6VI

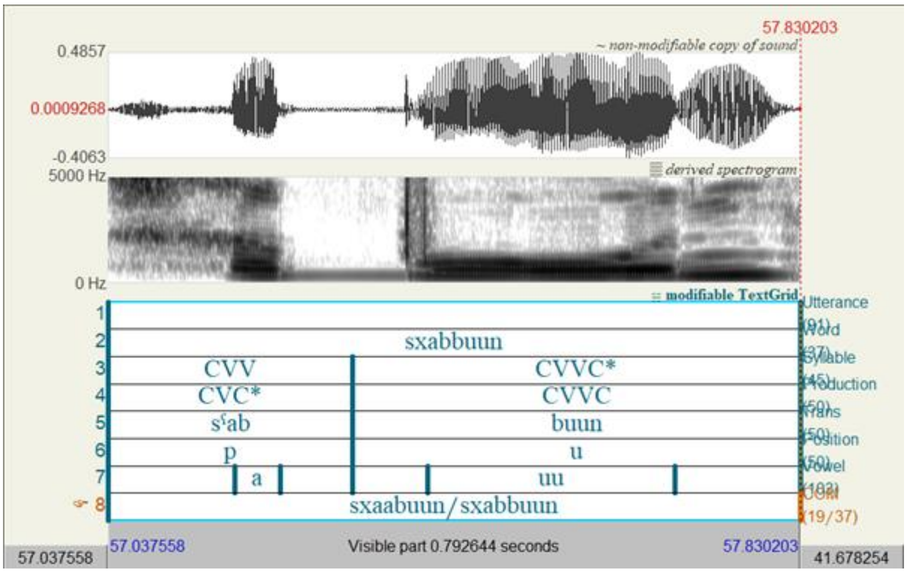


(b) /sakakiin/ Speaker 5PB

Figure 4.62: Spectrograms of superheavy syllables in age group 38–44

Age group 45–51

By 45–51 months, participants produced superheavy syllables more frequently in monosyllabic words (n = 40, 67.8%), compared to disyllabic (n = 17, 28.8%), and multisyllabic (n = 2, 3.4%) words. For monosyllabic words, superheavy syllables appeared in three sub-structures, including CVCC (n = 15, 37.5%), CVVC (n = 15, 37.5%), and CCVVC (n = 10, 25%). As for disyllabic words, two sub-structures were reported, CVCC (n = 2, 11.8%), and CVVC (n = 15, 88.2%), with CVVC being more frequent. An expansion in the number of superheavy syllables in disyllabic words (n = 17, 28.8%) is attested in this age group. Unstressed final CVVC syllables in disyllabic words emerged in this age group. Figure 4.63 demonstrates an unstressed final CVVC syllable in the word /sˤabbuun/ ('soap', CVC.CVVC, Speaker G9I). Nevertheless, in multisyllabic words, only CVVC was observed in the word final stressed environment. The developmental pattern diverting from the initial stressed final superheavy syllable production comprises the change in the stress assignment status of the syllable first and not the syllable position.

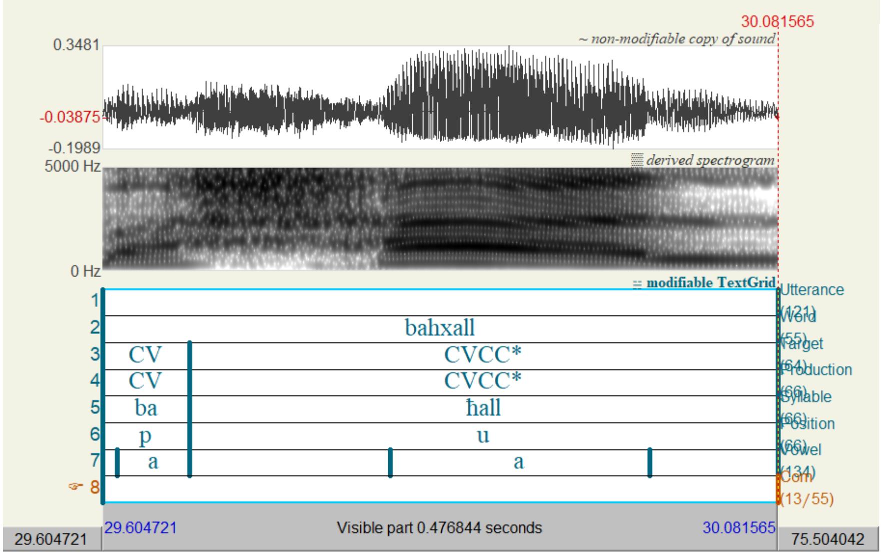


/sˤabbuun/ Speaker G9I

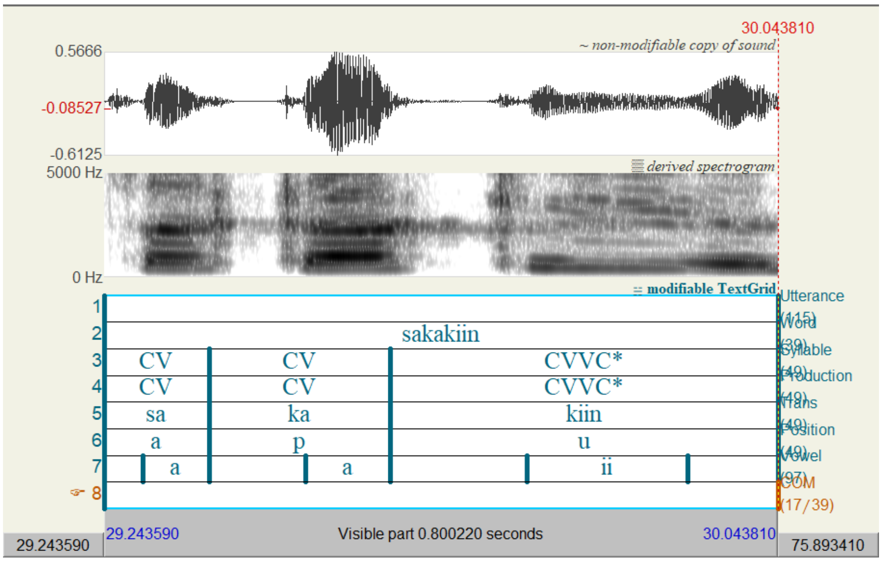
Figure 4.63: Spectrogram of a superheavy syllable in age group 45–51

Age group 52–58

By 52–58 months, participants preserved producing superheavy syllables across all word counts, with a slight increase in the number of productions in multisyllabic words (n = 4, 7.3%). Similar to the pattern attested in the younger groups, all sub-structures appeared in monosyllabic words, including CVCC (n = 10, 27.8%), CVVC (n = 12, 33.3%), and CCVVC (n = 14, 38.9%); only two appeared in disyllabic words, including CVCC (n = 2, 13.3%) and CVVC (n = 13, 86.7%) syllables; while only CVVC syllables appeared in multisyllabic words. Figure 4.64/a demonstrates a stressed final CVCC syllable in the word /baħall/ (‘shop’, CV.CVCC, Speaker MD5). Moreover, Figure 4.64/b demonstrates a stressed final CVVC in the word /sakakiin/ (‘knives’, CV.CV.CVVC, Speaker 18T). CVVC syllables are the most occurring compared to other sub-structures. All superheavy syllables in disyllabic and multisyllabic words were in the stressed word final environment.



(a) /baħall/ Speaker MD5

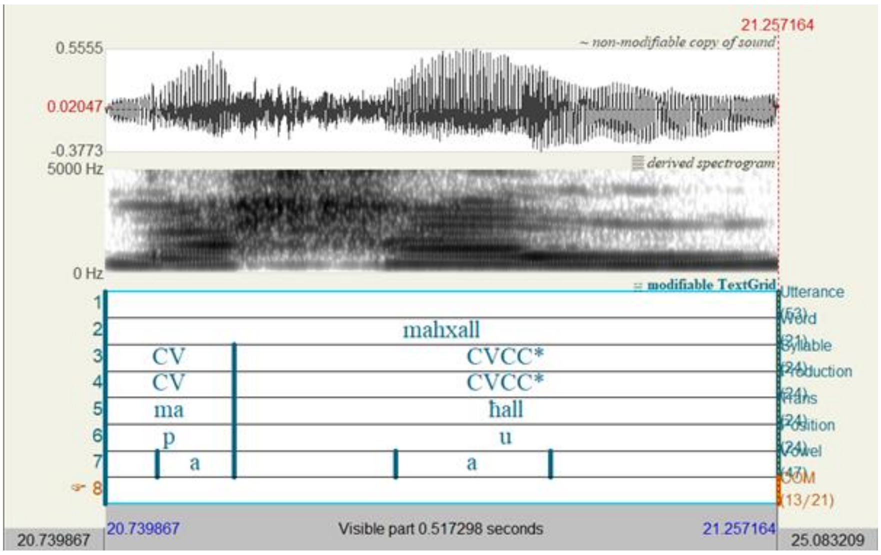


(b) /sakakiin/ Speaker 18T

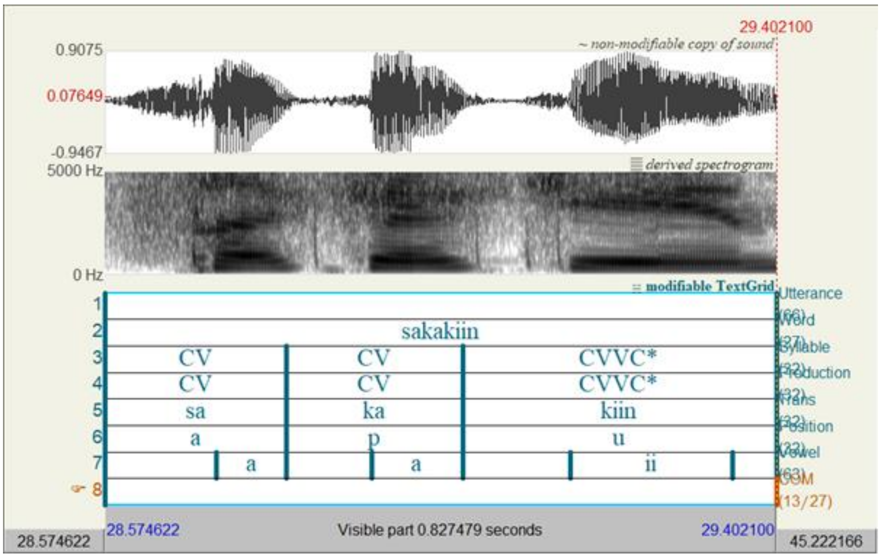
Figure 4.64: Spectrograms of superheavy syllables in age group 52–58

Age group 59–65

By 59–65 months, superheavy syllables are produced in monosyllabic, disyllabic, and multisyllabic words. Superheavy syllables in monosyllabic words are the most frequent (n = 29, 76.3%), followed by syllables in disyllabic words (n = 8, 21.1%) and then multisyllabic words (n = 1, 2.6%). Participants produced CCVVC syllables (n = 14, 48.3%) with a higher frequency than other sub-structures for monosyllabic words. CVVC and CVCC syllables appeared for disyllabic words, with CVVC (n = 7, 87.5%) being more commonly produced, while for multisyllabic words, only CVVC syllables were observed. Figure 4.65/a demonstrates the production of a stressed final CVCC syllable in the word /maħall/ (‘shop’, CV.CVCC, Speaker 6V9). Also, Figure 4.65/b demonstrates the production of a stressed final CVVC syllable in the word /sakakiin/ (‘knives’, CV.CV.CVVC, Speaker FDW). All productions in disyllabic and multisyllabic words were in the stressed word final environment.



(a) /maħall/ Speaker 6V9

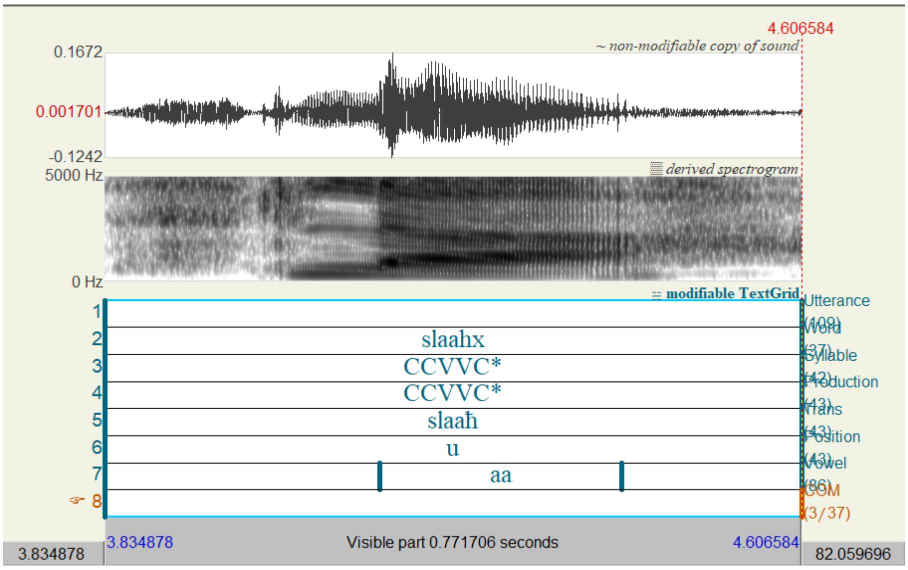


(b) /sakakiin/ Speaker FDW

Figure 4.65: Spectrograms of superheavy syllables in age group 59–65

Age group 66–72

At 66–72 months, superheavy syllables in monosyllabic (n = 55, 71.4%) and disyllabic words (n = 19, 24.7%) were more frequently produced than those in multisyllabic words (n = 3, 3.9%). Similar to the younger age groups, for monosyllabic words, all sub-structures were attested; for disyllabic words, CVVC and CVCC syllables were observed; but for multisyllabic words, only CVVCs were produced. An expansion in the frequency of CCVVC syllables (n = 23, 41.8%) was noted in monosyllabic words, which indicates an advancement of superheavy complexity across the age groups. Figure 4.66 demonstrates a stressed final CCVVC syllable in the word /slaaħ/ (‘weapon’, Speaker MMA). All productions stressed superheavy syllables in the word final position.

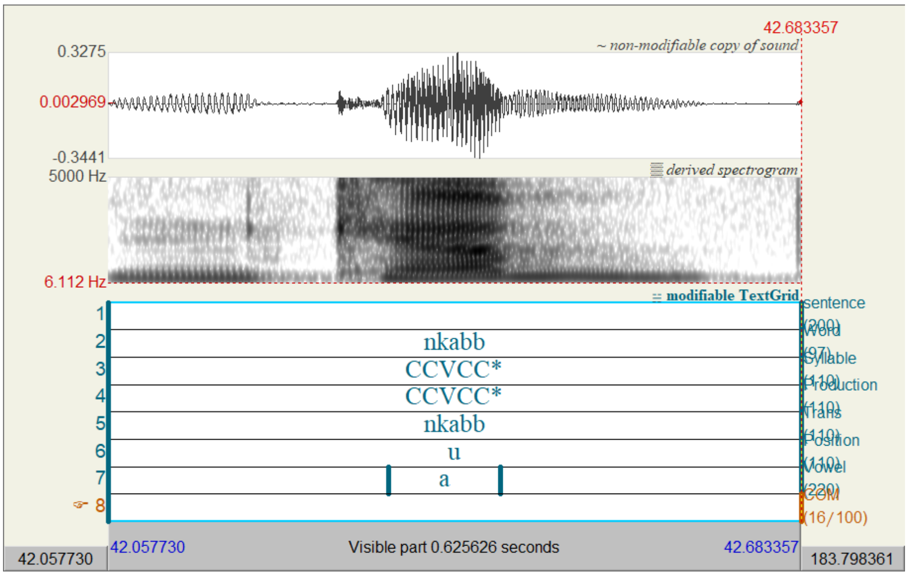


/slaaħ/ Speaker MMA

Figure 4.66: Spectrogram of a superheavy syllable in age group 66–72

Adults

Similar to the patterns observed in age groups 52–58, 59–65, and 66–72 months, the adult group produced superheavy syllables across all word counts. Superheavy syllables in monosyllabic words were the most frequent (n = 94, 77.7%), while they were the least frequent in multisyllabic words (n = 4, 3.3%). All sub-structures were attested for monosyllabic words, with CCVVC syllables (n = 50, 53.2%) being the most frequent. In addition, a fourth sub-structure, CCVCC, was observed (n = 3, 3.2%) (Note that this sub-structure was not attested in any of the child groups). Figure 4.67 demonstrates a stressed final CCVCC syllable in the word /nkabb/ (‘it spilled’, Speaker BMG). CVCC, CVVC, and CCVVC syllables were observed for disyllabic words, with CVVC syllables being the most frequent (n = 19, 82.6%). As for multisyllabic words, only CVVCs were produced. All superheavy syllables were in the stressed word final environment in disyllabic and multisyllabic words.



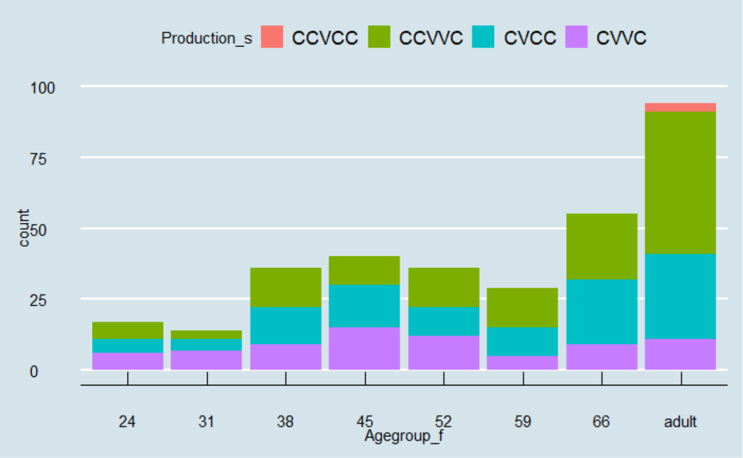
/nkabb/ Speaker BMG

Figure 4.67: Spectrogram of a superheavy syllable in the adult group

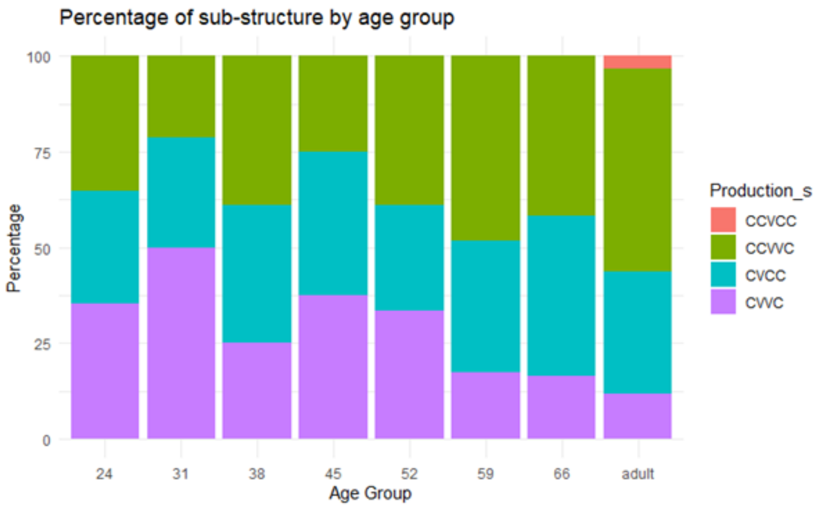
4.4.2.2.Bayesian Models and Descriptive Statistics

Superheavy Syllables in Monosyllabic Words

Figure 4.68/a demonstrates the frequency of superheavy syllables in monosyllabic words across the age groups. The number of superheavy syllables produced shows an overall increase as age increases. The syllabic complexity becomes more evident with age as more CCVVC syllables are produced, and CCVCC syllables appear in the adult group. Figure 4.68/b exhibits the percentages of sub-structural productions across the age groups. Until 52–58 months of age, CVVC syllables are the most frequently produced. However, by 59–65 months, the percentage of CVVC syllables slightly declines as the percentages of other sub-structures increase.



(a) Frequency



(b) Percentage

Figure 4.68: Frequency and percentage of superheavy syllables in monosyllabic words

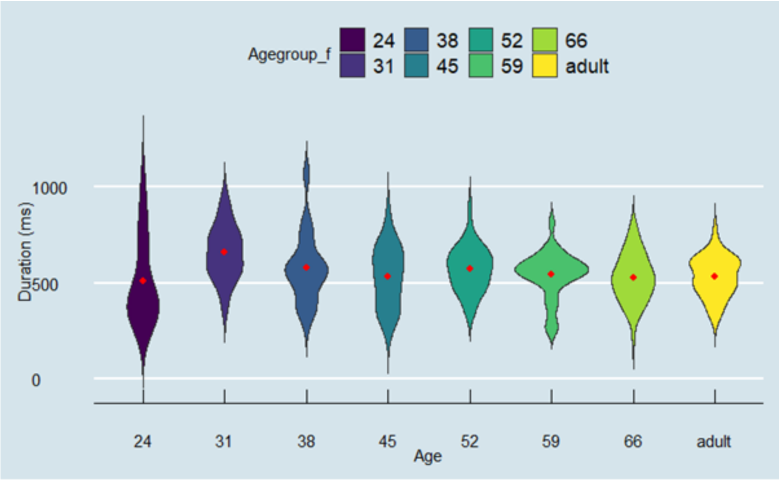
For superheavy syllables in monosyllabic words, a total of 321 syllables were analysed. For the syllable duration model, the Bayesian analysis was performed using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Since superheavy syllables in monosyllabic words only appear in stressed final environments, the model was fitted only to determine the effect of age group and sub-structure on syllable duration. Table 4.26 summarizes the Bayesian model output for superheavy syllables in monosyllabic words.

Table 4.26: Bayesian model output for superheavy syllables in monosyllabic words/ Syllable durations

Population–Level Effects	Estimate	Est. Error	l–95% CI	u–95% CI	Bulk_ESS	Tail_ESS
Intercept	590.12	37.92	516.89	666.43	8559	14441
Sub-structure (CVVC)	–35.72	29.38	–92.82	22.19	11325	17960
Sub-structure (CCVVC)	59.43	26.27	7.46	110.75	9952	16353
Sub-structure (CCVCC)	91.66	78.29	–62.52	245.03	23477	24339
Age group	–12.52	7.09	–26.68	1.42	7084	12369

Age Group

The model output indicates that age group is not a strong predictor for syllable duration (β = –12.52, CI [–26.68, 1.42]). The youngest two age groups have the highest mean duration, while the durations decrease by 38–44 months as demonstrated in Figure 4.69. The mean syllable duration in the youngest age group is 510.4 ms (SD = 224.0), then peaks for 31–37 months, reaching 659.3 ms (SD = 140.5). However, the mean durations negligibly fluctuate between 525 ms to 580 ms across the older age groups, with the oldest child group averaging 527.7 ms (SD = 126.7) and the adult group averaging 533.6 ms (SD = 106.7).

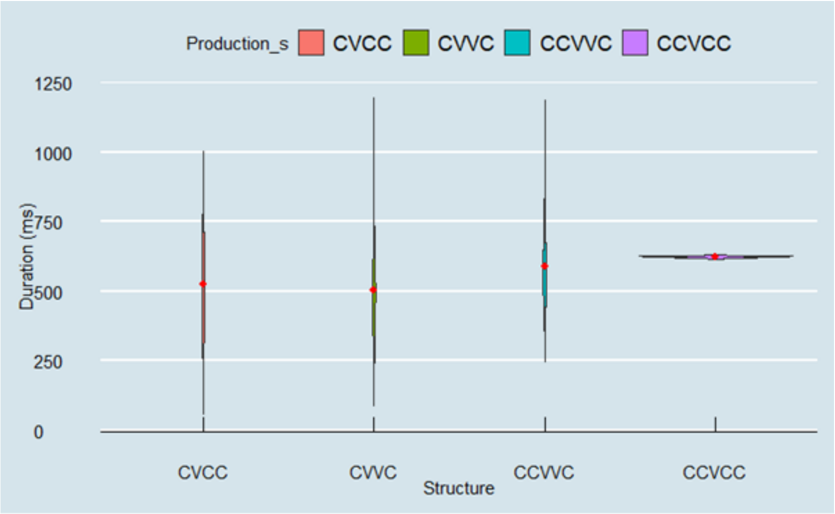


Syllable duration

Figure 4.69: The distribution and the mean syllable duration (ms) for age group in monosyllabic words

Sub-Structure

The model output suggests that sub-structure is only a strong predictor for syllable duration (CCVVC, β = 59.43, CI [7.46, 110.75]), but not for (CVVC, β = –35.75, CI [–92.82, 22.19]), or (CCVVC, β = 91.66, CI [–62.52, 245.03]). Figure 4.70 shows that CCVCC syllables have the highest mean durations (Mean = 623 ms, SD = 3.5), followed by CCVVC syllables (Mean = 587.7 ms, SD = 123.8), then CVCC syllables (Mean = 525.8 ms, SD = 143.1), while CVVC syllables are the shortest (Mean = 501.9 ms, SD = 148.6). On average, CCVCC syllables are 1.1 times longer than CVCC syllables and 1.2 times longer than CVVC and CVCC syllables.



Syllable duration

Figure 4.70: The distributions and the mean syllable and vowel durations (ms) for sub-structure in monosyllabic words

Vowel Duration

As for the vowel duration model, the Bayesian analysis was performed with a total of 321 vowels using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Since superheavy syllables in monosyllabic words only appear in stressed final environments, the model was fitted only to determine the effect of age group and sub-structure on the durations. Table 4.27 summarizes the Bayesian model output for superheavy syllables in monosyllabic words.

Table 4.27: Bayesian model output summary vowel duration in monosyllabic words

Population–Level Effects	Estimate	Est. Error	l–95% CI	u–95% CI	Bulk_ESS	Tail_ESS
Intercept	184.75	21.52	143.53	228.97	10882	15227
Sub-structure (CVVC)	113.07	12.59	89.88	139.75	12193	14428
Sub-structure (CCVVC)	115.98	10.76	94.50	138.01	15336	14079
Sub-structure (CCVCC)	–22.27	34.15	–88.93	44.84	33434	24123
Age group	–9.88	4.30	–18.54	–1.44	8370	11525

Age Group

The model output suggests that age group is a strong predictor for vowel duration (β = –9.88, CI [–18.54, –1.44]). Figure 4.71 demonstrates that vowel durations in superheavy syllables slightly decrease with age, with the youngest two age groups exhibiting the highest mean values. The youngest age group has a mean vowel duration of 233.7 ms (SD = 79.9), followed by an increase in age group 31–37 months (Mean = 314.9 ms, SD = 107.1). The onset of the decrease is observed in age group 38–44 months (Mean = 227 ms, SD = 123.5). The adult group has the shortest mean vowel duration (Mean = 189.2 ms, SD = 73.5).

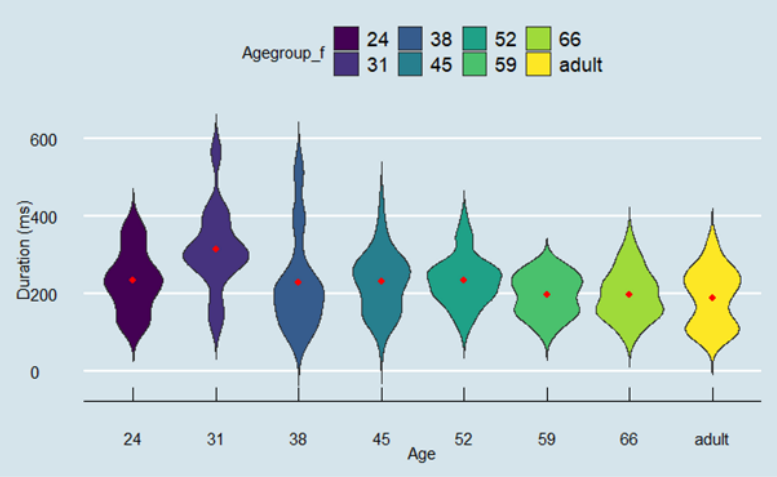


Figure 4.71: Posterior predictive plot for vowel durations across the age groups in monosyllabic words

Sub-Structure

Sub-structure is a strong predictor for vowel duration in superheavy syllables in monosyllabic words (CVVC, β = 113.07, CI [–89.88, 139.75]), and (CCVVC, β = 115.98, CI [94.50, 138.01]), but not for (CCVCC, β = –22.27, CI [–88.93, 44.84]). Figure 4.72 displays that vowels in CVVC syllables are the longest (Mean = 262.3 ms, SD = 76.6), then vowels in CCVVC syllables (Mean = 247.4 ms, SD = 69.4), followed by vowels in CVCC syllables (Mean = 141.9 ms, SD = 45.7), while vowels in CCVCC syllables are the shortest (Mean = 106.2 ms, SD = 21.8). Approximately, vowels in CVVC syllables are 1.1 times longer than vowels in CCVVC syllables, 1.8 times longer than vowels in CVCC syllables, and 2.5 times longer than in CCVCC.

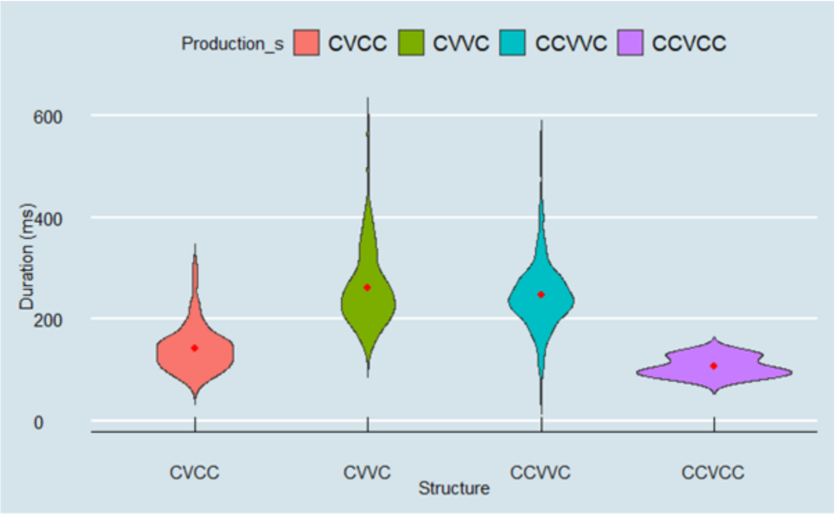
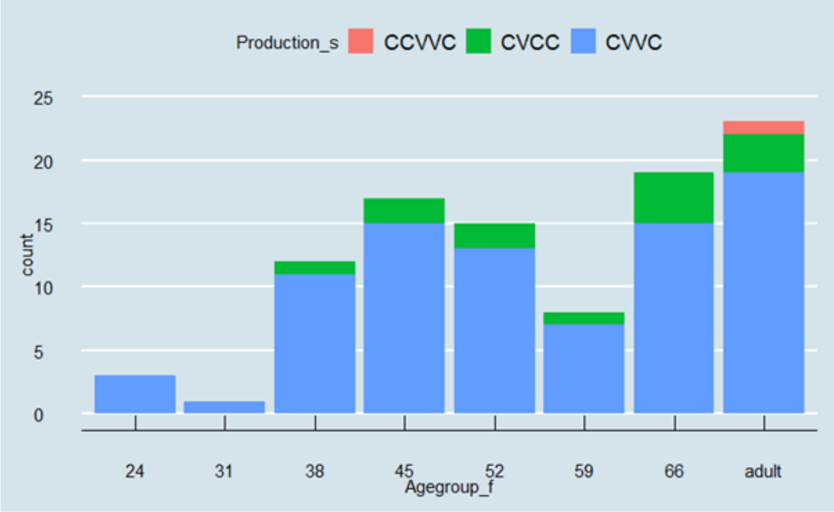


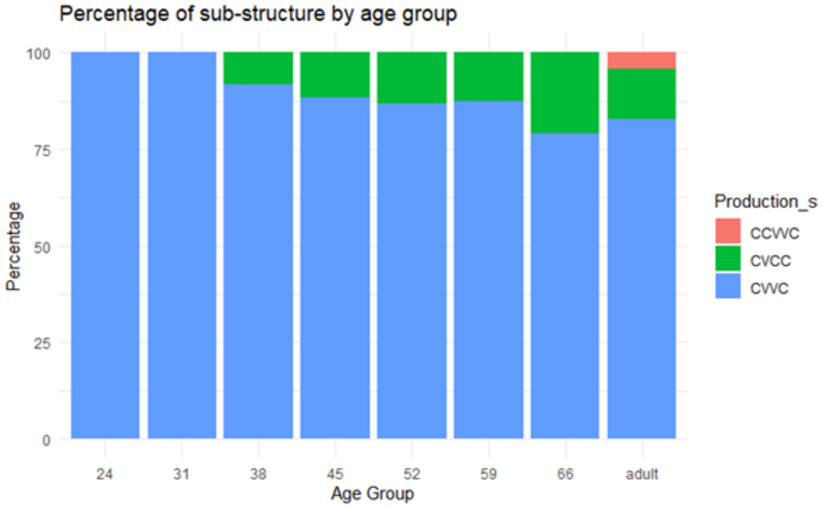
Figure 4.72: Posterior predictive plot for vowel duration for sub-structure in monosyllabic words

Superheavy Syllables in Disyllabic and Multisyllabic Words

Figure 4.73 demonstrates the frequency and percentage distributions of superheavy sub-structures in disyllabic words across the age groups. CVVC syllables are the most frequently produced, with a notable decrease in the percentage of these syllables by the age of 38–44 months, where the percentages of CVCC and CCVVC syllables increase. Only adults produced three superheavy sub-structures in their disyllabic words including CVCC, CVVC, and CCVVC.



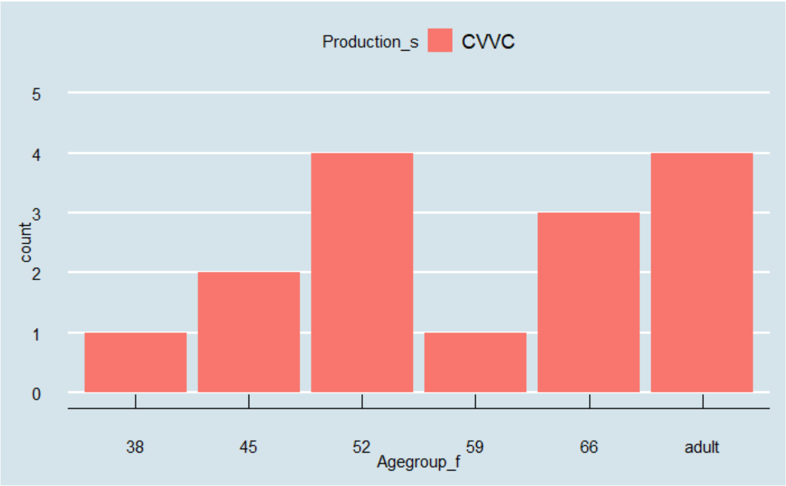
(a) Frequency



(b) Percentage

Figure 4.73: Frequency and percentage of superheavy syllables in disyllabic words

Figure 4.74 demonstrates the frequency distribution of CVVC syllables, as they were the only attested sub-structure. No observations of superheavy syllables in multisyllabic words were recorded for age groups 24–30 and 31–37 months. All superheavy syllable productions were in stressed final environments.



(a) Frequency

Figure 4.74: Frequency of superheavy syllables in multisyllabic words

Syllable Duration

For superheavy syllables in disyllabic (n = 98) and multisyllabic words (n = 19), a total of 130 syllables were analysed. Predictors such as age group, stress, and sub-structure were assigned as independent variables. Syllable position was not assigned as a predator as all tokens were observed in the word final position. For the syllable duration model, the Bayesian analysis was performed using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Table 4.28 summarizes the Bayesian model output for superheavy syllables in disyllabic and multisyllabic words.

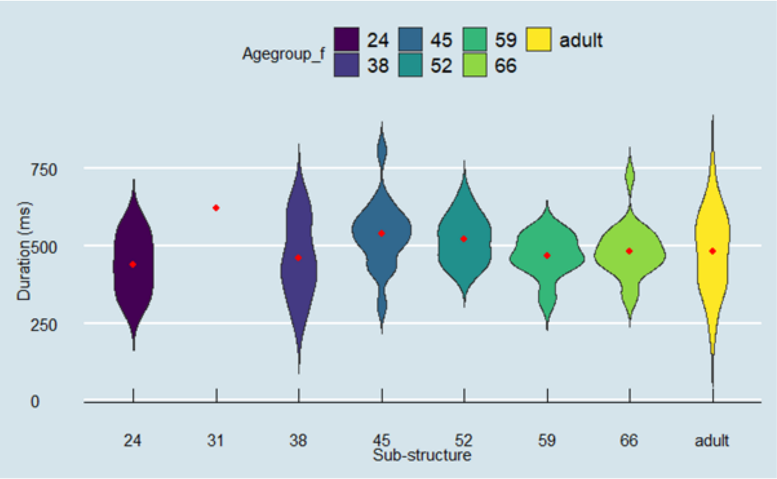
Table 4.28: Bayesian model output summary for syllable duration in di/multisyllabic words

Population–Level Effects	**Estimate**	**Est. Error**	**l–95% CI**	**u–95% CI**	**Bulk_ESS**	**Tail_ESS**
Intercept	408.31	77.46	255.85	560.43	12818	18741
Sub-structure (CVVC)	70.42	43.91	–16.16	158.42	12853	18212
Sub-structure (CCVVC)	272.82	96.21	83.09	462.57	17429	20412
Stress (Stressed)	41.64	54.39	–65.16	148.24	18813	22607

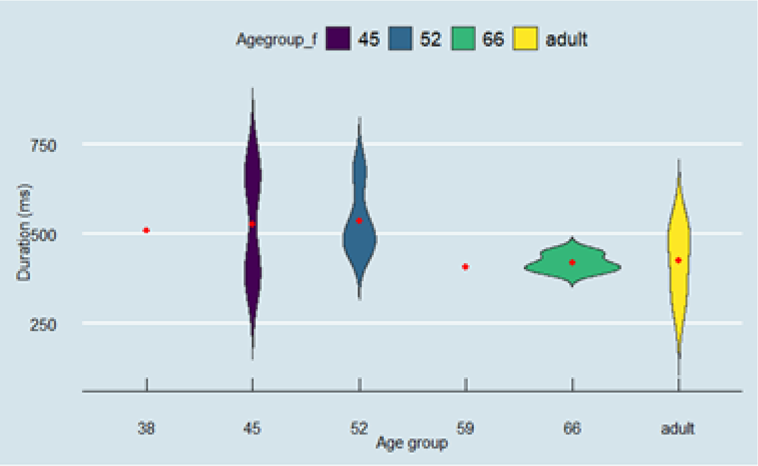
Age group	–4.66	8.07	–20.40	11.67	6165	9965

Age Group

The model output suggests that age group is not a strong predictor for syllable duration (β = –4.66, CI [–20.40, 11.67]). Figure 4.75/a shows that the mean syllable duration in the youngest age group is 440.3 ms (SD = 105.8), followed by a noticeable increase by 31–37 months, reaching 621.9 ms (SD = NA). In age group 45–51 months (Mean = 538.6 ms, SD = 109.8), the durations start declining, reaching the adult group mean (Mean = 481.9 ms, SD = 126.1). Superheavy syllables in multisyllabic words were not attested in age groups 24–30 months and 31–37 months. Figure 4.75/b shows that in age group 38–44 months, the mean syllable duration is 508.2 ms (SD = NA), and a consistent durational increase is observed until age group 52–58 months (Mean = 535.6 ms, SD = 99.0). By 59–65 months, the mean duration decreases (Mean = 405.6 ms, SD = NA), reaching the adult group mean (Mean = 424.1 ms, SD = 98.7).



(a) Superheavy syllables in disyllabic words



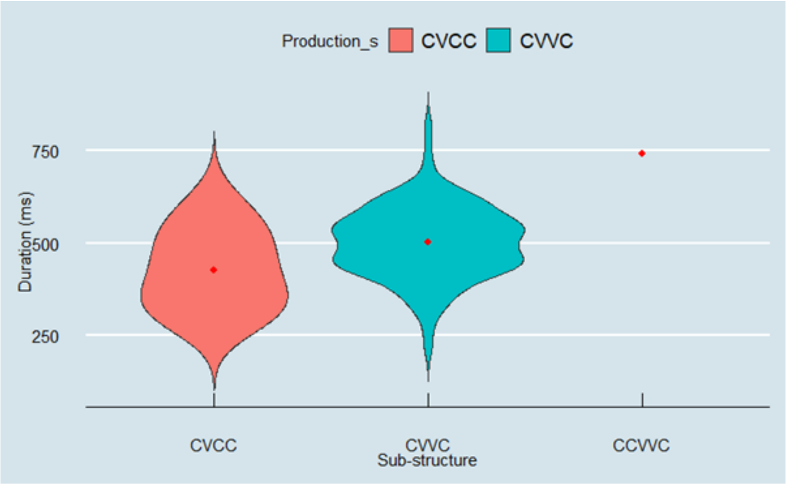
(b) Superheavy syllables in multisyllabic words

Figure 4.75: The distributions and the mean syllable durations (ms) for age group in disyllabic and multisyllabic words

Sub-Structure

The model output shows that sub-structure is a strong predictor for syllable duration for (CCVVC, β = 272.82, CI [83.09, 462.57]), but not for (CVVC, β = 70.42, CI [–16.16, 158.42]). Figure 4.76 shows that for superheavy syllables appearing in disyllabic words, CCVVC syllables are the longest (Mean = 741.8 ms, SD = NA), followed by CVVC syllables (Mean = 502.3 ms, SD = 97.9), while CVCC syllables are the shortest (Mean = 426.9 ms, SD = 110.6).

The violin plot for CCVVC syllables in disyllabic words was not generated in the below figure as there was only an occurrence of CCVVC syllable.

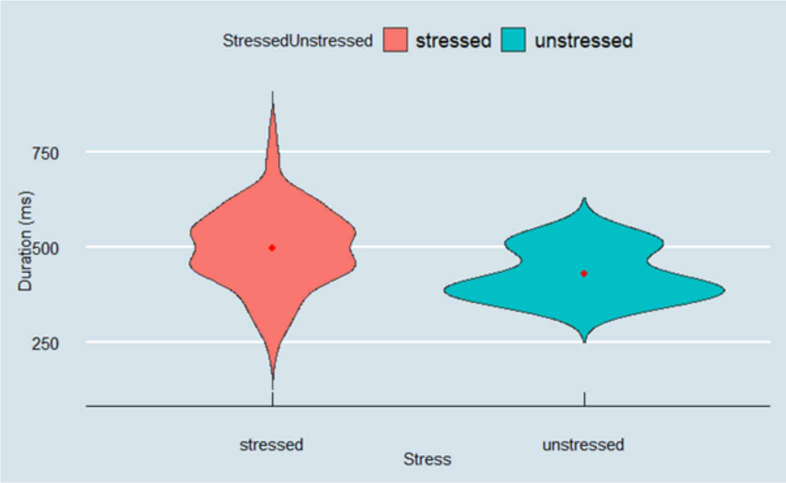


Syllable duration

Figure 4.76: The distributions and the mean syllable duration (ms) for sub-structure in disyllabic words

Stress

Moreover, the model output suggests that stress assignment is not a strong predictor for syllable duration (β = 41.64, CI [–65.16, 148.24]). Stressed superheavy syllables in disyllabic words (Mean = 496.8 ms, SD = 104.9) are approximately 1.2 times longer than unstressed syllables (Mean = 427.9 ms, SD = 75.1). Vowels in stressed superheavy syllables (Mean = 212.0 ms, SD = 60.8) are only 1.1 times longer than their counterparts (Mean = 196.1 ms, SD = 80.9).



Syllable duration

Figure 4.77: The distributions and the mean syllable duration (ms) for stress in disyllabic words

Vowel Duration

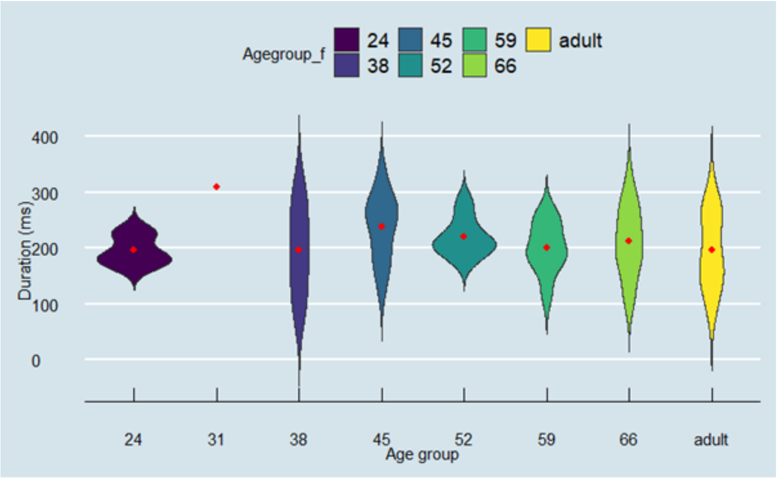
For the vowel duration model (n = 130), the Bayesian analysis was performed using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Table 4.29 summarizes the Bayesian model output for superheavy syllables in disyllabic and multisyllabic words.

Table 4.29: Bayesian model output summary for vowel duration in di/multisyllabic words

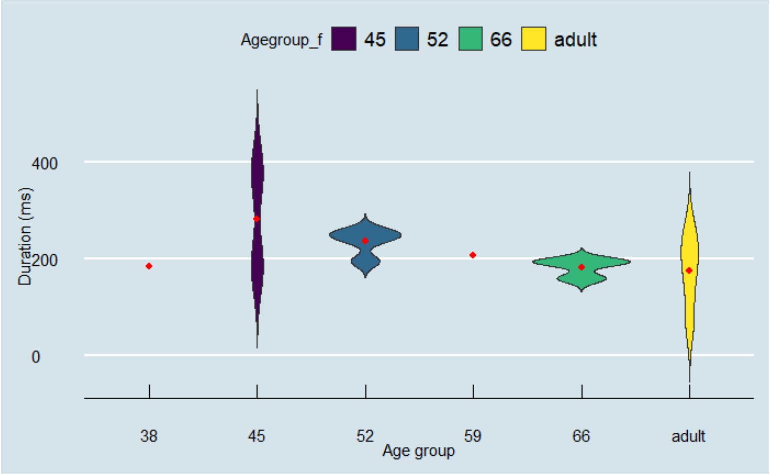
Population–Level Effects	**Estimate**	**Est. Error**	**l–95% CI**	**u–95% CI**	**Bulk_ESS**	**Tail_ESS**
Intercept	145.18	39.65	66.08	222.61	13818	19727
Sub-structure (CVVC)	91.70	19.96	53.18	132.20	15403	17143
Sub-structure (CCVVC)	122.94	48.71	25.64	217.81	21838	21541
Stress (Stressed)	2.57	27.98	–52.42	58.01	24992	22604
Age group	–3.13	4.62	–12.17	6.16	7468	11357

Age Group

The model output indicates that age group is not a strong predictor for vowel duration (β = –3.13, CI [–12.17, 6.16]). Figure 4.78/a demonstrates variability in mean vocalic durations is attested in the younger age groups ranging from 24–30 months (Mean = 195.6 ms, SD = 28.9), 31–37 months (Mean = 310.1 ms, SD = NA), 38–44 months (Mean = 196.5 ms, SD = 78.8), to 45–51 months (Mean = 237.2 ms, SD = 59.1). At 52–58 months (Mean = 219.7 ms, SD = 34.6), the mean durations demonstrate a general decreasing trend, reaching the adult group mean (Mean = 196.3 ms, SD = 68.2). As for multisyllabic words, Figure 4.78/b shows that the mean vowel durations reach a maximum in age group 52–58 months (Mean = 235.2 ms, SD = 28.2). However, a consistent decrease is shown after 59–65 months of age, reaching the adult group mean (Mean = 174.1 ms, SD = 78.6).



(a) Superheavy syllables in disyllabic words

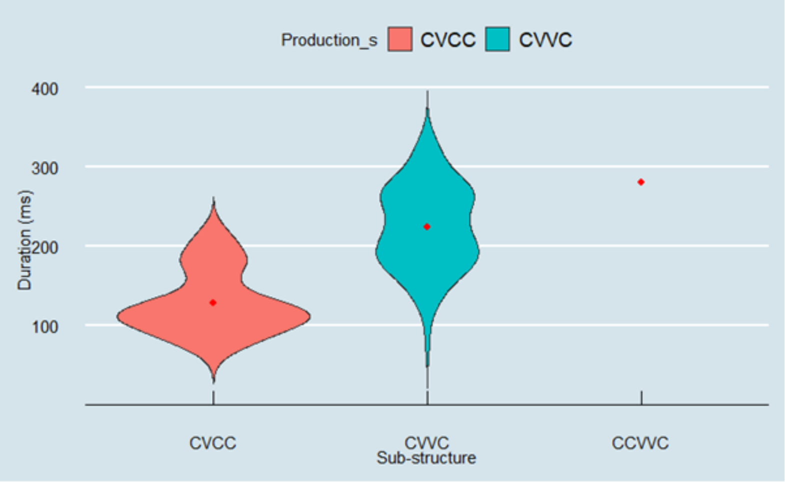


(b) Superheavy syllables in multisyllabic words

Figure 4.78: The distributions and the mean vowel durations (ms) for age group in disyllabic and multisyllabic words

Sub-Structure

The model output suggests that sub-structure is a strong predictor for vowel duration (CVVC, β = 91.70, CI [53.18, 132.20]), (CCVVC, β = 122.94, CI [25.64, 217.81]). Figure 4.79 shows that vowels in CCVVC syllables are the longest (Mean = 279.9 ms, SD = NA), then vowels in CVVC syllables (Mean = 223.5 ms, SD = 53.4), whereas vowels in CVCC syllables are the shortest (Mean = 129.2 ms, SD = 40.2).

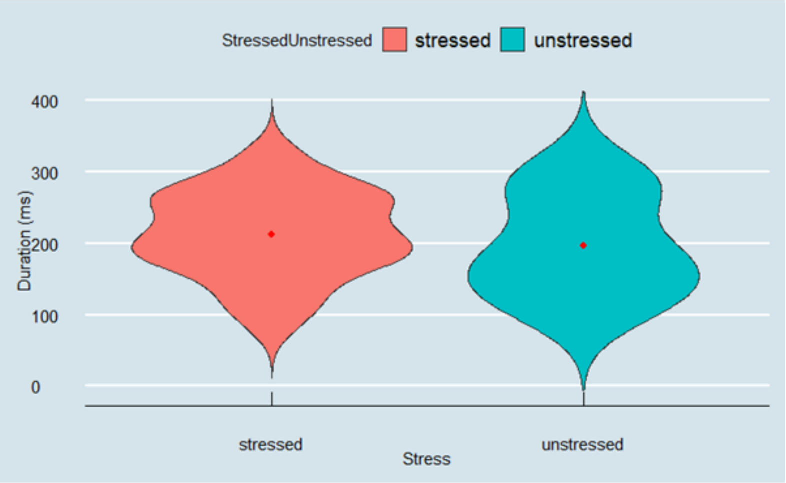


Superheavy syllables in disyllabic words

Figure 4.79: The distributions and the mean vowel duration (ms) for sub-structure in disyllabic words

Stress

As for stress, the model output shows that it is not a strong predictor for vowel duration (β = 2.57, CI [–52.42, 58.01]). Vowels in stressed superheavy syllables (Mean = 212.0 ms, SD = 60.8) are only 1.1 times longer than their counterparts (Mean = 196.1 ms, SD = 80.9).



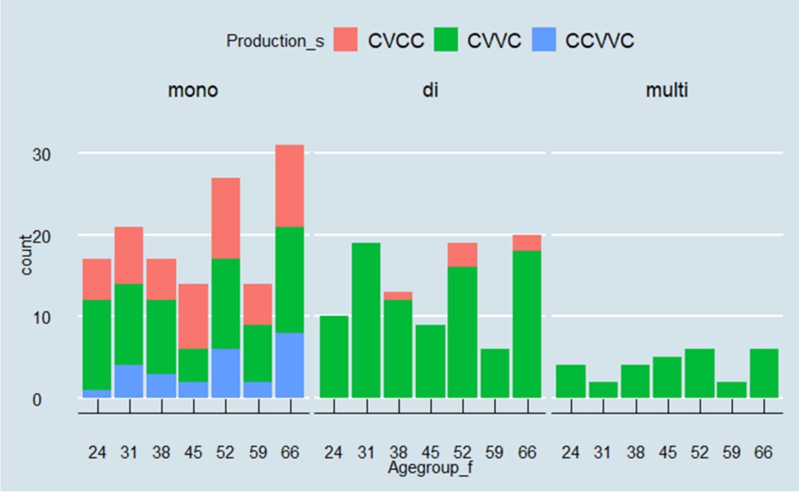
Superheavy syllables in disyllabic words

Figure 4.80: The distributions and the mean vowel duration (ms) for stress in disyllabic words

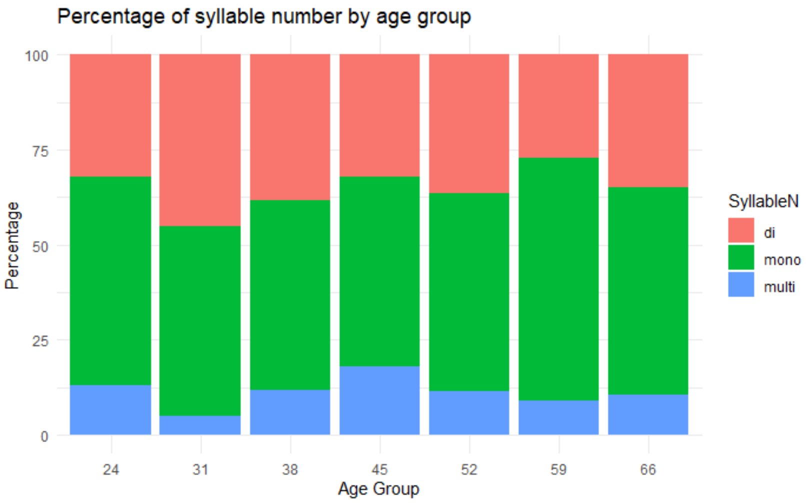
4.4.3. Picture Task

4.4.3.1.Superheavy Syllable Frequency

Figure 4.81/a demonstrates the frequency of superheavy syllables in monosyllabic, disyllabic, and multisyllabic words across the age groups ranging from 24–30 months to 66–72 months. Superheavy syllables are the most frequent in monosyllabic words, followed by disyllabic words, and they are the least frequent in multisyllabic words. Additionally, more sub–structural varieties of superheavy syllables are attested in monosyllabic words compared to disyllabic and multisyllabic words. Figure 4.81/b suggests that JA children performed similarly across the age group, producing the superheavy syllables in the three-word lengths with comparable percentages.



(a) Frequency



(b) Percentage

Figure 4.81: Frequency and percentage of superheavy syllables in PT

Additionally, Table 4.30 presents the following variables: age group (children: 24–30, 31–37, 38–44, 45–51, 52–58, 59–65, and 66–72 months), word length (monosyllabic, disyllabic, and multisyllabic), sub–structure (CVCC, CVVC, and CCVVC), stress assignment (stressed, unstressed), and syllable position within a word (final, nonfinal).

Table 4.30: Superheavy syllables PT

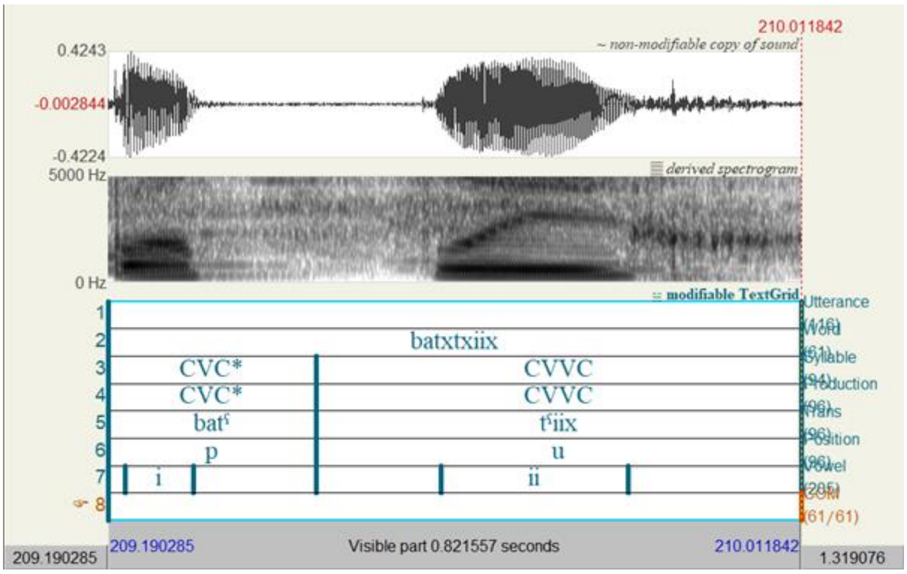
Age Group	Word Length	Sub-Structure	Stress	Position	Frequency (n)
24–30	Monosyllabic	CVCC	Stressed	Final	5

		CVVC	Stressed	Final	11
		CCVVC	Stressed	Final	1
	Disyllabic	CVVC	Unstressed	Final	2
			Stressed	Final	8
	Multisyllabic	CVVC	Stressed	Final	4
31–37	Monosyllabic	CVCC	Stressed	Final	7
		CVVC	Stressed	Final	10
		CCVVC	Stressed	Final	4
	Disyllabic	CVVC	Unstressed	Final	4
			Stressed	Final	15
	Multisyllabic	CVVC	Unstressed	Final	1
			Stressed	Nonfinal	1
38–44	Monosyllabic	CVCC	Stressed	Final	5
		CVVC	Stressed	Final	9
		CCVVC	Stressed	Final	3
	Disyllabic	CVCC	Stressed	Final	1
		CVVC	Unstressed	Final	2
			Stressed	Final	10
	Multisyllabic	CVVC	Stressed	Nonfinal	1
				Final	3
45–51	Monosyllabic	CVCC	Stressed	Final	8
		CVVC	Stressed	Final	4
		CCVVC	Stressed	Final	2
	Disyllabic	CVVC	Unstressed	Final	4
			Stressed	Final	5
	Multisyllabic	CVVC	Unstressed	Final	1
			Stressed	Final	4
52–58	Monosyllabic	CVCC	Stressed	Final	10
		CVVC	Stressed	Final	11
		CCVVC	Stressed	Final	6
	Disyllabic	CVCC	Unstressed	Nonfinal	2
			Stressed	Final	1
		CVVC	Unstressed	Final	5
			Stressed	Final	11
	Multisyllabic	CVVC	Stressed	Nonfinal	1
				Final	5
59–65	Monosyllabic	CVCC	Stressed	Final	5
		CVVC	Stressed	Final	7
		CCVVC	Stressed	Final	2
	Disyllabic	CVVC	Unstressed	Final	3
			Stressed	Final	3

	Multisyllabic	CVVC	Stressed	Final	2
66–72	Monosyllabic	CVCC	Stressed	Final	10
		CVVC	Stressed	Final	13
		CCVVC	Stressed	Final	8
	Disyllabic	CVCC	Unstressed	Nonfinal	2
		CVVC	Unstressed	Final	9
			Stressed	Final	9
	Multisyllabic	CVVC	Stressed	Final	6

Age group 24–30

At 24–30 months, superheavy syllables in monosyllabic words are the most frequent (n = 17, 54.8%), followed by syllables in disyllabic words (n = 10, 32.3%), while syllables in multisyllabic words (n = 4, 12.9%) are the least frequent. A total of three sub-structures were attested in monosyllabic words, including CVCC, CVVC, and CCVVC, with CVVC syllables (n = 11, 64.7%) being the most frequent. As for disyllabic words, only CVVC syllables were observed. The majority of CVVC productions (n = 8, 80%) were in stressed word-final environments. However, unstressed word final CVVCs emerged in this age group. Figure 4.82 demonstrates an unstressed final CVVC syllable in the /battiix/ (‘watermelon’, CVC.CVVC, Speaker YRO). Similarly, only CVVC syllables were observed in multisyllabic words, as they were attested in stressed final environments.

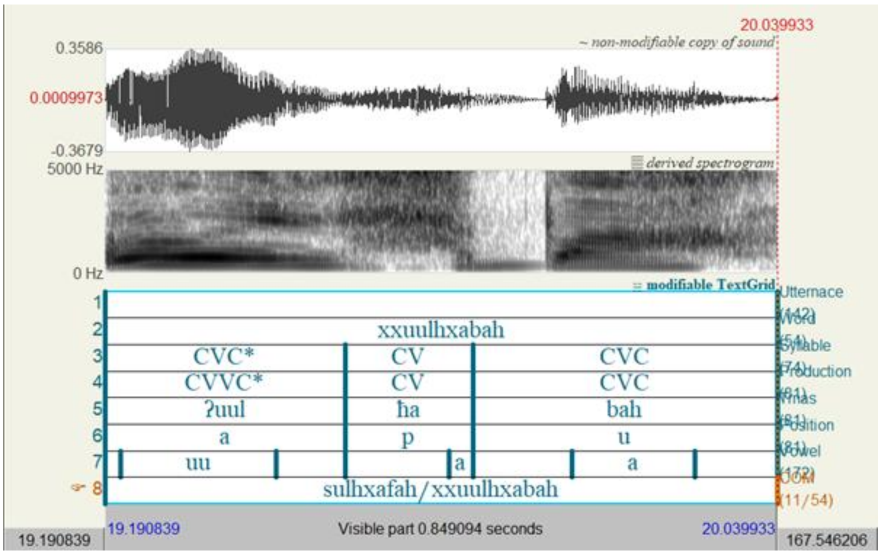


/battiix/ Speaker YRO

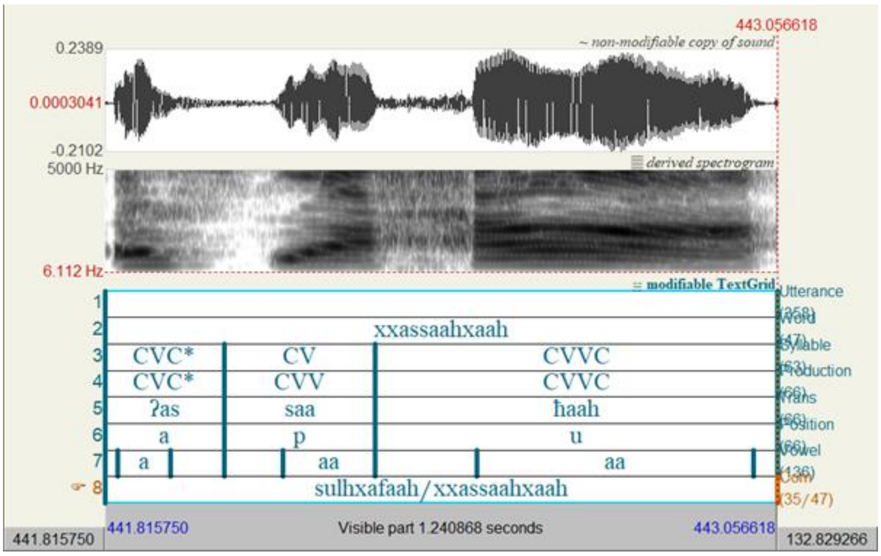
Figure 4.82: Spectrogram of a superheavy syllable in age group 24–30

Age group 31–37

At 31–37 months, superheavy syllables in monosyllabic words (n = 21, 50%) were the most frequent, and the least frequent in multisyllabic words (n = 2, 4.8%). Similar to the youngest age group, all three sub-structures including CVCC (n = 7, 33.3%), CVVC (n = 10, 47.6%), and CCVVC (n = 4, 19%), were observed in monosyllabic words. As for disyllabic words, only CVVC syllables were attested. CVVC syllables appeared in stressed final and unstressed final environments. The production of superheavy syllables in multisyllabic words was limited to CVVC syllables. These syllables were observed in stressed non–final environments, as Figure 4.83/a demonstrates the word /ʔuulħabah/ (‘turtle’, CVVC.CV.CVC, Speaker M37). The emergence of unstressed final CVVC syllables in multisyllabic words was also reported. Figure 4.83/b demonstrates an unstressed final CVVC syllable in the word /ʔassaaħaah/ (‘turtle’, CVC.CVV.CVVC, Speaker 2GS).



(a) /ʔuulħabah/ Speaker M37

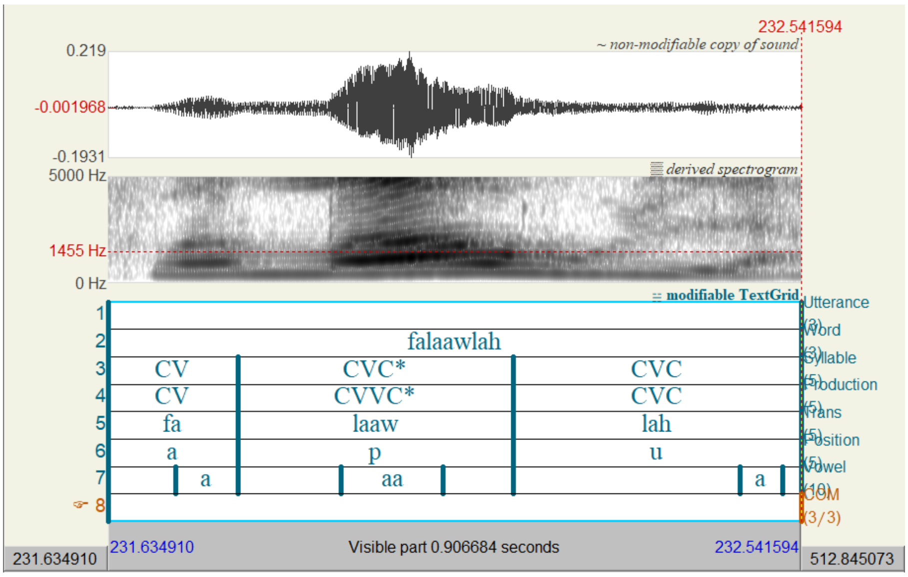


(b) /ʔassaaħaah/ Speaker 2GS

Figure 4.83: Spectrograms of superheavy syllables in age group 31–37

Age group 38–44

At 38–44 months, participants consistently produced superheavy syllables across all word lengths. Superheavy syllables in monosyllabic (n = 17, 50%) and disyllabic words (n = 13, 38.2%) were more frequently produced compared to syllables in multisyllabic words (n = 4, 11.8%). All sub-structures were attested in monosyllabic words, including CVCC (n = 5, 29.4%), CVVC (n = 9, 52.9%), and CCVVC (n = 3, 17.6%). Disyllabic words had an expansion in the observed structures as in addition to CVVC syllables (n = 12, 92.3%), a CVCC token (n = 1, 7.7%) was reported. The majority of superheavy productions in disyllabic words were in the stressed final environment (n = 10, 83.3%); however, unstressed final CVVC syllables were observed (n = 2, 16.7%). As for multisyllabic words, only CVVC syllables were attested appearing in stressed final and non–final environments. Figure 4.84 demonstrates a stressed non–final CVVC syllable in the word /falaawlah/ CV.CVVC.CVC (‘strawberry’, CV.CVVC.CVC, Speaker 6VI). The number of superheavy syllables in disyllabic and multisyllabic words has slightly increased compared to the younger groups.



/falaawlah/ Speaker 6VI

Figure 4.84: Spectrogram of a superheavy syllable in age group 38–44

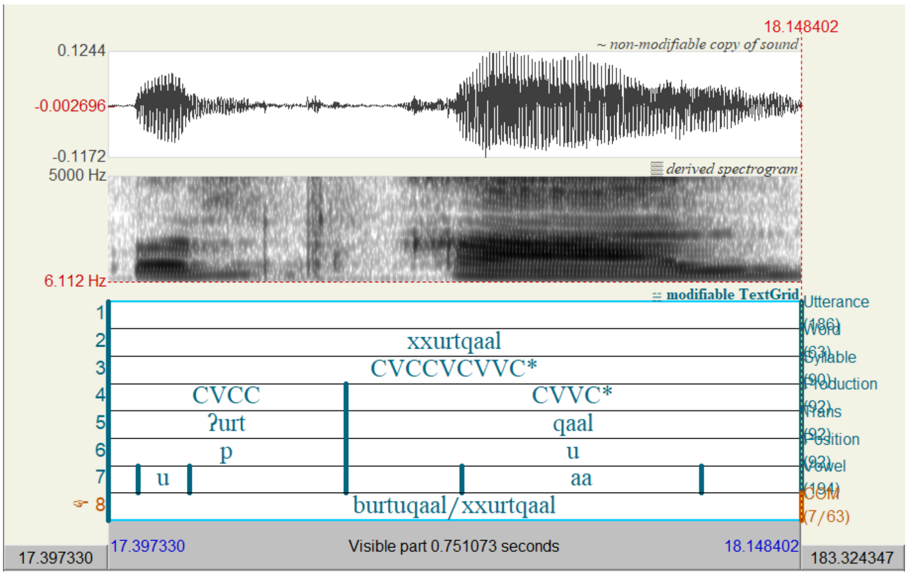
Age group 45–51

By 45–51 months, participants produced superheavy syllables across all word lengths. Superheavy syllables in monosyllabic words were the most frequent (n = 14, 50%), followed by disyllabic words (n = 9, 32.1%), while they were least frequent in multisyllabic words (n = 5, 17.9%) were the least frequent. All sub-structures appeared in monosyllabic words, including

CVCC (n = 8, 57.1%), CVVC (n = 4, 28.6%), and CCVVC (n = 2, 14.3%). CVVC was the only sub-structure of superheavy syllables observed in the data for disyllabic and multisyllabic words. Similar to the younger groups, participants produced stressed and unstressed final CVVC syllables in disyllabic words. Moreover, in stressed and unstressed final environments, superheavy productions in multisyllabic words were attested. It is noticed that the linguistic structures appear in disyllabic words first, then in multisyllabic words.

Age group 52–58

By 52–58 months, the frequency of superheavy syllables increased across all word lengths, with superheavy syllables being the most frequent in monosyllabic words (n = 27, 51.9%), followed by disyllabic words (n = 19, 36.5%), then multisyllabic words (n = 6, 11.5%). Three sub-structures CVCC (n = 10, 37%), CVVC (n = 11, 40.7%), and CCVVC syllables (n = 6, 22.2%) appeared in monosyllabic words. Two sub-structures CVCC (n = 17, 54.8%) and CVVC (n = 17, 54.8%) appeared in disyllabic words. The majority of superheavy syllables in disyllabic words were in stressed final environments (n = 11, 68.8%). However, this age group used unstressed non–final CVCC syllables (n = 5, 31.2%). Figure 4.85 demonstrates an unstressed non–final CVCC syllable in the word /ʔurtqaal/ (‘orange’, CVCC.CVVC, Speaker N8E). Also, unstressed final CVVC syllables were observed. As for multisyllabic words, only CVVC syllables were observed, which appeared in stressed final and non–final environments.



/ʔurtqaal/ Speaker N8E

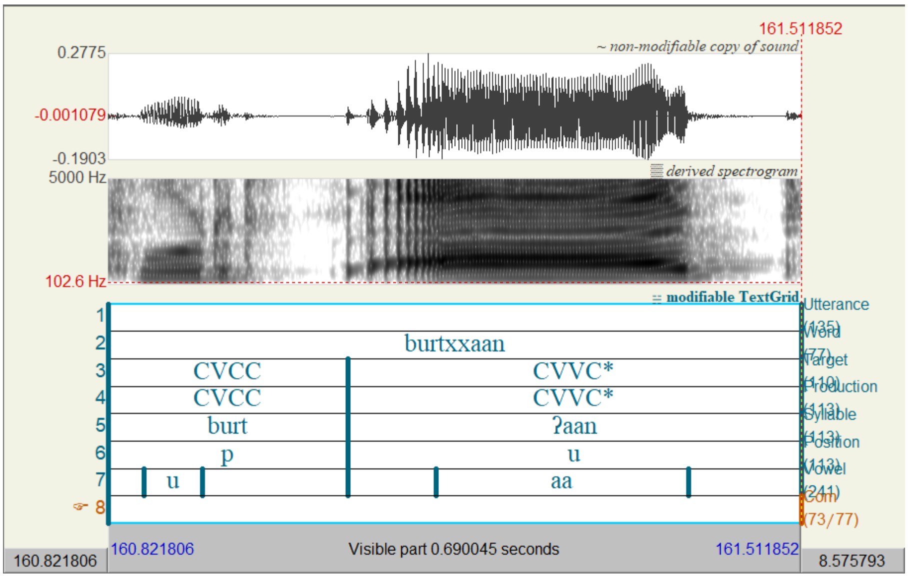
Figure 4.85: Spectrogram of a superheavy syllable in age group 52–58

Age group 59–65

At 59–65 months, the consistency of superheavy syllable production across the word lengths was attested. Superheavy syllables were the most frequent in monosyllabic words (n = 14, 63.6%), followed by disyllabic words (n = 6, 27.3%), while they were the least frequent in multisyllabic words (n = 2, 9.1%). Three sub-structures were produced in monosyllabic words, including CVCC (n = 5, 35.7%), CVVC (n = 7, 50%), and CCVVC (n = 2, 14.3%). Participants produced only CVVC syllables in disyllabic words appearing in stressed final and unstressed final environments. Likewise, only CVVC syllables appeared in multisyllabic words and were produced in the stressed final environment.

Age group 66–72

By 66–72 months, superheavy syllable production patterns are similar to those found in the younger age groups. Superheavy syllables were the most frequent in monosyllabic words (n = 31, 54.4%), and there was a notable expansion in the frequency of these syllables in disyllabic (n = 20, 35.1%) and multisyllabic words (n = 6, 10.5%) was attested. Three sub-structures appeared in monosyllabic words including CVCC (n = 10, 32.3%), CVVC (n = 13, 41.9%), and CCVVC (n = 8, 25.8%). Disyllabic words had two sub-structures including CVCC (n = 18, 90%) and CVVC syllables (n = 2, 10%). The production patterns included stressed final CVVC syllables, unstressed final CVVC syllables, and unstressed non–final CVCC syllables. Figure 4.86 demonstrates an unstressed non–final CVCC syllable in the word /burtʔaan/ (‘orang’, CVCC.CVVC, Speaker MMA). As for multisyllabic words, only CVVCs were attested in the stressed final environment.



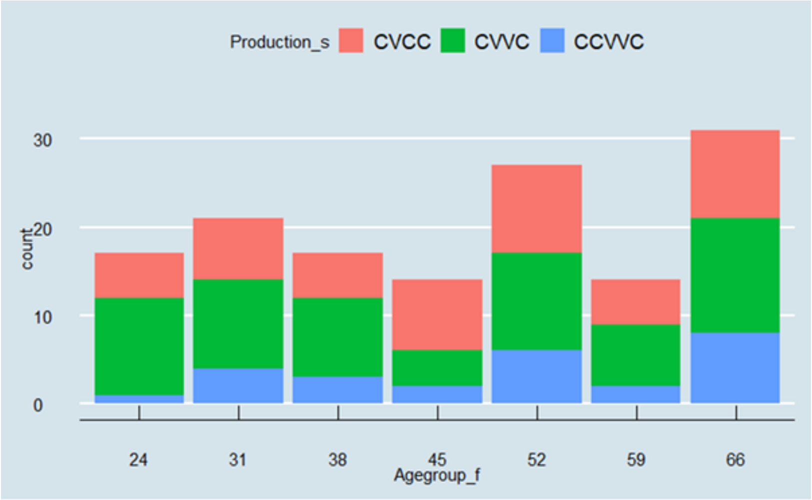
/burtʔaan/ Speaker MMA

Figure 4.86: Spectrogram of a superheavy syllable in age group 66–72

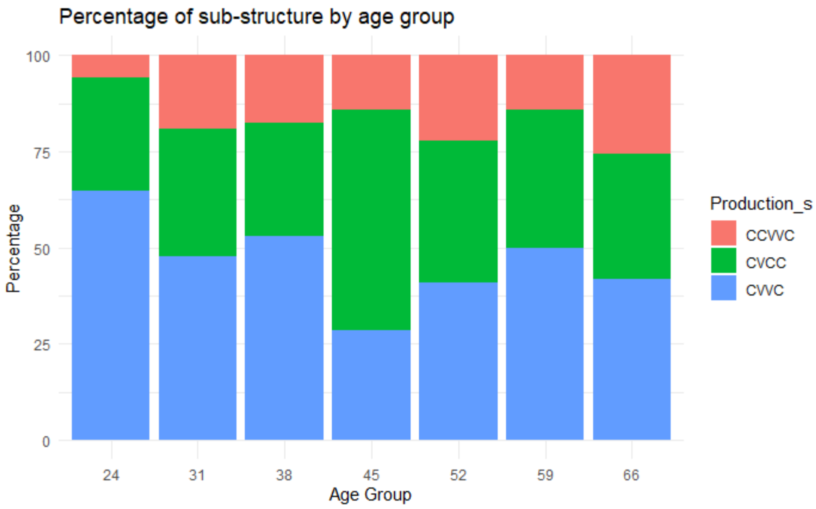
4.4.3.2.Bayesian Models and Descriptive Statistics

Superheavy Syllables in Monosyllabic Words

Figure 4.87/a shows the frequency distribution of superheavy sub-structures in monosyllabic words across the age groups, ranging from 24–30 months to 66–72 months. The frequency of superheavy syllables generally increases with age. All groups produced a total of three sub-structures, including CVCC, CVVC, and CCVVC. Moreover, Figure 4.87/b shows the percentages of sub-structure production across the age groups, with CVVC being the most frequent, followed by CVCC, and CCVVC being the least frequent.



(a) Frequency



(b) Percentage

Figure 4.87: Frequency and percentage of superheavy syllables in monosyllabic words

Syllable duration

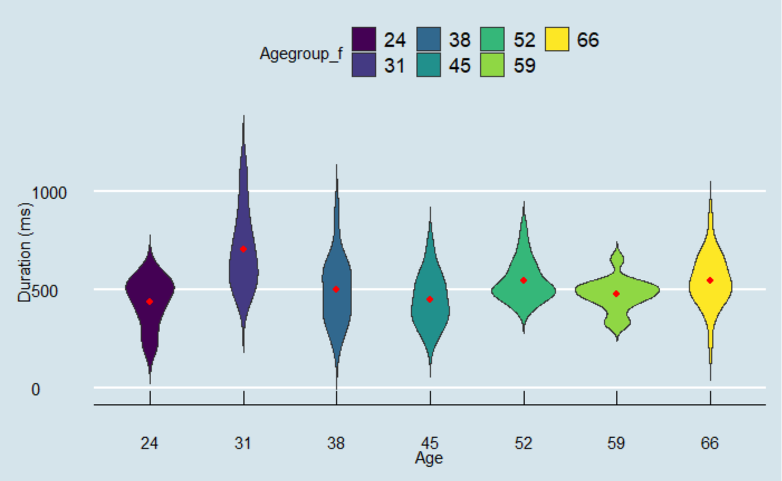
For superheavy syllables in monosyllabic words, a total of 141 syllables were analysed. For the syllable duration model, the Bayesian analysis was performed using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Since superheavy syllables in monosyllabic words only appear in stressed final environments, the model was fitted only to determine the effect of age group and sub-structure on the duration. The interaction between age group and sub-structure was also added. Table 4.31 summarizes the Bayesian model output for superheavy syllables in monosyllabic words.

Table 4.31: Bayesian model output summary for syllable duration in monosyllabic words

Population–Level Effects	Estimate	Est. Error	l–95% CI	u–95% CI	Bulk_ESS	Tail_ESS
Intercept	505.01	51.38	401.77	604.78	9570	15352
Sub-structure (CVVC)	–1.42	51.91	–101.91	104.00	12849	17319
Sub-structure (CCVVC)	277.89	80.21	120.88	437.31	13448	18730
Age group	–6.33	13.46	–31.62	21.54	7534	12818
Age group: Sub-structure (CVVC)	0.50	12.37	–24.49	24.36	11261	16946
Age group: Sub-structure (CCVVC)	–27.27	15.22	–57.19	2.84	12335	17533

Age Group

The model output indicates that age group is not a strong predictor for syllable duration (β = –6.33, CI [–31.62, 21.54]). Figure 4.88 demonstrates that the mean syllable duration in the youngest age group is 439.1 ms (SD = 122.2), then it increases, reaching 702.3 ms for 31–37 months (SD = 194.8). However, the mean durations decrease by 38–44 months, and they fluctuate across the age groups, with the oldest child group averaging 545.6 ms (SD = 132.4).



Syllable duration

Figure 4.88: The distribution and the mean syllable duration (ms) for age group in monosyllabic words

Sub-Structure

The model output indicates that sub-structure is a strong predictor for syllable duration for (CCVVC, β = 277.89, CI [120.88, 437.31]), but not for (CVVC, β = –1.42, CI [–101.91, 104.00]). CCVVC syllables have the highest mean syllable durations (Mean = 676.0 ms, SD = 170.9), followed by CVCC syllables (Mean = 505.2 ms, SD = 128.8), while CVVC syllables are the shortest (Mean = 498.6 ms, SD = 141.9). CCVVC syllables are approximately 1.4 times longer than CVVC syllables, while CVCC syllables are only one time longer than CVVC syllables.

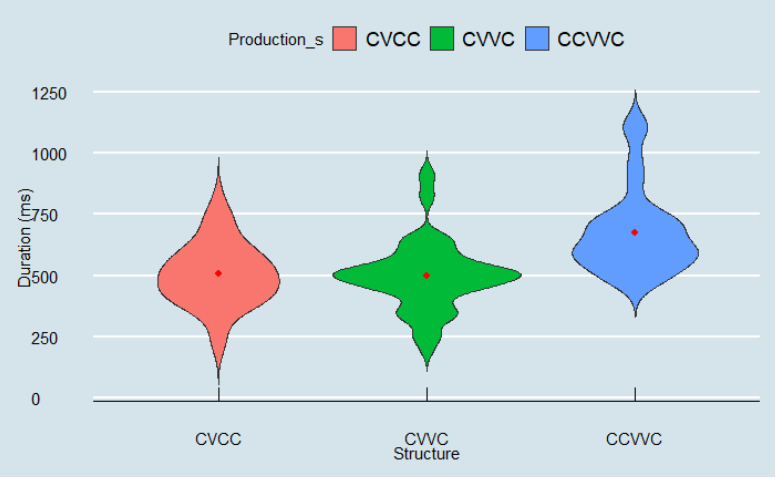


Figure 4.89: The distributions and the mean syllable durations (ms) for sub-structure in monosyllabic words

Interaction between Age Group and Syllable Structure

As for the interaction between age group and sub-structure, the model output suggests that this interaction is not a strong predictor for syllable duration (Age-CVVC, β = .50, CI [–24.49, 24.36), (Age-CCVVC, β = –27.27, CI [–57.19, 2.84]). Figure 4.90 shows that the syllable duration of CVCC and CVVC syllables remain consistent across the age groups, where the difference in the median durations is negligible. Nevertheless, the syllable duration of CCVVC syllables remarkably decreases with age.

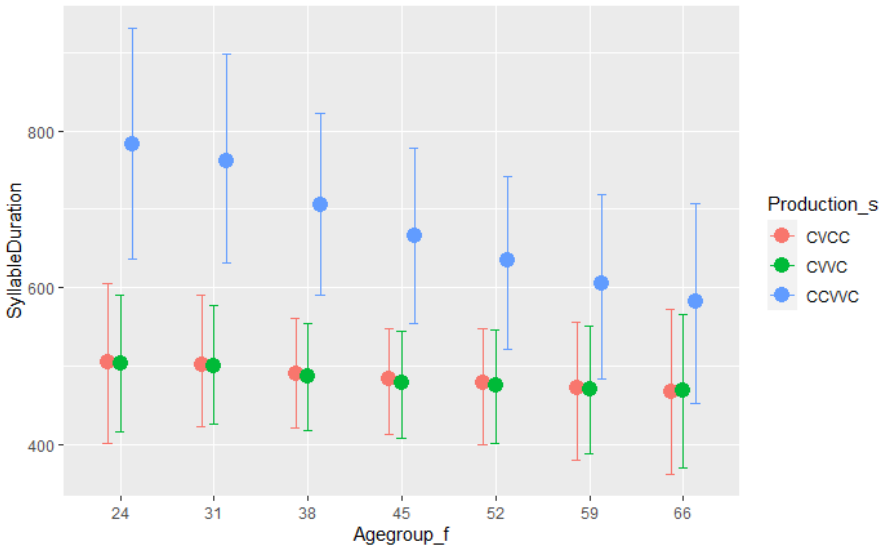


Figure 4.90: Posterior predictive plot for syllable duration for the interaction between age group and sub-structure in monosyllabic words

Vowel Duration

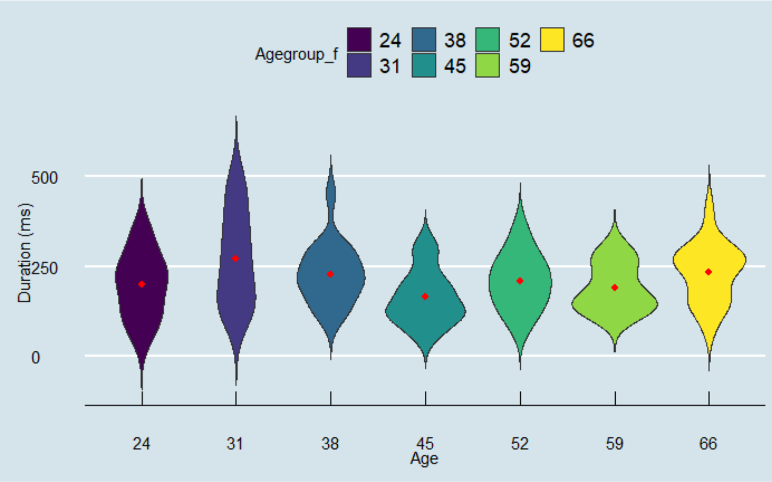
As for the vowel duration model, the Bayesian analysis was performed on a total of 141 vowels using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Since superheavy syllables in monosyllabic words only appear in stressed final environments, the model was fitted only to determine the effect of age group and sub-structure on the duration. Table 4.32 summarizes the Bayesian model output for superheavy syllables in monosyllabic words.

Table 4.32: Bayesian model output summary for vowel duration in monosyllabic words

Population–Level Effects	Estimate	Est. Error	l–95% CI	u–95% CI	Bulk_ESS	Tail_ESS
Intercept	123.88	24.14	74.16	170.04	13857	19188
Sub-structure (CVVC)	129.80	26.81	77.26	183.10	16556	20855
Sub-structure (CCVVC)	227.37	42.44	144.72	311.18	17310	20560
Age group	3.80	6.32	–8.33	16.33	12357	19108
Age group: Sub-structure (CVVC)	–3.55	6.78	–16.51	10.26	13063	20571
Age group: Sub-structure (CCVVC)	–24.15	8.38	–40.21	–7.22	15927	21157

Age Group

The model output suggests that age group is not a strong predictor for vowel duration (β = 3.80, CI [–8.33, 16.33]). Figure 4.91 shows that vowels in superheavy syllables slightly decrease with age. The youngest age group has a mean vowel duration of 200.3 ms (SD = 91.4), followed by an increase in age group 31–37 months (Mean = 272.5 ms, SD = 128.8). By age group 38–44 months, the durations demonstrate a general declining trend, reaching the 66–72 months age group mean (Mean = 232.9 ms, SD = 81.6).

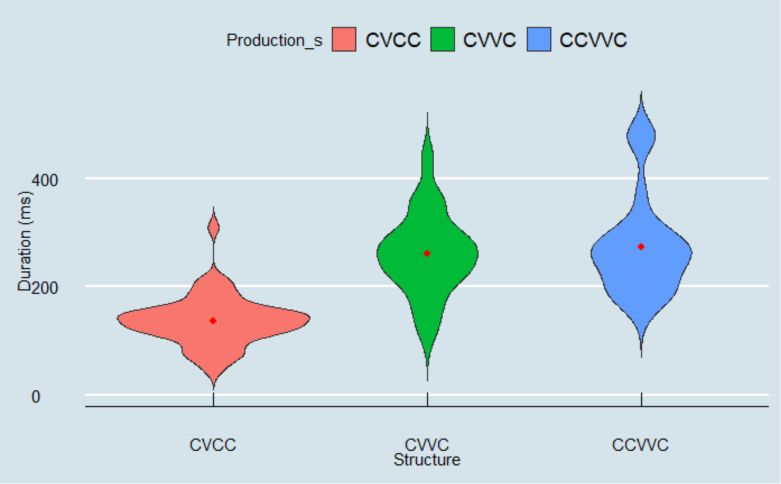


Vowel duration

Figure 4.91: The distribution and mean vowel duration (ms) for age group in monosyllabic words

Sub-Structure

As for sub-structure, the model output indicates that it is a strong predictor for vowel duration for (CVVC, β = 129.80, CI [77.26, 183.10]), (CCVVC, β = 227.37, CI [144.72, 311.18]). Vowels in CCVVC syllables are the longest (Mean = 273.5 ms, SD = 90.6), followed by vowels in CVVC syllables (Mean = 260.7 ms, SD = 75.2), and vowels in CVCC syllables are the shortest (Mean = 136.6 ms, SD = 45.3). On average, vowels in CCVVC syllables are two times longer than in CVCC syllables, while vowels in CVVC syllables are 1.9 times longer than in CVCC syllables.



Vowel duration

Figure 4.92:The distribution and the mean vowel duration (ms) for sub-structure in monosyllabic words

Interaction between Age Group and Sub-Structure

The interaction between age group and sub-structure is a strong predictor for vowel duration (Age-CCVVC, β = –24.15, CI [–40.21, –7.22), but not for (Age-CVVC, β = –3.55, CI [–16.51, 10.26]). Syllable duration of CVCC and CVVC syllables remain consistent across the age groups, where the difference in the median durations is negligible (Figure 4.93). Nevertheless, the syllable duration of CCVVC syllables remarkably decreases with age. From age group 24–30 months to 45–51 months, vowels in CVCC syllables are the shortest, followed by vowels in CVVC syllables, while vowels in CCVVC syllables are the longest. Nonetheless, by 52–58 months, vowels in CCVVC syllables become shorter than those in CVVC syllables.

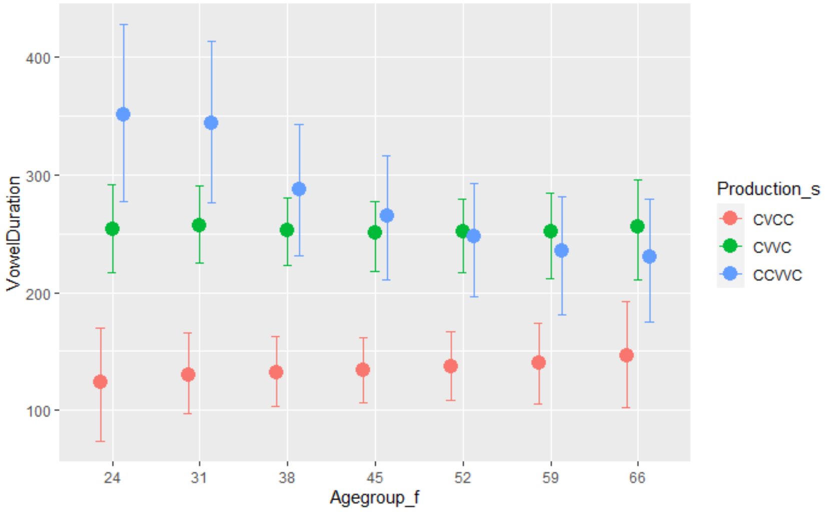
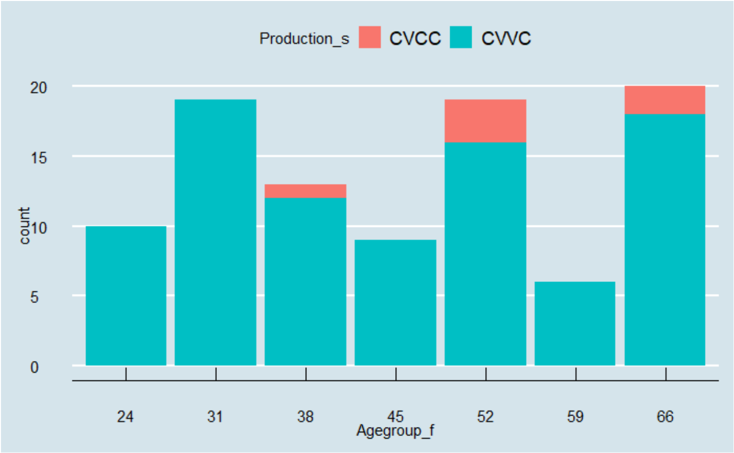


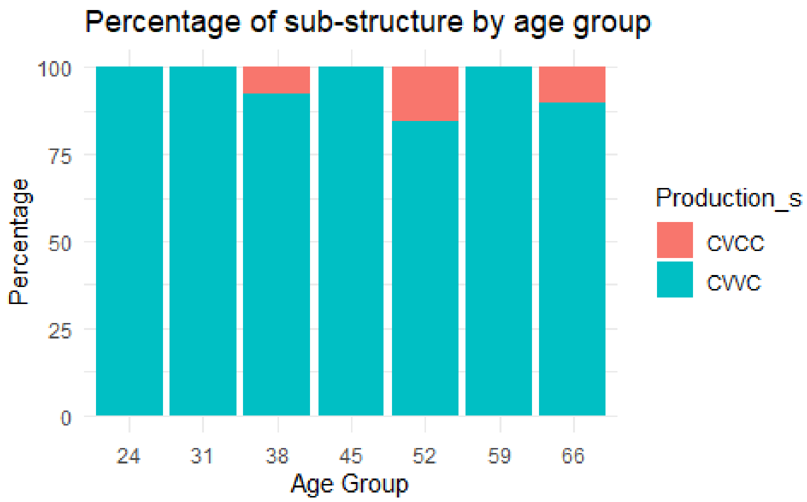
Figure 4.93: Posterior predictive plot for vowel durations for the interaction between age group and sub-structure monosyllabic words

Superheavy Syllables in Disyllabic and Multisyllabic Words

Figure 4.94 demonstrates the frequency distributions of superheavy sub-structures in disyllabic words. CVVC syllables are consistently produced across the age groups, while CVCC syllables are produced in age groups 38–44, 52–58, and 66–72 months. CVVC syllables are the most frequently produced in disyllabic words.



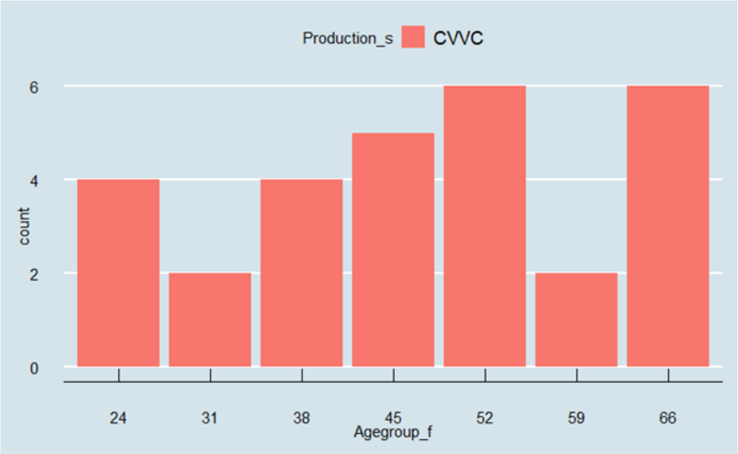
(a) Frequency



(b) Percentage

Figure 4.94: Frequency and percentage of superheavy syllables in disyllabic words

Figure 4.95 displays the frequency distribution of CVVC syllables as the only sub-structure observed in multisyllabic words across the age groups.



(a)

Figure 4.95: Frequency of superheavy syllables in multisyllabic words

Syllable Duration

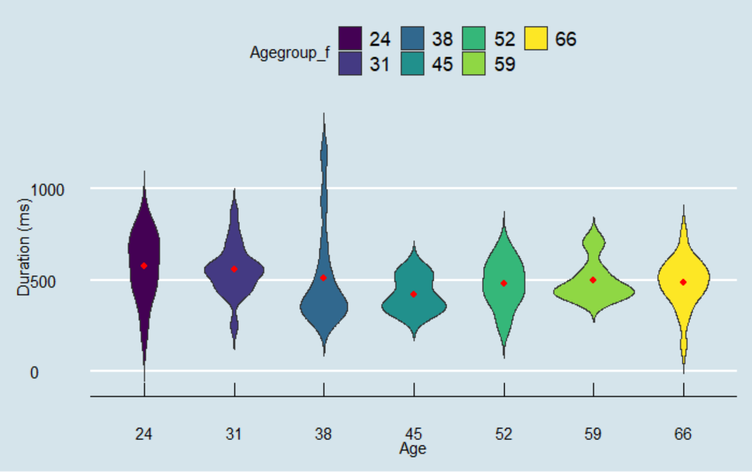
For superheavy syllables in disyllabic (n = 96) and multisyllabic words (n = 29), a total of 125 syllables were analysed. Predictors such as age group, syllable position, stress, and sub-structure were assigned independent variables to investigate their effects on superheavy syllable duration. For the syllable duration model, the Bayesian analysis was performed using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Table 4.33 summarizes the Bayesian model output for superheavy syllables in disyllabic and multisyllabic words.

Table 4.33: Bayesian model output summary for syllable duration in di/multisyllabic words

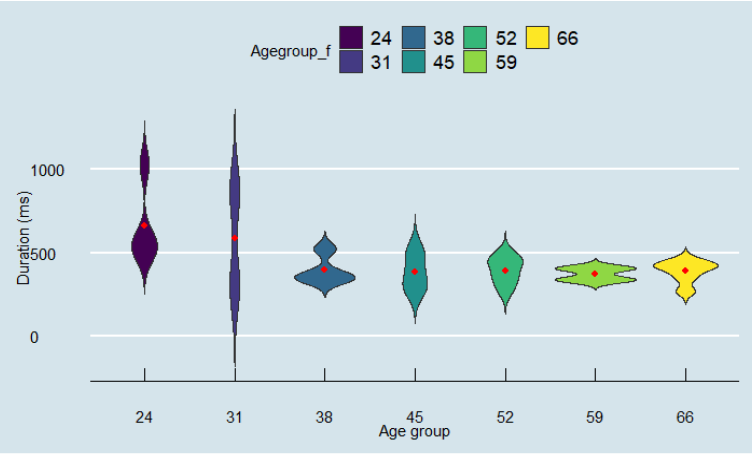
Population–Level Effects	**Estimate**	**Est. Error**	**l–95% CI**	**u–95% CI**	**Bulk_ESS**	**Tail_ESS**
Intercept	370.73	79.05	216.49	527.56	32206	27652
Sub-structure (CVVC)	–51.59	77.27	–205.27	100.01	48872	26201
Stress (Stressed)	–22.69	36.14	–91.67	50.65	38233	24600
Position (Final)	240.24	70.98	101.73	380.58	49176	26699
Age group	–15.44	10.52	–35.23	6.71	19867	19548

Age Group

The model output indicates that age group is not a strong predictor for syllable duration (β = –15.44, CI [–35.23, 6.71). For superheavy syllables in disyllabic words, Figure 4.96/a shows a decrease in durations until 45-51 months of age where an increase is evident (e.g., 24-31 months, Mean = 576.8 ms, SD = 173.5; 45-51 months, Mean = 419.5 ms, SD = 92.3; 66-72 months, Mean = 484.9 ms, SD = 134.2). As for multisyllabic words, mean superheavy syllable durations decrease with age (Figure 4.96/b). The youngest age group has the highest mean duration (Mean = 660.6 ms, SD = 244.4), followed by a decrease in age group 31–37 (Mean = 584.0 ms, SD = 395.9). By age group 38–44 months and older, the mean syllabic durations are around 350 ms and 395 ms. The mean syllable duration in the oldest age group, 388.7 ms (SD = 66.9), is approximately 1.7 times shorter than the mean duration of the youngest age group.



(a) Superheavy syllables in disyllabic words

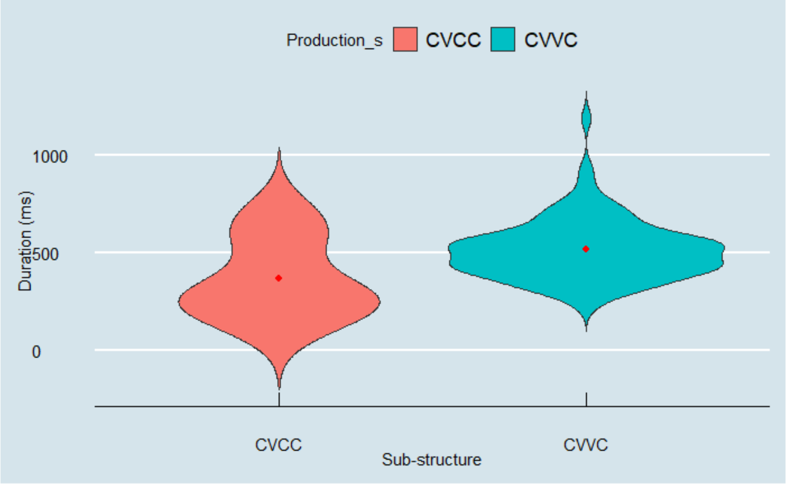


(b) Superheavy syllables in multisyllabic words

Figure 4.96: The distributions and the mean syllable durations (ms) for age group in disyllabic and multisyllabic words

Sub-Structure

As for sub-structure, the model output suggests it is not a strong predictor for syllable duration (β = –51.59, CI [–205.27, 100.01). For disyllabic words, Figure 4.97 demonstrates that CVVC syllables (Mean = 514.8 ms, SD = 210.4) are longer than CVCC syllables (Mean = 367.6 ms, SD = 153.7) by 1.4 times on average. Only CVVC syllables were produced for multisyllabic words.

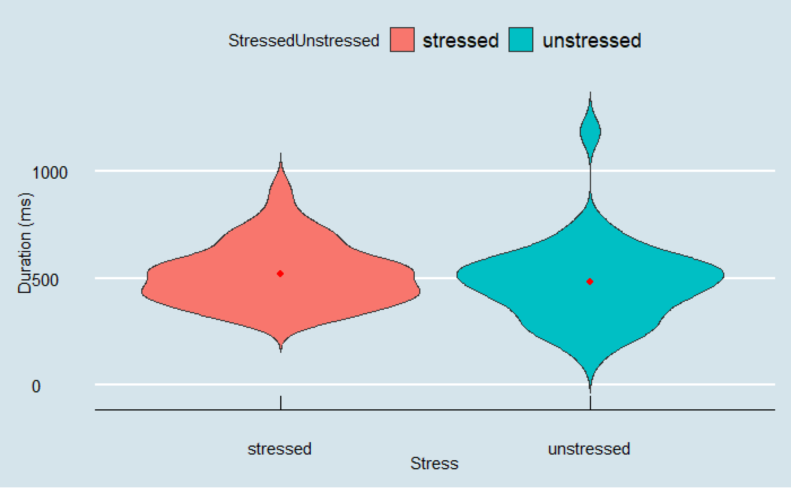


Superheavy syllables in disyllabic words

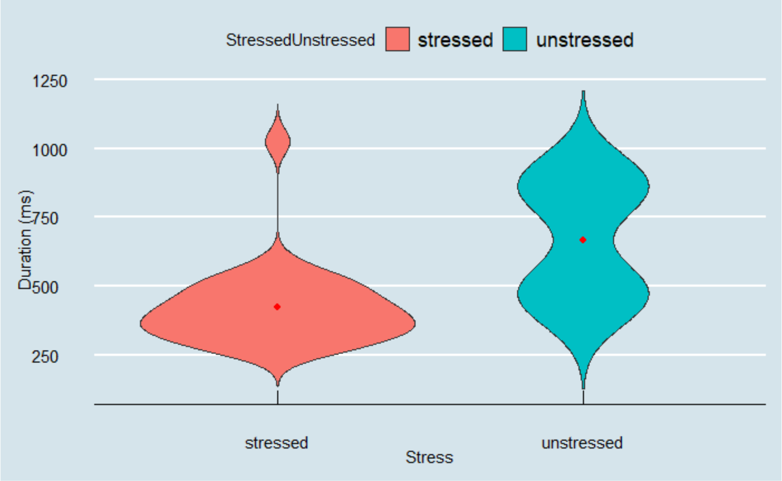
Figure 4.97: The distribution and the mean syllable duration (ms) for sub-structure in disyllabic words

Stress

Additionally, stress assignment is not a strong predictor for syllable duration ($\beta = -22.69$, CI [–91.67, 50.65). Figure 4.98/a shows that stressed superheavy syllables in disyllabic words (Mean = 518.1 ms, SD = 142.1) are approximately 1.1 times longer than unstressed syllables (Mean = 481.8 ms, SD = 190.8). However, stress assignment does not affect syllabic durations of superheavy syllables in multisyllabic words (Figure 4.98 /b). Unstressed syllables (Mean = 667.7 ms, SD = 277.6) are longer than stressed syllables (Mean = 423.3 ms, SD = 147.9). On average, unstressed syllables are reported to be 1.6 times longer than their counterparts.



(a) Superheavy syllables in disyllabic words

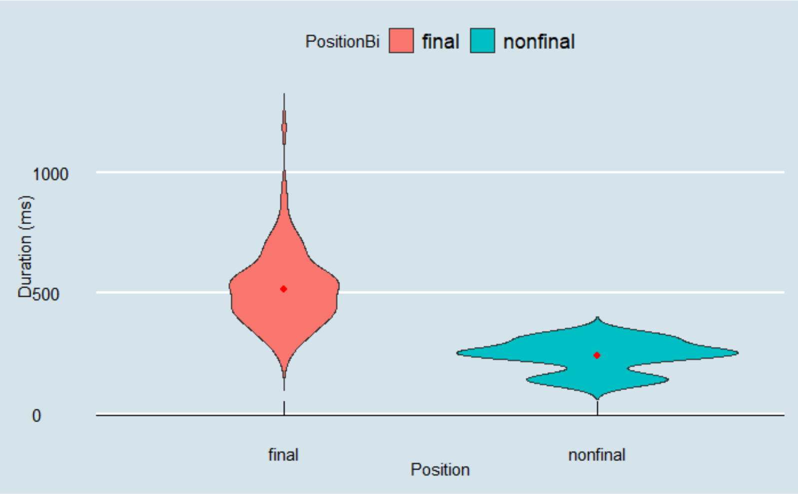


(b) Superheavy syllables in multisyllabic words

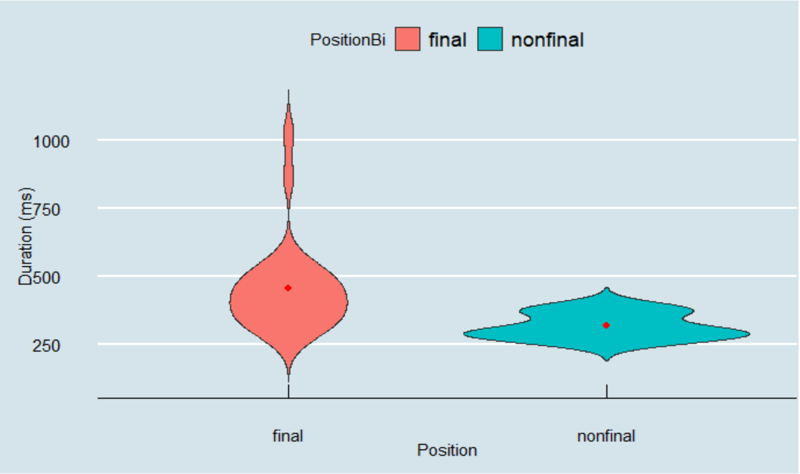
Figure 4.98: The distributions and the mean syllable durations (ms) for stress in disyllabic and multisyllabic words

Syllable Position

Moreover, the model output suggests that syllable position is a strong predictor for syllable duration (β = 240.24, CI [101.73, 380.58). Word final lengthening is observed in superheavy syllables produced in disyllabic words (Figure 4.99 /a). Word final syllables (Mean = 517.2 ms, SD = 153.2) are longer than word non-final syllables (Mean = 240.3 ms, SD = 72.8) by 2.2 times on average. Figure 4.99/b shows that in multisyllabic words, syllables in the word-final position (Mean = 454.2 ms, SD = 167.4) are 1.4 times longer than syllables in word non–final positions (Mean = 317.8 ms, SD = 53.5).



(a) Superheavy syllables in disyllabic words



(b) Superheavy syllables in multisyllabic words

Figure 4.99: The distributions and the mean syllable durations (ms) for syllable position in disyllabic and multisyllabic words

Vowel Duration

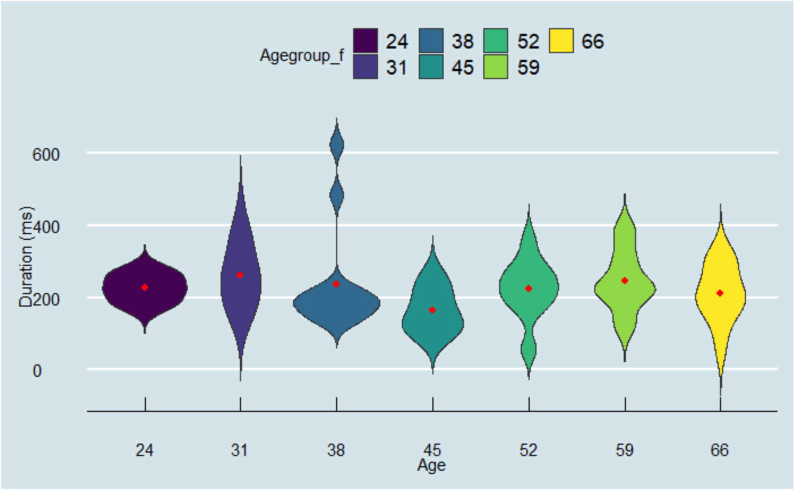
As for vowel durations, a total of 125 tokens were analysed. The Bayesian analysis was performed using four chains running for 10,000 iterations and a warmup period of 2000 iterations, yielding 32,000 post-warmup draws. Table 4.34 summarizes the Bayesian model output for superheavy syllables in disyllabic and multisyllabic words.

Table 4.34: Bayesian model output summary for vowel duration in di/multisyllabic words

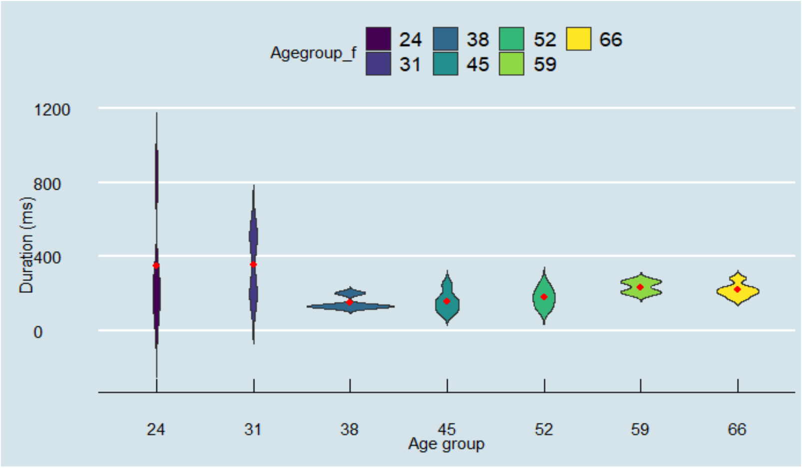
Population–Level Effects	**Estimate**	**Est. Error**	**l–95% CI**	**u–95% CI**	**Bulk_ESS**	**Tail_ESS**
Intercept	87.83	51.74	–13.95	189.12	35640	26570
Sub-structure (CVVC)	62.90	53.33	–42.16	168.87	47175	26203
Stress (Stressed)	–3.91	24.27	–53.13	42.48	37790	21061
Position (Final)	97.10	48.87	1.29	192.52	46387	25894
Age group	–4.44	5.95	–15.70	7.88	23483	20595

Age Group

The model output suggests that age group is not a strong predictor for vowel duration (β = -4.44, CI [-15.70, 7.88). Figure 4.100/a shows that the youngest age group, 24–30 months, has a mean duration of 225.9 ms (SD = 38.4), then the durations peak at 38–44 months (Mean = 261.2 ms, SD = 92.3). At 45–51 months, the mean durations slightly decrease, reaching 164.5 ms (SD = 60.6). By 52–58 months (Mean = 224.1 ms, SD =80.9), the durations increase, and they fluctuate between 224.1 ms and 212.4 ms in the oldest three age groups. On the other hand, the superheavy vowel durations in multisyllabic words show a U–shaped pattern (Figure 4.100/b). The two youngest age groups have mean durations of 350.4 ms (SD = 329.9), and 353.6 ms (SD = 221.6), respectively. This is followed by a sharp decrease by age 38–44 months (Mean = 148.9 ms, SD = 35.9) until 52–58 months (Mean = 178.1 ms, SD = 49.7). Subsequently, the mean vowel durations increase by 59–65 months (Mean = 235.4 ms, SD = 40.2), while the oldest age group has a mean duration of 218.6 ms (SD = 37.5).



(a) Superheavy syllables in disyllabic words

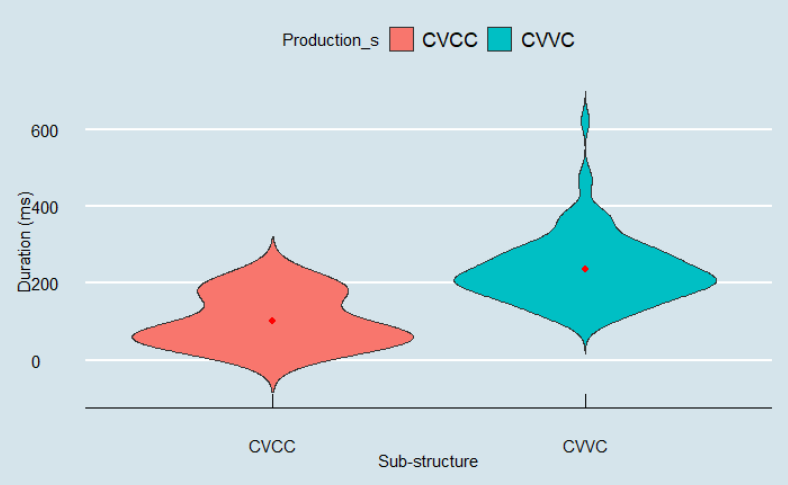


(b) Superheavy syllables in multisyllabic words

Figure 4.100: The distributions and the mean vowel durations (ms) for age group in disyllabic and multisyllabic words

Sub-Structure

The model output shows that sub-structure is not a strong predictor for vowel duration (β = 62.90, CI [–42.16, 168.87). For disyllabic words, Figure 4.101 shows that vowels in CVVC syllables (Mean = 235.1 ms, SD = 86.7) are longer than vowels in CVCC syllables (Mean = 101.9 ms, SD = 66.7) by approximately 2.3 times.

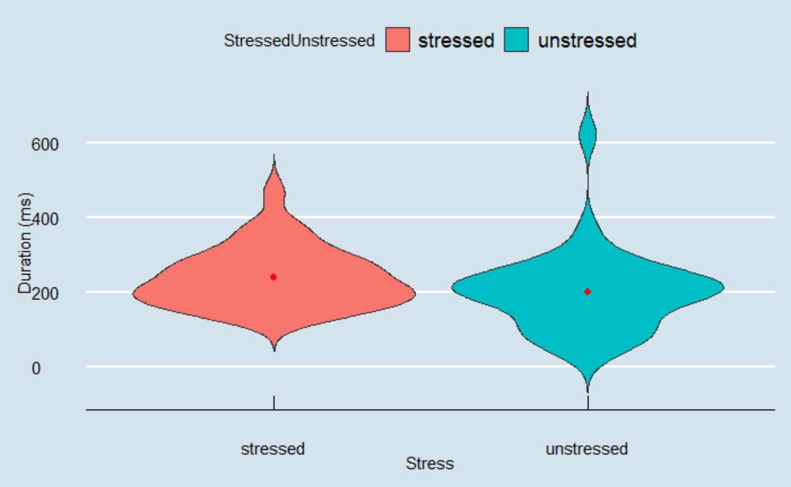


Superheavy syllables in disyllabic words

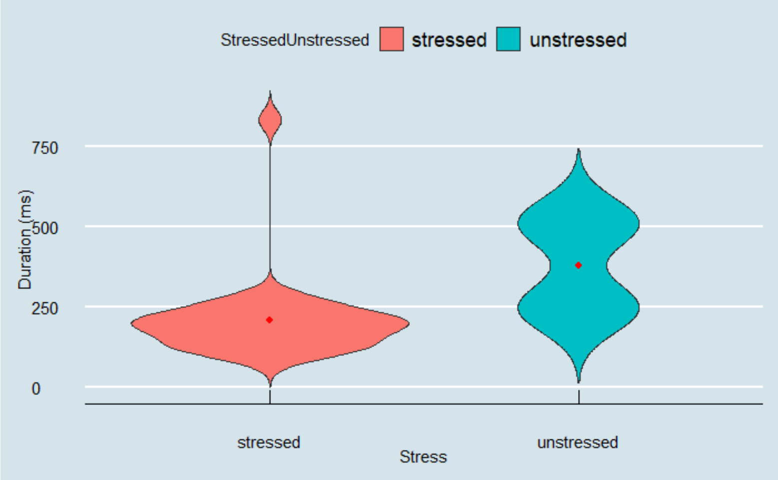
Figure 4.101: The distributions and the mean vowel duration (ms) for sub-structure in disyllabic words

Stress

As for stress, the model output indicates that stress assignment is not a strong predictor for vowel duration (β = –3.91, CI [–53.13, 42.48). For superheavy syllables in disyllabic words, Figure 4.102/a demonstrates that vowels in stressed syllables (Mean = 240.0 ms, SD = 79.1) are, on average, 1.2 times longer than their counterparts (Mean = 201.6 ms, SD = 107.7). As for superheavy syllables in multisyllabic words, Figure 4.102/b shows that mean vowel durations in unstressed syllables (Mean = 378.1 ms, SD = 186.9ms) are 1.8 times longer than durations in stressed syllables (Mean = 207.2 ms, SD = 135.5).



(a) Superheavy syllables in disyllabic words



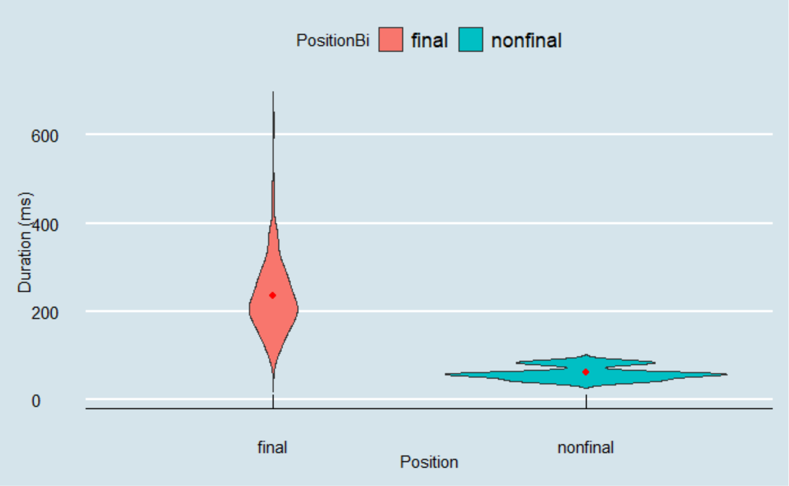
(b) Superheavy syllables in multisyllabic words

Figure 4.102: The distributions and the mean vowel durations (ms) for stress in disyllabic and multisyllabic words

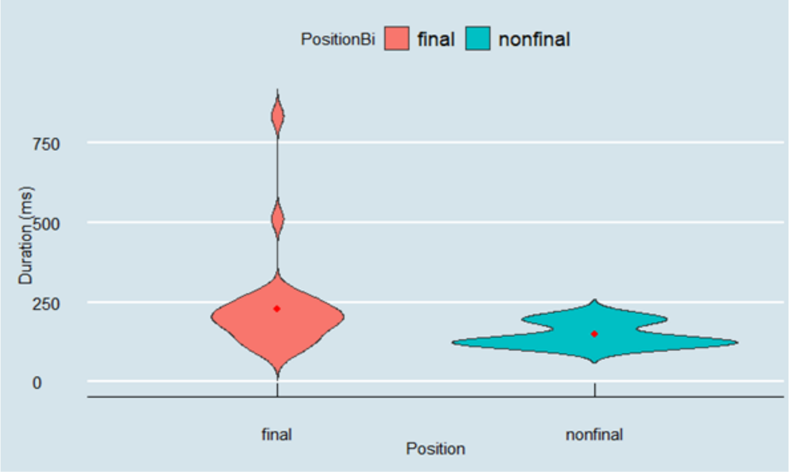
Syllable Position

Furthermore, the model output indicates that syllable position is a strong predictor for vowel duration (β = 97.10, CI [1.29, 192.52). In superheavy syllables appearing in disyllabic words, Figure 4.103/a shows that vowels in word-final syllables (Mean = 234.1 ms, SD = 86.0) are remarkably longer than vowels in word non–final syllables (Mean = 59.9 ms, SD = 17.2). Approximately, vowels in word-final positions are 3.9 times longer than their counterparts. As for multisyllabic words (Figure 4.103/b), vowels in superheavy word-final syllables (Mean =

227.2 ms, SD = 147.8) are longer than vowels in word non–final positions (Mean = 148.1 ms, SD = 42.4). Approximately, word-final vowels are 1.5 times longer than their non-final counterparts.



(a) Superheavy syllables in disyllabic words



(b) Superheavy syllables in multisyllabic words

Figure 4.103: The distributions and the mean vowel durations (ms) for syllable position in disyllabic and multisyllabic words

4.4.4. Summary

Table 4.35 summarizes the production of superheavy syllables across the tasks based on age group, word length, sub-structure, and emerging new forms.

Table 4.35: Summary of superheavy syllable productions across tasks (Mono=monosyllabic; Di=disyllabic; Multi=multisyllabic).

Age group	Criteria	ST	RT	PT
24–30	Word length	Mono, di, multi	Mono, di	Mono, di, multi
	Sub-structure	Mono: CVCC, CVVC Di: CVVC Multi: CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC Multi: CVVC
	New forms	-	-	Unstressed final CVVC in di
31–37	Word length	Mono, di	Mono, di	Mono, di, multi
	Sub-structure	Mono: CVVC, CVCC Di: CVVC, CVCC	Mono: CVCC, CVVC, CCVVC Di: CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC Multi: CVVC
	New forms	Unstressed final CVCC in di	-	Unstressed final CVVC in multi
38–44	Word length	Mono, di, multi	Mono, di, multi	Mono, di, multi
	Sub-structure	Mono: CVCC, CVVC, CCVVC Di: CVCC, CVVC Multi: CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC, CVCC Multi: CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC, CVCC Multi: CVVC
	New forms	CCVVC in mono-stressed Nonfinal CVVC in multi Nonfinal unstressed in di	CVCC in di CVVC in multi	CVCC in di
45–51	Word length	Mono, di	Mono, di, multi	Mono, di, multi
	Sub-structure	Mono: CVCC, CVVC Di: CVCC, CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC, CVCC	Mono: CVCC, CVVC, CCVVC Di: CVVC

			Multi: CVVC	Multi: CVVC
	New forms	-	Unstressed final CVVC	-
52–58	Word length	Mono, di, multi	Mono, di, multi	Mono, di, multi
	Sub-structure	Mono: CVCC, CVVC, CCVVC Di: CVCC, CVVC Multi: CVCC, CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC, CVCC Multi: CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC, CVCC Multi: CVVC
	New forms	Stressed nonfinal CVCC in multi	-	Unstressed nonfinal CVCC in di
59–65	Word length	Mono, di, multi	Mono, di, multi	Mono, di, multi
	Sub-structure	Mono: CVCC, CVVC, CCVVC Di: CVCC, CVVC Multi: CVCC, CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC, CVCC Multi: CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC Multi: CVVC
	New forms	-	-	-
66–72	Word length	Mono, di, multi	Mono, di, multi	Mono, di, multi
	Sub-structure	Mono: CVCC, CVVC, CCVVC Di: CVCC, CVVC Multi: CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC, CVCC Multi: CVVC	Mono: CVCC, CVVC, CCVVC Di: CVVC, CVCC Multi: CVVC
	New forms	-	-	-
Adults	Word length	Mono, di, multi	Mono, di, multi	-
	Sub-structure	Mono: CVCC, CVVC, CCVVC Di: CVCC, CVVC, CCVVC Multi: CVCC, CVVC	Mono: CVCC, CVVC, CCVVC, CCVCC Di: CVVC, CVCC, CCVVC Multi: CVVC	-
	New forms	CCVVC in di	CCVCC in mono CCVVC in di	-

The results show that superheavy syllables are produced more frequently in monosyllabic words, but with maturation, children start producing these syllables in disyllabic and multisyllabic words as the frequency of these word lengths increases over time. Superheavy syllables in monosyllabic and disyllabic words appear by 24–30 months and consistently by 38–44 months in multisyllabic words. More sub-structural varieties emerge with age, as older age groups demonstrate more variability and complexity in their superheavy productions. Superheavy syllables with coda clusters, such as CVCC syllables, emerge in age group 31–37 months, followed by superheavy syllables with onset clusters, such as CCVVC, in age group 38–44 months. CVVC syllables are the most commonly produced across all word lengths, CVVC syllables appear in all prosodic patterns first, followed by other sub-structures.

In ST, results reveal that the durations of superheavy syllables in monosyllabic words are longer in younger age groups and decrease around 45-51 months. The longest syllables are CCVVC syllables, while the shortest are CVCC syllables. Vowels in CVVC syllables are the longest and shortest in CVCC syllables. The Bayesian model output indicates that age group is the only strong predictor for syllable duration, while sub-structure and the interaction between age group and sub-structure are not. Sub-structure is the only strong predictor for vowel duration. For disyllabic words, younger children produce longer syllables and durations decrease by 45-51 months. Syllable and vowel durations are the longest in CVVC syllables and the shortest in CCVVC syllables. Word final lengthening is observed as word final syllables are 1.5 times longer than their counterparts, and word final vowels are 2.3 times longer than word non-final vowels. Stressed syllables are 1.9 times longer than unstressed ones, while vowels in stressed environments are 1.1 times longer than their counterparts. The Bayesian model output for superheavy syllables in di/multisyllabic words shows that age group, syllable position, and sub-structure are strong predictors, while stress is not.

Second, RT results show that superheavy syllables are most frequent in monosyllabic words and least frequent in multisyllabic words. Syllabic complexity and variability increase with age, with the youngest two age groups producing superheavy syllables in two-word lengths only. By 38-44 months, JA children start producing superheavy syllables in multisyllabic words. Younger age groups exhibit the longest productions, and durations decrease with age. Durations decrease at later age groups as word length increases. In monosyllabic words, CCVVC syllables are the longest, while CVVC syllables are the shortest. Vowels in CVVC are the longest, while they are the shortest in CCVVC syllables. In disyllabic words, syllable and vowel durations are the longest in CCVVC syllables and the shortest in CVCC syllables. Stress effects are evident, with stressed syllables being approximately 1.2 times longer than their counterparts, and vowels being 1.1 times longer in stressed environments. In multisyllabic words, only CVVC syllables in final stressed environments are produced.

Third, PT results demonstrate similar frequency patterns attested in ST and RT for monosyllabic and disyllabic words. Superheavy syllables are most frequently produced in monosyllabic words in CVVC syllables, with more sub-structural varieties appearing as age increases. Superheavy productions in disyllabic and multisyllabic words appear as early as 24-30 months of age. Unstressed final superheavy syllables emerge before stressed non-final ones. Syllable and vowel durations are the longest in CCVVC syllables and the shortest in CVCC syllables. Word final lengthening is evident, with final syllables being 2.2 times longer than their counterparts and vowels being 3.9 times longer than non-final ones. In multisyllabic words, word final syllables are 1.4 times longer than their counterparts, and on average, word final vowels are 1.5 times longer than word non-final ones. Stress marking is not evident in superheavy syllables in multisyllabic productions. The Bayesian model output shows that only sub-structure (CCVVC) is a strong predictor for syllable duration, while age group and the two-way interaction between age group and sub-structure are not.

The following table summarizes the Bayesian model outputs of target predictors in monosyllabic and di/multisyllabic words across the tasks.

Table 4.36: Summary of Bayesian model outputs in superheavy syllable productions (S: Strong, NS: Not strong)

Task	**ST**		**RT**		**PT**	
Duration	**Syllable**	**Vowel**	**Syllable**	**Vowel**	**Syllable**	**Vowel**
Mono						
Age group	S	NS	NS	S	NS	NS
Sub-structure	NS	S	S: CCVVC only	S: CVVC, CCVVC only	S: CCVVC only	S
Age x sub-structure CVVC	NS	NS	-	-	NS	NS
Age x sub-structure CCVVC	NS	NS	-	-	NS	S
Di/multi						
Age group	S	S	NS	NS	NS	NS
Stress	NS	NS	NS	NS	NS	NS
Syllable position	S	S	-	-	S	S
Sub-structure	NS	S: CVVC only	S: CCVVC only	S	NS	NS

The following is a summary of predictions for superheavy syllable productions and whether they were met or not based on the results.

a) Younger children are expected to produce longer superheavy syllables compared to older age groups ✓

b) Stress will affect superheavy syllables with stressed syllables exhibiting longer durations than unstressed ones ✓

c) Word-final lengthening will be observed in superheavy syllable productions, with final syllables being longer than non-final ones **X**

d) Superheavy syllables are bimoraic and not trimoraic as vowel shortening is anticipated to occur in non-final superheavy syllables ✓

4.5. Phonological Processes

Table 4.37 shows the phonological processes occurring in JA child forms. These include cases where child production forms do not match target adult forms across the tasks. The most frequent phonological processes are segment modification (i.e., metathesis, assimilation, gliding, stopping, backing, fronting, de-emphasis), vowel epenthesis, assimilation and gemination, syllable deletion, cluster reduction, weak syllable deletion, coda deletion, and syncope (Amayreh & Dyson 1998, Amayreh 2000, Alqattan 2015, Mashaqba et al., 2019). A total of 266 phonological processes were observed: 89 in ST, 128 in PT, and 55 in RT. ST had the highest accuracy rate with 89.8%, followed by PT with 87.6%, then RT with 86.3% (accuracy rate per task = number of mismatches / total number of syllables × 100).

Table 4.37: Phonological processes in JA child speech

Phonological process	**24–30**	**31–37**	**38–44**	**45–51**	**52–58**	**59–65**	**66–72**	**Total**
Segment modification	14	20	16	13	12	0	0	75
Vowel epenthesis	11	9	8	7	11	7	5	58
Assimilation and gemination	12	9	5	5	4	4	2	41
Syllable deletion	15	10	4	2	2	0	0	33
Cluster reduction	6	4	11	8	4	1	0	33
Weak syllable deletion	5	4	1	3	4	0	0	17
Coda deletion	2	1	1	0	0	0	0	4
Syncope	0	0	1	0	2	1	0	4
Total	65	57	47	38	39	13	7	266

The analysis shows that the number of phonological processes not matching target adult forms generally decreases with maturation. The percentage of phonological processes occurring across age groups is as follows: 24–31 months (18.4%), 31–37 months (15.0%), 38–44 months (11.9%), 45–51 months (6.4%), 52–58 months (9.6%), 59–65 months (5.5%), 66–72 months (4.0%) (Percentage of phonological processes = number syllables with processes / total number of syllables × 100). Results indicate that the three youngest age groups have the highest percentages of phonological processes deviating from the target form, yet the percentages decrease until 45–51 months. Remarkably, the 52–58 months age group shows a slight stabilization in the trend with a total of 39 occurrences compared to 38 occurrences in the previous group. However, a notable decrease is observed in the 59–65 months age group, where only 13 occurrences are observed. The oldest age group has the least occurrences with a total of seven processes only. Thus, with maturation, there is a general decreasing trend of the phonological processes' percentage.

Examining individual phonological processes, segment modification is the most frequent in the early stages appearing in the youngest age group and persisting until 52–58 months of age. Segment modification peaks at 20 occurrences in the 31–37 months age group, and gradually decreases to zero by 59–65 months, with no occurrences in the two oldest age groups. Vowel epenthesis shows a relatively consistent presence across all age groups with 11 occurrences in the youngest age group, followed by a decrease until 45–51 months. However, an increase in the number of occurrences is evident in age group 52–58 months, followed by another decrease in the oldest two age groups. The process of assimilation and gemination demonstrates a consistent occurrence across all age groups. The youngest age group has the highest number of occurrences, followed by a notable decline by 38–44 months, while the oldest age group has two occurrences only. Syllable deletion appeared in the younger age groups and persisted until 52–58 months with a declining trend. No occurrences of syllable deletion were evident in the oldest two age groups ranging from 59–65 to 66–72 months. Additionally, cluster reduction exhibits an irregular pattern, where the youngest age group has six occurrences and the following group has four. This was followed by a notable increase in age group 38–44 months, with 11 occurrences. By 45–51 months, the number of cluster reductions decreases until age group 59–65 months, while this process is not reported in the oldest age group. As for weak syllable deletion, the process did not appear in the two oldest age groups ranging from 59–65 to 66–72 months. In the younger age groups, from 24–30 months to 52–58, the process exhibits a decreasing trend from 24–30 months to 38–44 months, but it increases again by 45–51 months until 52–58 months. Coda deletion was not frequent, and it only appeared in the youngest three age groups ranging from 24–30 months to 38–44 months. Syncope was the least frequent process with four reported occurrences, and it only appeared in age groups 38–44, 52–58, and 59–65 months.

A further analysis was conducted to assess the occurrences based on syllable structure (light, heavy, superheavy), syllable position within a word (final, non-final), and lexical stress (stressed, unstressed). First, the data exhibits a developmental trend in the distribution of phonological processes among the age groups according to syllable structure (Figure 4.104). Overall, 97 phonological processes appeared in superheavy syllables, followed by 92 processes appearing in heavy syllables, while 77 processes appeared in light syllables. Children aged 24–30 and 31–37 months produce slightly more phonological processes in light and heavy syllables compared to superheavy syllables. However, a shift in the distributional pattern is attested, where in the older age groups, ranging from 45–51 to 66–72 months, there is a notable decrease in the number of processes occurring in light and heavy syllables, while the majority of processes occurred in superheavy syllables.

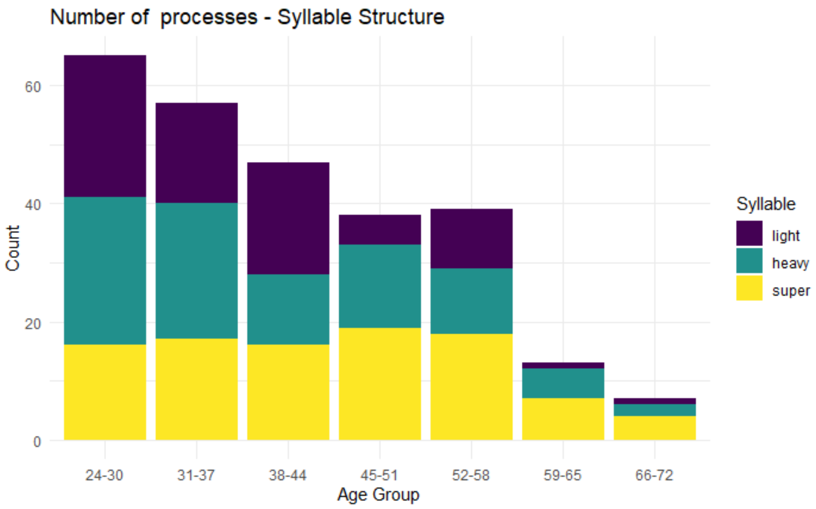


Figure 4.104: Number of processes – syllable structure

Second, the data distribution based on lexical stress shows that processes are more evident in stressed syllables across the age groups (Figure 4.105). A total of 156 phonological processes appeared in stressed syllables while 110 processes appeared in unstressed. However, deviating from the other groups, age group 38–44 demonstrates more processes in unstressed syllables compared to stressed ones.

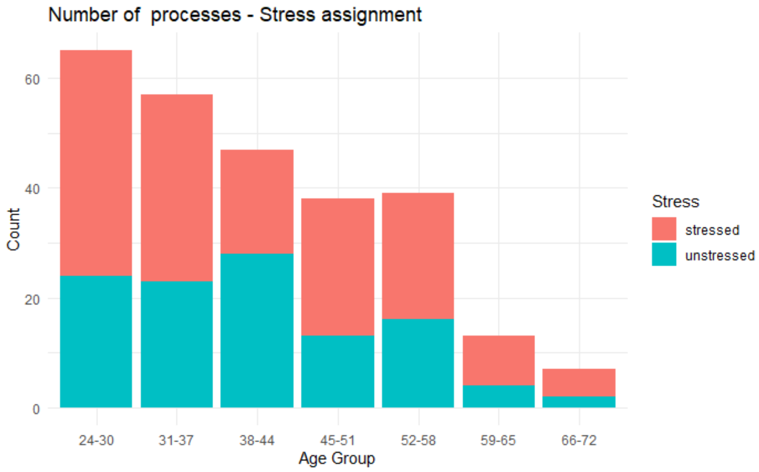


Figure 4.105: Number of processes – lexical stress

Third, across the age groups, slightly more phonological processes occurred in word final syllables, with a total of 138 processes compared to 128 processes in non-final ones. Higher occurrences of phonological processes in word nonfinal syllables compared to final ones are evident in age groups 24–30 months and 31–38 months (Figure 4.106). Nevertheless, age groups from 38–44 months to 66–72 months exhibit the tendency to produce more processes in word final syllables compared to non-final ones.

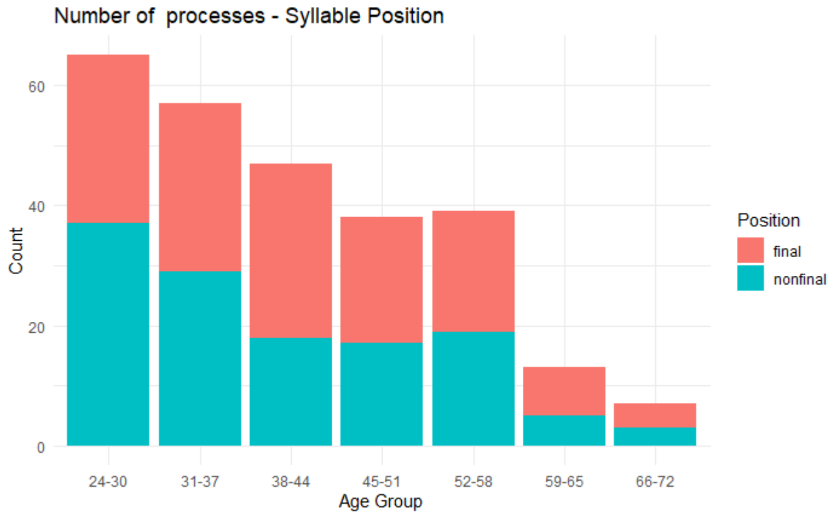


Figure 4.106: Number of processes – syllable position

Furthermore, the distributional trend of phonological processes in Figure 4.107 varies between younger and older age groups. In younger children aged 24–30 to 31–37 months, a greater frequency of processes is observed in stressed final syllables, followed by stressed nonfinal syllables. Nonetheless, by 38–44 months, while processes remain most prevalent in stressed final syllables, there is a noticeable decline in processes occurring in stressed nonfinal syllables. Processes found in unstressed final and stressed nonfinal syllables also decrease notably with age. By 59–65 months, no processes are observed in unstressed final syllables, with stressed final syllables continuing to exhibit the highest occurrence of processes.

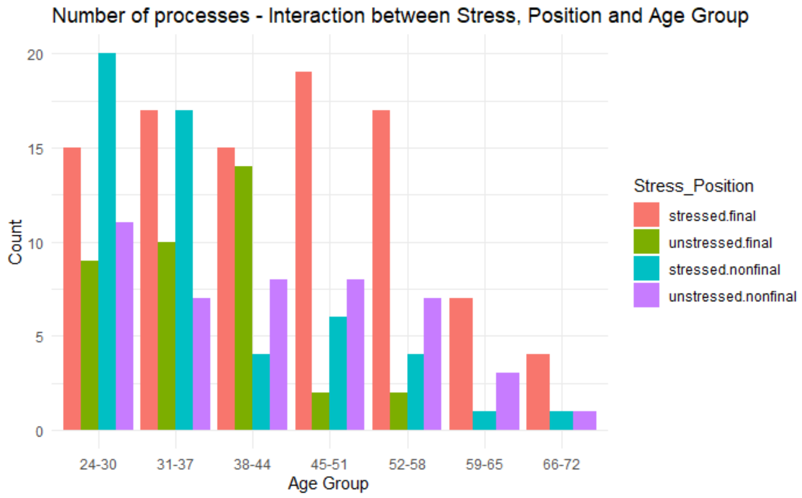


Figure 4.107: Number of processes – interaction between stress, position and age group

Summary

A total of 266 phonological processes deviating from the target adult forms were identified across the three tasks, including segment modification, vowel epenthesis, assimilation and gemination, syllable deletion, cluster reduction, weak syllable deletion, coda deletion, and syncope. The younger age groups exhibited a higher frequency of processes compared to older age groups. In the younger age groups, segment modification was the most frequent process, while only vowel epenthesis and assimilation/gemination appeared across all the age groups.

Further analysis examined the distribution of phonological processes according to syllable structure (light, heavy, superheavy), syllable position within a word (final, non-final), and lexical stress (stressed, unstressed). The data showed that phonological processes appeared more frequently in superheavy syllables. Stressed syllables and word-final syllables also exhibited a higher frequency of processes across the age groups, although younger age groups showed more processes in non-final syllables. The interaction between stress, syllable position, and age group highlighted a developmental trend where phonological processes in stressed final syllables persisted, while those in unstressed and non-final syllables declined with age.

5. Chapter Five: Discussion

Chapter 5. Discussion

Section 5.1 discusses the age group effect on durations and when values start exhibiting adult-like productions in JA child speech. Section 5.2 explores the relationship between lexical stress and durations, and assesses the feasibility of employing duration as an acoustic cue for stress marking. Section 5.3 discusses the word final lengthening and its pivotal role in shaping rhythmic and stress patterns. Section 5.4 explores syllable structure and its role in determining durations, in addition to discussing the trimoraicity and the bimoraicity constraint. Section 5.5 shows the development of superheavy syllables, highlighting the frequency distribution and the effect of age group, stress, syllable position within a word, and sub-structure on durations. The final section includes theoretical and clinical implications in addition to future research suggestions.

5.1.Age Group

The effect of age group on syllable and vowel durations was in line with the predictions (Section 2.3) that (1.a) a decreasing trend in durations would be evident, with younger age groups producing longer syllables compared to older age groups; and (1.b) the oldest age group would exhibit durations that are close to the adult ones, but the durations of the two groups would not intersect. This durational decreasing trend aligns with developmental theories suggesting that as children age, their articulatory control becomes more refined, leading to shorter productions (Green & Wilson 2006, Nip & Green 2013).

First, results showed that syllable and vowel durations were the longest in the youngest three age groups, ranging from 24–30 to 38–44 months, consistently across the tasks. This finding aligns with studies suggesting that younger children exhibited longer durations due to less developed articulatory control, and the broader understanding of universal language learning, where linguistic experience leads to shorter durations with maturation (Smith, Kenny &

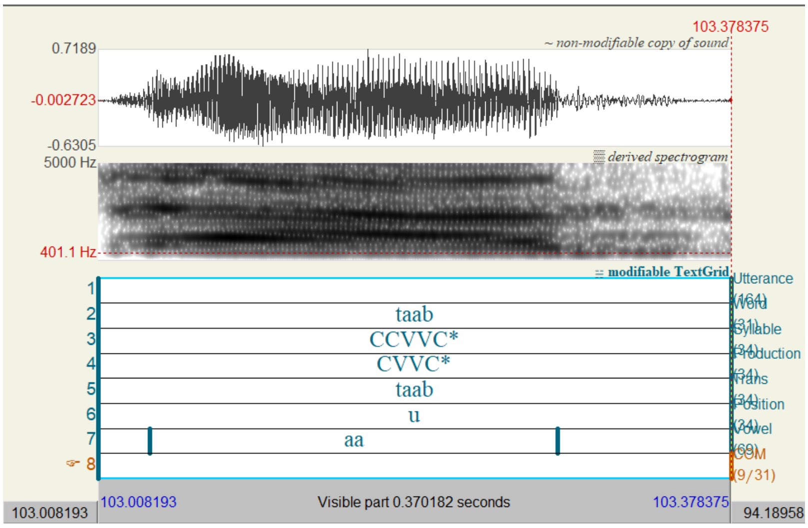
Hussain 1996, Green & Wilson 2006, Payne et al., 2012, Nip & Green 2013). The decrease in durations as children grew older (i.e., an increase in speaking rate) potentially stemmed from motor and neuromuscular maturation, reflecting a universal phonetic pattern (Green & Wilson 2006). This pattern of decreasing durations supports phonological development being characterized by a gradual refinement of motor skills and cognitive control of speech production. A transition was evident around 45–51 months, where durations began to approximate adult productions, which is consistent with studies in English and French-speaking children where the transition was not observed before 2;6 to 4;6 years of age (Smith 1978, Canault et al., 2020). Such findings could contribute to the ongoing debates concerning the age at which children begin to exhibit adult-like patterns, where language-specific factors may contribute as additional features influencing this transition.

Second, the present study showed that syllable and vowel durations in children's speech approximated the adults' values; nevertheless, they did not intersect. The results can be explained by the age groups recruited for this study, ranging from 24–30 to 66–72 months. Previous studies have described that durations continue to decrease until 10–12 years of age before reaching adult-like targets (Kent and Forner 1980, Smith, Sugarman & Long 1983). Thus, children younger than 10 years of age may lack adult-like control and durations (Menyuk 1971, Chermak & Schneiderman 1985). Although older children produce shorter syllables as they approximate adult-like targets, the developmental trajectory showed that the decreasing trend was not linear. The non-linear trajectory supports the view that speech development involves complex interactions between cognitive, linguistic, and motor development (Smith & Thelen 2003). Thus, while there was a noticeable decrease in durations between the youngest and oldest age groups, consecutive age groups may not always show consistent declines (Sharkey & Folkins 1985, von Hofsten 1989, Smith & Zelaznik 2004 and Green & Wilson 2006).

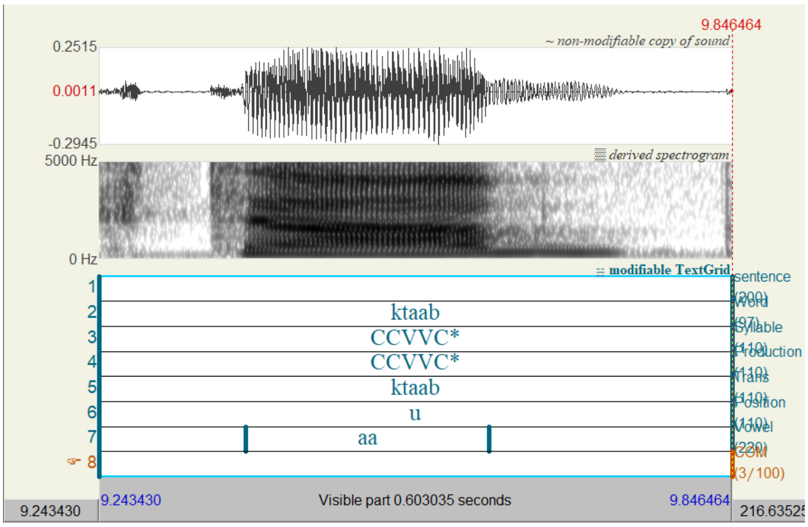
A discrepancy between two sets of findings occurred where the distribution of raw data showed that age group plays a role in speech development (Figure 4.1) while the Bayesian model outputs showed that age group was not consistently a strong predictor for syllable or vowel durations (Table 4.4). This discrepancy suggests that variability in durational patterns cannot be attributed to age alone but is influenced by a range of factors (i.e., motor, linguistic, and cognitive) that interact simultaneously as children develop their speech, reflecting a universal phonetic tendency (Smith & Thelen 2003, Canault et al., 2020). Such results are comparable to Smith, Kenny & Hussain`s (1996) longitudinal study arguing that age was not a determining factor for durational variability in speech as younger participants did not produce syllables with significantly longer durations than older ones. Similarly, Canault et al. (2020) reported that only the youngest and oldest age groups showed significant differences in duration, while other inter-group comparisons did not, highlighting the complexity of durational development.

For task-related effects, the Bayesian model output in RT differed from PT and ST results. While the durations decreased with maturation in ST and PT, syllable and vowel durations increased in RT. The results could be attributed to the task design being mainly tailored for superheavy syllables suggesting that older children produced these syllables more accurately than the younger ones. Superheavy syllables are considered challenging for children due to their (1) structural complexity: containing long vowels or coda clusters requiring the maintenance of high muscular effort in the vocal tract for longer periods (Gay 1978), and (2) they interact with stress assignment and syllable position rules in JA, where these syllables are produced in stressed word final environments (Watson 2002, Al huneety et al., 2023). Thus, with the structural and language-specific properties, superheavy syllables are prone to more phonological processes across the child groups, mainly the younger ones, as evident in the phonological processes analysis (Section 4.5).

Within this context, phonological processes contributing to the durational patterns across the age groups emerged, such as cluster reduction and vowel epenthesis. These processes highlight the strategies children used to manage the complexities of JA syllable structures, which may have also influenced the observed durational patterns. First, cluster reduction was reported in the results across age groups 24–30 to 59–65 months, except for the oldest child group. Figure 5.1: Spectrograms of the production form /taab/ and the target form /ktaab//a shows Speaker M0N (24–30 months) producing /ktaab/ ('book', CCVVC) as /taab/ (CVVC, 370 ms), simplifying the onset cluster. Figure 5.1: Spectrograms of the production form /taab/ and the target form /ktaab//b demonstrates the adult production (Speaker BMG) of the same word (ktaab CCVVC 603 ms), with a difference of 230 ms from the child production. Since superheavy syllables contain more segments than other syllabic structures, deleting a segment reduces the number of articulatory movements required, thus reducing the duration. Cluster reduction has been reported in previous Arabic studies on child production. For example, in Al huneety et al.’s (2023) study on AA speaking children, participants in the 2;1–2;6 years age group tended to reduce clusters. In Kuwaiti Arabic, Ayyad (2011) highlighted that simplification strategies to avoid clusters, such as vowel epenthesis and cluster reduction, were evident in children between 3;10 to 5;2 years of age.



(a) /taab/ Speaker M0N

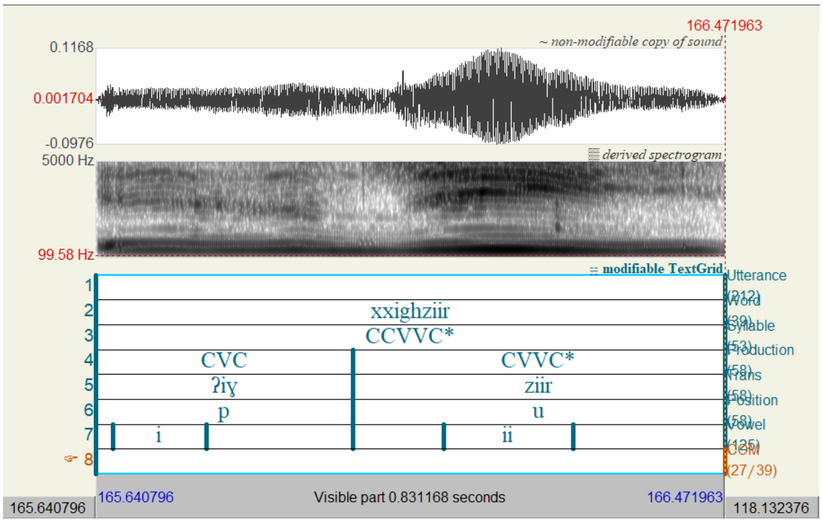


(b) /ktaab/ Speaker BMG

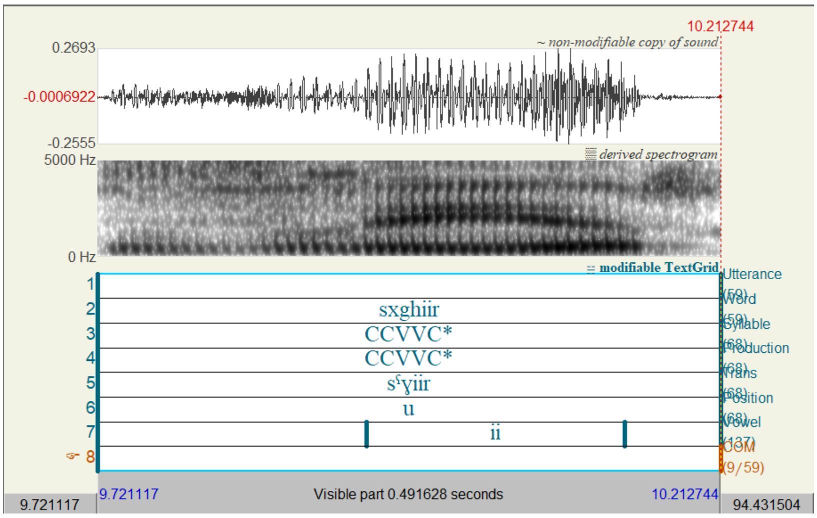
Figure 5.1: Spectrograms of the production form /taab/ and the target form /ktaab/

Second, vowel epenthesis was employed to simplify clusters (Watson 2002, Kiparsky 2003, Watson 2007) and persisted across all the child age groups. The persistence of vowel epenthesis across age illustrates how children navigate challenging syllable forms, reflecting both the developmental trajectory of articulatory control and the influence of language-specific phonological patterns. Figure 5.2/a demonstrates the word /ʔiɣziir/ ('small', CVC.CVVC, Speaker YRO, 24–30 months), while Figure 5.2/b demonstrates the target form /zɣiir/

(CCVVC, Speaker 6ED, adult). Based on JA being a VC dialect (Section 2.1.3), the vowel /i/ was epenthesized before the cluster /i.zɣiir/, where the syllable /ʔiz/ emerged with the first segment of the cluster /z/ appearing in the coda position. The glottal stop was produced since JA does not allow onset-less syllables to surface (Hayes 1995). The second segment of the cluster /ɣ/ became the onset of the second syllable, /ɣiir/. The child production of the word /ʔiɣziir/ (CVC.CVVC) had syllable durations of 334.7 ms and 339.8 ms, respectively. On the other hand, the adult production of /zɣiir/ (CCVVC) had a syllable duration of 491.6 ms, which is significantly longer than the child productions. Recalling the nature of RT, producing epenthetic vowels resulted in the emergence of heavy syllables, which were shorter in duration. Thus, it can be assumed that the number of syllables not superheavy in their structure due to phonological processes, particularly in the younger age groups, contributed to the shorter durations of syllables in this task.



(a) /ʔiɣziir/ Speaker YRO



(b) /zɣiir/ Speaker 6ED

Figure 5.2: Spectrograms of the production form /ʔiɣziir/ and the target form /zɣiir/

5.2.Lexical Stress

The results of lexical stress effect on durations did not fully support the prediction (1.c) that stress would influence syllables resulting in longer durations for stressed vowels/syllables compared to unstressed counterparts. A significant discrepancy was observed between the raw data distribution and the Bayesian probability outputs. First, the raw data distribution supported the prediction that stressed environments are longer than their unstressed counterparts. The difference in mean durations between stressed and unstressed syllables in ST was 72 ms, while the difference in mean vowel duration was approximately 33 ms; 44.5 ms and 31.9 ms for vowels in PT; and 309 ms and 117 ms for vowels in RT, respectively. Notably, a higher difference in the mean durations of stressed and unstressed environments in RT was evident, which may be attributed to the task's nature including superheavy syllables. Since superheavy syllables were longer than other syllable structures, the durational difference between stressed and unstressed targets was more pronounced. These results are in line with de Jong and Zawaydeh's (2002) results of AA, showing that stressed syllables are approximately 20 ms longer than unstressed syllables.

The results showing longer vowel durations in stressed environments align with previous Arabic studies on stress such as Zawaydeh & de Jong (1999), de Jong & Zawaydeh (1999), and Chahal (2003) who emphasized that stress role is evident in vowel durations only rather than syllable durations. Cross-linguistically, de Manrique & Signorini (1983) reported that in Spanish, stress effects are more evident in vowel durations compared to consonant durations. In English, Pollock, Brammer & Hageman (1993) suggested that children as young as 2 to 4 years of age demonstrated the ability to produce longer vowels to mark stressed targets. Such observations support findings from cross-linguistic studies where stressed syllables were reported to be longer than unstressed ones (Fry 1955, Sluijter & van Heuven 1995, Turk & Sawusch 1997, Turk & White 1999, Zec 2007).

Second, the probability distribution results showed that lexical stress is not a strong predictor for syllable or vowel durations across all tasks, which is not consistent with the prediction. In ST (Figure 4.2), RT (Figure 4.14), and PT (Figure 4.24) two patterns were observed in the probability distribution plots: (1) stressed syllables were shorter than unstressed ones, and (2) vowels in stressed syllables were longer than vowels in unstressed ones. The Bayesian model results, with overlapping error bars in the probability plot (Figure 5.3), suggest high variability in speaker-specific patterns which may have contributed to the lack of significant differences between stressed and unstressed syllables. This high variability might have been a result of speaker-specific articulatory skills in producing stressed targets, further complicating the ability to detect consistent durational differences across age groups and tasks.

Stress marking being a weak predictor for syllable and vowel durations could be attributed to two main reasons. Firstly, as suggested in previous studies, Arabic may not strictly adhere to the stress-timed classification or show the same rhythmic regularity as English (Bertinetto 1989, Zawaydeh, Tajima & Kitahara 2002). Although JA is classified as a VC dialect within the stress-timed language category (Miller 1984), the predictability of stress rules in Arabic may result in less distinctive durational manifestations of stress (Watson 2002, Zawaydeh, Tajima & Kitahara 2002, Vogel et al., 2017). Ahn (2000), de Jon & Zawaydeh (2002), and Vogel et al. (2017) highlighted that durations do not suffice as a manifestation of prominence in Arabic as duration could potentially obscure the role of lexical contrasts. This was further supported by the Bayesian model outputs for the two-way interaction between age group and lexical stress, wherein although the difference between stressed and unstressed syllables increased with maturation (Figure 5.3), none of the models demonstrated that this interaction was a strong predictor for syllable or vowel durations across the tasks (Section 4.1, 4.2, 4.3). Thus, the reliability of stress marking as a predictor for durations is challenged (Roach 1982, Heliel 1982).

Focusing on Arabic-specific properties, the position of the stressed syllable within a word may have been a more decisive factor affecting syllabic productions. Despite lexical stress not being a strong predictor for syllable or vowel durations, the current results showed that the two-way interaction between lexical stress and syllable position in ST, and the three-way interaction in ST and PT were strong predictors for durations. Such results align with Allen & Hawkins's (1980) study specifying that increased durations of marking stress are only employed if stress occurs in the word-final position. This interweaves with the examination of English stress by Kehoe, Stoel-Gammon & Buder (1995) who stated that the magnitude of the difference between stressed and unstressed syllables should not be in intra-word observations as this comparison does not serve as an appropriate measure. In intra-word comparisons, the inherent ineliminable variables, such as word-final lengthening, contribute to durations. In inter-word comparisons, the magnitude of phonological comparisons with eliminated uncontrolled variables, such as position, allows for a better understanding of any developmental pattern emergence. Since the current study did not separate stressed and unstressed targets according to inter-and-intra word measures, it is assumed that unstressed final syllables may have contributed to the increased durations of unstressed targets.

Secondly, stress assignment being not a strong predictor for syllable or vowel durations could be attributed to children's motor and neuromuscular developing abilities not exhibiting sufficient durational control to differentiate stressed and unstressed syllables (Hawkins 1979, Allen and Hawkins 1980, Nip & Green 2013, Canault et al., 2020). The posterior predictive plot (Figure 5.3) suggests a refinement of temporal control of stress targets over time. The youngest two age groups, 24–30 and 31–37 months, did not exhibit adequate measures of temporal control to create a contrast between stress and unstressed syllables. Nevertheless, by 38–44 months, unstressed syllabic durations decrease to become shorter than stressed ones (Allen and Hawkins 1980, Pollock, Brammer & Hageman 1993, Zec 2007). Children in the early stages of development may not efficiently reduce unstressed syllable durations compared to stressed syllables as much as adult speakers do (Allen & Hawkins 1980, Gerken, Landau, Remez 1990, Gerken 1991, Pollock, Brammer & Hageman 1993, Ballard et al., 2012). The child's inability to deconstruct the syllabic information of unstressed syllables (i.e., identify and isolate the syllables) is evident in children not efficiently reducing the duration of these syllables (Allen & Hawkins 1980, Ballard et al., 2012). Ballard et al. (2012) indicated that English-speaking children continue to refine the durations of unstressed targets at least until seven years of age with the most noticeable difference occurring by 2–3 years of age. Arguably, Pollock, Brammer & Hageman (1993) reported that stressed syllable durations remain consistent over time but unstressed syllable durations continue to decrease with maturation. Similarly, Allen & Hawkins (1980) claimed that the reduction of weak unstressed syllables is mastered at later stages of development, and only by 4–5 years of age does the rhythm become more adult-like. In AA, Al huneety et al. (2023) suggested that children up to 2 years of age do not fully acquire stress rules, but by three years, children produce stress in an adult-like manner. Moreover, some studies argued that other acoustic parameters may be more prominent for stress marking in child productions, such as intensity and pitch (Pollock, Brammer & Hageman 1993). Kehoe, Stoel-Gammon & Buder (1995) suggested that while English-speaking children as young as 18–30 months use similar acoustic cues as adults, such as pitch and intensity, their ability to consistently apply duration as a stress marker shows more developmental patterns with maturation.

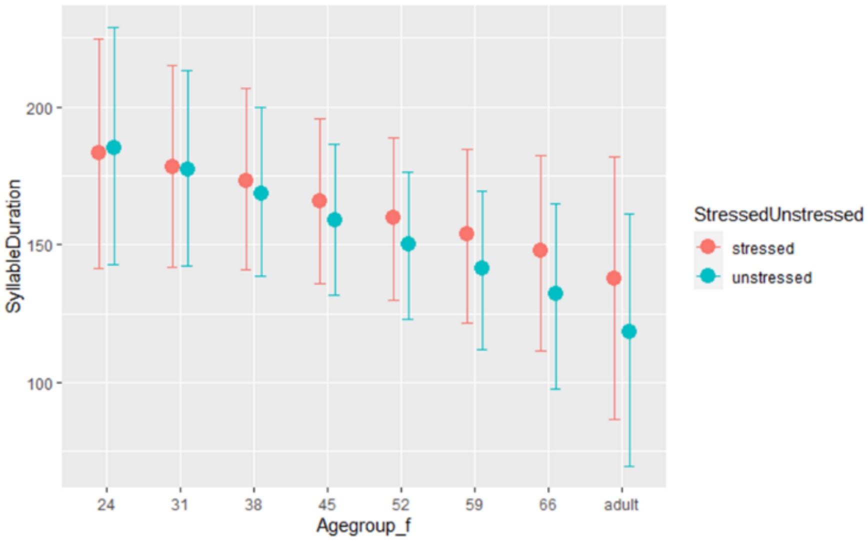


Figure 5.3: Posterior predictive plot for syllable duration of the interaction between stress and age group in ST

5.3.Syllable Position

The current results supported prediction (1.d) that word final lengthening would be evident in longer durations for word final syllables and vowels compared to non-final ones. The results demonstrated that syllable position within a word was a strong predictor for syllable and vowel durations across all tasks, except for vowel duration in PT (Table 4.18). Word final syllables were 1.6 times longer than word non-final syllables in ST, 2.7 times longer than their counterparts in RT, and 1.8 times longer in PT. Vowels in word final syllables were 1.5 times longer than vowels in word non-final syllables in ST, 2.4 times longer than their counterparts in RT, and 1.6 times longer in PT.

These findings are consistent with the well-documented tendency of word final lengthening across languages. Scholars such as Delattre (1966), Smith (1978), Robb & Saxman (1990), Beckman, Edwards & Fletcher (1991), and Halle, Boysson-Bardies & Vihman (1991) suggested that word final lengthening emerges in children's vocalizations. Robb & Saxman (1990) specified that word final lengthening in English is evident in children as early as 8–14 months and during their pre-word vocalization stage. Similar patterns were documented in French-speaking children by Konopczynski (1986), Fletcher (1991), and Allen (1983), postulating that two-year-old French children produced final lengthening in a similar pattern found in adult productions. Final lengthening affecting vowels in JA child speech also supports previous reports on English by Zec (2007) and Gordon (2007), de Jong & Zawaydeh (1999), and Yeou (2005).

The manifestations of syllable final lengthening are proposed to vary with age, either becoming more pronounced in some languages or repressed in others (Halle, Boysson-Bardies, and Vihman 1991). Figure 5.4 below exhibits a JA-specific pattern, where the manifestation of word final lengthening becomes repressed with age (i.e., younger children exhibit more word final lengthening compared to older children). The two-way interaction between age group and syllable position was a strong predictor for syllable and vowel durations across all tasks (except for vowel durations in RT). The probability plot (Figure 5.4) presents the developmental pattern of syllable final lengthening in JA speech. The durations of word final syllables decreased with maturation whereas the non-final syllable durations did not exhibit much durational difference. Younger children, aged 24–30 to 38–44 months, seemed to show greater word final lengthening than older age groups; however, the age group 45–51 months and older demonstrated more overlap between the durations of word final and non-final syllables. The narrower durational difference between the two environments coincides with the proposition that the increased motor learning process results in more control of word final and non-final durations (Oller 1973, Klatt 1975, Smith 1978, Cooper & Paccia-Cooper 1980, Nip & Green 2003). The question remains whether syllable final lengthening results from the organizational constraints imposed upon the articulators to produce meaningful units by phonological learning or intrinsic biological aspects of speech production (Smith, Kenny & Hussain 1996, Canault et al., 2020). The findings of this study do not allow for the formation of a definitive conclusion but they indicate that lengthening effects decrease with maturation.

Another possible explanation for this durational pattern could be attributed to younger children producing more monosyllabic words than disyllabic and multisyllabic words. The utterance boundaries coincided with word boundaries and monosyllabic word durations, coded as word final syllables, may have skewed durational distributions. Nonetheless, older children produced more disyllabic and multisyllabic words with word position contrasts allowing for observing final lengthening (Table 4.23, Table 4.33). Then, the durational patterning of syllables according to syllable position was not as skewed as the patterning evident in the younger groups.

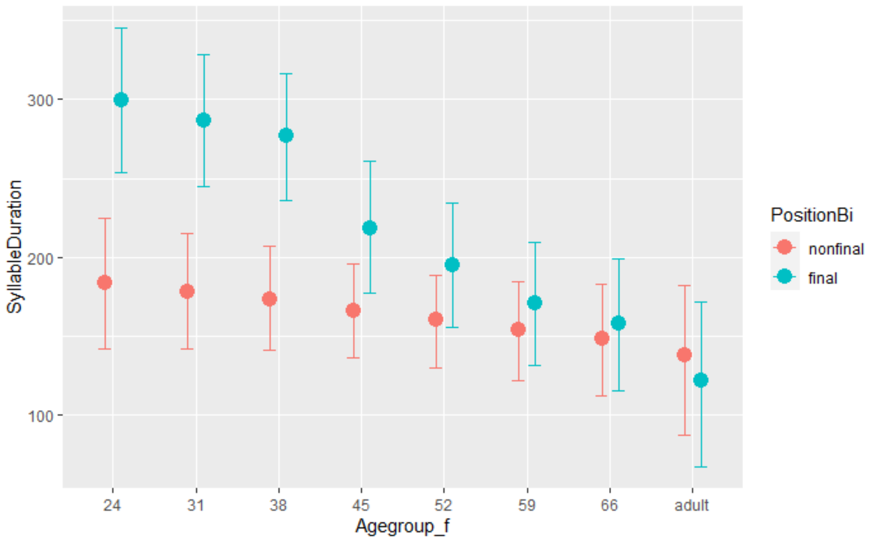


Figure 5.4: Posterior predictive plot for syllable duration for syllable position across the age groups in ST

Final syllable lengthening serves as a key indicator of normal phonological development in JA child speech (Smith 1978, Kubaska & Keating 1981, Allen & Hawkins 1980, Crystal 1986). A notable finding was the alignment of utterance boundaries with word boundaries (i.e., younger children produced more monosyllabic words compared to older children). This suggests that children begin developing an awareness of word boundaries as a prosodic feature by marking the durations of their constituents from an early age. The consistent occurrence of syllable final lengthening across tasks emphasizes its importance as a robust, language-specific phonological feature in early prosodic development. Deficits in word final lengthening may be associated with developmental disorders as it plays a significant role in the child's ability to segment and store target rhythms (Snow 1998b, Yeou 2005). The absence or misapplication of this feature could obscure/delay the child's capacity to perceive and produce correct rhythmic structures, leading to further communicative challenges (Lehiste 1977, Echols and Newport 1992). This notion is supported by studies suggesting that word-final lengthening and prosodic marking reflect the interaction between language-specific experience and age-related motor skills in typically developing children (Delattre 1966, Oller & Smith 1977, Nelson et al., 1989, Robb &

Saxman 1990, Morgan 1996, Morgan & Demuth 1996). Al Huneety et al. (2019) further supported this view in their study of AA-speaking children, mentioning that perceptual saliency of word-final syllables serves as an entry point for identifying words in perception and production (Echols & Newport 1992).

The results showed a significant three-way interaction between age group, stress, and syllable position but distinct patterns were reported for syllable and vowel durations. In children aged 24–44 months (Figure 5.5), word final stressed syllables were longer than final unstressed syllables. However, by 45–51 months, this trend shifted with unstressed final syllables becoming longer than stressed final ones. For non-final stressed syllables, stressed syllables were consistently longer than their unstressed counterparts with the durational difference increasing with maturation. As for vowel durations, vowels in stressed final syllables were longer than those in unstressed final syllables across age groups. Conversely, vowels in stressed non-final syllables decreased in duration over time, reaching values comparable to vowels in unstressed non-final syllables by 52–58 months. Notably, the shift in durational patterns for syllables occurred earlier than for vowels, with significant changes observed by 45–51 months for syllables and 52–58 months for vowels. Younger children produced shorter stressed non-final syllables but by 59–65 months, these syllables became increasingly longer than unstressed non-final syllables displaying a clear developmental trajectory toward adult-like patterns. These findings support Snow's (1994) work which observed a similar interaction in English-speaking children aged 2 to 4 years. Snow's results demonstrated that durations of non-final stressed syllables decreased with age and by 4 years children's durations (Mean = 188ms) closely resembled adult-like productions (Mean = 178ms). Correspondingly, Smith's (1978) analysis of the temporal aspect of English productions showed that the performance of child productions in the 3;0–4;6 years age group was similar to that of adults in the effects of stress and syllable position on vowel durations. Further supporting this, Schwartz et al. (1996)

demonstrated that child vowel durations in stressed non-final syllables were longer than their counterparts, and vowels in stressed final syllables were longer than vowels in unstressed final ones. As for syllables, Schwartz and colleagues proposed that children produced longer stressed non-final syllables than their counterparts, and final stressed syllables were longer than final unstressed syllables. Although the durational differences between syllable position and stress combinations were similar to the current study’s results, their analysis revealed that this interaction was not significant for the child group, with an age range of 22–26 months, which contrasted with the current model outputs.

Durational patterns of older children not aligning with those of younger children support Snow's (1992) argument regarding continuity and discontinuity in children's use of final syllable timing and stress assignment patterns. Snow highlighted that fluctuations in utilizing previously learned rhythmic patterns, coupled with the emergence of new forms, reflect development in the prosodic control and durational contrasts. Similarly, Lahey (1974) noted due to the acquisition of syntax, the prosodic skills evident at 22 months were not merely extensions of the surface forms observed at 16 months but involved distinct mechanisms and constraints that signify a complex pattern of linguistic development.

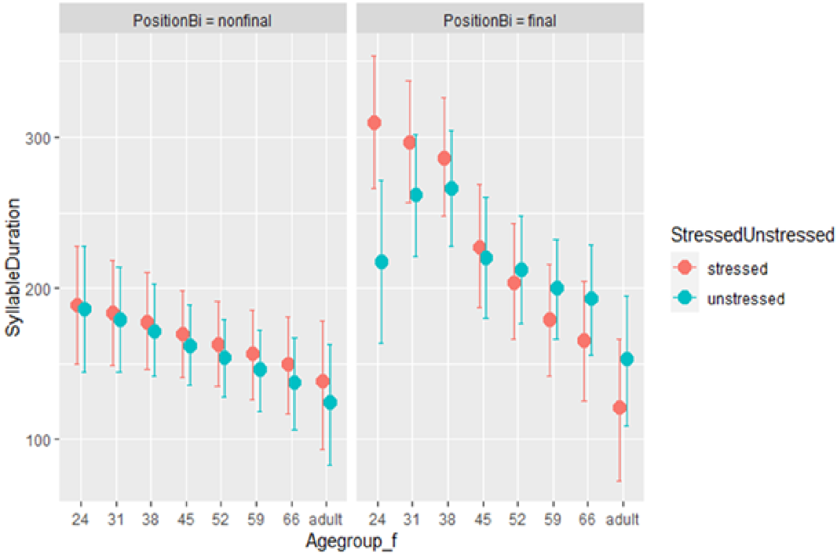


Figure 5.5: Posterior predictive plot for syllable duration for the three-way interaction

5.4.Syllable Structure and Moraicity

The current results support the general prediction that core syllables, such as CVC, CVV, CV, and CVVC, would be produced more frequently than marginal syllables, such as CCVVC and CCVC. In ST, the frequency distribution showed that core syllables were more frequent than marginal ones. The core syllables accounted for 89.6 % of the syllables, while the marginal ones accounted for 10.4% only. Similarly in PT, core syllables (93.4%) were more frequently produced than marginal ones (6.6%). As for RT, since the task's nature is tailored for superheavy syllables, no assessment of syllabic frequency could be performed. Second, results support the general prediction that younger children would exhibit less variability in their syllabic structures but gradually exhibit adult-like patterns with maturation. Table 4.1 shows that younger children predominantly use simpler syllable structures with limited variability. With maturation, the frequency of more complex structures, such as CCVVC and CVVC, increased, reflecting a gradual shift toward adult-like syllabic patterns. This progression in syllable complexity aligns with Jakobson's (1968) universal order of syllable acquisition, Fikkert's (1994) parametric theory of syllable structure, and Demuth & Fee's (1995) theory of children's early word shapes. The results further validate Levelt, Schiller & Levelt's (2000) work emphasizing that a greater variety of syllabic structures is exhibited with maturation, reflecting a shift toward adult-like patterns.

These findings are consistent with studies on Arabic child speech such as Alqattan (2015) who suggested that Kuwaiti children produced increasingly complex syllable shapes in the older age groups compared to younger ones. Mashaqba et al. (2019) reported that JA children exhibited a universal developmental pattern with older children producing a greater variety of syllabic structures. They also highlighted that CVC, CV, CVVC, and CVV were the most frequently occurring syllables in JA child speech, with more complex forms, such as CCCVV and CVVCC, being less frequent.

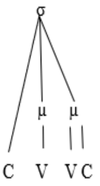
Results support prediction (1.e) that syllable structure complexity would influence durations, with syllables containing more constituents displaying longer duration. The Bayesian model outputs demonstrated that syllable structure was a robust predictor for syllable and vowel durations (Table 4.4, Table 4.11, Table 4.16). Across all tasks, light syllables were the shortest, followed by heavy syllables, while superheavy syllables were the longest. Similarly, vowels in light syllables were the shortest, followed by vowels in heavy syllables, while vowels in superheavy syllables were the longest. The robustness of syllable structure across languages supports the proposition that the number of constituents and their complexity determine syllable and vowel durations in child productions (Newman 1972, Sen 2012, Khattab & Altamimi 2013). The composition of syllable structure provides crucial insights into the durational patterns of JA productions. Generally, more complex structures were expected to be longer than their counterparts as it was argued that the assigned moras contributed to the syllabic and vocalic durations, creating distinctions between syllable structure types (Khattab and Al-Tamimi 2013). The syllable structure constituents set the number of moras assigned based on the moraic constraints such as WPR, extrametricality, and extrasyllabicity (McCarthy and Prince 1986, Watson 2000, Abdoh 2011, Huneety 2015).

The issue of trimoraic superheavy syllables does not arise they appear in the word final position as final codas do not receive a mora (Hayes 1995, Watson 2002). However, these syllables become problematic to the theoretical account as they appear in word non-final positions, as they are assigned a mora contributing to the weight of the syllable by WPR (Hayes 1995). Durations of word final and non-final superheavy syllables provide insights into their behaviour as either bimoraic or trimoraic. The hypothesis is if superheavy and heavy syllables are bimoraic, then their durations should be comparable. In heavy CVV and superheavy CVVC syllables, the moras are assigned to the long vowels, and thus, comparing both structures' vowel durations in word non-final positions is necessary.

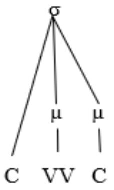
The question arises whether the phonological context of these syllables leads to differences in duration, and if so, how to determine if it is due to syllable position or structure. The analysis in Section 4.1 was carried out and included ST only as it had more data points compared to PT and RT. Results showed that the interaction between syllable position and syllable structure was not a strong predictor for syllable duration. Nonetheless, this interaction was a strong predictor for vowel durations. The model output supported prediction (2.d) that superheavy syllables would be bimoraic and not trimoraic as vowel shortening was anticipated to occur in non-final superheavy syllables. Vowels in non-final superheavy syllables were shorter than vowels in final superheavy syllables, contrary to vowels in heavy or light syllables that did not show notable differences. This observation suggests a JA-specific phonetic property, where there is an adherence to restricting non-final superheavy syllables to being bimoraic instead of trimoraic. This is achieved by producing shorter vowels in non-final positions to avoid contributing to the weight of the syllable. Such results are consistent with the analyses of Newman (1972), Clements and Keyser (1983), Watkins (2001), and Zec (2007). These scholars argued that vowel shortening occurs in languages prohibiting CVVC syllables from being trimoraic, where long vowels are blocked from closed syllables to not disturb the bimoraicity condition. The current study's results are consistent with the durational pattern observed in Broselow et al.'s (1995) analysis of Syrian and Lebanese speakers producing CVVC syllables. They proposed that vowels in a CVVC syllable are shorter than a long vowel in an open syllable and longer than a short vowel in an open syllable. The current observations are also consistent with Bamakhramah's (2010) OT analysis of Meccan Arabic, where changing the templatic form of non-final CVVC syllables (i.e., incorporating the final consonant into the onset of the following syllable) was evident. However, the mechanism of vowel epenthesis that Meccan Arabic employs to avoid trimoraicity is inconsistent with the current JA productions that employ vowel shortening instead. Vowel shortening is consistent with another OT analysis by Gordon (2007) who suggested that truncation of overlong syllables occurs by vowel shortening and preserving the coda consonant.

Nonetheless, such results are inconsistent with Crossley’s (2023) analysis. He indicated that while there is evidence of coda shortening in JA superheavy syllable productions, the lack of vowel shortening suggests that JA does not exhibit evidence for the phonetic effects of a trimoraic ban. Crossley's argument surpassed the syllabic level analysis as superheavy syllables in the CVVG configuration were observed to occur due to affixation. He suggested that if shortening is attributable to affixation alone, then shortening is expected in all codas before consonants. The discrepancy in findings might be attributed to focusing on syllables containing geminates, which were not attested in the current data, which may have influenced pre-and-post vowel durations (Khattab & Al-Tamimi 2014, Al-Deaibes 2016).

The bimoraic and trimoraic discrepancy in the literature does not warrant definite conclusions to the attested current results. Figure 5.6 could serve as a possible moraic representation of superheavy syllables in JA. Vowel shortening could be a manifestation of mora sharing with the coda consonant in superheavy non-final syllables. Nonetheless, since the current study did not examine segmental durations of codas, the answer remains unclear. The Bayesian model output for syllable duration shows a possibility that shortening effects do not influence consonants in superheavy syllables in word non-final positions. Thus, a coda in a non-final superheavy syllable may still be assigned a mora, contributing to the weight of the syllable by the WPR. However, the vocalic model output showed that remarkably shorter vowels are produced in superheavy non-final syllables compared to final ones. Since a long vowel in a non-final CVVC syllable had a comparable vowel duration to a short vowel in a non-final CVC, it is hypothesized that the latter does not contribute to syllable weight by two moras. Instead, the vocalic constituent has durational manifestations contributing to a single mora, as expected in a CVC syllable. Nonetheless, only traces of vowel shortening can be suggested, but without analysing segmental durations, no definitive conclusions of mora sharing can be offered.



(a) Mora sharing



(b) Vowel shortening

Figure 5.6: Metrical trees for superheavy syllable representation in word non-final positions

5.5.Superheavy Syllables

The development of superheavy syllables in JA child speech is exemplified by children demonstrating an increased ability to produce these syllables in a variety of sub-structures, word lengths, and stress patterns. This ability is influenced by aspects of articulatory control and phonological development (Smith 1978, Green & Wilson 2006, Nip & Green 2013). Enhanced articulatory control coupled with increased language experience allows for the production of more complex structures (e.g., long vowels and clusters) and longer word lengths (e.g., expanding from monosyllabic to multisyllabic words) (Tingley & Allen 1975, Robb & Saxman 1990). As children's phonological awareness and understanding of phonological rules advance, they produce more complex superheavy productions (Ferguson & Farewell 1975, Fikkert 1994). The discussion below concerns the development of superheavy syllables in the domains of sub-structure frequency, frequency across word lengths, and prosodic patterns in addition to durational development.

The development of superheavy syllables and frequency distribution

First, the analysis revealed that superheavy syllable sub-structures emerge and expand with maturation. The results supported the general prediction that despite their complexity, superheavy syllables were anticipated to emerge as early as two years of age. In ST, the production of superheavy syllables in CVVC and CVCC sub-structures emerged at 24–30 months, restricted to monosyllabic words. By 31–37 months, the production of CVCC syllables expanded to disyllabic words. By 38–44 months, a third sub-structure, CCVVC, appeared in monosyllabic words. In PT, although all sub-structures were observed even in the youngest age group, a clear developmental trend showed that older children produced a higher number of superheavy syllables. In RT, in the age group 24–30 months, superheavy syllables were produced in CVCC, CVVC, and CCVVC sub-structures in monosyllabic words. With maturation, around age group 38–44 months, superheavy syllables in disyllabic and multisyllabic words became more frequent. For ST, coda clusters (i.e., CVCC) appeared earlier than onset clusters (i.e., CCVVC) in JA child speech. This observation is consistent with the results of Mashaqba et al.'s (2019) study of AA, where coda clusters appeared as early as 2;0–2;6 years of age in monosyllabic words, while onset clusters emerged by 2;7–3;0 years of age. Studies on Kuwaiti Arabic child speech reported this observation as Ayyad (2011) and Alqattan (2015) suggested that onset clusters are less frequent compared to coda clusters, and they do not appear in the younger age groups due to their complexity. The acquisition and development of coda clusters earlier than onset clusters may be attributed to the influence of a language-specific property, wherein Arabic has a higher frequency of superheavy syllables with coda clusters than onset clusters (Broselow et al., 1995, Watson 2002). Additionally, the role of coda clusters differs from onset clusters as codas contribute to the moraic weight of the syllable by WPR (Section 2.1.3.2.2). Thus, it could be argued that more superheavy syllables with coda clusters appeared in the child productions due to speakers maintaining the bimoraicity condition (Hayes 1995, Watson 2002, Davis & Ragheb 2014).

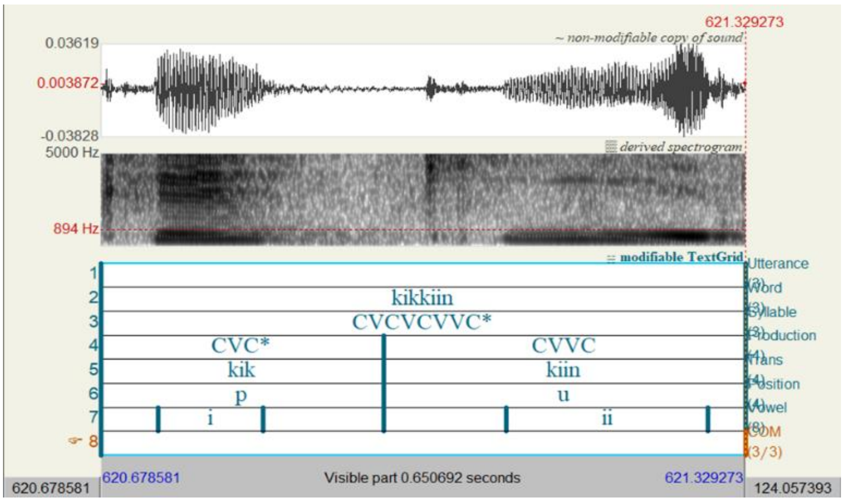
Second, word length played a significant role in superheavy syllable development. Children exhibited an overall preference for producing superheavy syllables in monosyllabic words (CVVC, CVCC, CCVVC). As more sub-structural varieties emerged, JA children developed their phonological skills becoming less reliant on producing superheavy syllables exclusively in monosyllabic words. Instead, they started incorporating them into disyllabic and multisyllabic words, supporting the universal syllable complexity and developmental trend (Jakobson 1968, Chomsky & Halle 1968, Fikkert 1994). This observation has been attested in Alqattan's (2015) analysis of Kuwaiti children exhibiting a preference for simpler monosyllabic word structures. In PT, superheavy syllables in disyllabic and multisyllabic words appeared as early as 24–30 months of age. However, superheavy syllables in multisyllabic words were observed at a later stage, by 38–44 months for ST and RT. This discrepancy could be attributed to the nature of PT, where visual stimulations of objects aided in the production of these syllables. The current findings align with Alqattan (2015), Mashaqba et al. (2019), Al Huneety et al. (2023), and Ammar (2002) who observed that older children could produce a greater variety of syllables across word lengths. Mashaqba et al. (2019) identified superheavy syllables in disyllabic words (CVC/CVVC) as early as 1;0–1;6 years and in multisyllabic words by 1;7–2;0 years. Al Huneety et al. (2023) found that superheavy syllables appeared in monosyllabic words by 1;0–1;6 years and expanded to more complex structures by 2;1–2;6 years, aligning with the current study. However, their study did not report the emergence of CCVVC or CVCC syllables in disyllabic or multisyllabic words, which the current study observed at 31–37 and 38–44 months, respectively. Ammar (2002) found that Egyptian children aged 2–3 years produced superheavy syllables (CVVC, CVCC), but not CCVVC, in monosyllabic words, while the current study observed all three forms. Differences in sub-structure emergence may reflect dialectal differences between JA and Egyptian Arabic.

Moreover, the current data demonstrated that two superheavy sub-structures were observed in disyllabic and multisyllabic words, including CVVC and CVCC, with CVVCs being more frequent and disyllabic words being more common than multisyllabic words. Such an observation was reported in Kuwaiti child speech, where Alqattan (2015) suggested that CVVC syllables are the most commonly produced superheavy syllables appearing across all word lengths, while CVCC, CCVCC, and CVVCC were less common. Al huneety et al. (2023) and Ammar (2002) argued that CVVC syllables being more frequent is attributed to children's finding clusters in CVCC or CCVVC challenging. Al huneety et al. (2023) added that compared to CVVC syllables, CVCC syllables are less evident due to vowel epenthesis from a Jordanian-dialect-specific perspective. Such results are similar to the current findings, where vowel epenthesis (2.1.3.2.4) was evident in CVCC productions (e.g., /kalb/-/ka.lib/ 'dog' and /ʔird/-/ʔi.rid/ 'monkey').

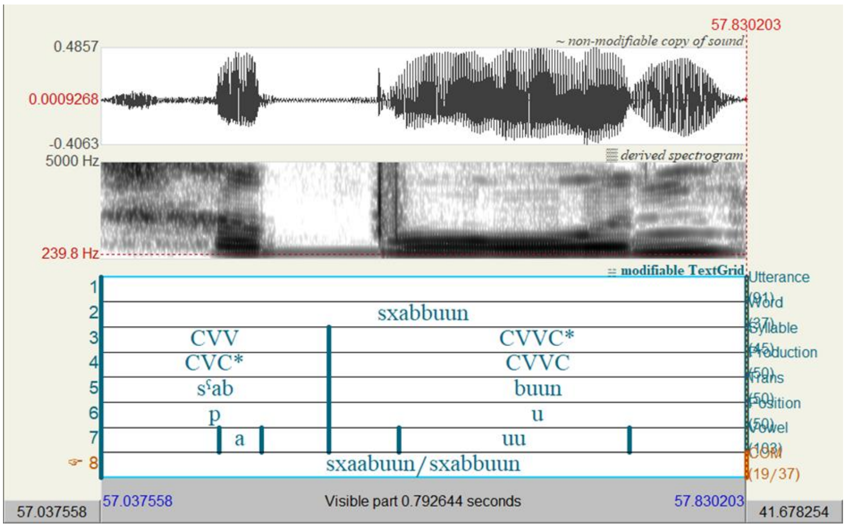
Third, there was an increase in variability and complexity of stress patterns and syllable position in superheavy syllable JA productions. Older child groups demonstrated more variability in their prosodic templates where superheavy syllables were produced in different word positions and stress assignment patterns. This development reflects a universal prosodic development, where children exhibit an initial preference for simpler prosodic structure, but with maturation, more complex patterns appear (Ferguson 1977, Vihman 1993, Abdoh 2010, Vihman & Croft 2007). Such an observation was evident in Alqattan's (2015) and al Huneety et al.'s (2023) studies, suggesting that with maturation, the prosodic templates of Ammani and Kuwaiti children expand. They attributed this to vocabulary expansion that occurs with maturation allowing for the gradual development of phonological and prosodic knowledge (Tobin 1997, Hua & Dodd, 2000).

The current data revealed that children's acquisition of superheavy syllables seems to be intertwined with stress placement and syllable position affecting their production. Unstressed final syllables appeared earlier than stressed non-final syllables. In ST, unstressed final syllables appeared in the age group 31–37 months, while stressed non-final syllables emerged by 59–65 months for disyllabic words and by 38–44 months for multisyllabic words. In PT, unstressed final syllables appeared as early as 24–30 months, while stressed non-final syllables appeared in the 31–37 months group. This development might be influenced by language-specific factors, such as the stress patterns and phonotactic constraints of Arabic, which may have determined the order in which these features emerged. The earlier appearance of unstressed final syllables over stressed non-final syllables could be attributed to their frequency and saliency in the input language. Similar patterns have been documented in English (Kehoe

2002) and Spanish (Demuth 2001). Another confounding factor for this pattern is the occurrence of the phonological process, assimilation and gemination, which persisted across age groups (Table 4.37). To demonstrate, Figure 5.7/a demonstrates Speaker 6VI from the age group 38–44 months producing the target form /sa.ka.kiin/ 'knives' CV.CV.CVVC as /kik.kiin/ CVC.CVVC. Similarly, Figure 5.7/b demonstrates Speaker G9I in the age group 45–51 producing the target form of /s^{ʕ}aa.buun/ 'soap' CVV.CVVC as /s^{ʕ}ub.buun/ CVC.CVVC. In the target form, the superheavy syllable in the final position is expected to attract and receive the main stress. However, after applying the phonological process in the production form, the stress is assigned to a non-final syllable instead, resulting in the superheavy unstressed final syllable. This process has been documented in AA children by Mashaqba et al. (2021) where this process was depicted as a lengthening strategy to preserve the underlying moraic weight of the target syllable (Davis and Ragheb 2014, Mashaqba et al., 2021).



(a) /kik.kiin/ Speaker 6VI



(b) /sˤub.buun/ Speaker G9I

Figure 5.7: Spectrograms of the words /kik.kiin/ and /sˤub.buun/

Durational patterns

Word length was incorporated as a variable due to differences observed in the three-way interaction results in syllable and vowel durations across the tasks (4.1, 4.2, 4.3). The descriptive statistics supported prediction (2.a) that younger children would be expected to produce longer superheavy syllables compared to older age groups. This prediction aligns with theoretical accounts of early phonological development, suggesting that younger children exhibit less efficient control, resulting in longer syllables (Smith, Sugarman & Long 1983; Green & Wilson 2006; Nip & Green 2013). Nonetheless, this prediction was not supported by the Bayesian model output, where age group was not consistently a strong predictor for superheavy syllable durations. Only ST showed evidence that age group affects syllable duration, where younger children produce significantly longer superheavy syllables than older age groups. Then, durational differences might be task-specific, where task complexity or familiarity, may play a significant role in shaping durational patterns. Given that age groups were divided into six-month intervals, the age gap between the groups may not have been substantial enough to detect significant durational variations in superheavy syllables. Therefore, identifying any potential statistical distinctions based on age might have been difficult, making it challenging to draw conclusions on the durational differences. Future research could consider having a larger age interval to account for superheavy syllable developmental patterns.

The descriptive statistics showed that in monosyllabic words, younger age groups exhibited longer syllable and vowel durations than older age groups. Such results are consistent with the findings in Section 5.1, suggesting that with maturation, syllables and vowels become shorter due to developed articulatory coordination, motor skills, and language experience (Green & Wilson 2006, Nip & Green 2013, Canault et al., 2020). However, the oldest child group, aged 66–72 months, did not exhibit adult-like productions. This suggests that JA superheavy syllable development continues to influence durations after 72 months, supporting that some aspects of phonological development may extend into later childhood (Kehoe 1997, Fikkert 1994). Between age groups 38–44 and 45–51 months, durations of superheavy syllables in monosyllabic words decreased notably and started trending towards the adult values in ST and PT. Nonetheless, RT did not follow this trend, where the durations decreased by 31–37 months, which is earlier than the other tasks. This may be attributed to the nature of the task, where children might have been influenced by the experimenters' durational patterns reflecting the role of imitation in early speech development (Vihman 1996).

Additionally, word length may be a factor in assessing durational trends. As the number of syllables in a target word increases, children demonstrate delayed development in reducing durations (Son & Santen 1997, Chu & Feng 2001, Hajek & Stevens 2008). This was evident in disyllabic words whereby age 45–51 months both syllable and vowel durations demonstrated a sharp decrease, which is at a later stage compared to monosyllabic words. This highlights that the acquisition of efficient articulatory patterns is gradual and is influenced by word complexity (Levelt 1994). For multisyllabic words, the durations created a U-shaped pattern, highlighting the non-linear nature of language development (Kuczaj 1977). In ST and PT, durations were longer in younger age groups, then durations decreased by 52–58 months but increased again by 59–65 months. The U-shaped pattern is a universal phenomenon in child language acquisition linked to cognitive development as children move from simple imitation to complex rule-based understanding, temporarily disrupting production accuracy (Ambridge & Lieven 2011, Marcus et al., 1992). Early mastery seems to be affected by a temporary regression before achieving stable and adult-like patterns (Karmiloff-Smith 1992). As for RT, the syllabic and vocalic durations decrease by 59–65 months of age. Recalling that in monosyllabic words, the age group for the decrease was around 38–44 months and 45–51 for disyllabic words; the current finding further supports the proposition that as the number of syllables in the target words increases, children learn to control their superheavy durations at a later stage of development (Nip & Green 2003).

The results did not support prediction (2.b) that lexical stress would affect superheavy syllables, with stressed superheavy syllables exhibiting longer durations than unstressed ones. Stress effects were not consistently evident in superheavy syllable productions in disyllabic words across the tasks. Confirming with the previous analysis of lexical stress assignment discussed in Section 5.2, duration did not play a decisive role in stress marking in superheavy syllables. Such results further support the proposition that in Arabic, stress assignment may primarily be predetermined by syllable structure and its position within a word rather than stress rules (Hayes 1995, Watson 2007). In stress-timed languages, it has been suggested that syllable structure tends to reinforce the distinction between stressed and unstressed syllables (Firth 1948, Dauer 1983, Fear, Cutler, & Butterfield 1995). Thus, in languages with a fixed stress system such as Arabic, the availability of explicit acoustic correlates of stress (e.g., duration) may be limited. Cutler, Dahan, and Van Donselaar (1997) argued that since stress is fully predictable in such languages, there may be less need for its explicit realization through the prolonged duration of the stressed syllable.

As for syllable position, the results supported prediction (2.c) that word-final lengthening would be observed in superheavy syllable productions, with final syllables being longer than non-final ones. The descriptive statistics showed that the mean durational differences according to syllable position were more pronounced in PT compared to ST. This could be attributed to the nature of PT, as visually stimulating pictures may have resulted in excited productions, contributing to the longer durations (Fletcher 1991). Word final lengthening effects were not discussed for RT as the task design only contained word final superheavy syllables. In agreement with the discussion in Section 5.3, the Bayesian model outputs demonstrated that syllable position was a strong predictor for syllable and vowel durations in superheavy syllables in disyllabic and multisyllabic across the three tasks. In the current study, syllable position has been the only consistently strong predictor for syllabic productions, denoting prosodic boundary marking (i.e., word final lengthening) as more salient than stress marking.

Regarding the influence of sub-structure on durations, the descriptive statistics revealed that superheavy syllables exhibited minimal variations in syllable durations. The Bayesian model outputs indicated that the sub-structure was not a strong predictor for superheavy syllable durations. However, the differences were more pronounced in the mean vocalic durations across various sub-structures. Vowels in CCVVC syllables consistently demonstrated the longest durations across the tasks, followed by vowels in CVVC syllables, while vowels in CVCC syllables were the shortest. This observation aligns with scholars such as Eilers et al. (1984), Campbell (1992), and Greenberg et al. (2003) who emphasized the importance of the syllabic peak's nature in distinguishing syllables within a child's linguistic repertoire. The findings imply that the temporal characteristics of vowels, influenced by the syllable's sub-structure, are a key factor in the developmental trajectory of child speech (McCarthy 1981, 1982, Hayes 1995, Hubbard 1994). Moreover, the longer durations of vowels in more complex structures such as CCVVC may suggest that children require additional articulatory effort and time to produce these syllables, reflecting the developing phonological skills (Sen 2012, Khattab & Altamimi 2013). This further emphasizes the importance of syllable structure in the acquisition and development of child speech.

5.6. Summary and Theoretical Implications

The present findings extend the theoretical accounts indicating the universal non-linearity and variability of linguistic development (Sharkey & Folkins 1985, Smith & Zelaznik 2004, Green & Wilson 2006). Results indicate that durations decrease with maturation, aligning with the universal tendency for improved articulatory coordination, neurological development, motor learning, and lexical expansion (Nip & Green 2013, Canault et al., 2020). Durational trends start exhibiting adult-like productions by 45–51 months, although the oldest child group did not match adult patterns. This supports evidence that durational values continue decreasing until 10–12 years (Kent and Forner, 1980, Smith et al., 1983). The present durational patterns demonstrate the interaction between phonetic and phonological development, where universal patterns, such as prosodic development, and language-specific factors, such as syllable structure, combine to shape the unique trajectory of speech maturation in JA children.

Regarding lexical stress, the data supported previous findings that stressed syllables tend to be longer than unstressed syllables (Fry 1955, Beckman 1986, Sluijter & van Heuven 1995, Turk & Sawusch 1997, Turk & White 1999). The Bayesian model outputs showed that stress was not a strong predictor for syllable or vowel durations and was not involved in significant two-way interactions such as stress × position and age group × stress. However, it was involved in the three-way interaction in ST and PT. Such results are consistent with studies arguing that although Arabic is considered to have lexical stress, duration does not serve as a strong acoustic correlate (Roach 1982, Heliel 1982, Bertinetto 1989). This is probably attributed to the predictability of Arabic stress and it being determined by factors, such as syllable weight and syllable position within a word, reducing the difference between stressed and unstressed targets (Section 1.1.2). According to acoustic studies, children do not show sufficient durational control to employ the contrastive role of stress but they employ other acoustic cues such as pitch and intensity (Pollock et al., 1993, Kehoe, Stoel-Gammon & Buder 1995). Scholars postulate that language and dialects vary in their stress marking, and one acoustic cue does not account for the discrepancy between the stressed and unstressed environments (Bertinetto 1989, Pollock et al., 1993). Then, the critical role of duration does not suffice to trace stress development patterns of JA child speech and other acoustic cues may be more effective.

The observation of word final lengthening in JA productions offers substantial evidence for prosodic boundary marking (Delattre 1966, Smith 1978, Beckman and Edwards 1990, Robb & Saxman 1990). Nevertheless, the manifestations of word final lengthening vary according to age group, where younger children exaggerate their durations. The data presented acoustic evidence that children demonstrate similar prosodic patterns to those found in adult productions

(Smith 1978), and that prosodic knowledge and sensitivity to word boundaries appear at an early stage of linguistic development (Allen 1983, Robb & Saxman 1990). This prosodic sensitivity is crucial for effective communication (Delattre 1966, Snow 1998, Yeou 2005), demonstrated in the attunement of rhythmic cues that manifest higher prosodic phenomena such as stress assignment (Hayes 1995, Watson 2007). The consistent final lengthening in JA speech substantiates the conventional assumption that boundary marking is a characteristic of normal linguistic child productions, as deficits in marking features may occur in disordered speech development (Allen & Hawkins 1980, Snow 1998).

Additionally, syllable structure plays a significant role in determining the temporal aspects of speech production with light syllables being the shortest, and superheavy syllables being the longest. Such results have implications on the distinctive and complex nature of superheavy syllables as they contain additional segments contributing to their moraic weight (Newman 1972, McCarthy & Prince 1986, Watson 2000). Moreover, the current results offer acoustic evidence supporting the trimoraic ban discussed in the literature. The patterning of superheavy syllables in JA abides by the bimoraic constraint through vowel shortening (Broselow et al., 1995, Khattab and Al-Tamimi 2014, Bamakhramah 2010).

With maturation, the child’s phonological skills expand, allowing for the production of variable and complex structures, as evident in superheavy sub-structures expansion. CVVC syllables were the most frequently produced compared to CVCC and CCVVC syllables. This supports claims in the literature concerning the permissibility of CVVC syllables in different word positions and environments such as derivatives and plural forms (Broselow1992, Watson 2011). CVCC syllables appeared earlier than CCVVC syllables, as coda clusters emerge before onset ones (Alqattan 2015, Al huneety et al., 2023). The coda cluster in CVCC contributes to the syllable's moraic weight, rendering it bimoraic, while the onset cluster in CCVVC is weightless (Hayes 1995, Watson 2007). Since children find it challenging to produce clusters, examples of vowel epenthesis emerged (e.g., /kalb/-/kalib/, /ktaab/-/kitaab/). This contributes to the literature analysing dialectical restructuring and simplification of superheavy syllables reducing their frequency (Ammar 2002).

The analysis of superheavy syllables shows that with maturation, children become less reliant on producing superheavy syllables exclusively in monosyllabic words; instead, they incorporate them in disyllabic and multisyllabic words (Alqattan 2015, Mashaqba et al., 2019, Al huneety et al., 2023). Increased variability and complexity with age are evident in previous accounts of Arabic child speech (Dyson & Amayreh 2000, Ammar 2002, Ayyad 2011, Alqattan 2015) and cross-linguistical ones (Stoel-Gammon 1987, Hua & Dodd 2000). Moreover, the prosodic templates of child speech increase with the gradual expansion of vocabulary (Hua & Dodd 2000, Alqattan 2015). Superheavy syllable analysis revealed that duration in Arabic is not a strong indicator of lexical stress as in West Germanic languages (Firth 1948, Dauer 1983). Acoustic data show that independent of lexical stress, syllable structure, and position effects can account for durational patterning in child speech.

Finally, the consistent word final lengthening in superheavy syllables observed in this study validates the importance of prosodic boundary marking in child phonological and phonetic development (Morgan 1986, Nelson et al., 1989, Allen & Hawkins 1980). Establishing that syllable position has been the only consistently strong predictor for superheavy syllable productions denotes prosodic boundary marking as more salient than stress marking. This can impact future research on Arabic child development to consider emphasizing the durational boundary-marking as an aspect of categorizing normal and abnormal productions.

5.7.Clinical Implications

The present findings have clinical implications related to employing durational measures to evaluate the phonological development of child speech. First, employing durational analysis can be used to diagnose abnormal or delayed speech productions. Monitoring durations enables the evaluation of prosodic skills and the detection of discrepancies not captured by qualitative observations. This approach provides a reliable statistical reference for quantifying and assessing the temporal aspects of child speech.

Longer durations are expected in normal children up to 38–44 months, then durations should start decreasing and trending towards adult-like productions. Children with deficits prolong their syllables and vowels, compensating for unclear or weak productions (Brown et al., 2005, Civier, Tasko & Guenther 2010). Current data can be used to establish baseline measurements to be used as a diagnostic tool and progress monitoring technique during initial therapy sessions. Tracking durations over time provides feedback on treatment effectiveness, particularly improving durational precision when applicable.

Second, assessing the child's superheavy syllable productions can be an early identification method of normal and abnormal speech, as difficulties in producing these syllables may be indicative of speech delays and deficits. Clinicians could employ superheavy syllables in their diagnosis and assessment, identifying delays or difficulties in their production. The phonological processes could be viewed as part of normal developmental trajectory rather than recognizing them as errors.

5.8.Limitations and Directions for Future Research

Multiple aspects of data collection have been impacted by the COVID-19 pandemic. First, the shift from in-person to online recordings restricted the experimenter’s ability to fully engage participants and ensure their attentiveness leading to potential distractions and boredom. Second, although the experimenter took measures to ensure data quality, holding online recording sessions has introduced limitations in addressing audio quality and clarity. Third, the restrictions on in-person interactions and data collection resulted in a smaller sample size, which may influence the generalizability of findings.

Tasks ranging from naturally and visually elicited speech to more controlled productions were chosen to minimize confounding factors of task effects. Nevertheless, the fixed task order influenced the participants' performances, as children started refusing to cooperate and repeat words in the last task, RT. For superheavy syllables, data analysis was carried out for syllables and vowels, but not for different segmental types. If segmental types were considered, this could have contributed to tracing cues of mora sharing, further supporting vowel shortening evident in non-final superheavy syllables. Further analysis of the trimoraicity ban and the phonological behaviour of these syllables can be carried out in the future.

The current data contributes to the scarce sources of acoustic-based research on Arabic child speech. Greater emphasis on acoustic data in Arabic linguistic research is essential to establish a reliable reference facilitating further investigations. The absence of Arabic-based acoustic corpora, particularly for child speech, poses possible limitations to researchers. Further investigations can focus on child productions across dialects, highlighting the effect of dialectical prosodic rules and constraints, such as cluster permissibility, epenthesis, and syncope on superheavy productions. Finally, future research should consider addressing superheavy syllable productions in participants younger than 24–30 months and older than 66–72 months.

References

Abd-el-Jawad, H.R. (1986) *The Emergence of an Urban Dialect in the Jordanian Urban Centers*. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Abdoh, E.M.A. (2011) *A study of the phonological structure and representation of first words in Arabic*. Doctoral dissertation, University of Leicester.

Abdul-Karim, K.W. (1980) *Aspects of the phonology of Lebanese Arabic*. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Abercrombie, D. (1967) *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Abou-Elsaad, T., Baz, H. and El-Banna, M. (2009) 'Developing an articulation test for Arabic-speaking school-age children', *Folia Phoniatrica et Logopaedica*, 61(5), pp. 275-282.

Abu Guba, M.N. (2018) *Syllable structure and syllabification in Ammani Arabic: External evidence from the adaptation of English loanwords*. Doctoral dissertation, University of Kansas.

AbuAbbas, K.H. (2003) *Topics in the phonology of Jordanian Arabic: An optimality theory perspective*. Doctoral dissertation, University of Kansas.

Abu-Abbas, K.H., Zuraiq, W.M. and Abdel-Ghafer, O.A. (2011) 'Geminates and long consonants in Jordanian Arabic', *International Journal of Linguistics*, 3(1), pp. 1-17.

Abu-Rabia, S. (2000) 'Effects of exposure to literary Arabic on reading comprehension in a diglossic situation', *Reading and Writing*, 13, pp. 147-157.

Abu-Salim, I.M. (1980) 'Epenthesis and geminate consonants in Palestinian Arabic', *Studies in the Linguistic Sciences*, 10(2), pp. 1-11.

Abu-Salim, I.M. (1982) *A reanalysis of some aspects of Arabic phonology: a metrical approach*. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Adra, M.A. (1999) *Identity effects and opacity in Syrian Arabic: An optimality theory analysis*. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Ahn, M. (2000) *Phonetic and functional bases of syllable weight for stress assignment*. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Al Huneety, A., Mashaqba, B., Alhala, M.A., Abu Guba, M.N. and Al-Shdifat, K.G. (2023) 'Acquisition of stress in the speech of Ammani Arabic-speaking children', *Ampersand*, 10, p. 1094.

Al Huneety, A.I., Mashaqba, B.M., Al-Shdifat, K.G., Khasawneh, E.A. and Thnaibat, B. (2023) 'Multisyllabic word production by Jordanian Arabic speaking children', *Speech, Language and Hearing*, 26(4), pp. 249-265.

Al Mashaqba, B.M. (2015) *The phonology and morphology of Wadi Ramm Arabic*. Doctoral dissertation, University of Salford (United Kingdom).

Al Tamimi, Y.A. and Al Shboul, Y. (2013) 'Is the phonotactics of the Arabic complex coda sonority-based?', *Journal of King Saud University - Languages and Translation*, 25(1), pp. 21-33.

Alahmari, M. (2021) 'Prosodic adaptation of superheavy syllables in Arabic loanwords in Turkish: An optimality-theoretic analysis', *Journal of King Abdulaziz University - Arts and Humanities*, 29(1), pp. 581-596.

AlAmro, M. (2016) 'Syllabification in Najdi Arabic: A constraint-based analysis', *Arab World English Journal (AWEJ)*, 6, pp. 1-17.

Al-Ani, S.H. (1970) *An acoustical and physiological investigation of the Arabic/E/*. Doctoral dissertation, Indiana University.

Al-Btoush, M.A. (2005) *The acquisition of stress in Jordanian Arabic and English among children*. MA Thesis, University of Jordan.

Al-Deaibes, M. (2016) *The phonetics and phonology of assimilation and gemination in Rural Jordanian Arabic*. Doctoral dissertation, University of Manitoba.

Alhammad, R. (2018) 'The role of the Syllable Contact Law-Semisyllable (SCL-SEMI) in the coda clusters of Najdi Arabic and other languages', *Proceedings of the Linguistic Society of America*, 3, pp. 1-17.

Alhoody, M. and Aljutaily, M. (2020) 'Some characteristics of syllable structure in Qassimi Arabic (Q.A.): An optimality theoretic framework', *International Journal of English Linguistics*, *10*(4), pp.193-202.

Ali, A., Rasheed, A., Siddiqui, A.A., Naseer, M., Wasim, S. and Akhtar, W. (2015) 'Non-parametric test for ordered medians: The Jonckheere Terpstra test', *International Journal of Statistics in Medical Research*, 4(2), p. 203.

Allen, G.D. and Hawkins, S. (1980) 'Phonological rhythm: Definition and development', in Yeni-Komshian, G.H., Kavanagh, J.F. and Ferguson, C.A. (eds.) *Child Phonology*. New York: Academic Press, pp. 227-256.

Allen, J.F. (1983) 'Maintaining knowledge about temporal intervals', *Communications of the ACM*, 26(11), pp. 832-843.

Al-Mozainy, H.Q. (1981) *Vowel alternations in a Bedouin Hijazi Arabic dialect: Abstractness and stress*. PhD dissertation, The University of Texas at Austin.

Alotaibi, Y. and Meftah, A. (2013) 'Review of distinctive phonetic features and the Arabic share in related modern research', *Turkish Journal of Electrical Engineering & Computer Sciences*, 21(5), pp. 1426-1439.

Alqattan, S. (2015) *Early phonological acquisition by Kuwaiti Arabic children*. Doctoral dissertation, Newcastle University.

Al-Saidat, E. and Al-Momani, I. (2010) 'Future markers in Modern Standard Arabic and Jordanian Arabic: A contrastive study', *European Journal of Social Sciences*, 12(3), pp. 397-408.

Al-Sughayer, K.I. (1990) *Aspects of comparative Jordanian and Modern Standard Arabic phonology*. PhD dissertation, Michigan State University.

Altakhaineh, A.R.M. and Zibin, A. (2014) 'Phonologically conditioned morphologically process in Modern Standard Arabic: An analysis of Al-ibdal "substitution" in ftaʕal pattern using prosodic morphology', *International Journal of English Language and Linguistics Research*, 2(1), pp. 1-16.

Al-Tamimi, F., Abu-Abbas, K. and Tarawnah, R. (2010) 'Jordanian Arabic final geminates: An experimental clinical phonetic study', *Poznań Studies in Contemporary Linguistics*, 46(2), pp. 111-128. doi: 10.2478/v10010-010-0006-6.

Al-Thamery, A.A. and Ibrahim, M.A. (2005) 'Word-stress in Arabic: A phonological study from a generative perspective', *ADAB AL-BASRAH*, (38), pp. 1-17.

Al-Wer, E. (2002) 'Jordanian and Palestinian dialects in contact: Vowel raising in Amman', *Contributions to the Sociology of Language*, 86, pp. 63-80.

Al-Wer, E. (2007) 'The formation of the dialect of Amman: From chaos to order', in Miller, C., Caubet, D. and Watson, J. (eds.) *Arabic in the City: Issues in Dialect Contact and Language Variation*. London: Routledge, pp. 69-90.

Al-Zabibi, M. (1990) *An acoustic-phonetic approach in automatic Arabic speech recognition*. PhD dissertation, Loughborough University.

Amayreh, M.M. and Dyson, A.T. (1998) 'The acquisition of Arabic consonants', *Journal of Speech, Language, and Hearing Research*, 41(3), pp. 642-653.

Amayreh, M.M. and Dyson, A.T. (2000) 'Phonetic inventories of young Arabic-speaking children', *Clinical Linguistics & Phonetics*, 14(3), pp. 193-215.

Amer, F.H., Adaileh, B.A. and Rakhieh, B.A. (2011) 'Arabic diglossia', *Argumentum*, 7, pp. 19-36.

Ammar, W. and Morsi, R. (2006) 'Phonological development and disorders: Colloquial Egyptian Arabic', in Ball, M.J. and Gibbon, F. (eds.) *Phonological Development and Disorders in Children*. Clevedon: Multilingual Matters, pp. 204-232.

Arias, J. and Lleó, C. (2009) 'Comparing the representation of iambs by monolingual German, monolingual Spanish, and bilingual German-Spanish children', in Pešková, A. and Kireeva, L. (eds.) *Convergence and Divergence in Language Contact Situations*. Hamburg: Helmut Buske Verlag, pp. 205-234.

Arvaniti, A. (2009) 'Rhythm, timing, and the timing of rhythm', *Phonetica*, 66(1-2), pp. 46-63.

Ayyad, H. and Bernhardt, B.M. (2009) 'Phonological development of Kuwaiti Arabic: Preliminary data', *Clinical Linguistics & Phonetics*, 23(11), pp. 794-807. doi: 10.3109/02699200903236493.

Ayyad, H.S. (2011) *Phonological development of typically developing Kuwaiti Arabic-speaking preschoolers*. PhD dissertation, University of British Columbia. Available at: http://hdl.handle.net/2429/34935.

Bagemihl, B. (1991) 'Syllable structure in Bella Coola', *Linguistic Inquiry*, 22(4), pp. 589-646.

Ballard, K.J., Djaja, D., Arciuli, J., James, D.G. and van Doorn, J. (2012) 'Developmental trajectory for production of prosody: Lexical stress contrastivity in children ages 3 to 7 years and in adults', *Journal of Speech, Language, and Hearing Research*, 55(6), pp. 1820-1835.

Bamakhramah, M.A. (2010) *Syllable Structure in Arabic Varieties with a Focus on Superheavy Syllables*. ERIC.

Bamakhramah, M.A. (2014) 'Superheavy syllables and the concept iltiqāʔ as-sākinayn in traditional Arabic linguistics', *Journal of Arts*, 26(2), pp. 1-15.

Banse, R. and Scherer, K.R. (1996) 'Acoustic profiles in vocal emotion expression', *Journal of Personality and Social Psychology*, 70(3), pp. 614-636.

Beckman, M.E. (1986) *Stress and Non-Stress Accent*. Dordrecht: Foris Publications.

Beckman, M.E., Yoneyama, K. and Edwards, J. (2003) 'Language-specific and language-universal aspects of lingual obstruent productions in Japanese-acquiring children', *Journal of the Phonetic Society of Japan*, 7(2), pp. 18-28.

Benhallam, A. (1980) *Syllable structure and rule types in Arabic*. PhD dissertation, University of Florida.

Berkovits, R. (1994) 'Durational effects in final lengthening, gapping, and contrastive stress', *Language and Speech*, 37(3), pp. 237-250.

Bernhardt, B.M. and Stemberger, J.P. (2020) 'Phonological development', in Schwartz, R.G. (ed.) *Child Bilingualism and Second Language Learning: Multidisciplinary Perspectives*. Philadelphia: John Benjamins Publishing, pp. 223-244.

Bertinetto, P.M. (1989) 'Reflections on the dichotomy "stress" vs. "syllable-timing"', *Revue de Phonétique Appliquée*, 91(93), pp. 99-130.

Bertoncini, J. and Mehler, J. (1981) 'Syllables as units in infant speech perception', *Infant Behavior and Development*, 4, pp. 247-260.

Best, C.T. (2017) 'Speech perception in infants: Propagating the effects of language experience', in Traxler, M.J. and Gernsbacher, M.A. (eds.) *The Handbook of Psycholinguistics*. Hoboken, NJ: Wiley, pp. 470-490.

Biadsy, F., Hirschberg, J.B. and Habash, N.Y. (2009) 'Spoken Arabic dialect identification using phonotactic modeling', *Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5337-5340.

Bion, R.A., Benavides-Varela, S. and Nespor, M. (2011) 'Acoustic markers of prominence influence infants’ and adults’ segmentation of speech sequences', *Language and Speech*, 54(1), pp. 123-140.

Boersma, P. and Weenink, D. (2020) *Praat: Doing phonetics by computer* [Computer program]. Version 6.1.16, retrieved 20 December 2020 from http://www.praat.org/.

Bokhari, H.A. (2020) *A Comprehensive Analysis of Coda Clusters in Hijazi Arabic: An Optimality-Theoretic Perspective*. PhD dissertation, King Abdulaziz University.

Broselow, E., Huffman, M., Chen, S. and Hsieh, R. (1995) 'The timing structure of CVVC syllables', *Amsterdam Studies in The Theory and History of Linguistic Science Series 4*, pp. 119.

Brown, S., Ingham, R.J., Ingham, J.C., Laird, A.R. and Fox, P.T. (2005) ‘Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies’, *Human Brain Mapping*, 25(1), pp. 105-117.

Btoosh, M.A. (2006) 'Constraint interactions in Jordanian Arabic phonotactics: An optimality-theoretic approach', *Journal of Language and Linguistics,* 5(2), pp. 102-221.

Bunta, F. and Ingram, D. (2007) 'The acquisition of speech rhythm by bilingual Spanish-and English-speaking 4-and 5-year-old children', *Journal of Speech, Language, and Hearing Research*.

Bürkner, P.C. and Charpentier, E. (2020) 'Modelling monotonic effects of ordinal predictors in Bayesian regression models', *British Journal of Mathematical and Statistical Psychology*, 73(3), pp. 420-451.

Bürkner, P.C. (2017) 'Advanced Bayesian multilevel modeling with the R package brms', *arXiv preprint,* arXiv:1705.11123.

Campbell, W.N. (1992) 'Syllable-based segmental duration', *Talking machines: Theories, models, and designs*, pp. 211-224.

Canault, M., Yamaguchi, N., Paillereau, N., Krzonowski, J., Johanna-Pascale, R., Dos Santos, C. and Sophie, K. (2020) 'Syllable duration changes during babbling: a longitudinal study of French infant productions', *Journal of Child Language*, 47(6), pp. 1207-1227.

Carlisle, R.S. (2001) 'Syllable structure universals and second language acquisition', *International Journal of English Studies*, 1(1), pp. 1-19.

Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P. and Riddell, A. (2017) 'Stan: A probabilistic programming language', *Journal of Statistical Software,* 76(1).

Catford, J.C. (1977) 'Mountain of tongues: the languages of the Caucasus', *Annual Review of Anthropology*, 6, pp. 283-314.

Chahal, D. (2003) 'Phonetic cues to prominence in Lebanese Arabic', In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2067-2070).

Chermak, G.D. and Schneiderman, C.R. (1985) 'Speech timing variability of children and adults', *Journal of Phonetics*, *13*(4), pp.477-80.

Chiang, D., Diab, M., Habash, N., Rambow, O. and Shareef, S. (2006) 'Parsing Arabic dialects', pp. 369.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. and Halle, M. (1968) *The Sound Pattern of English*. New York: Harper & Row.

Chomsky, N. (1995) 'Language and nature', *Mind*, 104(413), pp. 1-61.

Chu, M. and Feng, Y. (2001) 'Study on factors influencing durations of syllables in Mandarin', *Journal of Chinese Linguistics*, pp. 15-37.

Civier, O., Tasko, S.M. and Guenther, F.H. (2010) ‘Overreliance on auditory feedback may lead to sound/syllable repetitions: simulations of stuttering and fluency-inducing conditions with a neural model of speech production’, *Journal of Fluency Disorders*, 35(3), pp. 246-279.

Clark, E.V. and Casillas, M. (2015) 'First language acquisition', in *The Routledge Handbook of Linguistics*. London: Routledge, pp. 311-328.

Clements, G.N. and Keyser, S.J. (1983) *CV Phonology: A Generative Theory of the Syllable*. Cambridge, MA: MIT Press.

Cleveland, R.L. (1963) 'A classification for the Arabic dialects of Jordan', *Bulletin of the American Schools of Oriental Research*, 171(1), pp. 56-63.

Cooper, W.E. and Paccia-Cooper, J. (1980) *Syntax and Speech*. Cambridge, MA: Harvard University Press.

Correia, S. (2009) 'The acquisition of primary word stress in European Portuguese', *Phonology*, 26(1), pp. 53-82.

Crossley, T. (2023) *Superheavy Syllables in Jordanian Arabic; A Case Against a Language-Wide Trimoraic Ban* (Doctoral dissertation, The University of Arizona).

Crystal, D. (1986) 'Prosodic development', in Fletcher, P. and Garman, M. (eds.) *Language Acquisition*. Cambridge: Cambridge University Press, pp. 174-197.

Cummins, F. (1999) 'Rhythmic constraints on stress timing in English', *Journal of Phonetics*, 27(2), pp. 145-171.

Cutler, A. (2005) 'Lexical stress', in Brown, K. (ed.) *Encyclopedia of Language and Linguistics*. 2nd edn. Oxford: Elsevier, pp. 128-134.

Cutler, A. and Swinney, D.A. (1987) 'Prosody and the development of comprehension', *Journal of Child Language*, 14(1), pp. 145-167.

Cutler, A. and Butterfield, S. (1992) 'Rhythmic cues to speech segmentation: Evidence from juncture misperception', *Journal of Memory and Language*, 31(2), pp. 218-236.

Cutler, A., Dahan, D. and Van Donselaar, W. (1997) 'Prosody in the comprehension of spoken language: A literature review', *Language and Speech*, 40(2), pp. 141-201.

Daana, H.A. (2009) *The development of consonant clusters, stress and plural nouns in Jordanian Arabic child language* (Doctoral dissertation, The University of Essex).

Daana, H.A. (2017) 'The development of coda consonants in the speech of a bilingual child: A case study', *International Journal of Linguistics*, 9(5), pp. 149. doi: 10.5296/ijl.v9i5.11747.

Daana, H.A. (2018) 'Handling coda clusters formed by geminates in different Arabic dialects', *International Journal of Linguistics and Literature (IJLL)*, 7, pp. 65-80.

Daana, H.A. and Khrais, S.M. (2018) 'The acquisition of English and Arabic onset cluster: A case study', *English Linguistics Research*, 7(1), pp. 13. doi: 10.5430/elr.v7n1p13.

Darwin, C.J. (1975) 'On the dynamic use of prosody in speech perception', in *Structure and Process in Speech Perception: Proceedings of the Symposium on Dynamic Aspects of Speech Perception held at IPO, Eindhoven, Netherlands, August 4–6, 1975*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 178-194.

Dauer, R.M. (1983) 'Stress-timing and syllable-timing reanalyzed', *Journal of Phonetics*, 11(1), pp. 51-62.

Davis, B.L., MacNeilage, P.F., Matyear, C.L. and Powell, J.K. (2000) 'Prosodic correlates of stress in babbling: An acoustical study', *Child Development*, 71(5), pp. 1258-1270. doi: 10.1111/1467-8624.00227.

Davis, S. (2011) 'Geminates', in van Oostendorp, M., Ewen, C.J., Hume, E. and Rice, K. (eds.) *The Blackwell Companion to Phonology*. Malden, MA: Wiley-Blackwell, pp. 837-859.

Davis, S. and Ragheb, M. (2014) 'Geminate representation in Arabic', *Perspectives on Arabic Linguistics XXIV-XXV*, 1, pp. 3-19.

De Jong, K. and Zawaydeh, B. (2002) 'Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration', *Journal of Phonetics*, 30(1), pp. 53-75.

de Jong, K. and Zawaydeh, B.A. (1999) 'Stress, duration, and intonation in Arabic word-level prosody', *Journal of Phonetics*, 27(1), pp. 3-22. doi: 10.1006/jpho.1998.0088.

de Jong, K. (1994) 'The correlation of linguistic structure and prosodic structure', *Journal of Phonetics*, 22(2), pp. 141-165.

de Manrique, A.M.B. and Signorini, A. (1983) 'Segmental duration and rhythm in Spanish', *Journal of Phonetics*, 11(2), pp. 117-128.

Delattre, P. (1966) 'A comparison of syllable length conditioning among languages', *International Journal of American Linguistics*, 32(4), pp. 272-280.

Demuth, K. (1995) 'Markedness and the development of prosodic structure', in Archibald, J. (ed.) *The Acquisition of Non-linear Phonology*. Amsterdam: John Benjamins, pp. 161-180.

Demuth, K. (1996) 'The prosodic structure of early words', in Morgan, J.L. and Demuth, K. (eds.) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 171-184.

Demuth, K. (2009) 'The prosody of syllables, words and morphemes', in *Cambridge Handbook on Child Language*. Cambridge: Cambridge University Press, pp. 183-198.

Demuth, K. and Fee, E.J. (1995) 'Minimal prosodic words in early phonological development', *Ms*, Brown University and Dalhousie University.

Demuth, K. and Morgan, J. (1996) 'The prosodic structure of early words', in Morgan, J.L. and Demuth, K. (eds.) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 171-184.

Diab, M. and Habash, N. (2007) 'Arabic dialect processing tutorial', In *Proceedings of the human language technology conference of the NAACL, companion volume: tutorial abstracts* (pp. 5-6).

Diab, M., Ghoneim, M. and Habash, N. (2007) 'Arabic diacritization in the context of statistical machine translation', in *Proceedings of Machine Translation Summit XI: Papers*.

Dickinson, D.K. and Tabors, P.O. (2002) 'Fostering language and literacy in classrooms and homes', *Young Children*, 57(2), pp. 10-19.

Dodd, B. and Gillon, G. (1997) 'The nature of the phonological deficit underlying disorders of spoken and written language', in *Cross-Language Studies of Learning to Read and*

Spell: Phonologic and Orthographic Processing. Dordrecht: Springer Netherlands, pp. 53-70.

Dodd, B., Holm, A., Hua, Z. and Crosbie, S. (2003) 'Phonological development: a normative study of British English-speaking children', *Clinical Linguistics & Phonetics*, 17(8), pp. 617-643.

Dresher, B.E. and Kaye, J.D. (1990) 'A computational learning model for metrical phonology', *Cognition*, 34(2), pp. 137-195.

Dresher, E.B. (2004) 'On the acquisition of phonological representations', in *Proceedings of the Workshop on Psycho-Computational Models of Human Language Acquisition*, pp. 43-50)

Dresher, B.E. (2019) 'Contrastive feature hierarchies in phonology: Variation and universality', in *Variable Properties in Language: Their Nature and Acquisition*. Georgetown University Press, pp. 13-26.

Duanmu, S. (1994) 'Syllabic weight and syllabic duration: A correlation between phonology and phonetics', *Phonology*, 11(1), pp. 1-24.

Dyson, A.T. (1988) 'Phonetic inventories of 2-and 3-year-old children', *Journal of Speech and Hearing Disorders*, 53(1), pp. 89-93.

Echols, C.H. and Newport, E.L. (1992) 'The role of stress and position in determining first words', *Language Acquisition*, 2(3), pp. 189-220. doi: 10.1207/s15327817la0203_1.

Edwards, J., Beckman, M.E. and Fletcher, J. (1991) 'The articulatory kinematics of final lengthening', *The Journal of the Acoustical Society of America*, 89(1), pp. 369-382.

Eilers, R.E., Bull, D.H., Oller, D.K. and Lewis, D.C. (1984) 'The discrimination of vowel duration by infants', *The Journal of the Acoustical Society of America*, 75(4), pp. 1213-1218.

Eimas, P.D. (1975) 'Auditory and phonetic coding of the cues for speech: Discrimination of the [rl] distinction by young infants', *Perception & Psychophysics*, 18(5), pp. 341-347.

Eimas, P.D., Siqueland, E.R., Jusczyk, P. and Vigorito, J. (1971) 'Speech perception in infants', *Science*, 171(3968), pp. 303-306.

Elashhab, S. (2018) 'Epenthesis and syncope in some Arabic dialects', *International Journal of Humanities and Social Science Invention*, 7(1), pp. 1-9.

Elgibali, A. (1996) 'Phonetics and language teaching', in *Proceedings of the KSPS Conference*, The Korean Society of Phonetic Sciences and Speech Technology, pp. 167-167.

Embarki, M., Yeou, M., Guilleminot, C. and Al Maqtari, S. (2007) 'An acoustic study of coarticulation in modern standard Arabic and dialectal Arabic: Pharyngealized vs. nonpharyngealized articulation', in *ICPhS XVI*, pp. 141-146.

Embarki, M. (2013) 'Phonetics', in Owens, J. (ed.) *The Oxford Handbook of Arabic Linguistics*. Oxford: Oxford University Press, pp. 23-44.

Espy-Wilson, C.Y. (1994) 'A feature-based semivowel recognition system', *The Journal of the Acoustical Society of America*, 96(1), pp. 65-72.

Farwaneh, S. (1992) 'Directional syllabification and syllable structure in Arabic dialects', in *66th Annual Meeting of the Linguistic Society of America*.

Fear, B.D., Cutler, A. and Butterfield, S. (1995) 'The strong/weak syllable distinction in English', *The Journal of the Acoustical Society of America*, 97(3), pp. 1893-1904.

Feldman, H.M. (2019) 'How young children learn language and speech', *Pediatrics in Review*, 40(8), pp. 398-411.

Ferguson, C.A. and Farwell, C.B. (1975) 'Words and sounds in early language acquisition: English initial consonants in the first fifty words', *Language*, 51(2), pp. 419-439.

Ferguson, C.A. and Slobin, D.I. (1973) *Studies of child language development*. New York: Holt, Rinehart and Winston.

Ferguson, C.A. (1983) 'Reduplication in child phonology', *Journal of Child Language*, 10(1), pp. 239-243.

Ferguson, C.A. (1962) 'The languages of children', *Scientific American*, 207(3), pp. 87-98.

Ferguson, C.A. (1977) 'Baby talk as a simplified register', in Snow, C.E. and Ferguson, C.A. (eds.) *Talking to Children: Language Input and Acquisition*. Cambridge: Cambridge University Press, pp. 209-235.

Ferguson, C.A., Menn, L. and Stoel-Gammon, C. (eds.) (1992) *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press.

Ferrat, K. and Guerti, M. (2017) 'An experimental study of gemination in Arabic language', *Archives of Acoustics*, 42(4), pp. 571-578.

Féry, C. (1998) 'German word stress in optimality theory', *Journal of Comparative Germanic Linguistics*, 2(2), pp. 101-142.

Fikkert, P. (1994) *On the Acquisition of Prosodic Structure*. Holland Institute of Generative Linguistics.

Firth, J.R. (1948) 'Sounds and prosodies', *Transactions of the Philological Society,* 47(1), pp. 127-152.

Fletcher, J. (1991) 'Rhythm and final lengthening in French', *Journal of Phonetics*, 19(2), pp. 193-212.

Fletcher, J. (2010) 'The prosody of speech: Timing and rhythm', in Hardcastle, W.J., Laver, J. and Gibbon, F.E. (eds.) *The Handbook of Phonetic Sciences*. Oxford: Wiley-Blackwell. Available at: https://doi.org/10.1002/9781444317251.ch15.

Fowler, C.A. (1990) 'Final lengthening and prosodic structure', in *Speech Production and Speech Modelling*, pp. 171-186.

Franke, M. and Roettger, T.B. (2019) 'Bayesian regression modeling (for factorial designs): A tutorial'.

Fry, C.L. (1966) 'Training children to communicate to listeners', *Child Development*, 37(2), pp. 675-685.

Fry, D.B. (1955) 'Duration and intensity as physical correlates of linguistic stress', *The Journal of the Acoustical Society of America*, 27(4), pp. 765-768.

Fry, D.B. (1958) 'Experiments in the perception of stress', *Language and Speech*, 1(2), pp. 126-152.

Gathercole, S.E. and Baddeley, A.D. (1990) 'Phonological memory deficits in language disordered children: Is there a causal connection?', *Journal of Memory and Language*, 29(3), pp. 336-360.

Gay, T. (1978) 'Articulatory units: Segments or syllables', in *Syllables and Segments*, pp. 121-132.

Gelman, A. and Shalizi, C.R. (2013) 'Philosophy and the practice of Bayesian statistics', *British Journal of Mathematical and Statistical Psychology*, 66(1), pp. 8-38.

Gelman, A., Hill, J. and Vehtari, A. (2021) *Regression and Other Stories*. Cambridge University Press.

Gerken, L. (1991) 'The metrical basis for children's subjectless sentences', *Journal of Memory and Language*, 30(4), pp. 431-451.

Gerken, L. (1994a) 'A metrical template account of children's weak syllable omissions from multisyllabic words', *Journal of Child Language*, 21(3), pp. 565-584. doi: 10.1017/S0305000900009466.

Gerken, L. (1994b) 'Young children's representation of prosodic phonology: Evidence from English-speakers' weak syllable productions', *Journal of Memory and Language*, 33(1), pp. 19-38.

Gerken, L. (1996) 'Prosody's role in language acquisition and adult parsing', *Journal of Psycholinguistic Research*, 25(2), pp. 345-356. doi: 10.1007/BF01708577.

Gerken, L., Landau, B. and Remez, R.E. (1990) 'Function morphemes in young children's speech perception and production', *Developmental Psychology*, 26(2), pp. 204-216.

Goad, H. and Brannen, K. (2003) 'Phonetic evidence for phonological structure in syllabification', in van de Weijer, J., van Heuven, V.J. and van der Hulst, H. (eds.) *The Phonological Spectrum, Volume II: Suprasegmental Structure*. Amsterdam: John Benjamins, pp. 3-30.

Goldsmith, J. (1990) *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell.

Gordon, M. (2007) *Syllable weight: Phonetics, phonology, typology*. Routledge.

Grabe, E. and Low, E.L. (2002) 'Durational variability in speech and the rhythm class hypothesis', in Gussenhoven, C. and Warner, N. (eds.) *Laboratory Phonology 7*. Berlin: Mouton de Gruyter, pp. 515-546.

Great Ormond Street Hospital for Children NHS Foundation Trust (2016) *Speech & language development (from birth to 12 months)*. NHS. Available at: https://www.gosh.nhs.uk/file/2924/download (Accessed: 12 February 2021).

Green, J.R., Moore, C.A. and Reilly, K.J. (2002) 'The sequential development of jaw and lip control for speech', *Journal of Speech, Language, and Hearing Research*, 45(1), pp. 66-79.

Green, J.R. and Wilson, E.M. (2006) 'Spontaneous facial motility in infancy: A 3D kinematic analysis', *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 48(1), pp. 16-28.

Greenberg, S., Carvey, H., Hitchcock, L. and Chang, S. (2003) 'Temporal properties of spontaneous speech—a syllable-centric perspective', *Journal of Phonetics*, 31(3-4), pp. 465-485.

Guba, M.N.A. (2016) *Phonological adaptation of English loanwords in Ammani Arabic*. Doctoral dissertation, University of Salford (United Kingdom).

Guenther, F.H. and Gjaja, M.N. (1996) 'The perceptual magnet effect as an emergent property of neural map formation', *The Journal of the Acoustical Society of America*, 100(2), pp. 1111-1121.

Gussenhoven, C. (2004) *The Phonology of Tone and Intonation*. Cambridge University Press.

Hajek, J. and Stevens, M. (2008) 'Vowel duration, compression and lengthening in stressed syllables in Central and Southern varieties of standard Italian', in *ISCA*.

Hall, T.A. (2002) 'The distribution of superheavy syllables in Standard German', *Linguistic Review*, 19(4), pp. 377-420. doi: 10.1515/tlir.2002.004.

Ham, W. (2001) *Phonetic and Phonological Aspects of Geminate Timing*. New York: Routledge.

Harrington, J. (2010) 'Acoustic phonetics', in Hardcastle, W.J., Laver, J. and Gibbon, F.E. (eds.) *The Handbook of Phonetic Sciences*. Oxford: Wiley-Blackwell, pp. 81-129.

Hawkins, J.A. (1979) 'Implicational universals as predictors of word order change', *Language*, 55(3), pp. 618-648.

Hayes, B. (1989) 'Compensatory lengthening in moraic phonology', *Linguistic Inquiry*, 20(2), pp. 253-306.

Hayes, B. (1995) *Metrical Stress Theory: Principles and Case Studies*. University of Chicago Press.

Hayes, P.J. (1981) 'The logic of frames', in *Readings in Artificial Intelligence*, pp. 451-458. Morgan Kaufmann.

Heliel, M.H. (1982) 'Stress timing in modern literary Arabic', *al-'Arabiyya*, 15(1), pp. 90-107.

Hochberg, J.G. (1988) 'Learning Spanish stress: Developmental and theoretical perspectives', *Language*, 64(4), pp. 683-706.

Hoff, E. (2003) 'The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech', *Child Development*, 74(5), pp. 1368-1378.

Hoijtink, H., Klugkist, I. and Boelen, P.A. (2008) *Bayesian Evaluation of Informative Hypotheses*. Vol. 361. New York: Springer.

Hua, Z. and Dodd, B. (2000) 'The phonological acquisition of Putonghua (Modern Standard Chinese)', *Journal of Child Language*, 27(1), pp. 3-42. doi: 10.1017/S030500099900402X.

Hubbard, K. (1994) *Duration in Moraic Theory*. PhD thesis. University of California, Berkeley.

Huehnergard, J. and Pat-El, N. (2019) 'Introduction to the Semitic Languages and their History', in *The Semitic Languages*, pp. 1-21.

Huneety, A. (2015) *The Phonology and Morphology of Wadi Mousa Arabic*. Doctoral dissertation, University of Salford.

Huneety, A. and Mashaqba, B. (2016) 'Stress rules in loan words in Bedouin Jordanian Arabic in the north of Jordan: A metrical account', *SKASE Journal of Theoretical Linguistics*, 13(3).

Hyman, L.M. (1985) 'Word domains and downstep in Bamileke-Dschang', *Phonology*, 2(1), pp. 47-83.

Ingram, D. (1989) *First Language Acquisition: Method, Description and Explanation*. Cambridge University Press.

Ingram, D. (1992) 'Early phonological acquisition: A cross-linguistic perspective', in *Phonological Development: Models, Research, Implications*, pp. 423-435.

Inkelas, S. and Zec, D. (1995) *The Phonology-Syntax Connection*. Chicago: University of Chicago Press.

Irshied, O. and Kenstowicz, M. (1984) 'Some phonological rules of Bani-Hassan Arabic: A Bedouin dialect', *Studies in the Linguistic Sciences*, 14(1), pp. 109-147.

Irshied, O.M. (1984) *The Phonology of Arabic: Bani Hassan - A Bedouin Jordanian Dialect*. University of Illinois at Urbana-Champaign.

Ito, J. (1986) *Syllable Theory in Prosodic Phonology*. New York: Garland Publishing.

Itô, J. (1989) 'A prosodic theory of epenthesis', *Natural Language & Linguistic Theory*, 7(2), pp. 217-259.

Jakobson, R. (1968) *Child Language: Aphasia and Phonological Universals*. No. 72. Walter de Gruyter.

Jusczyk, P.W., Houston, D.M. and Newsome, M. (1993) 'The beginnings of word segmentation in English-learning infants', *Cognitive Psychology*, 25(2), pp. 204-234.

Jusczyk, P.W., Cutler, A. and Redanz, N.J. (1993) 'Infants' preference for the predominant stress patterns of English words', *Child Development*, 64(3), pp. 675-687.

Jusczyk, P.W., Hirsh-Pasek, K., Kemler Nelson, D.G., Kennedy, L.J., Woodward, A. and Piwoz, J. (1992) 'Perception of acoustic correlates of major phrasal units by young infants', *Cognitive psychology*, *24*(2), pp.252-293.

Jusczyk, P.W., Houston, D.M. and Newsome, M. (1999) 'The beginnings of word segmentation in English-learning infants', *Cognitive Psychology*, 39(3-4), pp. 159-207.

Jusczyk, P.W., Luce, P.A. and Charles-Luce, J. (1994) 'Infants' sensitivity to phonotactic patterns in the native language', *Journal of Memory and Language*, 33(5), pp. 630-645.

Kager, R. and Zonneveld, W. (1986) 'Schwa, syllables, and extrametricality in Dutch', in *Proceedings of the 12th Annual Meeting of the Berkeley Linguistics Society*, pp. 272-282.

Kager, R. (1999) *Optimality Theory*. Cambridge University Press.

Kager, R.W.J. (1995) 'The metrical theory of word stress', in *Blackwell Handbooks in Linguistics*, 1, pp. 367-402.

Kalaldeh, R. (2018) 'Acoustic analysis of Modern Standard Arabic vowels by Jordanian speakers', *International Journal of Arabic-English Studies*, 18(1), pp. 95-107.

Karmiloff-Smith, A. (1992) ‘Nature, nurture and PDP: Preposterous developmental postulates?’, *Connection Science*, 4(3-4), pp.253-269.

Kehoe, M. (1994) 'Beyond the trochaic constraint: An examination of rhythmic processes in English children's productions of multisyllabic words', in *Proceedings of the 18th Annual Boston University Conference on Language Development*, pp. 350-361.

Kehoe, M. (2001) 'Prosodic Structure and Phonetic Processing in Early Speech Development', *Journal of Child Language*, 28(2), pp. 251-285.

Kehoe, M. and Stoel-Gammon, C. (1997) 'The acquisition of prosodic structure: An investigation of current accounts of children's prosodic development', *Language*, 73(1), pp. 113-144. doi: 10.2307/416597.

Kehoe, M. and Lleó, C. (2003) 'The acquisition of syllable types in monolingual and bilingual children learning German and Spanish', in van de Weijer, J., van Heuven, V.J. and van der Hulst, H. (eds.) *The Phonological Spectrum, Volume II: Suprasegmental Structure*. Amsterdam: John Benjamins, pp. 291-310.

Kehoe, M., Stoel-Gammon, C. and Buder, E.H. (1995) 'Acoustic correlates of stress in young children's speech', *Journal of Speech, Language, and Hearing Research*, 38(2), pp. 338-350.

Kehoe, M.M. (2001) 'Prosodic patterns in children’s multisyllabic word productions', in *Proceedings of the 25th Annual Boston University Conference on Language Development*, pp. 329-340.

Kehoe, M.M. and Stoel-Gammon, C. (2001) 'Development of syllable structure in English-speaking children with particular reference to rhymes', *Journal of Child Language*, 28(2), pp. 393-432. doi: 10.1017/S030500090100469X.

Kenstowicz, M. (1986) 'Notes on syllable structure in three Arabic dialects', *Revue québécoise de linguistique*, 16(1), pp. 101-127.

Kent, R.D. and Read, C. (2002) *The Acoustic Analysis of Speech*. Singular /Thomson Learning (2)

Kent, R.D. (1976) 'Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies', *Journal of Speech and Hearing Research*, 19(3), pp. 421-447.

Kent, R.D. (1992) 'The biology of phonological development', in *Phonological Development: Models, Research, Implications*, pp. 65-90.

Kent, R.D. and Forner, L.L. (1980) 'Speech segment durations in sentence recitations by children and adults', *Journal of Phonetics*, 8(2), pp. 157-168.

Khamis-Dakwar, R. and Froud, K. (2006) 'Lexical processing in two language varieties', in *Perspectives on Arabic Linguistics: Papers from the annual symposium on Arabic linguistics. Volume XX: Kalamazoo, Michigan*, Vol. 290, p. 153.

Khamis-Dakwar, R., Froud, K. and Gordon, P. (2012) 'Acquiring diglossia: Mutual influences of formal and colloquial Arabic on children's grammaticality judgments', *Journal of Child Language*, 39(1), pp. 61-89.

Khattab, G. and Al-Tamimi, J. (2013) 'Influence of geminate structure on early Arabic templatic patterns', in *The Emergence of Phonology: Whole-Word Approaches and Cross-Linguistic Evidence*, p. 374.

Khattab, G. and Al-Tamimi, J. (2014) 'Geminate timing in Lebanese Arabic: The relationship between phonetic timing and phonological structure', *Laboratory Phonology*, 5(2), pp. 231-269. doi: 10.1515/lp-2014-0009.

King, R.D. (1967) 'Functional load and sound change', *Language*, 43(4), pp. 831-852.

Kiparsky, P. (2003) 'Syllables and Moras in Arabic', in *The Syllable in Optimality Theory*. Cambridge University Press, pp. 147-182.

Klatt, D.H. (1975) 'Vowel lengthening is syntactically determined in a connected discourse', *Journal of Phonetics*, 3(3), pp. 129-140.

Klatt, D.H. (1976) 'Linguistic uses of segmental duration in English: Acoustic and perceptual evidence', *The Journal of the Acoustical Society of America*, 59(5), pp. 1208-1221.

Konopczynski, G. (1986) *Du Prélangage au Langage: Acquisition de la Structuration Prosodique*. Doctoral dissertation, Strasbourg 2.

Kubaska, C.A. and Keating, P.A. (1981) 'Word duration in early child speech', *Journal of Speech, Language, and Hearing Research*, 24(4), pp. 615-621.

Kuhl, P.K. (1994) 'Learning and representation in speech and language', *Current Opinion in Neurobiology*, 4(6), pp. 812-822.

Kuhl, P.K. (1979) 'Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories', *The Journal of the Acoustical Society of America*, 66(6), pp. 1668-1679.

Kuhl, P.K. (2004) 'Early language acquisition: Cracking the speech code', *Nature Reviews Neuroscience*, 5(11), pp. 831-843. doi: 10.1038/nrn1533.

Kuhl, P.K. and Meltzoff, A.N. (1996) 'Infant vocalizations in response to speech: Vocal imitation and developmental change', *The Journal of the Acoustical Society of America*, 100(4), pp. 2425-2438.

Kuhl, P.K. (1991) 'Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not', *Perception & Psychophysics*, 50(2), pp. 93-107.

Ladd, D.R. (2008) *Intonational Phonology*. Cambridge University Press.

Ladd, D.R. (1996) *Intonational Phonology*. Cambridge: Cambridge University Press.

Lahey, M. (1974) 'Use of prosody and syntactic markers in children’s comprehension of spoken sentences', *Journal of Speech and Hearing Research*, 17(4), pp. 656-668.

Laks, L. and Berman, R.A. (2014) 'A new look at diglossia: Modality-driven distinctions between spoken and written narratives in Jordanian Arabic', in *Handbook of Arabic Literacy: Insights and Perspectives*, pp. 241-254.

Lee, C.S. and Todd, N.P.M. (2004) 'Towards an auditory account of speech rhythm: Application of a model of the auditory “primal sketch” to two multi-language corpora', *Cognition*, 93(3), pp. 225-254.

Lee, S., Potamianos, A. and Narayanan, S. (1997) 'Analysis of children's speech: Duration, pitch and formants', in *Fifth European Conference on Speech Communication and Technology*.

Lee, W.S. and McAngus Todd, N.P. (2004) 'Between-speaker variation in the temporal coordination of consonant clusters in Cantonese', *Journal of the International Phonetic Association*, 34(1), pp. 1-23.

Lehiste, I. (1964) 'Compounding as a phonological process', in *Proceedings of the 9th International Conference of Linguists*, ed. by Horace Lust. The Hague: Mouton.

Lehiste, I. (1977) 'Isochrony reconsidered', *Journal of Phonetics*, 5(3), pp. 253-263.

Leitenstorfer, F. and Tutz, G. (2007) 'Knot selection by boosting techniques', *Computational Statistics & Data Analysis*, 51(9), pp. 4605-4621.

Levelt, C.C., Schiller, N.O. and Levelt, W.J. (2000) 'The acquisition of syllable types', *Language Acquisition*, 8(3), pp. 237-264.

Levelt, C.C. and Van de Vijver, R. (2004) 'Syllable types in cross-linguistic and developmental grammars', in *Constraints in Phonological Acquisition*, pp. 204-218.

Levey, S. and Schwartz, R.G. (2002) 'Syllable omission by two-year-old children', *Communication Disorders Quarterly*, 23(4), pp. 168-176.

Liberman, M. and Prince, A. (1977) 'On stress and linguistic rhythm', *Linguistic Inquiry*, 8(2), pp. 249-336.

Liberman, M.Y. (1975) *The Intonational System of English*. Doctoral dissertation, Massachusetts Institute of Technology.

Lively, S.E. (1993) 'An examination of the perceptual magnet effect', *The Journal of the Acoustical Society of America*, 93(4), pp. 2423.

Lively, S.E. (1993) 'An examination of the perceptual magnet effect', *The Journal of the Acoustical Society of America*, 93(4 Supplement), pp. 2423-2423.

Lleó, C. and Prinz, M. (1996) 'Consonant clusters in child phonology and the directionality of syllable structure assignment', *Journal of Child Language*, 23(1), pp. 31-56.

Macken, M.A. (1993) 'Developmental changes in the acquisition of phonology', in *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, pp. 435-449. Dordrecht: Springer Netherlands.

MacNeilage, P.F. and Davis, B.L. (2001) 'Motor mechanisms in speech ontogeny: Phylogenetic, neurobiological and linguistic implications', *Current Opinion in Neurobiology*, 11(6), pp. 696-700.

MacNeilage, P.F. (1970) 'Motor control of serial ordering of speech', *Psychological Review*, 77(3), pp. 182-196.

Maddieson, I. (1980) 'Phonological generalizations from the UCLA Phonological Segment Inventory Database', *UCLA Working Papers in Phonetics*, 50, pp. 57-68.

Majorano, M., Vihman, M.M. and DePaolis, R.A. (2014) 'The relationship between infants' production experience and their processing of speech', *Language Learning and Development*, 10(2), pp. 179-204.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review, 79*(6), 487–509. https://doi.org/10.1037/h0033467.

Mashaqba, B. and Huneety, A. (2018) 'Emergence of iambs in Eastern Arabic: Metrical iambicity dominating optimal nonfinality', *SKASE Journal of Theoretical Linguistics*, 15(3), pp. 1-21.

Mashaqba, B., Huneety, A., Al-Khawaldeh, N. and Thnaibat, B. (2021) 'Geminate acquisition and representation by Ammani Arabic-speaking children', *International Journal of Arabic-English Studies*, 21(1), pp. 219-242.

Mashaqba, B.M., Al-Shdifat, K.G., Al Huneety, A.I., Alhala, M.A. (2019) 'Acquisition of syllable structure in Jordanian Arabic', *Communication Sciences & Disorders*, 24(4), pp. 953-967.

Massaro, D.W. (1972) 'Preperceptual images, processing time, and perceptual units in auditory perception', *Psychological Review*, 79(2), pp. 124-145.

Mattock, K. and Burnham, D. (2006) 'Chinese and English infants' tone perception: Evidence for perceptual reorganization', *Infancy*, 10(3), pp. 241-265.

Maye, J., Werker, J.F. and Gerken, L. (2002) 'Infant sensitivity to distributional information can affect phonetic discrimination', *Cognition*, 82(3), pp. B101-B111.

McCarthy, J. and Prince, A. (1986) *Prosodic Morphology*. Manuscript. University of Massachusetts, Amherst, and Brandeis University.

McCarthy, J.J. (1979) 'On stress and syllabification', *Linguistic Inquiry*, 10(3), pp. 443-465.

McCarthy, J.J. (2008) 'The Serial Interaction of Stress and Syncope', *Natural Language and Linguistic Theory*, 26(3), pp. 499-546. doi: 10.1007/s11049-008-9051-3.

McCarthy, J.J. and Prince, A. (1994) 'Prosodic morphology', *Linguistics Department Faculty Publication Series*, pp. 1-15.

McCarthy, J.J. and Prince, A.S. (1990) 'Foot and word in prosodic morphology: The Arabic broken plural', *Natural Language & Linguistic Theory*, 8(2), pp. 209-283.

McCarthy, J.J. (1981) 'A prosodic theory of nonconcatenative morphology', *Linguistic Inquiry*, 12(3), pp. 373-418.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J. and Amiel-Tison, C. (1988) 'A precursor of language acquisition in young infants', *Cognition*, 29(2), pp. 143-178.

Menggo, S. (2017) 'Phonology Acquisition at the Holophrastic Stage', *Jurnal Pendidikan dan Kebudayaan Missio*, 9(1), pp. 70-74.

Menyuk, P. (1971) *The Acquisition and Development of Language*. Englewood Cliffs, NJ: Prentice Hall.

Miller, M. (1984) 'On the perception of rhythm', *Journal of Phonetics*, 12(1), pp. 75-83.

Mitchell, J.C. (1990) 'Type systems for programming languages', in *Formal Models and Semantics*, pp. 365-458. Elsevier.

Mitchell, T.F. (1969) 'David Abercrombie, Elements of General Phonetics. Edinburgh: Edinburgh University Press, 1966. Pp. 203', *Journal of Linguistics*, 5(1), pp. 153-164.

Morgan, J.L. and Demuth, K. (1996) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Morgan, J.L. (1996) 'Prosody and the roots of parsing', *Language and Cognitive Processes*, 11(1-2), pp. 69-106.

Mowrer, D.E. and Burger, S. (1991) 'A comparative analysis of phonological acquisition of consonants in the speech of 2½-6-year-old Xhosa- and English-speaking children', *Clinical Linguistics & Phonetics*, 5(2), pp. 139-164.

Mustafawi, E. (2017) 'Arabic phonology', in *The Routledge Handbook of Arabic Linguistics*, pp. 11-31.

Na'eem, H., Abudalbuh, M. and Jaber, A. (2020) 'Medial tri-consonantal clusters in Urban Jordanian Arabic', *Jordan Journal of Modern Languages and Literatures*, 12(1), pp. 45-56.

Nelson, D.G.K., Hirsh-Pasek, K., Jusczyk, P.W. and Cassidy, K.W. (1989) 'How the prosodic cues in motherese might assist language learning', *Journal of Child Language*, 16(1), pp. 55-68.

Nelson, D.G.K., Jusczyk, P.W., Mandel, D.R., Myers, J., Turk, A. and Gerken, L. (1995) 'The head-turn preference procedure for testing auditory perception', *Infant Behavior and Development*, 18(1), pp. 111-116.

Nespor, M. and Vogel, I. (1986) *Prosodic Phonology*. Dordrecht: Foris.

Nespor, M., Shukla, M. and Mehler, J. (2011) 'Stress-timed vs. Syllable-timed Languages'.

Newman, D. (2002) 'The phonetic status of Arabic within the world's languages: The uniqueness of the lughat al-daad', *Antwerp Papers in Linguistics*, 100, pp. 65-75.

Newman, P. (1972) 'Syllable weight as a phonological variable', *Studies in African Linguistics*, 3(3), pp. 301-323.

Nip, I.S. and Green, J.R. (2013) 'Increases in cognitive and linguistic processing primarily account for increases in speaking rate with age', *Child Development*, 84(4), pp. 1324-1337.

Nip, I.S., Green, J.R. and Marx, D.B. (2009) 'Early speech motor development: Cognitive and linguistic considerations', *Journal of Communication Disorders*, 42(4), pp. 286-298.

Nip, I.S., Green, J.R. and Marx, D.B. (2011) 'The co-emergence of cognition, language, and speech motor control in early development: A longitudinal correlation study', *Journal of Communication Disorders*, 44(2), pp. 149-160.

Nolan, F. and Jeon, H.S. (2014) 'Speech rhythm: A metaphor?', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1658), p. 20130396.

O'Connor, J.D. and Trim, J.L.M. (1953) 'Vowel, consonant, and syllable—a phonological definition', *Word*, 9(2), pp. 103-122.

Oller, D.K. (1980) 'The emergence of the sounds of speech in infancy', in Yeni-Komshian, G., Kavanagh, J.F. and Ferguson, C.A. (Eds.) *Child Phonology* (Vol. 1, pp. 93-112). Academic Press.

Oller, D.K. (1971) 'Position-in-Utterance Duration Effects', *The Journal of the Acoustical Society of America*, 49(1A), pp. 105.

Oller, D.K. and Smith, B.L. (1977) 'Effect of final-syllable position on vowel duration in infant babbling', *The Journal of the Acoustical Society of America*, 62(4), pp. 994-997.

Oller, D.K. (1973) 'The effect of position in utterance on speech segment duration in English', *The Journal of the Acoustical Society of America*, 54(5), pp. 1235-1247.

Ota, M. (1999) *Phonological Theory and the Acquisition of Prosodic Structure: Evidence from Child Japanese*. Georgetown University.

Ota, M. (2001) 'Phonological theory and the development of prosodic structure: Evidence from child Japanese', *Annual Review of Language Acquisition*, 1(1), pp. 65-118. doi: 10.1075/arla.1.03ota.

Ota, M., Lidz, J., Snyder, W. and Pater, J. (2016) 'Prosodic phenomena: Stress, tone, and intonation', in *The Oxford Handbook of Developmental Linguistics*, pp. 68-86.

Owens, J. (2013) *The Oxford Handbook of Arabic Linguistics*. Oxford University Press.

Palva, H. (1984) 'A general classification for the Arabic dialects spoken in Palestine and Transjordan', *Studia Orientalia Electronica*, 55, pp. 357-376.

Payne, E., Post, B., Astruc, L., Prieto, P. and Vanrell, M.D.M. (2012) 'Measuring child rhythm', *Language and Speech*, 55(2), pp. 203-229.

Pham-Gia, T. and Hung, T.L. (2001) 'The mean and median absolute deviations', *Mathematical and Computer Modelling*, 34(7-8), pp. 921-936.

Pierrehumbert, J.B. (2003) 'Phonetic diversity, statistical learning, and acquisition of phonology', *Language and Speech*, 46(2-3), pp. 115-154.

Pierrehumbert, J.B. (1980) *The Phonology and Phonetics of English Intonation*. Doctoral dissertation, Massachusetts Institute of Technology.

Piironen, J. and Vehtari, A. (2017) 'Comparison of Bayesian predictive methods for model selection', *Statistics and Computing*, 27, pp. 711-735.

Pike, K.L. (1945) *The Intonation of American English*. Ann Arbor: University of Michigan Press.

Polka, L. and Bohn, O.-S. (2003) 'Asymmetries in vowel perception', *Speech Communication,* 41(1), pp. 221-231.

Polka, L. and Bohn, O.S. (2011) 'Natural referent vowel (NRV) framework: An emerging view of early phonetic development', *Journal of Phonetics,* 39(4), pp. 467-478.

Pollock, K.E., Brammer, D.M. and Hageman, C.F. (1993) 'An acoustic analysis of young children’s productions of word stress', *Journal of Phonetics*, 21(3), pp. 183-203.

Pons, F. and Bosch, L. (2010) 'Stress pattern preference in Spanish-learning infants: The role of syllable weight', *Infancy*, 15(3), pp. 223-245.

Port, R.F., Dalby, J. and O’Dell, M. (1987) 'Evidence for mora timing in Japanese', *The Journal of the Acoustical Society of America*, 81(5), pp. 1574-1585.

Price, P.J., Ostendorf, M., Shattuck-Hufnagel, S. and Fong, C. (1991) 'The use of prosody in syntactic disambiguation', *Journal of the Acoustical Society of America*, 90(6), pp. 2956-2970.

Prieto, P. (2006) 'The acquisition of prosody in Catalan and Spanish: The interaction of universal and language-specific constraints', in Buder, E.H. and Renwick, E. (Eds.) *Proceedings of the 8th International Conference on Laboratory Phonology*, pp. 115-134.

Prince, A. and Smolensky, P. (2003) 'Optimality Theory in phonology', *International Encyclopedia of Linguistics*, 3, pp. 212-22.

Pye, C., Ingram, D. and List, J. (1987) 'The acquisition of word-final consonant clusters in Spanish-speaking children', *Journal of Child Language*, 14(2), pp. 367-387.

R Core Team (2013) *R: A language and environment for statistical computing*. Available at: http://www.R-project.org (Accessed: April 2022).

Rakhieh, B.A. (2009) *The Phonology of Ma'ani Arabic: Stratal or Parallel OT*.

Ramus, F., Nespor, M. and Mehler, J. (1999) 'Correlates of linguistic rhythm in the speech signal', *Cognition*, 73(3), pp. 265-292.

Repetti, L. (1994) 'Degenerate syllables in Friulian', *Linguistic Inquiry*, 25(1), pp. 186-193.

Ridouane, R. (2007) 'Gemination in Tashlhiyt Berber: An acoustic and articulatory study', *Journal of the International Phonetic Association*, 37(2), pp. 119-142.

Roach, P. (1982) 'On the distinction between ‘stress-timed’ and ‘syllable-timed’ languages', in Crystal, D. (ed.) *Linguistic Controversies*. London: Edward Arnold, pp. 73-79.

Robb, M.P. and Saxman, J.H. (1990) 'Syllable durations of preword and early word vocalizations', *Journal of Speech, Language, and Hearing Research*, 33(3), pp. 583-593.

Rojczyk, A. and Porzuczek, A. (2012) 'Selected aspects in the acquisition of English phonology by Polish learners: Segments and prosody'.

Rouas, J.L., Farinas, J., Pellegrino, F. and André-Obrecht, R. (2005) 'Rhythmic unit extraction and modelling for automatic language identification', *Speech Communication*, 47(4), pp. 436-456.

Roy, P. & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2- to 4- year-olds: Evidence from typically developing children. Journal of Speech, Language, and Hearing Research, 47(1), pp. 223-234. doi: 10.1044/1092-4388(2004/019).

Ryding, K.C. (2005) *A Reference Grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.

Saffran, J.R., Aslin, R.N. and Newport, E.L. (1996) 'Statistical learning by 8-month-old infants', *Science*, 274(5294), pp. 1926-1928.

Saiegh-Haddad, E. and Henkin-Roitfarb, R. (2014) 'The structure of Arabic language and orthography', in K. H. E. and M. A. (Eds.) *Handbook of Arabic Literacy*. Dordrecht: Springer Netherlands, pp. 3-28.

Sakarna, A.K. (2005) 'The linguistic status of the modern Jordanian dialects', *Arabica*, pp. 522-543.

Saleh, M., Shoeib, R., Hegazi, M. and Ali, P. (2007) 'Early phonological development in Arabic Egyptian children: 12–30 months', *Folia Phoniatrica et Logopaedica*, 59(5), pp. 234-240. doi: 10.1159/000104461.

Salim, J.A. and Mehawesh, M. (2014) 'Stages in language acquisition: A case study', *English Language and Literature Studies*, 4(4), p. 16.

Sander, E.K. (1972) 'When are speech sounds learned?', *Journal of Speech and Hearing Disorders*, 37(1), pp. 55-63.

Savin, H.B. and Bever, T.G. (1970) 'The non-perceptual reality of the phoneme', *Journal of Verbal Learning and Verbal Behavior*, 9(3), pp. 295-302.

Sawaie, M. (2007) 'Modern Arabic: Structures, functions, and varieties', *Language*, 83(2), pp. 461-461.

Schwartz, R.G., Petinou, K., Goffman, L., Lazowski, G. and Cartusciello, C. (1996) 'Young children’s production of syllable stress: An acoustic analysis', *The Journal of the Acoustical Society of America*, 99(5), pp. 3192-3200.

Selkirk, E. (1981) 'Epenthesis and degenerate syllables in Cairene Arabic', in *Theoretical Issues in the Grammar of Semitic Languages*, M.I.T. Working Papers in Linguistics. Department of Linguistics and Philosophy. Massachusetts Institute of Technology, Cambridge, Mass., 3, pp. 209-232.

Selkirk, E.O. (1982) *The Syntax of Words*. Cambridge, MA: MIT Press.

Sharkey, S.G. and Folkins, J.W. (1985) 'Variability of lip and jaw movements in children and adults: Implications for the development of speech motor control', *Journal of Speech, Language, and Hearing Research*, 28(1), pp. 8-15.

Shriberg, L.D., Tomblin, J.B. and McSweeny, J.L. (1999) 'Prevalence of speech delay in 6-year-old children and comorbidity with language impairment', *Journal of Speech, Language, and Hearing Research*, 42(6), pp. 1461-1481.

Skoruppa, K., Pons, F., Christophe, A., Bosch, L., Dupoux, E., Sebastián-Gallés, N., Limissuri, R.A. and Peperkamp, S. (2009) 'Language-specific stress perception by 9-month-old French and Spanish infants', *Developmental Science*, 12(6), pp. 914-919.

Sluijter, A.M. and Van Heuven, V.J. (1996) 'Spectral balance as an acoustic correlate of linguistic stress', *The Journal of the Acoustical Society of America*, 100(4), pp. 2471-2485.

Sluijter, A.M. and van Heuven, V.J. (1995) 'Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch', *Phonetica*, 52(2), pp. 71-89.

Sluijter, A.M.C. (1995) *Phonetic Correlates of Stress and Accent*. PhD Thesis, University of Leiden.

Smith, A. and Zelaznik, H.N. (2004) 'Development of functional synergies for speech motor coordination in childhood and adolescence', *Developmental Psychobiology*, 45(1), pp. 22-33.

Smith, B.L. (1978) 'Temporal aspects of English speech production: A developmental perspective', *Journal of Phonetics*, 6(1), pp. 37-67.

Smith, B.L. and Kenney, M.K. (1994) 'Variability control in speech production tasks performed by adults and children', *The Journal of the Acoustical Society of America*, 96(2), pp. 699-705.

Smith, B.L. and Thelen, E. (2003) 'Development as a dynamic system', *Trends in Cognitive Sciences*, 7(8), pp. 343-348.

Smith, B.L., Kenney, M.K. and Hussain, S. (1996) 'A longitudinal investigation of duration and temporal variability in children's speech production', *The Journal of the Acoustical Society of America*, 99(4), pp. 2344-2349.

Smith, B.L., Sugarman, M.D. and Long, S.H. (1983) 'Experimental manipulation of speaking rate for studying temporal variability in children's speech', *The Journal of the Acoustical Society of America*, 74(3), pp. 744-749.

Snow, D. (1994) 'Phrase-final syllable lengthening and intonation in early child speech', *Journal of Speech, Language, and Hearing Research*, 37(4), pp. 831-840.

Snow, D. (1998) 'A Prominence Account of Syllable Reduction in Early Speech Development: The Child's Prosodic Phonology of Tiger and Giraffe', *Journal of Speech, Language, and Hearing Research*, 41(5), pp. 1171-1184. doi: 10.1044/jslhr.4105.1171.

Snow, D. (1998b) 'Prosodic markers of syntactic boundaries in the speech of 4-year-old children with normal and disordered language development', *Journal of Speech, Language, and Hearing Research*, 41(5), pp. 1158-1170.

Son, R.J.v. and Santen, J.P.v. (1997) 'Strong interaction between factors influencing consonant duration'. *5th European Conference on Speech Communication and Technology*, Greece.

Stan Development Team (2016) *Stan Reference Manual, Version 2.14.0*. Available at: https://mc-stan.org (Accessed: 23 August 2024).

Stan Development Team (2016) *RStan: The R interface to Stan*. R package version 2.26.24. Available at: https://mc-stan.org/ (Accessed: 23 August 2024).

Steriade, D. (2001) 'Directional asymmetries in place assimilation: A perceptual account', in *The Role of Speech Perception in Phonology*, pp. 219-250.

Stoel-Gammon, C. (2011) 'Relationships between lexical and phonological development in young children', *Journal of Child Language*, 38(1), pp. 1-34.

Stoel-Gammon, C. and Williams, A.L. (2013) 'Early phonological development: Creating an assessment test', *Clinical Linguistics & Phonetics*, 27(4), pp. 278-286.

Studdert-Kennedy, M. (1976) 'Speech perception', in Lass, N.J. (ed.) *Contemporary Issues in Experimental Phonetics*. Academic Press.

Sudhoff, S., Lenertova, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., Richter, N. and Schließer, J. (2006) *Methods in Empirical Prosody Research*. Berlin, Boston: De Gruyter. doi: 10.1515/9783110914641.

Suleiman, M.F. (1993) *A Study of Arab Students' Motivations and Attitudes for Learning English as a Foreign Language*. Arizona State University.

Tabain, M. (2003) 'Effects of prosodic boundary on /aC/ sequences: Articulatory results', *Journal of the Acoustical Society of America*, 113(6), pp. 2834-2849.

Thelen, E. (1991) 'Motor aspects of emergent speech: A dynamic approach', in *Biological and Behavioral Determinants of Language Development*, pp. 339-362.

Thnaibat, B. (2019) *Acoustics of Geminates in Jordanian Children with Cleft Palate*. Unpublished MA Thesis, The Hashemite University, Zarqa, Jordan.

Tingley, B.M. and Allen, G.D. (1975) 'Development of speech timing control in children', *Child Development*, pp. 186-194.

Tobin, Y. (1997) 'Phonology as human behavior: Theoretical implications and clinical applications', in *Sound and Meaning: The Roman J.*

Trask, R.L. (1996) *A Dictionary of Phonetics and Phonology*. Routledge.

Trubetzkoy, N.I. (1969) *Principles of Phonology*. Berkeley: University of California Press.

Tsao, F., Liu, H. and Kuhl, P.K. (2004) 'Speech Perception in Infancy Predicts Language Development in the Second Year of Life: A Longitudinal Study', *Child Development*, 75(4), pp. 1067-1084. doi: 10.1111/j.1467-8624.2004.00726.x.

Turk, A. and Shattuck-Hufnagel, S. (2000) 'Word-boundary-related duration patterns in English', *Journal of Phonetics,* 28(4), pp. 397-440.

Turk, A., Nakai, S. and Sugahara, M. (2006) 'Acoustic segment durations in prosodic research: A practical guide', in *Methods in Empirical Prosody Research*, pp. 1-28.

Turk, A.E. and Sawusch, J.R. (1997) 'The domain of accentual lengthening in American English', *Journal of Phonetics*, 25(1), pp. 25-41.

Turk, A.E. and White, L. (1999) 'Structural influences on accentual lengthening in English', *Journal of Phonetics*, 27(2), pp. 171-206.

van den Berg, R. (2012) *Syllables Inside Out: A Longitudinal Study of the Development of Syllable Types in Toddlers Acquiring Dutch: A Comparison between Hearing Impaired Children with a Cochlear Implant and Normally Hearing Children*. University of Antwerp.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B. and Bürkner, P. (2021) 'Rank-normalization, folding, and localization: An improved R for assessing convergence of MCMC (with discussion)', *Bayesian Analysis*, 16(2), pp. 667-718.

Veneziano, E. (1981) 'Early language and nonverbal representation: A reassessment', *Journal of Child Language*, 8(3), pp. 541-563.

Vihman, M. (2015) 'Perception and Production in Phonological Development', in *The Handbook of Language Emergence*, Hoboken, NJ: John Wiley & Sons, Inc, pp. 437-457.

Vihman, M.M. and Velleman, S.L. (2000) 'Phonetics and the origins of phonology', in Burton-Roberts, N., Carr, P. and Docherty, G. (eds.) *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford University Press, pp. 305-339.

Vihman, M.M. (2018) 'The development of prosodic structure', in *The Development of Prosody in First Language Acquisition*, 23, pp. 185.

Vihman, M.M. and McCune, L. (1994) 'When is a word a word?', *Journal of Child Language*, 21(3), pp. 517-542.

Vihman, M.M. (1991) 'Ontogeny of phonetic gestures: Speech production', in *Modularity and the Motor Theory of Speech Perception: Proceedings of a Conference to Honor Alvin M. Liberman*, Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 69-84.

Vogel, I., Athanasopoulou, A., Pincus, N. and Ouali, H. (2017) 'Acoustic properties of prominence and foot structure in Arabic'. In *Perspectives on Arabic Linguistics XXIX* (pp. 55-88). John Benjamins.

Warren, R.M. (1971) 'Perceptual restoration of missing speech sounds', *Science*, 176(4034), pp. 392-393.

Watkins, M.A. (2001) 'Variability in vowel reduction by Brazilian speakers of English'.

Watson, J. (2011) *Word Stress in Arabic*. Oxford: John Wiley & Sons, Ltd.

Watson, J.C. (2002) *The Phonology and Morphology of Arabic*. Oxford University Press.

Watson, J.C. (2007) 'Syllabification patterns in Arabic dialects: Long segments and mora sharing', *Phonology*, 24(2), pp. 335-356.

Werker, J.F. and Tees, R.C. (1984) 'Cross-language speech perception: Evidence for perceptual reorganization during the first year of life', *Infant Behavior and Development*, 7(1), pp. 49-63.

Werker, J.F., Gilbert, J.H., Humphrey, K. and Tees, R.C. (1981) 'Developmental aspects of cross-language speech perception', *Child Development*, pp. 349-355.

White, L. and Turk, A. (2010) 'English words on the Procrustean bed: Polysyllabic shortening reconsidered', *Journal of Phonetics*, 38(3), pp. 459-471.

White, L. (2014) ‘Communicative function and prosodic form in speech timing’, *Speech Communication*, 63, pp.38-54.

Wightman, C.W., Shattuck-Hufnagel, S., Ostendorf, M. and Price, P.J. (1992) 'Segmental durations in the vicinity of prosodic phrase boundaries', *The Journal of the Acoustical Society of America*, 91(3), pp. 1707-1717.

Winter, B. (2013) 'Linear models and linear mixed effects models in R with linguistic applications', *arXiv preprint* arXiv:1308.5499.

Yeou, M. (2005) 'Variability of F0 peak alignment in Moroccan Arabic accentual focus'.

Zamuner, T.S. (2003) *Acquisition of Syllable Structure in a Second Language: A Case Study of Spanish-Speaking Children Learning English*. Doctoral dissertation, University of Toronto.

Zawaydeh, B.A. and de Jong, K. (1999) 'Stress, phonological focus, quantity, and voicing effects on vowel duration in Ammani Arabic', *Regents of the University of California Berkeley,* CA, pp. 451.

Zawaydeh, B.A., Tajima, K. and Kitahara, M. (2002) 'Discovering Arabic rhythm through a speech cycling task', *Amsterdam Studies in the Theory and History of Linguistic Science Series 4,* pp. 39-58.

Zec, D. (2007) 'The syllable', in *The Cambridge Handbook of Phonology,* pp. 161-194.

Zharkova, N. (2004) 'Strategies in the acquisition of segments and syllables in Russian-speaking children', in *Developmental Paths in Phonological Acquisition*. Special issue of *Leiden Papers in Linguistics*.

Zughoul, M.R. (1980) 'Diglossia in Arabic: Investigating solutions', *Anthropological Linguistics*, 22(5), pp. 201-217.

Zuraiq, W. (2005) *The Production of Lexical Stress by Native Speakers of Arabic and English and by Arab Learners of English*. University of Kansas.

Appendix: Repetition Task Word List

Word	Gloss	Syllable structure
Ma.laak	Angel	CV.CVVC
Ktaab	Book	CCVVC
Slaaħ	Weapon	CCVVC
Mak.tuub	Written	CVC.CVVC
Ma.ħall	Shop	CV.CVCC
dʒa.waab	Answer	CV.CVVC
Sa.kaa.kiin	Knives	CV.CVV.CVVC
ʕamm	Uncle	CVCC
Sadd	Dam	CVCC
Baab	Door	CVVC
Bint	Girl	CVCC
sˤaa.buun	Soap	CVV.CVVC
Dars	Lesson	CVCC
Dubb	Bear	CVCC
Kbeer	Big	CCVVC
Sɣeer	Small	CCVVC
Nkabb	Spilled	CCVCC
Muftaaħ	Key	CVC.CVVC
sˤħuun	Dishes	CCVVC
ʃriit	Cassette	CCVVC
ħraam	Blanket	CCVVC
Zlaam	Men	CCVVC
Traab	Soil	CCVVC
Ftuur	Breakfast	CCVVC
Kfuuf	Gloves	CCVVC
Klaab	Dogs	CCVVC
Rfuuf	Shelfs	CCVVC
Dmuuʕ	Tears	CCVVC
ħmaar	Donkey	CCVVC
Snaan	Teeth	CCVVC

Fuut	Enter	CVVC
ħsˤaan	Horse	CCVVC
Ktiir	Many	CCVVC
Mniiħ	Good	CCVVC
Nuur	Light	CVVC
ʕa.riis	Spouse	CVVC
ʕa.ruus	Wife	CVVC
Druus	Lessons	CCVVC
Wlaad	Boys	CCVVC
Ba.naat	Girls	CVVC
Talj	Snow	CVCC
Ramz	Symbol	CVCC
ʃams	Sun	CVCC
Kanz	Treasure	CVCC
Ward	Roses	CVCC
ħarf	Letter	CVCC