A Methodology For Analysing Improvised Jazz: A Computer-Aided Approach

David Blackwell

Submitted in fulfilment of the requirements

for the degree of

Doctor of Philosophy

Elder Conservatorium of Music

Faculty of Arts, Business, Law, and Economics

The University of Adelaide

February 2023

Contents

Abstract i

Declaration ii

Acknowledgements iii

List of Tables vii

List of Figures xvii

Definition of Terms xviii

1 Introduction 1

Part I
In Search of a New Methodology 12

2 Literature Review 12

2.1 Analysis of Improvised Jazz 12

2.2 Computer-aided Musical Analysis 15

2.3 Computer-aided Jazz Analysis 17

2.4 Machine Learning in Music 19

2.5 Grant Green 21

2.6 Summary 23

3 Approaching A New Methodology 25

3.1 Data Selection and Preparation 26

3.1.1 Data Selection 26

3.1.2 Transcription 33

3.1.3 Data Extraction 44

3.1.4 Treatment of Data 45

3.2 Analysis of Grant Green 50

3.3 Performer Classification and Comparative Analysis 55

Part II
Analysis of Grant Green's Improvisational Style 60

4 Feature Information from Grant Green's Transcriptions 60

5 Pitch Domain 68

5.1 Raw Pitch 68

5.2 Tonal Pitch Class 82

5.3 Chordal Pitch Class 87

5.4 Intervals 116

5.5 Examples 130

5.5.1 Surrounding Note Figures 130

5.5.2 Voice Leading 140

5.6 Green's Improvisational Style In The Pitch Domain 145

6 Rhythm Domain 148

6.1 Note lengths 148

6.2 Beat Distribution 158

6.3 Metrical Weight 163

6.4 Rests 166

6.5 Examples 178

6.5.1 Rhythmic Variety 178

6.5.2 Metrical Density 187

6.6 Green's Improvisational Style In The Rhythm Domain 196

7 Micro Domain 199

7.1 Swing 199

7.2 Micro Timings 210

7.3 Note Placement 220

7.4 Green's Improvisational Style In The Micro Domain 228

8 Macro Domain **230**

8.1 Phrases . . . 230

8.2 Large Scale Feature Trends . . . 250

8.3 Green's Improvisational Style In The Macro Domain . . . 260

9 Findings of Grant Green Analysis **261**

Part III
Performer Classification and Comparative Analysis **266**

10 Feature Information of Additional Performer for Classification and Comparative Analysis **266**

11 Feature Selection **269**

11.1 Consistent Features . . . 274

11.2 Solo Level . . . 277

11.3 Phrase Level . . . 277

11.4 Bar Level . . . 278

11.5 Note Level . . . 279

11.6 Feature Selection Summary . . . 280

12 Model Results **281**

12.1 Model Performance Metrics . . . 283

12.1.1 C4.5-like Results . . . 289

12.1.2 C5.0 Results . . . 290

12.1.3 Random Forest Results . . . 292

12.1.4 Model Performance Metrics Summary . . . 293

12.2 Feature Analysis . . . 294

12.2.1 n-way . . . 297

12.2.2 One-vs-All . . . 301

12.2.3 One-vs-One . . . 307

12.2.4 Feature Analysis Summary . . . 316

12.3 Model Results Summary . . . 316

13 Comparative Analysis **318**

13.1 Pitch Domain 319

13.2 Rhythm Domain 326

13.3 Macro Domain 330

13.4 Comparative Analysis Summary 334

14 Findings of Performer Classification and Comparative Analysis **336**

15 Conclusion **339**

List of Sources **344**

Appendices **368**

Appendix A Additional Graphs **368**

Appendix B Code Examples **372**

Appendix C Transcription Details **378**

Appendix D Summary Performance Metrics **384**

Appendix E Data **400**

E.1 Transcription Files 400

E.2 Data Extraction and Setup Files 400

E.3 Machine Learning Setup Files 401

E.4 Machine Learning Models 401

E.5 Machine Learning Confusion Matrices 401

E.6 RDS Data Files 401

E.7 Thesis Files 402

E.8 Supplementary Files 402

Abstract

This musicological research employs a computer-aided statistical approach to the analysis of improvised jazz. The main aim of this study is to develop a new methodology for systematically analysing improvisational style, with this research completed in two parts. This is done first through an analysis of Grant Green's improvisational style, based on transcriptions of forty improvisations between 1960 and 1965. Green (1935–1979) was a prolific but underrated jazz guitarist, and was the unofficial house guitarist for Blue Note Records between 1960 and 1965. This research aims to explore, analyse, and explain Green's improvisational style with reference to his use of pitch, rhythm, micro, and macro features. Secondly, the results of this analysis are used to inform performer classifications and comparative analysis between Grant Green, John Coltrane, Miles Davis, and Charlie Parker. Tree based machine learning algorithms are utilised to complete the performer classification tasks, with the comparative analysis based upon the features found to classify the performers. This research built upon previous work from the Jazzomat Research Project (2012–2017), based out of the University of Music Franz Liszt Weimar. This research uses methods and software developed by the Jazzomat Research Project to transcribe and extract the data from Green's solos, with the data for the other three performers in the comparative analysis coming from their Weimar Jazz Database. The analyses, and training and evaluation of the machine learning classifiers, were undertaken in the *R* programming language. The results of this study found that Green conformed to many of the improvisational conventions of the time, with these results confirming the validity of the developed methodology. Findings from the classification task found that the C5.0 classifier was the most efficient and performant when classifying the improvisers. The results of this research contribute to the field of computational musicology and the analysis of improvised jazz. The methodology developed through this research will allow future investigations to thoroughly explore the improvisational style of other musicians.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Signature: David Blackwell

Date: 2022/02/06

Acknowledgements

I would like to acknowledge the support and guidance offered by my supervisors, Dr. Luke Harrald and Professor Aaron Corn. I would also like to thank the excellent current and previous staff at the Elder Conservatorium of Music for their support throughout my music education, especially Greg Wain, Steven Knopoff, Associate Professor Kimi Coaldrake, and Professor Charles Bodman Rae. Thank you to my fellow post graduate candidates for their friendship throughout the years, with special thanks to Dr. Iran Sanadzadeh for her guidance and encouragement. I would like to thank Dr. Klaus Frieler of the Jazzomat Research Project for providing example files and instructions for generating the database used throughout my research. Finally, I am indebted to Lee, Rob, and Rebecca Blackwell and Madi Grant for their unwavering love, support, and assistance, throughout the highs and lows of this journey.

List of Tables

2.1 Number of transcriptions in each of the surveyed studies. 13

3.1 Mean and standard deviation for the year and tempo for all 147 vs. the selected songs. 29

3.2 Title, Album, Tempo, and Solo times for all improvisations that were transcribed to form the corpus used within the study. Tracks that contained two solos have the second solo listed under the first, with separate tempo and timestamps listed. 31

3.3 Summary of chord type simplifications, showing the number of bars, with at least one note event, associated with each chord type. 42

4.1 Duration in seconds of Green's improvisations. 60

4.2 Green's general corpus details. 61

4.3 Distribution of time signatures in Green's corpus. 61

4.4 Distribution of key signatures in Green's corpus. 62

4.5 Distribution of tonality modes in Green's corpus. 63

4.6 Distribution of recording year in Green's corpus. 64

4.7 Distribution of chord types and tonics in Green's corpus. 65

4.8 Distribution of tempo classes in Green's corpus. 65

5.1 Octave distribution for Green vs. guitarists in the WJazzD. 73

5.2 Frequency of octave sequence lengths (note count) in Green's corpus. . . 76

5.3 Distribution of notes per octave, following a note played in a specific octave. 79

5.4 TPC weight distribution for each tonality mode in Green's corpus. . . . 84

5.5 Distribution of TPC Weight transitions around a NDT. D = Diatonic Tone, N = Non-Diatonic Tone. 86

5.6 Translation of CPC to CDPCX for $\Delta 7$, 7, m7, ø7, and ∘7 chords in reference to a C root. 88

5.7 Percentage of $\mathrm{CPC}_{\mathrm{Weight}}$ that had a chromatic approach, departure, approach and departure, and approach and/or departure. 108
5.8 Distribution of Δ7 USTs. 115
5.9 Distribution of m7 USTs. 115
5.10 Distribution of 7 USTs. 115
5.11 Interval Descriptors. 117
5.12 Distribution of Parsons code in Green's corpus. 117
5.13 Statistical descriptors for descending and ascending intervals in Green's corpus. 118
5.14 Proportion of major and minor 3rds for each chord type where both notes were part of the extended arpeggio. 124
5.15 Distribution of SNF interval patterns in Green's corpus. 131
5.16 Distribution of SNF interval sequences in Green's corpus. 131
5.17 Distribution of SNF interval sequences over Δ7, 7, and m7 chords in Green's corpus. 133
5.18 Distribution of $\mathrm{CPC}_{\mathrm{Weight}}$ for each SNF interval sequence target note and chord type in Green's corpus. 135
5.19 Distribution of SNF arpeggio target notes for Δ7, 7, and m7 chords in Green's corpus. 136
5.20 Distribution of metrical weights for notes which were or were note the target note of a SNF. 137
5.21 Distribution of Parsons codes for notes that were or were not voice led. . 141
5.22 Distribution of interval size classes for notes that were or were not voice led. 142

6.1 Duration and IOI classes and boundaries. Based on table found in the Jazzomat Research Project website's 'Transformation' page (Jazzomat Research Project 2017). 149
6.2 Proportion of notes with a duration or IOI class of very short, short, or medium in Green's corpus vs. the WJazzD. 150
6.3 Distribution of relative and fuzzy duration and IOI classes in Green's corpus. 151
6.4 Distribution of metrical weights for notes played at tempos ≤ 170 BPM and > 170 BPM in Green's corpus. 166

6.5 Count, mean, standard deviation, and range for sequences of notes that were or were not followed by a rest for Green, Coltrane, Davis, Parker, Metheny, Martino, and Abercrombie. 169

6.6 Results of statistical tests analysing the frequency of rests under different situations in Green's improvisations. The columns show the statistic score, the degrees of freedom, the p-value, and effect sizes. 171

6.7 Results of statistical tests analysing the duration of rests (as $\text{rest}_{\text{prop}}$) played by Green in various situations. The columns show the statistic scores, degrees of freedom, p-values, and effect sizes. 176

6.8 Metrical divisions: fully and partially occupied beats. 180

7.1 Distribution of Green's BUR for beats with two notes into Snap, Even, and Swing. 202

7.2 Distribution of note placement categorical features in Green's corpus. . . 222

8.1 Phrase shape (contour) descriptors for both reduced and full codes. . . . 232

8.2 Abesser code frequency of phrases played by Green. 234

8.3 Statistical descriptors for both measures of Green's phrase lengths. . . . 236

8.4 Abesser code vs. phrase length TukeyHSD p-values. 239

8.5 Distribution of Green's notes in each phrase position depending on if the following beat had a chord change. 249

8.6 Frequency of feature trend mode shapes in Green's corpus. 253

10.1 Summary Information for Coltrane, Davis, and Parker in the WJazzD. . 267

11.1 Distribution of features selected as input variables for the ML classifiers by domain. 270

11.2 Frequency of selected features for each abstraction level. 272

12.1 Equations for Summary Statistics. 285

12.2 Summary statistics for models which significantly outperformed the NIR. 286

12.3 MCC results for C4.5-like models. 289

12.4 MCC results for C5.0 models. 291

12.5 MCC results for random forest models. 292

12.6 Select OHE simplified feature classes found to be important in n-way comparisons. 299
12.7 Select OHE simplified feature classes found to be important in one-vs-all comparisons. GG: Grant Green. JC: John Coltrane. CP: Charlie Parker. MD: Miles Davis. 303
12.8 Select OHE simplified feature classes found to be important in one-vs-one comparisons. GG: Grant Green. JC: John Coltrane. CP: Charlie Parker. MD: Miles Davis. 309

D.1 C4.5-like solo level summary performance metrics. 385
D.2 C4.5-like phrase level summary performance metrics. 386
D.3 C4.5-like bar (4, 2) level summary performance metrics. 387
D.4 C4.5-like bar (2, 1) level summary performance metrics. 388
D.5 C4.5-like note level summary performance metrics. 389
D.6 C5.0 solo level summary performance metrics. 390
D.7 C5.0 phrase level summary performance metrics. 391
D.8 C5.0 bar (4, 2) level summary performance metrics. 392
D.9 C5.0 bar (2, 1) level summary performance metrics. 393
D.10 C5.0 note level summary performance metrics. 394
D.11 Random forest solo level summary performance metrics. 395
D.12 Random forest phrase level summary performance metrics. 396
D.13 Random forest bar (4, 2) level summary performance metrics. 397
D.14 Random forest bar (2, 1) level summary performance metrics. 398
D.15 Random forest note level summary performance metrics. 399

List of Figures

3.1 Screenshot of the program written to randomly select the improvisations. 27
3.2 Screenshot from a completed transcription in *Sonic Visualiser*. 34
3.3 Beat transcription in *Sonic Visualiser*. Left: Plain spectrogram. Right: Beat transcription over spectrogram. 39
3.4 Example of four possible chord progressions over a blues. 41
3.5 Distribution of Green's notes within the metrical context for improvisations in $\frac{4}{4}$. Left: raw note placement data. Right: quantised metrical circle map data (n=48). 54
3.6 Depiction of the connections between the abstraction levels, machine learning classifiers, and comparisons in the performer classification tasks. 56

4.1 Distribution of major, minor, and blues keys in the datasets of Green, WJazzD and *iReal Pro*. 63
4.2 Density distribution of raw tempos for each beat in Green's corpus. . . . 67

5.1 Pitch class circle map of Green's corpus. 69
5.2 Pitch class circle maps for each tonality mode in Green's corpus. 70
5.3 Distribution of NITP for Green vs. guitarists in the WJazzD. 71
5.4 Diagram showing the layout of notes on a standard guitar. 72
5.5 Distribution of Green's pitches in each octave. 74
5.6 Distribution of how long Green spent in each octave. 75
5.7 Examples of octave sequences in Green's improvisations. 77
5.8 Count of most frequent pitch vs. octave sequence length. 78
5.9 Distribution of fuzzy intervals used in octave transitions. 80
5.10 Example of a large ascending interval between phrases, *Round About Midnight* (1961), bars 34–37. 80
5.11 $\text{IOI}_{\text{BeatProp}}$ distribution for each octave transition. 81
5.12 TPC distribution for major, minor, and blues tonality modes. 83

5.13 Example of blues language in Green's improvisations, *Miss Ann's Tempo* (Solo 1, 1961), bars 73–76. . . . 83
5.14 $\mathrm{IOI}_{\mathrm{BeatProp}}$ distribution for each TPC weight. . . . 85
5.15 Distribution of absolute interval sizes (< 7 semitones) played into and out of a note for each TPC weight. . . . 86
5.16 Distribution of chordal pitch descriptors for every note in Green's corpus. CDPCX excludes ∘7 chords. . . . 90
5.17 CDPCX distribution of Δ7 chords in Green's corpus. . . . 92
5.18 Green's playing over IVΔ7 (IV) chords in *Red River Valley* (1962). Bar 11, 0 NHTs; bar 59, 1 NHT; bar 91, 2 NHTs. . . . 92
5.19 CDPCX distribution of m7 chords in Green's corpus. . . . 93
5.20 Green's distribution of CDPCX for the first note of a m7 chord vs. all other notes in the chord. . . . 94
5.21 m7 CDPCX distribution compared against a m7 and 7 substitution when the following chord was a 7. . . . 95
5.22 Examples of Green substituting a 7 for a m7 chord. Top: *I'll Remember April* (1961), bar 0. Bottom: *Miss Ann's Tempo* (Solo 1, 1961), bar 48. . . 95
5.23 CDPCX distribution for 7 chords in Green's corpus. . . . 96
5.24 Dominant 7 CDPCX distribution for each tonality mode in Green's corpus. 97
5.25 Example of a resolution from the ♭6 (♯5) of a 7 chord to the 3rd of a Δ7 chord, *The Surrey With The Fringe On Top* (1963), bars 12–13. . . . 98
5.26 Example of a resolution from the ♭2 (♭9) of a 7 chord to the 5th of a m7 chord, *Green With Envy* (Solo 2, 1961), bars 14–15. . . . 98
5.27 CDPCX distribution of each 7 function in a blues harmony. . . . 99
5.28 CDPCX distribution for ø7 chords in Green's corpus. . . . 100
5.29 CDPCX distribution of ø7 chords in the beat before a dominant resolution to a V7 chord, compared to the 7 chord. . . . 100
5.30 Example of an anticipation of a V7 chord over a ø7 chord in a minor ii–V pattern, *Brazil* (Solo 1, 1962), bars 23–25. . . . 101
5.31 CPC distribution for ∘7 chords in Green's corpus. . . . 102
5.32 Example of a diminished line, *Green With Envy* (Solo 2, 1961), bar 44. . 102
5.33 Distribution of $\mathrm{CPC}_{\mathrm{Weight}}$ for each metrical weight for Δ7, m7, and 7 chords. . . . 104

5.34 Conditional odds of metrical weight distribution of the 4^{th} for $\Delta7$, m7, and 7 chords. 105
5.35 CPC_{Weight} distribution of $\Delta7$, 7, and m7 chords depending on whether or not they were played in the beat before a chord change. 106
5.36 $IOI_{BeatProp}$ distribution for each CPC_{Weight}. 109
5.37 Articulation distribution for each CPC_{Weight}. 110
5.38 Frequency of altered chord tensions over 7 chords in Green's corpus. . . . 111
5.39 Proportion of NHTs in the beats before a chord change. 112
5.40 CDPCX transition around ♭3/♮3 over 7 chords. 113
5.41 CDPCX Transition around ♭6 over 7 chords. 113
5.42 Example of a large ascending interval followed by a descending line, *The Song Is You* (1961), bars 112–117. 118
5.43 $IOI_{BeatProp}$ distribution for repeated notes in Green's corpus. 119
5.44 Distribution of Parsons code for the 200 intervals following a repeated note in Green's corpus and distribution of repeated Parsons for the 200 intervals following a repeated note in Davis, Parker, and Coltrane's corpus.120
5.45 Distribution of Parsons code for the 200 intervals following a descending or ascending interval in Green's corpus. 121
5.46 Distribution of raw interval values in Green's corpus. 122
5.47 Example of a large (two octave) interval between phrases, *The Surrey With The Fringe On Top* (Solo 1, 1963), bars 19–21. 123
5.48 Distribution of fuzzy interval classes in Green's corpus. Left: Fuzzy interval classes. Right: Absolute fuzzy interval classes. 125
5.49 Distribution of $IOI_{BeatProp}$ and $duration_{BeatProp}$ for each fuzzy interval class in Green's corpus. 126
5.50 Distribution of NITP for each fuzzy interval class in Green's corpus. . . . 128
5.51 Example of a SNF, *Tico-Tico* (Solo 1, 1962), bars 15–16. 130
5.52 Distribution of SNF target tones in $\Delta7$, 7, and m7 chords in Green's corpus. 132
5.53 `2 1 ab` SNF interval sequence targeting a note over a $\Delta7$ chord, *Oleo* (Solo 1, 1962), bar 43. 134
5.54 `2 1 ab` SNF interval sequence targeting a note over a 7 chord, *Sonny-moon For Two* (1960), bars 84–85. 134

5.55 `2 1 ab` SNF interval sequence targeting a note over a m7 chord, *The Surrey With The Fringe On Top* (Solo 1, 1963), bars 1–2. 134

5.56 $\mathrm{IOI}_{\mathrm{BeatProp}}$ distribution for surrounding notes in a SNF. 137

5.57 Example of a SNF with a quaver-equivalent rhythm, *Green With Envy* (Solo 2, 1961), bars 47–49. 138

5.58 Distribution of target tone $\mathrm{IOI}_{\mathrm{BeatProp}}$ in SNFs. 139

5.59 $\mathrm{IOI}_{\mathrm{BeatProp}}$ transitions for SNFs, where each trigram occurred more than ten times. 139

5.60 Example of Green playing SNFs throughout his improvised lines, *Green With Envy* (Solo 2, 1961), bars 11–15. 140

5.61 Distribution of pairwise metrical weight sequences for bigrams which were or were not voice led. 143

5.62 Distribution of $\mathrm{CPC}_{\mathrm{Weight}}$ depending on whether when a note was the target of voice leading. 144

6.1 *MeloSpy*'s relative note length class distribution against the note length ratios and standard symbolic notation rhythms. 150

6.2 Distribution of $\mathrm{duration}_{\mathrm{BeatProp}}$ in Green's improvisations, the inset shows $\mathrm{duration}_{\mathrm{BeatProp}}$ greater than two. 152

6.3 Distribution of $\mathrm{IOI}_{\mathrm{BeatProp}}$ in Green's improvisations, the inset shows $\mathrm{IOI}_{\mathrm{BeatProp}}$ greater than two. 153

6.4 Proportion of notes that were labelled as short, medium, or long for each tempo class, for fuzzy duration and fuzzy IOI classes. 154

6.5 Proportion of fuzzy duration and IOI classes Green played in each metrical weight. 156

6.6 Distribution of $\mathrm{CPC}_{\mathrm{Weight}}$ for each fuzzy duration and IOI class in Green's corpus. 157

6.7 Metrical circle map for improvisations in $\frac{4}{4}$ in Green's corpus. 159

6.8 Beat distribution for improvisations in $\frac{4}{4}$ in Green's corpus displayed using raw metrical values. 160

6.9 Beat distribution for improvisations in $\frac{3}{4}$ in Green's corpus. Left: Metrical circle map. Right: Raw metrical values. 160

6.10 Metrical circle map and raw metrical values for notes played in each tempo range. Left: BPM ≤ 170. Right: BPM > 170. 161

6.11 Examples of differences in Green's beat distribution at both tempo ranges. a) Low tempo and highly varied beat distribution, *Blues In Maude's Flat* (1961), bars 54–59. b) High tempo with lower variety of beat distributions, *Oleo* (1962), bars 57–63. 162

6.12 Distribution of simulated metrical weights against rhythmic values in $\frac{3}{4}$ and $\frac{4}{4}$. 163

6.13 Distribution of metrical weights for improvisations in $\frac{3}{4}$ and $\frac{4}{4}$ in Green's corpus. 164

6.14 Example of a phase with a metrical weight distribution similar to Green's overall distribution in $\frac{4}{4}$, *I'll Remember April* (1961), phrase 33, bars 109–112. 165

6.15 Distribution of $\text{rest}_{\text{prop}}$ in Green's corpus. 167

6.16 Frequency of sequence lengths for notes that were not followed by a rest vs. those that were. The solid bar in the foreground shows the frequency for each sequence length, the lighter bar in the background shows the cumulative frequency. 168

6.17 Distribution of rest frequency for each phrase position in Green's corpus. 172

6.18 Example of Green using rests within a phrase, *Wives And Lovers* (1964), bars 131–135. 173

6.19 Example of a likely swallowed note resulting in a rest value being labelled, *The Song Is You* (1962), bars 65–68. 173

6.20 Proportion of intra-phrase fuzzy intervals that were or were not followed by a rest. 174

6.21 Distribution of the number of rests per bar in Green's corpus. 175

6.22 Distribution of $\text{rest}_{\text{prop}}$ for each phrase position in Green's corpus. 177

6.23 Mean metrical density of the most frequent beat divisions across the range of tempos. The size and opacity of the points showed the relative frequency of that division for each tempo segment. 180

6.24 Distribution of divisions in Green's corpus. 181

6.25 Single Beat MCM circle graphs for divisions 1, 2, 3, and 4 in Green's corpus. Top left: division 1. Top Right: division 2. Bottom Left: division 3. Bottom Right: division 4. 182

6.26 Distribution of divisions in each tempo range in Green's corpus. 184

6.27 Single beat metrical circle maps showing sub-beat movements for both tempo ranges in Green's corpus. Left: BPM ≤ 170. Fight: BPM > 170. . 184

6.28 Distribution of divisions in Green's corpus dependent on the tonality mode of the improvisation. 185

6.29 Examples of the difference in rhythmic variety of Green's improvisations in a major and blues tonality mode. Top: *I Wish You Love* (1964), bars 7–11. Bottom: *Freedom March* (Solo 1, 1961), bars 20–23. 186

6.30 Distribution of the metrical density, as number of notes per bar, in Green's corpus. 189

6.31 Comparison of metrical densities between Green and Parker, Coltrane, and Davis. 190

6.32 Distribution the number of notes per bar for the binary tempo ranges in Green's corpus. 191

6.33 Distribution of proportion of chromatic intervals for each metrical density in Green's corpus. 192

6.34 Examples of chromatic movement in metrically dense bars. a) Chromatic approach and target tones, *I'll Remember April* (1961), bar 11. b) Alternating chromatic intervals *At Long Last Love* (1965), bar 21. 193

6.35 Distribution of proportions of NHTs for each metrical density in Green's corpus. 194

6.36 Percentage of bars for a given metrical density that had at least one non-harmonic tone. 195

7.1 Symbolic notation equivalent of three possible BURs ranging from no swing (1:1) to heavy swing (3:1). 200

7.2 Example of a potential swing pair (bar thirteen, beat three) when using the duration of the second note, but not when using the IOI, *At Long Last Love* (1965), bars 13–14. 201

7.3 Distribution of BURs in Green's corpus. 203

7.4 Distribution of BURs for Green, Parker, and Metheny. 204

7.5 Distribution of Green's BUR for each of the binary tempo ranges. 206

7.6 The proportion of beats for each tonality mode in each tempo range that were classified as swung beats for which a BUR was calculated. 207

7.7 Distribution of Green's BUR for each tonality mode in both tempo ranges. 208

7.8 Distribution of note articulations in Green's corpus. 211

7.9 Proportion of notes in each articulation class played by Green in both tempo ranges 212

7.10 Example of a phrase at a lower tempo (BPM = 120) with most notes played with long articulation, *Round About Midnight* (1961), bars 31–32. 213

7.11 Example of a phrase at higher tempo (BPM = 240) with most notes played with short articulation, *The Song Is You* (1962), bars 25–28. . . . 213

7.12 Distribution of articulation values for each absolute fuzzy interval class in Green's corpus. 214

7.13 Example of Green varying the articulation of notes in a repeated sequence, *Freedom March* (Solo 1, 1961), bar 55. Top: Symbolic notation. Bottom: *SonicVisualiser* transcription. 215

7.14 Distribution of micro-gaps between notes in Green's corpus. Left: 0–0.01 beats. Right: 0.01–0.3 beats. 216

7.15 Distribution of micro-gaps for the four most frequent beat densities (notes per beat) in Green's corpus. 217

7.16 Distribution of micro-gaps for each of the binary tempo ranges in Green's corpus. 219

7.17 Distribution of note onset differences in Green's corpus. Left: raw difference in seconds. Right: difference as proportion of surrounding beat length. 221

7.18 Distribution of proportional onset differences for BPM $\leq$ 170 and BPM $>$ 170 in Green's corpus. 223

7.19 Distribution of proportional onset differences for divisions that contained more than 1000 note events in Green's corpus. 224

7.20 Distribution of proportional note offsets for notes played on or off the beat in Green's corpus. 225

7.21 Distribution of Green's onset difference for the first note of a swing pair, the second note of a swing pair, and non-swung notes. 227

8.1 Normalised note onset vs. pitch for all phrases combined in Green's corpus with line of best fit indicating overall trend. 233

8.2 Normalised onset vs. pitch for each contour code in Green's corpus. Line of best fit from fourth-order polynomial. 235

8.3 Distribution of phrase lengths, as both note count and beat sum, for phrases with thirty or fewer notes in Green's corpus. 237

8.4 Example of a phrase with the most common contour and length, *Born To Be Blue* (1961), phrase 6, bars 11–12. 237

8.5 Distribution of Green's phrase lengths for each Abesser contour code. . . 238

8.6 Example of a phrase with a repeated motif, *Seepin'* (Solo 1, 1960), bars 176–180. 239

8.7 Distribution of frequency of unique pitches for each fuzzy phrase length. Left: Unique pitch count. Right: Unique pitch proportion. 240

8.8 Distribution of CPC_{Weight} for notes that began a phrase compared to all other notes in Green's corpus. 242

8.9 Distribution of CPC_{Weight} for notes that ended a phrase compared to all other notes in Green's corpus. 243

8.10 Distribution of CDPCX for notes that ended phrases in Green's corpus. . 244

8.11 Parsons distribution for each phrase position in Green's corpus. 245

8.12 Distribution of raw intervals Green played in each phrase position. 246

8.13 MCM beat placement and bigram transitions for notes of Green's phrases. Top: first note. Bottom: all other notes. Left: quadruple time. Right: triple time. 247

8.14 Beat placement circle map for the final note in Green's phrases. Left: quadruple time. Right: triple time. 248

8.15 Overall bar-wise trends over the course of an improvisation for the metrical density, median interval, median pitch, NDT proportion, and NHT proportion. 251

8.16 Examples of the five feature trend mode shapes. The improvisations the data were drawn from were (top left to bottom right): *Brazil* (Solo 2, 1962); *Take These Chains From My Heart* (1963); *Sunday Mornin'* (Solo 1, 1961); *Our Miss Brooks* (1961); *Seepin'* (Solo 1, 1960). 252

8.17 Examples of common metrical density mode shapes in Green's corpus. . . 254

8.18 Examples of common median interval mode shapes in Green's corpus. . . 255

8.19 Examples of common median pitch mode shapes in Green's corpus. . . . 256

8.20 Examples of common proportion of NDT mode shapes in Green's corpus. 257

8.21 Examples of common proportion of NHT mode shapes in Green's corpus. 259

9.1 Summary of findings into Green's improvisational style 264

11.1 Condensed features that were selected as input variables for each abstraction. Left: one or two abstractions. Right: three or four abstractions. . . 275

11.2 Distribution of tempos for all performers in classification task. 276

11.3 Distribution of raw articulation values for all performers in classification task. 276

12.1 Decision tree for the n-way C4.5-like solo level model. 282

12.2 An example confusion matrix. 284

12.3 Number of models, for each comparison type, that performed significantly better than the NIR. 287

12.4 Phrase misclassifications between Green and Davis. a) Phrases from Davis misclassified as Green, *Eighty-one* (1964), bars 28–29. b) Phrase from Green misclassified as Davis, *Sonnymoon For Two* (1960), bars 140–141. 288

12.5 Frequency of simplified features across the abstraction levels from the C5.0 classifier in the n-way classification task. 298

12.6 Frequency of simplified features across the abstraction levels from the C5.0 classifier in the one-vs-all classification tasks. 302

12.7 Frequency of simplified features across the abstraction levels from the C5.0 classifier in one-vs-one classification tasks. 308

13.1 Distribution of NITP for the four performers compared to their main instrument in the WJazzD. 320

13.2 Distribution of interaction between NITP and octaves for each performer. 321

13.3 Distribution of TPC classes identified as important in the feature analysis for each performer. 322

13.4 Distribution of $\flat 3/\natural 3$ and $\natural 3/\flat 3$ CDPCX classes for each performer. 323

13.5 Example of phrases from Davis and Coltrane highlighting the use of the altered 3rd more common in their data. a) Davis, *Blues By Five* (1956), bars 85–89. b) Coltrane, *Nature Boy* (1963), bars 460–464. 324

13.6 Fuzzy Interval distribution for each performer. 325

13.7 Distribution of divisions for each performer. Left: all data combined. Right: at the $\text{bar}_{4|2}$ abstraction level. 326

13.8 Example of Parker playing a double time phrase with beats of division 8, *Embraceable You* (1947), bars 5–8. 327

13.9 Distribution of metrical densities for each performer. Overall distributions across all data, and distributions of mean metrical densities for the phrase and bar levels. 328

13.10 Proportion of notes played on each beat weight for all performers. 329

13.11 Example of Coltrane playing a high proportion of notes in the first beat of a bar, *Blue Train* (1957), bars 87–88. 330

13.12 Distribution of starting notes of phrases across the beat weights for each performer. 331

13.13 Distribution of fuzzy intervals into the final note of a phrase for each performer. 332

13.14 Distribution of fuzzy IOI for the last note of a phrase for each performer. 333

13.15 Examples showing specific beginning and ending phrase features in Green and Parker's improvisations. a) Green: *Tico-Tico* (1962), bars 93–96. b) Parker: *Yardbird Suite* (1946), bars 21–23. 334

A.1 Bigram Pitch Class Distribution for entire corpus of Green's improvisations. 368

A.2 Pitch class circle maps for each blues key signature in Green's corpus. . . 369

A.3 CDPCX distribution for m7 in the beat before a dominant resolution to 7, compared to both chord types, in Green's corpus. 369

A.4 Most common divisions for full bars in $\frac{4}{4}$, in Green's corpus. 370

A.5 Most common divisions for full bars in $\frac{8}{4}$, in Green's corpus. 370

A.6 Most common divisions for full bars in $\frac{3}{4}$, in Green's corpus. 371

C.1 Adaptive Spectrogram Settings. 378

C.2 Melodia - Melody Extractions Settings. 379

C.3 Bar and Beat Tracker Settings. 379

C.4 Example of potential chord annotation issues, *Red River Valley* (1963), bars 3–5. 382

Definition of Terms

Absolute: The size of a feature regardless of the direction.

Articulation: The ratio of the duration to the IOI (duration $\div$ IOI). When the duration and IOI were equal, the articulation was 1.

Beat Weight: Classifies a note as being played in the first beat of a bar, the last beat of a bar, or any other beat in a bar.

Chordal Pitch Class (CPC): Classifies a note in relation to the chord of the moment ($n \in [0 : 11]$, where 0 is the tonic).

Comping: Abbr. accompanying. The accompanying harmonic, rhythmic, and melodic material played behind the soloist. For a guitarist, the chords played in the head and solo sections.

$\mathbf{CPC_{Weight}}$: Classifies a note as an arpeggio tone, scale tone, or non-harmonic tone in reference to the chord of the moment.

Diatonic Tones (DT)/Non-Diatonic Tones (NDT): Classifies notes as either diatonic or non-diatonic to the tonality mode of the piece. Based on the ionian mode for major tonality modes, aeolian for minor tonality modes, and a nine-note scale combining the tonic and relative minor blues for blues tonality modes.

Double-stop: A technique where two notes, often on adjacent strings, are played simultaneously.

Duration: Time (seconds) between a notes onset and offset ($\text{Note}_{\text{offset}} - \text{Note}_{\text{onset}}$).

Fuzzy: A term applied to other features to indicate a variable where multiple levels were combined (e.g. fuzzy intervals, fuzzy IOI).

Gradient: Ratio of the change in pitch (ΔP) over the number of intervals (ΔT) between pitch extrema ($\Delta P \div \Delta T$).

Harmonic Tones (HT)/Non-Harmonic Tones (NHT): Classifies notes as either harmonic (arpeggio or scale) or non-harmonic to the nominal chord they were played over. Based on the ionian mode for $\Delta 7$, mixolydian for 7, dorian or phrygian for m7, locrian for ø7, and 8-note diminished scale (whole-half) for $\circ 7$ chords.

Inter-Onset Interval (IOI): Time (seconds) between subsequent note onsets ($\text{Note}^{+1}_{\text{onset}} - \text{Note}_{\text{onset}}$).

$\text{IOI}_{\text{BeatProp}}$ and $\text{Duration}_{\text{BeatProp}}$: The IOI or duration as a proportion of the surrounding beat length. A proportional value of 1 was equivalent to the length of a crotchet in simple time (IOI *or* Duration ÷ Beat Length).

Metrical Weight: Classifies a note as being played on a metrically strong beat (e.g. 1 and 3 in $\frac{4}{4}$), metrically weak beat (e.g. 2 and 4 in $\frac{4}{4}$), or off-beat.

Mode: The overall tonality of a piece, e.g. major or minor.

Note Placement: Classifies a note as being played ahead of the beat, behind the beat, or directly on the beat.

Onset Difference: The difference between a note's actual onset (o_a) and its nominal onset (o_n), as a proportion of the surrounding beat length (($o_n - o_a$) ÷ Beat Length).

Overfit: A process where a machine learning algorithm tunes the parameters of a model too closely to the training data such that it is unable to successfully classify the testing data, or other new data.

Parsons Code: Classifies an interval as ascending, descending, or repetition.

Percentage Point (PP): Used to describe the absolute change in percentages, instead of a proportional change in percentages. For example, an increase from 10% to 20% would be an increase of 100% or 10 PPs.

Phrase Position: Classifies a note as being the first note of a phrase, the last note of a phrase, or any other note in the phrase.

Pitch Class: Classifies a note compared to the pitch of C ($n \in [0:11]$, where 0 is C and 11 is B).

Pitch Extrema: Turning points in an improvisation, a note surrounded by notes that were both higher or lower in pitch.

Plurality: The most frequent class or group within a set, but where that class did not have a majority (>50%). For example, if Green only played three notes – C, E, and G – and the distributions were – C: 40%, E: 35%, G: 25% – Green played the note C the most but not more than half, so the plurality of notes played were C.

Tonal Pitch Class (TPC): Classifies a note in relation to the key of the piece ($n \in [0:11]$, where 0 is the tonic).

Tonality Mode: Specifies if a piece was in a major, minor, or blues key. A combination of *MeloSpy*'s `tonality_type` (e.g. Blues or Functional) and `mode` (e.g. major or minor) features.

Transcription: Within this research transcription does not refer to a symbolic notation of an improvisation. Instead, a broader definition of transcription is considered: the transfer of data from an audio format into another format that can be read or understood by a musician, researcher, or computer. For this research, the format of the transcription is one that a computer can read and understand. The Jazzomat Research Project coined the term “performance-based” (Frieler 2016b, slide 5) transcriptions for the highly detailed and descriptive – rather than prescriptive – transcriptions generated through their transcription and annotation process.

Mathematical Symbols

$\bar{x}$: Sample mean. Used with $\pm$ (plus or minus) to show the mean and standard deviation of a particular feature (except for when described in text, where it may be written as ‘the mean was found to be [mean] $\pm$ [standard deviation]’).

χ: Chi. Used in χ^2-tests.

η: Eta. Used in η^2 as an effect size for ANOVA.

p: p-value. Used to indicate the significance.

V: Cramer’s V. Used as an effect size for χ^2-tests.

d: Cohen’s d. Used as an effect size for t-tests.

List of Acronyms

Below is a list of acronyms and abbreviations used throughout the thesis.

Acronym	Description
ADSR	Attack, Decay, Sustain, Release
ANOVA	Analysis of Variance
BUR	Beat-Upbeat Ratio
CDPCX	Chordal Diatonic Pitch Class Extended
CPC	Chordal Pitch Class
CSV	Comma Separated Value
DT	Diatonic Tone
HT	Harmonic Tone
IOI	Inter-Onset Interval
ML	Machine Learning
NDT	Non-Diatonic Tone
NHT	Non-Harmonic Tone
NIR	No Information Rate
NITP	Normalised Instrument Tessitura Pitch
PC	Pitch Class
PP	Percentage Point
RNG	Random Number Generator
SD	Standard Deviation
SNF	Surrounding Note Figure
TPC	Tonal Pitch Class
TT	Tritone
UST	Upper Structure Triad
WJazzD	Weimar Jazz Database

Chapter 1

Introduction

The analysis of improvised jazz has been employed in the transmission of jazz language since its early beginnings, and is still an integral part of the pedagogical process at music institutions (Berliner 1994). The academic analysis of improvised jazz began in earnest in the 1950s with an increasing number of detailed studies being undertaken since the 1970s (Pfleiderer and Frieler 2010, 2). A large proportion of this research was undertaken using a small sample size of improvisations, often ten or less. With advances in computational processing power and the development of new software tools, analysing music using computational and statistical methods, a field sometimes referred to as "computational musicology" (Frieler 2020, 124), is an area of research that is experiencing considerable growth. These tools and approaches also allow for a broader investigation into a performer's improvisational style, with more data to support the conclusions that are drawn. Improvisational style has been defined as "those melodic, harmonic, rhythmic, and technical aspects that characterize an individual's improvisations[, referring] to those elements which [sic] are repetitively found in an individual's improvisations and, taken together, identify a solo as coming from [them]" (Brooks 2008, 1).

This research falls within the field of computational musicology, focusing on computer-aided musical analysis, and draws upon the authors training as a jazz guitarist. Computational approaches to music analysis in the field of folk and Western Classical Art music have a long history (Cuthbert and Ariza 2010; Huron 1991, 1994, 1996), with research out of the Jazzomat Research Project providing rapid advances to the computer-aided analysis of improvised jazz (Jazzomat Research Project 2017). Historically, most jazz analyses have been undertaken by hand (close-reading), and due to the substantial amount of time and effort required to complete detailed transcriptions and analyse them, most studies used a small sample size of improvisations. In contrast, the approach taken by the Jazzomat Research Project, and followed in this research, can best be thought of as distant-reading (corpus analysis), which looked for "general trends and ... patterns" (Frieler 2020, 124) in the corpus being studied.

This research built upon the previous work of the Jazzomat Research Project to develop a methodology to systematically explore the improvisational style of a single musician. To successfully develop a new methodology for undertaking the proposed analysis, the correct performer had to be selected, and for this research guitarist Grant Green was chosen. Specifically, forty improvisations from Green between 1960 and 1965 were selected to form the corpus to be analysed. Guitarists have been under-represented in these new approaches to analysis, and by being a predominantly monophonic improviser, Green's improvisations were well suited to this form of analysis. Green was a prolific musician throughout the 1960s and 1970s, and was the unofficial house guitarist for Blue Note Records from 1960 to 1965. Although Green influenced jazz guitar improvisation, and was talented and respected as an improviser (Erlwine 2017; Waggoner 2002), he did not fundamentally change jazz. For this research, where the focus was on developing a methodology, Green's conformity to the improvisational expectations of the styles in which he played was an advantage. It allowed for the analysis to provide insight into Green's improvisational style, validity of the methodology, and set a baseline for future analyses.

Beyond the use of the data to study an individual performer's improvisational style, the results from an analysis could be used to classify performers, with benefits to music information retrieval applications and musicological study (Herlands et al. 2014, 276). The results of the analysis of Green was subsequently used to inform the development of a methodology for undertaking performer classification and comparative analysis. This used additional improvisational data of John Coltrane, Miles Davis, and Charlie Parker from the Jazzomat Research Project's Weimar Jazz Database (WJazzD) (Pfleiderer, Frieler, et al. 2017). These three improvisers are amongst the largest names in jazz, with Parker also being referenced as a substantial influence upon Green (Green 1999, 6). They were also the most represented improvisers within the WJazzD, each with more than fifteen improvisations transcribed. Tree based machine learning (ML) algorithms were utilised within the performer identification task, with the comparative analysis based upon the musical features found to best classify the performers.

1.1 Aims and Research Questions

The overarching aim of this research was to develop a new methodology for systematically analysing the improvisational style of a performer, through the use of computer-aided and statistical methods to analyse their improvised lines. More specific aims are below, followed by the research questions developed to interrogate these aims. Further to the overarching aim, this research aimed to:

- Further the methods first applied by researchers in the Jazzomat Research Project by developing a new methodology for the in-depth analysis of a single musician's improvisational style;
- Continue developing the methods and knowledge of computer-aided statistical analysis of music;
- Create a database of 40 detailed transcriptions of Grant Green's improvisations;
- Define Green's improvisational style broadly with reference to his use of pitch, rhythm, micro, and macro features;
- Undertake a performer classification task through the use of machine learning algorithms, informed by the results of the prior analysis into Green;
- Develop a methodology for accurately classifying improvisers
- Classify musicians based solely on their improvisational content with a high degree of accuracy;
- Investigate which features are useful to successfully classify performers;
- Provide a meaningful interpretation of the results to describe similarities and differences between the performers.

The research questions developed to interrogate these aims were:

1. What new insights into improvisations could a computer-aided statistical approach to analysis provide?
   i) How could these methods be used to provide a systematic approach to jazz analysis, especially in defining a performer's improvisational style?
   ii) How could these methods be used to describe Grant Green's improvisational style between 1960 and 1965?
2. Which features were useful for describing the improvisational style of an individual musician?
   i) Which features could be extracted directly from the transcriptions?
   ii) Which features needed to be created from the data stored within the transcriptions?
   iii) Which independent variables were useful in describing how a musician's improvisational style changes under differing conditions?

3. How well did the results of the analysis explain Grant Green's improvisational style?
4. Within a jazz context, how well could the results of the analysis be used to inform a performer classification task through the use of machine learning?
   i) Which features were useful in accurately differentiating between performers?
   ii) Which machine learning classifiers were most useful for this task?
5. How could the results be interpreted in a meaningful musicological manner to aid in a comparative analysis of the performers?
6. What was the highest accuracy that could be achieved in a performer classification task using tree-based interpretable classifiers?

The terms 'improvisations' or 'solos' and 'improvisational style' had a specific meaning within this project. Based on standard approaches and consistent use of terminology, improvisations or solos referred to the sections of a song when a single performer was playing a series of lines such that they would generally be considered an improvisation. This definition explicitly excluded the improvisatory nature of the accompanying instruments, and any interactions between them and the soloist. However, this did not deny the substantial improvisatory nature of the rhythm section. Improvisational style, as defined above, referred to the elements or features of a performer's improvisations – note choice, note placement, swing, etc. – that had some consistency between solos such that their combination described a particular 'style' of that performer.

This study was not aiming to, or arguing for, this approach to be a replacement for the close analysis of jazz improvisations. As will be discussed in the literature review, although the close analysis of a small number of improvisations has limitations and drawbacks regarding generalisations about a performer's improvisational style, there are situations where that method is valid and appropriate for the task at hand. Instead, this study aimed to supplement and extend the possible approaches to analysing improvised jazz, including the development of a methodology that allowed for broader generalisations about a performer's improvisational style. It also aimed to develop a more efficient analytical methodology that could provide analysts another starting point for their analyses. This research was also not aiming to investigate and analyse the audio file or multiple instruments within the improvisational section of each recording. Although there is much to be gained from the analysis of an improvised line in the context of the other musicians, this research followed the approach of the vast majority of prior studies that focused on only the improvised lines. This was a necessity not only due to the scope of this project, but also due to the current limitations in storing and analysing this kind of multivariate data. It was not an analysis of the recording or the performance, but a computer-aided analysis of a transcription of the improvisation. This project drew upon the rich history of jazz transcription and analysis, while employing more recent technologies with a different approach. Additionally, the results of the comparative analysis did not imply that the similarities and differences observed were the only valid results from a comparative analysis, or that the comparisons were valid for all the musicians across their entire career. Due to the limited data from the musicians, the results of the comparative analysis were limited to the available data.

1.2 Treatment of Source Materials

The primary source materials used for this research were the collection of forty improvisations played by Green between 1960 and 1965. These improvisations were subsequently transcribed, collated, and had data extracted from them using methods developed by the Jazzomat Research Project. The transcriptions, unlike those using standard symbolic notation, were generated by annotating directly onto the waveform of the improvisation. These descriptive, performance-based transcriptions were highly detailed and generated a vast amount of data to be extracted and analysed (Pfleiderer, Frieler, et al. 2017, 73; Frieler 2016b, slide 5). Transcription is used here in the broadest definition of the transfer of data from an audio format into another that can be read or understood by a musician, researcher, or computer. For this research, the format of the transcription is one that a computer can read, understand, and extract data from. It was this data, and the subsequent processing of it, that formed the basis of the analyses within this research. The specifics of the selection of improvisations and following transcription can be found in the sections 'Data Selection' (3.1.1) and 'Transcription' (3.1.2) of Chapter 3. The transcriptions were chosen through a pseudo-random processes, with selection criteria that ensured they were somewhat representative of Green's output between 1960 and 1965. There was also supplemental data drawn from the WJazzD. This additional data was used sparingly throughout the analysis of Green to situate him, where appropriate, within the context of the available data. The improvisatory data of Coltrane, Davis, and Parker was used in the performer classification and comparative analysis.

1.3 Method of Research

Building upon four bodies of literature – traditional jazz analysis, computer-aided jazz analysis, computer-aided non-jazz musical analysis, and machine learning in music – this research developed a new computer-aided methodology for undertaking an analysis of a performer's improvisational style with the aid of statistical tools. The results of this methodology, the analysis of Green's improvisational style, were then applied to a performer classification and comparative analysis task. Software and methods developed by the Jazzomat Research Project allowed for the transcription of the improvisations and the extraction of data from the transcriptions. The analysis of Green's data, training and evaluation of the models used in the performer classification, and the comparative analysis were all completed in the *R* Statistical Language (R Core Team 2017).[1] The analysis of the features focused on univariate and bivariate statistics. Due to the vast number of features and interactions that could be analysed, part of the development of the methodology focused on a limiting mechanism to focus the analysis on areas with the greatest interest or importance.

1.4 Significance

This research demonstrates doctoral significance and originality in the following areas:

- computational musicology, specifically the computer-aided statistical analysis of music;
- filling a significant gap in jazz studies – the scarcely analysed work and improvisational style of influential guitarist Grant Green, including the contribution of a large number of new, highly detailed, transcriptions of Green's improvisations;
- the development of a new methodology for the in-depth computer-aided statistical analysis of a single performer's improvisational style;
- contribution to and development of methods in jazz analysis;
- building upon the analysis of Green, the development of a methodology for undertaking a performer classification task based solely on the improvisational content of solos.

As the computer-aided analysis of improvised jazz is a field in its infancy, the methodologies developed by this research contribute to a foundation for future

[1] Version 4.0.2.

research to build upon and develop. The results of the analyses contribute a proof-of-concept of the methodology's validity, practicality, and effectiveness, and a baseline against which to compare future analyses.

The development of a new methodology to undertake a detailed computer-aided statistical analysis of a single performer's improvisational style

Academic research into jazz gained traction throughout the second half of the 1900s, and is now a large field of study. This thesis contributes a new methodology for undertaking an analysis of a performer's improvisational style by employing computer-aided statistical methods of analysis. Through its top-down analytical methodology and structuring of domains, this research presents an approach that can scale to analyses of varying sizes, and allowing for both broad and deep investigations into improvisational style.

The contribution of an in-depth analysis of influential jazz guitarist Grant Green

Limited literature has been published about Green and his improvisational style, resulting in a significant gap that warrants serious academic study. Green was "underrated … during his lifetime" (Erlwine 2017) and "overshadowed by Wes Montgomery" (Waggoner 2002, 88). Green has been described as "a vital link … from Charlie Christian through Barney Kessel to Kenny Burrell and Wes Montgomery" (Green 1961i), and his "single-note lines … helped change the way [the] guitar [was] played" (Green 1999, xi). Green was also "the unofficial house guitarist for Blue Note [records]" (Scott 2012) in the early 1960s, and recorded 22 albums as a leader for Blue Note between 1960 and 1965, as well as recording prolifically as a sideman. The time period of interest in this research, 1960 to 1965, was selected not only because of the amount of output Green produced in this time, but also because it has been described as the period of time when his improvisational style was most mature (Waggoner 2002, 88).

The contribution of a large number of new Grant Green transcriptions

There are very few published transcriptions of Green's improvisations. Therefore, the forty transcriptions made for this research are a substantial new contribution of material for jazz researchers, students of jazz guitar, and musicologists interested in computer-aided analysis. As these transcriptions follow the same methods used to generate the WJazzD, they presented a significant contribution to a growing dataset of detailed jazz transcriptions that are required for future research. The current dataset of computer-readable detailed jazz transcriptions is also biased towards horn instruments, and contain few transcriptions of guitarists.

The development of a methodology to classify jazz musicians, based solely on the content of their improvisations, using machine learning algorithms

There have been many approaches to analysing or categorising music through the use of machine learning algorithms, including categorising music into different styles, genres, artists, and composers. Combining the new data from Green's improvisations with that of the four most represented musicians in the WJazzD, this research developed a new methodology for classifying performers based solely on the content of their improvisations. By using only the data from the transcriptions, the results of the classification can be used not only to identify the performers, but also be employed to undertake a comparative analysis of their improvisational styles.

The development of a methodology to use machine learning algorithms to provide meaningful musicological insight into the similarities and differences between the improvisational styles of jazz musicians

Prior research within Western Classical Art music has undertaken composer classification using machine learning tools while also attempting to provide musicological meaning from the results (Herlands et al. 2014; Herremans, Martens, and Sörensen 2016). These previous studies showed that it was possible to use machine learning models to accurately classify music while also providing meaningful musicological results. This approach has yet to be fully realised within the jazz idiom. This research, drawing upon the previous literature and results of the analysis into Green, presents a new methodology for undertaking this approach to classification and analysis that is specifically focused on jazz improvisation.

The contribution of an exploration into the most distinguishing improvisational features between Grant Green, John Coltrane, Miles Davis, and Charlie Parker

There have been previous studies that have investigated the differences between musicians, and due to their cachet as “jazz giants” they frequently have featured Coltrane, Davis, or Parker. This is the first study that employed machine learning to identify which features to analyse between these performers. This presents a novel approach for investigating the improvisational differences between performers.

1.5 Organisation of Thesis

This thesis is separated into three Parts, each covering a distinct process in the development of the new methodology for the computer-aided analysis of improvised jazz lines.

Part I: In Search Of A New Methodology, covers the background for the research, starting with the literature review in Chapter Two. The literature review is separated into five parts, focusing on the different areas of literature relevant to this study. Chapter Three initially discusses the overarching approach to developing the new methodology. This is followed by discussion of the steps that were required to be completed before the analysis could start, specifically: the choice of the improvisations to transcribe; the transcription process; and the extraction and treatment of the data.

Part II: Analysis of Grant Green’s Improvisational Style, is separated into six chapters, the first, Chapter Four, examines the general data that comprises the corpus of Green’s improvisational data. The following four chapters, Chapters Five through Eight, each investigate one of the four domains of improvisational style: Pitch domain; Rhythm domain; Micro domain; and Macro domain. Chapters Four and Five, the Pitch and Rhythm domains, are the largest of the domains, containing the most features as well as being the focal point of many prior studies. They examine feature-categories within their domain, finishing with an investigation into two specific examples. The Micro and Macro Domains, having a smaller subset of features from which to draw, are focused on the investigation of specific feature examples within each domain. Part II closes with Chapter Nine, summarising and bringing together the results from the four domains.

Part III: Performer Classification and Comparative Analysis, has five chapters and begins, in Chapter Ten, with a general discussion of the data used within the classification and comparative analysis. The next two chapters, Chapters Eleven and Twelve, relate directly to the classification task. Chapter Eleven discusses the features chosen to be included in the classification task at each abstraction level. Chapter Twelve reports on the results of the trained models, first investigating the performance metrics of the models, followed by an analysis of the features that were most frequently used to successfully classify the performers. Chapter Thirteen shows the results of the example comparative analysis based on a subset of the features from the feature analysis. Finally, Chapter Fourteen brings together and discusses the results from the chapters in Part III.

The thesis ends with the Conclusion, summarising the findings from Parts II and III. It discusses the overall results and findings from the analyses and reflects on the development of the methodologies. It finishes with a discussion on future work for computer-aided study of improvised jazz.

Following the thesis are five appendices. Appendices A through D are found entirely within the document. Appendix A includes additional graphs that were supplemental to the findings presented. Appendix B contains examples of specific code used throughout the thesis, with a complete list of code found in Appendix E.2. Appendix C contains more specific details regarding the transcription and annotation process. Appendix D lists the complete metrics for models that were found to significantly outperform the baseline accuracy. The final appendix, Appendix E, can be found entirely within the additional attached data, and contains: *Sonic Visualiser* transcriptions files, automatically generated symbolic notation transcriptions as a PDF, and the SQLite3 database of the transcriptions; files used to extract the data from the transcriptions, the extracted data, and *R* scripts to clean and process the data; files used to set up the data and train the ML models; the trained models; confusion matrices of the models on the training and testing data; RDS data files containing the processed data used throughout the research; files used to compile this document; and supplementary files.

A note on musical examples

All the musical examples from Green were written as they would be for a guitarist, one octave lower than concert pitch (𝄠). All musical examples were generated automatically using the custom `dataToLilypond` function (Appendix E.2), and were unedited unless other specified.[2] The examples were based on the raw transcription data, and therefore have more complicated rhythms than would be expected from a standard symbolic notation transcription.

[2]The `dataToLilypond` function automatically converted music in $\frac{8}{4}$ to $\frac{4}{4}$.

Part I
In Search of a New Methodology

Chapter 2

Literature Review

The prior research that forms the foundations for this thesis came from four separate, but interrelated, fields of study: jazz analysis; computer-aided music research, inclusive of Western Classical Art music (classical) and Western folk music (folk); computer-aided music research in jazz; and machine learning in music. This literature review highlights the current strengths and weaknesses of the prior work, drawing upon these to inform the development of the methodology presented within this research.

Additionally, an overview on literature related to Green was presented. This included his selection as the case study performer on which Part II was based. As this research focused on the analysis of improvisational style, and not a biography of Green, only limited biographical information was included. For a more detailed biography on Green, see *Grant Green: Rediscovering the Forgotten Genius of Guitar* by Sharony Andrews Green (Green 1999). S. A. Green followed up her book with a documentary, *The Grant Green Story*, that premiered at the 11$^{\text{th}}$ Annual Harlem International Film Festival in New York City, September 2016, published online in 2017 (Green 2017).

2.1 Analysis of Improvised Jazz

Throughout the history of jazz, the analysis of improvised lines has been used in the pedagogical process and transmission of jazz language (Berliner 1994). Although academic studies of improvised jazz have been undertaken since the 1950s, more frequent and detailed research has been published since the 1970s (Pfleiderer and Frieler 2010, 2). Masters and doctoral theses comprise a substantial amount of the available literature, with two common styles of study. The first were purely academic investigations into a performer (e.g. Martin 1996; Moore 2022; Ostercamp 2018; Owens 1974; Porter 1985; Weir 2006). The second aimed to analyse solos to inform the researcher's own improvisational style (e.g. Angus 2014; Hunt 2015; McEvoy 2014; Schnabel 2021).

A common approach has been to analyse improvisations to investigate a single, or small number, of features of a performer's improvisational style. These could include harmonic, rhythmic, and melodic patterns, as in Brooks' 2009 dissertation on Terence Blanchard (2008), or more specific features as in Reyman's 2011 dissertation on Andrew Hill (2011).[1] The number of improvisations used as the dataset in these studies varied, with some using more thirty (e.g. Angus 2014; Brooks 2008; Owens 1974; Weir 2006), while others had fewer than ten (e.g. Reyman 2011; Stewart 1973; Van der Bliek 1987; Zimmer 2016). A survey across thirty-six studies found that the mean number of improvisations used in an analysis was 13.06 ± 11.69, with twenty-four using ten or fewer. This included theses and published articles from 1973 to 2022.[2] Table 2.1 shows the number of improvisations used in each of the thirty-six studies.

Table 2.1: Number of transcriptions in each of the surveyed studies.

Author	Year	Transcriptions	Author	Year	Transcriptions
Stewart	1973	1	Angus	2014	34
Smith	1983	7	McEvoy	2014	10
Van der Bliek	1987	6	Hunt	2015	24
Kelly	1997	8	Zimmer	2016	8
Moore	1999	5	Stillman	2017	19
Korman	1999	4	Bechtel	2018	25
Kenny	1999a	9	Doyle	2018	10
Wild	2002	6	Ostercamp	2018	3
Freedy	2003	6	Sample	2019	5
Weir	2006	32	Baldwin	2020	8
Scott	2006	3	Pilzer	2020	22
Hodges	2007	22	Satterthwaite	2020	3
Brooks	2008	35	Gabric	2020	11
Gardiner	2008	10	Granville	2020	3
Alton-Lee	2010	10	Heyo	2021	4
Heyer	2011	49	Schnabel	2021	14
Love	2011	34	Moore	2022	10
Reyman	2011	7	Marcus	2022	3

[1] Including: approach and neighbour tones; enclosures; chromaticism; swing; note placement; and syncopation.

[2] Owens (1974) was excluded from these calculations due to the substantial outlier of 250 improvisations that no other study has approached.

The focus for the majority of the prior research was on specific features of a musician's improvisational style. However, there were studies that attempted to provide a broader overview; for example, Timothy Weir's 2006 PhD thesis on Kenny Dorham. Weir used 32 transcriptions to analyse and study the development of Dorham's vocabulary and improvisational style across his twenty-five year recording career. Weir's study suffered from two main issues regarding the sample size and selection of improvisations. The first related to how representative the improvisations were of Dorham across his whole career. Weir stated that "[in] order to complete this longitudinal analysis ... a selection of representative solos [were] taken from each ... [approximately] five-year section" (Weir 2006, 11) of Dorham's career. As Weir selected the solos and provided no discussion on the requirements for a solo to be selected as 'representative', it is unclear what bias may have been introduced in selecting improvisations that fitted his hypothesis. The second issue related to how well thirty-two transcriptions could capture the nuances of an improviser's style over such a substantial period of time. It was unlikely that such a relatively limited sample size would be able to capture anything but the largest changes in an improviser's style, and would struggle to be truly representative of their playing.

A continuing issue in these previous studies has been the sample size and selection criteria of the improvisations used to form generalised descriptions of a performer's improvisational style. These issues were more apparent in studies that aimed to investigate a larger number of features or a wider time span with a limited corpus. Sample size is important across research disciplines as it allows for confidence in the results of analysis and interpretation of results. Although larger sample sizes can alleviate some sampling selection bias, it is not itself enough to overcome this issue.[3] This review of the prior jazz research found that, historically, many studies were undertaken with small sample sizes (≤ 10) of improvisations. With advances in computational processing power and the development of new software tools it is now substantially easier to to analyse a larger number of improvisations, and apply statistical methods to these analyses. These tools and methods also allow for a broader investigation into a performer's improvisational style, with more data to support the conclusions that are drawn.

[3]Many of the issues regarding sample size and selection bias do not apply to performance-based studies, where the aim of the research was to use the analysis to influence the researcher's own improvisational style. In these cases, careful selection of improvisations that focus on the elements a researcher wants to incorporate into their own style was necessary. This concept was explored thoroughly in Schnabel's 2021 DMA dissertation *Using Solo Transcription To Develop A Personal Jazz Improvisational Style* (Schnabel 2021). Issues arise when these analyses make broader declarations about the studied performer's improvisational style.

2.2 Computer-aided Musical Analysis

The broadest definition of computer-aided musical analysis includes any analysis in which a computer was used to facilitate research. This could be very localised – using a computer to slow down and loop a section for transcription – or as in this research, used throughout the whole analytical process. This more specific definition, where the computer is critical to and extensively used throughout the research, is a more appropriate definition of what is more widely considered computer-aided musical analysis. Inherent in this definition, and discussion of computational musicology, is the use of software, and mathematical and statistical tools throughout the analytical process. In the preface to the 2003 book *Statistics In Musicology*, Beran, looking to the future of musicology, hypothesised that "[s]tatistics is likely to play an essential role in the future developments of musicology [because]: statistics is concerned with finding structures in data; ... statistical methods and structures are mathematical, and can often be carried over to various types of data ...; and ... musical data are massive and complex – and therefore basically useless, unless suitable tools are are applied to extract essential features." (Beran 2003, vii) In the two decades since this was published the use of computers and statistics in musicology and musical analysis has seen widespread use. As Meredith discussed in his preface to *Computational Music Analysis* (2016) computer-aided music analysis can absorb a number of "subdisciplines with names like 'mathematical music theory', 'computer music', 'systematic musicology', 'musical information retrieval', 'computational musicology', 'digital musicology', 'sound and music computing', and 'music informatics'." (Meredith 2016, v) However, "the extent to which the computer is doing 'music analysis' (as understood by musicologists) is uncertain ... [and it is better to] think of computational music analysis ... more like [a] forensic science" (Marsden 2016, 24). It can be used to answer "important and relevant questions for music analysis, but the final musical judgements ... will be made by people" (Marsden 2016, 24).

Some early applications in applying computers to music were from composers, including Milton Babbitt and David Cope, who used them for creative and compositional aims, instead of purely analytical; with work in computational composition ongoing and current. Babbitt published at article, "The Use of Computers in Musicological Research" in 1965, but with his own caveat that "[he was] not a computer expert, [nor was he] a musicologist" (Babbitt 1965, 74). This showed that even from the earliest days of general purpose computing becoming more widespread, musicians were investigating and pursuing the potential applications to music creativity and research.

Computational approaches to music analysis in the field of classical and folk music have a long history. An early computational approach was the Humdrum Toolkit, developed in 1986 by David Huron, although now largely disused due to its complicated notation system and syntax. Studies that used this system include an analysis of the use of dynamics by fourteen composers in 435 piano pieces (Huron 1991) and an investigation into the contour of melodic phrases in 36 705 folk song phrases (Huron 1996). The Humdrum Toolkit also inspired the creation of a more recent and widely used tool, music21, developed in 2008 by Michael Cuthbert and Christopher Ariza (Cuthbert and Ariza 2010). Music21 is a toolkit developed in python for computer-aided musicological research, and works with many common file types including MusicXML and MIDI. This allows music21 to be used to analyse a wide range of publicly available data, as there exists myriad classical and folk music freely available in these formats. However, these software tools were designed to predominantly study classical music and folk songs, and are not suited for the analysis of improvised jazz.

A substantial difference currently exists in the available sample size between classical music and improvised jazz. For example, a study of set-classes by Agustin Martorell and Emilia Gomez (2016). Their corpus included used more than 16 000 MIDI tracks from a range of classical composers, “anonymous medieval pieces, church hymns, and the Essen Folksong collection” (Martorell and Gomez 2016, 104). Music21 includes 27 collections of songs, including the works of J.S. Bach, Beethoven, the Essen Folksong Collection, and a collection of Fourteenth-Century Italian music. In comparison, the largest collection of computer-readable detailed jazz transcriptions is the Jazzomat Research Project’s WJazzD, containing 456 improvisations (Pfleiderer, Frieler, et al. 2017). This presented a clear need for the creation of more transcriptions to increase the available sample size of improvisations to be used in the analysis of improvised jazz.

Although improvised jazz could be studied using programs such as music21, there are issues with how the musical data is stored. MusicXML stores its data as symbolic notation, while MIDI contains little more than pitch, note onset, and duration information. Symbolic notation is the storage of musical data as it would appear on a score. The symbolic representation of music has been criticised for being “inadequate in realising many of jazz improvisation’s subtle nuances” (Kenny 1999b, 73). The problem being that to fully “represent [an improvisation] within [the symbolic notation] system of ‘grids’ [requires the addition of] substantial annotations . . . rendering the resulting [transcription] difficult to read and interpret” (Kenny 1999b, 73). Although this related to physical scores of symbolic notation, the same concept applies to digital representations such as MusicXML.

Therefore, a better approach was to develop a new purpose-built system for transcribing and storing improvised jazz.

The Jazzomat Research Project's solution was a "performance-oriented" (Frieler 2016a) representation of music. Their transcription and storage system was based on the audio file of an improvisation, with annotations of notes, beats, chords, form, and articulations placed on top. This allowed for the "subtle nuances" (Kenny 1999b, 73) of jazz to be fully represented. These performance-oriented transcriptions were designed to be computer-readable rather than human-readable. The data from these transcriptions could then be used for computer-aided analysis.[4] This raw representation of the musical data is better suited to the computer-aided analysis of improvised jazz. Therefore, their software solution, *MeloSpy*, and methods were selected to use in this research[5]

2.3 Computer-aided Jazz Analysis

Introducing the Jazzomat Research Project, the authors highlighted issues with traditional methods in jazz analysis. They referenced Ekkehard Jost's "two-step methodology for jazz analysis" (Jost in Pfleiderer and Frieler 2010, 3) of "first listen[ing] to all available recordings of a musician or ... group, and then choos[ing] the typical pieces and analys[ing] them in detail" (Jost in Pfleiderer and Frieler 2010, 3). They argued: "why [listen] to all the records and ... [choose only a few] typical examples? Why not rely analytically on as many improvisations as possible ... [following the logic] that the more examples one can rely on analytically the more valid are the results of the analysis" (Pfleiderer and Frieler 2010, 3). Historically, most analyses have been undertaken by hand, and due to the substantial amount of time and effort required to complete detailed transcriptions and analyse them, most studies used a small sample size of improvisations. As technology has improved, it has become more practical to use more improvisations during research. Ideally, an analysis would include all improvisations recorded by a musician. However, considering the current impracticality of this approach, a relatively large and representative sample size of improvisations should produce a statistically valid analysis of a musician's improvisational style. Aside from

[4]This data could be converted to symbolic notation, but would then suffer the same issues as other symbolic representations of music.

[5]*MeloSpy* is used throughout this document to refer to the software suite developed by the Jazzomat Research Project. This included both the *MeloSpyGUI* and *MeloSpySuite*, a command line interface (CLI) of the same software. Although there are some minor differences between the GUI and CLI, they both contain the three main tools developed by the Jazzomat Research Project: "**melconv** for converting melodic file formats, **melfeature** for feature extraction, and **melpat** for pattern mining" (Jazzomat Research Project 2017, sic). These tools are based on a Python library, *MeloSpyLib*, which has yet to be publicly released.

advantages in speed, efficiency, and a reduction of bias from an increased sample size, a computer-aided approach to analysis does not rely on "the [researchers] listening experiences … and [their] ability to find and remember significant features within the music" (Pfleiderer, Zaddach, and Frieler 2016, 17). Instead a computer can immediately find and "remember" all features of an improvisation, and any connections or similarities between other pieces.

Jazzomat Research Project

To date, the most mature computer-aided jazz research has been published by researchers associated with the Jazzomat Research Project, or using their database or tools. Through their research, they developed a new system for transcribing, storing, and extracting data from improvisations. This included the development of their software suite *MeloSpy* and their database of transcriptions, WJazzD. These methods were then applied to sample studies to demonstrate the applications of their research, including: exploring the structure of phrases (Frieler 2014b; Frieler, Zaddach, and Abeßer 2014); automatic style classification using rhythm, tempo, and tonality (Eppler et al. 2014); an investigation into feature trends across the timespan of their database (1925–2009) (Frieler 2018); a comparison of the improvisational styles of Miles Davis and John Coltrane (Frieler 2016b, 2020); and an investigation into the distribution of swung notes (Corcoran and Frieler 2021). This prior research was required to explore and demonstrate the potential applications of their methods. No research employing their methods to undertake a large and detailed study of an individual musician has been released so far.

The majority of the published research has been from those associated with the project. However, other research has used elements of their work, most notably the WJazzD, including: research into description logic (Kantarelis et al. 2021); investigations into the relationship between interval patterns and metrical positions (Cross and Goldman 2021); applications of linguistic corpus analysis tools to jazz improvisation (Norgaard and Römer 2022); and using transcribed material to improve chord prediction models (Driever et al. 2022).

At the end of 2017, the Jazzomat Research Project published their end of project book, *Inside the Jazzomat* (2017). Within the book the authors covered the need to develop computational tools for the analysis of jazz, their transcription method, and discussed the types of features that could be extracted from the improvisations. Nine case studies were presented within the book, demonstrating applications of their methods. Nearly all of the case studies investigated only a single improvisation, with the largest dataset including fourteen solos, eight from Paul

Desmond and six from Chet Baker (Pfleiderer, Frieler, et al. 2017, 175). Within the nine case studies, the most frequently analysed features were: raw pitch features (pitch range and chromaticism); chordal pitch class; interval size; and metrical density, (notes per bar, second, or phrase). These case studies were the most detailed and practical use of the software and methods so far. However, they were still limited in scope, focusing on small samples sizes and a limited number of features. These case studies provided a starting point for developing a methodology for an in-depth analysis of a single improviser.

2.4 Machine Learning in Music

There have been many approaches to analysing or categorising music through the use of ML algorithms. These algorithms have been used to categorise music into different styles or genres (Abeßer, Dittmar, and Großmann 2008; Corrêa and Rodrigues 2016; Cuthbert, Ariza, and Friedland 2011; Nasridinov and Park 2014), artists (Abeßer, Dittmar, and Großmann 2008; Ramirez, Maestre, and Serra 2010; Zanon and Widmer 2003), and composers (Abeler 2015; Cuthbert, Ariza, and Friedland 2011; Hajj, Filo, and Awad 2017; Hedges, Roy, and Pachet 2014; Herlands et al. 2014; Herremans, Martens, and Sörensen 2016; Kaliakatsos-Papakostas, Epitropakis, and Vrahatis 2010; Lee 2008). Most of these studies used data from symbolic notation, often extracted from MIDI files, or used features extracted from raw audio. Excluding the style or genre categorisation tasks, which by necessity used a broad range of music styles, most focused on classical music.

There have been few uses of ML to classify jazz, with this area of study hampered by the lack of many high quality transcriptions. In examining broad differences between styles, instruments, and performers in Pfleiderer and Abeßer (Pfleiderer, Frieler, et al. 2017, 87) used a random forest to select features that differed most significantly between the two datasets they constructed. These features were then used to describe some differences between the groups. In Abeßer et al. (2008) the authors extracted features from “excerpts … of 20 to 40 seconds” (Abeßer, Dittmar, and Großmann 2008, 2) from both MIDI and raw audio files to perform genre and artist classification. Within the artist classification they undertook two experiments, one based on classifying four jazz saxophonists, and one classifying four electric guitarists. Using thirty excerpts, a support vector machine achieved an accuracy of 58.8% for the guitarists and 56.0% for the saxophonists (Abeßer, Dittmar, and Großmann 2008, 4).

In Ramirez et al. (2010) the authors use "sound analysis techniques ... in order to extract features such as pitch, timing, amplitude, and timbre" (Ramirez, Maestre, and Serra 2010, 1523). They used these features to categorise three saxophonists playing a strict rendition of four melody lines without ornamentation recorded within a controlled studio environment and four saxophonists from commercial recordings (Ramirez, Maestre, and Serra 2010, 1521). For the studio recordings they had a total of 792 note events per performer, and 820 note events per performer for the commercial recordings. The authors used decision tress, support vector machines, artificial neural networks, k-Nearest Neighbour, and ensemble methods to classify the musicians based on three abstraction levels, single notes, short phrases, and long phrases. In all cases, the single note level performed worse (maximum accuracy of 44.9% for the studio recordings and 29% for the commercial recordings). The short and long phrases performed equally as well, and substantially better than the single note (approximately 98% for the studio recordings and approximately 76% for the commercial recordings) (Ramirez, Maestre, and Serra 2010, 1522). This indicated that classification with a high degree of accuracy should be possible with larger datasets and more detailed features.

Within classical music there has been research examining the use of composer classification, while also attempting to extract musicological meaning from the results. In Herlands et al. (2014) the authors categorised the works of Haydn and Mozart based on melodic and rhythmic features (Herlands et al. 2014, 279) using a support vector machine, naive Bayes, decision trees and random forest, and ensemble classifiers (Herlands et al. 2014, 280). They achieved an an average accuracy of up to 80% when classifying based on global features.[6] By investigating the features that were most discriminating between the two composers, the authors found evidence that the music of Haydn contained "more virtuosic writing for the first violin" (Herlands et al. 2014, 281), which probably "emanate[d] from [the] personal and social circumstances" (Herlands et al. 2014, 281) of the composers.

Similarly, Herremans et al. (2016) used decision trees, rulesets, logistic regression, support vector machines, and naive Bayes to classify the music of Bach, Haydn, and Beethoven (Herremans, Martens, and Sörensen 2016, 370). They chose these machine learning models because "most ... are comprehensible and offer insight into the styles of the composers ... [although] a few black-box models were also built as a performance benchmark"(Herremans, Martens, and Sörensen 2016, 370). They selected twelve features to be extracted from a symbolic representation of 1045 movements from the three composers combined.[7] The authors achieved the highest

[6] The entire first movement of a sonata (Herlands et al. 2014, 277).

[7] Including: the proportion of chromatic intervals; most common melodic interval; most common pitch; proportion of repeated notes; and frequency of an interval of size two.

accuracy with the black-box models (93%), but their lowest accuracy for the interpretable decision tree was still quite high at 80% (Herremans, Martens, and Sörensen 2016, 376). Based upon the results of their comprehensible models, the authors were able to determine that, at least within the sampled dataset, "Beethoven [did] not focus on using one particular interval, in contrast to Haydn or Bach" (Herremans, Martens, and Sörensen 2016, 384). These two studies highlighted the possibility of achieving high accuracy in a performer identification task, while also providing musicological results.

2.5 Grant Green

Green was a prolific jazz guitarist, working in the early 1960s in a hard bop and post-bop style, and returning in the 1970s as a pioneer of the jazz-funk movement. From 1960 to 1965, the era this research focused on, Green was "the unofficial house guitarist for Blue Note [records]" (Scott 2012). He recorded twenty-two albums as a solo artist for them, and thirty-three more as an accompanist, with artists such as Lou Donaldson, Hank Mobley, Stanley Turrentine, and Herbie Hancock. In this period he also recorded *His Majesty King Funk* for Verve in 1965, *Reaching Out* for Black Lion in 1961, as well as being an accompanist on sixteen other non-Blue Note albums.

There has so far been limited serious study of Green's improvisational style. In total, there have been: two papers published by Andrew Scott (2006, 2009); a biography by his daughter-in-law Sharony Andrews Green (1999); a documentary that followed up the biography also by Sharony Andrews Green (2017); two "method" books, one published by Wolf Marshall (2004) for Hal Leonard in the 'Signature Licks' series and the other by Corey Christiansen (2003) for Mel Bey in the "Essential Jazz Lines" series; and three theses where Green was either the focus of the study (Bechtel 2018; Wild 2002), or was included amongst other guitarists (Kaiser 2013). Two of the theses were Masters (Kaiser 2013; Wild 2002) while the other was a Doctor of Musical Arts dissertation (Bechtel 2018). At the beginning of this research it was to be the first major analytical research into Green's improvisations, but since then there has been the concurrent research published by Bechtel (2018).

Wild's research focused on Green's playing in six solos over ii-V-I progressions, one of the most common chord progressions in jazz. Wild described the construction of Green's lines as "comprising [of] three phases: establishment of target notes, skeletal framework, and embellishment." (Wild 2002, 15) Although this was an efficient way of investigating the construction of Green's lines, it was not specific to Green. Wild's analysis provided a description of improvisational tools, using the solos of

Green as an example, rather than a specific analysis of "Green's Approach to Improvisation" (Wild 2002). Wild finishes off his analysis by providing examples of patterns that Green used within the studied improvisations.

Kaiser's research investigated "seven solos from seven jazz guitar legends" (Kaiser 2013, 3), with one solo from Green, "You Stepped Out Of A Dream" from 1961. His analysis focused on Green's use of upper structure harmonies and chromatic movements in his lines. Kaiser found that Green "used chromaticism to begin phrases and approach chord tones" (Kaiser 2013, 17). He also found that Green's use of diminished seven harmony over a dominant sonority was "particularly notable" (Kaiser 2013, 19) and its use "creat[ed] a strong resolution to the tonic" (Kaiser 2013, 19). He summarised Green's improvisation over "You Stepped Out Of A Dream" as "lines which [sic] emphasize[d] extended harmony and connect[ed] chord tones" (Kaiser 2013, 22), and that improvisers including Green "[were] responsible for the continuation of the bebop language throughout the 1960s" (Kaiser 2013, 22).

The most substantial prior study into Green was Bechtel's 2018 DMA dissertation "Grant Green: An Analysis of the Blue Note Guitarist's Musical Vocabulary". In comparison to this study, which focused on the analysis of Green's improvisational style, Bechtel's research aimed to "create an in-depth analysis of the improvisational material used by ... Green" (Bechtel 2018, 12) and "[establish] the key characteristics of his musical vocabulary" (Bechtel 2018, 13). Bechtel's research focused on the same time period as this study, 1960–1965, and of the twenty-five solos he transcribed, only six overlapped with those of the current project.[8] Bechtel's selection of improvisations followed that of much of the prior research, which "was ... based on historical significance of the performance, critical acclaim, and quality of the recorded material" (Bechtel 2018, 15).

Bechtel based much of his analysis of Green's musical vocabulary from Jerry Coker's *Elements of the Jazz Language for the Developing Improviser*, selecting "[five] elements from Green's transcriptions as improvisational devices that he uniquely adapted to the guitar" (Bechtel 2018, 15). Three of these elements were reoccurring phrases ("Honeysuckle Rose", 3-♭9, CESH[9]), while the other two were the use of the blues and digital patterns (Bechtel 2018, 16). Following the harmonic vocabulary analysis, Bechtel analysed Green's rhythmic vocabulary, specifically his time feel and rhythmic motives, and his sound concept.[10] Having selected the devices from

[8] *I'll Remember April* (Green 1961k), *I'm An Old Cowhand* (Green 1964b), *Minor League* (first solo only, Green 1964d), *Sonnymoon For Two* (Green 1960c), *Take These Chains From My Heart* (Green 1963f), and *Wives and Lovers* (Green 1964h).

[9] Contrapuntal Embellishment of Static Harmony

[10] "Through the analysis of photographs and ... limited ... video, [Bechtel established] how [Green's] unique sound concept was produced through [his] use of equipment and technique." (Bechtel 2018, 16)

Coker's book, Bechtel searched through Green's improvisations to find examples of these techniques. Although Bechtel was able to find many examples, it was as if the devices from Coker led the analysis, instead of Green's transcriptions leading to the identification of musical vocabulary. The issue of deciding on musical vocabulary and then looking for it in Green's improvisations is epitomised in Bechtel's analysis of digital patterns. By his own admission, as Green "developed as a musician in the pre-Giant Steps era, it can be assumed that [he] did not make conscious effort to focus on … these patterns" (Bechtel 2018, 67). As a result, Bechtel's analysis of Green's musical vocabulary would be useful for a musician trying to apply improvisational concepts they are already aware of to emulate some elements of Green's improvisations. It did not succeed as an investigation into the elements of Green's musical vocabulary.

2.6 Summary

The methods and features used throughout this research were informed by four bodies of literature. Part II, which focused on the analysis of Green's improvisational style, was informed by: research associated with the Jazzomat Research Project[11]; theses and articles with a focus on jazz analysis[12]; and computer-aided research into classical, folk, and pop/rock music[13]. The jazz analysis, while not computer-aided, identified analytical tasks frequently undertaken in more traditional analyses, including features frequently investigated, which were then adapted to a computer-aided and statistical methodology. Similarly, although the non-jazz computer-aided research could not inform analytical tasks or features, they did inform the statistical approach. Finally, research out of the Jazzomat Research Project was able to inform both the analytical tasks and features as well as

[11] Abeßer (2016); Abeßer, Frieler, Pfleiderer, et al. (2013); Abeßer, Frieler, Cano, et al. (2017); Beaty et al. (2022); Corcoran and Frieler (2021); Dittmar, Pfleiderer, and Muller (2015); Eppler et al. (2014); Frieler (2014b); Frieler (2014a); Frieler (2020); Frieler (2016a); Frieler (2016c); Frieler (2018); Frieler, Pfleiderer, Zaddach, et al. (2013); Frieler, Pfleiderer, Abeßer, et al. (2016); Frieler, Zaddach, and Abeßer (2014); Janssen and Kranenburg (2014); Lartillot (2014); Pfleiderer (2014); Pfleiderer (2016); Pfleiderer, Zaddach, and Frieler (2016); Pfleiderer, Frieler, et al. (2017); Pfleiderer and Frieler (2010); Zaddach (2016); Zaddach and Pfleiderer (2016)

[12] Alton-Lee (2010); Angus (2014); Bechtel (2018); Bellaviti (2005); Brooks (2008); Butterfield (2011); Cook (2012); Cross and Goldman (2021); Freedy (2003); Friberg and Sundström (1997); Friberg and Sundström (2002); Gardiner (2008); Hernandez (2020); Heyer (2011); Hodges (2007); Hunt (2015); Kaiser (2013); Kenny (1999b); Kenny (1999a); Korman (1999); Larsen (2021); Martin (1996); McEvoy (2014); Moore (1999); Moore (2022); Ostercamp (2018); Owens (1974); Potter (1992); Prince and Schmuckler (2014); Reyman (2011); Salmon (2011); Satterthwaite (2020); Scott (2006); Scott (2009); Smith (1983); Solstad (2015); Stewart (1973); Stillman (2017); Weir (2006); Van der Bliek (1987); Zimmer (2016)

[13] Burgoyne, Wild, and Fujinaga (2013); Cuthbert (2017); Cuthbert and Ariza (2010); Giomi and Ligabue (1991); Huron (1991); Huron (1996); Marsden (2016); Martorell and Gomez (2016); Meredith (2016); Rolland and Ganascia (2002); Temperley and Clercq (2013)

computer-aided and statistical approaches to undertaking the analyses. The fourth body of literature, associated with machine learning in music, informed Part III, the performer classification and comparative analysis.[14]

These bodies of literature provided a foundation on which to develop a new methodology for a computer-aided and statistical analysis of improvised jazz. The development of this methodology, along with the prior literature, also formed the basis of the example performer classification and comparative analysis. Recent research has shown a continued interest in the improvisations of Green, and his importance to the ongoing legacy of improvised jazz guitar; however, there are still substantial gaps to be filled in the understanding of his improvisational style.

[14] Ali and Siddiqui (2017); Choi et al. (2017); Corrêa and Rodrigues (2016); Cuthbert, Ariza, and Friedland (2011); Driever et al. (2022); Frederico (2006); Giraldo and Ramirez (2017); Hajj, Filo, and Awad (2017); Hedges, Roy, and Pachet (2014); Herlands et al. (2014); Herremans, Martens, and Sörensen (2016); Ishwaran (2007); Kaliakatsos-Papakostas, Epitropakis, and Vrahatis (2010); Lee (2008); Nasridinov and Park (2014); Ng and Breiman (2005); Ramirez, Maestre, and Serra (2010); Widmer (2000); Zanon and Widmer (2003)

Chapter 3

Approaching A New Methodology

This chapter describes the overarching approach used throughout the research, beginning with discussion of: the approach; data setup; analysis of Green; and performer classification and comparative analysis. Each of the steps are discussed, including limitations and issues. The top-down methodology for the analysis of Green's improvisational style is then presented, along with the statistical methods used to complete the analysis. The process of this methodology and the results of the analysis are presented throughout Part II. Finally, the methodology for completing the performer classification and comparative analysis is discussed, with this presented throughout Part III.

Overarching Approach

The overarching approach of this research was comprised of the following steps:

1. Selection of performer and improvisations;
2. Transcription of improvisations;
3. Extraction of improvisational data;
4. Analysis of Green's improvisational style through a domain-based top-down methodology;
5. Training of ML models for performer classification;
6. Evaluation of model accuracies and feature importance;
7. Comparative analysis informed by feature importance.

The prior literature review informed the methods for analysis and the features to target. These steps are expanded upon in the sections below. Section 3.1, Data Selection and Preparation, focuses on steps 1, 2, and 3, Section 3.2, Analysis of Grant Green, on step 4, and Section 3.3, Performer Classification and Comparative Analysis on steps 5, 6, and 7. The methodologies for the analysis of Green and the performer classification and comparative analysis are presented, with the results and application of these methodologies explored throughout Parts II and III.

3.1 Data Selection and Preparation

3.1.1 Data Selection

The first step was the selection of the performer and improvisations to be transcribed, which formed the dataset for the analyses. The selection of the performer was based on the criteria, with the performer needing to be:

- a guitarist, based on the author's own experience as a guitarist and the limited number of guitarists in the WJazzD;
- a predominantly monophonic improviser, as limitations in *MeloSpy* excluded many guitarists;
- an improviser who had little to no serious studies examining their improvisational style;
- an improviser the author was familiar with, but had not previously studied closely, to limit bias in the analyses.

The selection criteria limited the number of available performers, with Green being the most suitable candidate. The improvisations to be transcribed were then drawn from Green's solos between 1960 and 1965, when he was the unofficial house guitarist for Blue Note Records (Scott 2012). Following the approach of Brooks in his thesis on Terence Blanchard (2008), only albums where Green was listed as the leader were considered to draw improvisations from.[1] This was because it could be "assumed that [the leader] has the most artistic control [in] those recordings and is not modifying [their] normal style to accommodate … another leader" (Brooks 2008, 19).

Green recorded twenty-four albums as a leader between 1960 and 1965 with a total of 147 songs. For this research thirty songs (20.41%) were selected from which to transcribe the improvisations. To ensure that the selection of the improvisations was not biased, the songs were selected through a pseudo-random process. The process was pseudo-random as the selected improvisations needed to be somewhat representative of Green's output in this period. Specifically, they needed to be representative of the number of songs recorded each year and the mean tempo of the 147 tracks. These two features were selected as they provided a broad overview of Green's improvisational output while only requiring limited pre-processing of the solos. As many of Green's albums were released long after they were recorded, often posthumously, the recording date was used to calculate the improvisations per year.

[1] The album *Reaching Out* was originally released in 1961 under Dave Bailey. It was subsequently re-released under Green's name in 1973 as *Green Blues* and again with the original title 1989. Since it was recorded between 1960 and 1965 and released under Green's name, the album was considered to draw tracks from.

The tempo of each song was calculated using the program *Transcribe!* (Robinson 2014), where the placement of beats were manually tapped during playback and the 'Calculate Tempo' feature was used to obtain a final tempo value.[2]

These two variables were calculated for each of the 147 tracks, with each track assigned an index for reference in the random selection process. The index number, year, album, title, and tempo (BPM) was stored for each track in a CSV file, which was used in the pseudo-random selection. A program was written in *C#* to randomly select thirty representative tracks, seen in Figure 3.1.[3]

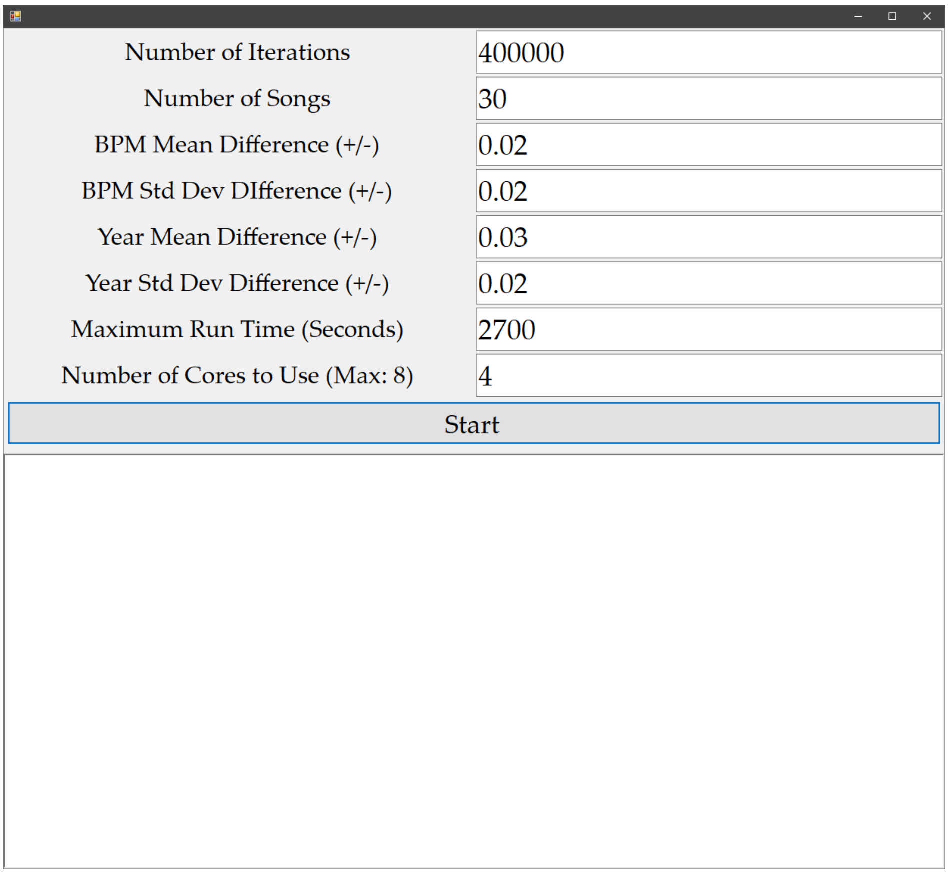

Figure 3.1: Screenshot of the program written to randomly select the improvisations.

[2]For songs where the improvisations were performed in double-time, the tempo of the head was calculated, instead of the double-time feel.

[3]The CSV file and random selection program are available in Appendix E.8. An exhaustive search of all thirty tracks was not possible due to the number of possible combinations (`147 choose 30` $= 1.64 \times 10^{31}$).

Selection criteria required the thirty tracks to closely match the mean and standard deviation (SD) of the tempo and recording year.[4] The program followed this procedure:

1. Calculate the mean and SD of the tempo and recording year for entire corpus, based on the input csv file;
2. Randomly select thirty tracks (Figure 3.1, option row two);
3. Calculate the mean and SD of the tempo and recording year for the current selection;
4. If the results fall within the limits set in the program (option rows three to six), print the index of the songs in the output field (bottom half of the program), otherwise nothing is printed;
5. Repeat for the number requested iterations (option row one).

This process is repeated until the maximum run time is reached (option row seven), or the 'Stop' button is pressed. The program was designed to run in batches of smaller iterations to allow for monitoring of the program in regards to the run time, and to allow for automatic stopping after a set period of time. The final option (row eight), 'Number of Cores to Use', reports the maximum number of CPU cores that can be used by the program, and by setting a number greater than one, allows for multi-threaded performance to increase the speed of the program.

Tuning the parameters for the allowable differences in the mean and SD changed how many valid sets were reported. The parameters were tuned through multiple runs so that after around forty-five minutes, six or fewer sets of improvisations would be reported.[5] From these sets, the one with the best fit to the overall mean and SD of recording year and tempo was selected. The final tuning variables were:

- BPM Mean Difference: ± 0.02
- BPM Standard Deviation Difference: ± 0.03
- Year Mean Difference: ± 0.03
- Year Standard Deviation Difference: ± 0.02

Running the program with these tuned variables returned three sets of thirty songs that were all equally close to the overall distribution. From these one was selected at random to form the initial selection of improvisations.

Following close inspection of the selected improvisations, three improvisations could not be accurately transcribed due to Green frequently playing double-stops. Investigation of the remaining twenty-seven songs also indicated that there was an

[4]Limitations of calculating the mean and SD of the year were noted, and adjustments subsequently made; however, it was a useful metric for initial selection criteria.

[5]These numbers refer to a PC running Windows 10 on an Intel Core-i7 6700K, 32GB RAM, using four cores.

over-representation of songs from 1964. Consequently, one song from that year was randomly removed. To replace these four songs, the remaining twenty-six were hard-coded as 'selected songs' within the random selection program, and the program was instructed to choose only four new songs to complete a new set of thirty.[6] For the additional selection the tuning variables were relaxed slightly. The results of the mean and SD for the tempo and recording year of the selected songs compared to all 147 is shown in Table 3.1. The final selection closely matched the All Songs distribution, indicating that the selected songs were representative of Green's output between 1960 and 1965, based on the tempo and recording year.

Table 3.1: Mean and standard deviation for the year and tempo for all 147 vs. the selected songs.

	Tempo	Year
All Songs		
Mean	147.32	1962.15
SD	63.02	1.36
Selection		
Mean	147.83	1962.10
SD	63.11	1.40

The improvisations within each of the thirty selected tracks were extracted using *Audacity* (Audacity Team 2015). Within ten of the thirty tracks Green improvised twice and, following the example of the Jazzomat Research Project, each improvisation was transcribed separately and labelled as improvisation one or two in order of their appearance.[7] In total, forty improvisations were selected to be transcribed to form the corpus of this research.

Issues with Data Selection

There were two main issues related to the data selection procedure. The first of theses issues was known prior to the data selection procedure, the second became apparent throughout the analyses. The first concerned selecting songs that could accurately be transcribed and imported into the *MeloSpy* system. Improvisations that contained frequent polyphony could not be used within *MeloSpy*. To limit bias

[6]The need to hard code values for subsequent selections limited the ability of the program to be easily distributed and used for other projects. The program needs to be re-written as an *R* function with options to easily set pre-determined indices.

[7]Trading improvisations or solo breaks within a head were not included.

in the data selection no songs from Green's corpus were removed for any reason, including frequent polyphony, prior to the first selection. Following the initial selection, closer listening to the selected tracks occurred, with only songs that had a high number of musical events that could not be accurately transcribed excluded.

The second concerned the distribution of independent features within the corpus. Although the forty improvisations represented a substantial proportion of Green's output as a leader between 1960 and 1965, they could not fully represent all facets of his improvisational style. This was most apparent in the distribution of independent features including time signatures, tonality types, and tempo ranges. For example, when considering the time signature, thirty-seven of the forty transcriptions were played in $\frac{4}{4}$ with only three in $\frac{3}{4}$. This imbalance in the data meant that it was difficult to compare features in different time signatures, as there was not enough $\frac{3}{4}$ data to determine if an effect was due to the time signature, or simply an element of the three improvisations. It would have been possible to record more features for each track before selection and try to balance these additional features. However, this had the potential to increase the bias through the selection or omission of features. Additionally, selecting improvisations to have balanced datasets for specific features could result in their over representation. For example, in a survey of the chordal structure of 227 jazz standards undertaken in the author's undergraduate studies, only 43 (18.50%) of the tunes analysed could be broadly considered to be in a minor key, with 16 (38.10%) of those oscillating between the relative major and minor keys. Similarly, many jazz standards are played in $\frac{4}{4}$. Consequently, creating a more balanced dataset of these features would result in the dataset not being representative of Green's output. The solution to this issue would be to increase the number of improvisations transcribed, which could then better represent the minority classes.

Selected Tracks

Information for all thirty tracks, and forty improvisations, are listed in Table 3.2. For the ten tracks that contained two solos, the second is listed with only its average tempo and timestamp.[8] The reference in the final column references to the Additional Online Sources section of the List of Sources, with links to a YouTube listing of each track.

[8] The tempos here and those in the csv used to select the improvisations vary slightly, with the data in Table 3.2 based on the transcribed solos.

Table 3.2: Title, Album, Tempo, and Solo times for all improvisations that were transcribed to form the corpus used within the study. Tracks that contained two solos have the second solo listed under the first, with separate tempo and timestamps listed.

Title	Album	Year	Tempo (BPM)	Solo Start	Solo End	Reference
At Long Last Love	I Want To Hold Your Hand	1965	112	01.05	03.26	(Green 1965a)
Blues In Maude's Flat	Grant Stand	1961	119	07.46	13.49	(Green 1961b)
Born To Be Blue	Ballads	1961	106	02.55	03.32	(Green 1961c)
Brazil	The Latin Bit	1962	220	01.15	02.14	(Green 1962b)
			220	03.31	04.30	(Green 1962b)
Freedom March	Sunday Mornin'	1961	142	00.40	03.04	(Green 1961d)
			141	06.28	07.49	(Green 1961d)
Go Down Moses	Feelin' The Spirit	1962	242	00.52	02.52	(Green 1962d)
			239	04.29	06.06	(Green 1962d)
Green With Envy	Green Street	1961	218	01.24	04.56	(Green 1961j)
			231	07.12	08.18	(Green 1961j)
Have You Ever Had The Blues	Blues For Lou	1963	125	01.00	02.02	(Green 1963c)
Idle Moments	Idle Moments	1963	69	01.54	05.40	(Green 1963e)
I'll Remember April	Remembering	1961	195	00.56	03.56	(Green 1961k)
I'm An Old Cowhand	Talkin' About	1964	150	00.52	03.27	(Green 1964b)
I Wish You Love	Street of Dreams	1964	186	02.35	05.21	(Green 1964a)
Little Girl Blue	Ballads	1961	108	02.54	04.48	(Green 1961l)
Minor League	Solid	1964	225	02.02	03.18	(Green 1964d)
			231	06.02	06.26	(Green 1964d)

Miss Ann's Tempo	Grant's First Stand	1961	248	00.23	02.09	(Green 1961m)
			257	03.50	04.57	(Green 1961m)
Moon River	Gooden's Corner	1961	130	01.16	02.38	(Green 1961n)
Nancy (With The Laughing Face)	Nigeria	1962	122	02.29	03.32	(Green 1962f)
Oleo	Oleo	1962	253	00.32	02.03	(Green 1962i)
			257	04.28	04.57	(Green 1962i)
Our Miss Brooks	Reaching Out	1961	88	03.03	04.15	(Green 1961o)
Red River Valley	Goin' West	1962	198	00.47	02.46	(Green 1962j)
Round About Midnight	Green Street	1961	120	02.03	06.24	(Green 1961r)
Seepin'	First Session	1960	125	00.00	06.11	(Green 1960b)
			116	10.03	11.43	(Green 1960b)
Sonnymoon For Two	First Session	1960	198	00.29	04.23	(Green 1960c)
Stella By Starlight	I Want To Hold Your Hand	1965	112	02.32	03.40	(Green 1965d)
Sunday Mornin'	Sunday Mornin'	1961	191	00.32	01.34	(Green 1961s)
			193	02.32	03.03	(Green 1961s)
Take These Chains From My Heart	Am I Blue	1963	109	01.26	02.39	(Green 1963f)
The Song Is You	Nigeria	1962	241	01.10	03.22	(Green 1962l)
The Surrey With The Fringe On Top	Blues For Lou	1963	239	00.38	01.53	(Green 1963g)
			244	03.40	04.15	(Green 1963g)
Tico-Tico	Tha Latin Bit	1962	130	01.34	06.27	(Green 1962m)
Wives And Lovers	Matador	1964	151	01.34	04.30	(Green 1964h)

3.1.2 Transcription

Following the selection of the improvisations to form the corpus, each of the forty solos were transcribed. The process for transcribing the improvisations was informed by the Jazzomat Research Project's guide for annotating solos, but adapted to suit the needs of this research.[9]

Preparing the Files

Each improvisation was extracted from the original audio file, beginning just prior to the first note of the solo, regardless of its position within the form, and ending after the offset of the final note. Each file was split into its left and right stereo components as most had Green louder in one channel with the backing instruments, especially the drum, louder in the other.[10] Splitting the audio files made it easier to isolate and transcribe Green's improvisations, while isolating the drum track improved the accuracy of the beat transcriptions. One of the forty solo tracks, *Our Miss Brooks* (Green 1961o), required the tuning to be adjusted, +40 cents, to match the 440Hz tuning standard.

With the files prepared, each soloist track was run through an automatic transcription process. This acted as a first pass and template for the manual transcription, reducing the time required to complete each transcription. This was completed using *Songs2See* (Dittmar, Cano, and Grollmisch 2017), which took in an audio file and output a MIDI file of the automated transcription. The transcription process then moved to *Sonic Visualiser* (SV), where the main transcription and annotation of the improvisations was completed (Cannam, Landone, and Sandler 2010). The general approach and steps undertaken to complete the transcriptions is explained below. A more detailed discussion, including settings for plugins and view options can be found in Appendix C. Figure 3.2 shows a screenshot from a completed SV transcription file. The screenshot has been edited and condensed slightly for clarity.

Transcribing the Solos

Each transcription was completed in three passes, with most of the work occurring in the second pass. The first pass was listening through the automated transcription alongside the actual improvisation to evaluate the overall accuracy. The track was then slowed down to 12.5% of normal speed, and the zoom in SV was increased to

[9] https://jazzomat.hfm-weimar.de/tutorials/sv/sv_tutorial.html

[10] Of the thirty tracks, only three had identical left and right channels, with these converted into a single mono file.

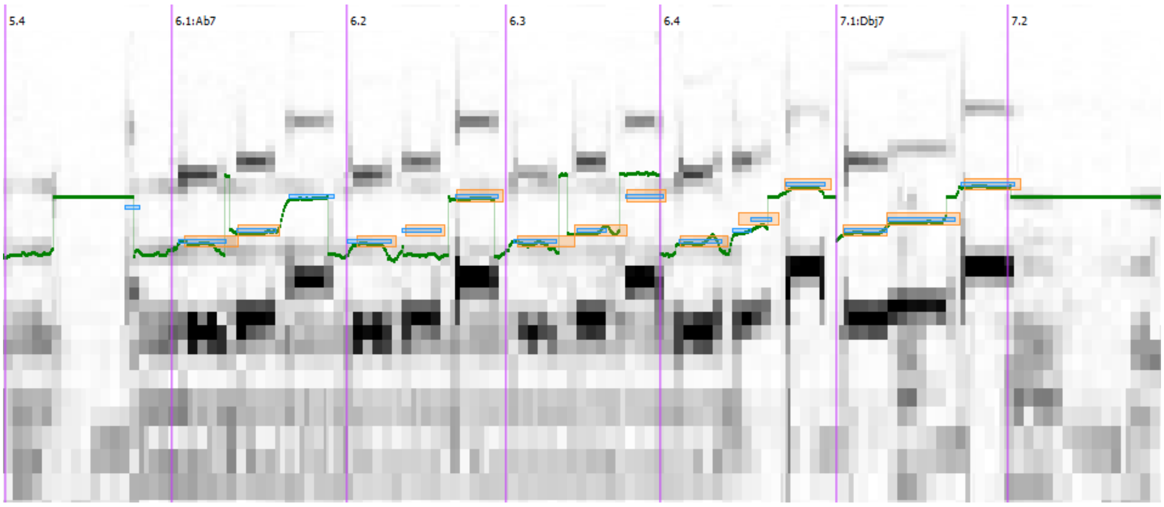

Figure 3.2: Screenshot from a completed transcription in *Sonic Visualiser*.

the maximum. Using the automated transcription and visual plugins as a guide and checking the pitch of notes against a piano or guitar, each note was transcribed into a new note layer using the 'draw' tool.

Each improvisation was transcribed in small batches of notes. Once the notes were transcribed at the slowest tempo and highest zoom, the speed was gradually increased to check (audibly and visually) each section of the improvisation over repeated listens, with any errors rectified. After one batch of notes was deemed to be accurately transcribed, the next section would undergo the same process. Once the whole improvisation was transcribed, the entire track would be played back at increasing speeds, starting at 12.5%, to check for accuracy and consistency across the whole transcription, with errors again rectified.[11] On average, it took over one hour to complete one minute of transcription. The final pass occurred after completing the annotation of the transcription (below). This was to allow time to pass between the completion of the transcription and the final check, creating a chance for a 'fresh' pair of ears to do the final check. Although this process was time consuming, the resulting transcriptions were highly precise.

Between the second and third pass of the transcription, the beat track was transcribed and annotated. The beat track acted as the metrical foundation of the transcription against which the transcribed notes were compared. Additionally, the time signature, form structure, metrical structure (bars and beat labels), and chords were annotated on the beat track layer. The initial transcription of the beats was completed in a separate SV file using the backing channel of the split audio files. The use of the accompaniment track for beat transcription improved the performance of the automatic beat tracker plugin. Additionally, the spectrogram

[11]The faster speeds were also played with the track zoomed further out, so that the notes did not pass by too quickly. All changes to the transcription were made with the highest zoom available.

clearly showed where the ride and hi-hat cymbals were played. These two cymbals were used to transcribe the true beat placements of the song, with these drawn onto a new beat layer.[12] A similar procedure of transcribing sections of the improvisation, checking the accuracy, and rectifying any errors, was undertaken. Since only one dimension (time) was being transcribed, in comparison to the two dimensions for notes (pitch and time), the beat transcription was able to be completed at a higher speed, and in larger sections. Before the beat track was annotated, it was imported into the original note transcription file, with a final check of the beat track played along with the completed note transcription.

Annotating the Solos

The annotation of the transcriptions stored data about the structure of the improvisations. The transcription of the notes and beats provided the content of the improvisations, with the annotations adding context to the data. The beat track data was exported from SV as a CSV file and annotated in *Microsoft Excel*. The annotations followed the form of: [time of beat],[bar].[beat]:[form-label]-[chord]-[time signature]. The time signature was only annotated on the first beat, or when there was a change in time signature, and form and chord annotations were only included when they changed. The four most common types of annotation looked like:

- First beat: 0.102312925,1.1:A1-Bbm7-4/4
- Form change: 16.73650794,17.1:Bbm7-A2
- Chord change: 2.187029478,3.1:Eb7
- Other beats: 0.363537415,1.2

The form labels and chord annotations were based on real book charts of each of the songs, from both physical real books and the *iReal Pro* app (Biolcati 2017). Issues and limitations of this approach are discussed in the Issues with Transcription section. The only exception was for *Oleo* (Green 1962h), which featured a reharmonisation. These chord changes were taken from the transcription in the book *Best of Grant Green: A Step-By-Step Breakdown of the Guitar Styles and Techniques of the Jazz Groove Master (Guitar Signature Licks)* (Marshall 2004). The completed annotations were then re-exported as a csv and imported back into the SV transcription file.

The final step was to annotate the phrases within each improvisation. Phrase annotations are influenced by the musical sensibilities of the transcriber, as no strict definitions of what constitutes a phrase exist. Hal Crook, in his book *How To*

[12]In the instances where there were not clear indications of a ride hit, the middle distance between two hi-hats was recorded as the beat.

Improvise, defined a phrase as: “a period of continuous, but not necessarily constant, melodic/rhythmic activity, which can vary in length from one beat to several measures depending on tempo” (1991, 26). This definition highlights the ambiguity around what a phrase is, providing a description but not setting any criteria. The beat track transcription and annotations added between thirty minutes and two hours to the transcription time, depending on the length and complexity of the track.

With the annotations completed, a final pass was undertaken with each transcription listened to again twice, once with a speed between 50% and 80% and again at 100% to check for any final errors in the note transcription, beat transcription, and phrase annotations. All these steps were repeated for each of the forty improvisations. Once a transcription file was considered complete, it was duplicated and had excess layers removed.[13] These clean SV files were used to create the final database. The final SV transcription files[14], the automatically generated symbolic notation version of the transcriptions, and the SQLite3 database of Green’s corpus can be found in Appendix E.1.

Issues with Transcription

Throughout the transcription process issues related to the transcription of the improvisations occurred. Although most of these issues were known beforehand, they still needed to be considered and dealt with. The issues predominantly contended with finding a balance between the transcription and representation of the improvisational data, and its use in the research. The five main issues that arose throughout the transcription process were:

1. Polyphony;
2. Guitar techniques (e.g. slides, tremolo, treatment of appoggiatura);
3. Beat track;
4. Double time feel;
5. Chord annotation.

The first two related more specifically to the guitar or similar instruments while the other three concerned issues related more broadly to jazz transcriptions.

An issue with the transcription process developed by the Jazzomat Research Project is that it is incapable of transcribing any polyphony. For the guitar, this is largely chord stabs or double-stops. Although this is a major limitation, and any update

[13]The excess layers included those used to aid in the transcription, including the visual plugins, and the automated beat and note transcriptions.

[14]With the audio files removed due to copyright.

that helped to overcome this issue would allow for a wider range of instruments and improvisers to be transcribed, it is an understandable limitation due to the structure of the data. This limitation related to how to treat features, such as measuring the interval distance between notes, when there are multiple notes. Green was partly selected due to his predominantly single note line improvisations. When he did occasionally play double-stops, the highest pitched note was considered the melody and transcribed.[15]

The second issue related to what was termed 'articulation techniques' by the Jazzomat Research Project. The available techniques for annotation were: bend; shake; vibrato; slide; fall-off; dead-note; dirty/split tone; and top tones (Jazzomat Research Project 2017). Other guitar techniques, including hammer-ons and pull-offs, were not valid annotations in the *MeloSpy* system. As this section related to instrument specific techniques, the issue of tremolo is also discussed. The techniques that required the most consideration were bends and tremolo picking, both of which occurred infrequently in Green's improvisations. Slides, which Green played more frequently, were transcribed with each single note transcribed individually. Due to the infrequency of available articulation techniques that could be annotated in Green's improvisations, none were annotated in the transcriptions. Bends, Green's rarest articulation events, were also transcribed as individual note events. For example, a bend with only two clearly defined notes would be transcribed as two notes, with the onset of the second note set as the point where that pitch became constant.

The final articulation technique considered was tremolo picking. There were two possible approaches to transcribing this technique. The first would be to transcribe the tremolo notes exactly as played, with each individual pick transcribed as its own note event. The advantage of this approach would be that it exactly captures what was played, and could therefore be argued to be the most accurate representation. The disadvantages were that it would not be musically meaningful to treat each individual pick as its own note, with the aural and musical effect being that of a single note articulated many times. The other disadvantage of this approach was that it would inflate the number of notes per bar and the number of repeated intervals in any bar where a tremolo note was played. As a result, those bars and note events would need to be removed from the dataset entirely for all analyses. The other approach to transcribing tremolo notes was to follow symbolic notation, where

[15] In Green's dataset, double-stops were always associated with periods of tremolo picking. These occurred at: *Blue's In Maude's Flat*, bars 139–144 and 150–152 (Green 1961b); *Tico-Tico*, bars 89–93 (Green 1962m); *Idle Moments* bars 15-17, 21, 38, and 63 (Green 1963e). There was one additional section of tremolo picking not related to polyphony, in bars 40–41 of *At Long Last Love* (Green 1965a).

each sequence of tremolo notes were transcribed as one long note, with a single onset and offset. The advantage of this approach was that it reflected the intent and resulting effect of the articulation technique, while also allowing the note to be included in analyses. The disadvantages were that it meant the resulting transcription was no longer entirely descriptive, and due to the lack of support of custom technique annotations, it could not be represented in the final transcriptions and analyses. The other disadvantage was that instead of including many short repeated notes, there were now additional extra long notes in the corpus. Since the occurrence of tremolo in Green's corpus was low, this approach had a smaller impact on the resulting analyses, with very long outliers easy to identify and exclude when necessary. For this research, the second approach was selected, with each tremolo section transcribed as a single note.

The beat track for each improvisation was transcribed with the aid of both automatic beat recognition plugins in SV and the spectrogram of the audio channel where the drummer was most prominent. This method differed substantially from that described by the Jazzomat Research Project[16], which suggested to tap the beats along with the recording and then adjust them afterwards. If the beats needed to be adjusted, it was more efficient to skip the manual tapping of the beats and to use plugins instead. The issue related to what should be considered the true location of the beat, to which the automatic beat transcription would be adjusted to match. The clearest solution, and the one selected, was to base the true beat location on the drummer. Traditionally, drummers play the ride cymbal on every beat, with a hi-hat on beats two and four (in a $\frac{4}{4}$ swing feel), providing a consistent set of beats to transcribe. These two cymbals showed up clearly on the spectrogram, seen in Figure 3.3, which shows two screenshots of the beat track transcription from SV. The left shows the spectrogram without the beat transcription, with the red lines on the right the transcribed beat locations.[17] The first visible beat is a hi-hat on beat four followed by a ride cymbal on the first beat of the following bar. In situations where cymbal pulses were obscured or missing, the beat locations were interpolated as mid-points between transcribed beats. These were then checked and adjusted as necessary, along with all the other beat transcriptions, through the checking procedures described previously.

The final two issues, double time feel and the choice of chord symbols, are considered in all types of jazz transcriptions. Double time is a musical device that is most commonly employed by soloists during their improvisations. It "is a special effect which [sic] occurs when one or more players make the tempo sound twice as

[16]https://jazzomat.hfm-weimar.de/tutorials/sv/sv_tutorial.html
[17]The transcribed beat lines were edited to display more clearly on the page.

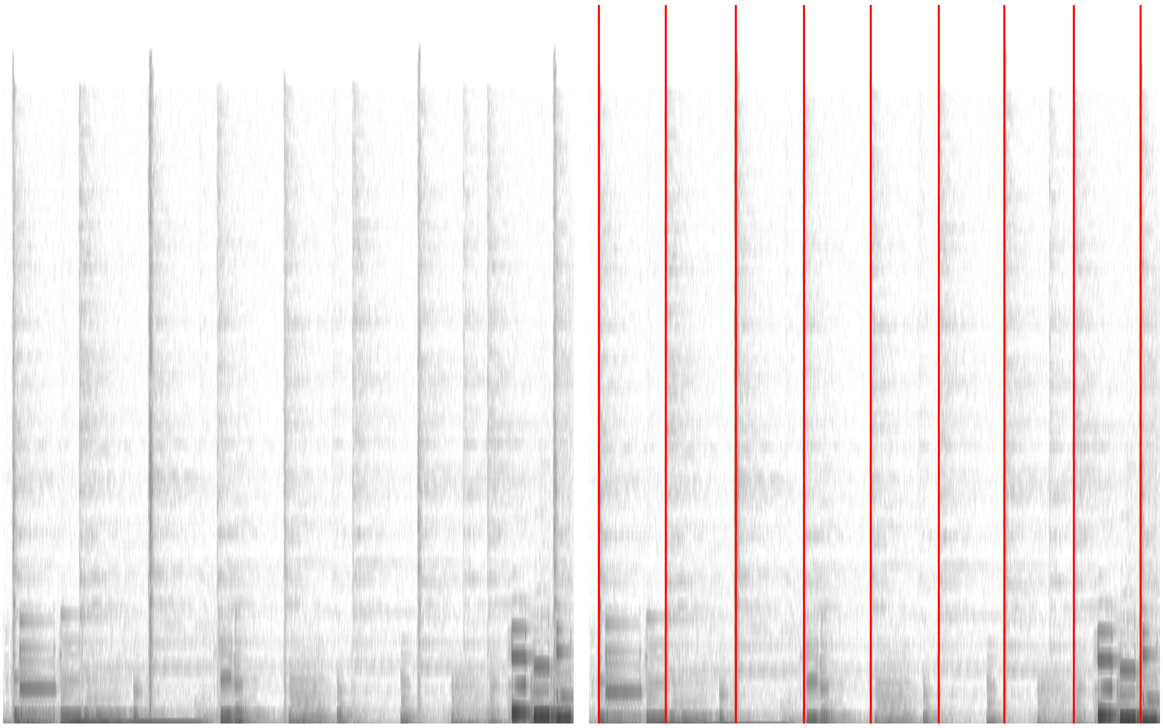

Figure 3.3: Beat transcription in *Sonic Visualiser*. Left: Plain spectrogram. Right: Beat transcription over spectrogram.

fast as the original tempo … [most often used] at slow to medium tempos" (Crook 1991, 140). During a double time section, while the aural effect is that of playing twice as fast as the original tempo, "the chords continue to change at the original tempo" (Crook 1991, 140). Often it is only the soloist who goes into a double time feel, but "at times the accompanying players will double the tempo with the soloist and the whole band will play the new tempo" (Crook 1991, 140). In situations where only Green played in a double time feel, it was transcribed as it sounded, with Green playing twice as fast compared to the surrounding harmonic progression. The issue that impacted the transcription process was how to deal with situations where the entire band played in a double time feel for the entirety of the improvisation.[18] There were three options for how to transcribe the double time feel:

1. Transcribe the solo in relation to the underlying chord movement (as played in the head, number of bars per chorus stays the same);
2. Transcribe the solo in relation to the played metrical structure (number of bars per chorus doubles, with each bar split into two);
3. Transcribe the solo with a new time signature, doubling the numerator of the original time signatures, e.g. $\frac{4}{4}$ to $\frac{8}{4}$ (number of bars per chorus stays the same, number of beats per bar doubles).

[18]In the corpus of selected Green improvisations, only solos in $\frac{4}{4}$ were played entirely in double time. As a result, this discussion focuses on quadruple time, but the same principle could be applied to other time signatures.

The first option did not reflect what was played by the musicians, while also artificially inflating the number of notes per beat and per bar. The second option did not have these limitations, but did not reflect the underlying form of the song. The final option, and the one selected for this research, did not artificially inflate the number of notes per beat and reflected the underlying form of the song. As with the second option, it did not alter the rhythmic structures played (a quaver-equivalent note was still transcribed the same), the third option also kept the tempo and chord movement the same as the head. However, it did inflate the number of beats per bar, and therefore the number of notes per bar.

There were no improvisations either in the WJazzD or in Green's corpus that were in $\frac{8}{4}$. Therefore, transcribing the double time $\frac{4}{4}$ improvisations as $\frac{8}{4}$ allowed for easy identification of them throughout the analyses. With the double time transcriptions easily identifiable, functions were written to deal with some of the disadvantages of this approach. Specifically, two functions were written that transformed the transcriptions in $\frac{8}{4}$ into the equivalent of option two listed above, in $\frac{4}{4}$ with double the number of bars.[19] This provided an approach that could consider the transcription as either of option two or three as presented above.

The final issue related to transcription is one of the most persistent throughout all jazz transcriptions, the selection of the chord changes. The issue was that there rarely exists "one definite set of chords for a piece" (Potter 1989, 40), with variations existing not only between performances, but also between choruses, forms, and even musicians. The chord changes on which an improvisation are based act as a template, instead of a set of instructions that must be followed. Even in a small ensemble with only four performers (e.g. soloist, bassist, drummer, and a comping instrument – piano or guitar), it is likely that there are three subtly different chord progressions played. This issue can best be highlighted through an example of one of the most common chord progressions, the blues, shown in Figure 3.4. The first three examples are all fairly common changes to a B♭ blues, with a combination of all three likely used in a single solo. The fourth example presents a sequence of chord changes, based on the previous chord sequences, but with extra chord alteration and substitution.[20] Although the entire progression is unlikely to be used, individual alterations and substitutions could be played.

[19]The code for the function can be found in Appendix B, code blocks B.1 and B.2. The code presented throughout this document has been cleaned up for clarity and from improvements in the author's own coding. No functional differences, unless specified, exist between the code used in the research, and that presented in this document.

[20]All chords are based on basic substitutions, including: diatonic substitution; addition or removal of ii–V; tritone substitution; secondary dominants; and diminished substitutions.

```
a) Basic blues
   I
|| Bb7     |    %    |    %    |    %    |
   IV                  I
 | Eb7     |    %    | Bb7     |    %    |
   ii        V         I         V
 | Cm7     | F7      | Bb7     | F7      ||

b) Common jazz blues
   I         IV        I         ii - V/IV
|| Bb7     | Eb7     | Bb7     | Fm7 Bb7 |
   IV                  I         ii - V/ii
 | Eb7     |    %    | Bb7     | Dø7 G7b9|
   ii        V         I    VI   ii   V
 | Cm7     | F7      | Bb7  G7 | Cm7  F7 ||

c) Common blues alternative
   I         IV        I         ii - V/IV
|| Bb7     | Eb7     | Bb7     | Fm7 Bb7 |
   IV        #IV       I   Desc. chromatic 7 chords leading to V/ii
 | Eb7     | Eo7     | Bb7  A7 | Ab7  G7 |
   ii        V         I    VI   ii   V
 | Cm7     | F7      | Bb7  G7 | Cm7  F7 ||

d) Blues with alterations
|| Bb7 Bo7 | Em7  A7 | Dm7 Gm7 | Fm7   E7  |
 | Eb7  A7 | Ebm7 Ab7| Bb7  A7 | Dø7 Db7#11|
 | Cm7 F#7 | F7   B7 | Bb7 Db7 | C7    B7  ||
```

Figure 3.4: Example of four possible chord progressions over a blues.

The main issue was selecting which set of chord changes should be annotated for each transcription and what impact that choice could have on the resulting analysis. A common approach is to use a set of “idealised chord changes” (Heyer 2011, 38) to analyse improvisations against. Idealised chord changes can be thought of as a “recording’s hypothetical harmonic lead sheet ... [with the] changes ... rarely played exactly as written, but ... provide [a set of changes] on which the musicians likely based their performance” (Heyer 2011, 38).

The idealised chord changes were drawn from two main sources: real book charts (Hal Leonard Corporation 2004a,b, 2006; Sher and Bauer 1988, 1991; Sher, Evergreen, and Dunlap 1995); and the *iReal Pro* app (Biolcati 2017). When multiple versions were available, the author used their own experience to select or combine the chord changes into a single set. The author’s own experience was used to select chords for common changes such as the blues or rhythm changes. Idealised chord changes presented a simplified version of the chords played by the rhythm section and those conceived of or outlined by an improviser. For the purposes of this research, this had three interrelated impacts on the annotation of the chords:

1. Simplification of chord types;
2. Simplification of chord extensions;
3. Harmonic variation and substitution.

Nine chord types were initially annotated: major triad; major 7 ($\Delta 7$); minor triad; minor 6; minor 7 (m7); minor$^{\Delta 7}$; dominant 7 (7); half-diminished 7 (ø7); and diminished 7 ($\circ 7$). Table 3.3 shows the number of bars (with at least one note event) associated with each chord type. This data showed that many of these nine types appeared infrequently throughout Green's corpus. Consequently, and to ensure there was enough data in each chord type class,the nine chord types were simplified into five: $\Delta 7$; 7; m7; ø7; and $\circ 7$.[21]

Table 3.3: Summary of chord type simplifications, showing the number of bars, with at least one note event, associated with each chord type.

	Simplified Chords				
Raw Chords	$\Delta 7$	7	m7	ø7	$\circ 7$
$\Delta 7$	460	-	-	-	-
Maj	72	-	-	-	-
7	-	1858	-	-	-
m7	-	-	1014	-	-
min	-	-	149	-	-
m6	-	-	11	-	-
m$^{\Delta 7}$	-	-	9	-	-
ø7	-	-	-	174	-
$\circ 7$	-	-	-	-	140
Total	532	1858	1183	174	140

Higher extensions, altered notes, or alternative voicings for chords are often written on lead sheets; however, it is accepted that the "rhythm section will (and should) ... employ [their own] extensions" (Heyer 2011, 38). Extensions do not generally change the function of a chord. One exception to this is a 7♭9 chord as part of a minor ii–V progression (ø7 – 7♭9), where the 7♭9 is used in place of a standard 7 chord for

[21] Issues related to the combination of the tonic minor into the m7 class, where the functions of the two chords were different, were noted. With the current dataset it was decided that this simplification was necessary to ensure a large enough sample size. All changes to the data post-transcription were non-destructive, allowing the raw chord data to exist alongside the simplified data.

smoother voice leading. It was decided to use main function of the chord as the basis of the transcription, resulting in all extensions and alterations being discarded.[22]

Along with extensions and altered notes, harmonic variation and substitution is used frequently throughout a song to vary the idealised chord changes.[23] Common harmonic substitutions include: diatonic substitutions; the addition or removal of ii–V; tritone substitution; secondary dominants; and diminished substitution. The two most pertinent substitutions that impacted the annotation of the idealised chord changes were the addition and removal of ii–V and the use of tritone substitutions.[24]

The use of add/remove ii–V harmonic substitutions provides for a wide variety of possible chord progressions.[25] This variety presented a serious issue regarding which set of changes should be annotated. The solution used was to take the chord changes from the real books or app as a starting point and then, using domain knowledge of ii–V progressions, alter the changes to reflect a common idealised set of changes that were neither simplistic or overly complex.

A tritone substitution is the substitution of a 7 chord with a 7 chord a tritone away; for example, substituting a G7 chord with a D♭7 chord. The issue with tritone substitutions is that although they are frequently used by both accompanists and improvisers, they are rarely written in chord charts. Even close inspection of an improvised line over a 7 chord does not always confirm whether or not a tritone substitution was being thought of by the soloist. The solution to this was that unless the tritone substitution was a clear part of the original composition, appearing in multiple sets of chord charts, a non-substituted 7 chord was annotated.[26]

The aim of the transcription process was to generate high quality, precise, and descriptive transcriptions while limiting issues that occur in both traditional and computer-aided transcription practices. The steps outlined here (and expanded upon in Appendix C) presented how the dataset for this study was generated.

[22] 7♭9 chords as part of a minor ii–V could be identified by searching for 7 chords preceded by a ø7 chord.

[23] This is distinct from re-harmonisation, where an entirely new set of changes is written for a song, which then become the idealised chord changes.

[24] Although the other substitutions were important, their impact could largely be contained to the discussion of ‘performers playing differing sets of changes’.

[25] The add/remove ii–V substitution allows for splitting a V7 into one, or multiple, ii–V progressions, or combining a ii-V into a single V7 chord.

[26] Attempts were made to write a function to identify whether or not a specific line over a 7 chord was based on the annotated chord or that of the tritone substitution. The results of this function did not perform well enough to be included in this research.

3.1.3 Data Extraction

Following the completion of the transcriptions, the SV files needed to be converted to an SQLite3 database that could be read by *MeloSpy*. It was this database that *MeloSpy* used to extract the features from the transcriptions, which formed the dataset used throughout this research. The documentation from the Jazzomat Research Project regarding the metadata and configuration files required by the `MelConvert` function of *MeloSpy* to convert the SV files was insufficient to successfully generate a database. To overcome this issue, the author reached out to the Jazzomat Research Project through their support email with an initial technical support enquiry prior to beginning this research in February of 2017, and a following enquiry in January of 2020.

The resulting *MeloSpy* SQLite3 database had twelve tables, four of which contained information critical to the analysis. The tables were (critical tables in bold): **beats**; composition_info; db_info; esac_info; **melody**; melody_type; popsong_info; record_info; **sections**; **solo_info**; track_info; and transcription_info. Although the conversion process was generally successful, with only minor changes made to the SV files for them to compile correctly, most of the metadata was not successfully imported. Of the tables listed above, only the solo_info table necessitated the addition of substantial metadata.[27] These edits were made using a trial version of *SQLite Expert Professional* (Coral Creek Software 2017).

The final step was extracting the improvisational features from the database. Data was extracted from both Green's database and the WJazzD. *MeloSpy* uses functions written in YAML called Feature Definition Files (FDFs) to process the improvisational data stored in the database and extract features used for analysis.[28] FDFs rely on basic transformations generated by the Jazzomat Research Project's python library MeloSpyLib. At the time of writing, this library has not been released; consequently, the exact workings of the MeloSpyLib are unknown. From the raw improvisational data the MeloSpyLib generates features including: chordal pitch class; tonal pitch class; interval classes; duration classes; and inter-onset interval classes. The FDFs used these as inputs to process the data and output the results to a specified file format, typically a CSV file.

Although custom FDFs could be written by end users, limitations in the inputs and transformation processes, and a lack of detail regarding the MeloSpyLib, meant this option was not explored. Instead, the raw improvisational features, including basic

[27]The tables composition_info, record_info, track_info, and transcription_info were all updated with correct metadata. These had no impact on the extraction of the transcription data or the analysis, and were edited for completeness of the database.

[28]More information regarding FDFs can be found on https://jazzomat.hfm-weimar.de/commandline_tools/melfeature/melfeature_features.html and accompanying pages.

transformations (e.g. tonal pitch class, chordal pitch class, intervals), were extracted from the database, with all further data manipulation and analysis undertaken in *R*.[29] Beat track and phrase data were extracted and subsequently combined with the raw data within *R*. Additionally, the beats table from the SQLite3 databases were extracted using *SQLite Expert Professional* as a CSV file.[30] The beats table data, in contrast with the beat information extracted with *MeloSpy*, contained all the beat track information (including for beats where no note onset events occurred). This data was combined with the raw data, as well as used in functions for manipulating and generating new features.

3.1.4 Treatment of Data

The data extracted from the transcriptions focused on fundamental improvisational features. As these did not provide all the features required to undertake the analysis of Green's improvisational style, custom functions were written in *R* to manipulate the data and generate new features. The list of functions used to manipulate and generate the new features can be found in Appendix E.2, Functions.R. Additionally, the list of *R* libraries used throughout the document can be found in Libraries.R, and the main data preparation files can be found in Green_DataPreparation.R and WJD_DataPreparation.R. The results of these last two files formed the datasets used in the analysis of Green and the performer classification and comparative analysis.[31] The new custom functions could be broadly split into four groups[32]:

1. functions for marking specific notes or aspects of the transcription;
2. functions that expanded upon features by *MeloSpy*;
3. functions that re-calculated features provided by *MeloSpy*;
4. functions that created new features.

This discussion focused on five functions to serve as an example of those written for this research.[33]

[29]The files (a .bat and .YAML file) used to extract the features from the databases used in this research, as well as the extracted CSV files, can be found in Appendix E.2.

[30]The transcription_info table for the WJazzD database was also extracted to clean the data for the performer classification and comparative analysis task.

[31]The resulting *R* datasets can be found as .RDS files in Appendix E.6.

[32]There were also quality of life functions written for automating data exploration, plotting of figures, or automatic generation of excerpts of symbolic notation for inclusion in this document.

[33]The functions are provided 'as is'. Many should be immediately transferable to other analyses; however, not all will be. They were written to fulfil the requirements for this research, and consequently dealt specifically with elements of Green's transcription data. The limited functions could be expanded to have broader use, but that was outside the scope of this research.

The function `notePlacement` is an example of the first type, marking specific notes or aspects of a transcription. `notePlacement` used the difference from the nominal metrical onset to label a note as being played behind, on, or ahead of the beat (with a threshold for what was considered on the beat), shown in Code block 3.1. The default threshold (0.1% of the surrounding beat length) was selected based on evaluation of Green's SV transcription files.

```
notePlacement <- function(n, threshold = 0.001) {
  # n = difference from the nominal metrical onset
  if (n < -threshold){j <- "before"}
  else if(n <= threshold & n >= -threshold){j <- "on"}
  else if(n > threshold){j <- "after"}
  return(j)
}
```

Code 3.1: Code for labelling the placement of a note in relation to their nominal metrical position.

A function of the second type, that expanded upon features already present in *MeloSpy*, was `expandedTempoClass`.[34] The tempo class feature from *MeloSpy* provided only one level for all improvisations with a tempo $>$ 180 BPM. `expandedTempoClass`, shown in Code block 3.2, included two additional classes for higher tempos.

```
expandedTempoClass <- function(n) {
  if (n < 60){ j <- "Slow"  }
  else if (n >= 60 && n < 100){ j <- "Medium Slow" }
  else if (n >= 100 && n < 140){ j <- "Medium" }
  else if (n >= 140 && n < 180){ j <- "Medium Up" }
  else if (n >= 180 && n < 220){ j <- "Up" }
  else if (n >= 220 && n < 260){ j <- "Quick" }
  else if (n >= 260){ j <- "Fast" }

  return(j)
}
```

Code 3.2: Code for generating the expanded tempo class feature.

[34] The function was named `bpmTC` in Functions.R and Data_Preparation.R, and was combined with `tcFactor` to label the tempo classes.

An example function of the third type, that re-calculated features provided by *MeloSpy*, was `manualSwing`. The author was unable to to replicate the swing values calculated by *MeloSpy*. Therefore, new swing values were calculated based on the raw duration and IOI data, shown in Code block 3.3.[35]

```
manualSwing <- function(df) {
  # Marks swing beats as 1, non-swing beats as 0
  df <- as.data.frame(df %>%
            dplyr::rename() %>%
            group_by(id, bar, beat) %>%
            mutate(swing = ifelse((division==2 & sum(tatum)==3), 1,
              ifelse((division==3 & sum(tatum)==4), 1,
              ifelse(((division==4 & sum(tatum)==5 & min(tatum)!=2) |
                   (division==4 & length(division)==2 & sum(tatum==4))),
                   1,0)))))
  # Calculates the swing ratio for eligible binary notes based on
  # the ioi of the first note and duration of the second note
  df$swingRatio <- (df %>%
                     dplyr::rename() %>%
                     group_by(id, bar, beat, division) %>%
                     mutate(duration1 = dplyr::lead(duration),
                           swingRatio = ifelse((swing==1),
                                      (ioi_raw/duration1), NA )) %>%
                     pull(swingRatio))
  # Removes the swing beat marker
  df <- subset(df, select = -(swing))
  # Puts swing ratios between 0.98 and 3.02 into a new column named swing
  df <- as.data.frame(df %>%
                      dplyr::rename() %>%
                      mutate(swing = ifelse((swingRatio < 0.98 |
                                          swingRatio > 3.02),
                                          NA, swingRatio)))
  data.frame(df)
}
```

Code 3.3: Code for manually generating the swing ratio of swung note pairs.

Two examples are given for the final type of function, those which generated new features that were not provided for by *MeloSpy*. The first, `restCalc`, created an entirely new feature, the length of the rest between notes, based upon the raw data from the transcription files. `restCalc` returned two versions of the rest value, a raw value (seconds) and as a proportion of the surrounding beat length, shown in Code block 3.4.

[35] There was an error in this code that erroneously calculated swing for a few non-swing notes. The updated function, `manualSwingMarkerANDBUR` found in Code block B.6, should be used instead. The code presented here, along with a separate marking function `swingMarker`, created the data used in this research.

```r
restCalc <- function(df) {
  # creates the new columns for the data to be entered into
  df$restDur <- NA
  df$restProp <- NA

  for(i in 1:length(df$id)){
    # checks if last row in data frame
    if(i != length(df$id)){
      # checks if last row in song
      if(df$id[i]==df$id[i+1]){
        # calculates the duration (seconds)
        df$restDur[i] <- df$onset[i+1]-df$offset[i]
        # calculates the duration (proportion of surrounding beat)
        df$restProp[i] <- df$restDur[i]/df$beatLength[i]
      }
    }
  }
  return(df)
}
```

Code 3.4: Code for calculating the rest values between notes.

The second function created a new feature, inspired by features available from *MeloSpy*. *MeloSpy*'s metrical weight feature simplified a note's metrical position into one of three classes: played on a metrically strong beat; played on a metrically weak beat; played off-beat. This inspired new features that shared the weight name, including beat weight, and $\mathrm{CPC}_{\mathrm{Weight}}$. The example function shown in Code block 3.5 is `cpcWeight`, which simplified the chordal pitch features into one of three categories: arpeggio tone (2); scale tone (1); non-harmonic tone (0).[36] The default assumption was that a m7 chord was functioning as a iim7, therefore the $\mathrm{CPC}_{\mathrm{Weight}}$ values were based on the dorian mode. The code from lines 10–21 checked if the m7 chord was a vi of the current key and was preceded by the IΔ7. If it was, it was considered a vim7, with the $\mathrm{CPC}_{\mathrm{Weight}}$ values based on the aeolian mode.

These examples highlighted the breadth of new functions that were required to be written for this research. The aim of the functions was to create features that were both musicologically meaningful, with many having direct comparison to features investigated in standard close-reading analyses, while also being computationally and statistically useful. The functions written for this research transformed the raw, or lightly processed, improvisational data from the transcriptions into features and data that could be used to explore Green's improvisational style.

[36]This code required extra features generated from the custom `chordSetup` function, which can be found in Functions.R.

```
cpcWeight <- function(chord, cpc, ctc, ctcPrev, chordTypePrev) {
  if (is.na(ctc)) {
    k <- NA
  }
  else{
    x <- cpc
    n <- ctc-(ctcPrev%%12)
    n[is.na(n)] <- 0
    if (is.na(chord)) {  }
    # Checks if m7 chord is likely a vi minor
    else if (chord == "-7" & n == 9 &
             (chordTypePrev == "j7" & !is.na(chordTypePrev))) {
      ifelse(x %in% c(0,3,7,10), k <- 2,
             ifelse(x %in% c(2, 5, 8), k <- 1,
                    k <- 0))
    }
    else if (chord == "-7") {
      ifelse(x %in% c(0,3,7,10),k <- 2,
             ifelse(x %in% c(2, 5, 9), k <- 1,
                    k <- 0))
    }
    else if (chord == "7") {
      ifelse(x %in% c(0,4,7,10), k <- 2,
             ifelse(x %in% c(2,5,9), k <- 1,
                    k <- 0))
    }
    else if (chord == "j7") {
      ifelse(x %in% c(0,4,7,11), k <- 2,
             ifelse(x %in% c(2,5,9), k <- 1,
                    k <- 0))
    }
    else if (chord == "m7b5") {
      ifelse(x %in% c(0,3,6,10), k <- 2,
             ifelse(x %in% c(1,5,8), k <- 1,
                    k <- 0))
    }
    else if (chord == "o7") {
      ifelse(x %in% c(0,3,6,9), k <- 2,
             ifelse(x %in% c(2,5,8,11), k <- 1,
                    k <- 0))
    }
  }
  return(k)
}
```

Code 3.5: Code for generating the $\mathrm{CPC}_{\mathrm{Weight}}$ feature.

3.2 Analysis of Grant Green

3.2.1 Methodology

The overarching aim for this research was to develop a methodology for investigating performers' improvisational style employing computer-aided and statistical techniques. Presented here is the methodology that was developed throughout this research. The chapters in Part II used this methodology to explore Green's improvisational style. These are presented as an example of how the methodology could be used in practice. Specific methods and approaches relevant to the feature investigated are included within the discussion in the analysis chapters.

For the purposes of this research and methodology, all features were separated into one of four domains: pitch; rhythm; micro; and macro.[37] There was some overlap and frequent interactions between the domains; however, the separation aided in filtering the available features and grouped similar features together.

The two domains most prominent in prior research were the pitch and rhythm domains. The micro domain benefited the most from the precise transcription method developed by the Jazzomat Research Project. Due to the prior lack of many high quality transcriptions, features in the micro domain were found the least in the literature. The macro domain focused on broader structural or large scales features of an improvisation; for example, phrases or how the median pitch per bar changed over the course of a solo. Although features associated with the macro domain could be included in other domains, they were categorised as such due to their relationship with the larger feel of an improvisations, rather than the specifics of each relevant domain.

Due to the complexity of improvisation, there were not only many individual improvisational features, but also a multitude of interactions between these features and other aspects of music. Therefore, it was not possible within the scope of this project to fully analyse all potential aspects of a performer's improvisational style. A balance was required between analysing a broad range of improvisational features in substantial depth to gain insights into Green's improvisational style, while fitting within the scope of this research. The result was the development of the following top-down analytical methodology:

[37]Features that were solely independent variables, e.g. tempo, were not assigned to a domain. Features from other domains were also used as independent variables.

1. Assign all features to their main improvisational domain: pitch, rhythm, micro, and macro;
2. Within each domain assign features to a broad feature-category (e.g. for the pitch domain, the categories were raw pitch, tonal pitch class, chordal pitch class, and intervals);
3. Begin analysing the features within the broad feature-categories, again starting at an overview level (e.g. overall distributions) before investigating specific situations and interactions.

How deep an analysis into a specific feature went depended on three criteria:

1. Its overall importance in jazz pedagogy;
2. Its appearance and importance in previous studies;
3. The results of the initial analyses.

With this methodology, features that had previously been found to be important, and those that showed interesting or promising results at a broader level, were analysed in more detail, In contrast, those that did not were only analysed at a general, less specific, level. However, these still provided some insight into that feature's impact on Green's improvisational style. This approach allows for translation of this methodology from one performer to another. While certain features were analysed deeper for Green, another performer with their own improvisational style would have a different set of features that garner more analytical attention. This methodology also allows for projects of different sizes to still follow the same approach, as either the number of features investigated deeply or how deep an analysis into a feature goes, can be scaled as required. It is noted that due to this methodology, there were likely interesting facets of Green's improvisational style that were not investigated fully.

3.2.2 Methods

The methods of analysis used in this research included descriptive and exploratory statistics, and statistical hypothesis testing. The descriptive and exploratory statistics (e.g. graphs and measures of centre) were most similar to some standard analytical practices in jazz analyses. The main statistical hypothesis tests used throughout this research were: χ^2-tests (chi-squared test); ANOVA (analysis of variance); t-tests; correlation; and linear regression. The analyses within this research focused on univariate and bivariate statistics. All analyses were undertaken in the *R* Statistical Language, with packages from CRAN and github.

Effect sizes

For all statistical tests presented within this document, the relevant effect size was reported along with the statistical significance (p-value).[38] Effect size describes how large the effect between two features was. As the size of a dataset increases, or the number of features involved in an analysis increase, the more likely it is that a given statistical test will find a statistically significant difference between two features. "For example, if a sample size is 10 000, a significant P value is likely to be found even when the difference in outcomes between groups is negligible" (Sullivan and Feinn 2012, 280). As Green's data consisted of more than 20 000 note events, many hypothesis tests returned statistically significant results (often with values of $p < .001$). For this reason, effect sizes were included for each statistical test.

Each statistical test had its own effect size, e.g. Cramer's V for χ^2-tests, η^2 (eta-squared) for ANOVA, or Cohen's d for t-tests. Each also had their own scale of what was considered a small, medium, or large effect size. The standard magnitudes for effect sizes come primarily from Cohen (1988) *Statistical Power Analysis for the Behavioural Sciences.* The effect size magnitudes for each class reported by Cohen were designed primarily for use in the psychological and behavioural sciences. However, as they are a widely accepted set of magnitudes and classes, they were used in this research following the guidelines of the University of Cambridge's MRC Cognition and Brain Sciences Unit FAQ on 'Rules of thumb on magnitudes of effect sizes' (University of Cambridge MRC Cognition and Brain Sciences Unit 2009).[39]

3.2.3 Limitations of the Analysis

There were limitations with the analysis of Green's improvisational style that were known of before the analysis started, as well as others that presented themselves throughout the research. Issues related to specific features are discussed in the relevant sections of Part II.

One issue related to how this research interacted with both standard musicological approaches to jazz analysis and the computer-aided statistical approaches. To be analytically thorough, an initial set of data would have been gathered on Green, with this data analysed to create hypotheses about his improvisational data. Following this, more data on Green would have been gathered, with this data used to test the prior hypotheses. However, this approach was not followed for two

[38]Statistical significance was taken to be $p < .05$ throughout this research.

[39]https://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize. The page also contains resources regarding the use of effect sizes in research, and links to a paper by Bakker et al. (2019), discussing effect size thresholds and interpretations.

reasons. First, due to the scope of this research, and the time-intensive nature of transcribing the improvisations, it was not practical to follow this approach for this project. Although the corpus of Green's data could have been split into an evaluation and testing set, there was concern regarding the amount of available data and the impacts this would have on the analyses. Second, this approach, while analytically rigorous, is not one traditionally used when analysing improvised jazz. Instead, hypotheses are traditionally based on prior research, institutional knowledge from study, and professional experience from performance. Although these *a priori* hypotheses likely contain biases, they provide a strong foundation for further research and new hypotheses, and formed the foundation on which many hypotheses in this research were based. Finally, although this study employed a computer-aided approach through the use of statistical tools, it is predominantly a musicological study, and the author has endeavoured to apply the statistical tools and methods diligently. It is hoped that the methodology presented, and comments on issues found throughout this document, can form the basis of more rigorous studies in the future.

Another issue that was ever-present throughout this research related to how to classify and categorise musical features. When analysing symbolic musical transcriptions – although there is an awareness that the transcription does not fully represent what was played – there are clear delineations in features (e.g. rhythm).[40] As a result of the highly precise descriptive transcriptions, the ideality of symbolic notation no longer existed, leaving the subtlety and variety of human performance. However, this created issues in analysis, boundaries that were clearly drawn in symbolic notation, were not present in the data from the *MeloSpy* transcriptions. Instead, the data from the transcriptions was often presented as a spectrum. Consequently, depending on the feature and the aims of an analysis, segmentation needed to be reimposed upon the data.

Edge cases always exist, and if each note was investigated individually, two analysts could categorise features into separate classes. The segmentation of the data into suitable classes required an understanding of the musical form and practices of the music studied. The answers for how to approach this were rarely clear cut. For example, Figure 3.5 shows two representations of the same distribution of notes in bars of $\frac{4}{4}$. The left graph shows the raw note placement data, while the right shows a quantised version of the data (the *MeloSpy* mcm_48 feature).[41] There were three main differences between the representations:

[40] Pitch is one such feature where the delineations are the same between symbolic notation and those created through the methods of *MeloSpy*, where pitch is based on MIDI note values.

[41] The mcm_48 feature is not a straight quantisation of the raw data, with the values based on the FlexQ algorithm (Pfleiderer, Frieler, et al. 2017, 319).

1. The raw data is continuous around the circle, while the quantised has six prominent classes per beat;
2. In comparison to the quantised data, the raw data appears to be rotated anti-clockwise by approximately 10 degrees;
3. Although the main beats are close in location between the two representations, what would be the quaver between two beats is spread out and rotated even more in the raw data.

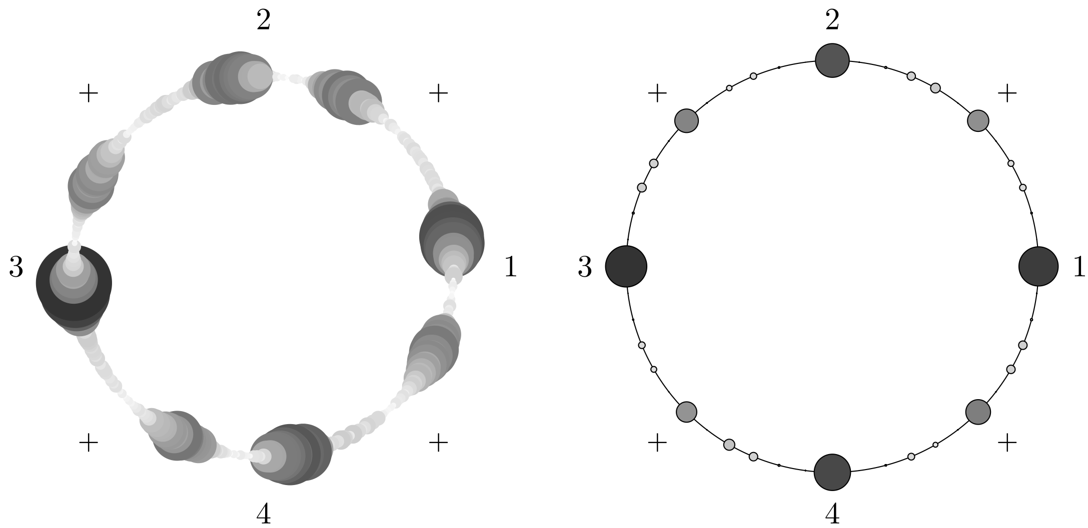

Figure 3.5: Distribution of Green's notes within the metrical context for improvisations in $\frac{4}{4}$. Left: raw note placement data. Right: quantised metrical circle map data (n=48).

Each of these had a simple explanation:

1. A quantisation cannot be continuous, the raw data will be grouped to the nearest class;
2. In general, Green is playing behind the beat, this resulted in an anti-clockwise rotation;
3. The displacement of the mid-beat quaver pulse is a result of swing, which was quantised to its nominal half-beat position.

The quantised version, while not looking like the raw data, was a reasonably accurate representation of the data, comparable to symbolic notation. This representation, while not accurate to what was played, was helpful in discussing concepts in a manner familiar to other musicians. Without the proper understanding of style and standard performance practices, quantised classes could be created that more closely match the underlying raw data, but were musically meaningless. The segmentation of the data in this research into classes needed to find the balance between representing what was played, while still having musicological meaning.

3.3 Performer Classification and Comparative Analysis

The performer classification and comparative analysis built upon the results of the analysis into Green's improvisational style. The aim of the performer classification task was to use the improvisational data to be able to identify the performers with a high degree of accuracy, using comprehensible ML algorithms. It also aimed to identify which ML algorithms performed the best, and how different abstraction levels influenced both the classification accuracy and the features used. From the results of the ML algorithms, specifically the features found to classify the performers, the comparative analysis provided an example of how this approach, and the results from it, could be applied to analysis. The performer classification and comparative analysis was based upon the same concept as the individual analysis, but used ML algorithms to act as the filter to select the features investigated. This section of the research was undertaken in four main steps:

1. Feature selection;
2. Training and evaluation of machine learning classifiers;
3. Analysis of features used to classify performers;
4. Example comparative analysis.

The selection of features was informed by the prior analysis into Green's improvisational style, and previous jazz and ML literature. The selected features focused on fundamental improvisational features, e.g. chordal pitch class, note lengths, or swing, rather than the instrument or year of recording. The model training and evaluation was undertaken with the `caret` package by Max Kuhn (Kuhn 2008). The ML algorithms used for this research were: C4.5-like decision trees (`J48`); C5.0 decision trees (`C5.0`); and random forest (`rf`). These models were selected as they were based on decision trees, which are broadly understandable, and reported the features that were most useful in the classification task.[42]

The `caret` package was used as it simplified the training and evaluation process, including data separation, cross-validation, and the tuning of parameters. The only parameter that could not be tuned by `caret` was 'ntree' when training the RFs. The RFs were first built with many (20 000–40 000) trees. The error rates, classification and out-of-bag, were then plotted against the number of trees. The optimal number of trees was then selected based upon the lowest overall stable error rate. When there were multiple comparisons, as in the one-vs-all and one-vs-one comparisons, the number of trees was determined by the lowest 'ntree' that best suited all of the models at that abstraction level.

[42]These were reported as variable importance metrics.

For each classifier the improvisatory data was collated at five abstraction levels and run through three types of comparison. The abstraction was the level at which the features were extracted and collated, for example, one row of data per note or per phrase. The five abstraction levels were:

- Solo level;
- Phrase Level;
- Two sliding windows at the Bar level ($Bar_{4|2}$: size 4, step 2; $Bar_{2|1}$: size 2, step 1);
- Note level.

The three types of comparisons were: n-way – all performers were classified against all other performers; one-vs-all – one performer was classified against the combined data of the other three performers; and one-vs-one – each of the four performers underwent pairwise classification. Figure 3.6 shows the five different abstraction levels, three classifiers, and three comparisons used, with the lines showing the connections between them. This resulted in 165 separately trained models.[43] Only models that were able to successfully classify the performers contributed to the features selected for the comparative analysis.

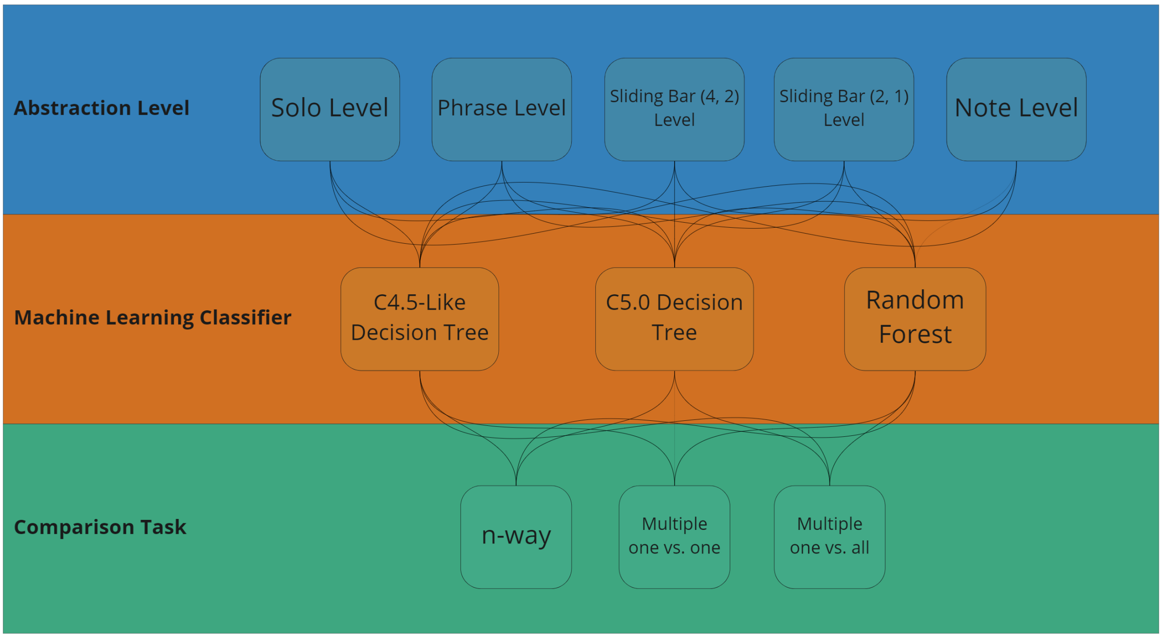

Figure 3.6: Depiction of the connections between the abstraction levels, machine learning classifiers, and comparisons in the performer classification tasks.

As with the prior section regarding the analysis of Green's improvisational style, much of the discussion related to specific selection of features for use in classifying, and those used in the comparative analysis are presented in their relevant chapter in Part III.

[43]For each abstraction level and classifier there was one n-way, six one-vs-one, and four one-vs-all.

3.3.1 Issues and Considerations

There were two main issues that needed to be considered, leading up to the training and evaluation of the ML models. The first related to dealing with the issue of using the data from the WJazzD. The second dealt with considerations on how to combine the data for the non-note level abstractions.

The issues with using the data from the WJazzD were two-fold. The first was that because the author did not select which improvisations were transcribed, the potential biases raised in the literature review could have been present in the selection of solos. There was also a substantial difference in the sample size of the three performers selected and that of Green. There was no solution for this as it was beyond the scope of this project to transcribe additional improvisations to deal with this issue. This additional data from the WJazzD represented the largest comparable dataset that could be used in this study, and was therefore the only option. Consequently, any discussion of feature differences found between the performers can only be considered for the available data, and should not be extrapolated. However, the approach taken, and the findings of the model evaluation, feature analysis, and comparative analysis still provided valid results, and can be used to inform future projects.

The second issue regarding the use of the WJazzD data related to differences in how aspects of the music were coded, in comparison to Green's corpus. The start of improvisations within the WJazzD, especially those that began at the end of a chorus, were frequently annotated with no chords (NC). Between Coltrane, Parker, and Davis there were twenty-seven improvisations where this occurred. For the non-phrase abstractions levels the bars with no chord related annotations could be excluded without substantial impact on the overall amount of data. For the phrase level, it meant that the entire data for at least one phrase would have to be excluded. Due to the smaller number of phrases that were played in each improvisation, this represented a large decrease in the available data. Although the chords were not annotated, the nominal changes could be found for each bar by looking at the "Database Content" page on the Jazzomat website (Jazzomat Research Project 2017). Using this information, the dataset was edited in *R* to list the correct chord changes for the NC bars. Relevant features, such as the chordal pitch class, were then recalculated for the bars with the new chord data.

The final consideration, which had to be made in conjunction with the feature selection process, was how to collate the data at the non-note abstraction levels.[44,45]

[44] The note level required no collation, only feature selection.

[45] The machine learning setup files, including the pre-processing done to create the base dataset for Part III, and the setup and training files for each abstraction, can be found in Appendix E.3.

For both the bar abstractions a custom function `slideFunct` was written to collate the data for each overlapping window. The function allowed the window and step size to be set, so it could be used for both bar abstractions.

A balance had to be found between variables that would be useful in the classification task, while still having meaningful musicological interpretations. For continuous data, this was achieved through measures of centre, specifically the mean, SD, and median. For categorical variables where there was likely to be only one class at a given abstraction level (e.g. the octave), or only one main class, the mode was used to select that class. For the other categorical variables, such as the distribution of intervals or chordal pitch classes played, the values were one-hot encoded (one feature for each class).[46] Each feature then reported the proportion of notes for that abstraction that contributed to each class.

For each of the non-note abstraction levels, the result was a dataset that summarised the improvisational data in each individual abstraction (e.g. each phrase) for every solo. This data was then used to train the ML models, with the variable importance metrics reporting which of the features were most useful in classifying the performers. The comparative analysis considered both the summarised versions of the features, as well as the raw features from the base datasets.

Code reproduction

This research started prior to the release of *R* version 3.6.0, which changed how the pseudo-random number generator worked.[47] The pseudo-random number generator (RNG) was used in the `set.seed()` function, which was used to split the data into training and testing sets for the classification task. The bug that this fixed only applied to very large populations of data, much larger than those used in this research. Therefore, there was no negative impact on this research with continuing to use the older version of the RNG. Code run with newer versions of *R* will generate slightly different results than those presented. To reproduce the code and results presented in this research, an additional command must be run: `RNGkind(kind = "Mersenne-Twister", normal.kind = "Inversion", sample.kind = "Rounding")`, which sets the RNG to match that which was originally used in this research.[48] As all of the models were trained using data split before *R* version 3.6.0, using the new RNG version would require all the models to be re-trained.

[46] One-hot encoding was preferred over dummy-encoding (where the resulting features is $n-1$ the number of classes) as it resulted in one feature for each class.

[47] https://bugs.r-project.org/bugzilla/show_bug.cgi?id=17494

[48] Only the `sample.kind` needs to be adjusted.

Part II
Analysis of Grant Green's Improvisational Style

Chapter 4

Feature Information from Grant Green's Transcriptions

Prior to the analysis it was necessary to have an understanding of the data that was contained within Green's corpus. This provided crucial information regarding which features could be analysed or used as independent variables. The corpus was drawn from forty improvisations played on thirty tracks, with Green improvising twice on ten of the tracks. Table 4.1 shows a summary of the duration (seconds) of the improvisations that comprised the corpus used throughout this research. The total duration of Green's corpus was 1 hour, 22 minutes and, 22.34 seconds, with a mean duration of 123.41 ± 86.86 seconds.

Table 4.1: Duration in seconds of Green's improvisations.

Min	Max	Med	Mean	SD	Sum
24.20	370.74	94.38	123.41	86.86	4936.50

The shortest improvisation was Green's second solo over *Minor League* (1964d) (24.20s), with five of the seven improvisations with a duration under a minute being the second solo in a song. The two exceptions to this were *Born To Be Blue* (1961c) (37.87s) and the first improvisation over *Brazil* (1962b) (59.52s). Two of Green's improvisations had a duration over six minutes, *Blues In Maude's Flat* (1961b) (6:02) and his first improvisation over *Seepin'* (1960b) (6:10). The majority (twenty-five) of Green's improvisations were between one and three minutes. Although the duration of each of Green's improvisations provided context to the variety of solos that Green played, it was of no further analytical use. A scaled duration, normalised onset, was used to investigate how Green's improvisational style changed throughout the course of a solo.

Table 4.2 shows that from the forty improvisations there were 20 478 note events from 2961 bars of music. As the data extracted from the database only reported details for beats with at least one note onset event, bars and chords were only counted when they were associated with note events. Within the entire corpus, there were only thirty bars with zero note events associated with them. There were 3869 chord annotations in Green's corpus, where multi-bar chords were counted for each bar in which they occurred. Only counting chords when a chord changed lowered the number of chords slightly to 3231. There were 1251 phrases annotated in of Green's corpus, ranging from a phrase with a single note to a phrase with 158 notes ($\bar{x} = 16.37 \pm 15.11$ notes).

Table 4.2: Green's general corpus details.

No. of Notes	No. of Chords	No. of Phrases	No. of Bars
20478	3869	1251	2961

Table 4.3 shows the distribution of time signatures in the corpus, with 92% of the improvisations being from songs in a quadruple time signature. As discussed in Section 3.1.2, transcriptions annotated with an $\frac{8}{4}$ time signature were tracks in $\frac{4}{4}$ that were played entirely in double-time by the whole ensemble. Due to the major imbalance in classes between quadruple and triple time signatures, the time signature of the improvisation could not be used as an independent variable in the analysis of Green's improvisational style. Consideration was given to differences between the time signatures, such as metrical stresses, throughout the analysis, especially within the Rhythm domain.

Table 4.3: Distribution of time signatures in Green's corpus.

Quadruple		Triple
$\frac{4}{4}$	$\frac{8}{4}$	$\frac{3}{4}$
28	9	3

The distribution of keys in the corpus is shown in Table 4.4. Similar to the time signatures, there is a substantial class imbalance between the major and minor keys, with a ratio of 3:1. Within this research, the broad category of tonal centre was called "mode", following the structure of keys in *MeloSpy* that were "coded in the form '`<NOTENAME>[-<MODE>]`' " (Jazzomat Research Project 2017). The mode included the general major and minor keys, as well as the standard modes of the major scale for modal compositions. There were no modal improvisations in Green's corpus. The major mode also contained all blues improvisations, despite blues not being in a major key. Within this corpus, the most prevalent key was B♭ major; although six of the ten B♭ major improvisations were B♭ blues.

Table 4.4: Distribution of key signatures in Green's corpus.

Major (Total: 30)							Minor (Total: 10)				
A♭	B♭	C	D♭	E♭	F	G	B♭	C	D	E♭	G
2	10	2	1	4	6	5	3	3	1	1	2

Figure 4.1 shows the distribution of key centres and modes for the improvisations of Green, the WJazzD, and a collection of songs used in the *iReal Pro* app (Biolcati 2017).[1] *iReal Pro* displays chord progressions for a variety of jazz, latin, and pop music, with sets of songs downloadable from the *iReal Pro* forum (Technimo 2017). For Figure 4.1 the "Jazz 1350" and "Brazilian 150" were combined into a single dataset of 1500 songs, containing the title, composer, feel, key, time signature, and chord changes. For this set of songs the key signature information was taken as a baseline distribution of keys with the following caveats:

- These are not "official" lead sheets or chord changes, so may contain errors;
- Although songs can be played in any key, the default key stored with each song was used (this may explain why the key of C Major is more prevalent than in either the Green or WJazzD data);
- Only major and minor keys are available, with no details on the number of blues.[2]

This graph showed that, with the exception of the key of C in the *iReal Pro* data, the general trend of major keys was similar between Green and the other two datasets. The general trend of the minor keys was also similar, although they were generally more prevalent in Green's corpus. Although the most frequent blues key for both Green and the WJazzD was the same, B♭ blues, there were no other similarities between these distributions.[3] Green improvised over as many B♭ blues as he did for his most frequent major key, F major (15.00%), while none of the blues keys in the WJazzD were as common as the most frequent major keys. The difference in the proportion of blues keys between the Green corpus and WJazzD is most likely due to two factors:

[1]Thirty-eight improvisations from the WJazzD are not represented in this figure; six chromatic keys – nine solos, one mixolydian mode, one dorian mode, and twenty-seven solos with no key recorded).

[2]There were fifty-six songs that have a duration of twelve bars, that of a standard blues, but they were not necessarily all blues. There are also many songs that use the blues form but aren't included with the "jazz 1350" song list, as one standard blues progression is the same as any other.

[3]Both 'major' and 'minor' blues were combined in the figure.

2. Sampling bias in the WJazzD, where fewer blues improvisations were transcribed;
3. The genres with which Green was most associated within the period of study were hard bop and post-bop, which drew heavily on the blues tradition.

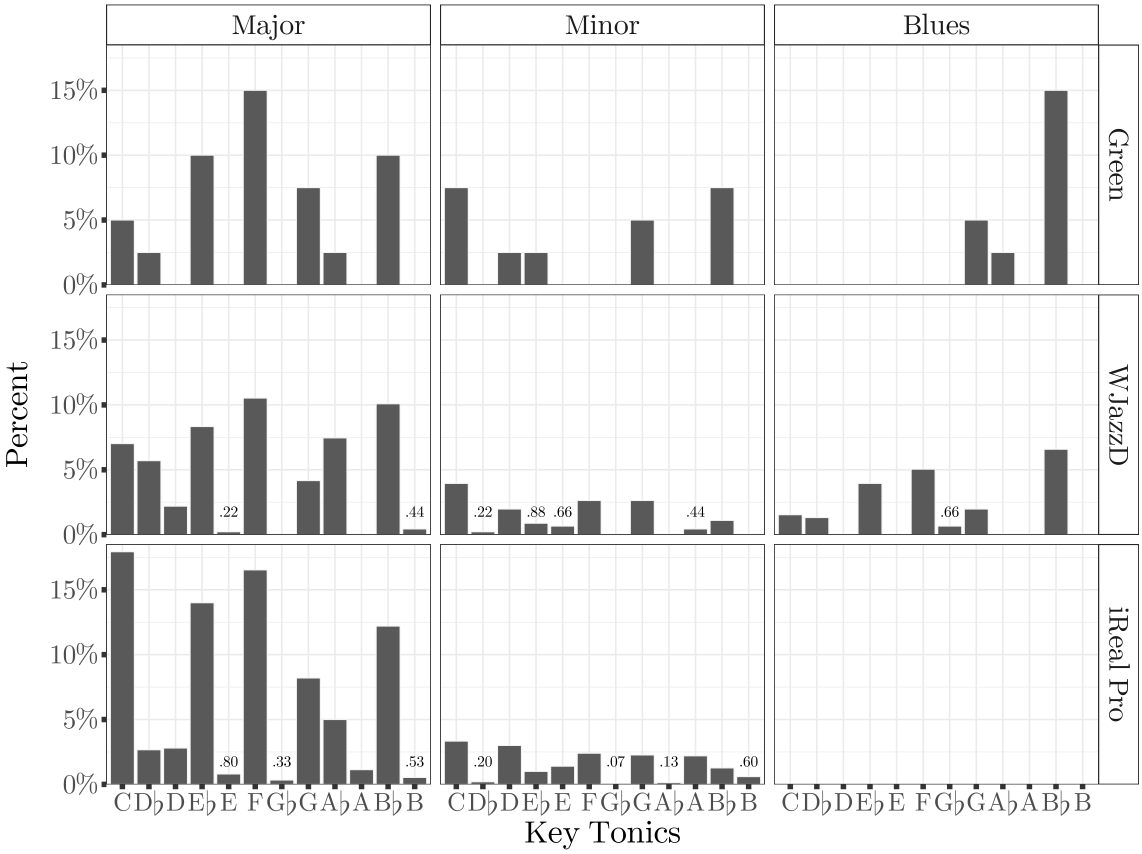

Figure 4.1: Distribution of major, minor, and blues keys in the datasets of Green, WJazzD and *iReal Pro*.

As the distribution of individual key signatures was too diverse it was not possible to use key signatures as an independent variable in the analyses. When considering the mode, the substantial class imbalance also meant that it was not practical to use the mode in the analyses. If blues were separated from the major mode, as in the above graph – combining the *MeloSpy* features "tonality type" and "mode" into a new feature "tonality mode" – the class imbalances were no longer as large, as seen in Table 4.5. There were still twice as many improvisations in a major tonality mode than either minor or blues. However, considering the importance of the tonality, the tonality mode feature was used in the analysis of Green's improvisational style.

Table 4.5: Distribution of tonality modes in Green's corpus.

Blues	Major	Minor
9	21	10

Table 4.6 shows the number of improvisations transcribed from each year in the period of study. The number of tracks selected for each recording year aimed to be roughly proportional to Green's output as a leader in those same years. As a result, there were substantially more transcriptions from 1961 and 1962, with these two years contributing more than half of the transcriptions and note events than any of the other years.[4] Due to this imbalance in the data, the year of recording was not used as an independent variable in the analysis of Green's improvisational style.

Table 4.6: Distribution of recording year in Green's corpus.

1960	1961	1962	1963	1964	1965
3	15	10	5	5	2

The distribution of chords, both quality and pitch, is displayed in Table 4.7. Above each chord symbol or group of chord symbols is the broader chord type class to which they were assigned in this research. The chords are shown distributed over the twelve chromatic pitches, with enharmonic equivalent ♭ and ♯ combined together. The chord tonics were not considered in the analysis as the chord based features were not dependent on the root note of the chord.

As the ii–V is a key building block of functional jazz harmony (Levine 1995, 19) the m7 and 7 chords were the most frequent in the data, with 1190 and 1912 occurrences respectively. Δ7 chords occurred approximately half as often as m7, with ø7 and $\circ$7 the least frequent. There was a substantial imbalance in the frequency of the chord types. However, chords are foundational for improvisation in functional jazz harmony, and had to be included in the analysis of Green's improvisational style, with the following caveats:

- due to the nature of jazz harmony these class imbalances would always exist;
- while each chord class could be analysed independently any comparative analysis between chord types focused on the three most frequent classes, 7, m7, and Δ7.

[4]The improvisations from 1961 and 1962 contribute 62.50% of the transcriptions, and 61.99% of the note events.

Table 4.7: Distribution of chord types and tonics in Green's corpus.

	C	D♭	D	E♭	E	F	G♭	G	A♭	A	B♭	B	Sum
Major 7													
$\Delta 7$	48	52	-	47	13	68	24	61	43	1	89	14	460
Maj	-	-	-	6	-	40	1	1	2	-	33	-	83
Total	48	52	0	53	13	108	25	62	45	1	122	14	543
Minor 7													
min	18	-	51	-	-	-	-	68	-	12	1	-	150
m6	11	-	-	-	-	-	-	-	-	-	-	-	11
m7	245	18	62	61	19	200	13	132	38	83	132	17	1020
$m^{\Delta 7}$	9	-	-	-	-	-	-	-	-	-	-	-	9
Total	283	18	113	61	19	200	13	200	38	95	133	17	1190
Dominant 7													
7	203	45	127	271	62	332	25	265	83	78	403	18	1912
Half Diminished 7													
ø7	39	-	73	-	-	1	5	1	-	24	-	31	174
Diminished 7													
$\circ 7$	-	12	15	6	59	-	8	-	-	-	2	38	140

The number of solos Green performed in each tempo class is shown below in Table 4.8 (BPM tempo range displayed beneath each class). Compared to the seven tempo classes used in this research, *MeloSpy* had only five tempo classes (Jazzomat Research Project 2017). The tempo classes Slow to Medium Up remain unchanged, while Up initially covered all tempos greater than 180 BPM. For the new tempo classes, following the ranges established by *MeloSpy*, Up was limited to tempos less than 220, with two new classes added, Quick ($220 \leq \text{BPM} < 260$) and Fast ($\text{BPM} \geq 260$).

Table 4.8: Distribution of tempo classes in Green's corpus.

Slow	Medium Slow	Medium	Medium Up	Up	Quick	Fast
$x < 60$	$60 \leq x < 100$	$100 \leq x < 140$	$140 \leq x < 180$	$180 \leq x < 220$	$220 \leq x < 260$	$260 \leq x$
0	2	13	4	8	13	0

Although this set of tempo classes was sufficient for Green, investigation of the WJazzD suggested that it would not be unreasonable to change Fast to cover BPM ranges of 260–300, with additional class for tempos ≥ 300 BPM. There were nineteen solos in the WJazzD with BPM ≥ 300, and only two where another class, BPM ≥ 340, would be required. Extensions to the tempo class feature, further to an additional class for BPM ≥ 300, would require transcribing many improvisations with a BPM of 260 to 360+, followed by analysis of whether there are identifiable differences in the improvisational style between these higher tempos.

Green had no improvisations in either the Slow or Fast tempo classes, with Green most frequently improvising over songs with a Medium or Quick tempo class.[5] Figure 4.2 shows for each beat in the transcriptions, the density distribution of the tempo, given by $60 \div$ Beat Length. The full-length vertical lines indicate the break points of the tempo classes and the rugs at the bottom show the mean tempo for each of Green's forty improvisations. There were eight outliers in the beat data, five where the tempo was less than 60 (BPM between 42.24 and 47.51, the last five beats of the second improvisation over *Seepin'* (1960b)), and three where the tempo was over 300. One, in the first beat of the second improvisation over *Minor League* (1964d) (382.81 BPM) and the other two as the last two beats of the first improvisation over *Go Down Moses* (1962d) (both 480.74 BPM). All outliers were due to issues in transcription, and in improvisations at the extremities of the tempo range. Aside from the beat in *Minor League* (1964d) that contained only a single note, all other outliers contained either no notes or only held notes at the end of the improvisation. The outliers were excluded from the graph.

The graph showed a split in the data around 170 BPM. This matched the data in Table 4.8, with approximately half of the improvisations (47.50%) played at a tempo ≤ 170 BPM, while the remaining 52.50% improvisations were played at a tempo > 170 BPM. As more than half of Green's improvisations were contained within only two classes (Medium and Quick), the use of tempo or tempo classes as an independent variable in analysis was troublesome. Splitting the data at 170 BPM created two approximately even groups of data for analysis. Tempo is a critical independent variable that had to be included in this study. Therefore, the binary tempo range was the default tempo feature used, with the raw tempo or tempo classes used when necessary with careful consideration due to the imbalances in the data.

[5] Over the entire WJazzD only eight improvisations (1.75%) had a BPM in the Slow tempo class.

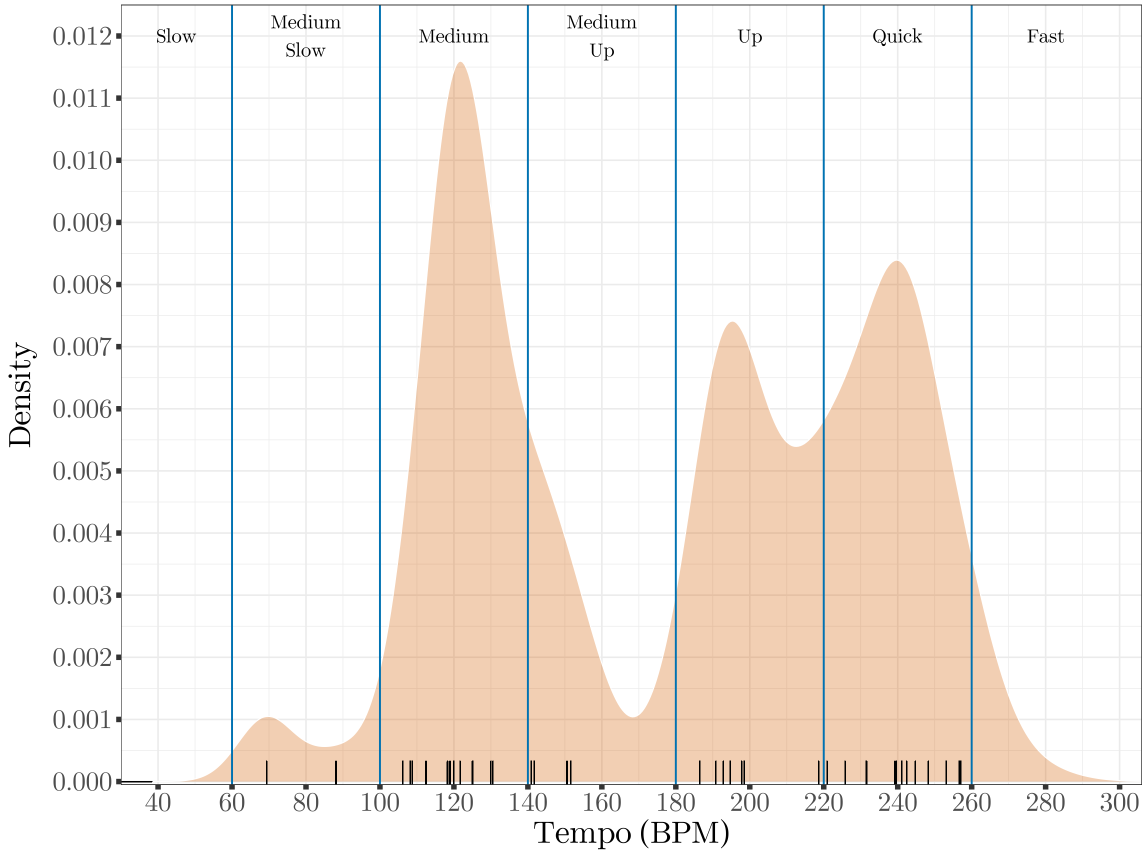

Figure 4.2: Density distribution of raw tempos for each beat in Green's corpus.

This chapter provided background data on common features of the songs over which Green improvised. This included general information of the tracks, the improvisations, the time signatures, key signatures, tonalities, years of recording, chords, and tempo. This also included discussion of any alteration or elimination of features in the analysis of Green's improvisational style. The following chapters present the analyses for each domain, starting with the Pitch Domain, followed by the Rhythm, Micro, and Macro domains, with Part II concluding with a summary of the findings.

Chapter 5

Pitch Domain

The pitch domain related to any feature associated with the pitch of the notes played. As a substantial amount of focus is placed upon the pitches played within an improvisation, the pitch domain is one of the most studied within improvised jazz. The pitch domain features were grouped into four broad feature classes: raw pitch; tonal pitch class (TPC); chordal pitch class (CPC); and intervals. The raw pitch focused on pitches without context to the key centre or the surrounding chords. The TPC dealt with the relationship of the pitches to the overall key centre of the improvisation. The CPC dealt with the relationship between the pitches and the nominal chords they were played over. Finally, the intervallic features related to the movement between pitches. From these high level features, two specific examples related to this pitch domain were investigated, surrounding note figures (SNF) and voice-leading.

5.1 Raw Pitch

The first pitch domain features to be analysed were related to the raw pitches with which Green improvised. The analysis of raw pitch features had limitations. The main limitation was that the analysis of raw pitch values existed in a void, without any context of the key or chord structure. It could not be used to compare different aspects of Green's improvisational style, nor compare Green against other performers. Analysis of the raw pitches provided some insight into the general key centres played by Green, and analysis of the pitch in comparison to the pitch range of the instrument allowed for investigations into where on the instrument a Green most frequently played.

5.1.1 Pitch Class

The pitch class (PC) was the categorical representation of the twelve chromatic pitches, irrespective of the octave. Figure 5.1 was inspired by the TPC Markov chain diagrams in the gallery on the Jazzomat website (Jazzomat Research Project 2017), and the chapter "Metrical Circle Map and Metrical Markov Chains" (Frieler 2008).[1] Figure 5.1 displays the overall pitch use (0^{th} order Markov chains, or a unigram representation) of the corpus as a circle map.[2]

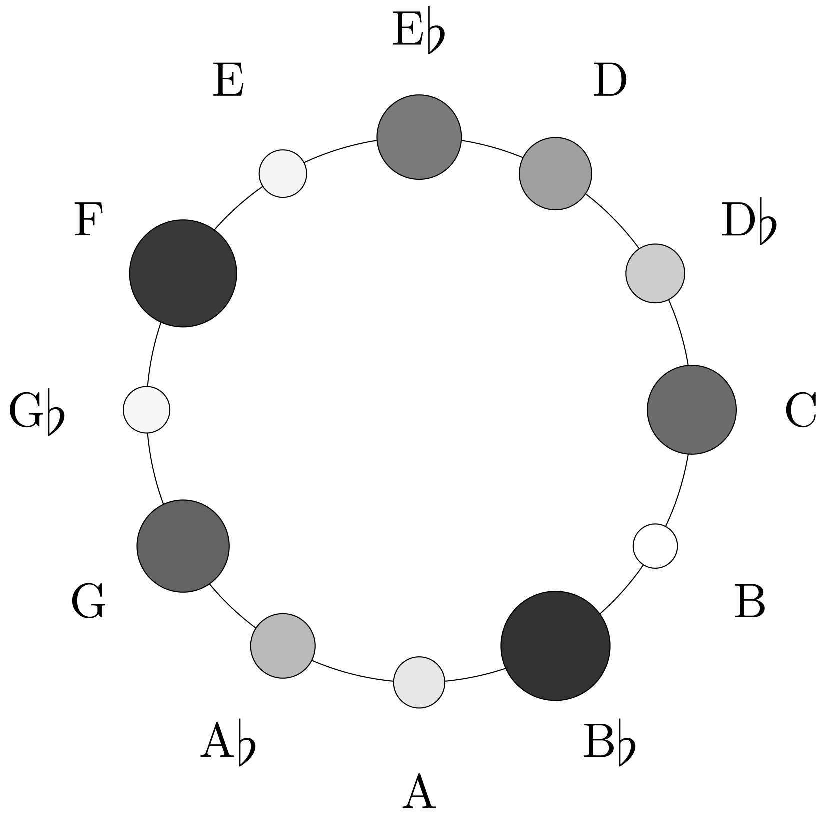

Figure 5.1: Pitch class circle map of Green's corpus.

The data here aligned with expectations based upon the distribution of key signatures in the corpus, with thirty-three of the forty improvisations (82.50%) played in 'flat' keys. This was evident through the high frequency of the notes B♭, F, E♭, C, and G, and the low frequency of the notes G♭/F♯, B♮, and E♮.

Figure 5.2 displays PC circle maps for each tonality mode. This figure shows a significant difference in the distribution of pitch classes used in these three tonality modes, with a small effect size ($\chi^2(22) = 680.79,\ p =< .001,\ V = .13$). The main

[1]The bigram (1^{st} order Markov chain) version of Figure 5.1 can be found in Appendix A, Figure A.1. The code to generate the circle maps can be found in Appendix B.

[2]The circles represented the frequency of each class through size and opacity, and were scaled so that the most frequent class was always the same size and had the same opacity; all other classes were scaled in proportion.

difference was that in the non-blues tonality modes Green played a broader distribution of PCs. The number of the blues improvisations played in B♭ had a substantial influence on the distribution of the PC. The blues PC circle map was highly representative of Green's note choice in a B♭ blues.[3] This data suggested that of the two blues notes, ♭3 and TT, Green strongly favoured the ♭3, with the TT rarely played.

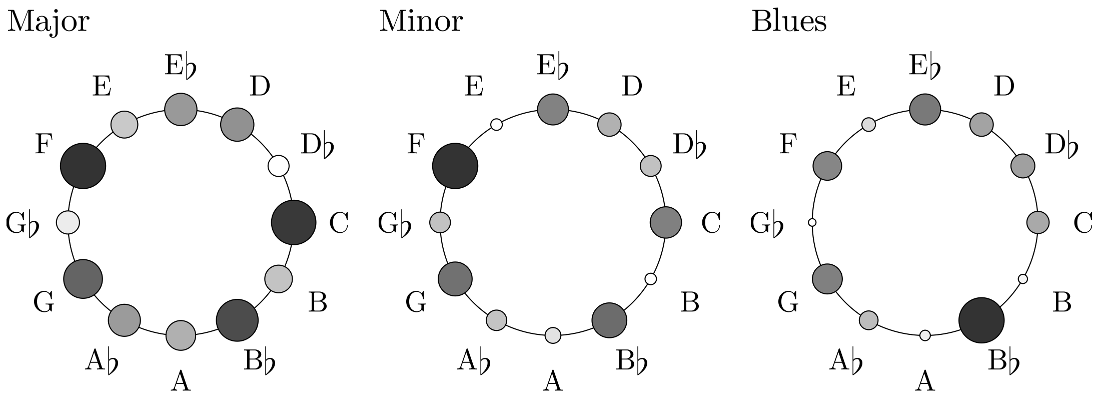

Figure 5.2: Pitch class circle maps for each tonality mode in Green's corpus.

5.1.2 Tessitura Pitch

More insightful than the PC was investigation into where on the guitar Green predominantly played. This data was represented both as octaves and through a scaled pitch feature, normalised instrument tessitura pitch (NITP). NITP is an expansion of *MeloSpy*'s tessitura normalised pitch (normalised pitch), which scaled all notes in an improvisation between 0 and 1 based on the lowest and highest pitches present in that solo. In contrast, the NITP was based on a standard pitch range of the instrument played. This normalisation allowed for better comparison across instruments and performers, as well as comparison between the same instrument across improvisations. For many instruments the lowest pitch is usually well defined while the highest pitch can be more flexible. In the case of the saxophone or trumpet the altissimo range is dependent on the performer; while for a guitarist it can be limited by the structure of the instrument – the number of frets or the design of the cut-out. The NITP pitch ranges were based on standard pitch ranges quoted in *Jazz Arranging Techniques* (Lindsay 2005). Consequently, there were pitches encoded with a NITP outside the 0 to 1 range. The range of the guitar was adjusted based on domain knowledge of the author, with 0 assigned to the open E♮ of the 6th string (E2), and 1 assigned to the C♮ on the 20th fret of the 1st string

[3]The unigram PC circle maps for each of the three blues keys are available in Appendix A, Figure A.2.

(C6), a 4th higher than the G5 in *Jazz Arranging Techniques* (Lindsay 2005, 17). Figure 5.3 shows the distribution of NITP for Green and the three guitarists represented in the WJazzD: John Abercrombie; Pat Martino; and Pat Metheny. This data indicated that Green tended to play in a higher register than the other guitarists.

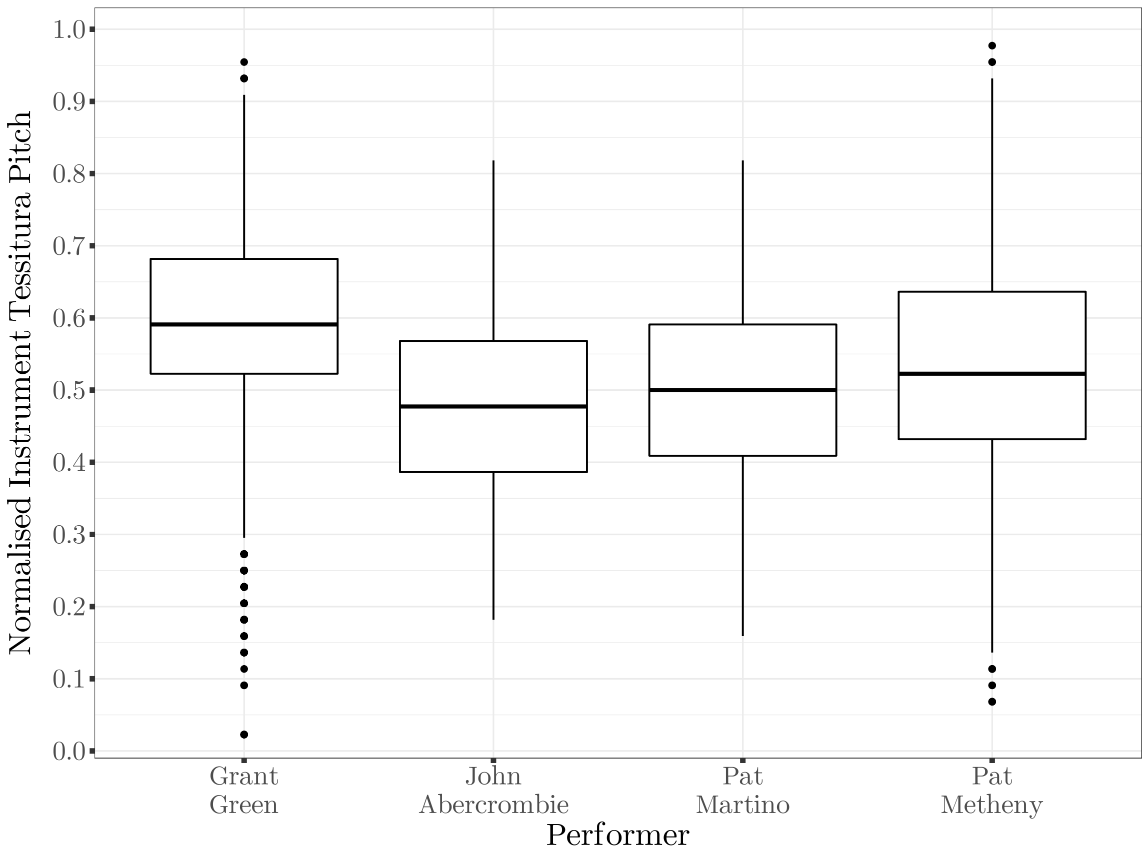

Figure 5.3: Distribution of NITP for Green vs. guitarists in the WJazzD.

Figure 5.4 shows two versions of the first twenty frets of a standard six string guitar. The top diagram shows the location of octaves across the fretboard (using octave numbers 2–6, based on Scientific Pitch Notation[4]), with middle C_4 on strings two to six emphasised in bold. The second diagram shows an approximate heat map of the most frequently played locations on the guitar within jazz improvisations.

This heat map was not drawn from any specific data, but based on the authors own experience. It is an approximation presented as an aid for a reader unfamiliar with the guitar. The numbers at the top list the fret numbers of the guitar, while the diagram in the middle shows the common location markers that often appear on the fretboard or along the top of the neck of the guitar (the markers are at frets 3, 5, 7, 9, 12, 15, 17, 19, with a double dot at the 12th fret to indicate one octave from the open strings).

[4]See Young (1939) for more details on Scientific Pitch Notation.

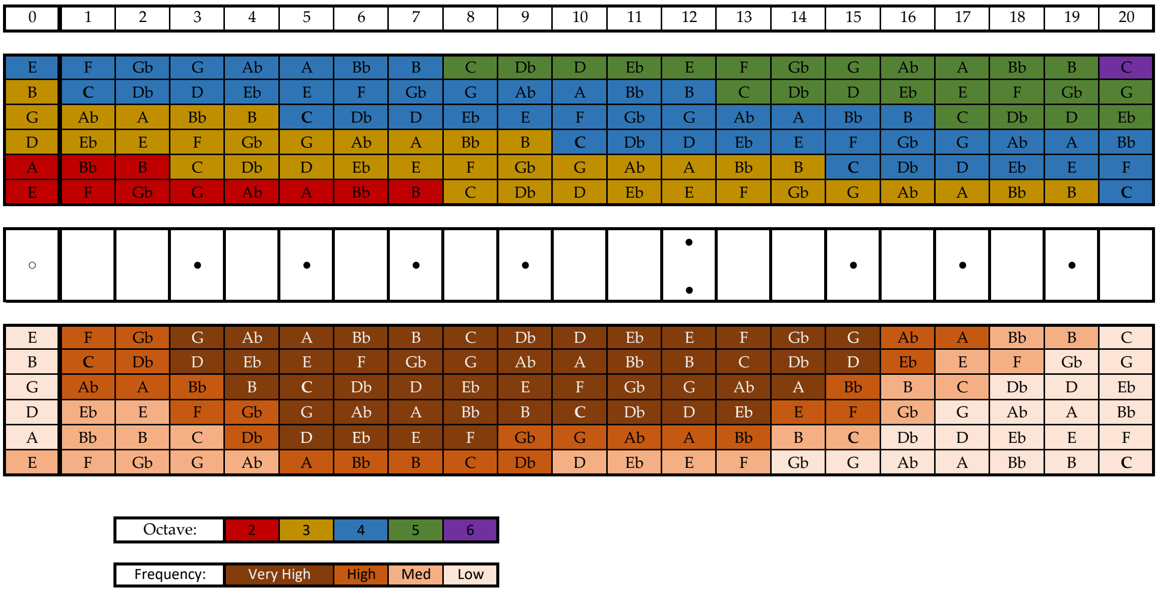

Figure 5.4: Diagram showing the layout of notes on a standard guitar.

Table 5.1 categorised the data from Figure 5.3 into the octave in which each note was played. This showed that 72.36% of notes Green played were in the 4th octave. The 5th octave was the second most frequent (16.71%), followed by the 3rd (10.83%), and 2nd (0.10%).[5] Overall, this distribution of octaves matched fairly close to the approximate heat map distribution, and suggested that a large proportion of Green's improvisation were played in the middle to upper registers (i.e. the top four strings, 5th fret and higher). The data also showed that the distribution of Green's notes amongst the octaves was significantly different from those of the other guitarists in the WJazzD, with a small effect size ($\chi^2(9) = 1028.26,\ p =< .001,\ V = .12$). Compared to the other guitarists, Green played substantially fewer notes in the 3rd octave, with Green favouring the 5th octave over the lower registers.

[5]The 2nd octave comprised a total of eleven frets and open strings at the lowest register of the guitar.

Table 5.1: Octave distribution for Green vs. guitarists in the WJazzD.

	2	3	4	5
Grant Green				
Count	21	2218	14818	3421
Percent	0.10%	10.83%	72.36%	16.71%
John Abercrombie				
Count	-	112	157	29
Percent	-	37.58%	52.68%	9.73%
Pat Martino				
Count	2	281	524	44
Percent	0.24%	33.02%	61.57%	5.17%
Pat Metheny				
Count	18	647	1365	344
Percent	0.76%	27.25%	57.50%	14.49%

Figure 5.5 shows Green's distribution of raw pitches in each octave, with the 2^{nd} octave's y-axis scaled to a one-hundredth of a percent, highlighting how infrequently Green played in that octave. Certain notes – F4, G4, B♭3, B♭4, and C5 – stood out as occurring substantially more frequently than other notes in their respective octaves. This matched the data in Figure 5.1, and was likely influenced by Green's octave preferences, and the transition points for octave numbering. As the vast majority of Green's notes were played in the 4^{th} octave, if Green wanted to play an F or G, it was nearly always an F4 or G4 as they occurred in the middle of the octave. In contrast, as the pitches C and B♭ were at the boundary of the octave numbers, those pitch classes were more evenly distributed across the octaves. This data also suggested that Green rarely played above the 12^{th} fret. Although there are notes in the 4^{th} octave above the 12^{th} fret, if he often played there the frequency of notes higher than F5 would likely be higher. Further investigation of the relation of the octave to Green's improvisational style focused on two areas: the length of time he spent in each octave; and octave transition properties.

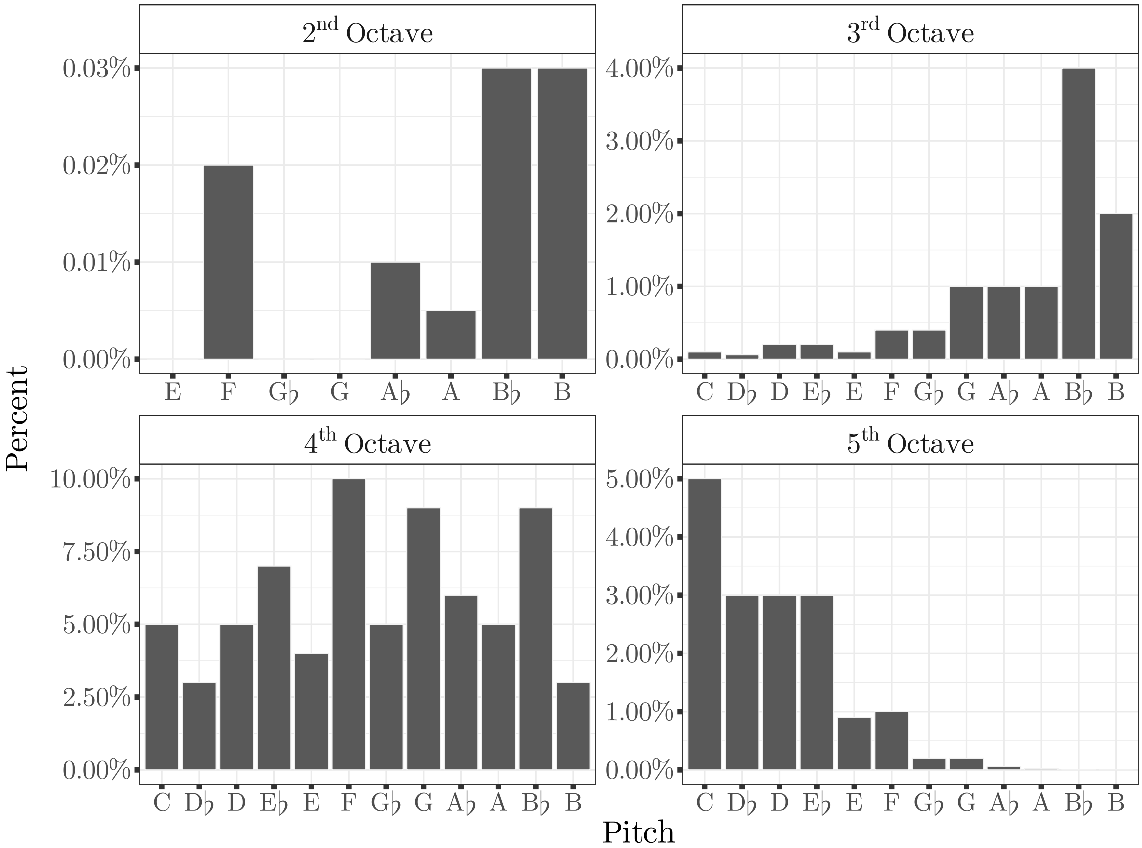

Figure 5.5: Distribution of Green's pitches in each octave.

Octave Duration

There were multiple methods for considering the length of a musical concept, with approaches differing depending on the musical features or analysis. Figure 5.6 depicts two methods for representing the octave duration.[6] The graph on the left shows the distribution of octave sequence lengths based on the number of notes played consecutively in the same octave. This data showed that the longest note sequences in a single octave occurred in the 4$^{\text{th}}$ ($\bar{x} = 6.75 \pm 8.42$ notes) and 5$^{\text{th}}$ ($\bar{x} = 2.74 \pm 4.12$ notes) octaves. The 3$^{\text{rd}}$ octave was most similar to the 5$^{\text{th}}$ ($\bar{x} = 2.30 \pm 2.14$ notes), while the 2$^{\text{nd}}$ octave had the shortest sequence of notes ($\bar{x} = 1.31 \pm 0.60$ notes). The graph on the right show the distribution of continuous time spent in each octave based on the length of the notes in relation to the surrounding beat length. For single note sequences, or the last note of a sequence, this was based on the duration, while for all other notes the inter-onset interval (IOI) was used.[7] This represented the number of beats from the onset of the first note of the sequence to the offset of the last note.

[6]The graphs focused on the range 0–20, containing 97.13% and 99.05% of the note count and duration sum data respectively.

[7]Duration was the time between a note's onset and offset while IOI was the time between the onset of two consecutive notes. $\text{Duration}_{\text{BeatProp}}$ and $\text{IOI}_{\text{BeatProp}}$ represented the duration or IOI as a proportion of the surrounding beat length. A proportional value of 1 was equivalent to a crotchet in simple time.

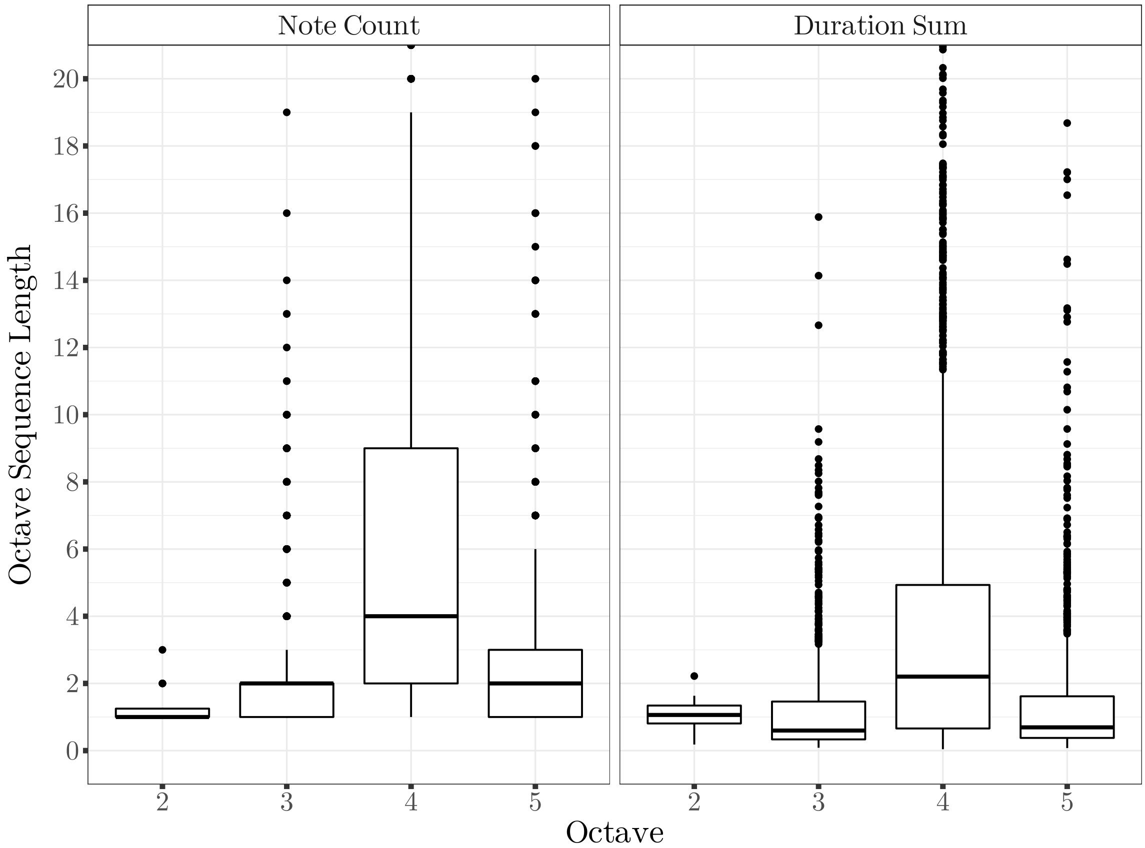

Figure 5.6: Distribution of how long Green spent in each octave.

The overall trend of the note length was similar to the number of notes. In comparison to the note count, the median beat duration for the 2nd octave was higher than both the 3rd and 5th octaves. This indicated that although the number of notes played consecutively in the 2nd octave was the lowest, the notes tended to have a longer length than those in the 3rd or 5th octaves. Table 5.2 lists for each octave, the frequency of each note count sequence length below ten, and then in groups of ten until fifty, with a final group for sequences with more than fifty notes.

Green's corpus contained 4426 octave sequences. 35.99% had an octave sequence length of one, with an additional 18.14% having a sequence of two notes before the octave change. The 2nd octave was substantially different from the other octaves, with 75.00% of sequences having a length of one. Additionally, Green never played more than three notes in a row in the 2nd octave. Nearly all of the very long (> 50) single octave sequences occurred in the 4th octave, with a small number also played in the 5th octave. Bechtel noted that the "use of elements such as repetition … [were] prevalent in Green's playing" (2018, 57).

Table 5.2: Frequency of octave sequence lengths (note count) in Green's corpus.

	Octave			
	2nd	3rd	4th	5th
1	12	481	487	613
2	3	245	319	236
3	1	69	171	155
4	0	64	170	71
5	0	26	154	56
6	0	28	135	37
7	0	12	111	18
8	0	12	96	19
9	0	13	91	8
10	0	8	59	4
11-20	0	8	284	23
21-30	0	0	72	5
31-40	0	0	22	2
41-50	0	0	11	1
>50	0	0	12	2

Figure 5.7 shows five octave sequences, including examples of repeated note sequences. These five sequences ranged from no structured repetition of notes to more complex repeated sequences:

a) *Blues In Maude's Flat* (Green 1961b), bars 55–56, no repeated note pattern, most common note 15.00%;
b) *Red River Valley* (Green 1962j), bars 40–43, single repeated note, most common note 86.67%;
c) *Miss Ann's Tempo* (Green 1961m), bars 96–100, two pitches in a three note sequence, most common note 69.57%;
d) *Tico-Tico* (Green 1962m), bars 83–86, three pitches in a four note sequence, most common note 50.00%;
e) *Stella By Starlight* (Green 1965d), bars 11–13, three pitches in a four note sequence repeated three times, followed by the same structure three times a perfect 4^{th} lower, most common note 24.00%.

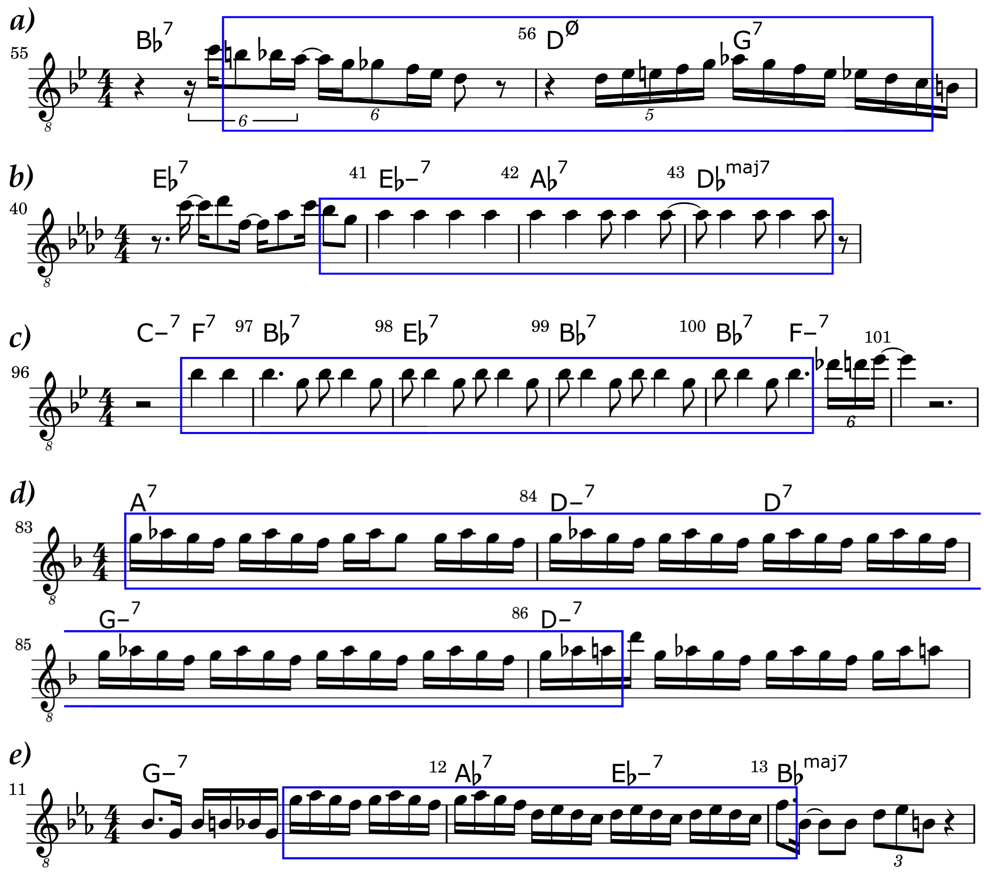

Figure 5.7: Examples of octave sequences in Green's improvisations.

It was hypothesised that longer phrases were most likely to be related to Green playing a repeated note sequence.[8] Figure 5.8 displays the count of the most frequent note against the length of the octave sequence, for sequences with more than ten notes. The line shows the intercept and slope from a line regression test (Adj $R^2 = .71$; $p < .001$). This data showed a substantial significant positive relationship between the length of an octave sequence and the number of times Green played the most frequent pitch. These results supported the hypothesis that as Green played longer in a single octave the occurrence of the most frequent note tended to increase, suggesting the use of repeated note patterns.

8 This would only apply to repeated note sequences contained within a single octave. Green also played repeated note sequences across octaves.

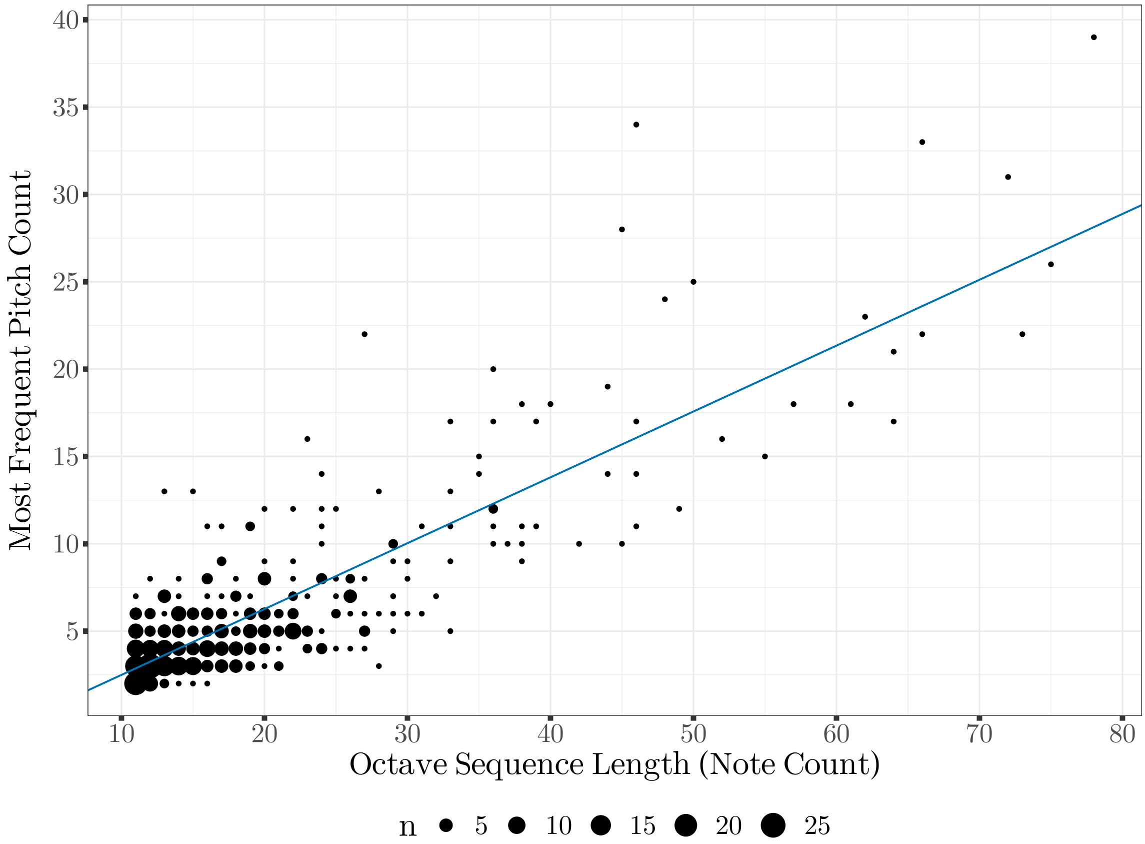

Figure 5.8: Count of most frequent pitch vs. octave sequence length.

Octave Transitions

The other investigation into the relationship of octaves to Green's improvisational style was the transition properties between octaves.[9] This focused on three transitional elements: the raw transition probabilities; the intervals played when changing octave; and the length of the note played when transitioning between octaves. Table 5.3 shows, for each octave, the octave distribution of the following note. 78.54% of notes Green played were followed by another note in the same octave. A further 21.32% changed from one octave to a neighbouring octave, with 0.14% moving to a note two octaves away.

Only in the 2nd octave was Green more likely to play a note in another octave (76.19%). In the 3rd octave, 56.83% of the notes Green played were followed by another note in the 3rd octave, with a similar trend observed when Green played in the 5th octave . In the 4th octave, the vast majority (85.33%) of the notes Green played were followed by another note in the same octave. Although this was substantially higher than the other octaves, it fit with the overall distribution of notes Green played.

[9]The term octave, when used to discuss transitions, referred to the octave number, instead of a leap of twelve semitones.

Table 5.3: Distribution of notes per octave, following a note played in a specific octave.

	Next Octave			
	2nd	3rd	4th	5th
2nd Octave				
Percentage	23.81%	52.38%	23.81%	-
Count	5	11	5	-
3rd Octave				
Percentage	0.64%	56.83%	41.90%	0.64%
Count	14	1252	923	14
4th Octave				
Percentage	0.01%	6.37%	85.33%	8.28%
Count	2	943	12624	1225
5th Octave				
Percentage	-	0.23%	36.29%	63.48%
Count	-	8	1241	2171

Excluding notes played in the same octave shows how likely Green was to transition from one octave to another. From the 2nd octave 68.75% of the transitions were to the 3rd octave, with only 31.25% jumping up to the 4th octave. In contrast, 97.06% of notes played by Green in the 3rd octave ascended to the 4th octave. From the 4th octave Green ascended to the 5th octave slightly more frequently than descending to the 3rd, 56.45% compared to 43.46%. In the entire corpus there were only two occurrences when Green played a note in the 4th octave followed by a note in the 2nd. Finally, from the 5th octave 99.36% of the transitions were Green descending down to the 4th, with only 0.64% descending two octaves to the 3rd.

Figure 5.9 shows the fuzzy interval distribution for each type of octave transition Green played: no octave; one octave; or two octave transition.[10] As all of the intervals played when transitioning between two octaves would be in the -5/5 fuzzy interval classes, the baseline for that graph was adjusted so that the fuzzy interval classes represent the same ranges but an octave higher or lower.

[10] Table 5.11 in the Intervals section fully details the fuzzy interval feature, briefly (fuzzy interval class (semitones)): 0 = Repetition (0); 1 = Step (1,2); 2 = Leap (3,4); 3 = Jump (5,6,7); 4 = Big Jump (8,9,10,11,12); 5 = $> 8^{ve}$ (≥ 13). The - sign indicated a descending interval.

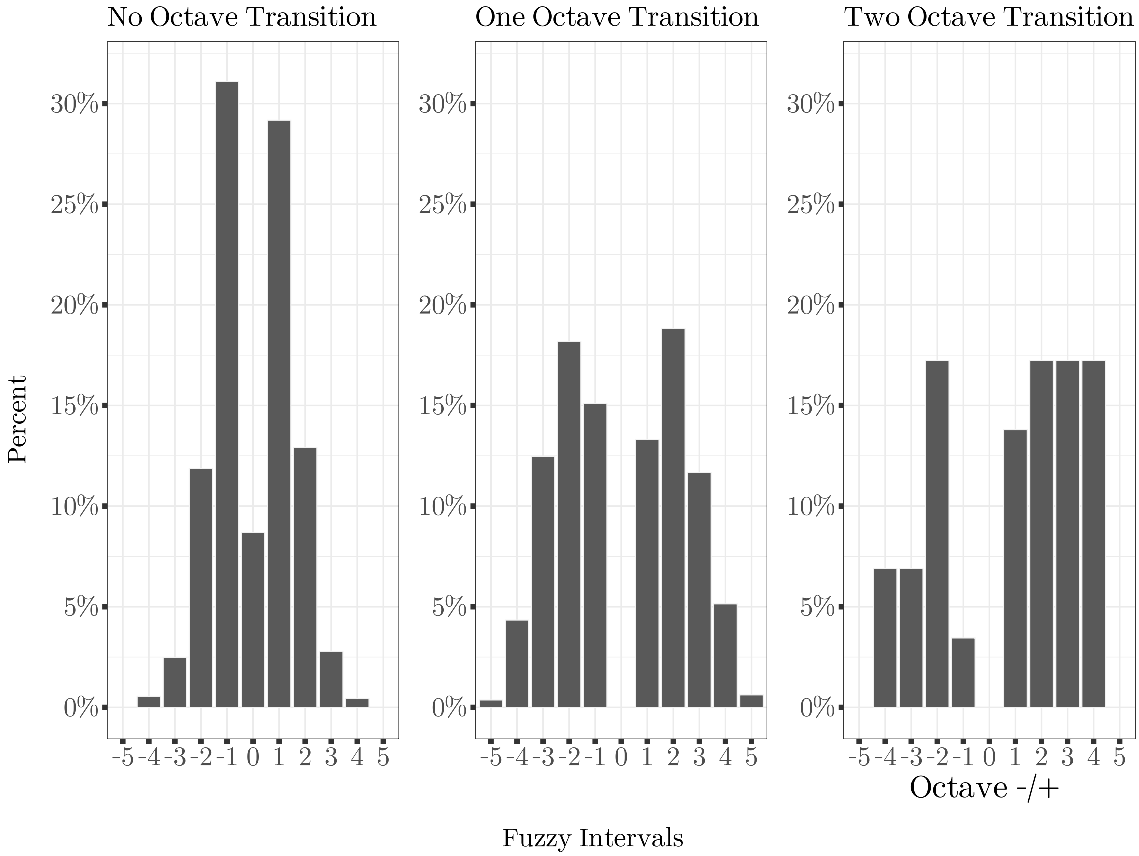

Figure 5.9: Distribution of fuzzy intervals used in octave transitions.

The 'No Octave Transition' graph shows that the majority of Green's intervals were an ascending or descending step. In contrast, when transitioning to a neighbouring octave, Green was more likely to play a leap (37.00%) than a step (28.41%). When Green transitioned between non-neighbouring octaves there were no clear trends in the intervals, with the graph indicating that Green was slightly more likely to have large octave leaps in an ascending direction. This was indicative of situations when Green played a descending phrase ending towards the lower register of the guitar, followed by his next phrase beginning in the middle to high range. This can be seen in Figure 5.10, an excerpt from Green's improvisation over *Round About Midnight* (Green 1961r).

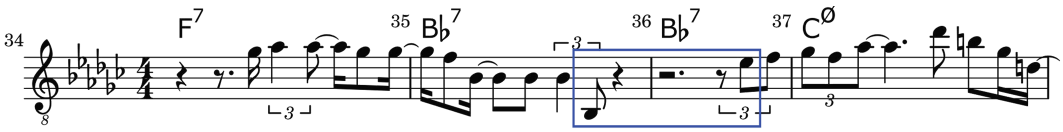

Figure 5.10: Example of a large ascending interval between phrases, *Round About Midnight* (1961), bars 34–37.

It was hypothesised that the larger leaps between octaves were associated with longer gaps between notes. This was supported by the data in Figure 5.11, which shows the $IOI_{BeatProp}$ distribution for each octave transition.[11] The $IOI_{BeatProp}$ when Green transitioned to a neighbouring octave was only marginally higher than no octave transition, median of 0.54 (IQR: 0.36–0.96)) beats compared to a median of 0.43 (IQR: 0.28–0.63) beats. When the transition was between non-neighbouring octaves the median $IOI_{BeatProp}$ was 2.57 (IQR: 1.97–3.03) beats. The median value of 2.57 equals approximately two and a half beats between the onset of the first note and the onset of the note two octaves away. In comparison, when Green wasn't transitioning between octaves, or only moving to the neighbouring octave, the median $IOI_{BeatProp}$ was around half a beat. This suggested that transitions to neighbouring octaves frequently happened through the course of Green playing a line, with larger leaps more likely to occur when there were longer gaps between notes (e.g. between phrases).

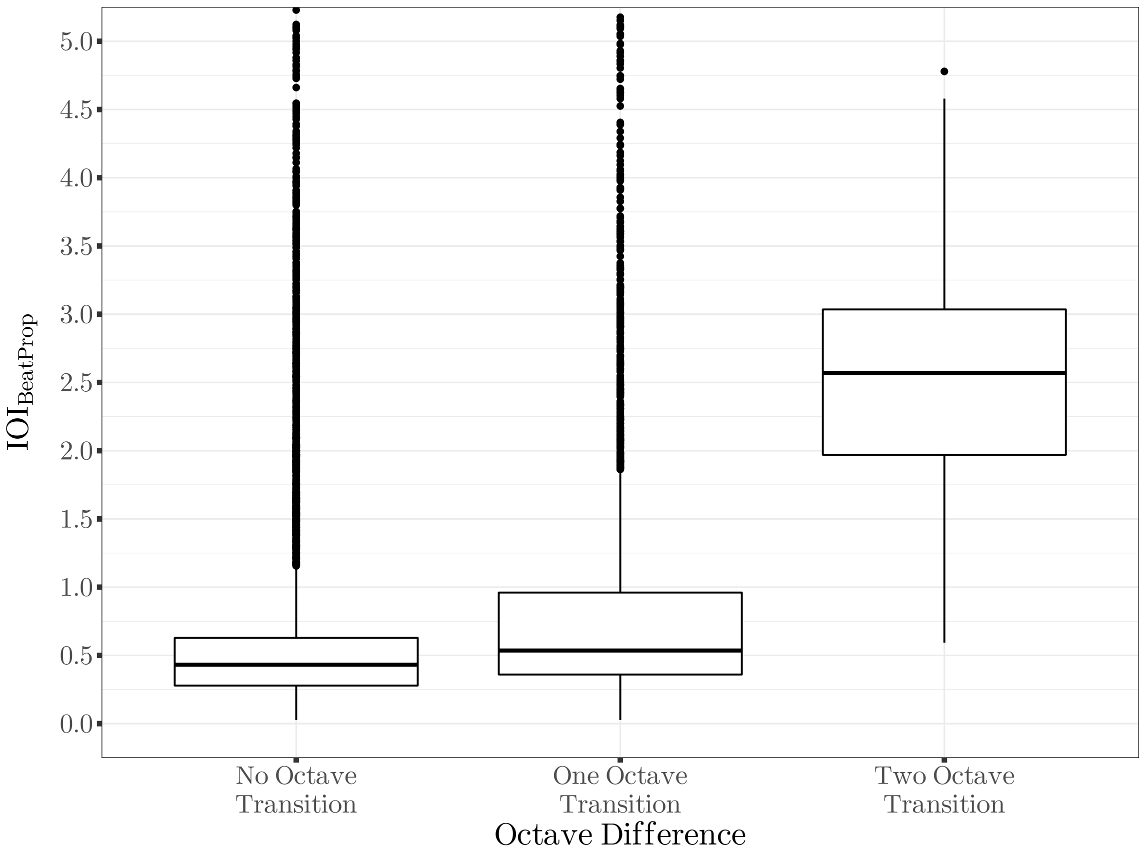

Figure 5.11: $IOI_{BeatProp}$ distribution for each octave transition.

[11]The graph focused on the data with an $IOI_{BeatProp}$ between zero and five, which contained 99.52% of the data.

5.1.3 Raw Pitch Summary

In summary, Green improvised predominantly in the 4th octave, and played more in the 5th octave than the other guitarists in the WJazzD. Green's longest octave sequences also occurred in the 4th octave. The longer Green spent improvising in the same octave the more likely it was that this coincided with a repeated note pattern, with only a few pitches contributing to the sequence. The majority of the note transitions played by Green were between notes in the same octave; when there was a transition it was most likely to be to a neighbouring octave. When Green did transition between neighbouring octaves, arpeggio motion was most common. Investigation into the length of the transition notes found that neighbouring octave transitions frequently occurred throughout the course of playing a line while larger leaps between non-neighbouring octaves were more likely to occur between Green's phrases.

5.2 Tonal Pitch Class

An early way musicians are taught to think about notes is through learning their scales, and relating the notes to the key centre of the piece they are playing. Similarly, an initial approach to improvisation can be to take the key centre of the tune or form and play the appropriate scale associated with that key centre. The TPC compares the pitches played to the overall key of an improvisation.[12] Although there were limitations to the TPC, it aided in providing a high level view of how Green's pitches related to the key centre. The graphs in Figure 5.12 show the TPC distribution for the major, minor, and blues tonality modes. The highlighted bars indicate diatonic tones (DT), with all other notes non-diatonic tones (NDT).[13] The TPC weight feature classified notes as either DTs or NDTs.

These graphs indicated that the majority of Green's notes were DTs (Major – 77.67%; Minor – 81.07%; Blues – 66.12%), with some NDTs indicative of elements of Green's improvisational style. Although the prevalence of DTs appeared to be less for blues when compared to the other tonalities, the blues scale contains only six notes, compared to the seven note major and minor scales. While not strictly diatonic to the blues scale, the ♮2, ♮3, and ♮6 are all consonant throughout a blues. Considering any one of these three notes as a DT raised Green's DT percentage in a blues to approximately 75% (♮2 – 74.29%; ♮3 – 74.64%; ♮6 – 76.37%), similar to that

[12] While the TPC accounted for annotated key changes within a piece, transient modulations were not considered.

[13] Based on the ionian mode for major, the aeolian mode for minor, and the ionian mode to label degrees for the blues scale.

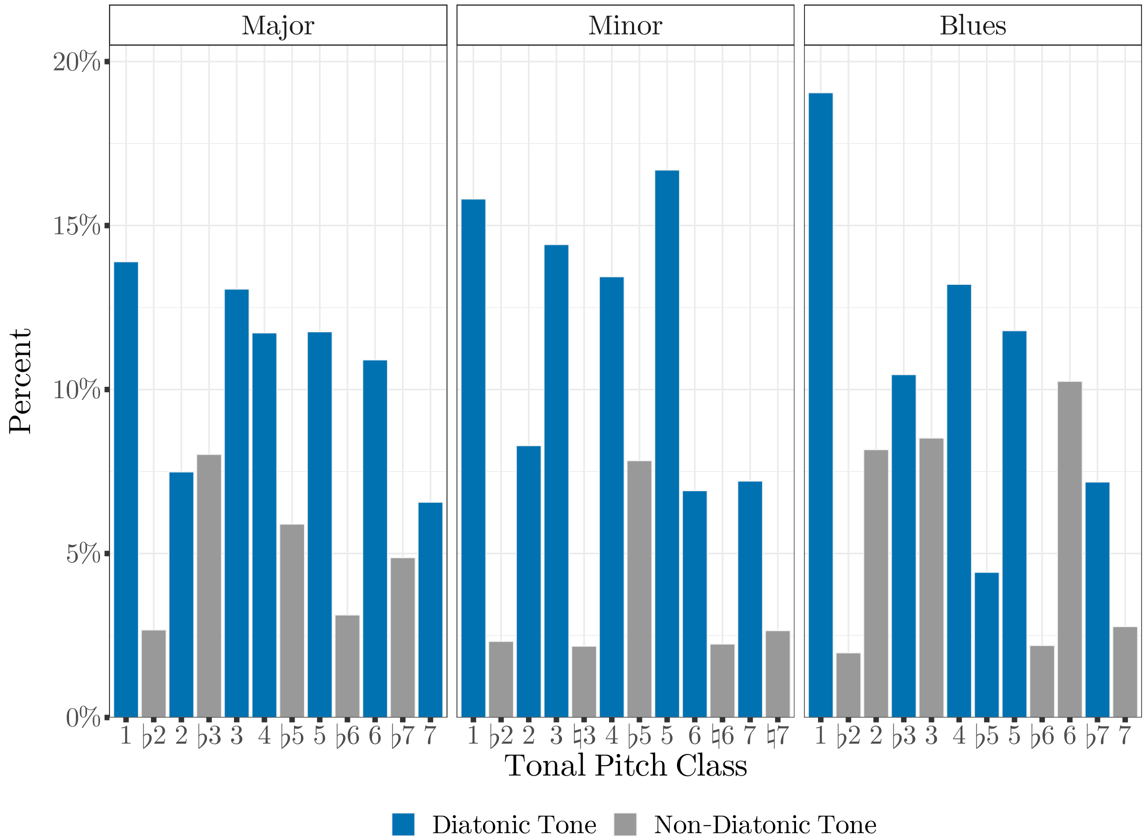

Figure 5.12: TPC distribution for major, minor, and blues tonality modes.

of the major or minor distribution. Therefore, the blues tonality mode was considered as a combination of the tonic blues and relative minor blues scales, which contributed the ♮2, ♮3, and ♮6, to form a scale with nine DTs.

Investigation of the major and minor distributions indicated that the prominent NDTs were the ♭5 for both major and minor, and the ♭3 and ♭7 in the major tonality. These tones suggested that Green employed blues language throughout his improvisations, agreeing with the findings from Bechtel (2018), Scott (2006), and Wild (2002). Figure 5.13 shows an excerpt from Green's improvisation over *Miss Ann's Tempo* (Green 1961m), playing a B♭ blues scales over the first four bars of the blues form.

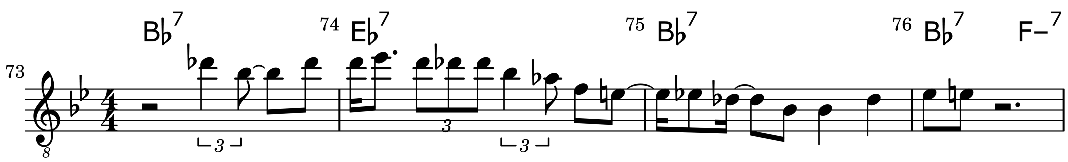

Figure 5.13: Example of blues language in Green's improvisations, *Miss Ann's Tempo* (Solo 1, 1961), bars 73–76.

As blues language was expected over a blues form, it was more insightful to investigate blues influences in Green's improvisations over the major tonality mode. The distribution of TPC for major tonalities in Figure 5.12, specifically the higher than expected use of ♭3, ♭5, and ♭7, could be explained by the use of the tonic blues scale in the major tonalities.[14] This suggested – along with the established influence the blues had on hard bop and jazz – that Green frequently employed blues influenced language within his improvisations, agreeing with the previous findings of Wild (2002), Scott (2006), and Bechtel (2018).

5.2.1 Treatment of Non-Diatonic Tones

The following section investigated Green's treatment of NDTs throughout his improvisations, focusing on their frequency of use, metrical placement, note length, the TPC weight of the surrounding notes, and the intervals used to move into and out of NDTs. Table 5.4 shows the TPC weight distribution for each tonality mode. In the entire corpus 17.76% of notes were NDTs. It was hypothesised that Green may have preferred to play NDTs off the beat, for example, as a chromatic passing tone. However, a χ^2-test found no significant difference in the TPC Weight distribution for on and off beat notes in Green's improvisations ($\chi^2(1) = 0.04$, $p = .843,\ V = .00$).

Table 5.4: TPC weight distribution for each tonality mode in Green's corpus.

	Diatonic Tone	Non-Diatonic Tone
Major		
Count	6777	2210
Percent	75.41%	24.59%
Minor		
Count	5063	1053
Percent	82.78%	17.22%
Blues		
Count	5002	373
Percent	93.06%	6.94%

Figure 5.14 shows the $\mathrm{IOI}_{\mathrm{BeatProp}}$ distribution for both TPC weights.[15] The line shows the mean value for each distribution. The data indicated that Green tended to play NDTs ($\bar{x} = 0.51 \pm 0.58$ beats) significantly shorter than DTs ($\bar{x} = 0.67 \pm 0.77$

[14]The tonic blues scale refers to when the key of the song matches the blues scale, e.g. B♭ blues for a song in the key of B♭ major. In comparison, a relative minor blues scale would be a the blues scale that matches the relative minor key of the song e.g. a G blues for a song in the key of B♭ major.

[15]Graphs focused on $\mathrm{IOI}_{\mathrm{BeatProp}}$ between 0 and 4 beats, representing 99.06% of the data.

beats), with a small effect size ($t(6636.70) = -14.64,\ p < .001,\ d = -0.36$).[16] The data showed that the majority of NDTs were played for half a beat or less, with a few played for one beat. In comparison, a higher proportion of DTs Green played went for one beat or more.

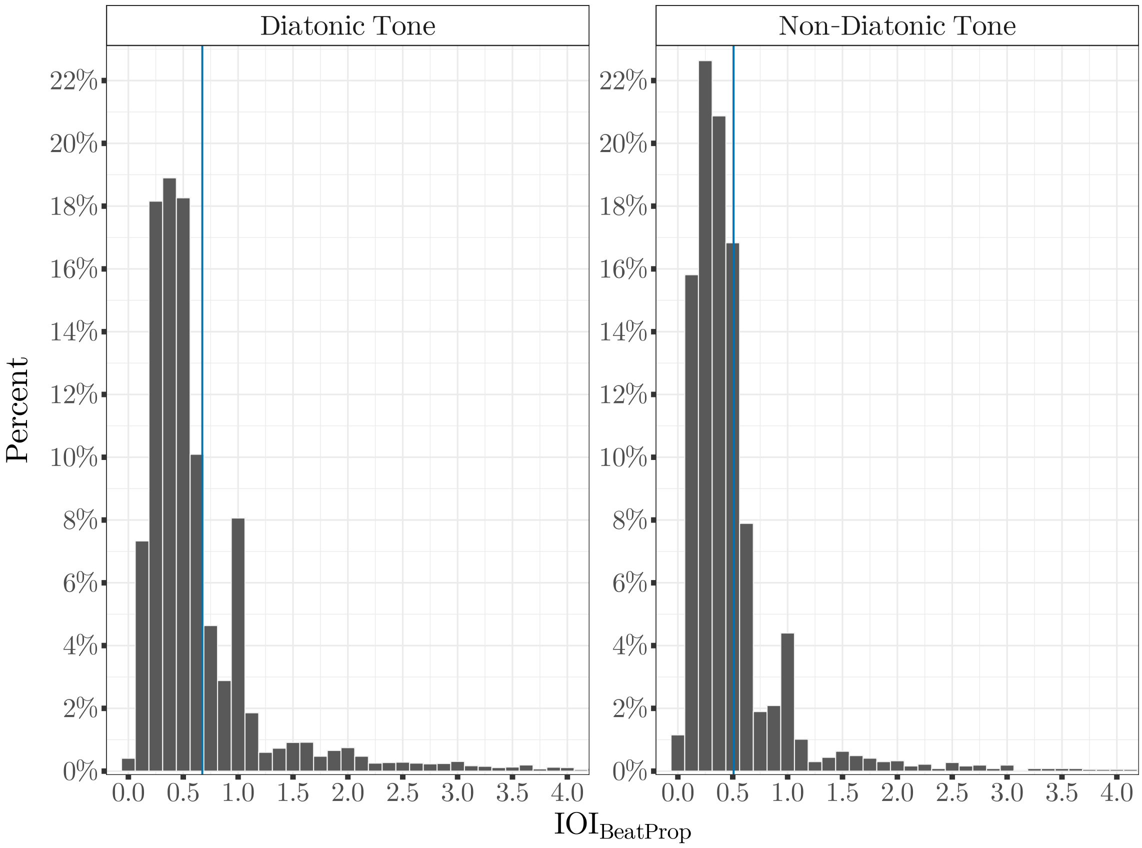

Figure 5.14: $\text{IOI}_{\text{BeatProp}}$ distribution for each TPC weight.

Table 5.5 shows for each NDT the number of times it was surrounded by zero, one, or two DTs. Nearly all (97.33%) of the NDTs Green played were proceeded or followed by a DT, with the vast majority being surrounded by DTs. A common approach for dealing with NDTs is the use of step-wise movement into or out of the tone, and it was hypothesised that Green utilised this approach in his improvisations. Figure 5.15 shows, for both DT and NDTs, the distribution of absolute intervals played into and out of the notes.[17] The data in this graph showed that Green strongly favoured step-wise, specifically chromatic movement, to move both into and out of NDTs. In total, Green used step-wise motion to move into a NDT 60.68% of the time, while moving out of a NDT step-wise 68.44% of the time. This was substantially more common than for DTs, where Green played step-wise motion into and out of the note 53.91% and 52.12% of the time respectively.

[16] An extreme outlier from a tremolo note in *Blues In Maude's Flat* (Green 1961b) with an $\text{IOI}_{\text{BeatProp}}$ of 22.58 beats was excluded from these calculations.

[17] Only absolute intervals of a perfect 5$^{\text{th}}$ or less were shown, containing 96.85% of all intervals (ascending or descending) played by Green.

Table 5.5: Distribution of TPC Weight transitions around a NDT.
D = Diatonic Tone, N = Non-Diatonic Tone.

	D-N-D	D-N-N	N-N-D	N-N-N
Count	2865	333	333	97
Percent	78.97%	9.18%	9.18%	2.67%

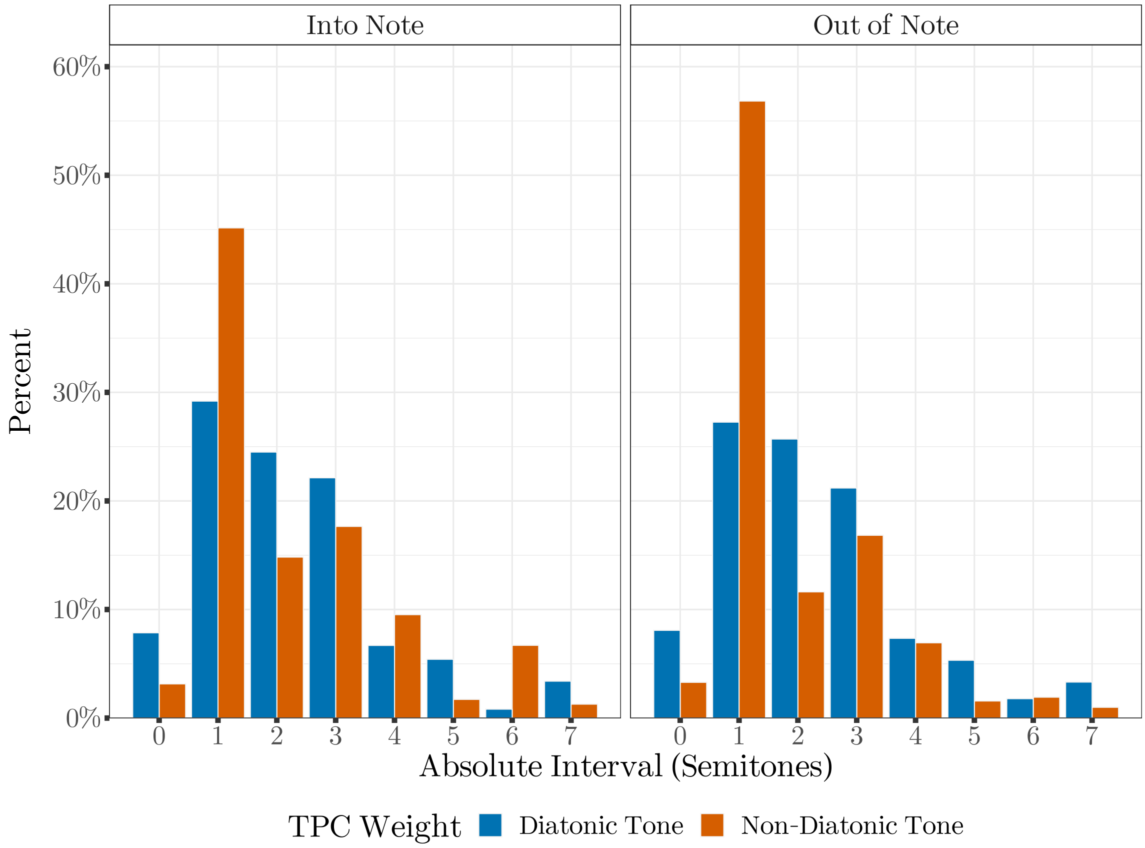

Figure 5.15: Distribution of absolute interval sizes (< 7 semitones) played into and out of a note for each TPC weight.

To compare the TPC weight in each situation, two χ^2-tests were run, finding significant differences in the distributions for the absolute intervals (< 7 semitones) played into ($\chi^2(7) = 1157.68,\ p =< .001,\ V = .24$) and out of ($\chi^2(7) = 1322.02$, $p =< .001,\ V = .26$) a note, both with a small effect size. These results supported the hypothesis that Green frequently used step-wise motion to move into and out of NDTs, with Green especially favouring chromatic transitions.

5.2.2 Tonal Pitch Class Summary

In summary, the majority of the notes that Green played in his improvisations were diatonic to the overall key of the piece. There was also evidence of blues influenced language in Green's improvisations. This was specifically notable in Green's frequent use of the ♭3 blues note in non-diatonic situations. Green also treated NDTs substantially different from DTs. Green rarely played three NDTs in a row, with the vast majority of NDTs being surrounded by DTs. Green also strongly preferred step-wise, specifically chromatic, movement to transition into and out of NDTs. There were limitations to the investigation of TPC and its relationship to improvisational style. TPC has limitations in its analytical ability as it cannot take into account transient modulations or the tonicisations of harmonies that moved outside the key centre. The CPC took these into account by comparing pitches to the nominal chord changes over which they were played.

5.3 Chordal Pitch Class

In comparison to the TPC, which compared all notes to the overall tonality, the CPC related the notes to the idealised chords they were played against (the chord of the moment).[18] While the TPC could be thought of as playing 'over the changes', the CPC reflected improvising 'through the changes'. Playing through the changes allows for the harmonic structure of a piece to be heard within an improvisation. Within *MeloSpy* there were two main features for analysing a note's relation to the chord of the moment: CPC, which labelled a note n as $n \in [0 : 11]$, with the root of the chord being 0, and the ♮7 being 11; and the chordal diatonic pitch class (with the extended version, CDPCX, used in this research, but referred to as the chordal diatonic pitch class[19]), which considered the diatonic scale of each chord type and classified the notes compared to that scale. These scales were: ionian for $\Delta 7$; mixolydian for 7; dorian for m7; locrian for ø7; and a tone-semitone 8-note diminished scale for ◦7. CPC was chord type agnostic – a CPC of 3 represented the note a minor 3rd above the root of the chord – while the CDPCX was chord type dependent – a CDPCX of 3 was the 3rd (♮ or ♭) for the chord of the moment. Additionally, the $\text{CPC}_{\text{Weight}}$ feature was created and frequently used in analyses. The $\text{CPC}_{\text{Weight}}$ grouped all notes into one of three categories: arpeggio tones; scale tones; and non-harmonic tones (NHT). Arpeggio tones and scale tones combined were harmonic tones (HT). The advantage of the $\text{CPC}_{\text{Weight}}$ was that it focused on the function of a note, rather than a specific CPC or CDPCX.[20] The CPC and CDPCX for the five most common chord types are shown in Table 5.6, with pitches compared to a C root.

[18] Most of the setup and data manipulation for chord-based features were completed through the `chordSetup` function, details of which can be found in Appendix C.

[19] The extended version of the chordal diatonic pitch class added variables for ♭2 and ♭6.

[20] For example, it was more informative to know the prevalence of arpeggio tones on metrically strong beats, rather than the specific note from the arpeggio. Where appropriate, the specific notes were also investigated. Similar to the overall methodology, this allowed the analysis to start broad, and then narrow down to focus on specific classes.

Table 5.6: Translation of CPC to CDPCX for Δ7, 7, m7, ø7, and ∘7 chords in reference to a C root.

		CDPCX				
	CPC	Δ7	7	m7	ø7[1]	∘7[2]
C	0	1	1	1	1	0
D♭	1	♭2	♭2	♭2	2	1
D	2	2	2	2	♭2 (♮2)	2
E♭	3	♭3/♮3	♭3/♮3	3	3	3
E	4	3	3	♮3/♭3	♮3/♭3	4
F	5	4	4	4	4	5
G♭	6	TT	TT	TT	5	6
G	7	5	5	5	TT (♮5)	7
A♭	8	♭6	♭6	♭6	6	8
A	9	6	6	6	♭6 (♮6)	9
B♭	10	♭7/♮7	7	7	7	10
B	11	7	♮7/♭7	♮7/♭7	♮7/♭7	11

[1] ♭2, TT, and ♭6 classes were treated as ‘altered’ 2, 5, and 6

[2] CDPCX encoding does not work for ∘7 chords, CPC encoding used instead

TT: Tritone

Within CDPCX some classes were only used for certain chord types, these were: ♭3/♮3; ♮3/♭3; ♭7/♮7; and ♮7/♭7. For example, in Δ7 chords the ♮3/♭3 and ♮7/♭7 classes could not exist, as ♭3 and ♭7 were not part of the chord structure. While the CDPCX worked for the most common chord types – Δ7, m7, and 7 – the table highlighted some issues when used over chords with more alterations. For ø7 chords the CDPCX levels were assigned so that the classes 1–7 reflected the notes of the locrian mode, with the classes ♭2, tritone (TT), and ♭6 used to reference the altered 2^{nd}, 5^{th}, and 6^{th} respectively. The discussion in this document referred to their actual value – ♮2, ♮5, and ♮6 – instead of their internally coded classes. As the CDPCX was designed for 7-note diatonic scales, the 8-note diminished scale could not be properly encoded, and therefore diminished chords were only analysed in reference to the CPC or $\text{CPC}_{\text{Weight}}$.

CDPCX and $\text{CPC}_{\text{Weight}}$ were preferred over the raw CPC values as it better allowed for comparison between the chord types. The CPC values were often used for statistical tests and when investigating a single chord type. The chordal pitch class analysis – through CPC, CDPCX, and CPC_{Weight} – provided insight into Green's treatment of notes in relation to their chords within the improvisation. It also provided insight into Green's treatment of NHTs, especially altered chord tensions over 7 chords, and allowed for investigation into the use of upper structure triads (UST).

5.3.1 Corpus Chordal Pitch Distributions

The distribution of chordal pitch descriptors for all the notes in Green's corpus are shown in Figure 5.16. The three descriptors shown are the CPC, CDPCX, and CPC_{Weight}. The CPC_{Weight} data indicated that 50.92% of the notes Green played came from the arpeggio of the surrounding chord, with 27.56% coming from the scale. This matched the data displayed in the CDPCX distributions, where the most frequently played notes were the tonic, 3rd, 5th, and 7th. The 4th, often considered a weak or diatonic non-harmonic tone over Δ7 and 7 chords (Crook 1991, 105), appeared prominently in both the CPC and CDPCX distributions. This could be explained by the distribution of the chord types, the metrical placement of the notes, as evidence of chordal anticipation or substitution, or simply an element of Green's improvisational style. The tonic and 5th were the two most commonly played tones; while the ♭7 was played more than twice as often as the ♮7. This was likely due to the only chord with a ♮7 as part of its arpeggio being Δ7. Considering more than half the chords in the corpus had a ♮3 as part of the arpeggio, Green playing ♭3 more often than ♮3 suggested the use of blues language in Green's improvisations, and possibly evidence of Green playing ♭9 over 7 chords. This grouped data, without consideration of the chord type or other features, provided only limited insight into Green's improvisational style. The following sections investigate the relationship of the chordal pitch to each chord type.

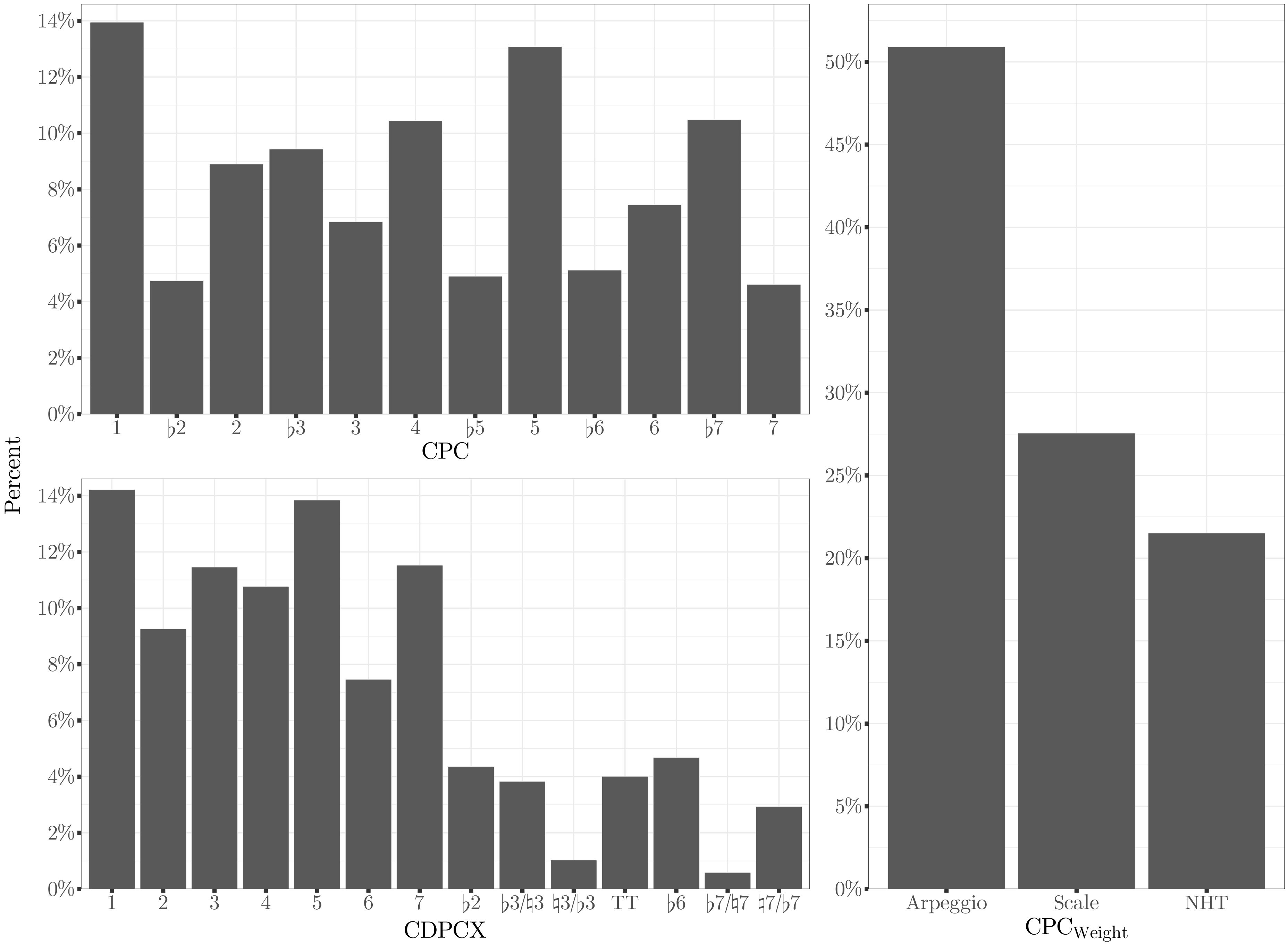

Figure 5.16: Distribution of chordal pitch descriptors for every note in Green's corpus. CDPCX excludes ∘7 chords.

5.3.2 Chord Type Chordal Pitch Distributions

The following sections investigated Green's improvisational style through analysis of his chordal pitch distributions over each of the five chord types present in the data. This was based on the hypothesis that Green's chordal pitch note choices differed depending on the nominal chords. A χ^2-test was run to investigate the relationship between the CPC distribution and the chord types. The results of this test found that Green played a significantly different distribution of CPC for each chord, with a medium effect size ($\chi^2(44) = 2986.39,\ p =< .001,\ V = .19$).[21]

Major Seven

Figure 5.17 shows Green's CDPCX distribution over $\Delta 7$ chords. The median number of notes Green played over a $\Delta 7$ chord was 5 (IQR: 3–8). Green's CDPCX distribution aligned closely with the $\text{CPC}_{\text{Weight}}$ distribution in Figure 5.16, with 55.58% of notes played from the arpeggio. Green played the 6^{th} at approximately the same rate as the 7^{th}. As the 6^{th} is a common substitution for the 7^{th} in a $\Delta 7$ chord, this suggested that Green treated them similarly over $\Delta 7$ chords. The most frequent notes came from the arpeggio, however there were both scale tones and NHTs that had a higher frequency than expected, including the: 4^{th}; ♭3/♮3; TT; and ♭7/♮7. These were indicative of blues language played over $\Delta 7$ chords.

In 35.63% of $\Delta 7$ chords in Green's corpus, he played only notes from the arpeggio or scale. In an additional 29.24% of chords, Green played only one NHT. In total, Green played ≤ 2 NHTs in 85.75% of $\Delta 7$ chords. Figure 5.18 shows examples of Green playing zero, one, and two NHTs over IV$\Delta 7$ chords in *Red River Valley* (Green 1962j). A correlation test found that the number of NHTs over a $\Delta 7$ chord was positively correlated with the total number of notes per chord, with a large effect size ($r = .54,\ t(405) = 12.97,\ p < .001,\ r^2 = .29$). This indicated that more metrically dense $\Delta 7$ chords were likely to have more NHTs. This analysis found that Green most frequently played arpeggio tones over $\Delta 7$ chords. Additionally, the higher than expected frequency of specific scale tones and NHTs indicated the use of blues language.

[21] Subsequent post-hoc tests found significant pairwise differences between all chord type comparisons at $p < .001$.

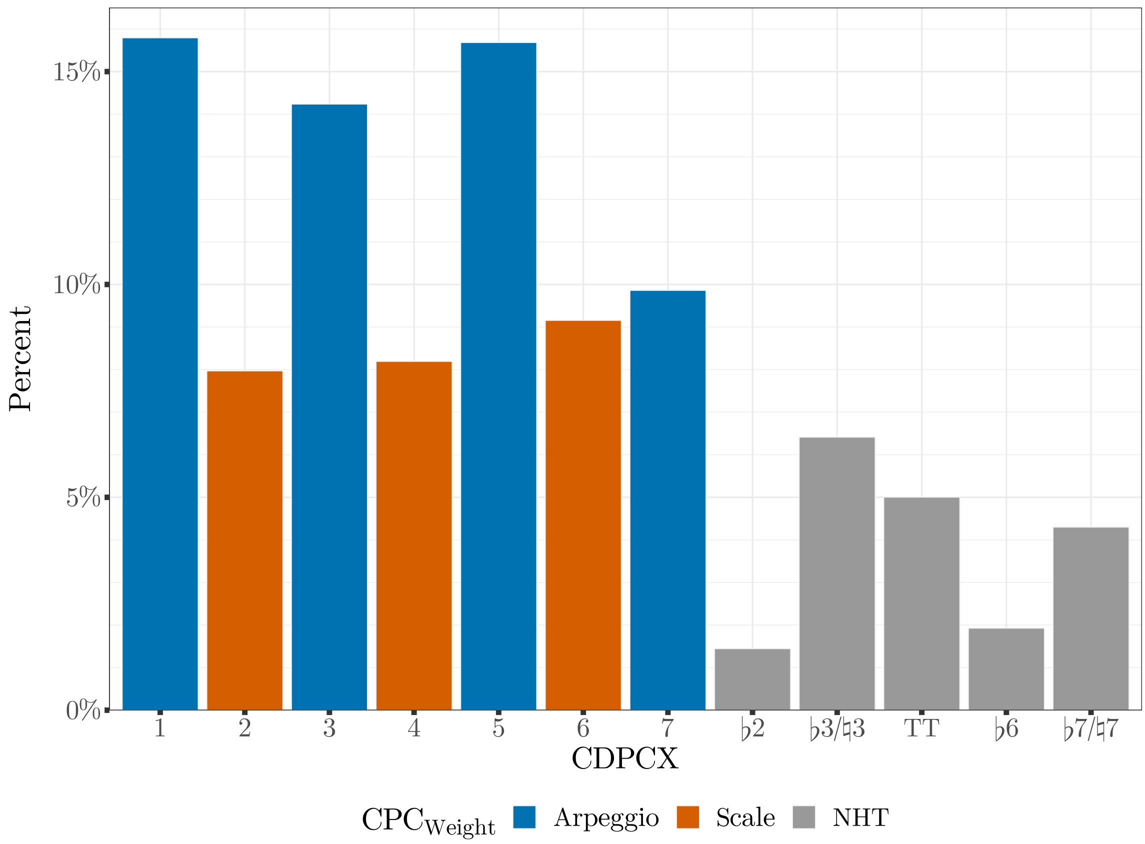

Figure 5.17: CDPCX distribution of Δ7 chords in Green's corpus.

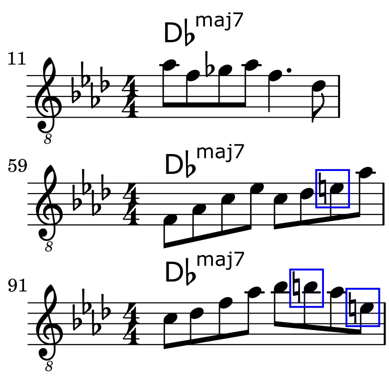

Figure 5.18: Green's playing over IVΔ7 (IV) chords in *Red River Valley* (1962). Bar 11, 0 NHTs; bar 59, 1 NHT; bar 91, 2 NHTs.

Minor Seven

Green's CDPCX distribution of m7 chords is shown in Figure 5.19. Similar to Green's Δ7 distribution, the most frequently played classes were the tones of the base arpeggio triad, although the 2nd and 4th scale tones were also played often. Green also played the 6th (13th) less frequently than either the ♭2 or TT NHTs. Green played a median of 5 notes (IQR: 3–8) over m7 chords. Overall, Green's CDPCX distribution indicated that he predominantly played diatonically over m7 chords.

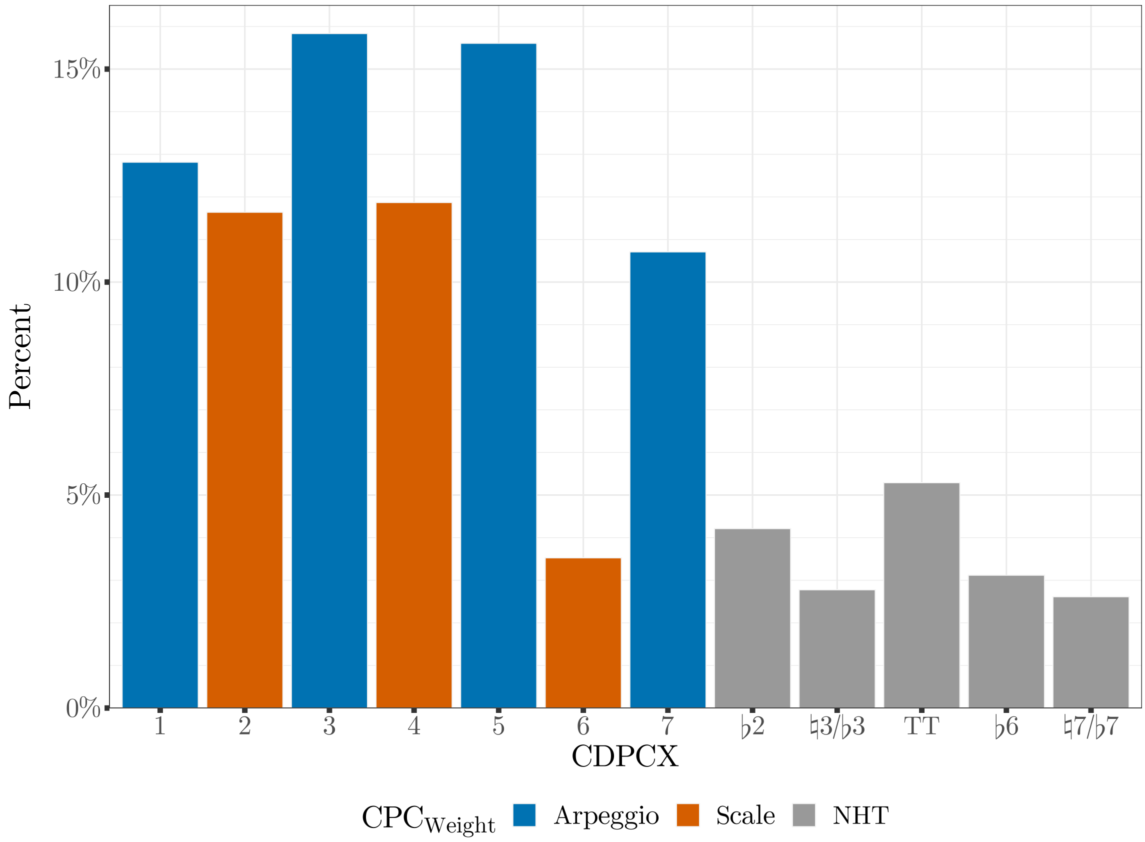

Figure 5.19: CDPCX distribution of m7 chords in Green's corpus.

Figure 5.20 shows a comparative distribution for the first note Green played over a m7 compared to the other notes played over the chord. The graphs showed a distinct preference for Green playing a 3rd as the first note. Of these, 75.12% were played on metrically strong beats. The second most common first note was the 4th, followed closely by the 5th. A hypothesis for the high frequency of the 4th in the m7 CDPCX distribution was that in a ii–V Green was either substituting the m7 with a 7, or resolving early to the 7. A 7 was the most frequent chord to occur after a m7 chord, occurring 76.29% of the time. Of those, 87.60% had a dominant, ii–V, resolution. In total, 66.83% of m7 chords in Green's corpus resolved to a 7 through a ii–V. To investigate these hypotheses the CDPCX distribution of m7 chords prior to a 7 chord (in a ii–V) were compared against both chord types.

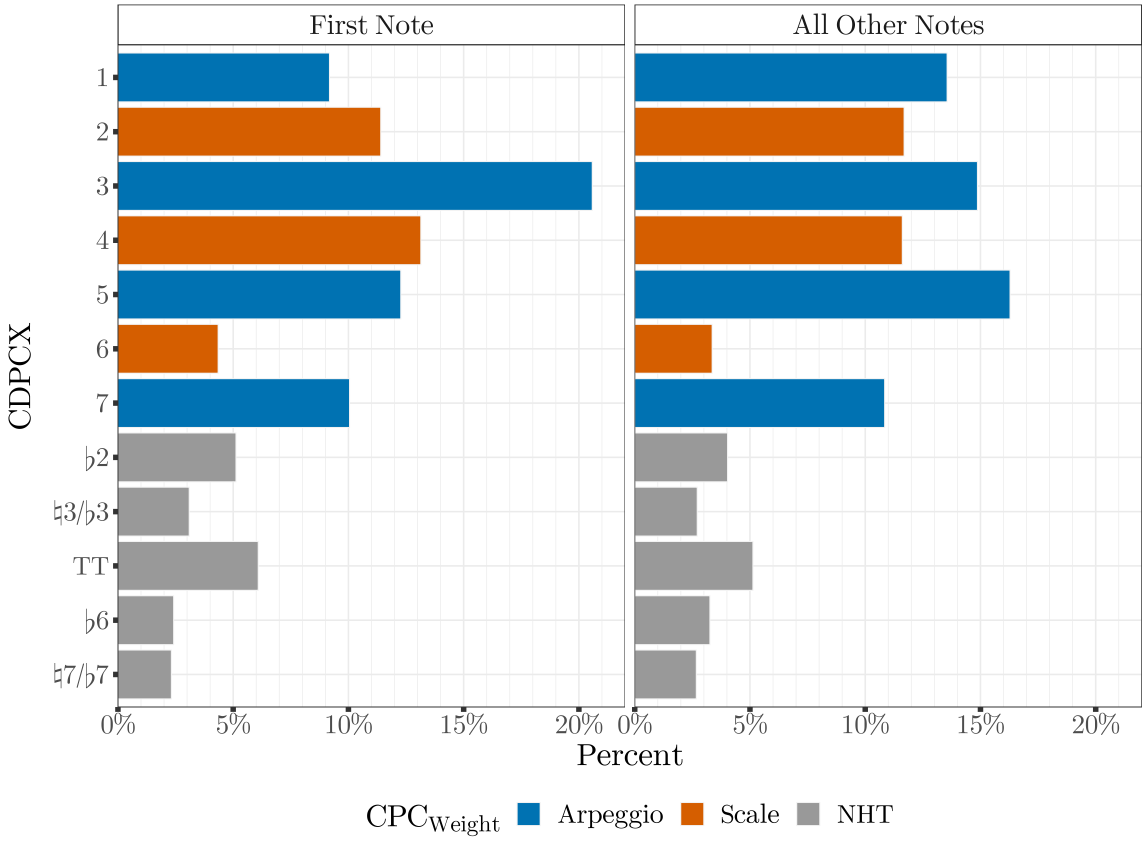

Figure 5.20: Green's distribution of CDPCX for the first note of a m7 chord vs. all other notes in the chord.

The graphs in Figure 5.21 show both these CDPCX distributions.[22] Due to the relationship between a m7 and 7 chord in a ii–V there were many overlapping classes, which hid differences between the distributions. For example, the high frequency of the 7$^{\text{th}}$ in 7 distribution of Figure 5.21 may have indicated a substitution; however, this was the equivalent of the 3$^{\text{rd}}$ of the m7, so provided no substantial evidence. Whereas, the very low frequency of the 3$^{\text{rd}}$ in the 7 distribution was indicative of Green not tending to substitute the m7 for a 7.

[22]Figure A.3 in Appendix A shows a similar comparative distribution, but only for the beat before the chord change. These distributions were not substantially different from Figure 5.21, and suggested that Green did not frequently anticipate the 7 resolution.

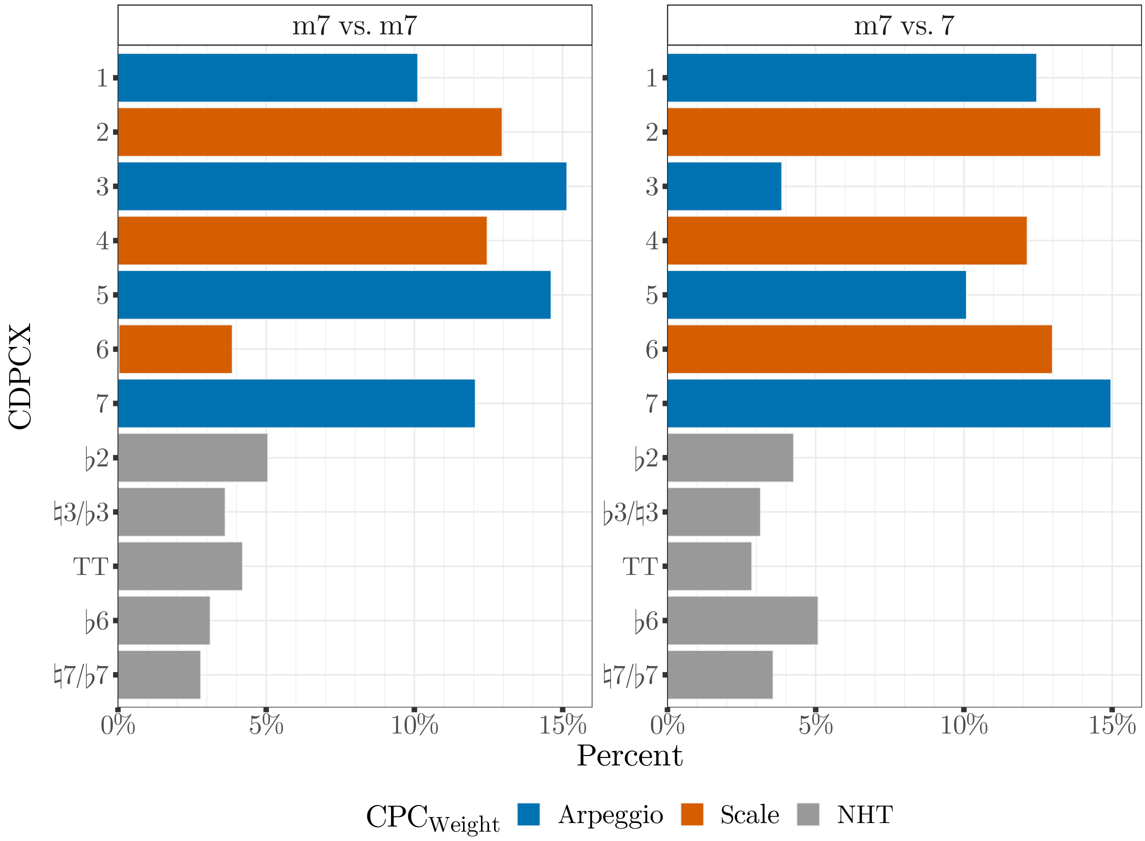

Figure 5.21: m7 CDPCX distribution compared against a m7 and 7 substitution when the following chord was a 7.

This was not to say that Green never substituted a 7 for a m7 chord. For example, Figure 5.22 shows two likely examples of when Green did as such. The top example, from Green's improvisation over *I'll Remember April* (Green 1961k) shows Green playing a descending D mixolydian scale over the Am7, starting on the 3rd of the D7, followed by the 13th, tonic, and 9th over the D7.[23] The other example is from Green's first solo over *Miss Ann's Tempo* (Green 1961m), where Green played an F7 arpeggio over the Cm7 chord, starting on the 3rd, followed by altered dominant language over the F7 chord.

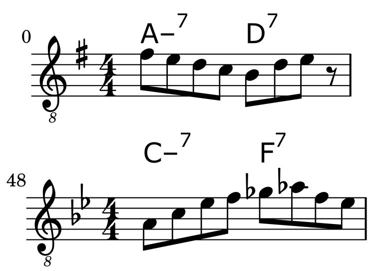

Figure 5.22: Examples of Green substituting a 7 for a m7 chord. Top: *I'll Remember April* (1961), bar 0. Bottom: *Miss Ann's Tempo* (Solo 1, 1961), bar 48.

[23]Bar 1 was always assigned to the first bar of the form, therefore any bars that occurred before this, for example at the end of the head as a pick up into the top of the form, were labelled backwards from one, e.g. 0, -1, -2, etc.

It was difficult to ascertain whether a specific example did show Green substituting or resolving early due to the close relationship between the iim7 and V7. The examples above were found manually, and not based on the distribution of the data. In future research it may be possible, with a larger dataset and a refined approach, to undertake broader searches for tell-tale markers of substitution and early resolution to 7 chords, and to better spot them in a large corpus of data. Green's data did not suggest that he frequently substituted or played dominant language from an accompanying 7 chord. The data did indicate that most of Green's notes came from the first five scale tones of the m7, as well as the 7$^{\text{th}}$, with Green preferring to play the 3$^{\text{rd}}$ as the first note of the chord.

Dominant Seven

Figure 5.23 shows the distribution of CDPCX Green played over 7 chords. The note density of 7 chords in Green's corpus was the same as m7 chords, with a median of 5 (IQR: 3–8) notes per chord. Compared to the Δ7 and m7 distributions, Green played a substantially higher proportion of NHTs (24.24% vs. just under 20%). Green also played fewer arpeggio tones, with this predominantly due to Green avoiding the 3$^{\text{rd}}$. The higher rate of NHTs over 7 chords was expected, as alterations of a 7 chord are common in jazz.

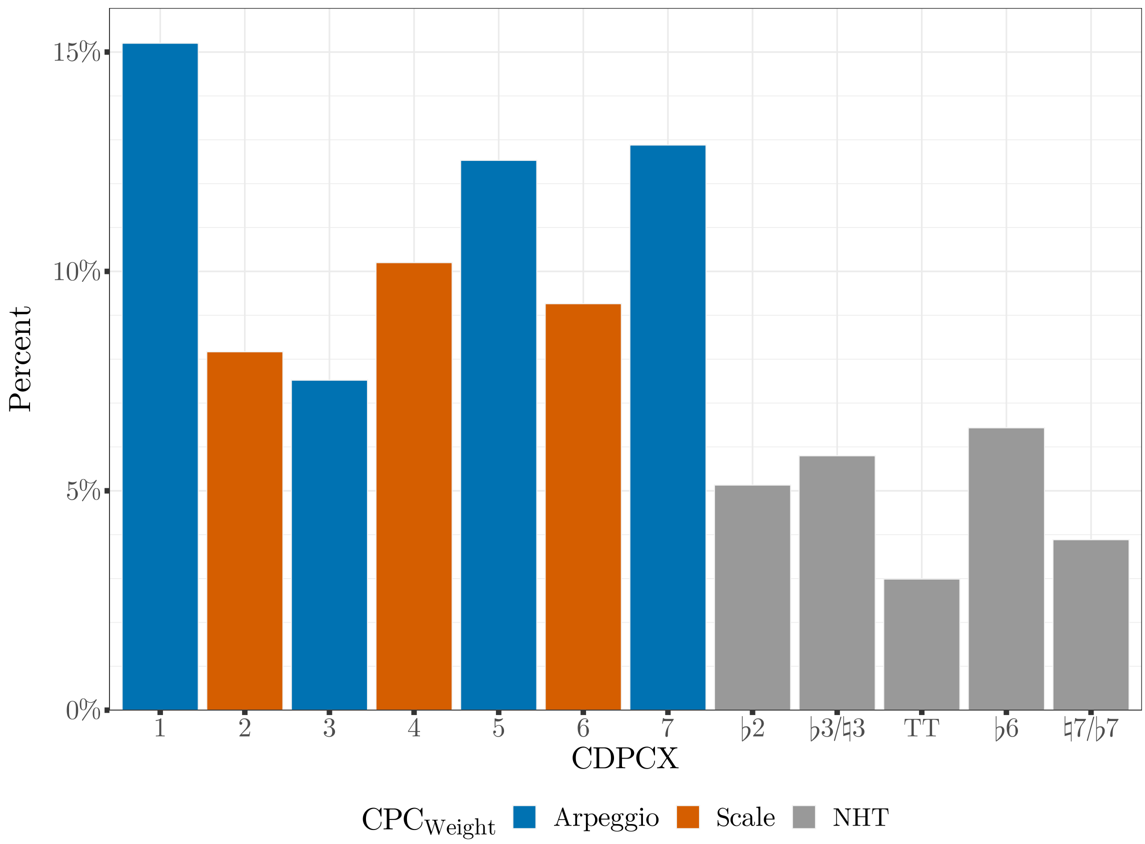

Figure 5.23: CDPCX distribution for 7 chords in Green's corpus.

Due to the multi-faceted roles that a 7 chord can occupy, it was hypothesised that Green's CDPCX distribution may have differed depending on the tonality mode of the improvisation. This was supported by a χ^2-test that found a significant difference in the CDPCX distributions of 7 chords between the tonality modes, with a small effect size ($\chi^2(22) = 231.06,\ p =< .001,\ V = .11$). The distributions for the major, minor, and blues tonality modes can be seen in Figure 5.24. Although the tonic, 5$^{\text{th}}$, and 7$^{\text{th}}$ were the most common notes in each tonality mode, the NHTs differed more. The most common NHTs for each tonality mode were ♭2 over the minor, ♭3/♮3 over the blues, and ♭6 over the major tonality mode.

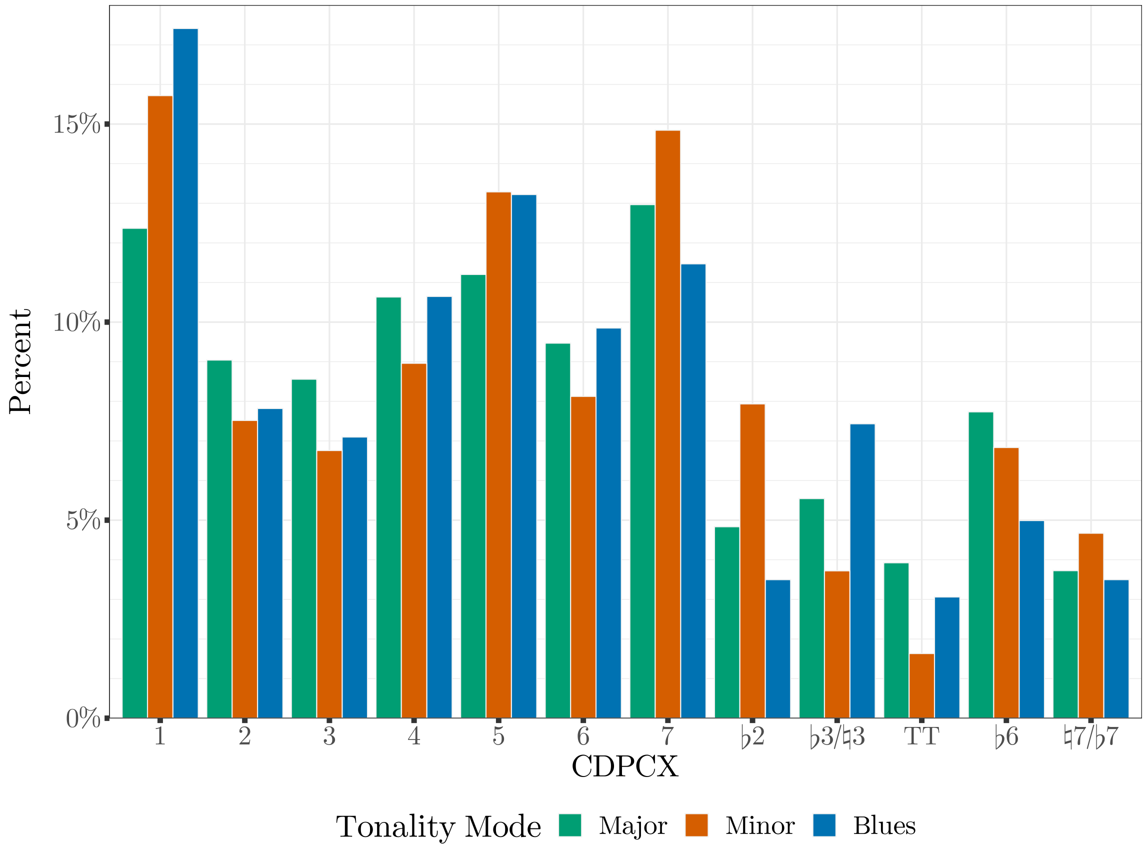

Figure 5.24: Dominant 7 CDPCX distribution for each tonality mode in Green's corpus.

The ♭6 (♯5) in major tonalities is a common alteration of a 7 chord, especially when resolving to a Δ7, where it is a semitone below the 3$^{\text{rd}}$ of the Δ7. In cases when the ♭6 was the last note before a Δ7 chord, 47.37% of the time Green followed this note with the 3$^{\text{rd}}$ of the Δ7 chord. An example of this can be seen in Figure 5.25, which shows an excerpt from Green's improvisation over *The Surrey With The Fringe On Top* (Green 1963g).

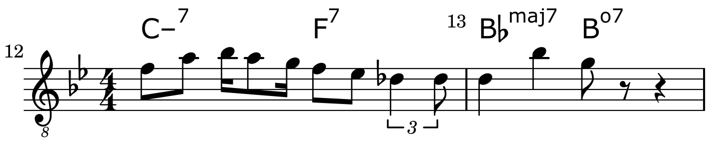

Figure 5.25: Example of a resolution from the ♭6 (♯5) of a 7 chord to the 3rd of a Δ7 chord, *The Surrey With The Fringe On Top* (1963), bars 12–13.

In a minor tonality the ♭2 (♭9) of a 7 chord is a standard alteration for a minor ii–V (iiø7 – V7$^{\flat 9}$ – i). An example of this can be seen in Figure 5.26, from Green's improvisation over *Green With Envy* (Solo 2, Green 1961j). Green starts on the ♭9, descends predominantly step-wise to the ♮3 to start a 3-5-♭7-♭9 arpeggio that resolved to the 5 of the B♭m7.

Figure 5.26: Example of a resolution from the ♭2 (♭9) of a 7 chord to the 5th of a m7 chord, *Green With Envy* (Solo 2, 1961), bars 14–15.

Dominant 7 chords in a blues were unlike those in major and minor tonalities, where the 7 chord had primarily a dominant function, often as part of a ii–V In a blues there were four different 7 functions[24]: tonic – I7; sub-dominant IV7; dominant – V7; and a secondary dominant – V7/ii. Only the latter two of these 7 chords functioned as dominant chords with a dominant resolution.[25] Figure 5.27 shows Green's CDPCX distribution for each 7 function in the blues harmony, with the colours indicating the tones that come from the tonic blues scale and additional notes from the relative minor blues scale.

The data continued to find evidence of Green's use of blues language, with Green frequently playing the ♭3. The only situations where the 3rd of the 7 chord was frequently played was when it also fit with the relative minor blues scales, specifically the I7 and IV7. A feature of Green's improvisations over a blues was that he did not frequently play the ♭5 blues note (TT in I7; ♭2 in IV7; ♮7/♭7 in V7; and 6 in V7/ii), with it consistently being the least frequently played note from either blues scale.

Overall, this analysis into Green's CDPCX distribution over 7 chords found that the most frequent tones were from the arpeggio. Green's least frequent HT was the 3rd, compared to Δ7 and m7 chords where the 3rd was very common. Compared to the Δ7 and m7 distributions, Green also played a higher proportion of NHTs over 7 chords. The analysis also found that Green treated 7 chords differently depending on the tonality mode of the improvisation, and for blues, the function of the 7 chord.

[24] In the blues progression used as the nominal chord changes in this research.

[25] Technically, the I7 in bar 4 as, iim7–V7/IV, did act as a dominant resolution, but most I7 chords in a blues did not.

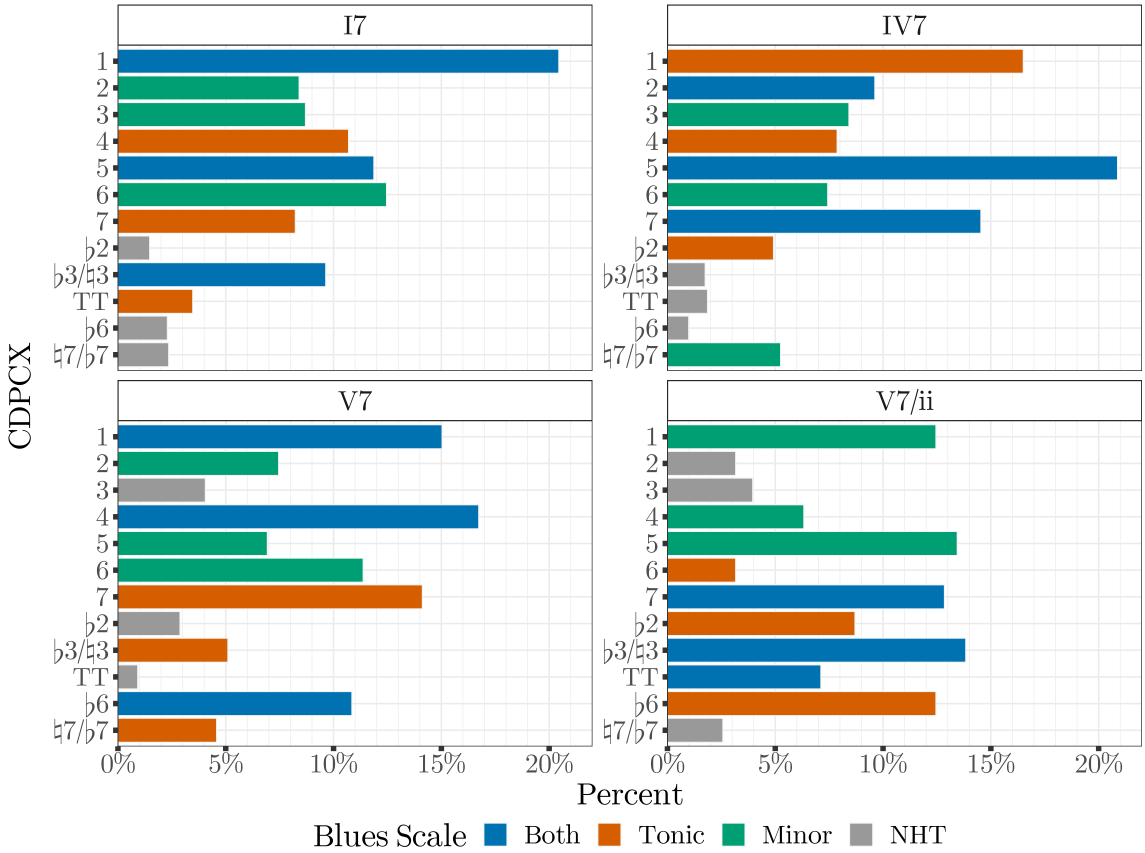

Figure 5.27: CDPCX distribution of each 7 function in a blues harmony.

Half-Diminished Seven

Within Green's corpus ø7 were the second least frequent of the five chord types, with only 150 occurrences in which Green played at least one note. The median number of notes that Green played over each ø7 chord was was 5 (IQR: 3–7), with the CDPCX distribution of these notes shown in Figure 5.28.

The two most common tones Green played over ø7 chords were the 4^{th} and 3^{rd} (approximately 18%); while the 5^{th} (♭5), 6^{th} (♭6), and 2^{nd} (♭2) were all played around 10% of the time. Similar to the m7 chords, the high occurrence of the 4^{th} could suggest that Green anticipated or substituted for V7 chord. In Green's corpus, 85.12% of ø7 chords resolved dominantly to a 7 chord, as part of a minor ii–V. The most frequent notes, if analysed against the V7 chord, would be the tonic, ♭7, ♭9, and ♯9, which fit well with a $V7^{(\flat 9)}$ chord. This hypothesis was supported by the data shown in Figure 5.29, which shows the ø7 CDPCX distribution in the beat before a V7 chord compared against the 7 chord.

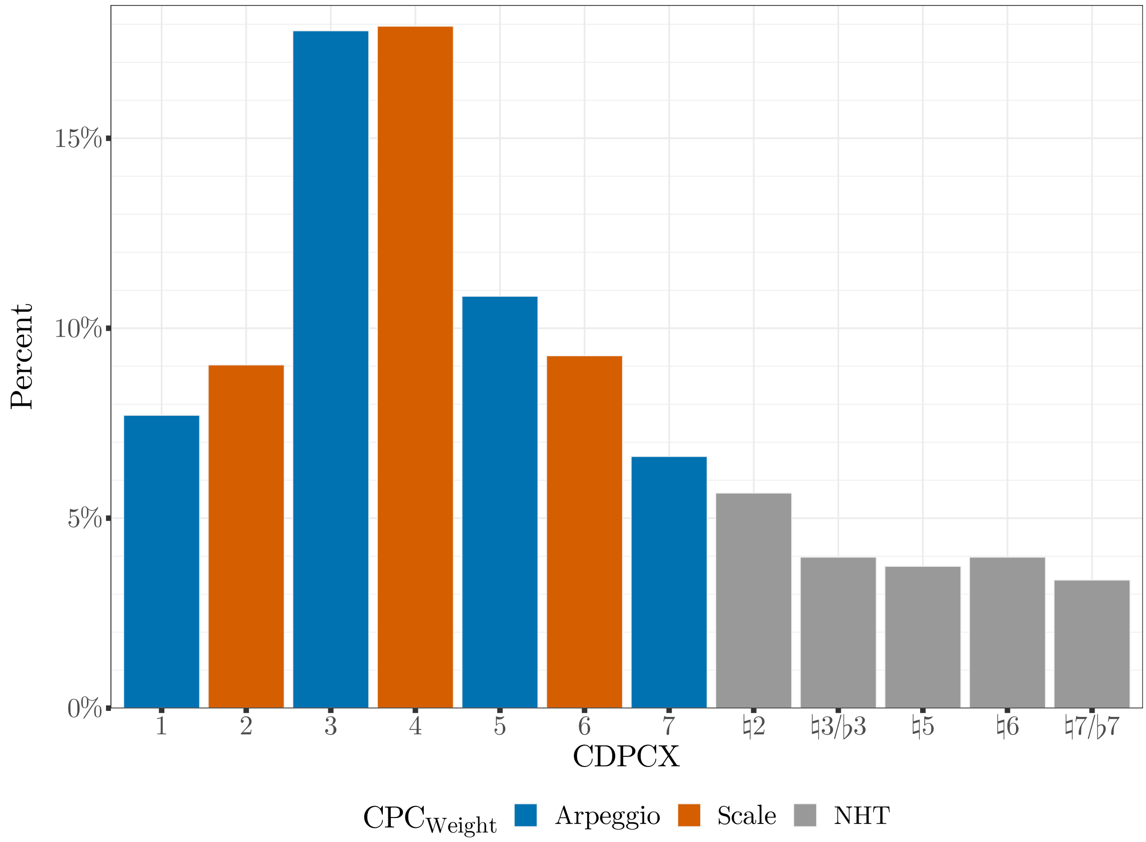

Figure 5.28: CDPCX distribution for ø7 chords in Green's corpus.

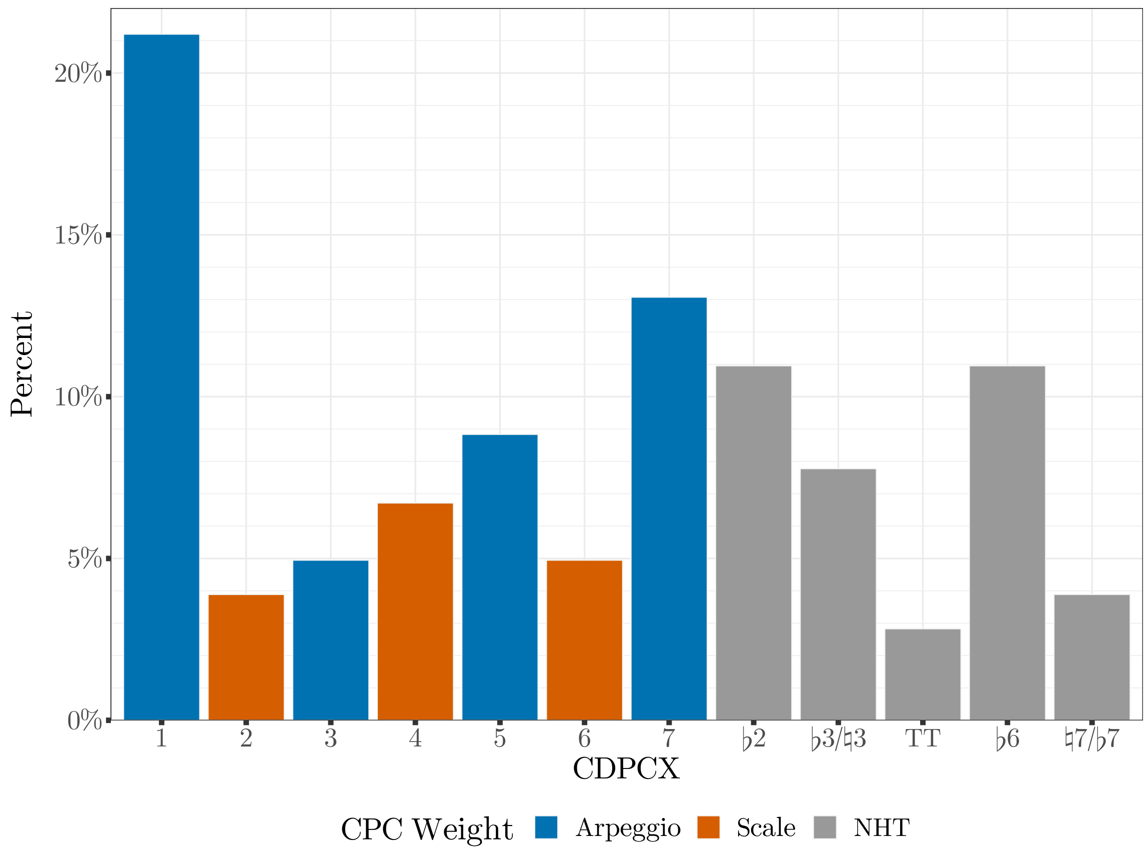

Figure 5.29: CDPCX distribution of ø7 chords in the beat before a dominant resolution to a V7 chord, compared to the 7 chord.

The ø7 distribution in the beat before a resolution to the 7 suggested that Green was anticipating the resolution. The most frequent notes were from a 7^{alt} chord; the tonic, 7^{th}, ♭9, and ♯5. An example of Green anticipating the 7 resolution can be seen in Figure 5.30. In this example, from Green's solo over *Brazil* (Solo 1, Green 1962b), Green outlines the dominant $\mathrm{E7}^{(\flat 9)}$ in the second bar of the Bø7, a bar before the nominal resolution.

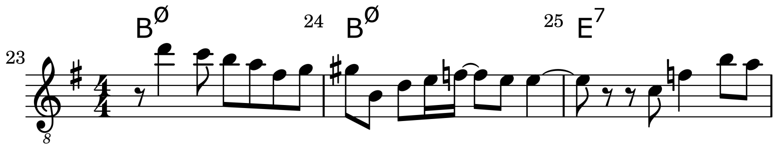

Figure 5.30: Example of an anticipation of a V7 chord over a ø7 chord in a minor ii–V pattern, *Brazil* (Solo 1, 1962), bars 23–25.

This analysis found evidence of Green substituting or anticipating the dominant 7 resolution when improvising over a ø7 chord. This was in contrast to the m7 chord, as part of a major ii–V, where no evidence of anticipation was found.

Diminished Seven

The final chord type present in the data was ∘7, of which there were 129 in which Green played at least one note. Over each ∘7 Green played a median of 6 notes (IQR: 3–8). Figure 5.31 shows the CPC distribution Green played for all ∘7 chords in the corpus. The tonic (1) was taken from the root of the ∘7 chord written in the harmony, but due to the cyclical nature of ∘7 chords any of the arpeggio tones could be equally considered the tonic. The distribution of notes in a ∘7 was difficult to analyse, both because of the cyclical nature of the chord, and due to ∘7 chords rarely functioning as a diminished chord in functional jazz harmony. Instead, they are frequently used as a chromatic passing chord, i.e. as a substitution for an altered 7 chord.

In Green's improvisations, 47.33% of the notes played over a ∘7 chord came from the arpeggio, a further 29.39% from the scale, and the remaining 23.28% from NHTs. These NHTs were mainly the ♭2 and ♮3, both of which could be considered altered tones of a 7 chord or from the scale of the surrounding key centre. Figure 5.32 shows an excerpt from Green's improvisation over *Green With Envy* (Solo 2, Green 1961j), in which he plays a descending diminished scale starting from the 6^{th}, with only a single NHT (a ♭2 E♭, instead of an E♮). Here, the D∘7 is a substitution for a $\mathrm{B}\flat 7^{(\flat 9)}$ in the progression D♭Δ7–D∘7($\mathrm{B}\flat 7^{(\flat 9)}$)–E♭m7 – A♭7.

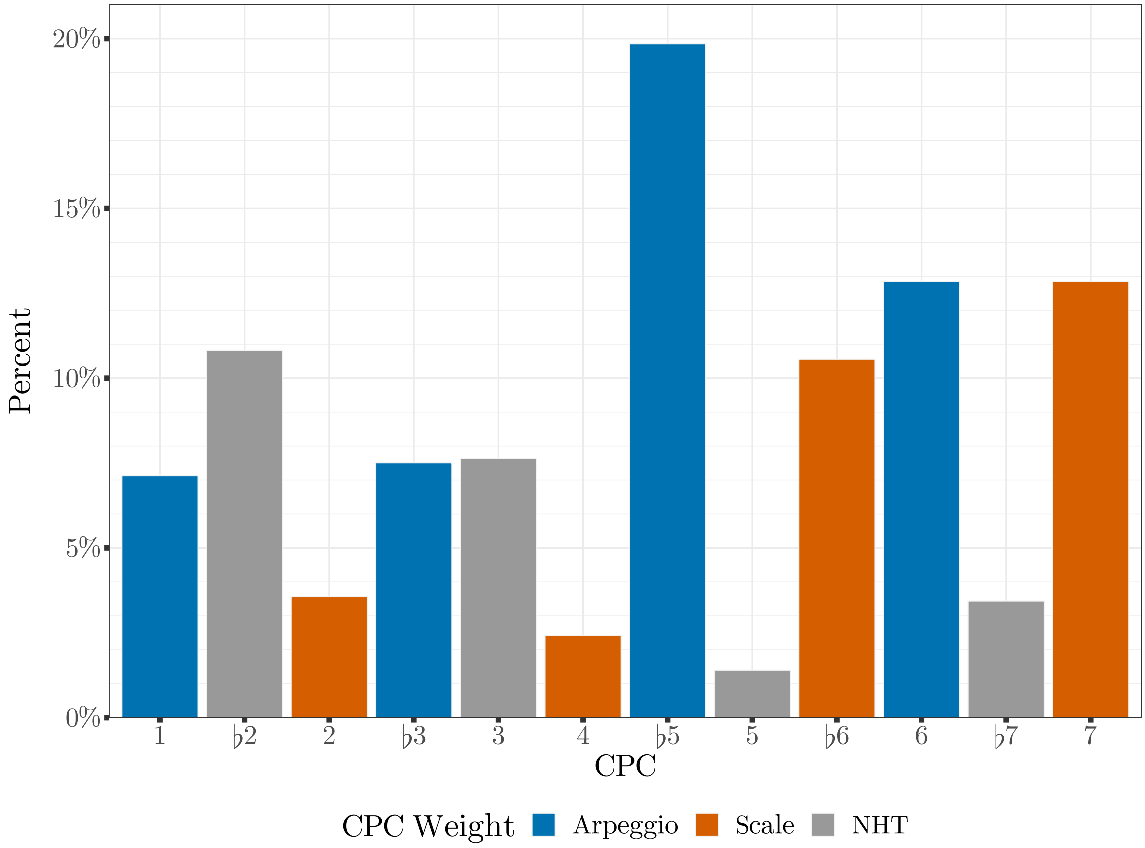

Figure 5.31: CPC distribution for ∘7 chords in Green's corpus.

Figure 5.32: Example of a diminished line, *Green With Envy* (Solo 2, 1961), bar 44.

While limited in scope, this analysis into Green's CPC over ∘7 chords found around half of the notes came from the arpeggio, with three-quarters being HTs. The most frequent NHTs, ♭2 and ♮3, could both be considered scale tones from an implied altered 7 chord or altered chord tones.

Chord Type Summary

This analysis investigated Green's chordal pitch distributions across the five chord types present in the data. The results of these analyses found that the majority of Green's notes came from the arpeggio of the chord, with only a small proportion of NHTs. NHTs were most common over 7 chords, where he frequently played the ♭9, ♯9, and ♯5. Across all chord types (except ∘7) Green played the 4th more frequently than may have been expected for a diatonic non-harmonic tone (Crook 1991, 105). For m7 and ø7 this may have been due to Green substituting or anticipating a dominant 7 resolution. While evidence of this was found for the minor ii−V (ø7), the m7 results were inconclusive. Therefore, the high frequency of the 4th was likely a combination of the blues, which was frequently used throughout Green's improvisations, and an element of Green's improvisational style.

The previous section investigated Green's chordal pitch distribution of notes for each chord, without much consideration of other musical elements. The following sections focused on analyses of Green's chordal pitch distributions against other improvisational elements, including: the metrical weight; the chord weight[26]; Green's treatment of NHTs, specifically dominant 7 altered chord tensions; and Green's use of USTs. These investigations focused on the three most common chord types – Δ7, m7, and 7 – with the CPC_{Weight} feature used for most of the analyses.

5.3.3 CPC_{Weight} vs. Metrical Weight

The following analysis investigated how the CPC_{Weight} differed between the metrical weights for the three common chord types. It was hypothesised that the HTs would be played more frequently on the beat, with a particular preference for arpeggio tones played on metrically strong beats. Consequently, it was hypothesised that NHTs would not frequently be played on the beat, with Green preferring to play them off-beat. Finally, the metrical weight distribution of the 4th for each chord types was analysed, to investigate the high frequency of the tone within Green's improvisations.

The distribution of CPC_{Weight} for each metrical weight and chord type is shown in Figure 5.33. The Δ7 and 7 distributions follow a similar trend, with the proportion of arpeggio tones decreasing as the metrical weight moved from metrically strong beats, to on-beat metrically weak beats, with the lowest proportion for off-beat notes. Both chord types also saw an increase in the proportion of scale tones when moving from metrically strong to weak beats. As hypothesised, the proportion of NHTs was highest for off-beat notes. In contrast, Green played the highest proportion of m7 arpeggio tones on metrically weak beats. Although the difference was smaller, Green did play more m7 NHTs off-beat than either of the on-beat metrical weights.

[26]Chord weight measured the proximity of a note to a previous or upcoming chord change.

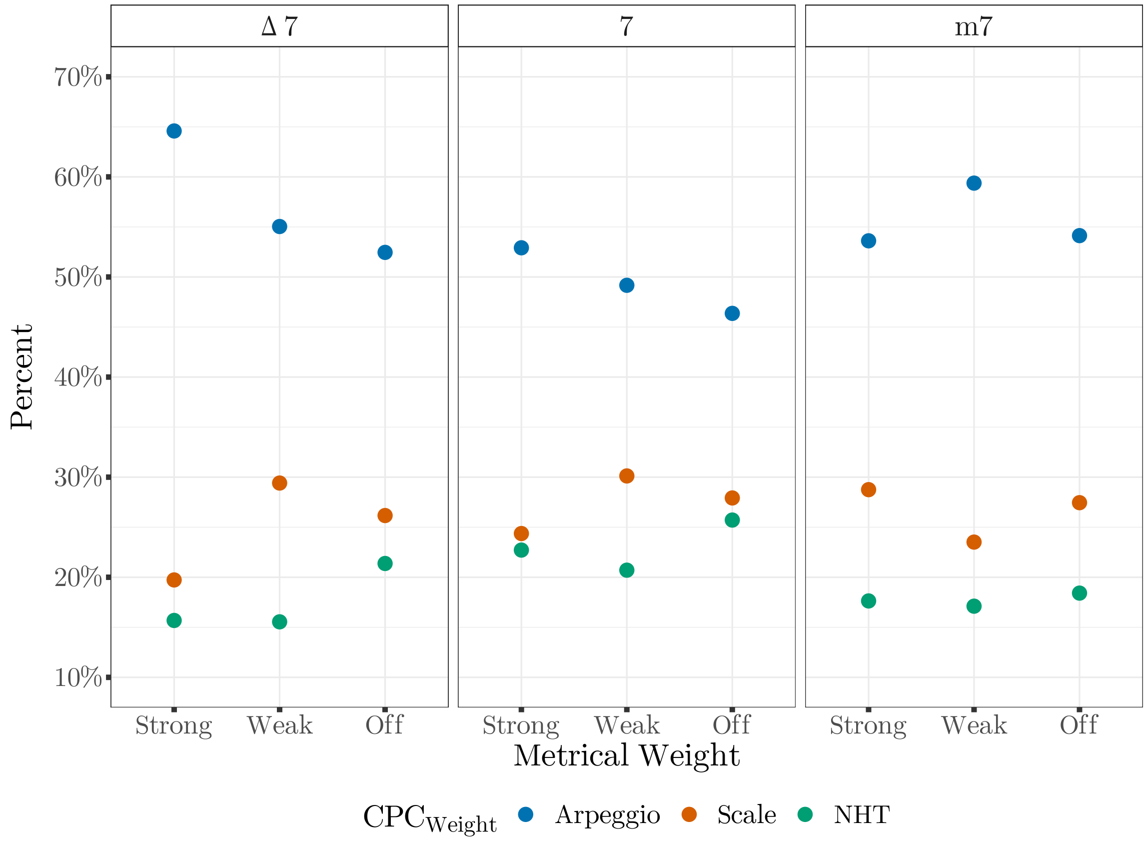

Figure 5.33: Distribution of $\text{CPC}_{\text{Weight}}$ for each metrical weight for Δ7, m7, and 7 chords.

A χ^2-test was run for each chord type separately to investigate the relationship between the $\text{CPC}_{\text{Weight}}$ and metrical weight distributions. A significant relationship was found for each chord type, each with a small effect size.[27] These results supported the hypothesis that Green's note choice, as $\text{CPC}_{\text{Weight}}$, was influenced by the metrical weight on which the note was played. Additionally, the results supported the hypothesis that Green was more likely to play NHTs off-beat.

The other metrical weight investigation focused on whether the high frequency of the 4$^{\text{th}}$ in Green's improvisations could be partly explained by the metrical weight. Specifically, the analyses focused on whether Green was treating the 4$^{\text{th}}$ more similar to a NHT, playing it more frequently on metrically weak beats or off-beat. The metrical weight distribution of the 4$^{\text{th}}$ for each chord type is shown in Figure 5.34. This graph used conditional odds to measure the frequency of the 4$^{\text{th}}$ in comparison to the overall metrical weight distribution for each chord type. The data in this graph showed that fourths had approximately equal odds of being played off-beat as any other note Green played. For Δ7 and 7 chords, fourths had lower odds of being played on metrically strong beats while having higher odds to be played on metrically weak beats. Only over m7 chords did fourths have even odds of being

[27] Δ7: $\chi^2(4) = 33.65$, $p =< .001$, $V = .08$; 7: $\chi^2(4) = 42.23$, $p =< .001$, $V = .05$; m7: $\chi^2(4) = 12.34$, $p = .015$, $V = .03$.

played by Green on metrically strong beats. The 4th is not generally considered to be a weak tone over m7 chords, which could explain its more frequent use on metrically strong beats. These results indicated that Green most frequently played the 4th off the beat. For Δ7 and 7 chords, Green was more likely to play it on metrically weak beats instead of strong beats. These results suggested that Green often treated the 4th as a diatonic non-harmonic tone (Crook 1991, 105), more similar to a NHT than a scale tone.

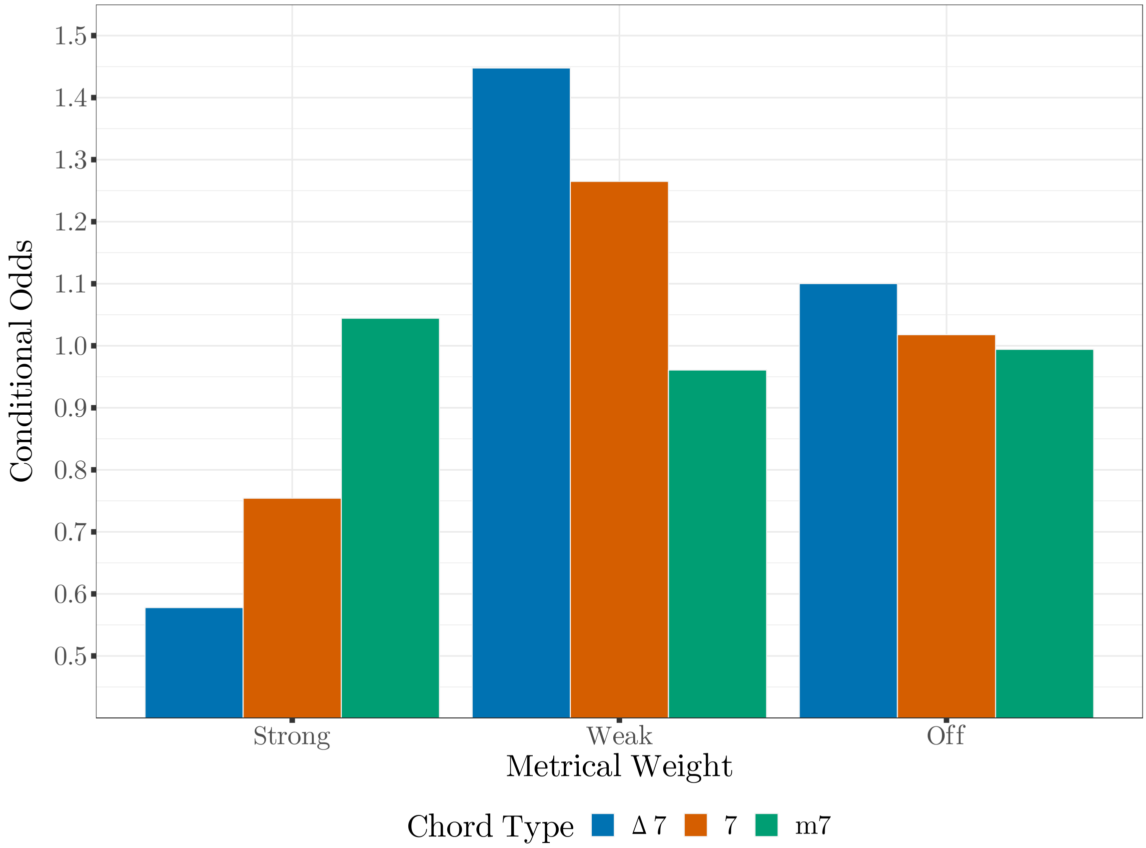

Figure 5.34: Conditional odds of metrical weight distribution of the 4th for Δ7, m7, and 7 chords.

5.3.4 $\text{CPC}_{\text{Weight}}$ vs. Chord Weight

To categorise a note's placement in relation to the surrounding nominal chord changes two new features were created. The first, chord weight, was a categorical variable with five classes (in order of priority):

- On beat – the beat of a chord change, including any off-beat notes played in the beat of a chord change;
- Beat between – the beat between two chord changes;
- Beat after – the beat after a chord change;
- Beat before – the beat before a chord change;
- No beat – no chord changes in the beat before or after the current beat.

The second feature was a numerical variable and counted the number of beats before a chord change. Notes played in the beat before a chord change had a value of one, with the count resetting on the beat of the chord change (there was no 0 value). The distribution of $\mathrm{CPC}_{\mathrm{Weight}}$ for the three chord types is shown in Figure 5.35, with the data split by whether or not the note was played in the beat before a chord change.

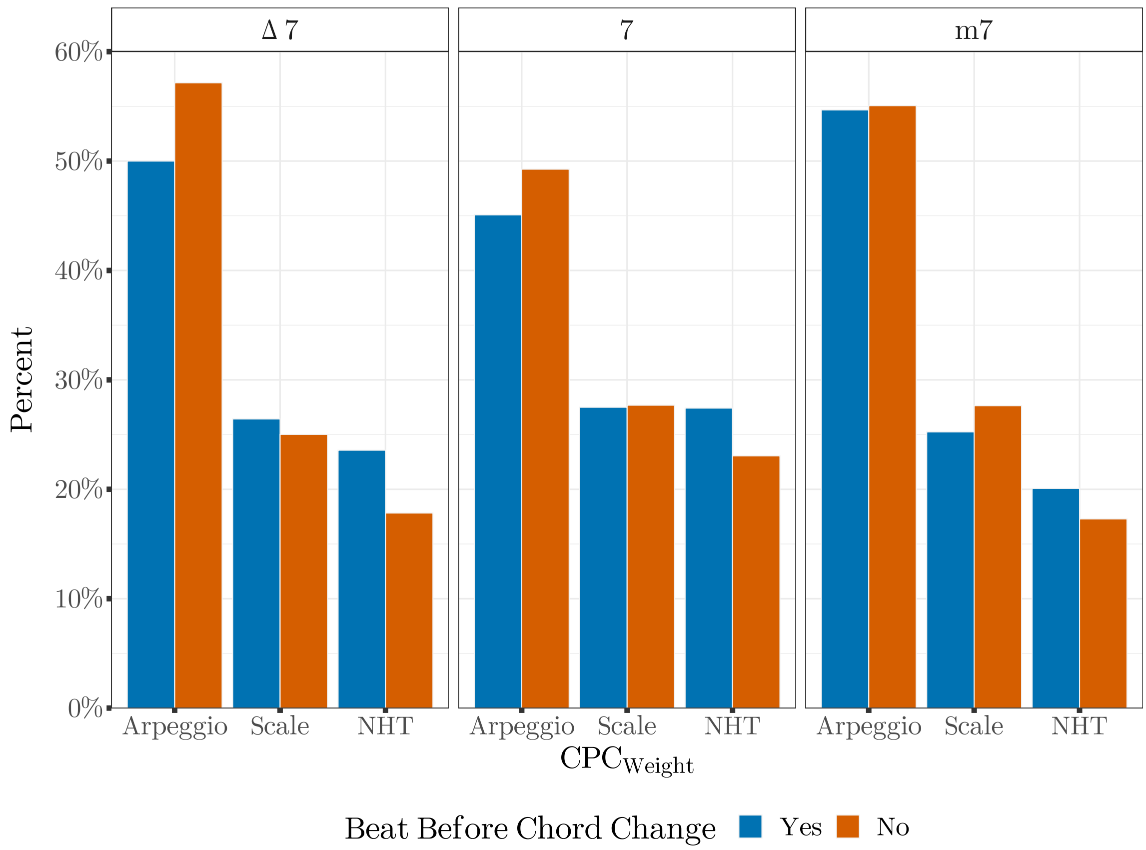

Figure 5.35: $\mathrm{CPC}_{\mathrm{Weight}}$ distribution of Δ7, 7, and m7 chords depending on whether or not they were played in the beat before a chord change.

This data showed that in the beat before a chord change, for all chord types, Green played more NHTs and fewer arpeggio tones. The only HT class where an increased frequency of notes played in the beat before a chord change was observed were scale tones over a Δ7 chord. The increase of NHTs may have been linked to an increase of chromatic intervals played by Green in the beat before a chord change. A χ^2-test found a significant relationship between whether a note was played in the beat before a chord change and if it moved chromatically, with a small effect size ($\chi^2(1) = 8.21,\ p = .004,\ V = .02$). However, an increase in chromatic intervals did not necessarily lead to an increase in NHTs. The results of this analysis indicated that Green played a higher proportion of NHTs in the beat before a chord change.

5.3.5 Treatment of Non-Harmonic Tones

The following section investigated Green's treatment of NHTs. It first analysed Green's broad treatment, focusing on the surrounding intervals and the duration of the NHTs. Specifically, the intervals used to transition into and out of NHTs were analysed, with a hypothesis that chromaticism may have been more common around NHTs. It was also hypothesised that the $\mathrm{IOI}_{\mathrm{BeatProp}}$ of NHTs played by Green would be shorter than the arpeggio or scale tones. Additionally, it was hypothesised that the articulation of NHTs would be lower than HTs, with Green playing them in a staccato manner.[28] Finally, an investigation into dominant 7 altered chord tensions was undertaken.

Non-Harmonic Tones – Chromaticism

Table 5.7 shows the percentage of notes that had a chromatic approach, departure, approach and departure, and approach and/or departure. These are shown for each of the $\mathrm{CPC}_{\mathrm{Weight}}$ individually, for the arpeggio and scales tones combined as HTs, and all notes combined. This data showed that 49.22% of Green's notes had a chromatic approach and/or departure. This was comprised of 18.33% that had only a chromatic approach, 18.25% with a chromatic departure, and 12.64% with both a chromatic approach and departure. The data also indicated that Green treated HTs and NHTs differently, with the largest differences in approach and departure, and therefore approach and/or departure. Separate χ^2-tests were run to compare the $\mathrm{CPC}_{\mathrm{Weight}}$ distributions in both these situations, with both tests finding a significant difference in the distributions, with a small effect size (approach and departure: $\chi^2(2) = 1451.11,\ p =< .001,\ V = .28$; approach and/or departure: $\chi^2(2) = 1034.48$, $p =< .001,\ V = .23$). Subsequent post-hoc tests for the approach and/or departure found significant pairwise differences for all comparisons ($p < .001$). The post-hoc tests for chromatic approach and departures found no significant difference between the arpeggio and scale tone ($p = .156$), with all other comparisons significant ($p < .001$).

[28] Articulation, investigated in Chapter 7, was the ratio of the duration to the IOI (duration÷IOI).

Table 5.7: Percentage of CPC_{Weight} that had a chromatic approach, departure, approach and departure, and approach and/or departure.

	Approach Only	Departure Only	Approach & Departure	Approach &/Or Departure
Arpeggio	20.54%	13.33%	7.58%	41.45%
Scale	18.36%	19.84%	8.25%	46.45%
HT	19.79%	15.58%	7.81%	43.18%
NHT	13.01%	27.97%	30.27%	71.26%
All	18.33%	18.25%	12.64%	49.22%

The data showed that Green both approached and departed HTs chromatically less than 10% of the time. In comparison, nearly a third of all NHTs had Green approaching and departing chromatically. Less than half of all HTs had either a chromatic approach and/or departure, while nearly three-quarters of the NHTs played by Green were either approached and/or departed chromatically. Green played a similar proportion of only chromatic approaches or departure for scale tones. For arpeggio tones, Green was more likely to approach the note chromatically rather than leave chromatically; while Green departed from NHTs chromatically more than twice as often as approaching chromatically. These results supported the hypothesis that Green played significantly more chromatic intervals to approach or depart NHTs compared to HTs. These results also indicated that Green had a preference for departing from NHTs chromatically compared to a chromatic approach, which was favoured for HTs.

Non-Harmonic Tones – Note Length

The following analyses investigated how Green's treatment of NHTs influenced both the length of the notes (as $IOI_{BeatProp}$) and the articulation of the notes. Figure 5.36 shows the $IOI_{BeatProp}$ distribution for each CPC_{Weight}, focused on the data between 0 and 2. This graph shows that the $IOI_{BeatProp}$ distributions for the arpeggio and scale tones were similar, while Green's NHT $IOI_{BeatProp}$ tended to be shorter. An ANOVA found that the $IOI_{BeatProp}$ did differ significantly depending on the CPC_{Weight} of the note, with a small effect size ($F(2, 18859) = 163.92,\ p < .001;\ \eta^2 = .02$). Post-hoc tests with Tukey's HSD procedure found significant pairwise differences for all comparison ($p < .001$). In Green's improvisations, both HTs had a median $IOI_{BeatProp}$ similar to a quaver-equivalent note (arpeggio: 0.50 beats; scale: 0.47

beats). In comparison, Green's NHTs had a median $\mathrm{IOI}_{\mathrm{BeatProp}}$ most similar to a quaver-triplet-equivalent note (NHT: 0.34 beats). These results supported the hypothesis that Green's NHTs tended to be shorter than his HTs.

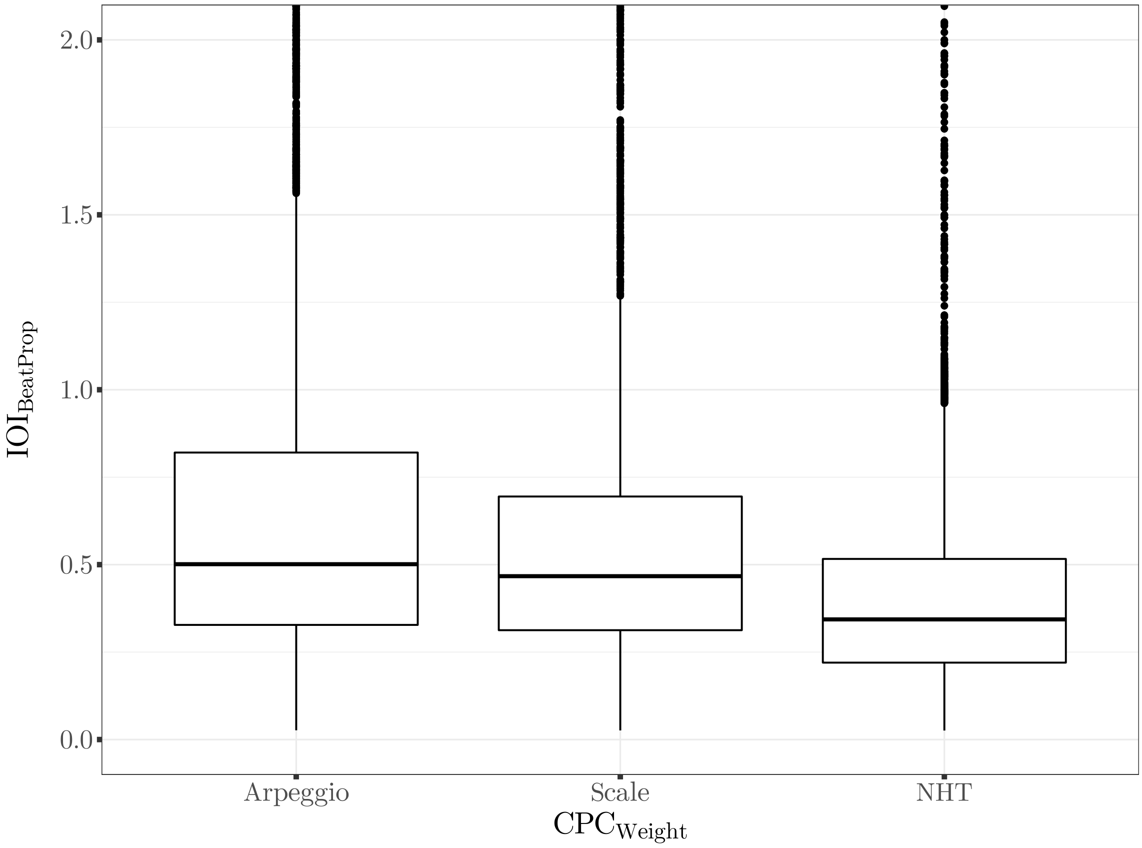

Figure 5.36: $\mathrm{IOI}_{\mathrm{BeatProp}}$ distribution for each $\mathrm{CPC}_{\mathrm{Weight}}$.

The second investigation related to the hypothesis that Green played NHTs with a lower articulation, e.g. their $\mathrm{IOI}_{\mathrm{BeatProp}}$ had a lower $\mathrm{duration}_{\mathrm{BeatProp}}$ when compared to HTs. Figure 5.37 shows the distribution of articulation values for each $\mathrm{CPC}_{\mathrm{Weight}}$. Although this data did show a substantial difference in the articulation of the NHTs compared to HTS, it was in the opposite direction than hypothesised. Instead of Green's NHTs having a shorter articulation than the HTs, they tended to have a longer articulation. An ANOVA found a significant difference in articulations dependent on the $\mathrm{CPC}_{\mathrm{Weight}}$ of the note, with a small effect size ($F(2, 18819) = 167.72,\ p < .001;\ \eta^2 = .02$). Subsequent post-hoc tests with Tukey's HSD procedure found significant pairwise differences for all comparisons ($p < .001$). Although the articulation for both the HTs were around 0.75 (arpeggio: 0.77; scale: 0.78), Green's NHT articulation was significantly higher (NHT: 0.85). The higher articulation for NHTs may have been related to their use as chromatic passing tones, or as part of a slur. These results did not support the hypothesis that Green played NHTs with a lower articulation than HTs, with the opposite found to be true. Combined, these results indicated that Green did tend to play NHTs for a shorter amount of time than his HTs, and that they more frequently had longer articulations.

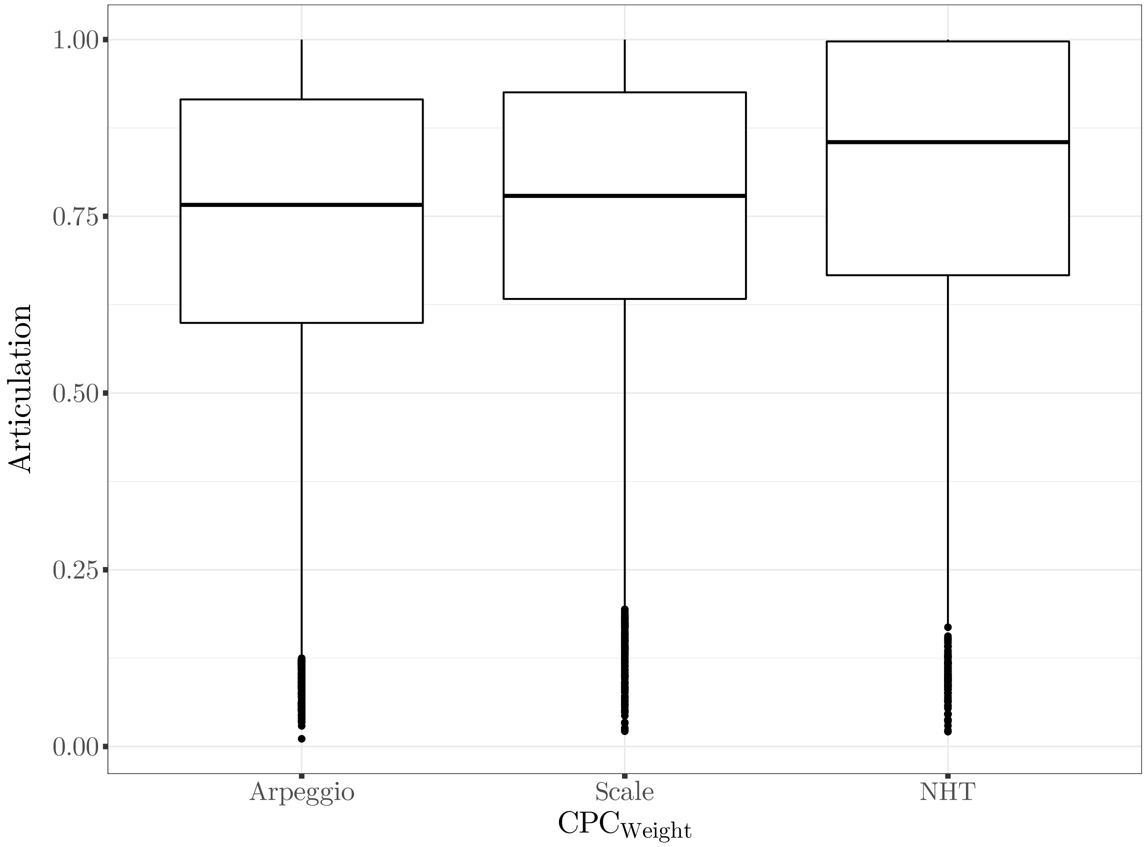

Figure 5.37: Articulation distribution for each CPC$_{\text{Weight}}$.

Dominant 7 Altered chord tensions

Dominant 7 chords had the highest frequency of NHTs within Green's corpus. This fit with standard jazz practices, where altered chord tensions tend to be played more frequently over 7 chords than other chord types. Altered chord tensions of a 7 referred to alterations of the 9$^{\text{th}}$, 11$^{\text{th}}$, 5$^{\text{th}}$, and 13$^{\text{th}}$; with the most common alterations being ♭9, ♯9, ♯11/♭5, and ♯5/♭13. Additionally, the ♮7 from the dominant bebop scale, while not an alteration played when comping, was an available tension note when improvising. Altered tensions, and NHTs more generally, worked as a mechanism of tension and release, providing resolutions to more stable tones (e.g. arpeggio). Therefore, their high frequency on an already unstable chord provided even heightened tension. Due to the nature of dominant harmonic progressions, altered tensions of a 7, when resolving to a Δ7 dominantly, could always resolve chromatically to an arpeggio tone.

Green's distribution of altered chord tensions can be seen in Figure 5.38. This data indicated that Green's most frequently played altered tensions were the ♯5/♭13 (♭6), ♯9 (♭3/♮3), and ♭9 (♭2). Around one-sixth of all 7 altered chord tensions were the ♮7, while just under one-eighth were TTs. This continued the trend of Green not frequently playing TTs in his improvisations. The previous analysis into chord weight found that Green played a higher proportion of NHTs in the beat before a chord change. To investigate how this differed for each specific altered chord tension, the proportions for each NHT was plotted against the numbers of beats until a chord change in Figure 5.39.[29] This data showed that in the beats leading up to a chord change, Green's proportion of NHTs that were ♯9, ♯5, and TTs increased.

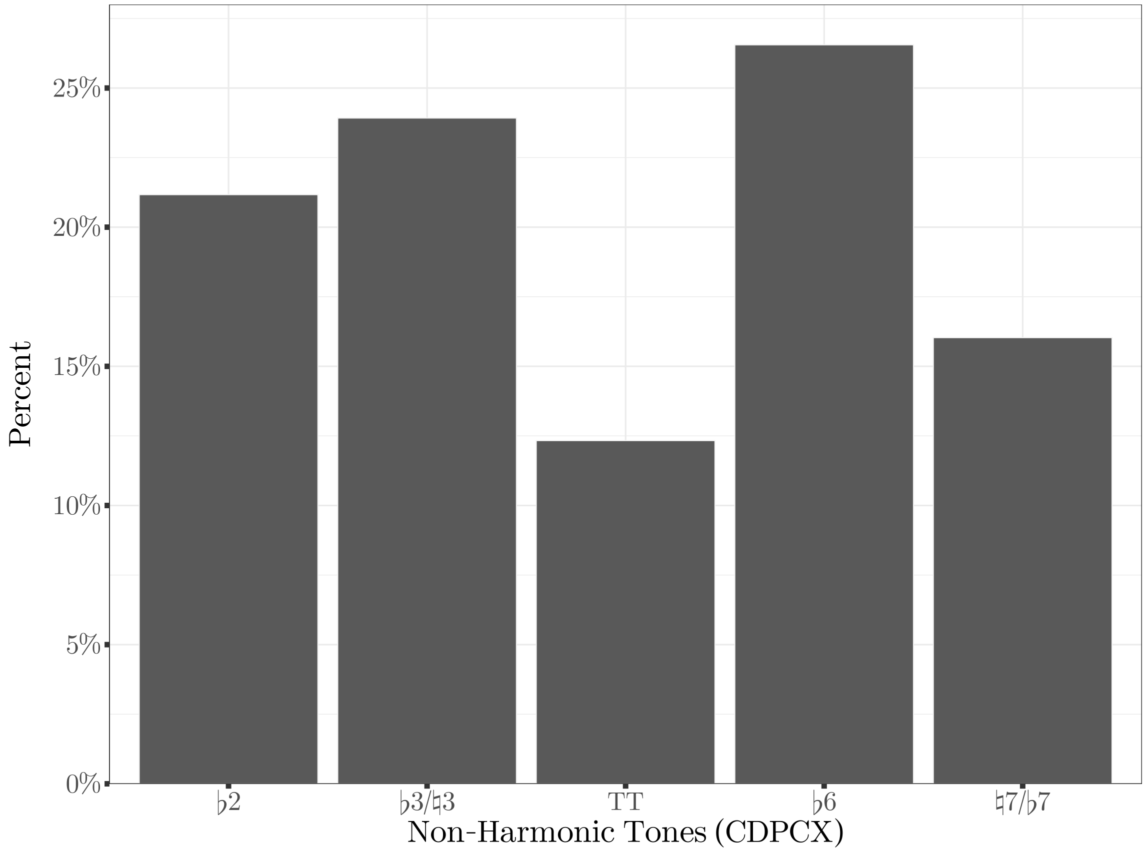

Figure 5.38: Frequency of altered chord tensions over 7 chords in Green's corpus.

[29]The percentages represented the proportion of NHTs, not the proportion of all tones.

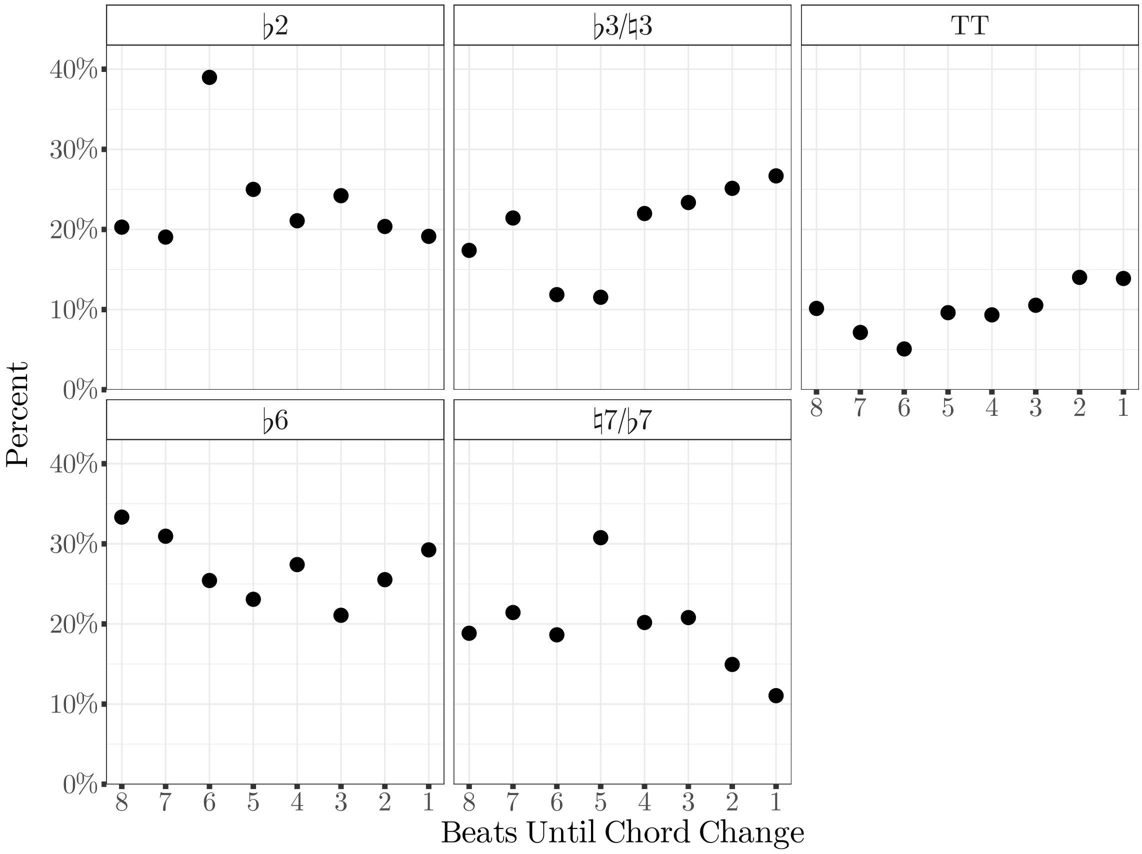

Figure 5.39: Proportion of NHTs in the beats before a chord change.

Figures 5.40 and 5.41 show the CDPCX transitions around the most common NHTs, the ♭3/♮3 and ♭6 respectively. The colours indicated which tone was played after the altered tone. The figures did not take into account the metrical weight of any of the notes or the chord type over which the note before and note after were played. However, they only included notes where the combined $\mathrm{IOI}_{\mathrm{BeatProp}}$ of the note before and altered tone was ≤ 1.1 beats. This ensured that the sequences were no longer than a crotchet-equivalent line. Only trigrams that occurred at least ten times in Green's corpus were plotted. The ♭3/♮3 and ♭6 had a wide variety of preceding and following tones; however, the majority of the notes came from tones that were close to the altered tone. The wider variety of tones that followed a NHT may have been explained by Green more frequently playing NHTs in the beats before a chord chord change, with the NHT used to target a variety of CDPCX classes of the following chord. With consideration of the results of Green's use of chromaticism around NHTs, it was expected that many of the leading and following notes were from tones close to the NHT.

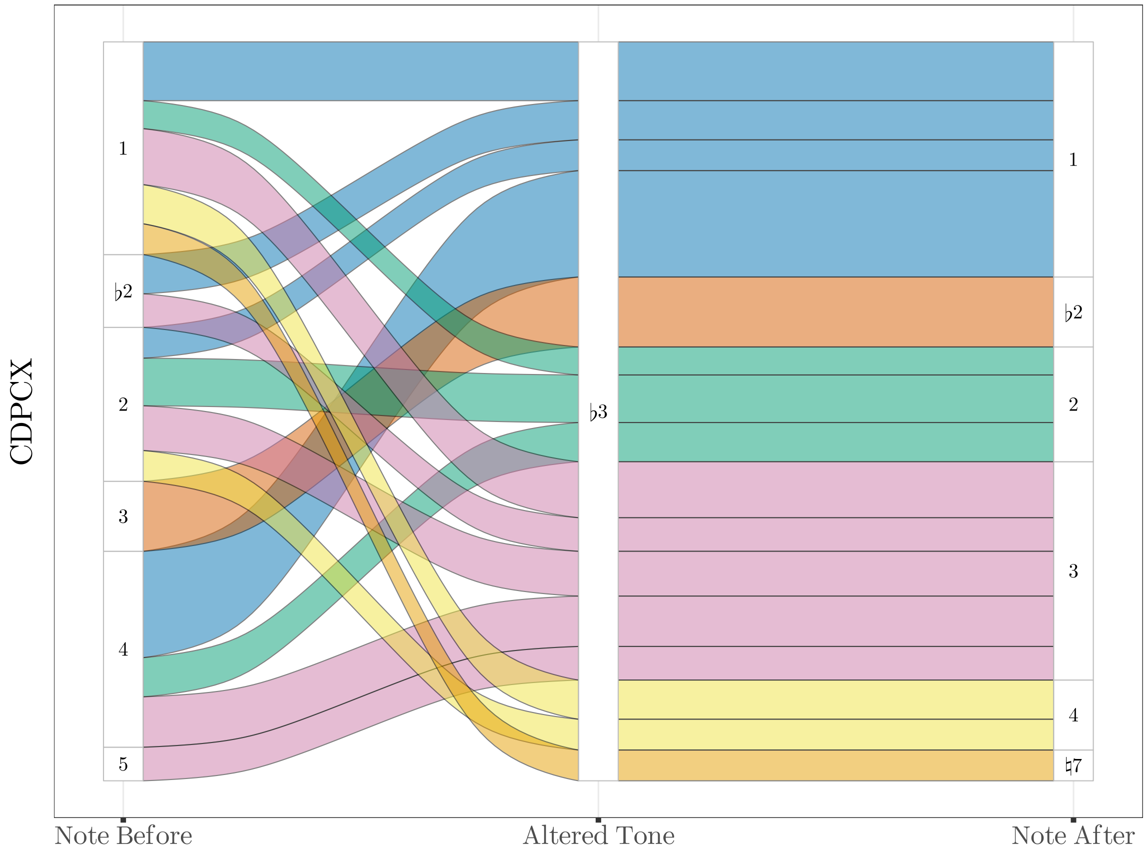

Figure 5.40: CDPCX transition around ♭3/♮3 over 7 chords.

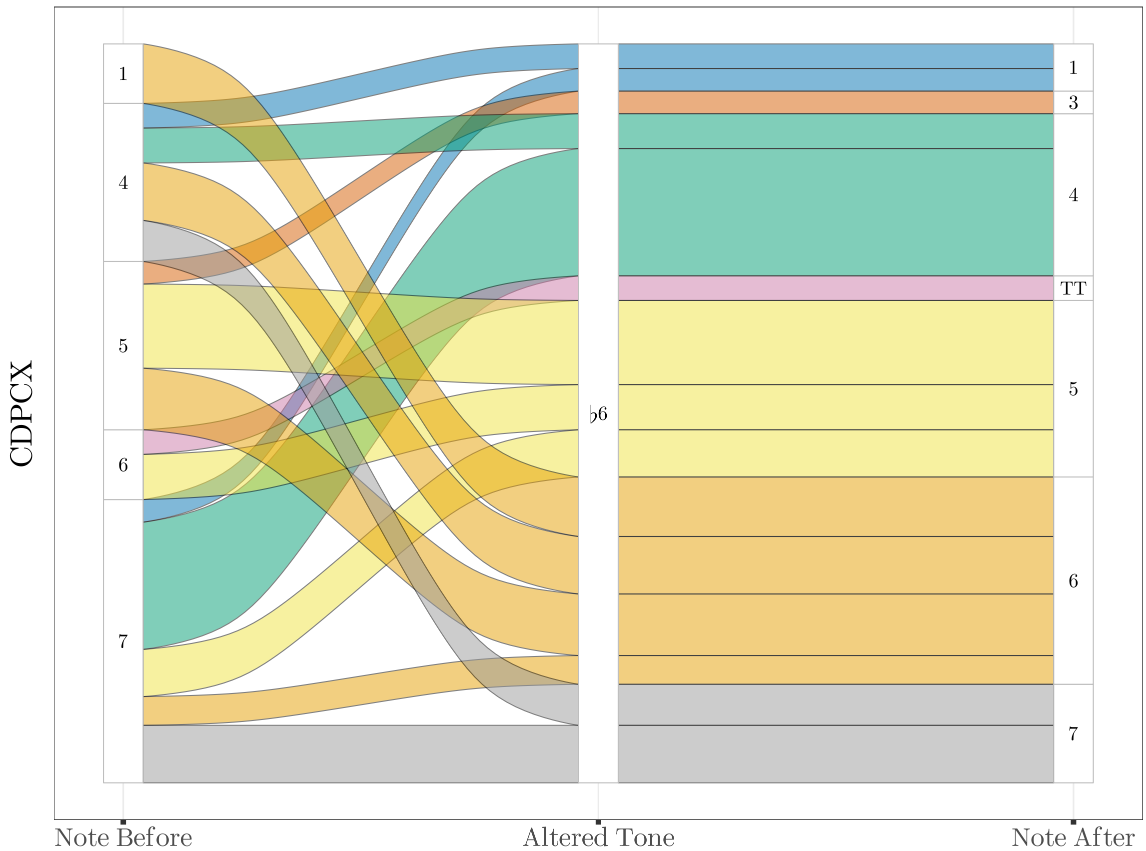

Figure 5.41: CDPCX Transition around ♭6 over 7 chords.

Treatment of Non-Harmonic Tones Summary

This analysis into Green's treatment of NHTs found that Green frequently used chromaticism around NHTs, with Green particularly favouring a chromatic departure. The analysis also found that Green's NHTs tended to be shorter than the HTs, although they often had longer articulations. The investigation into Green's use of dominant 7 altered chord tensions found that Green most frequently played the ♯5, ♯9, and ♭9 notes, with the first two being more common in the beats leading up to a chord change. In summary, these results found that Green treated NHTs substantially differently from HTs.

5.3.6 Upper Structure Triads

Upper structure triads (USTs) referred to the triadic structure of notes built from the extensions of a chord's arpeggio, i.e. from notes beyond the 7th. For example, the chord tones, including extensions, of a CΔ7 chord are: C–E–G–B–D–(F)–A (with the F being a weak, diatonic non-harmonic, chord tone (Crook 1991, 105)). Other available extensions, from alterations to those tones, could include G♯ and F♯ (♯5 and ♯11 respectively). In this example an UST would be any triad that included at least one of the notes D, A, F♯, or G♯. These triads included (treating CΔ7 as I): G (V); Am (vi); E (III)[30]; Bm (vii); and D (II). Each of the USTs would add at least one colour or altered tone to the base CΔ7 chord in an improvisation.

Although a wide variety of USTs were available, for this project only the following USTs were considered ([chord]: [triad] [(tones)], with uppercase for major triads and lowercase for minor):

- Δ7: II (2, ♯4, 6); III (3, ♯5, 7); V (5, 7, 9); vi (6, 1, 3); vii (7,9,♯4)
- 7: II (2, ♯11, 13); ♭III (♯9, 5, 7); ♭V (♯11, 7, ♭9); ♭VI (♯5, 1, ♯9); VI (13, ♭9, 3)
- 7^{sus}: ♭II (♭9, 11, ♯5); IV (11, 13, 1); ♭VII (7, 9, 11)
- m7: IV (7, 9, 11); ♭VII (11, 13, 1)

A function was written to search for the presence of USTs in Green's improvisations. This first looked at every trigram for each unique chord change in Green's improvisations and checked if the $IOI_{BeatProp}$ of the first and second notes were each less than 1.3 beats each, and that the combined $IOI_{BeatProp}$ of these two notes was less than 2 beats.[31] This was to ensure that each note of the trigram was at least marginally connected. Each trigram that passed this test was then checked to identify if the three notes formed a major or minor triad (in any inversion). If it did,

[30] Technically, although the III triad over a IΔ7 is not an UST as all notes were from the base chord, albeit with alterations, it is usually considered and grouped with the other USTs.

[31] Chords with less than three notes were not considered.

the root note was compared to the tonic of the surrounding chord, with both the type of triad and relation to the chord tonic recorded.

The frequency of the USTs listed above for each of the chord types are shown in Tables 5.8, 5.9, and 5.10 (both 7 and 7^{sus}). The count showed the raw frequency of the UST in Green's improvisations. The percentage was based on how frequently Green played each UST as a proportion of all triads found in that chord type. For example, this included the base triad for each chord type, as well as all other possible major and minor variants, regardless of whether they would be considered an UST in a broader search for USTs.

Table 5.8: Distribution of Δ7 USTs.

	II	III	V	vi	vii
Count	1	1	15	18	3
Percent	0.66%	0.66%	9.93%	11.92%	1.99%

Table 5.9: Distribution of m7 USTs.

	IV	♭VII
Count	3	16
Percent	1.01%	5.37%

Table 5.10: Distribution of 7 USTs.

	♭II	II	♭III	IV	♭V	♭VI	VI	♭VII
Count	5	4	7	44	2	8	4	58
Percent	1.25%	1.02%	1.79%	11.22%	0.51%	2.04%	1.02%	14.80%

This data indicated that Green did not frequently play trigrams that were comprised of the notes from a UST. Even the additional notes from the most frequent USTs, the vi for Δ7 and IV and ♭VII for 7, were more easily explained by simple substitution (♮6 instead of ♮7 for Δ7), extension (13^{th} for 7), or blues influenced language (11^{th} for 7). If there was evidence of Green frequently playing other USTs, these notes could fit within that explanation. With so little overall use, the results suggested that the use of USTs was not a frequent or integral part of Green's improvisational style.

5.3.7 Chordal Pitch Class Summary

This analysis investigated Green's use of chordal pitch classes. It found that the majority of Green's notes came from the arpeggio of the chord of the moment, with NHTs being most common over 7 chords. Green was most likely to play HTs on-beat, while NHTs were more frequently played off-beat. The analysis found that the 4^{th} was a frequent tone in Green's improvisations. While it was less likely to be played by Green on metrically strong beats over Δ7 or 7 chords, it was often played over m7 chords. The high frequency of the 4th was likely likely influenced by the use of blues language in Green's improvisations.

The analysis found that Green played a higher proportion of NHTs in the beat before a chord change. Nearly three-quarters of all NHTs Green played were either approached and/or departed from chromatically, with nearly a third having both a chromatic approach and departure. Green's NHTs tended to have a shorter $IOI_{BeatProp}$ when compared to HTs, but their articulation was often longer. Over 7 chords Green's most common NHTs were the ♭2, ♭3/♮3, and ♭6, with the latter two being more frequent in the beat before a chord change. The analysis into USTs found that they were not a common improvisational tool used by Green in his improvisations. Overall, the investigation into NHTs found that Green did treat them substantially different from HTs. As shown in the treatment of non-harmonic tones, it was important to consider not only the distribution of pitches played, but how Green transitioned between the notes. The following section investigated Green's use of intervals and intervallic structures within his improvisations.

5.4 Intervals

Intervals are a measure of the distance in pitch between two notes; within *MeloSpy* these were reported as the number of semitones. The interval value was attached to only the first note of each pair that formed the interval. Consequently, the last note event in each of the forty transcriptions contained no interval values. There were three features used to describe the intervals between two notes, from most to least coarse: Parsons code; fuzzy intervals; and raw intervals.[32] The Parsons code consisted of only three classes – ascending, repetition, or descending – while the fuzzy interval feature had eleven classes – five descending, repetition, and five ascending (Jazzomat Research Project 2017). Table 5.11 shows the interaction between the three interval descriptions, with the column names the Parsons codes, the row names the fuzzy interval classes, and the numbers in the cells the corresponding raw semitone interval values. The fuzzy interval feature used throughout this section was an expanded version of the *MeloSpy* feature. The original feature did not include the >8ve ascending or descending classes. This section focused on the analysis of the Parsons, the use of raw intervals, and the use of fuzzy intervals within Green's improvisations. The raw intervals were used to analyse both the overall distribution and a specific investigation into Green's use of arpeggios and triadic structures. The fuzzy intervals analysis investigated the overall distribution, and the interaction between fuzzy intervals, note length, and NITP.

[32] A common investigation of intervals was to search for interval patterns, e.g. licks or interval n-grams (Cross and Goldman 2021; Dixon et al. 2017). A full investigation of interval patterns was outside the scope of this research.

Table 5.11: Interval Descriptors.

	Ascending	Repetition	Descending
5: >8ve↑	13, 14, 15, 16, ...		
4: Big Jump Up	8, 9, 10, 11, 12		
3: Jump Up	5, 6, 7		
2: Leap Up	3, 4		
1: Step Up	1, 2		
0: Repetition		0	
-1: Step Down			-1, -2
-2: Leap Down			-3, -4
-3: Jump Down			-5, -6, -7
-4: Big Jump Down			-8, -9, -10, -11, -12
-5: >8ve↓			-13, -14, -15, -16, ...

5.4.1 Parsons Code

The distribution of the Parsons codes in Green's corpus is shown in Table 5.12. This data showed that Green played a similar number of ascending and descending intervals, with a slight preference for descending intervals. The use of repeated notes fairly rare in Green's improvisations. For comparison, the distribution of the Parsons code from the WJazzD was 49.25% descending, 5.12% repetition, and 45.62% ascending. This suggested that Green's slight preference for descending intervals was not distinct. However, Green's Parsons code distribution was significantly different from the WJazzD, with a small effect size ($\chi^2(2) = 131.13$, $p =< .001$, $V = .02$). This indicated that while Green did show a slight preference for descending intervals, his distribution was more even than the WJazzD.

Table 5.12: Distribution of Parsons code in Green's corpus.

	Descending	Repetition	Ascending
Count	9592	1395	9451
Percent	46.93%	6.83%	46.24%

For there to a higher proportion of descending intervals, the ascending intervals would be expected to be larger, to allow room for longer descending lines. Table 5.13 shows the absolute raw semitone interval mean, standard deviation, skewness, and kurtosis for all of the descending and ascending intervals that Green played. This data showed that the mean ascending interval was only fractionally larger than the descending. However, the skewness and kurtosis indicated that larger ascending intervals were more common in Green's improvisations than larger descending intervals.

Table 5.13: Statistical descriptors for descending and ascending intervals in Green's corpus.

	Mean	S.D	Skewness	Kurtosis
Descending	2.67	1.95	2.20	7.69
Ascending	2.68	2.17	2.59	11.14

Figure 5.42 shows an example, from Green's improvisation over *The Song Is You* (Green 1962l), of large ascending intervals followed by a predominantly descending line. These ascending intervals were between the first G in bar 112 to the E in bar 113 (ascending major 6th) and between the E♭ at the end of bar 114 and B in bar 115 (ascending minor 6th).

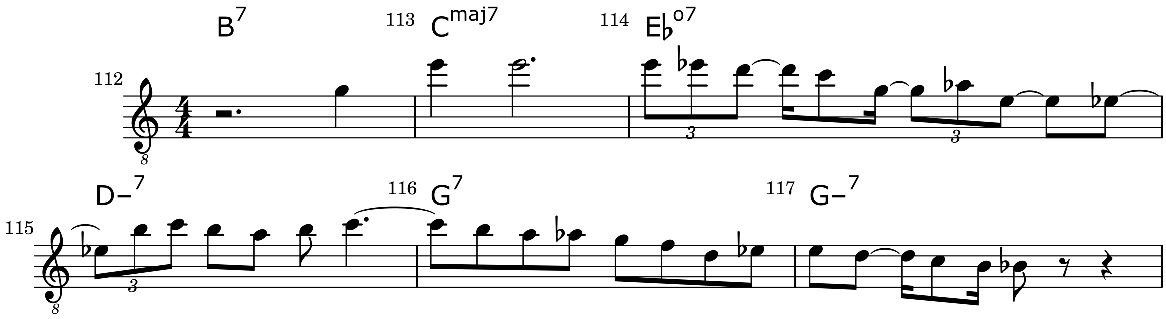

Figure 5.42: Example of a large ascending interval followed by a descending line, *The Song Is You* (1961), bars 112–117.

In the corpus there were 1021 sequences of repeated notes, with the vast majority, 881 (82.76%), being a single repeated note. There were 111 (10.87%) occurrences of two repeated intervals (three repeated notes). The two longest sequences of repeated intervals were twenty and thirty-one, both of which occurred once each in Green's corpus. These counts did not take into consideration the distance between the repeated notes, but as can be seen in Figure 5.43 the majority of repeated notes had an $\mathrm{IOI}_{\mathrm{BeatProp}} \leq 1$ beat, and would be considered true repeated notes.

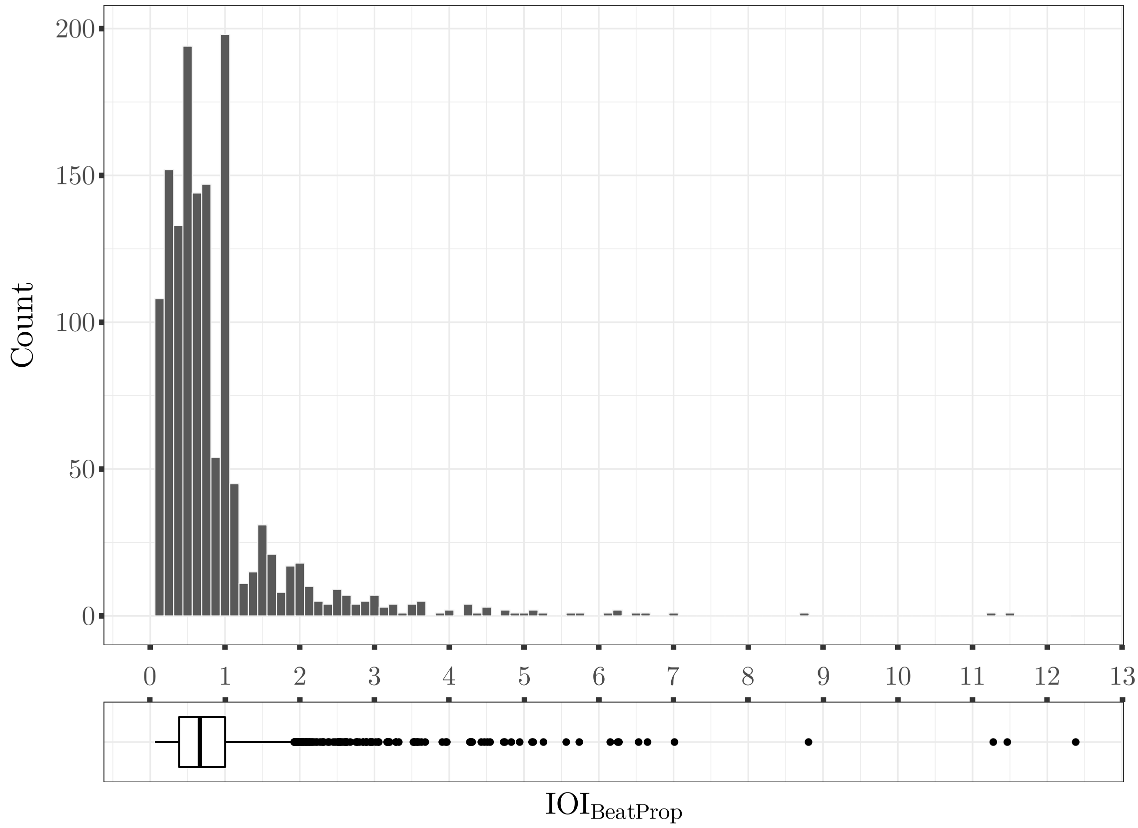

Figure 5.43: $\mathrm{IOI}_{\mathrm{BeatProp}}$ distribution for repeated notes in Green's corpus.

Figure 5.44 shows the distribution of the Parsons classes for 200 intervals following a repeated note. For comparison, the frequency of only the repeated notes for Davis, Parker, and Coltrane were included. The length of 200 intervals was selected as 90% of all songs from the four performers had at least 196 notes. The horizontal line in each plot shows the baseline mean of that Parsons for each performer, and a smoothed line (using LOESS – locally estimated scatterplot smoothing – moving average) was placed over each scatterplot.

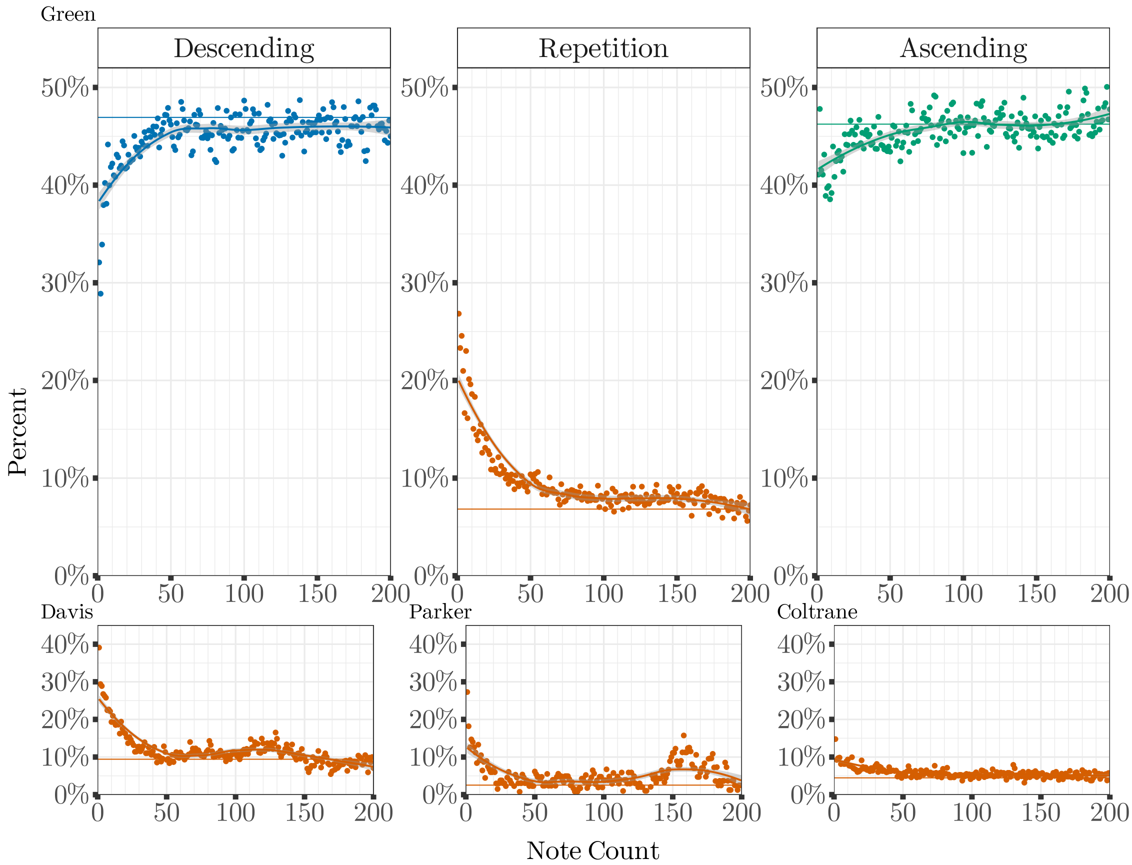

Figure 5.44: Distribution of Parsons code for the 200 intervals following a repeated note in Green's corpus and distribution of repeated Parsons for the 200 intervals following a repeated note in Davis, Parker, and Coltrane's corpus.

The graph showed that after Green repeated a note, he was more likely to play another repeated note in the next 200 notes, returning to the baseline after around 200 intervals. The proportion of ascending intervals returned to baseline probabilities after around eighty intervals, while the descending intervals stabilised slightly lower than the baseline proportion after around sixty intervals post-repetition. The proportion of repeated intervals decreased quickly over the first fifty intervals, with a slower decline over the next 150 intervals. The same increase in repetition probability can be seen in Davis, Parker, and Coltrane, but to different extents. Both Davis and Parker see a similar initial increase in repetition intervals as Green, returning to their baseline proportions after around fifty intervals. They both see an additional increase later, Davis around 100 intervals and Parker at 140, followed by a quick return to the baseline. In contrast, Coltrane only sees a minimal increase in repeated intervals proportions, returning to the baseline proportions after around fifty intervals, with little variation thereafter.

In comparison to the repeated intervals, Figure 5.45 shows the Parsons distributions for the 200 intervals following a descending or ascending interval in Green's improvisations. In these situations, large deviations from the baseline proportions were not observed. This indicated that Green's playing of an ascending or descending interval did not change the overall probability of other Parsons classes. These results suggested that there were improvisations where Green had a higher tendency to play repeated notes. Therefore, in those solos, once Green repeated a note, he was then more likely to play more repeated notes. In support of this was that in twenty-seven of the improvisations there were fewer than thirty repetition sequences each. Therefore, 64.35% of the repetition sequences were contained in the remaining thirteen improvisations.

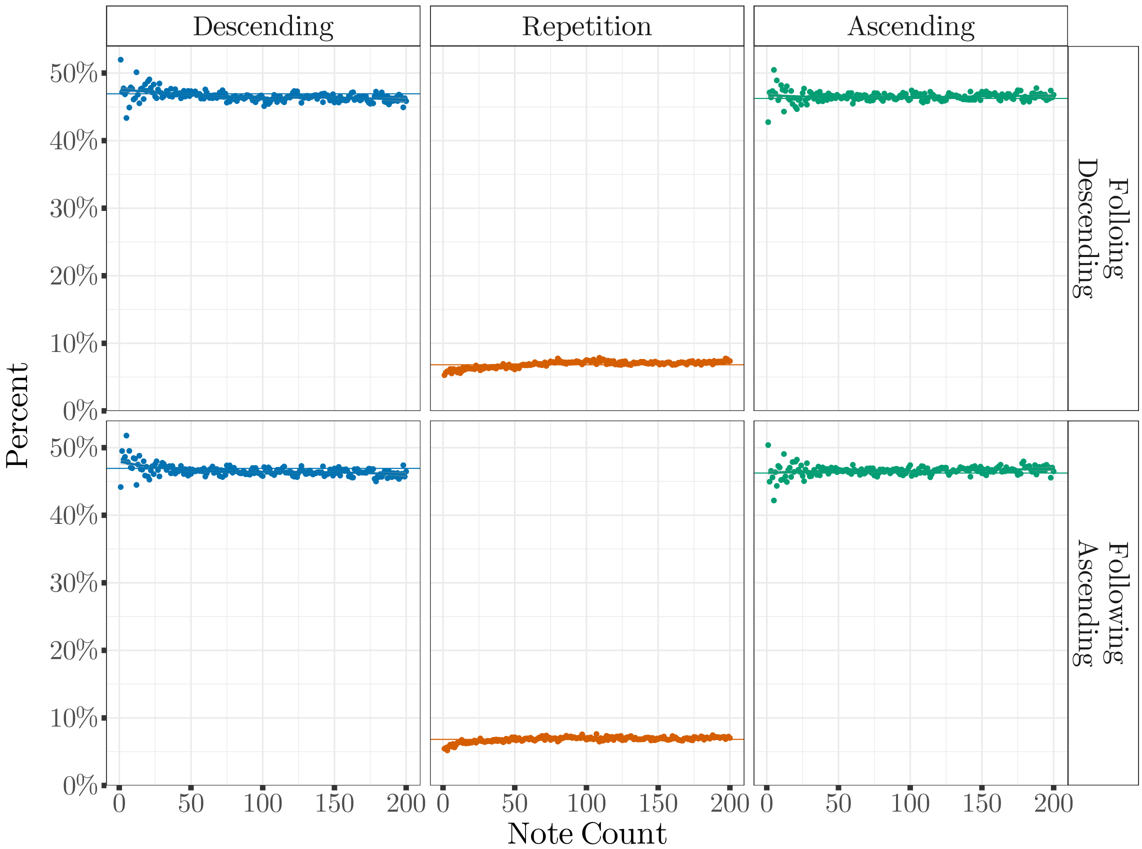

Figure 5.45: Distribution of Parsons code for the 200 intervals following a descending or ascending interval in Green's corpus.

5.4.2 Raw Interval Values

The distribution of raw semitone intervals in Green's improvisations is shown in Figure 5.46, with the vertical lines indicating the breaks in the fuzzy interval classes. Plot a) shows all intervals within one octave, with ascending and descending intervals larger than an octave shown in plots b) and c) respectively. These two graphs were split out due to their low occurrence within Green's improvisations, all being played less than 0.06% of the time.

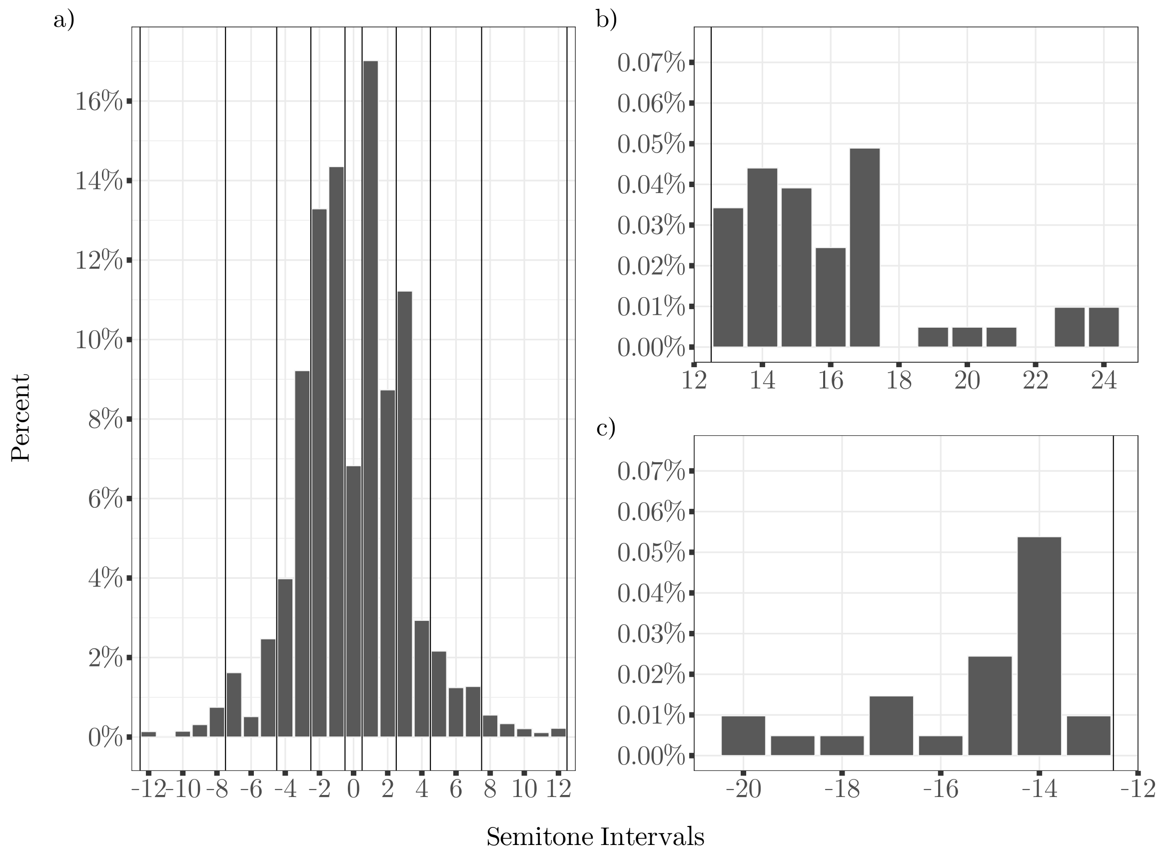

Figure 5.46: Distribution of raw interval values in Green's corpus.

The data in plots b) and c) also supported the Parsons hypothesis that Green's ascending intervals tended to be larger than his descending intervals. Although a descending interval of an octave and a tone (-14 semitones) was the most frequent single interval larger than one octave, the other large descending intervals were rarely played. In comparison to the twenty-six large (> 12 semitones) descending intervals, Green played thirty-nine ascending intervals between 13 and 17 semitones. Green's largest ascending interval, two octaves, was also a major 3rd larger than his descending intervals, an octave and a minor 6th.

Plot a) showed that the vast majority (80.65%) of Green's intervals were between a descending and ascending minor 3rd (3 semitones). The data showed that Green played ascending semitones nearly twice as often as ascending tones, while both descending semitones and tones were played equally. Green played ascending minor thirds more frequently than descending minor thirds and ascending tones; descending major thirds were also more frequent than ascending major thirds.

There were very few large intervals (greater than an octave) observed in Green's improvisations, only seventy-two in the entire corpus (0.35%). As large intervals would often break the flow of a line, it was most likely that these would have occurred between phrases (inter-phrase). This was observed in Green's data, where 88.89% of intervals greater than an octave occurred between phrases, with zero played at the beginning of a phrase. However, since Green played large intervals so infrequently, they only comprised 5.28% of all inter-phrase intervals. Figure 5.47 shows an example of Green playing a large interval between phrases, from his improvisation over *The Surrey With The Fringe On Top* (Solo 1, Green 1963g).

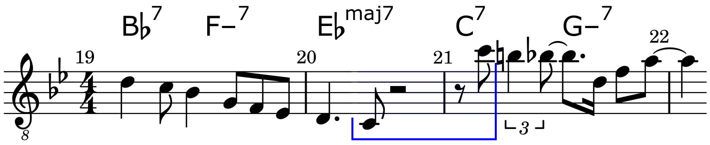

Figure 5.47: Example of a large (two octave) interval between phrases, *The Surrey With The Fringe On Top* (Solo 1, 1963), bars 19–21.

Triads and Arpeggios

The two most frequent types of transitions in Green's improvisations were seconds and thirds. Intervals of a third are frequently associated with triadic or arpeggio based movements. This section analysed the use of thirds within Green's improvisations to investigate whether he used them to play the arpeggio of the surrounding chord. This section focused on the absolute value of the intervals, as the direction of the arpeggio movement was not considered important. The data in Figure 5.46, above, showed that Green played substantially more minor thirds than major thirds. This was likely influenced by the structure of the chords, with minor thirds being more frequent than major thirds in the arpeggios of the three main chord types.[33] This preference for minor thirds over major thirds could be explained by Green predominantly playing thirds in arpeggios.

When Green played a third, 83.47% of the time, both notes were played over the same chord.[34] This was strong indication that Green likely used thirds for arpeggio based movement within his improvisations. Table 5.14 shows the percentage of minor and major thirds, where both notes were played over the same chord and both notes were part of the extended arpeggio (1, 3, 5, 7, 9, 13). For the extended arpeggios of the three main chord types there were three minor thirds and two major thirds. The data in this table showed that the majority of thirds, major or minor, that Green played over Δ7 and m7 chords were between two notes of the arpeggio. For 7 chords, Green rarely played major thirds between arpeggio tones, while around half of his minor thirds were played between two arpeggio tones. It was probable that the high frequency of altered chord tensions over 7 chords was related to the lower

[33]For Δ7 chords there were two major thirds and one minor thirds, while the opposite was true for both 7 and m7 chords. ø7 chords also contained one major third and two minor thirds, while ∘7 chords contained three minor thirds.

[34]Arpeggios could be played between chords, but this analysis focused on arpeggios played over a single chord.

use of standard arpeggio movement. These results indicated that Green frequently used thirds as part of arpeggio movements, most frequently over $\Delta 7$ and m7 chords.

Table 5.14: Proportion of major and minor 3^{rd}s for each chord type where both notes were part of the extended arpeggio.

	$\Delta 7$	7	m7
Major 3^{rd}			
Arpeggio	68.97%	33.57%	63.83%
Other	31.03%	66.43%	36.17%
Minor 3^{rd}			
Arpeggio	64.42%	48.27%	64.82%
Other	35.58%	51.73%	35.18%

5.4.3 Fuzzy Interval Values

The fuzzy interval values were a grouped form of the raw interval values. In general, the fuzzy interval features was preferred as a descriptor for two reasons:

1) As with many of the coarser features, the fuzzy intervals provided the function of the intervals (step, leap, jumps, etc.) without concern for the specifics;
2) There there were fewer classes, each containing more note events, which aided in statistical analyses.

Figure 5.48 shows two distributions of the fuzzy intervals that Green played. The left graph shows the fuzzy intervals from -5 ($> 8^{ve}$ down) to 5 ($> 8^{ve}$ up), while the right graph shows the absolute fuzzy interval classes. This data showed that Green played a similar number of ascending and descending fuzzy interval classes. Descending steps were slightly more likely than ascending, while ascending leaps were more common. More than half of all notes Green played moved by step-wise motion. The following sub-sections focused on specific hypotheses regarding how Green's use of fuzzy intervals were related to other features. The analyses focused on features related to the first note of the pair of notes that comprised each interval, specifically the note length and NITP.

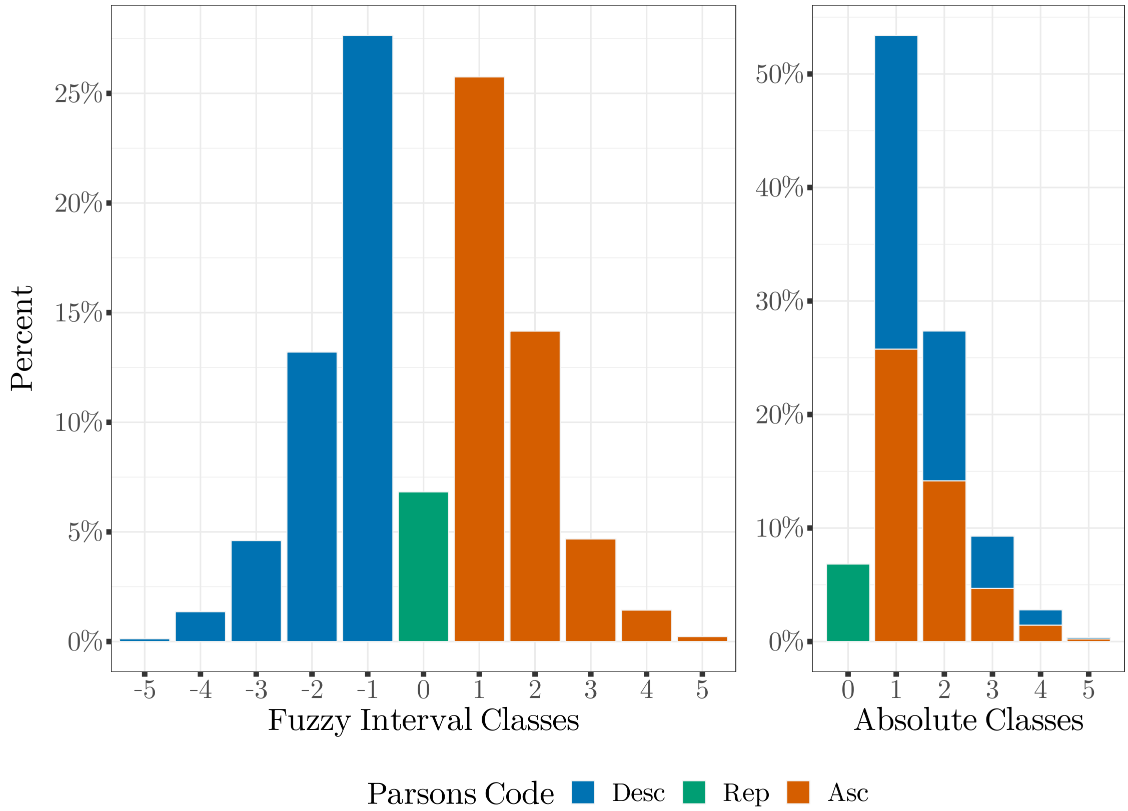

Figure 5.48: Distribution of fuzzy interval classes in Green's corpus. Left: Fuzzy interval classes. Right: Absolute fuzzy interval classes.

Fuzzy Interval Classes vs. Note Length

This analysis focused on the interaction between the length of the first note in a pair of intervals. It was hypothesised that larger fuzzy interval classes would be related to longer notes, especially longer $\text{IOI}_{\text{BeatProp}}$. These longer note lengths were expected due to larger intervals being more likely to be played between phrases, with mean inter-phrase $\text{IOI}_{\text{BeatProp}}$ longer than intra-phrase $\text{IOI}_{\text{BeatProp}}$ (2.54 vs. 0.52 beats). It was also hypothesised that there would be a relationship between the fuzzy intervals and the $\text{duration}_{\text{BeatProp}}$, although to a lesser degree.

These two distributions are shown in Figure 5.49, $\text{IOI}_{\text{BeatProp}}$ on the left and $\text{duration}_{\text{BeatProp}}$ on the right. To investigate the relationship between the two note length descriptors and the fuzzy interval classes an ANOVA was run for each comparison. A significant difference was found for each comparison, with the $\text{IOI}_{\text{BeatProp}}$ comparison having a medium effect size ($F(10, 20427) = 334.66$, $p < .001$; $\eta^2 = .14$), while the $\text{duration}_{\text{BeatProp}}$ comparison had a small effect size

($F(10, 20427) = 73.54,\ p < .001;\ \eta^2 = .03$).[35] These results supported the hypotheses that there was a relationship between the length of the notes and the fuzzy intervals between notes, and that the effect was more pronounced for $\mathrm{IOI}_{\mathrm{BeatProp}}$.

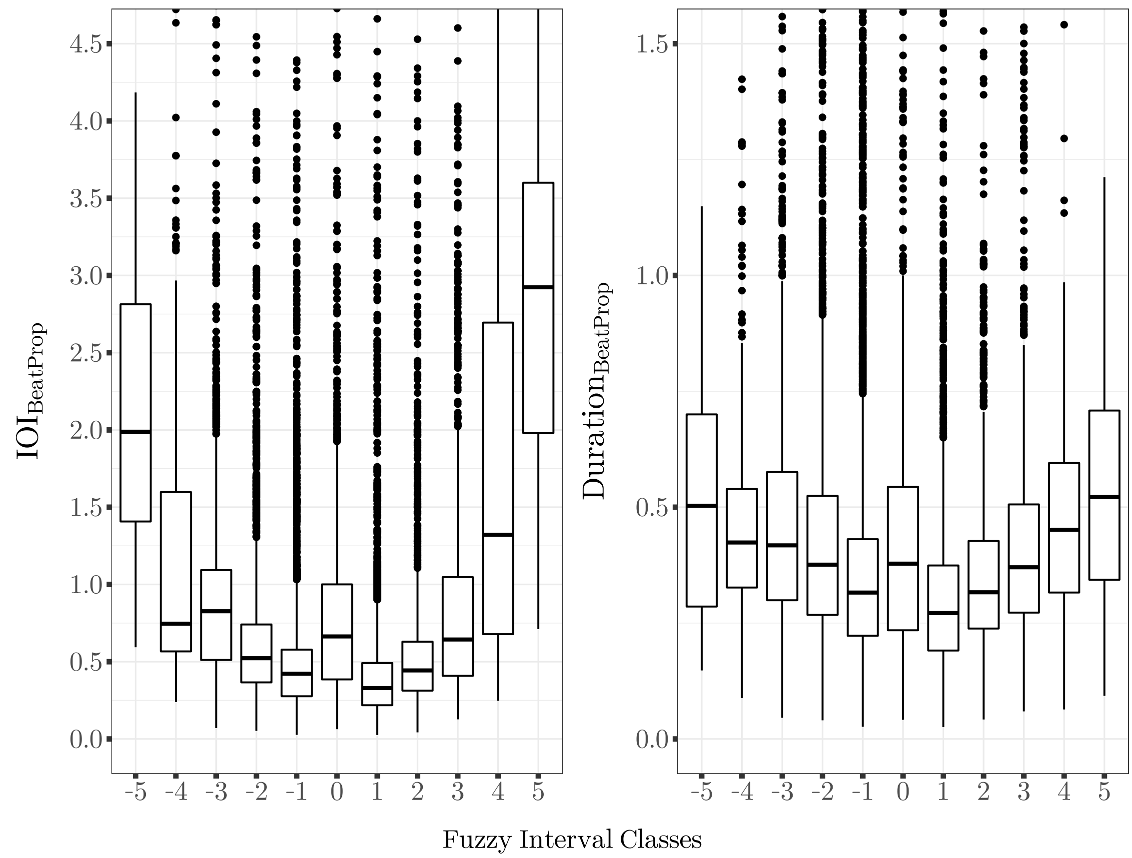

Figure 5.49: Distribution of $\mathrm{IOI}_{\mathrm{BeatProp}}$ and $\mathrm{duration}_{\mathrm{BeatProp}}$ for each fuzzy interval class in Green's corpus.

The overall trend for each distribution was similar, with longer notes at the largest intervals, tending to decrease towards the smallest intervals, with a small increase for repeated notes. However, there were substantial differences between the two distributions.[36] The largest differences were in the extremes of the interval classes, with the $\mathrm{IOI}_{\mathrm{BeatProp}}$ classes of -5, 4, and 5 being substantially longer than any of the other classes. In comparison, while the same shape of distributions was seen for the $\mathrm{duration}_{\mathrm{BeatProp}}$, the absolute differences were less extreme. The large values observed in the extreme $\mathrm{IOI}_{\mathrm{BeatProp}}$ classes were influenced by the fact that larger intervals were more likely to be played between phrases.

[35] $\mathrm{IOI}_{\mathrm{BeatProp}}$: Subsequent post-hoc tests with Tukey's HSD procedure found significant pairwise differences for all comparisons at $p < .001$ except 3 vs. 0 ($p = .0015$) and 3 vs. -4 ($p = .0047$). 4 vs. -5, 2 vs. -1, and -3 vs. 3 and -4 were not found to be significantly different. $\mathrm{Duration}_{\mathrm{BeatProp}}$: Subsequent post-hoc tests with Tukey's HSD procedure found significant pairwise differences at $p < .05$ except: 5 vs. -5, -4, -3, -2, -1, 0, 3, 4; -5 vs. -4, -3, -2, -1, 0, 3, 4; 4 vs. -4, -3, 3; -4 vs. -3, -2, 0, 3; 3 vs. -2, 0; and -2 vs. 0.

[36] The $\mathrm{IOI}_{\mathrm{BeatProp}}$ of a note was always the same or longer than the $\mathrm{duration}_{\mathrm{BeatProp}}$, as the IOI measured both the notes duration and any time before the following note.

The general trend for Green's $\text{IOI}_{\text{BeatProp}}$ and $\text{duration}_{\text{BeatProp}}$ was that he tended to play notes with smaller intervals for a shorter amount of time. The largest intervals had the longest lengths, while Green's repeated notes tended to be longer than his smallest intervals, with note lengths between leaps and jumps. The differences between the $\text{IOI}_{\text{BeatProp}}$ and $\text{duration}_{\text{BeatProp}}$ indicated that while Green's notes had overall similar durations. However, when he played larger intervals there was a longer length of time between the onsets of the two notes that comprised the interval. The results of this analysis supported the hypothesis that there was a relationship between the fuzzy interval and the note length of the first note of the interval. The analysis found that larger intervals between notes tended to have longer $\text{IOI}_{\text{BeatProp}}$, with the $\text{duration}_{\text{BeatProp}}$ tending to be only slightly longer.

Fuzzy Intervals vs. Normalised Instrument Tessitura Pitch

The following analysis focused on the relationship between the fuzzy intervals and the NITP of the notes Green played. As discussed in Section 5.1.1, sub-section on octave transitions, at the extreme ranges of the guitar there was a higher chance of Green changing octaves. Therefore, it was hypothesised that larger intervals were more likely to be played at the extreme ranges of the guitar. Specifically, that large descending intervals were more common at the higher register, and large ascending intervals were more common in the lower register. The distribution of NITP for each fuzzy interval class is displayed in Figure 5.50. This data indicated that there was little difference in the NITP distribution of the descending fuzzy interval classes. The median NITP ranged from 0.70, for >8ve↓, to 0.61 for step downs. The median for repeated notes, step, and leap ups all had a median NITP only slightly lower, 0.57. In comparison, the three largest ascending intervals had the lowest median values, and the largest difference between the classes (jump up: 0.52; big jump up: 0.45; >8ve↑: 0.34).

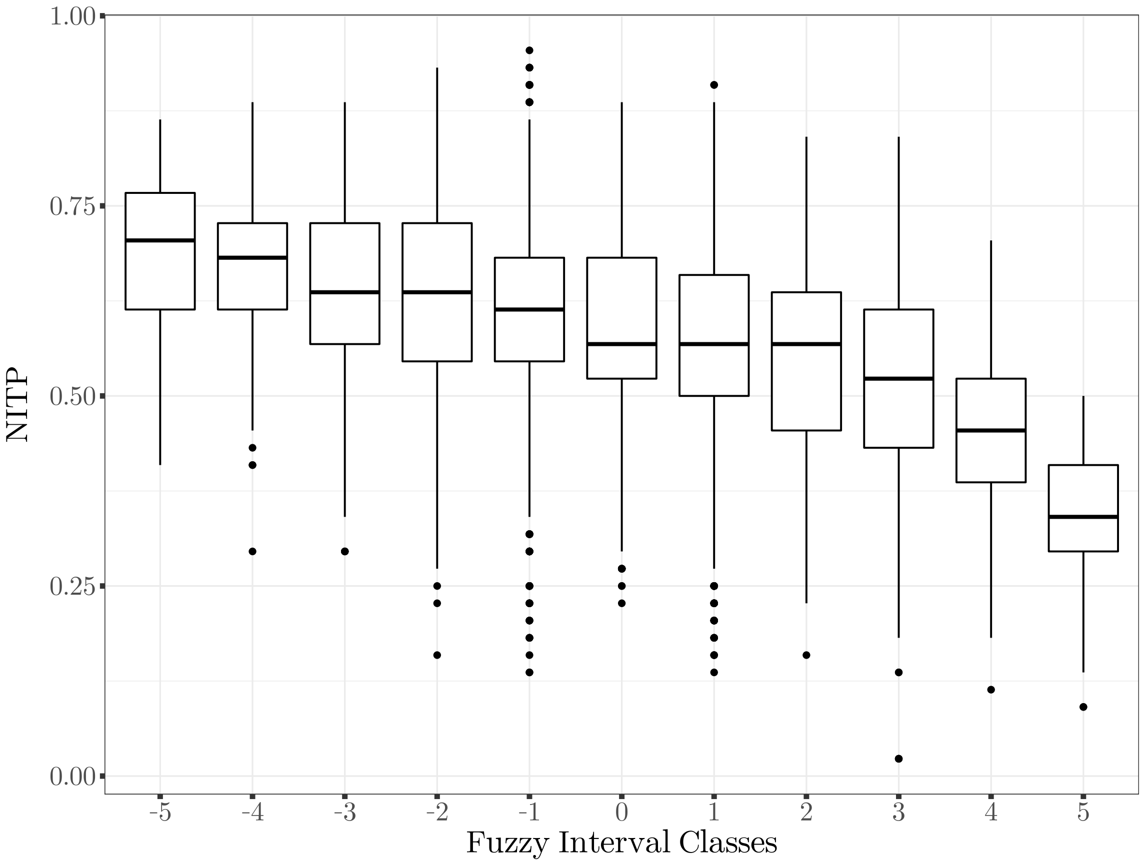

Figure 5.50: Distribution of NITP for each fuzzy interval class in Green's corpus.

An ANOVA was run to investigate the relationship between the fuzzy interval between two notes and the NITP of the first note. The ANOVA found a significant relationship between these two features, with a medium effect size ($F(10, 20427) = 218.99,\ p < .001;\ \eta^2 = .10$). Subsequent post-hoc tests using Tukey's HSD procedure found no significant differences between the following pairwise comparisons: `-5` vs. `-4`, `-3`, `-2`, and `-1`; `-3` vs. `-2` and `-4`; and `1` vs. `0`. Significant difference were found for `1` vs. `-5` at $p = .005$ and `0` vs. `-5` at $p = .006$, and for the remaining forty-six (of fifty-five) comparisons at $p < .001$. These results indicated that although Green did play descending intervals more frequently in the upper register of the guitar, there was no preference for one interval size over another. The lower Green played on the guitar, the more likely it was that he played a large ascending interval. The smallest intervals, either ascending or descending, were played throughout nearly the entire range of the guitar, with their median NITP being marginally lower than Green's overall median NITP (0.59).

5.4.4 Intervals Summary

This section investigated the use of intervals and intervallic structures in Green's improvisations. Comparing Green's Parsons distribution to the WJazzD found a significant difference between the two. While both distributions showed a preference for descending over ascending intervals, Green played a higher proportion of ascending intervals. The analysis also found that Green's ascending intervals tended to be larger than his descending intervals. Combined, these results suggested that Green likely started in a higher register, played a predominantly descending line, and the leapt up to descend again. Although repeated notes were relatively rare in Green's improvisations the analysis found that when Green played a repeated note, repeated intervals were more likely be played in the following 200 notes. This suggested that in improvisations where Green played repeated notes, this was a frequent element used throughout that improvisation, while other improvisations had limited repetition.

Within Green's corpus, around 80% of the raw intervals played had a value between a descending minor 3rd and an ascending minor 3rd. The two most frequent intervals Green played were an ascending and descending semitone. Of the smaller intervals ($<$ major 3rd), the least frequent in Green's improvisations were repeated notes, ascending seconds, and descending minor thirds. The analysis also found indications of the use of arpeggio structures within Green's improvisations.

The final section investigated how Green's fuzzy intervals were related to, or influenced by, other musical structures. The results of the analysis found a connection between both the duration$_{\text{BeatProp}}$ and IOI$_{\text{BeatProp}}$ with the fuzzy interval size. It found that when Green played a larger interval, or a repeated note, the note's length tended to be longer. The differences were larger for the IOI$_{\text{BeatProp}}$ than the duration$_{\text{BeatProp}}$. The majority of Green's notes had a fairly similar duration, often less than half a beat; however, when Green played a larger interval, there was often at least a beat between note onsets. The analysis also found that Green was more likely to play any size of descending interval at the higher registers of the guitar, while fuzzy intervals of a step were equally likely across the range of the guitar. The lower Green played on the guitar, the more likely he was to play a large ascending interval.

5.5 Examples

Building upon the previous analyses into Green's improvisational style in the pitch domain, the following sections investigated two specific examples of pitch-related improvisational features, surrounding note figures and voice leading.

5.5.1 Surrounding Note Figures

Surrounding note figures (SNF), or enclosures, are a common melodic improvisation technique where a target note is surrounded by pitches above and below. An example of an SNF is shown in Figure 5.51, from Green's improvisation over *Tico-Tico* (Solo 1, Green 1962m), with the SNF highlighted. There are multiple definitions of an SNF; however, for the purposes of this analysis SNFs were defined as:

- A series of three notes, all of different pitch;
- The final (target) note was surrounded by the pitches of the two previous notes, one above and one below;
- The interval difference between each surrounding note and the target note was a semitone or tone;
- The total $\mathrm{IOI}_{\mathrm{BeatProp}}$ of the surrounding notes was no greater than two beats.[37]

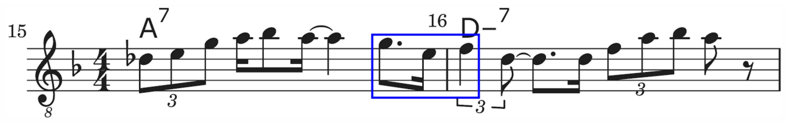

Figure 5.51: Example of a SNF, *Tico-Tico* (Solo 1, 1962), bars 15–16.

Using this definition, there were 1134 SNFs within the corpus of Green's improvisations. Therefore, 16.15% of all notes Green played were part of a SNF, either as a target tone or a leading tone. Although SNFs were often considered only in the context of targeting a chord tone (Hodges 2007, 139), this analysis initially considered all notes that were surrounded as defined. Of the 1134 SNFs within the corpus, 609 (53.70%) targeted an arpeggio tone, 251 (22.13%) targeted a scale tone, and 274 (24.16%) targeted NHTs. Therefore, the majority of SNFs that Green played were used to target chord tones.

The overall distribution of SNFs in Green's improvisations was investigated first. Following the definition above, there were four SNF patterns: a target tone surrounded either side by a semitone (`1 1`, e.g. D♭ →B →C); a semitone then a tone (`1 2`, e.g. D♭ →B♭ →C); a tone then a semitone (`2 1`, e.g. D →B →C); or by two

[37]When selecting the notes in code there was a small, 5%, buffer for the total $\mathrm{IOI}_{\mathrm{BeatProp}}$.

tones (`2 2`, e.g. D →B♭ →C).[38] Table 5.15 shows the distribution of the surrounding note interval patterns Green played before the target note. The most frequent interval pattern was `2 1`, a tone above or below the target note followed by a semitone approach (46.30%). The majority of SNFs resolved to the target tone chromatically (`1 1` or `2 1`), as the tension and release of chromatic resolutions resulted in a strong and musically pleasing resolution.

Table 5.15: Distribution of SNF interval patterns in Green's corpus.

	Interval Pattern[1,2]			
	`1 1`	`1 2`	`2 1`	`2 2`
Count	189	228	525	192
Percent	16.67%	20.11%	46.30%	16.93%

[1] Order of intervals is n m, where n is two notes before target tone and m is one.
[2] `1`: Semitone, `2`: Tone

The SNF interval patterns were then separated into interval sequences based on the size and direction of the surrounding notes. This split the initial four classes into eight, with each SNF interval pattern splitting into an above then below (`ab`) and below then above (`ba`) interval sequence.[39] Table 5.16 shows Green's distribution of the SNF interval sequences. The data indicated that the two most frequent SNF interval sequences Green played were `2 1 ab` and `2 1 ba`. The third most frequent interval sequence, `1 1 ab`, was drawn from the least frequent interval pattern (`1 1`).

Table 5.16: Distribution of SNF interval sequences in Green's corpus.

	Interval Sequence[1,2]							
	`1 1 ab`	`1 1 ba`	`1 2 ab`	`1 2 ba`	`2 1 ab`	`2 1 ba`	`2 2 ab`	`2 2 ba`
Count	170	19	119	109	336	189	67	125
Percent	14.99%	1.68%	10.49%	9.61%	29.63%	16.67%	5.91%	11.02%

[1] Order of intervals is n m where n is two notes before target tone and m is one.
[2] `1`: Semitone, `2`: Tone, `ab`: above then below target note, `ba`: below then above target note.

[38] While all examples given followed the pattern of a note above the target note followed by a note below, the inverse was also valid.
[39] For example, the `2 1` interval pattern became `2 1 ab` (D →B →C) and `2 1 ba` (B♭ →D♭ →C).

As SNFs could be played in a variety of situations, there were features that were likely to have influenced Green's use of SNFs. Two features of interest were: the relationship between the target note and the chord of the moment; and metrical information regarding the notes of the SNF, including the metrical placement of the target note, and the rhythm of all three notes in the SNF. The chords investigation also included analysis of the frequency of SNFs over the three main chord types, and the $\text{CPC}_{\text{Weight}}$ and CDPCX of the target notes.

Surrounding Note Figures and Chords

The initial analysis investigated whether Green played the target notes of SNFs more frequently over certain chord types. The hypothesis was that if Green favoured SNFs over specific chords, there would be a higher proportion of target notes over those chords. The proportion of notes of the three main chord types that were or were not SNF target notes is shown in Figure 5.52. This data showed that around half of each chords notes were SNFs, and a χ^2-test found no significant difference between the distributions ($\chi^2(2) = 3.58,\ p = .167,\ V = .01$). These results indicated that Green did not play target notes of SNFs more frequently over one chord type.

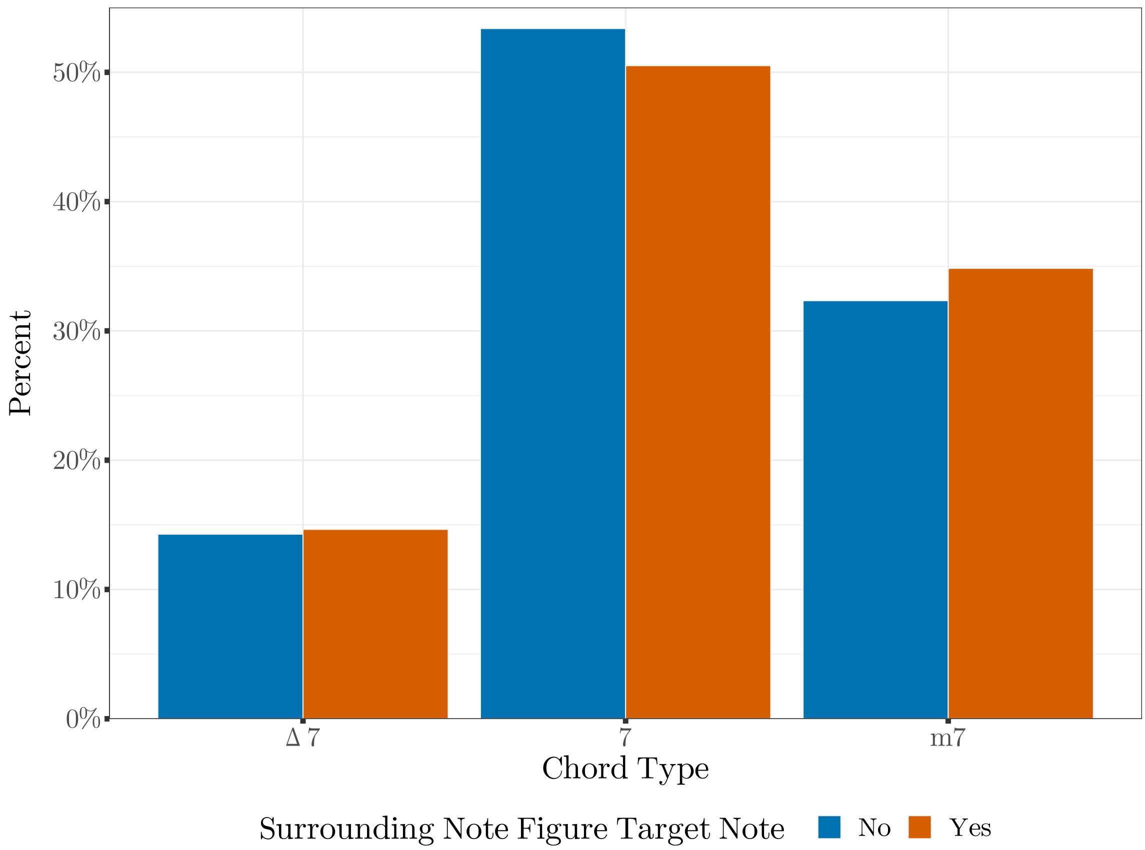

Figure 5.52: Distribution of SNF target tones in Δ7, 7, and m7 chords in Green's corpus.

As SNFs were equally likely to be played by Green over each chord type, the distribution of SNF interval sequences against the chord types was investigated, this data is shown in Table 5.17. The `2 1 ab` was the most frequent SNF for each of the chord types, while the next most common (> 10%) SNFs for each chord were:

- Δ7: `1 1 ab`, and `1 2 ba`;
- 7: `2 1 ba`, `2 2 ba`, `1 1 ab`, and `1 2 ab`;
- m7: `2 1 ba`, `1 1 ab`, `1 2 ab`, and `1 2 ba`.

Although Green played the `1 1 ab` SNF often over each chord type, the least frequently played SNF was the inverse `1 1 ba`. A χ^2-test found a significant difference in the distribution of SNF interval sequences between the chord types, with a small effect size ($\chi^2(14) = 39.32,\ p =< .001,\ V = .14$). Subsequent post-hoc tests found significant pairwise differences for all comparisons (Δ7 vs. 7: $p = .007$; Δ7 vs. m7: $p = .023$; 7 vs. m7: $p = .010$). These results indicated that although the most and least frequent SNF interval sequence Green played was the same across the chord types, Green did favour certain SNFs over specific chords. These differences were likely influenced by the scalic and arpeggio structures of the different chord types.

Table 5.17: Distribution of SNF interval sequences over Δ7, 7, and m7 chords in Green's corpus.

	Interval Sequences							
Chord Type	`1 1 ab`	`1 1 ba`	`1 2 ab`	`1 2 ba`	`2 1 ab`	`2 1 ba`	`2 2 ab`	`2 2 ba`
Δ7								
Count	34	1	11	20	50	14	12	14
Percent	21.79%	0.64%	7.05%	12.82%	32.05%	8.97%	7.69%	8.97%
7								
Count	79	10	58	45	138	100	28	80
Percent	14.68%	1.86%	10.78%	8.36%	25.65%	18.59%	5.20%	14.87%
m7								
Count	44	7	44	38	128	59	23	28
Percent	11.86%	1.89%	11.86%	10.24%	34.50%	15.90%	6.20%	7.55%

Figures 5.53, 5.54, and 5.55 show examples of Green's most frequent SNF interval pattern, `2 1 ab`, over each of the chord types. Figure 5.53 is an excerpt over a B∆7 chord from Green's improvisation over *Oleo* (Solo 1, Green 1962i). Figure 5.54 shows an excerpt of a SNF starting on an F7 chord and targeting the ♭7 of the following B♭7 chord from Green's improvisation over *Sonnymoon For Two* (Green 1960c). Finally, Figure 5.55 shows the `2 1 ab` SNF targeting the ♭3 of the Cm7 chord, starting on the previous B○7, from Green's improvisation over *The Surrey With The Fringe On Top* (Solo 1, Green 1963g).

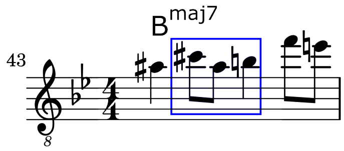

Figure 5.53: `2 1 ab` SNF interval sequence targeting a note over a ∆7 chord, *Oleo* (Solo 1, 1962), bar 43.

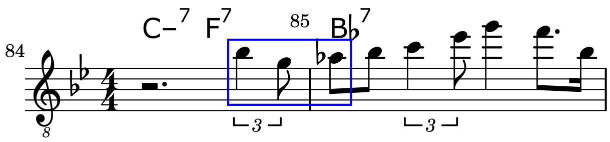

Figure 5.54: `2 1 ab` SNF interval sequence targeting a note over a 7 chord, *Sonnymoon For Two* (1960), bars 84–85.

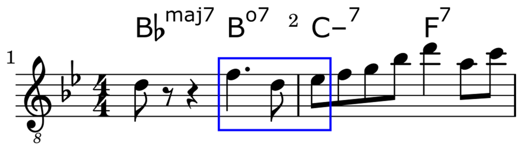

Figure 5.55: `2 1 ab` SNF interval sequence targeting a note over a m7 chord, *The Surrey With The Fringe On Top* (Solo 1, 1963), bars 1–2.

The following analyses focused on the six most frequent SNF interval sequences, `1 1 ab`, `1 2 ab`, `1 2 ba`, `2 1 ab`, `2 1 ba`, and `2 2 ba`. The distribution of $\text{CPC}_{\text{Weight}}$ of the target notes in the SNFs, shown in Table 5.18 was investigated next. This table shows, for each of the most frequent SNFs of each chord type, the proportion of target notes of each $\text{CPC}_{\text{Weight}}$.[40] This data showed that for most SNFs and chord types either the majority or plurality of target notes came from the arpeggio of the chord. The exceptions to this were: `1 1 ab` for both 7 and m7 chords, where NHTs were more frequently targeted; and `1 2 ab` and `2 1 ba` for m7 chords, where scale tones were more frequently targeted. Overall, 56.31% of target notes came from the arpeggio, with arpeggio tones slightly more frequent over ∆7 (64.42%) and 7 (60.00%) chords, while just under half of (48.24%) SNFs over m7 chords were from

[40]The blank (-) data indicated SNFs where the interval sequence did not occur > 10% of the time over that chord type.

the arpeggio. Scale tones were rarely targeted by Green over Δ7 and 7 chords, with the targeting of NHTs likely incidental within Green's improvisations.

Table 5.18: Distribution of $\mathrm{CPC}_{\mathrm{Weight}}$ for each SNF interval sequence target note and chord type in Green's corpus.

	`1 1 ab`	`1 2 ab`	`1 2 ba`	`2 1 ab`	`2 1 ba`	`2 2 ba`
Δ7						
Arpeggio	82.35%	-	80.00%	46.00%	-	-
Scale	2.94%	-	10.00%	28.00%	-	-
NHT	14.71%	-	10.00%	26.00%	-	-
7						
Arpeggio	40.51%	70.69%	-	44.93%	67.00%	88.75%
Scale	13.92%	18.97%	-	10.87%	21.00%	5.00%
NHT	45.57%	10.34%	-	44.20%	12.00%	6.25%
m7						
Arpeggio	22.73%	6.82%	81.58%	68.75%	32.20%	-
Scale	20.45%	88.64%	0.00%	5.47%	49.15%	-
NHT	56.82%	4.55%	18.42%	25.78%	18.64%	-

As arpeggio tones were generally the most frequent targeted tones, the final investigation analysed how frequently each chord tone was played as the target of an SNF.[41] Table 5.19 shows, for each of the common SNFs of each chord type, the frequencies of chord tones that were targeted by Green. Over Δ7 and m7 chords Green most frequently targeted the tonic and 3$^{\text{rd}}$, while all 7 chord tones were equally targeted, although with a slight preference for the third. The 7$^{\text{th}}$ was the least commonly targeted chord tone, with the 5$^{\text{th}}$ also not frequently targeted by Green over Δ7 and m7 chords. A χ^2-test found a significant difference in the distribution of arpeggio target tones between the chords, with a medium effect size ($\chi^2(6) = 66.87,\ p =< .001,\ V = .26$).[42]

[41] Future research should investigate how each individual SNF targeted different chord tones. The limited data for each combination of chords, SNFs, and arpeggio target notes meant there was not enough data to draw meaningful conclusions in this research.

[42] Subsequent post-hoc tests found significant pairwise differences for all comparisons at $p < .001$.

Table 5.19: Distribution of SNF arpeggio target notes for $\Delta7$, 7, and m7 chords in Green's corpus.

	Arpeggio CDPCX			
	1	3	5	7
$\Delta7$				
Count	37	22	2	6
Percent	55.22%	32.84%	2.99%	8.96%
7				
Count	65	80	70	58
Percent	23.81%	29.30%	25.64%	21.25%
m7				
Count	49	74	22	6
Percent	32.45%	49.01%	14.57%	3.97%

This analysis found that although SNFs were equally likely to be played by Green over all chord types, the chord types did have an influence on how Green constructed his SNFs. This included the frequency of different SNF interval sequences, and the CDPCX of the the target tones. Overall, Green also most frequently targeted arpeggio tones with his SNFs. Of the arpeggio tones, Green most frequently targeted the tonic and 3rd over $\Delta7$ and m7 chords. The 3rd was slightly favoured over 7 chords, although the distribution of chord tones was more even.

Surrounding Note Figures and Metre

This sub-section focused on the interaction between metre-based features and Green's use of SNFs. It was hypothesised that target tones were more likely to be played on the beat, and on metrically important beats. Following this analysis, the rhythm of Green's SNFs, as $\text{IOI}_{\text{BeatProp}}$, were investigated. The distribution of metrical weights for notes that were or were not the target notes of an SNF are shown in Table 5.20. This data showed that, although the majority of SNF target notes were played off the beat, Green was more likely to play target notes on metrically strong beats compared to non-target notes. While nearly two-thirds of non-target notes were played off the beat, just over half of target notes were. Consequently, just under 50% of target notes were played on the beat, with the largest differences observed in metrically strong beats. A χ^2-test found a significant difference in the metrical weight distribution between non-target and SNF target notes, with a small effect size ($\chi^2(2) = 65.23,\ p =< .001,\ V = .06$). These results supported the hypothesis that Green was more likely to play SNF target notes on metrically strong beats, compared to the other notes he played.

Table 5.20: Distribution of metrical weights for notes which were or were note the target note of a SNF.

	Strong	Weak	Off
Non-target	19.28%	17.57%	63.15%
SNF Target	27.69%	20.46%	51.85%

Following the analysis of the metrical placement of the target notes, the $\mathrm{IOI}_{\mathrm{BeatProp}}$ of the three notes that comprised the SNFs was investigated. Due to the selection criteria of SNFs, there were limits on the lengths of notes that were considered. Figure 5.56 shows the distribution of $\mathrm{IOI}_{\mathrm{BeatProp}}$ for each of the surrounding notes individually and summed.

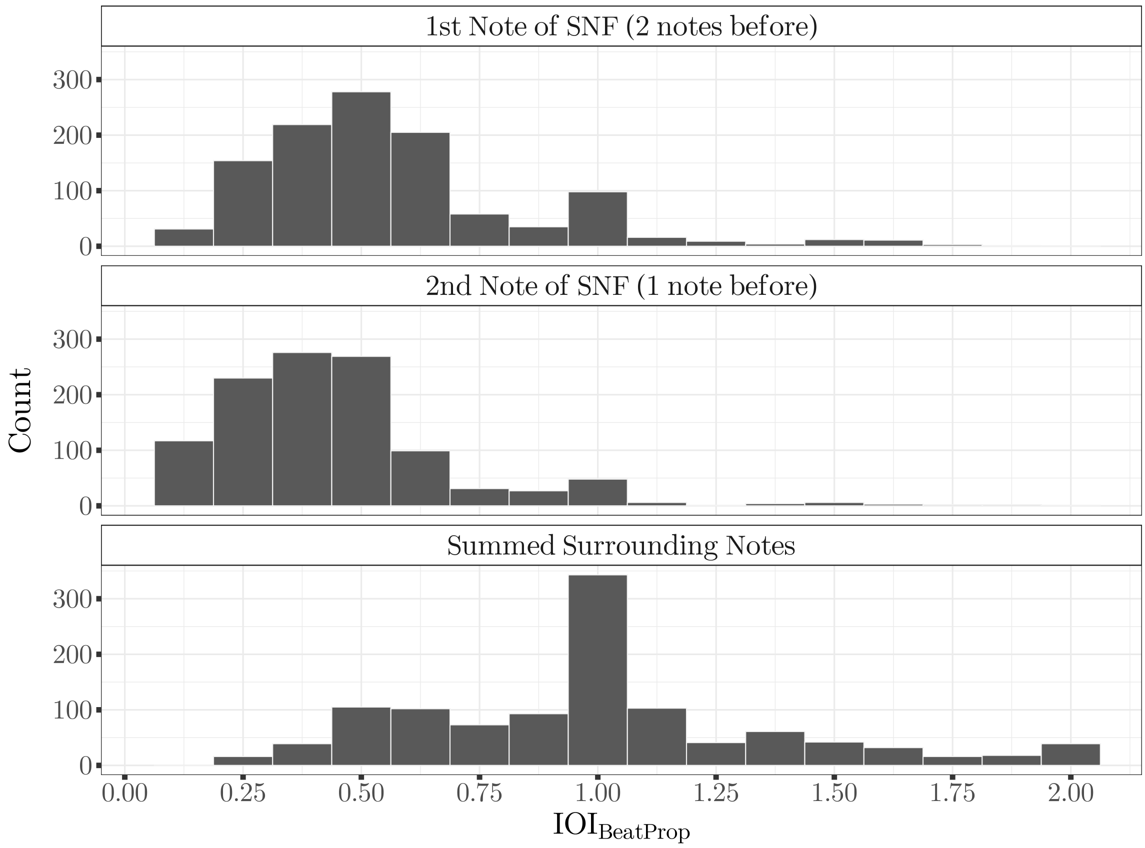

Figure 5.56: $\mathrm{IOI}_{\mathrm{BeatProp}}$ distribution for surrounding notes in a SNF.

The surrounding notes most frequently went for around half a beat, with the first note tending to be slightly longer, likely due to swing (1st note: $\bar{x} = 0.57 \pm 0.29$

beats; 2nd note: $\bar{x} = 0.44 \pm 0.26$ beats). The combined data indicated that Green's surrounding notes tended to be around one beat ($\bar{x} = 1.01 \pm 0.40$ beats), half the limit set by the selection criteria. Figure 5.57 shows an example of a SNF with this most frequent rhythm, from Green's improvisation over *Green With Envy* (Solo 2, Green 1961j).

Figure 5.57: Example of a SNF with a quaver-equivalent rhythm, *Green With Envy* (Solo 2, 1961), bars 47–49.

The final investigation focused on the $IOI_{BeatProp}$ of the target note. It was hypothesised that the target note of the SNF would tend to have a similar or slightly longer length compared to the leading notes, as they were most likely played throughout the course of an improvised line. Figure 5.58 shows the distribution of SNF target tones, with an $IOI_{BeatProp}$ between 0 and 2, which included 96.47% of all target tone data. This data showed that the majority of target notes (65.34%) had an $IOI_{BeatProp}$ between a semiquaver-equivalent (0.25 beats) and dotted quaver-equivalent (0.75 beats), with a mean $IOI_{BeatProp}$ of 0.67 ± 0.71 beats. The data also showed a small spike of data around an $IOI_{BeatProp}$ of 1, indicating that the target note did not infrequently have a crotchet-equivalent length.

Finally, Figure 5.59 shows the most frequent $IOI_{BeatProp}$ transitions that occurred within Green's improvisations. The thickness of the lines indicated the frequency of each trigram and colour represented the $IOI_{BeatProp}$ of the target note. Only trigrams that occurred more than ten times were plotted, with the $IOI_{BeatProp}$ binned to a demisemiquaver-equivalent centred on the nominal note length (e.g. 0.125 ± 0.0625). This data showed that the most frequent trigram was for all three notes of the SNF to have an $IOI_{BeatProp}$ of 0.5, a quaver-equivalent line. Also common was the swung version of the quaver-equivalent line (0.626→0.375→0.625 beats), and semiquaver-equivalent line (0.25 beats each). These results indicated that Green most frequently played his SNFs throughout his improvised lines. Figure 5.60 shows a phrase from Green's improvisation over *Green With Envy* (1961j), which contained four SNFs.

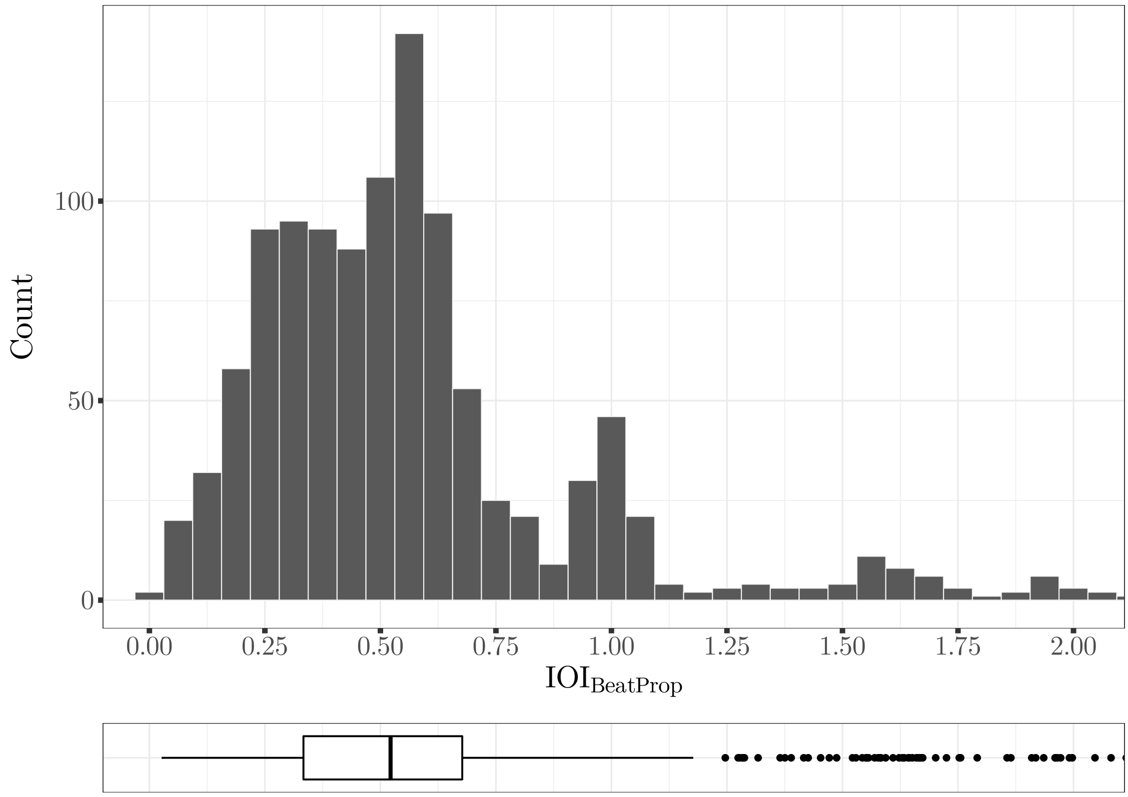

Figure 5.58: Distribution of target tone $IOI_{BeatProp}$ in SNFs.

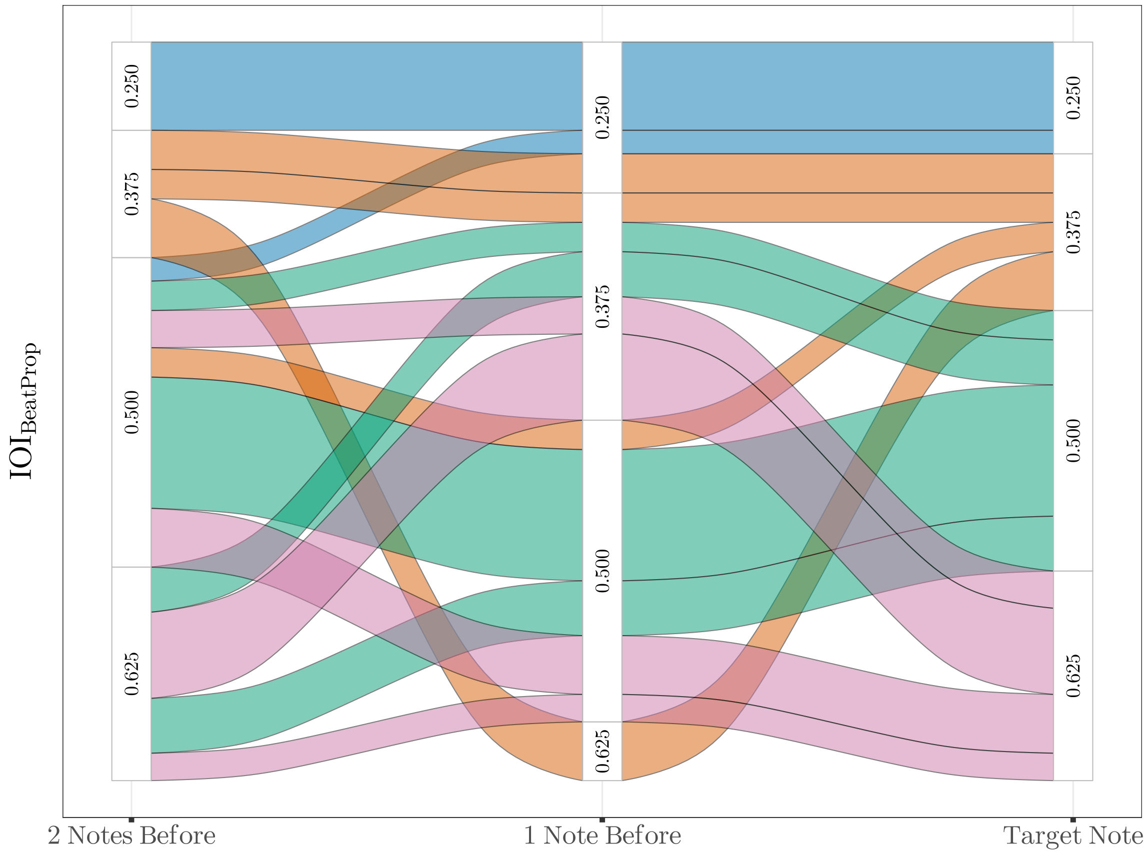

Figure 5.59: $IOI_{BeatProp}$ transitions for SNFs, where each trigram occurred more than ten times.

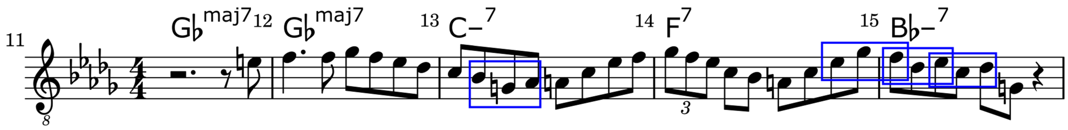

Figure 5.60: Example of Green playing SNFs throughout his improvised lines, *Green With Envy* (Solo 2, 1961), bars 11–15.

This analysis found that although the majority of target tones in a SNF were played off the beat, when compared to non-target tones, Green played a higher proportion on metrically strong and weak beats. This analysis also found that the $IOI_{BeatProp}$ of both the leading notes and target note tended to be around half a beat, suggesting that Green played SNFs throughout his improvised lines.

Surrounding Note Figure Summary

These analyses investigated the use of SNFs within Green's improvisations. While not exceptionally common within Green's corpus, around one in six notes were part of a SNF, as either a leading tone or as a target tone. The most frequent CPC_{Weight} of the target tone in Green's improvisations was an arpeggio tone. For Δ7 and m7 chords, the tonic and 3rd were most commonly targeted, while every arpeggio tone was equally targeted by Green over 7 chords. The most frequently played SNF pattern was `2 1 ab`, with `2 1 ba` second most common for both 7 and m7 chords, and the `1 1 ab` SNF pattern was second for Δ7 chords. When compared to non-target tones, Green was more likely to play a target tone on the beat, with more than a quarter played on metrically strong beats. This analysis also found that Green usually played SNFs within a line, with the notes tending to have a quaver-equivalent $IOI_{BeatProp}$.

5.5.2 Voice Leading

The term voice leading in this project was used to refer to the transitions between notes at the point of chord changes. This section investigated how Green used voice leading within his improvisations, focusing on how the intervals, metrical weight, and CPC_{Weight} differed depending on whether or not notes were identified as being part of a voice leading pair.

From all the notes Green played in his improvisations, there were 3245 (15.85%) occurrences of a note followed by another note that was played over a different nominal chord. Not all of these could have been considered voice leading as the following note, while played over a different chord, may not have been played over the subsequent chord, or the length of the note or space between notes, may have been too long. To determine which of the notes were valid points of voice leading a combination of features were used to refine the selection, these criteria were[43]:

- The following note was played over the next chord, according to the nominal chord changes;
- The note was not the last note in a phrase;
- The note's $\text{IOI}_{\text{BeatProp}}$ was ≤ 2;
- The note's $\text{Duration}_{\text{BeatProp}} \div \text{IOI}_{\text{BeatProp}}$ was ≥ 0.5 (i.e. the note was played for at least half of the time between the onset of the first and second note).

For the analysis of these notes only events where both notes were played over a Δ7, m7, or 7 chord were included. With these restrictions, only 1723 voice leading events remained, 53.10% of those originally detected. The analysis of the voice leading focused on the following features: the Parsons code and intervals used; the metrical weight of the notes; and the $\text{CPC}_{\text{Weight}}$ of the notes.

Voice Leading and Intervals

It was hypothesised that the intervals Green played at moments of voice leading would be different from the other intervals he played. Specifically, it was hypothesised that Green would play fewer repeated notes or large intervals (> perfect 4th) . The Parsons code of the notes was investigated first, with the distribution of Parsons for notes that were part of a voice leading pair compared to the distribution of all other notes. These distributions can be seen in Table 5.21. This data suggested that, as hypothesised, Green was less likely to play repeated notes at points of voice leading. Instead, Green was slightly more likely to play descending intervals, with just under half (49.10%) of all voice led notes descending, while he played a similar proportion of ascending intervals. A χ^2-test found a significant difference in the Parsons code distributions for voice led notes compared to all other notes, with a small effect size ($\chi^2(2) = 17.21,\ p =< .001,\ V = .03$).

Table 5.21: Distribution of Parsons codes for notes that were or were not voice led.

	Descending	Repetition	Ascending
Voice Leading	49.10%	4.47%	46.43%
No Voice Leading	46.73%	7.04%	46.22%

To investigate the other hypothesis, that Green played smaller intervals more frequently, the intervals had to be categorised. For this analysis, three categories

[43]The criteria applied only to the first note of the two notes involved in the voice leading.

were used: repeated notes; small intervals (< perfect 4^{th}); and large intervals ($\geq$ perfect 4^{th}). The distribution of these classes dependent on whether or note the note was part of a voice led pair is shown in Table 5.22. The data in this table showed that, although the majority of all notes moved through small intervals, Green was slightly more likely to use small intervals when voice leading (85.72% vs. 80.29%). A χ^2-test found a significant difference in these distributions, with a small effect size ($\chi^2(2) = 31.47,\ p =< .001,\ V = .04$). These results supported the hypothesis that Green played fewer large intervals or repeated notes at moments of voice leading.

Table 5.22: Distribution of interval size classes for notes that were or were not voice led.

	Repetition	Small	Large
Voice Leading	4.47%	85.72%	9.81%
No Voice Leading	7.04%	80.29%	12.67%

Voice Leading and Metrical Weight

The following analysis focused on the metrical weight of the target tone in Green's voice leading. It was hypothesised that when voice leading, the target tones played by Green were more likely to occur on metrically strong beats. Of all the target tones in Green's corpus 96.69% were played in the beat of the chord change, with 82.59% of these played on the beat. Consequently, 79.86% of all target tones were played on the beat of a chord change. As the majority of chord changes occurred on metrically strong beats, 79.80% of all target notes were played on metrically strong beats. In contrast, 87.35% of all the leading notes were played off the beat, with nearly all others played on a metrically weak beat.[44]

The leading tone to target tone metrical weight bigrams for voice led notes in Green's improvisations, when compared to every other pairwise metrical weight sequence, is shown in Figure 5.61. This data showed that the vast majority of all metrical weight sequences for voice led notes were from an off-beat leading note to a metrically strong target note. The metrical weight sequence distributions for voice led tones compared to all other notes was significantly different, with a large effect size ($\chi^2(8) = 4650.32,\ p =< .001,\ V = .48$). In comparison to the voice led notes, when Green played notes that were not part of a voice leading sequence, the plurality were off-beat to off-beat.

[44] There were only seven occurrences where Green played the leading note on a metrically strong beat.

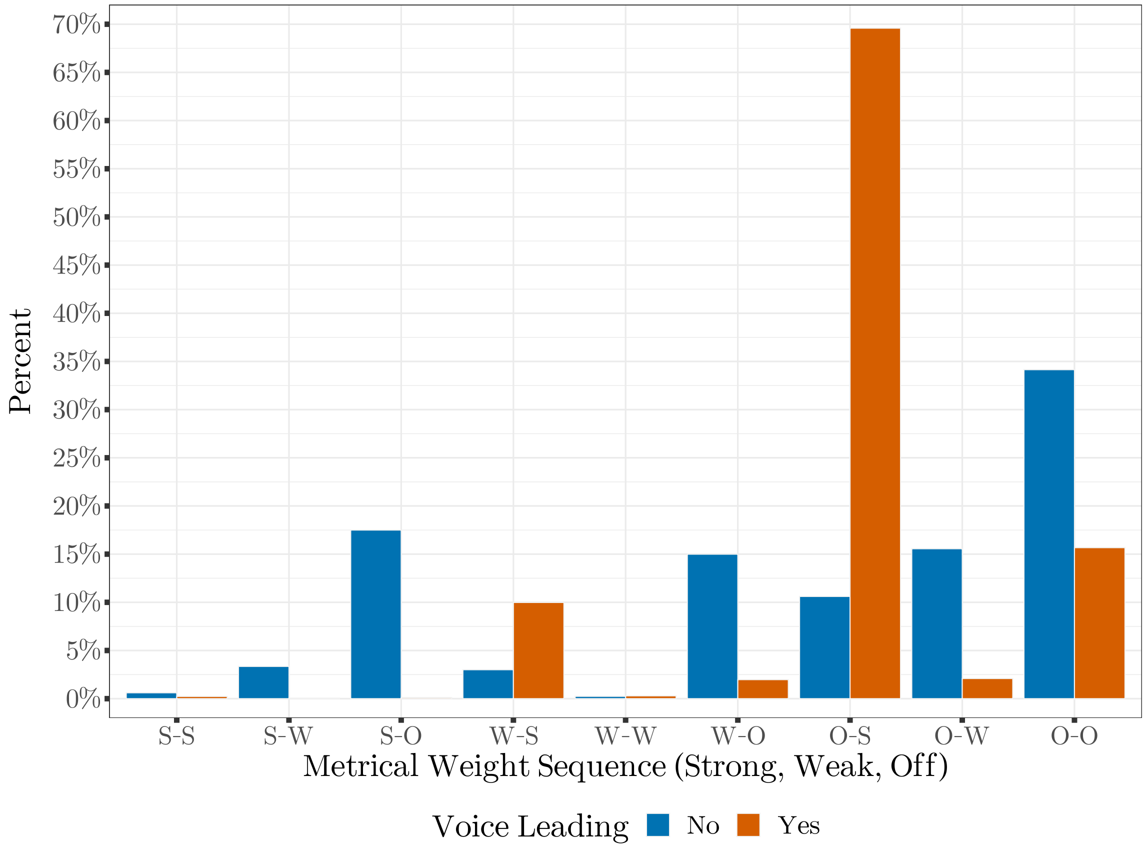

Figure 5.61: Distribution of pairwise metrical weight sequences for bigrams which were or were not voice led.

These results indicated that the metrical weights played by Green for notes identified as being in a voice leading pair were significantly different than when he was normally improvising. Green rarely targeted off-beat notes, as the vast majority of sequences moved from an off-beat note to a note on a metrically strong beat. However, these results were heavily influenced by the voice leading criteria set for this research. While a broader definition of voice leading would affect these results, a similar trend would be expected. This would be due to the vast majority of all notes being played off the beat, and target notes often associated with metrically strong positions. These results confirmed the hypothesis that the target tones of a voice leading pair are most frequently played on a metrically strong beat, approached by a note played off the beat, and that Green conformed to this practice.

Voice Leading and $\mathbf{CPC_{Weight}}$

This section focused on analysing Green's use of $\mathrm{CPC_{Weight}}$ transitions when voice leading. It was hypothesised that Green's target notes were more likely to be arpeggio tones. The $\mathrm{CPC_{Weight}}$ distribution of the target tones compared to non-target tones is shown in Figure 5.62. This data showed that Green played slightly more arpeggio tones as target notes, with the largest decrease coming from scale tones. However, a χ^2-test did not find a significant difference in the $\mathrm{CPC_{Weight}}$ distribution dependent on whether or not a note was the target tone in a voice leading pair ($\chi^2(2) = 5.15,\ p = .076,\ V = .02$). Therefore, the data did not support the hypothesis that Green targeted arpeggio tones when voice leading at a significantly higher rate compared to other notes when improvising.

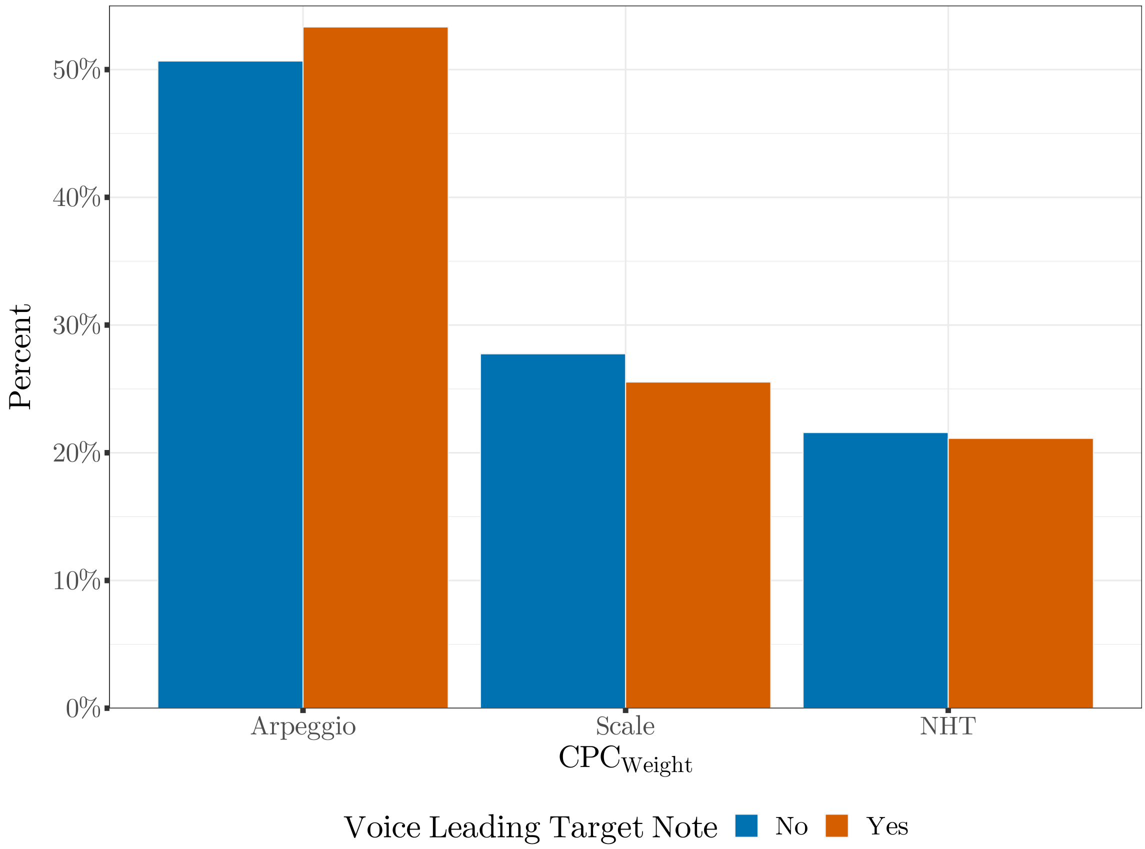

Figure 5.62: Distribution of $\mathrm{CPC_{Weight}}$ depending on whether when a note was the target of voice leading.

Voice Leading Summary

This section investigated the use of voice leading in Green's improvisations. It specifically focused on analysing how Green's use of intervals, metrical weight, and $\mathrm{CPC_{Weight}}$ differed depending on whether the notes were part of a voice leading pair. The analyses found that Green played a higher proportion of descending intervals at points of voice leading. Green also more frequently played smaller intervals when voice leading. Green rarely targeted off-beat notes when voice leading, with most voice leading target tones played on metrically strong beats. Finally, while Green was marginally more likely to play arpeggio tones as the target note in a voice leading pair, the $\mathrm{CPC_{Weight}}$ distribution for target tones compared to all other notes was not significantly different. In summary, this analysis found that in many cases

Green did treat notes when voice leading slightly differently from the rest of the notes he played, although these differences were often subtle or influenced by the selection criteria.

5.6 Green's Improvisational Style In The Pitch Domain

This chapter focused on features of Green's improvisational style within the pitch domain. Following the developed methodology, this chapter first presented analyses on broader pitch feature categories, including the raw pitch, TPC, CPC, and intervals. This was followed by more detailed analyses of feature interactions within those categories. Finally, the chapter concluded with the analysis of two specific pitch based examples, SNFs and voice leading.

The raw pitch analysis found that Green predominantly improvised in the 4^{th} octave, with a substantial number of notes also played in the 5^{th} octave. The analysis found that the longer Green stayed in a single octave, the more likely it was that Green was playing a repeated note pattern with a small number of unique pitches, a common element of Green's improvisational style. While most note transitions occurred within octaves, octave transitions nearly always occurred between neighbouring octaves, often through arpeggio movement. The analysis also found that non-neighbouring octave transitions most frequently occurred between phrases, rather than within a phrase.

The TPC analysis found that the majority of the notes Green played were diatonic to the overall key of the piece. With many of Green's improvisations being in the hard bop or post-bop style, there was also evidence of blues influenced language within his improvisations. This was evident by the frequent use of the ♭3 blues note throughout Green's improvisations. However, Green rarely played TTs, the other blues note. Green also rarely played a sequence of three NDTs in a row, with most NDTs surrounded by DTs. Green favoured step-wise motion, often chromatic, to transition into and out of NDTs.

Similar to the TPC analysis, the CPC analysis found that the vast majority of all notes Green played were harmonic to the chord of the moment. Specifically, the majority of the notes came from the arpeggio of the chord. Around one-fifth of all notes were NHTs, with Green more frequently playing NHTs over 7 chords. HTs were most frequently played on the beat in Green's improvisations, with NHTs played off-beat. Of the HTs, Green played the 4^{th} more frequently than was expected for a diatonic non-harmonic tone, with the analysis suggesting that it was frequently played off-beat, and was likely due to the influence of the blues in Green's improvisations. The frequency of NHTs increased in Green's improvisations in the beats leading up to a chord change. Similar to NDTs, Green also favoured chromatic approaches and departures around NHTs, with three-quarters having a chromatic approach and/or departure. Of the altered chord tensions played over a 7 chord, Green favoured the ♭9, ♯9, and ♯5. Combined, these results found that Green treated NHTs significantly different from his HTs.

The analysis of intervals found that although Green slightly favoured descending intervals, he played around the same proportion of descending and ascending intervals. Green's ascending intervals were more likely to be larger, indicating predominantly descending movement followed by an ascending leap. Repeated notes were relatively rare within Green's improvisations. In the cases where Green did play a repeated note, he was substantially more likely to play another repeated note in the next 200 notes. This indicated that while repetition was not often present in all of his improvisations, there were specific solos where repetition was an important improvisational tool. The majority of Green's intervals were small, between an ascending and descending minor 3rd, with chromatic movement most common. Most of the thirds Green played were played over the same chord, indicative of arpeggio structures within his improvisations. The analysis also found that for larger intervals the gaps between the two note onsets tended to be longer, which suggested that Green most frequently played large intervals between phrases. Descending intervals were most common at the higher registers of the guitar, while large ascending intervals were most common in the lower registers; ascending and descending steps were equally likely across the range of the guitar.

The investigation into Green's use of SNFs found them to be frequent, if not common, in his improvisations, with Green often playing them throughout a line. The target note of a SNF was most frequently a note from the arpeggio of the surrounding chord, with the tonic and 3rd being the most common chord tones. The most frequent SNF patterns played by Green started a tone away from the target note, followed by a chromatic approach into the target note. A chromatic surround was also common in Green's improvisations. SNF target tones were also more commonly played on metrically strong beats in Green's improvisations when compared to other non-target tones.

The voice leading analysis focused on specific transitions that occurred around the points of chord change within Green's improvisations. It found that Green frequently descended into voice led target notes, consistent with the previous interval findings. Partly due to the constraints on classifying voice lead pairs, Green rarely targeted off-beat notes when voice leading, with the majority played on metrically strong beats. Green frequently targeted arpeggio tones when voice leading in his improvisations.

In summary, this analysis found that Green's improvisational style within the pitch domain was influenced by both the tonality mode of the piece and the chords he was improvising over. While the majority of all notes were diatonic and harmonic to the situations he was improvising over, the blues also had a strong influence on Green's note choice. Green favoured smaller intervals when improvising, with the majority of movements being step-wise or arpeggio based. SNFs were found fairly frequently within Green's improvisations, with many of the elements of the SNFs fitting with the previous analyses into pitch and interval usage. This was also true for Green's note choices at points of voice leading, with most of the target tones coming from the arpeggio of the surrounding chord while being approached chromatically. As was evident throughout this chapter with the inclusion of comparisons to note length and metrical weight, analysis of the pitches without context of their rhythmic features could only provide so much insight into Green's improvisational style. The following chapter investigated these rhythmic features in more detail.

Chapter 6

Rhythm Domain

The rhythm domain referred to any feature that was associated with the rhythmic details of the notes played. Along with the pitch domain, the rhythm domain was one of the predominant domains investigated in analysis and improvisational pedagogy. Though separated in this research, many features of the micro and macro domains could be considered a subset of the rhythm domain. A feature's inclusion in this or later chapters was not intended to exclude their effect or importance in those respective domains, but reflected where they were best situated within this research.

> It's not what you play, it's how you play it.
>
> Mary Lou Williams (McCann 2017, 75)

This quote, attributed to jazz great Mary Lou Williams, highlighted the importance of the features within the rhythm domain in shaping an improvisation. The rhythm domain encompassed features that described a wide range of rhythmic phenomena, including: the length of the notes; the placement and distribution of notes in the beat structure of the bars; the use of rests; the variety of the rhythmic sub-divisions played; and the density of notes played.

6.1 Note lengths

The length of the notes could be described by eight features, grouped into two categories: the duration of the notes; and the inter-onset interval (IOI) of the notes. The duration of the note was the time (seconds) from the onset of the note to the offset of the same note ($\text{note}_{\text{offset}} - \text{note}_{\text{onset}}$). The IOI was the time from the onset of one note to the onset of the following note ($\text{note}^{+1}_{\text{onset}} - \text{note}_{\text{onset}}$), equal to the duration of the note and any rest between notes. These two note length descriptors could each be described by four features:

- the raw length (seconds) for each note event;
- a categorical variable of note length with five classes, compared to an absolute time value of 0.5 seconds, equivalent to a single beat at 120 bpm (absolute duration or IOI class);
- a categorical variable of note length with five classes, compared to the local beat duration (relative duration or IOI class);
- the ratio between the raw note length and the local beat location ($\text{duration}_{\text{BeatProp}}$ or $\text{IOI}_{\text{BeatProp}}$).[1]

This research focused on the relative categorical class descriptors and the $\text{duration}_{\text{BeatProp}}$ and $\text{IOI}_{\text{BeatProp}}$ ratio values. The raw note lengths and the absolute categorical classes were not useful in the analysis of Green's improvisational style as they did not allow for meaningful comparison of note lengths between tempos. The five classes can be seen in Table 6.1, which also lists the ratios of some standard rhythmic values for each class.

Table 6.1: Duration and IOI classes and boundaries.
Based on table found in the Jazzomat Research Project website's 'Transformation' page (Jazzomat Research Project 2017).

Class Name	Class Border	Class Label	Note Values
Very Short	< 35%	-2	Semiquavers: 25%, Quaver Triplets: 33%
Short	35% – 70%	-1	Quavers: 50%
Medium	70% – 140%	0	Crotchet: 100%
Long	140% – 280%	1	Minim: 200%
Very Long	> 280%	2	Dotted Minim: 300%, Semibreve: 400%

Figure 6.1 shows where common musical rhythmic values – from semiquavers to semibreves as standard, dotted, and triplets – fit within *MeloSpy*'s relative note length classes, from very short to very long classes. The graph indicated there was substantially more distinction between shorter notes, for example the long class contained the same relative note lengths as very short, short, and medium classes combined. This was not a substantive issue, as the vast majority of notes played were classified as one of the shorter classes. Table 6.2, shows the proportion of notes – in both the corpus of Green's improvisations and the entire WJazzD – that had an absolute or relative note length classes of very short, short, or medium.

[1] These values were the same used to calculate the relative duration and IOI classes.

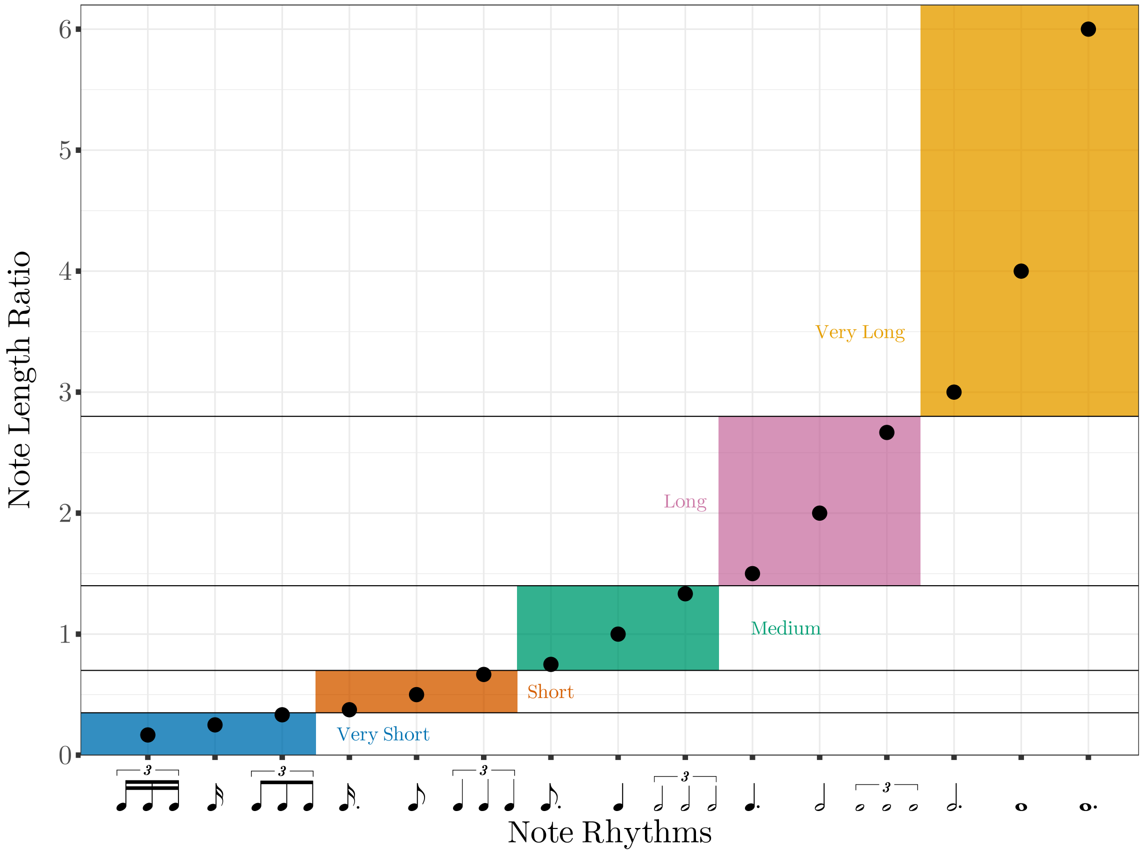

Figure 6.1: *MeloSpy*'s relative note length class distribution against the note length ratios and standard symbolic notation rhythms.

Table 6.2: Proportion of notes with a duration or IOI class of very short, short, or medium in Green's corpus vs. the WJazzD.

	Green	WJazzD
Duration		
Relative	98.28%	97.88%
Absolute	99.11%	98.89%
Inter-Onset Interval		
Relative	92.84%	92.85%
Absolute	95.47%	95.21%

6.1.1 Relative Note Lengths

The relative note length classes and the note length ratios described the same data, but the first was a categorical variable while the second was continuous.[2] In addition to the relative classes listed previously, condensed versions were also frequently used, called fuzzy duration or fuzzy IOI classes (Jazzomat Research Project 2017). The fuzzy classes condensed the five levels down to three – short, medium, and long – combining the two outer levels (very short, short, and long, very long) into single levels. The new classes had ratios of, Short: < 0.7 beats, Medium: 0.7–1.4 beats, and Long > 1.4 beats. As with other fuzzy features, these allowed for the broad intent of the note to be analysed, while more detailed analysis of the note lengths could be accomplished using the ratio features. The distribution of Green's relative classes and their fuzzy versions can be seen in Table 6.3.

Table 6.3: Distribution of relative and fuzzy duration and IOI classes in Green's corpus.

	Very Short	Short	Medium	Long	Very Long
Relative Duration					
Count	7309	11451	1366	338	14
Percent	35.69%	55.92%	6.67%	1.65%	0.07%
Fuzzy Duration					
Count		18760	1366	352	
Percent		91.61%	6.67%	1.72%	
Relative IOI					
Count	4156	10899	3957	1195	231
Percent	20.33%	53.33%	19.36%	5.85%	1.13%
Fuzzy IOI					
Count		15055	3957	1426	
Percent		73.66%	19.36%	6.98%	

This data showed that for all features the majority of notes Green played fell within the short class. The consolidation of notes within the short class became more extreme in the fuzzy features. Less than 10% of note durations had a fuzzy duration

[2] One extreme outlier, with a $\text{duration}_{\text{BeatProp}}$ of 22.57 and $\text{IOI}_{\text{BeatProp}}$ of 22.58, was excluded from the analyses into relative note length. This was a tremolo note from Green's improvisation over *Blues In Maude's Flat* (Green 1961b) that began in bar 138 and was transcribed as a single long note, as discussed in the Transcription section of Chapter 3.

of medium or long. In comparison, nearly 20% of fuzzy IOIs had a medium length, while an additional 6.98% had a long fuzzy IOI. These differences were due to how the note lengths were calculated, with this exacerbated by longer gaps between note onsets, such as at the end of phrases, not being fully occupied by a played note.

Figures 6.2 and 6.3 show the distributions of the $\text{duration}_{\text{BeatProp}}$ and $\text{IOI}_{\text{BeatProp}}$ respectively, with the colours indicating the fuzzy classes. Due to the small number of notes with a ratio greater than two, they were included as an inset in each figure.

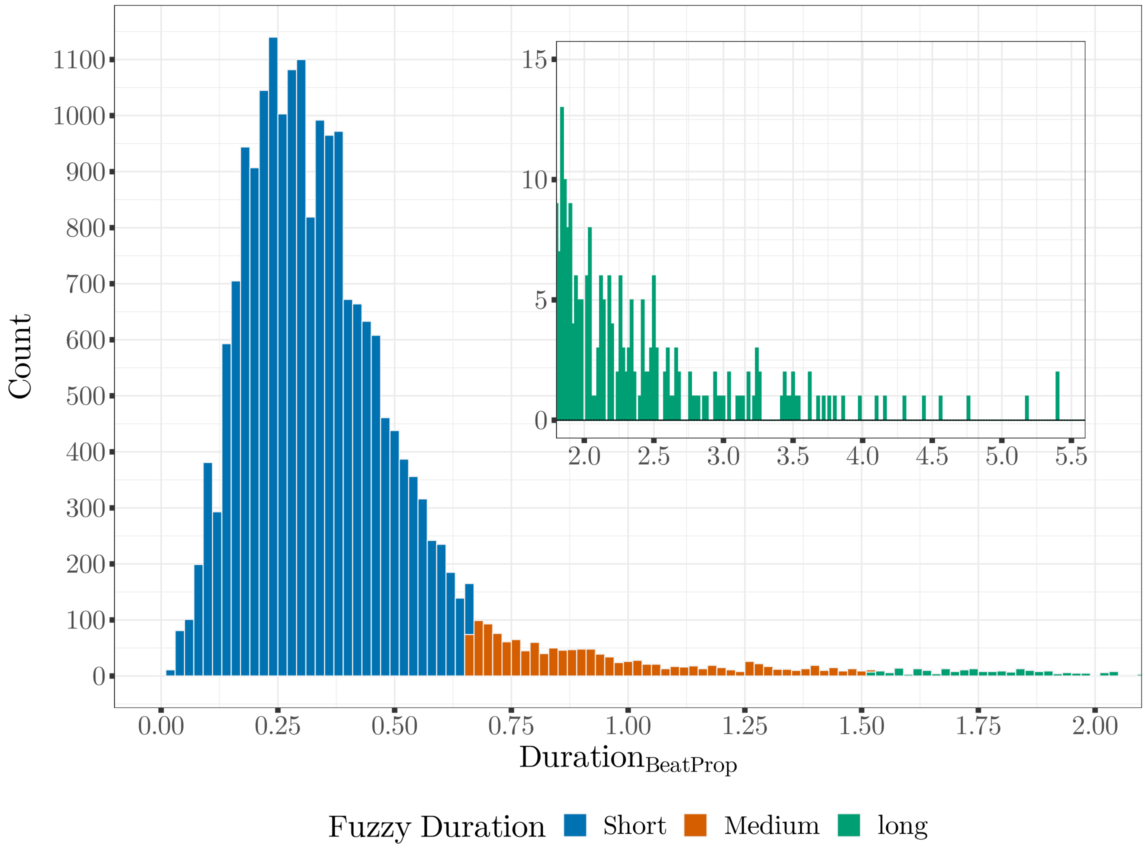

Figure 6.2: Distribution of $\text{duration}_{\text{BeatProp}}$ in Green's improvisations, the inset shows $\text{duration}_{\text{BeatProp}}$ greater than two.

The vast majority of Green's notes had a short $\text{duration}_{\text{BeatProp}}$, with the data in Figure 6.2 indicating that the majority (62.87%) of notes had a $\text{duration}_{\text{BeatProp}}$ between 0.2 and 0.5 beats. This suggested that most of Green's notes were equivalent to notes between a semiquaver and quaver. However, the distribution in Figure 6.3 suggested a slightly different interpretation. While there were no peaks in the $\text{duration}_{\text{BeatProp}}$ data, the $\text{IOI}_{\text{BeatProp}}$ distribution showed peaks in the data around 0.33, 0.5, and 1 beats. This indicated that although the majority Green's notes had a short fuzzy IOI, instead of being evenly distributed they tended to group around the equivalent of quaver triplet or quaver note lengths. The data also indicated that Green played many notes that were equivalent to a crotchet. Although these results were expected – the majority of notes in improvisations tend

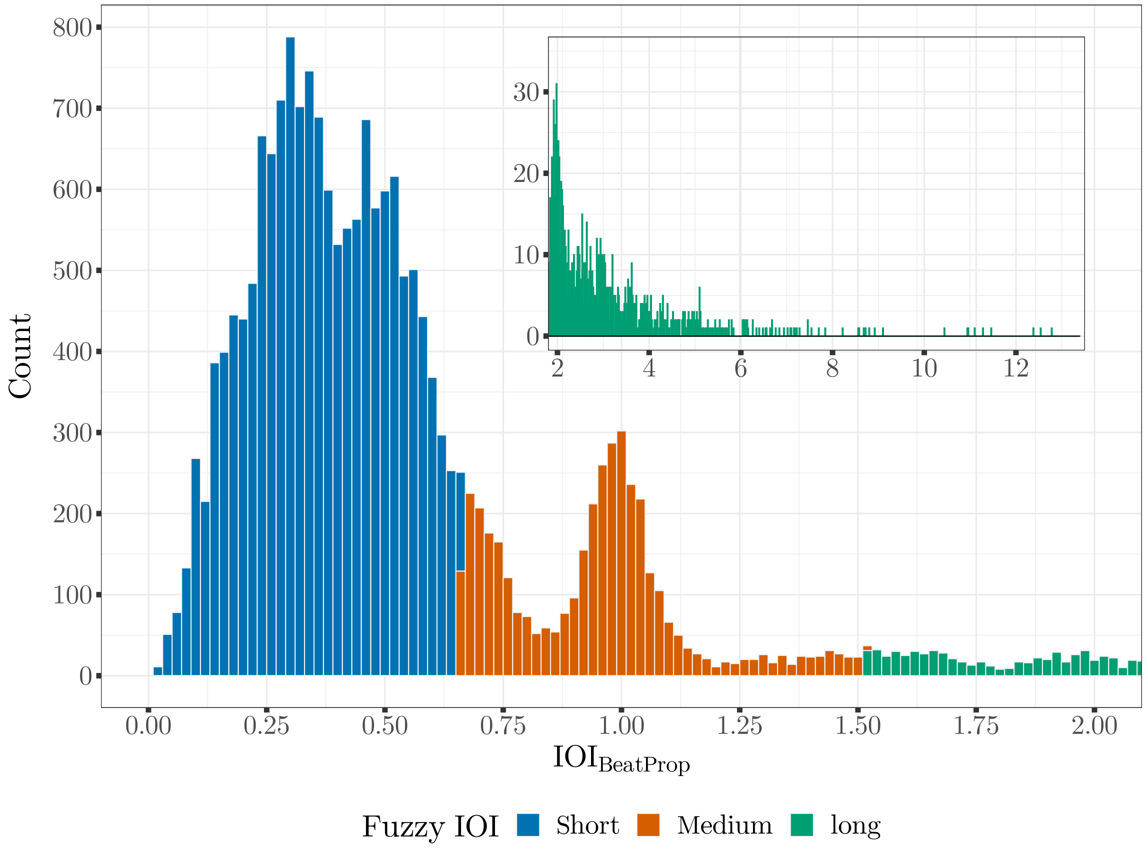

Figure 6.3: Distribution of $\text{IOI}_{\text{BeatProp}}$ in Green's improvisations, the inset shows $\text{IOI}_{\text{BeatProp}}$ greater than two.

to have rhythms between a semiquaver and crotchet – Green's data suggested that he rarely held medium or long IOI notes for their entire length. Additionally, the data suggested that Green did not shorten his notes in a systematic manner. The differences between Green's $\text{duration}_{\text{BeatProp}}$ and $\text{IOI}_{\text{BeatProp}}$ had enough variability that the peaks at specific note lengths in the $\text{IOI}_{\text{BeatProp}}$ data were not present in the $\text{duration}_{\text{BeatProp}}$ data.[3] The following sub-sections focused on the analysis Green's note lengths in comparison to the tempo, metrical weight, and $\text{CPC}_{\text{Weight}}$.

Note Length vs. Tempo

As the relative note lengths accounted for the tempo, this analysis investigated how the distribution of the relative note lengths changed with the tempo. The hypothesis was that as the tempo increased there would be fewer short fuzzy notes. The full set of tempo classes was used for this analysis, instead of the binary feature, as it was necessary to observe the changes throughout the tempo range. However, the results of the medium slow and medium up tempo classes should be considered cautiously, as they contained the fewest data points.

[3]A full examination of the differences between the $\text{duration}_{\text{BeatProp}}$ and $\text{IOI}_{\text{BeatProp}}$ can be found in the Articulation section of Chapter 7.

Figure 6.4 shows the proportion of notes in each tempo class that were classified as short, medium, or long fuzzy duration and fuzzy IOI. These graphs showed a trend of short notes being less frequent at higher tempos while medium and long notes were played more often. A χ^2-test was run for each feature, with the tests finding a significant difference in the proportion of fuzzy durations ($\chi^2(8) = 68.13,\ p =< .001,\ V = .04$) and fuzzy IOIs ($\chi^2(8) = 276.29,\ p =< .001,\ V = .08$) across the tempo classes, both with a small effect size. The graphs and these statistics showed the difference in note proportions was larger for the fuzzy IOI.

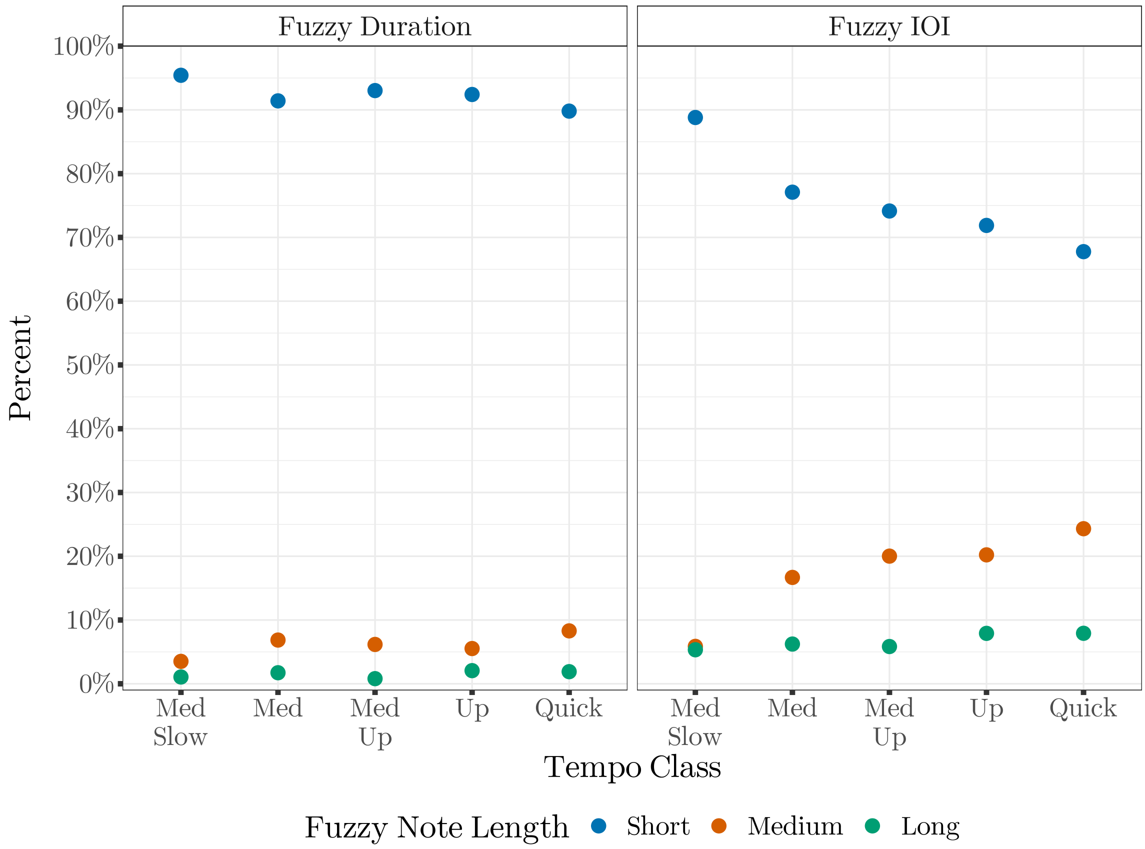

Figure 6.4: Proportion of notes that were labelled as short, medium, or long for each tempo class, for fuzzy duration and fuzzy IOI classes.

For the fuzzy duration, there was only a slight difference in the proportion of classes across the tempo range, with short notes played slightly less frequently at higher tempos. This data indicated that Green's fuzzy duration distribution was not heavily dependent on the tempo class. In comparison, the data showed that Green's notes as fuzzy IOI did change substantially as the tempo increased. Even excluding the Medium Slow data, which had the least data, a trend was observed with the proportion of short IOI notes decreasing from 77.09% at Medium tempos to 67.77% at Quick tempos. As with the fuzzy duration, medium IOI notes increased the most across the tempo classes, with a difference of 7.64PP between Medium and Quick tempos. There was also a very small increase (1.69PP) in the proportion of long IOI notes from the Medium to Quick tempo classes. The difference between the trends of the fuzzy duration and fuzzy IOI suggested that although the notes' relative duration remained fairly constant, there were larger gaps between notes at higher tempos. These results supported the hypothesis that Green played fewer short notes as the tempo increased, with this most apparent in the fuzzy IOI of the notes.

Note Length vs. Metrical Weight

The analysis into the effect of the metrical weight on Green's use of note lengths focused on how their distribution changed depending on their metrical weight. The hypothesis was that longer notes were more likely to be played on the beat, specifically on metrically strong beats.

Figure 6.5 shows the distribution of Green's metrical weights for both the fuzzy duration and fuzzy IOI. This data indicated that there were large differences in the distribution of fuzzy note lengths between notes played on the beat and those played off-beat. A pair of χ^2-tests were run comparing both the fuzzy duration and fuzzy IOI distribution to their respective metrical weight distributions. A significant relationship between both features and the metrical weight were observed (fuzzy duration: $\chi^2(4) = 486.83,\ p =< .001,\ V = .11$; fuzzy IOI: $\chi^2(4) = 1103.31$, $p =< .001,\ V = .16$), with small effect sizes for both features. The observed effect was slightly larger with the fuzzy IOI distribution.[4]

The plot of fuzzy durations showed a decrease in notes with a medium duration across the metrical weights (Strong: 13.20%; Weak: 8.92%; Off: 3.97%). This was matched almost directly with an increase in notes with a short duration (Strong: 84.50%; Weak: 88.85%; Off: 94.65%), with the proportion of long notes approximately the same for each metrical weight. For the fuzzy IOI classes, the difference in on-beat and off-beat distributions was larger; however, there was only a small difference between the metrically strong and weak beats. For both on-beat metrical weights, around 60% of notes had a short fuzzy IOI with an additional 30% with a medium IOI. In comparison, 81.42% of off-beat notes had a short IOI, with the proportion of medium IOI notes dropping to 12.99%. Green played long IOI most frequently on strong beats (10.29%), and weak beats (8.20%). Green played long IOI notes off-beat at half the rate he played them on metrically strong beats (5.58%). These results supported the hypothesis that the metrical weight impacted the length of the note Green played, with longer notes more common on metrically strong beats.

[4] Subsequent post-hoc tests found significant differences for all pairwise comparisons of the fuzzy duration classes ($p < .001$). For the fuzzy IOI pairwise comparisons, all were found to be significant, with the strong vs. off-beat and weak vs. off-beat significant at $p < .001$, and the strong vs weak comparison was significant at $p = .007$.

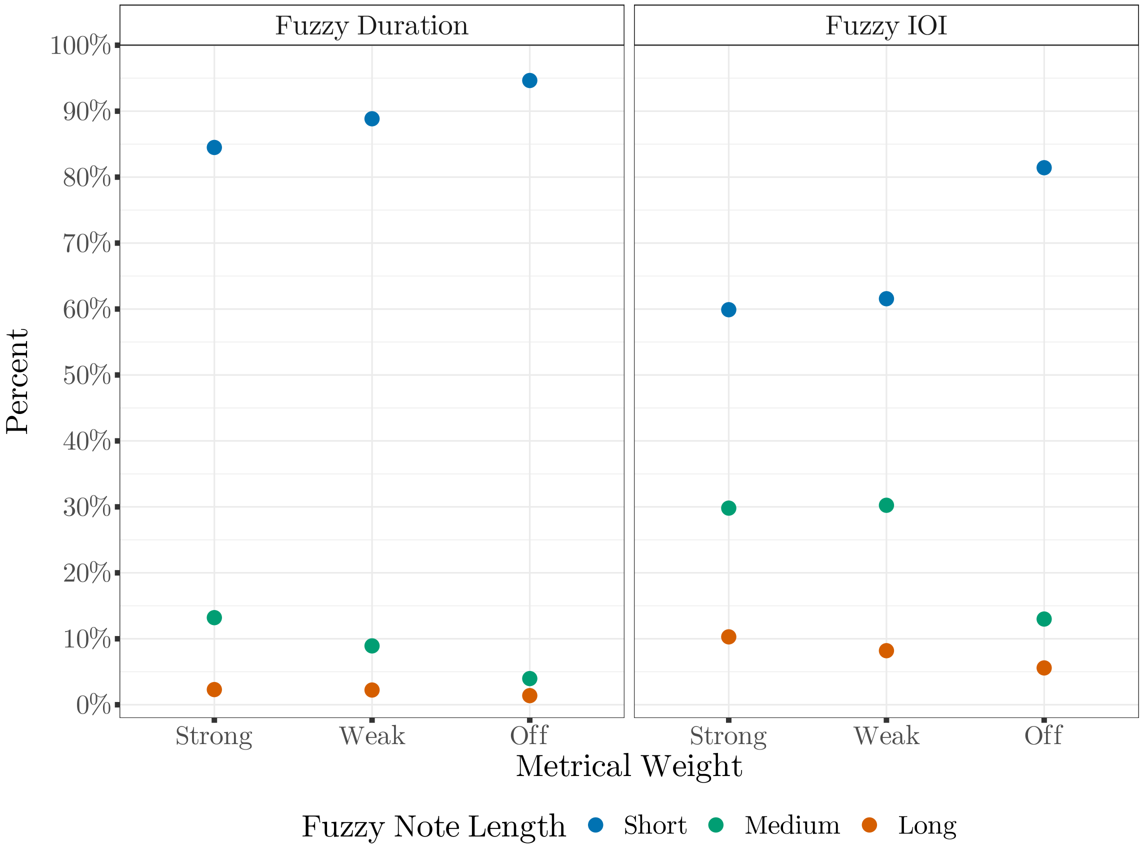

Figure 6.5: Proportion of fuzzy duration and IOI classes Green played in each metrical weight.

Note Length vs. $\mathrm{CPC}_{\mathrm{Weight}}$

The following analysis focused on the interaction between the $\mathrm{CPC}_{\mathrm{Weight}}$ of the notes Green played and their fuzzy note length. The hypothesis was that Green played HTs for longer, while NHTs more commonly had a short note length. Figure 6.6 shows the distribution of $\mathrm{CPC}_{\mathrm{Weight}}$ of each fuzzy note length, for both fuzzy durations and IOIs. These graphs showed that Green played scale tones at around the same proportion across all note lengths, with approximately 27% of all notes being scale tones. The largest observed differences were from arpeggio tones and NHTs. For the fuzzy duration, the proportion of long and medium notes was consistent for these tones. Green played a higher proportion of short notes as NHTs, with fewer being arpeggio tones. The fuzzy IOI also saw the largest change for short notes, with 24.68% of short notes being NHTs, while 47.84% were arpeggio tones.

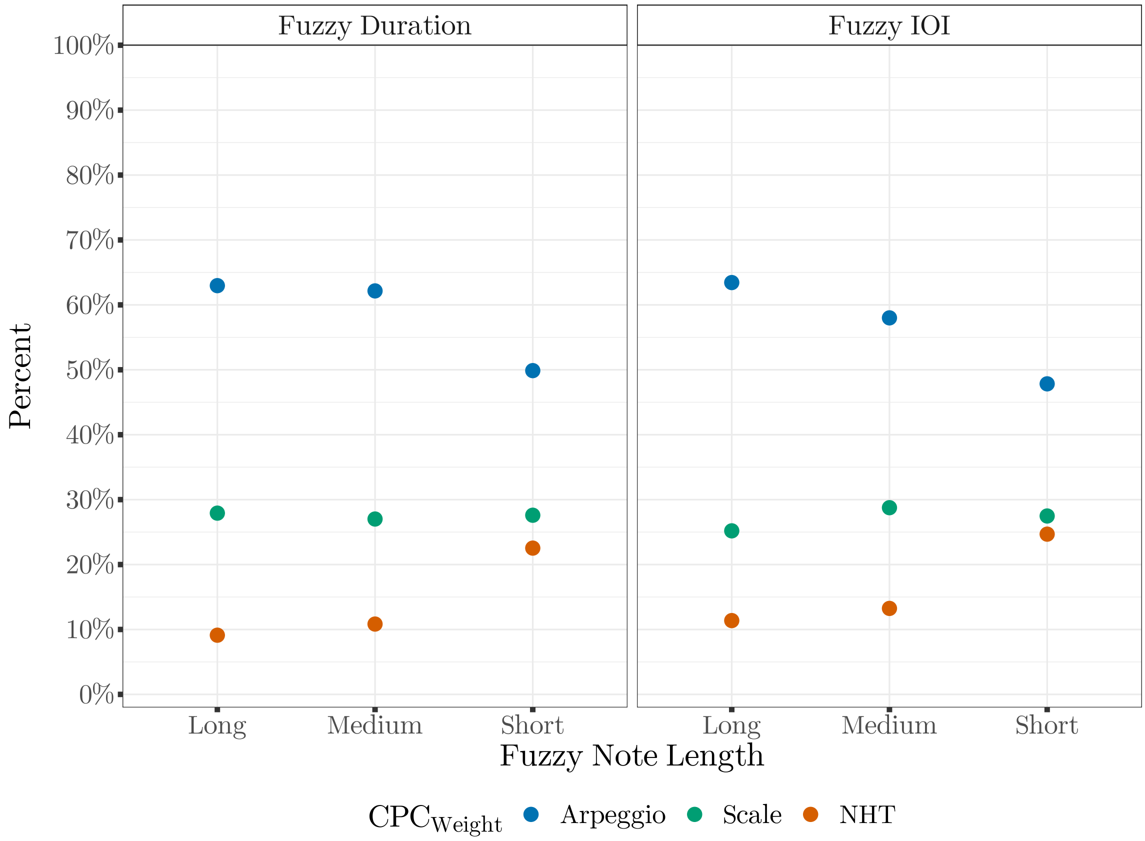

Figure 6.6: Distribution of $\mathrm{CPC}_{\mathrm{Weight}}$ for each fuzzy duration and IOI class in Green's corpus.

Two χ^2-tests were run to compare each fuzzy note length feature to the $\mathrm{CPC}_{\mathrm{Weight}}$, with both finding a significant interaction with a small effect size (fuzzy duration: $\chi^2(4) = 154.55,\ p =< .001,\ V = .06$; fuzzy IOI: $\chi^2(4) = 379.90,\ p =< .001,$ $V = .10$).[5] This analysis found that, although around half of short notes played by Green came from the arpeggio, NHTs were more frequently played with a short note length. In comparison, relatively few long notes were NHTs, with more than 85% of long notes being HTs. These results supported the hypothesis that there was a relationship between the $\mathrm{CPC}_{\mathrm{Weight}}$ of the notes Green played and their length.

6.1.2 Note Lengths Summary

This analysis of note lengths found evidence that most of the notes that Green played had a $\mathrm{duration}_{\mathrm{BeatProp}}$ equivalent to between a semiquaver and quaver. Green's $\mathrm{IOI}_{\mathrm{BeatProp}}$ tended to be slightly longer, between that of a quaver triplet or quaver. When compared to other features, Green's $\mathrm{duration}_{\mathrm{BeatProp}}$ was found to be

[5] For the fuzzy duration, post-hoc tests found significant differences between short notes and both long and medium notes ($p < .001$), while no significant difference was found between long and medium notes ($p = .638$). Post-hoc tests of the fuzzy IOI found significant differences between all pairwise comparisons (short vs. long and short vs. medium: $p < .001$; long vs. medium: $p = .002$).

not heavily dependent on the tempo of the improvisation. The $\mathrm{IOI}_{\mathrm{BeatProp}}$ of Green's notes were more dependent, with Green playing fewer notes with a short fuzzy IOI at higher tempos. This indicated that, although the note's duration remained fairly constant, there were longer gaps between notes at higher tempos. There was a general trend observed between the metrical weight and the length of a note. Green played more notes with a long or medium length on-beat compared to those played off-beat. Finally, an interaction was observed between the $\mathrm{CPC}_{\mathrm{Weight}}$ and the length of the note, with fewer NHTs being played for medium or long durations or IOI. Arpeggio tones were less likely to be short, while the proportion of notes that belonged to the scale was not found to be dependent on how long the note was played.

6.2 Beat Distribution

The following section investigated how Green's notes were distributed within the beat structures of his improvisations. Green's overall distribution was investigated first, through features including *MeloSpy*'s metrical circle map (MCM). This feature quantised every note to one of forty-eight bins in a bar. Additionally, a raw metre feature was created, which provided a non-quantised look at Green's beat distribution.

The main tool for visually assessing Green's beat distributions were circle maps, as seen in the Pitch Domain chapter. Two forms of the circle maps were generated; the first showed only the frequency at each beat and sub-beat location (unigram), while the second showed both the frequency and transitional properties (bigram).[6] Taken as a clock, the starting position is at 3 o'clock with the data moving anti-clockwise. The size and colour of the circles and lines represented their relative frequency. This can be seen in Figure 6.7, which shows the MCM beat distribution of notes played by Green over improvisations that were in $\frac{4}{4}$, with the main beats labelled as 0, 12, 24, and 36. The quavers were half-way between the crotchets, with the semiquavers again dividing the gap in two (e.g. at positions 0, 3, 6, 9 for the first beat). Notes with a quaver triplet rhythm have a nominal position between two MCM numbers (e.g. 3 and 4 or 8 and 9), and were therefore split between these two quantised values. From this figure, there was strong evidence that the majority of notes played by Green fell on either the beat or the quaver off-beat. The bigram lines indicated strong quaver and crotchet movement. Investigation into the rhythmic variety of Green's playing can be found in Section 6.5.1.

[6]Equivalent to 0^{th} order and 1^{st} order markov chains.

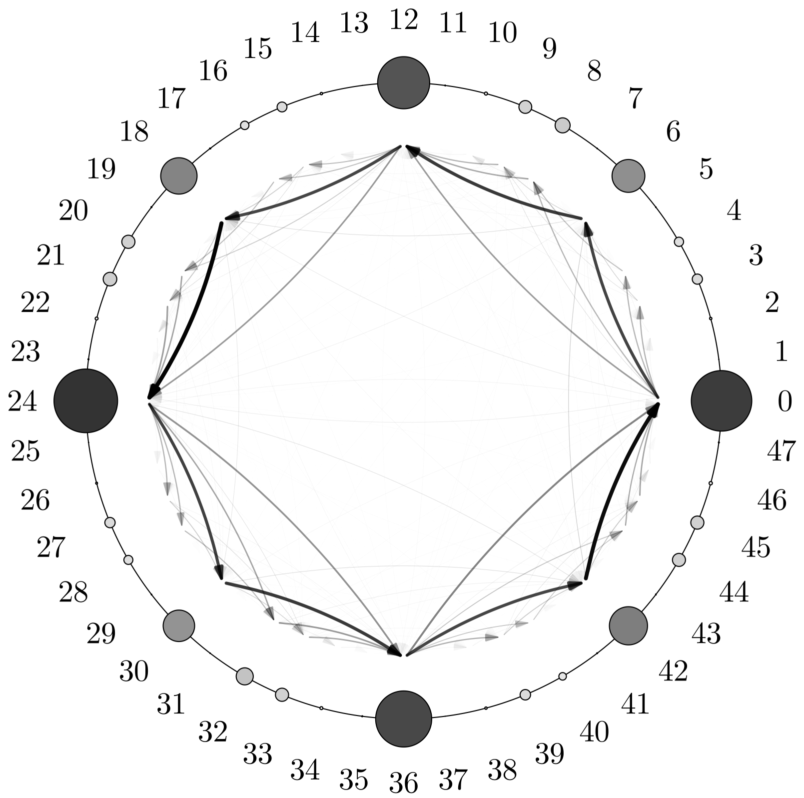

Figure 6.7: Metrical circle map for improvisations in $\frac{4}{4}$ in Green's corpus.

Figure 6.8 shows the same distribution of notes in $\frac{4}{4}$, but based on the raw data; the MCM labels were not changed for consistency.[7] Two main differences were observed in comparison to the previous figure. First, all the main beats notes were rotated anti-clockwise by around one MCM label, indicative of Green playing behind the beat. This will be investigated fully in Chapter 7, Micro Domain. Second, where previously the quaver beats were half-way between the down beats, here they were two-thirds along. This showed the swing pulse that underlies jazz and Green's improvisations. Due to how frequently Green's lines were comprised of swung quavers, the triplet pulse was largely obscured. There was only a small increase in density at the first triplet position, with the final note of the triplet entirely obscured by the swung quavers. This figure also highlighted the variability in Green's note placements, even for the most rigid down beat placements.

[7]Some quantisation was still necessary when plotting the data, but with many more bins. Only the unigram version is displayed due to plotting issues with generating the bigram data.

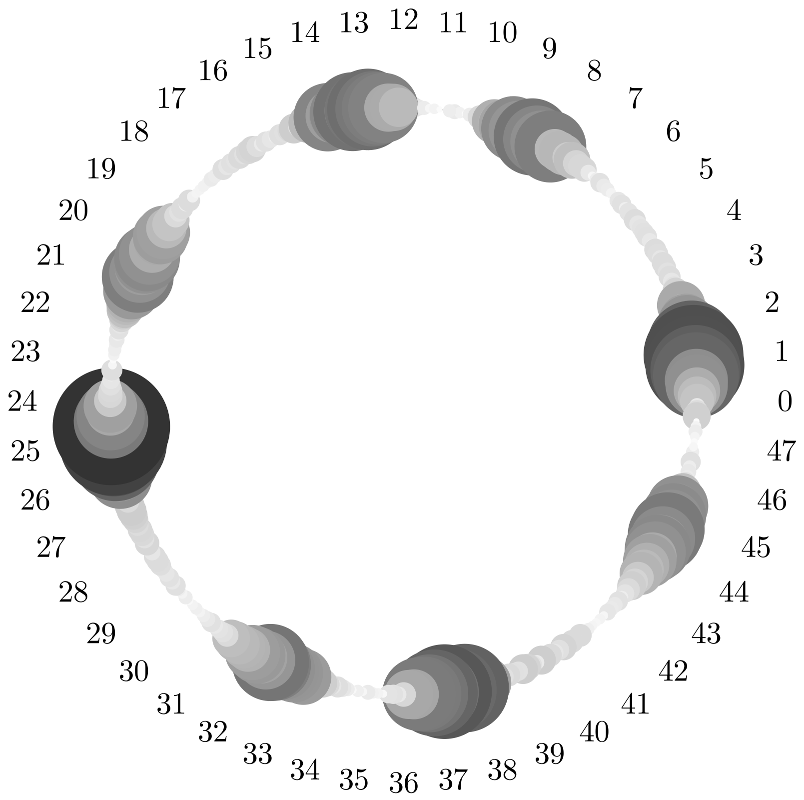

Figure 6.8: Beat distribution for improvisations in $\frac{4}{4}$ in Green's corpus displayed using raw metrical values.

The data for Green's $\frac{3}{4}$ improvisations were similarly plotted, with the MCM labels replaced with the nominal crotchet and quaver locations, in Figure 6.9. The same general trends can be seen in Green's playing over $\frac{3}{4}$, with most of the notes falling around the crotchet or quaver-equivalent positions. Due to the relatively few improvisations Green played over $\frac{3}{4}$ in the dataset, and the extra space due to there being one fewer beat, the triplet and semiquaver pulse was slightly more apparent.

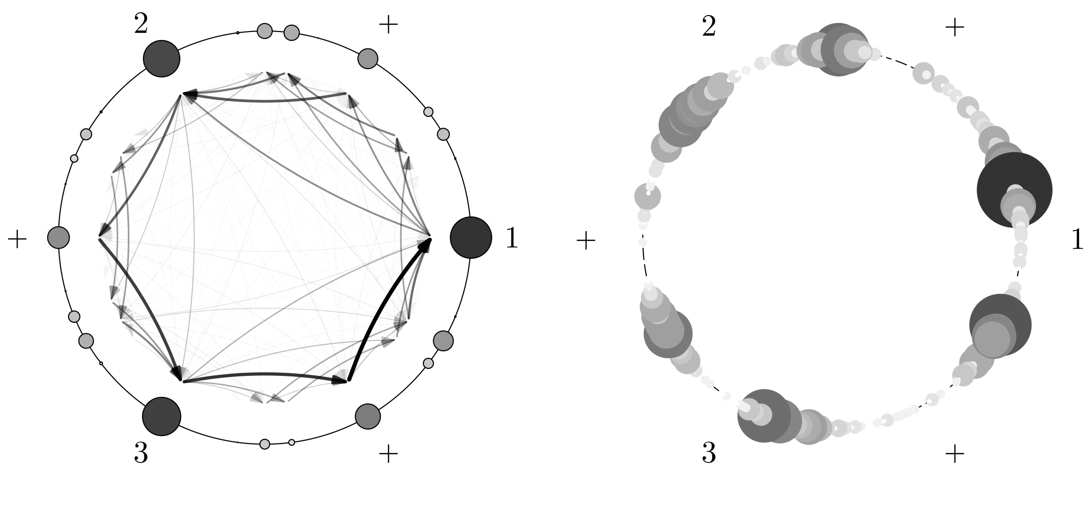

Figure 6.9: Beat distribution for improvisations in $\frac{3}{4}$ in Green's corpus. Left: Metrical circle map. Right: Raw metrical values.

The tempo was hypothesised to have a large influence on Green's distribution of notes with the beats of a bar, with slower tempos likely allowing for greater rhythmic flexibility. Due to the limited available data for Green outside of improvisations in $\frac{4}{4}$, this analysis focused on only Green's improvisations in $\frac{4}{4}$. Figure 6.10 shows the beat distribution of Green's notes for each tempo range, with the graphs on the left showing BPM $\leq$ 170 while the graphs on the right showed BPM $>$ 170. Both the MCM (top) and raw metrical values (bottom) were displayed for comparison. These graphs showed that when improvising at BPM $\leq$ 170 Green was more likely to place notes throughout the rhythmic sub-divisions between the beats. Additionally, a strong underlying quaver pulse was not observed within the MCM graph (top left). In comparison, the BPM $>$ 170 graphs showed Green playing more notes on the beat and on quaver off-beats, with fewer notes placed in other sub-beat positions.

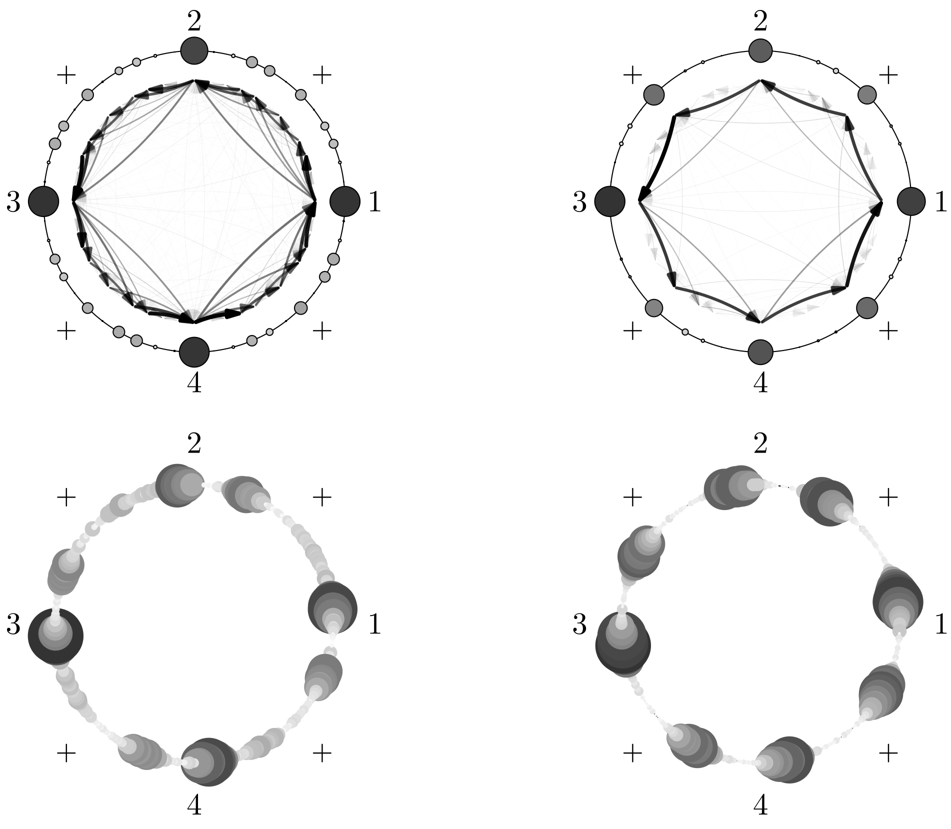

Figure 6.10: Metrical circle map and raw metrical values for notes played in each tempo range. Left: BPM $\leq$ 170. Right: BPM $>$ 170.

The beat distribution differences of Green's playing can be seen in two musical examples shown in Figure 6.11. The top example was from Green's improvisation over *Blues In Maude's Flat* (Green 1961b), which had a tempo of 119 BPM. The bottom example was from Green's improvisation over *Oleo* (Solo 1, Green 1962i), with a tempo of 252.6 BPM. The example from *Blues in Maude's Flat* showed Green using a wide variety of rhythms, from crotchets to demisemiquavers and quintuplets, resulting in the notes being placed in many sub-beat positions. In comparison, the example from *Oleo* showed Green improvising predominantly with crotchets and quavers.

Figure 6.11: Examples of differences in Green's beat distribution at both tempo ranges. a) Low tempo and highly varied beat distribution, *Blues In Maude's Flat* (1961), bars 54–59. b) High tempo with lower variety of beat distributions, *Oleo* (1962), bars 57–63.

6.2.1 Beat Distribution Summary

In summary, this analysis into the general beat distribution of Green's notes found that he frequently placed notes on the beat and on quaver equivalent off-beat positions. A difference in beat distribution was observed between tempo ranges when looking at improvisations in $\frac{4}{4}$. For improvisations with a BPM ≤ 170, Green's improvisations had greater distribution of notes within the beat structure of the bar, with greater variety of sub-beat placements. When improvising at BPM > 170, Green's notes tended to more frequently land on the beat or quaver off-beat. Slower tempos provided more flexibility to play more complex rhythms and varied note placements. To further analyse Green's approach to beat placement, his metrical weight distribution was investigated.

6.3 Metrical Weight

Metrical weight categorised a note based on its placement within the metrical structure of a bar and had three classes: metrically strong beat (beats 1 and 3 in $\frac{4}{4}$, beat 1 in $\frac{3}{4}$); metrically weak beat (beats 2 and 4 in $\frac{4}{4}$, beats 2 and 3 in $\frac{3}{4}$); and off-beat.[8] Figure 6.12 shows simulated distributions of metrical weights for triple and quadruple time signatures. The distributions show fully occupied bars with rhythm groupings from crotchets, with zero off-beats, to semiquavers, with three off-beats to every on-beat note.[9] This data showed the expected behaviour of the proportion of metrically strong and weak beats decreasing with the addition of more sub-divisions, as the number of down beats remained constant. However, there were very few bars in the corpus that were both fully occupied and occupied by the same division.

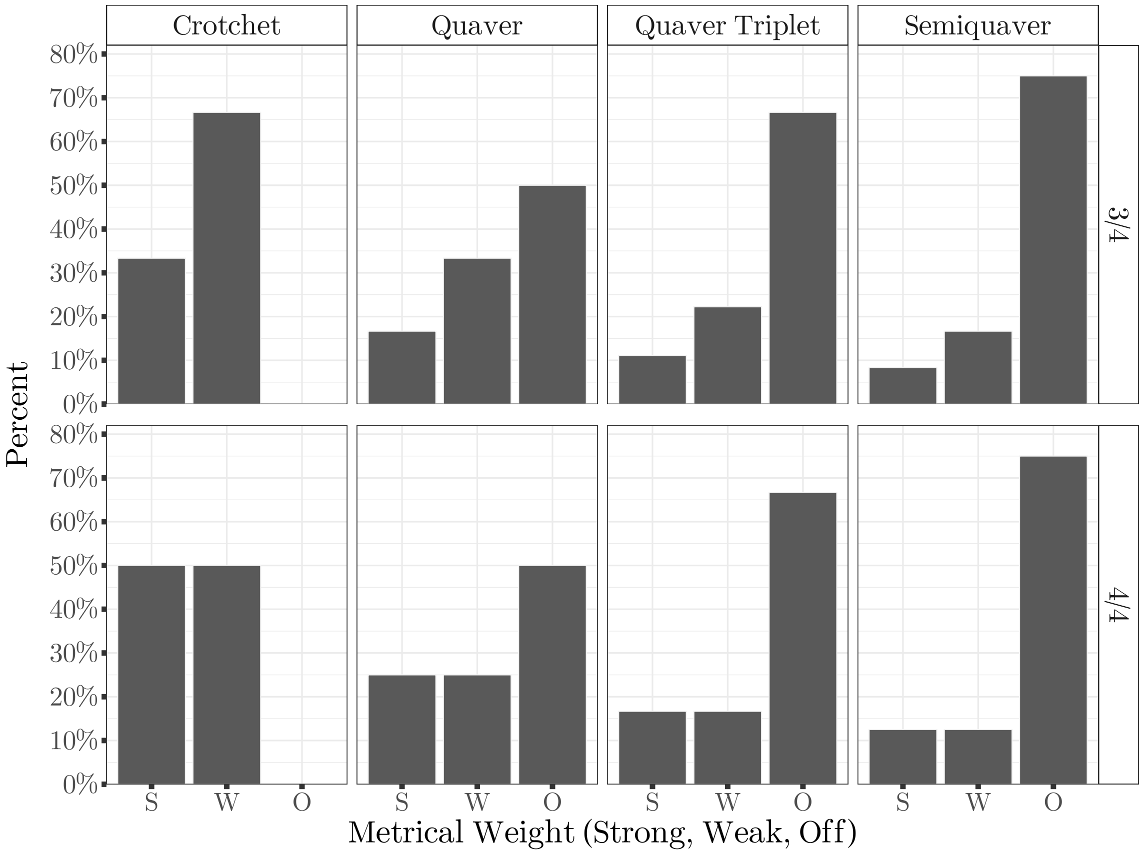

Figure 6.12: Distribution of simulated metrical weights against rhythmic values in $\frac{3}{4}$ and $\frac{4}{4}$.

This investigation into metrical weight distributions predominantly focused on how Green's distribution compared to these simulated fully-occupied distributions. This comparison also provided some insight into Green's use of rhythmic pulses. Figure

[8]Both quadruple time signatures, $\frac{4}{4}$ and $\frac{8}{4}$, were considered together as the metrical weight regarded beats 5 and 7 as metrically strong and 6 and 8 as metrically weak. This section used both quadruple time and $\frac{4}{4}$ to refer to this combined data.

[9]Crotchets were used here to represent any note that occupied one or more beats and played on a down beat, as the metrical weight did not treat a crotchet or dotted semibreve any different.

6.13 shows the distribution of metrical weight from Green's improvisations. Comparing Green's distribution to the simulated distributions, both the $\frac{3}{4}$ and $\frac{4}{4}$ data was most similar to that of the fully occupied quaver triplet bars. In $\frac{3}{4}$ 65.39% of Green's notes were played off-beat, with 22.28% played on a weak beat, and the remaining 12.33% played on a strong beat, closely matching the simulated triplet data. Comparatively, in $\frac{4}{4}$ 62.36% of Green's notes were played off-beat, 17.47% on a metrically weak beat, and 20.17% on strong beats, which also mostly aligned with the simulated quaver triplet data. Green's $\frac{4}{4}$ data had two substantial differences from the simulated data. These differences were observed in the strong and off-beat metrical weights, where Green played slightly fewer off-beat notes and more on metrically strong beats.

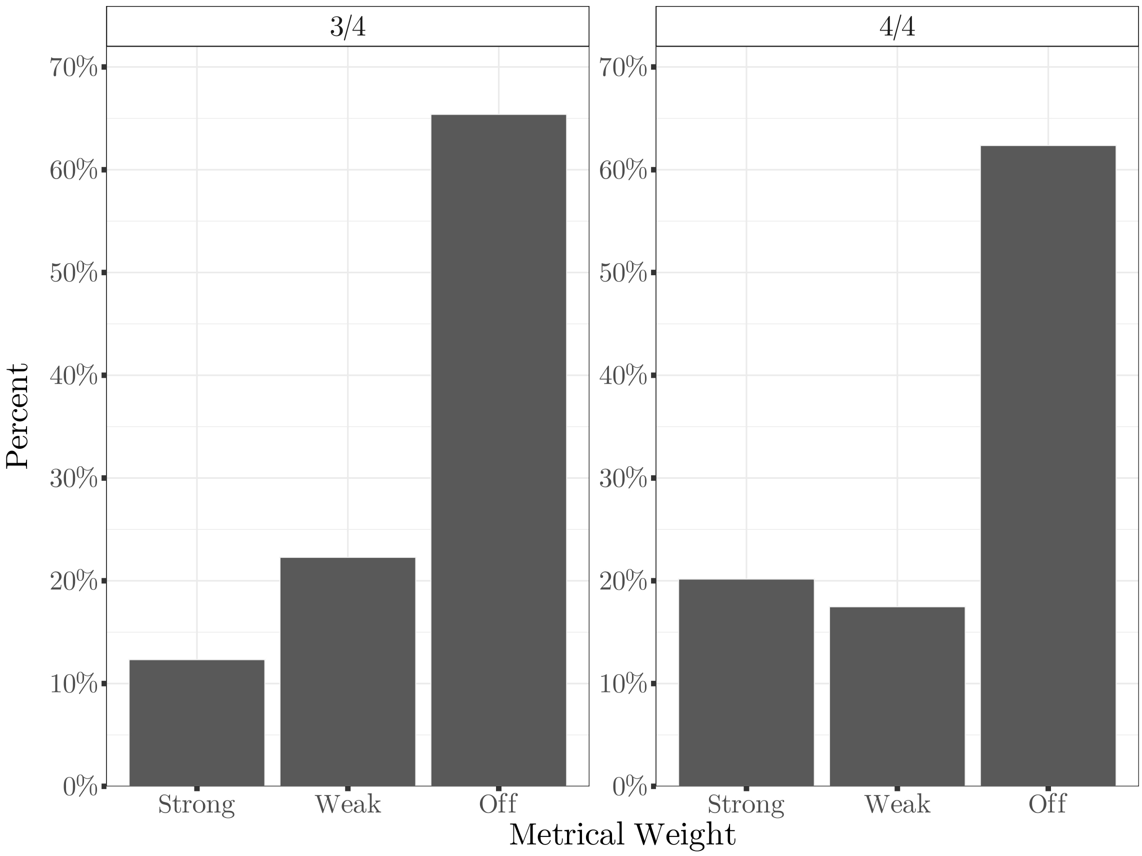

Figure 6.13: Distribution of metrical weights for improvisations in $\frac{3}{4}$ and $\frac{4}{4}$ in Green's corpus.

The differences between Green's actual data and the simulated data was due to variance in the length of notes and that many beats and bars were not fully occupied. These non-fully occupied beats would mainly decrease the frequency of off-beat notes, and consequently increase the proportion of on notes played on metrically strong and weak beats. The length of notes not only applied to sub-beat notes (e.g. dotted quavers), but also to notes that spanned multiple beats (e.g. minims). These longer notes were not well accounted for in the internal representation of metrical structure used by *MeloSpy*, as they considered notes only within the metrical framework based on their onset time.[10] These initial comparisons suggested that the majority of Green's improvisations were constructed with rhythmic building blocks of quavers, quaver triplets, semiquavers, with few long notes.

An example of how Green's metrical weight distribution translated to a musical example can be found in Figure 6.14. This shows a phrase from Green's improvisation over *I'll Remember April* (Green 1961k), which closely matched Green's distribution of metrical weights.[11] This phrase had 60.87% (14) of notes played off-beat, with 17.39% (4) and 21.74% (5) of notes played on the metrically weak and strong beats respectively. This indicated that per phrase, the true difference between the metrically weak and strong beats was only one or two notes. This then translated to an approximate 3PP difference between the two metrical weights weights at the corpus level.

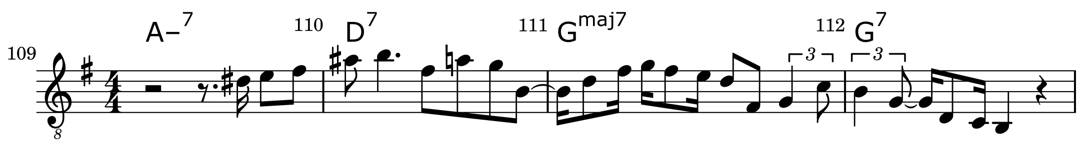

Figure 6.14: Example of a phase with a metrical weight distribution similar to Green's overall distribution in $\frac{4}{4}$, *I'll Remember April* (1961), phrase 33, bars 109–112.

As the metrical weight was a useful comparative variable for other features, it was analysed against features in other domains, and to avoid repetition, these were not included here. A feature that was hypothesised to have a substantial influence on the metrical weight distribution of Green's notes was the tempo of the improvisations. The distribution of Green's metrical weight for each of the binary tempo ranges is shown in Table 6.4. This data showed that Green played a higher proportion of notes off-beat at slower tempos, matching the findings of the previous Beat Distribution analysis. Additionally, at slower tempos Green's proportion of notes played on metrically strong and weak beats were nearly identical. The metrical weight distribution for BPM > 170 also closely matched that of the MCM graph in Figure 6.10. That figure indicated that approximately half the notes were played on-beat with the other half played off-beat, which this data supported. A significant difference was found between the metrical weights and the tempo range, with a small effect size ($\chi^2(2) = 364.47,\ p =< .001,\ V = .13$). These results supported the hypothesis that the tempo of a song had a significant influence on the metrical distribution of Green's notes.

[10]The metrical structure representation, which uses the FlexQ algorithm to determine sub-beat arrangements of division and tatums is discussed in the Rhythmic Variety sub-section (Section 6.5.1.

[11]Due to the way the musical examples were generated, data in the first and last bars that were not part of the phrase were omitted.

Table 6.4: Distribution of metrical weights for notes played at tempos ≤ 170 BPM and > 170 BPM in Green's corpus.

	Strong	Weak	Off
BPM ≤ 170			
Count	1641	1585	7128
Percent	15.85%	15.31%	68.84%
BPM > 170			
Count	2403	2046	5675
Percent	23.74%	20.21%	56.05%

6.3.1 Metrical Weight Summary

This analysis into the metrical weight of Green's notes found that the majority of his notes were played off-beat. While Green played a similar proportion of metrically weak and strong beats, notes on metrically strong beats were slightly more likely in $\frac{4}{4}$ or at higher tempos. The analysis also confirmed the visual observation of the beat distribution graphs, with the tempo range having a significant impact on Green's distribution of metrical weights. At higher tempos there was an approximately even distribution of on-beat and off-beat notes; however, when Green improvised at slower tempos, more than two-thirds of his notes were played off-beat.

6.4 Rests

Rests, which were not natively available within *MeloSpy*, were an important and useful feature for analysing rhythmic elements of Green's improvisational style, especially his use of space. After a note was played (at the time of the note's offset) one of three situations occurred:

1) the note was immediately followed by another note without any gap;[12]
2) the note was followed by a rest;
3) the note was followed by a micro-gap.

For the purposes of this research, rests were defined as the time between the offset of the current note and the onset of the following note. The function used to calculate

[12] In Green's corpus there were only 103 notes that were followed directly by another note, without any gaps in between.

rests also encapsulated micro-gaps between notes. These micro-gaps between notes will be investigated in Chapter 7, Micro Domain. For this study the separation between a rest and a micro-gap between notes was defined by the length of the rest as a proportion of the surrounding beat length ($\text{rest}_{\text{prop}}$); with values > 0.3 beats considered a rest, while those ≤ 0.3 beats considered a micro-gap. The limit was chosen to allow for quaver triplet rests (with a nominal $\text{rest}_{\text{prop}}$ of $0.\overline{33}$ beats) with a small buffer, with the acknowledgement that this excluded semiquaver-equivalent rests, and other shorter true rests. Due to the structure of the data, rest data was always attached to the note played directly before the rest.

Of all the notes in Green's improvisations 3583 (17.50%) were followed by a rest. Figure 6.15 shows the $\text{rest}_{\text{prop}}$ distribution in Green's data. Green had a mean $\text{rest}_{\text{prop}}$ equivalent to a crotchet rest ($\bar{x} = 1.06 \pm 1.18$ beats). However, the data was very skewed (skewness: 3.54) and had a long tail (kurtosis: 18.00). The median $\text{rest}_{\text{prop}}$ played by Green was closer to a quaver-equivalent rest (0.60 beats). The vast majority of Green's notes had a $\text{rest}_{\text{prop}} < 1$ beat (71.23%), with 90.57% having a $\text{rest}_{\text{prop}} < 2.5$ beats.

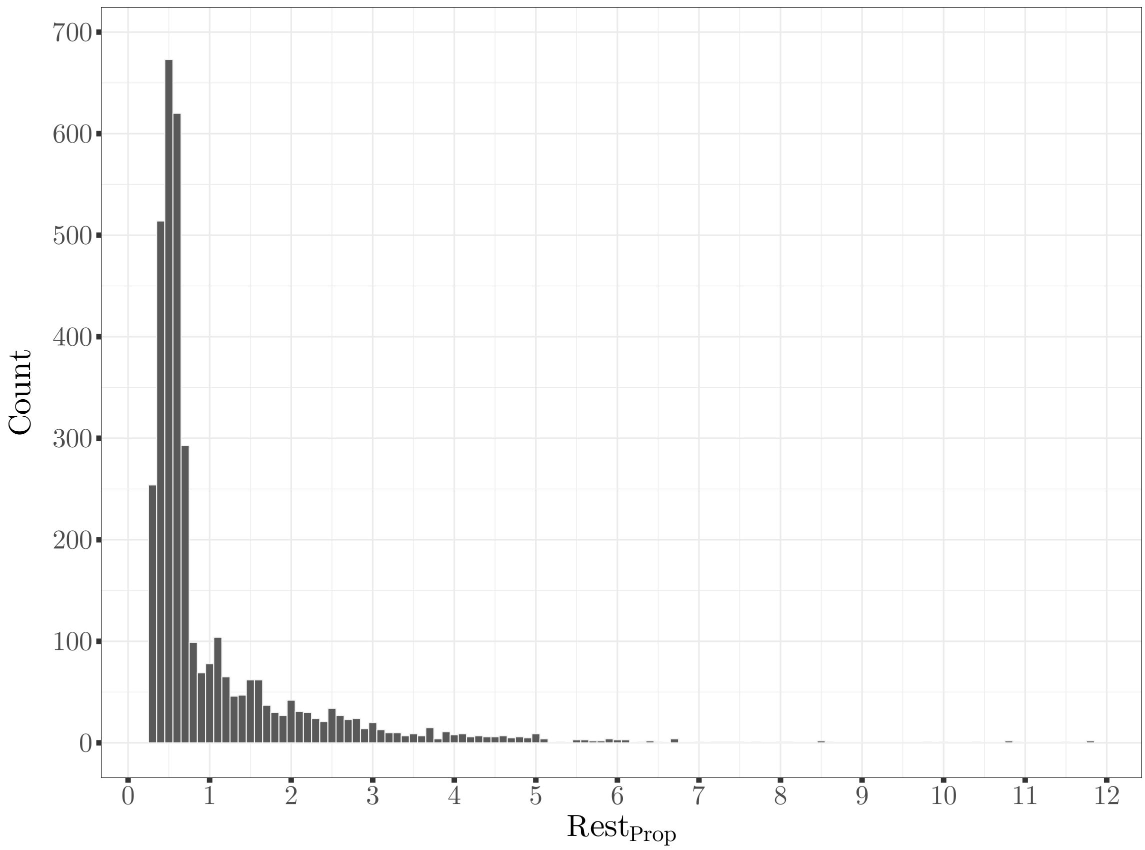

Figure 6.15: Distribution of $\text{rest}_{\text{prop}}$ in Green's corpus.

6.4.1 Rest Sequences

This section investigated sequences of notes that were or were not followed by a rest in Green's improvisations. Figure 6.16 shows the frequency of sequences for notes that were not followed by a rest (no rest sequence) and those that were (rest sequence). The foreground bar showed the frequency of that sequence length while the background bars showed the cumulative frequency. Due to the way rest data was stored, the no rest sequence lengths was one lower than the number of notes between rests. For example, a no rest sequence of length five would have five notes without rests plus the final note that was followed by a rest, resulting in six notes between rests.

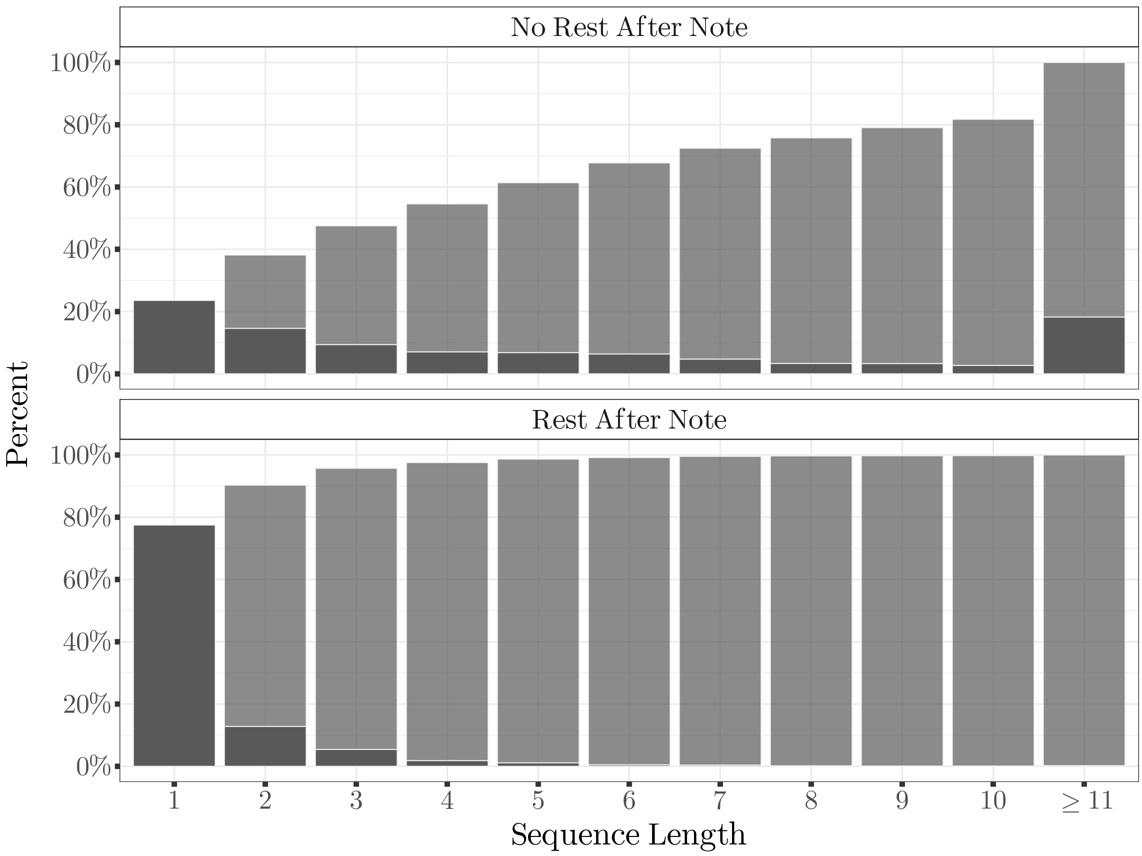

Figure 6.16: Frequency of sequence lengths for notes that were not followed by a rest vs. those that were. The solid bar in the foreground shows the frequency for each sequence length, the lighter bar in the background shows the cumulative frequency.

It was expected that in a flowing improvisation there would be few long rest sequences. Instead, the improvisations would tend to have long no rest sequences punctuated by the occasional rest. This situation was observed in Green's data, with Figure 6.16 showing substantial differences in the lengths between no rest and rest sequences. In Green's improvisations 77.50% of notes that had a rest were not followed by another note with a rest. In contrast, 76.42% of all no rest sequences were longer than one (three or more notes between rests), with 18.24% of sequences having a length of twelve or more notes between rests (≥ 11). On average, Green played 6.67 ± 9.11 notes in a row that were not followed by a rest before playing a note that was. In comparison, his mean sequence length for notes followed by a rest was much lower, at 1.43 ± 1.14 notes. The longest observed no rest sequence was 110 notes, while for rest sequences the longest observed was twenty notes.

To examine how Green's use of rests fit within their broader use in jazz, his sequence data was compared to those of Parker, Davis, Coltrane, Abercrombie, Martino, and Metheny.[13] The stats for all of these performers, including the number of sequences of each type, the mean, standard deviation, and minimum and maximum sequence lengths are shown in Table 6.5. As all notes belonged to one of the two types of sequences, each performer had approximately the same number of note sequences that were or were not followed by a rest, with slight deviations due to improvisations starting or ending with a specific sequence.[14]

Table 6.5: Count, mean, standard deviation, and range for sequences of notes that were or were not followed by a rest for Green, Coltrane, Davis, Parker, Metheny, Martino, and Abercrombie.

	N	Mean	SD	Min	Max
No Rest Sequences					
Grant Green	2528	6.67	9.11	1	110
John Coltrane	1832	9.48	10.17	1	136
Miles Davis	869	6.00	7.82	1	92
Charlie Parker	396	13.27	9.49	1	50
Pat Metheny	287	6.78	10.74	1	91
Pat Martino	39	20.72	18.42	1	77
John Abercrombie	25	10.80	12.74	1	64
Rests After Note Sequences					
Grant Green	2507	1.43	1.14	1	20
John Coltrane	1815	1.13	0.41	1	7
Miles Davis	856	1.36	0.85	1	10
Charlie Parker	379	1.06	0.31	1	5
Pat Metheny	286	1.48	0.99	1	7
Pat Martino	38	1.11	0.39	1	3
John Abercrombie	25	1.08	0.40	1	3

[13]The three guitarists were included to determine if there were instrumental differences in the use of rests.

[14]Due to the low sequence counts for especially Martino and Abercrombie, any conclusion draw from their data should be considered as only a possible indication.

The data in the table indicated that Green's no rest sequences tended to be on the shorter side, and were most similar to Davis and Metheny. In comparison, Coltrane and Abercrombie played slightly longer continuous sequences. Parker played sequences that were on average twice as long as Green's, while Martino's sequences were 50% longer again. The rest sequence lengths showed a smaller degree of variation, with Green – along with Davis and Metheny – playing slightly longer sequences on average. Green also played the longest sequence of notes followed by a rest of these performers.

To examine how Green's use of rest and no rest sequences differed from the other performers a set of pairwise t-tests were run, comparing Green's mean sequence length with each of the other improvisers. In no rest sequences, no significant differences were observed between Green and Metheny ($t(334.41) = -0.17,$ $p = 0.864;\ d = -0.02$) or Abercrombie ($t(24.24) = -1.62,\ p = 0.119;\ d = -0.66$). Significant differences, with small effect sizes, were found between Green and Davis ($t(1737.81) = 2.09,\ p = 0.037;\ d = 0.10$) and Coltrane ($t(3678.85) = -9.40,$ $p < .001;\ d = -0.31$), with large effect sizes found between Green and Parker ($t(515.57) = -12.93,\ p < .001;\ d = -1.14$) and Martino ($t(38.29) = -4.75,$ $p < .001;\ d = -1.54$). These results indicated that Green's no rest sequences were only slightly different to those of Davis and Coltrane's, while Parker and Martino played substantially longer no rest sequences on average.

For rest sequences, no significant difference was found between Green and Davis ($t(1979.69) = 1.91,\ p = 0.056;\ d = 0.09$) or Metheny ($t(376.84) = -0.85,\ p = 0.397;$ $d = -0.09$). Significant differences were observed between Green and the other performers, with small to medium effect sizes found between Green and Coltrane ($t(3344.78) = 12.20,\ p < .001;\ d = 0.42$) and Parker ($t(2136.73) = 13.33,\ p < .001;$ $d = 0.58$), with large effect sizes observed with Martino ($t(47.30) = 4.84,\ p < .001;$ $d = 1.41$) and Abercrombie ($t(28.05) = 4.20,\ p < .001;\ d = 1.59$). Although there were sample size issues with Martino and Abercrombie, the large effect sizes increased the confidence that the observed differences were true differences in improvisational style. These results indicated that on average Green's rest sequences tended to be on the longer side, though in terms of actual notes only slightly longer.

Overall Green's rest and no rest sequences notes were most similar to Davis and Metheny's. Green had amongst the shortest no rest sequences, with a note followed by a rest played on average after six or seven notes. Green also had the second highest mean rest sequence length. These results did not indicate any particular instrumentation difference related to rest and no rest sequences. Based on these results, the use of rests in Green's improvisational style could be described as short bursts of non-rest notes broken up by one or two notes that were followed by a rest.

Additionally, the results from these performers indicated that those who played longer no rest sequences also tended to play shorter rest sequences.

6.4.2 Rests vs. Other Features

There were many factors in an improvisation that may have influenced Green's use of rests, including: their phrase position[15]; their beat weight[16]; whether the note prior to the rest was played on or off the beat; the tempo range; the chord type; the metrical density, as notes per bar; the interval size between the notes surrounding the rest; and the $\mathrm{CPC}_{\mathrm{Weight}}$ of the note prior to a rest. There were two facets of Green's use of rests that were analysed against these features, the frequency of rests and the length of the rests.

Rest Frequency vs. Features

To investigate the first facet, a series of statistical tests were run to compare the frequency of rests against the listed features. The results of these tests can be found in Table 6.6. These results showed that Green's frequency of rests had a significant relationship with all of the listed features except the beat location and chord type over which the note prior to the rest was played. Of the significant features, the greatest effect sizes were observed in the phrase position, fuzzy intervals, and metrical density.

Table 6.6: Results of statistical tests analysing the frequency of rests under different situations in Green's improvisations. The columns show the statistic score, the degrees of freedom, the p-value, and effect sizes.

χ^2-test	χ^2	d.f	p-value	V
Phrase Position	5619.93	2	$< .001$	.52
Beat Location	4.58	2	0.10	.01
On or Off Beat	259.93	1	$< .001$	.11
Tempo Range	211.60	1	$< .001$	.10
Chord Type	0.77	2	0.68	.01
Fuzzy Intervals	1603.54	8	$< .001$	.29
$\mathrm{CPC}_{\mathrm{Weight}}$	196.06	2	$< .001$	.10
t-test	t	d.f	p-value	d
Metrical Density	15.86	1514.43	$< .001$	0.82

[15] Start, middle, or end of a phrase.
[16] First beat of the bar, last beat of the bar, or a middle beat of the bar.

Rest Frequency vs. Phrase Position

The distribution of Green's rests for each phrase position can be seen in Figure 6.17. This data showed that phrases nearly always (96.94%) ended with a rest. In Green's playing, only thirty-seven of 1251 phrases did not end with a rest. In comparison, only 14.80% of Green's phrases began with rests, with 12.37% of notes within a phrase followed a rest.[17]

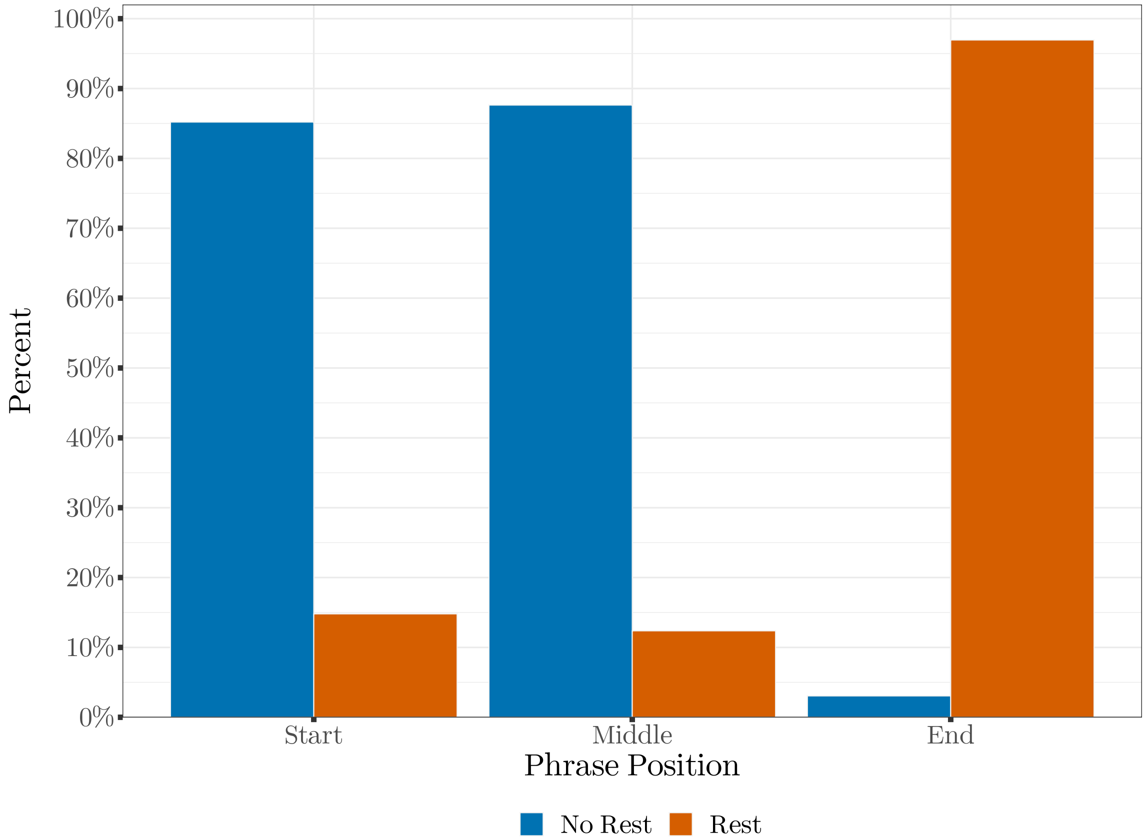

Figure 6.17: Distribution of rest frequency for each phrase position in Green's corpus.

As the vast majority (87.79%) of all notes were not at the start or end of a phrase, the majority of notes followed by rests (62.07%) also occurred within a phrase. However, as a proportion of all intra-phrase notes, those followed by rests were the smallest proportion. An example of Green using rests within a phrase can be seen in Figure 6.18. This phrase occurred towards the end of Green's improvisation over *Wives And Lovers* (Green 1964h), with the rests occurring between the C and G in bar 133 and between the G and C in bar 134.

[17] Post-hoc tests found significant differences between the pairwise groups, at $p < .001$ for starting and middle notes against ending notes and $p = 0.013$ between beginning and middle notes.

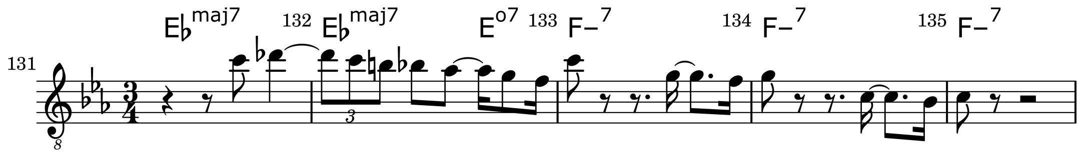

Figure 6.18: Example of Green using rests within a phrase, *Wives And Lovers* (1964), bars 131–135.

It was possible that some of rests that occurred within the phrases were not deliberate, but fumbles or misplays, where a held note was accidentally cut short. An example of this can be seen in Figure 6.19, which shows a phrase in Green's improvisation over *The Song Is You* (Green 1962l).[18] The swallowed note was the A in bar 67, beat 3, which had an $\mathrm{IOI}_{\mathrm{BeatProp}}$ of 1.10 beats but a $\mathrm{duration}_{\mathrm{BeatProp}}$ of 0.14 beats, resulting in a $\mathrm{rest}_{\mathrm{prop}}$ of 0.96 beats. As intent cannot be derived from these analyses, the results indicated an area where close analysis may be able to provide more nuanced interpretations.[19]

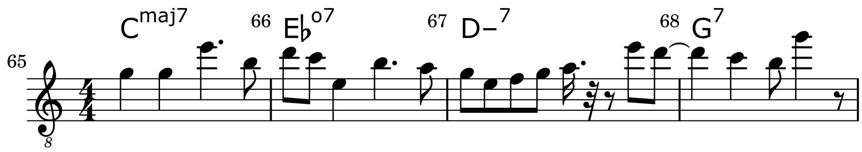

Figure 6.19: Example of a likely swallowed note resulting in a rest value being labelled, *The Song Is You* (1962), bars 65–68.

In summary, this analysis found that although most of Green's rests occurred within a phrase, they made up the lowest proportion of that phrase position. Rests were most common at the end of a phrase, with nearly all phrases ending with a rest.

[18]The process used to generate the figures attempted to minimise rests in the output, therefore the intermediary lilypond file had to be edited to show a more accurate approximation of what Green played.

[19]Whether any analysis can truly determine the intent of an improviser is in itself questionable. Analyses, corpus level or close, can determine what happened and can suggest possible intent, but any conclusive determination about what a performer intended to do should be considered with a high degree of scepticism.

Rest Frequency vs. Fuzzy Intervals

As previous analyses had found that large intervals mostly occurred between phrases, where rests were also frequent, this analysis focused on intra-phrase rests. Figure 6.20 shows, for each fuzzy interval class, the proportion of intervals that did or did not contain a rest between the notes.[20] This data showed that rests rarely occurred in Green's playing between steps or leaps, with rests being more frequent between repeated notes or intervals of a 4th or greater.[21]

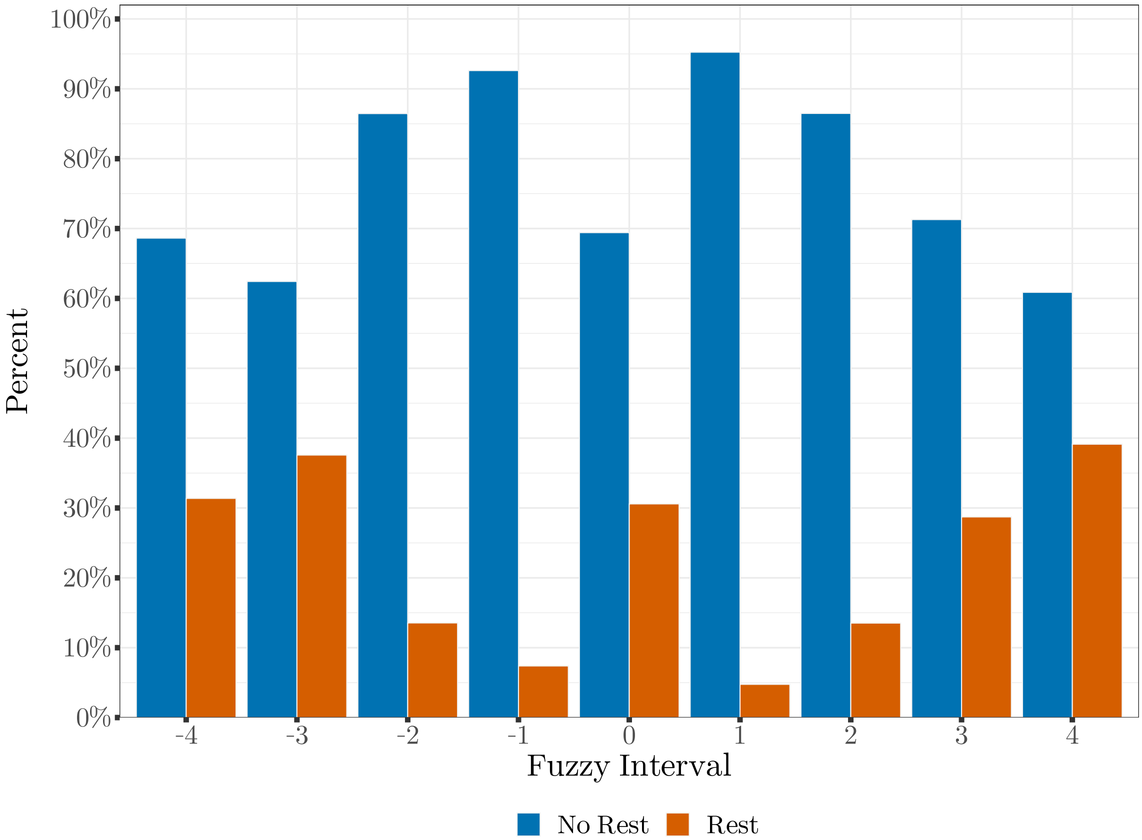

Figure 6.20: Proportion of intra-phrase fuzzy intervals that were or were not followed by a rest.

This indicated that the fuzzy interval classes could be broadly split between infrequent rest use (fuzzy intervals ± 1, 2) and semi-frequent rest use (fuzzy intervals ± 3, 4, and 0). On average, 8.59% of Green's smaller intervals had rests between them, in comparison to 32.34% of repeated notes or larger intervals. This suggested that Green was far less likely to play rests when moving step-wise or in thirds, with around a third of larger intervals containing a rest between the notes.

[20] The default fuzzy interval class, with nine levels from -4 to 4 was used in this analysis instead of the expanded version used previously to reduce statistical errors from small or empty classes.

[21] Post-hoc tests found significant differences between all pairwise comparisons of classes at $p <$ .001 except: -2 vs 2 ($p = 1.00$), -4 vs 0 ($p = .830$), -4 vs 3 ($p = .534$), -4 vs 4 ($p = .191$), -4 vs -3 ($p = .125$), -3 vs 4 ($p = .822$), 0 vs 3 ($p = .446$), 0 vs 4 ($p = .053$), which were not found to be significantly different; and 3 vs 4 ($p = .020$), -3 vs 0 ($p = .002$), which were found to be significantly different.

Rest Frequency vs. Metrical Density

The last feature to be investigated against Green's frequency of rests was the metrical density, as notes per bar. The correlation between the metrical density and frequency of rests was unsurprising, as the playing of a rest would reduce the available space in a bar to play more notes. Therefore, it was expected that bars that had more frequent or longer rests would have a lower metrical density. In Green's corpus there were 1035 bars that contained no rests at all, while 2284 contained at least one rest.[22]

Figure 6.21 shows the distribution of the number of rests that occurred in each bar. This figure showed that the plurality of bars contained one rest. The majority (58.14%) of Green's bars in which he played any rests contained only a single rest. On average, Green played 1.08 ± 0.97 rests per bar. For the bars without rests, the mean metrical density was 7.52 ± 3.60 notes per bar; for bars with at least one rests, the mean metrical density significantly lower at 5.55 ± 2.53 notes per bar.

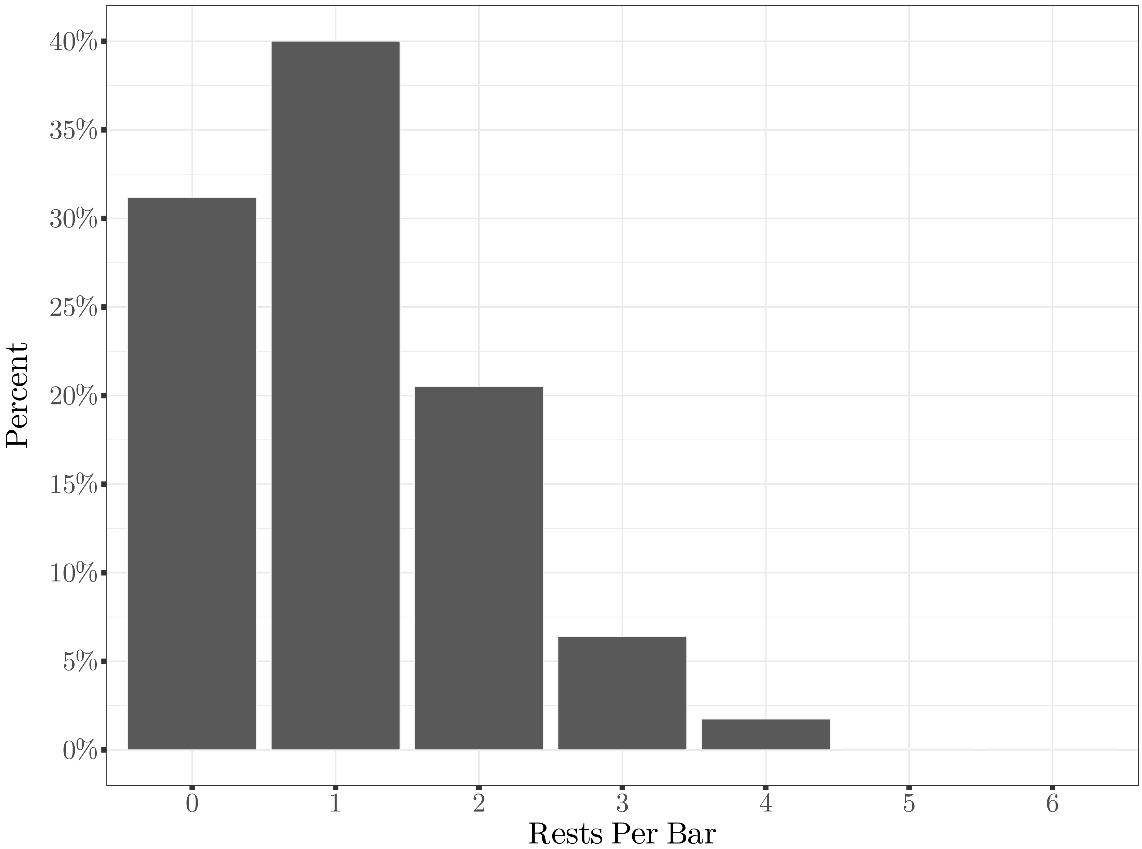

Figure 6.21: Distribution of the number of rests per bar in Green's corpus.

[22]This included the conversion of bars in $\frac{8}{4}$ to two bars of $\frac{4}{4}$.

Rest Frequency Summary

In summary, significant relationships were found between how frequently Green played rests and most of the features investigated. The largest effect sizes occurred between the phrase position of the note prior to a rest, the fuzzy interval between the notes surrounding a rest, and the metrical density. The subsequent analyses on these specific features found unsurprising results, with rests more common between phrases and larger intervals, while having a rest in a bar lowered the metrical density.

Rest Length vs. Features

To investigate the second facet regarding the $\text{rest}_{\text{prop}}$ of Green's rests a second set of statistical tests were run, analysing the $\text{rest}_{\text{prop}}$ against the listed features. The results of the analyses comparing the $\text{rest}_{\text{prop}}$ duration to the selected features can be found in Table 6.7.

Table 6.7: Results of statistical tests analysing the duration of rests (as $\text{rest}_{\text{prop}}$) played by Green in various situations. The columns show the statistic scores, degrees of freedom, p-values, and effect sizes.

ANOVA	F	d.f	p-value	η^2
Phrase Position	987.29	2, 3580	$< .001$	.36
Beat Location	16.24	2, 3580	$< .001$	.01
Chord Type	5.84	2, 3380	.003	.00
Fuzzy Intervals	2.78	8, 2400	.005	.01
$\text{CPC}_{\text{Weight}}$	5.31	2, 3380	.005	.00
t-test	t	d.f	p-value	d
On or Off Beat	4.21	3559.64	$< .001$	0.14
Tempo Range	-4.62	2380.87	$< .001$	-0.19
Correlation	t	d.f	p-value	r
Metrical Density	-13.99	3317	$< .001$	-.24

The results in Table 6.7 showed that each feature was found to have a statistically significant effect on Green's $\text{rest}_{\text{prop}}$ duration. Nearly all of the features had small effect sizes, with the only feature indicating a large effect size being the phrase position of the note preceding the rest.

Rest Length vs. Phrase Position

As the previous analysis found, nearly all of Green's phrases were separated by rests. Consequently, the rest after the final note of a phrase also described the time between phrases. Therefore, it was unsurprising to find a significant difference in the mean duration of rests played after notes at different phrase positions. Green's $\text{rest}_{\text{prop}}$ distribution for each phrase position can be seen in Figure 6.22.

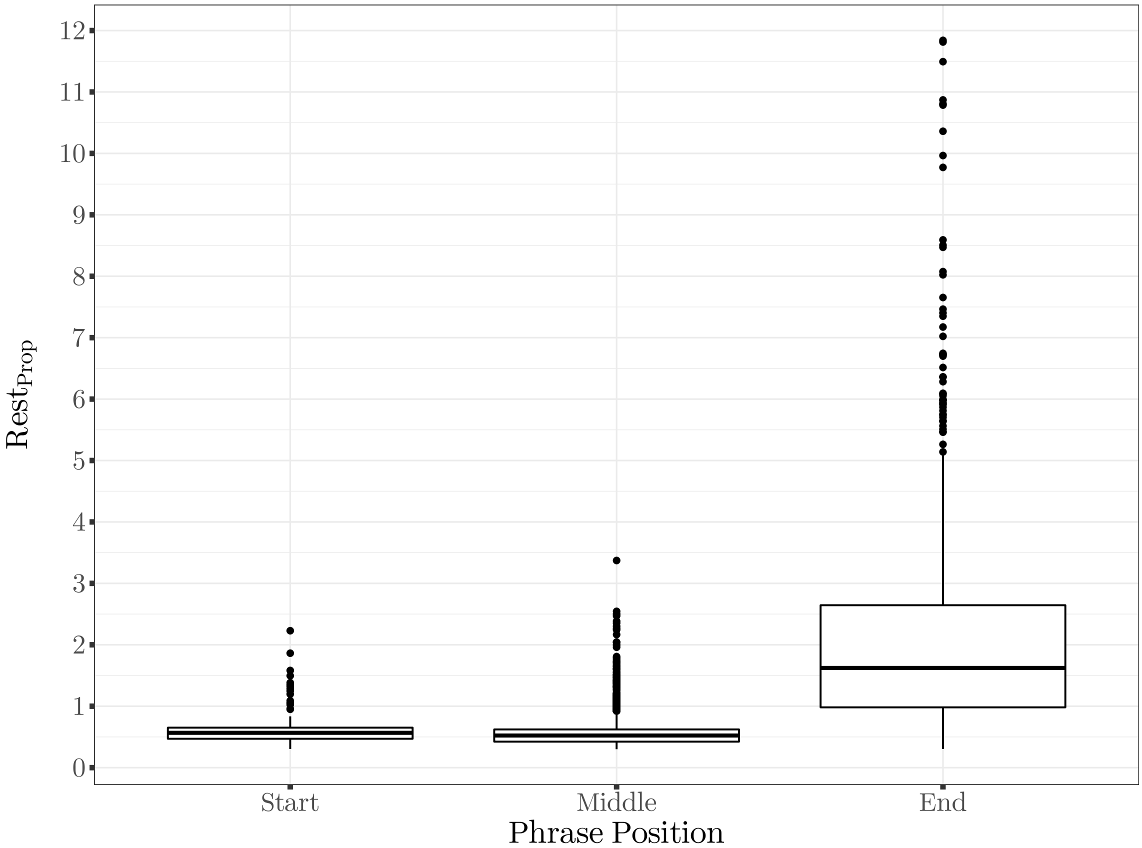

Figure 6.22: Distribution of $\text{rest}_{\text{prop}}$ for each phrase position in Green's corpus.

Post-hoc tests using Tukey's HSD procedure found a significant difference between Green's $\text{rest}_{\text{prop}}$ at the end of a phrase and those at the start or middle of a phrase (end vs. start and end vs. middle: $p < .001$). No significant difference was observed between Green's $\text{rest}_{\text{prop}}$ at the start of a phrase or any rests played throughout a phrase (start vs. middle: $p = .859$). Therefore, Green's use of $\text{rest}_{\text{prop}}$ within phrases can be condensed into two levels, inter-phrase and intra-phrase. On average, Green played approximately two beats of rest between each phrase ($\bar{x} = 2.07 \pm 1.61$ beats), while his average intra-phrase $\text{rest}_{\text{prop}}$ was around half a beat ($\bar{x} = 0.57 \pm 0.26$ beats). This indicated that the length of Green's inter-phrase rests were on average nearly four times longer than his intra-phrase rests.

Rest Length Summary

The results of the analysis into Green's $\text{rest}_{\text{prop}}$ suggested the only feature that had a large influence on the length of Green's rests was the phrase position of the prior notes. The analysis found that Green's inter-phrase rests tended to be around four times as long as his intra-phrase rests.

6.4.3 Rests Summary

These analyses focused on Green's use of rests within his improvisations. On average, Green's rests went for around a beat, with intra-phrase rests being around half a beat, while inter-phrase rests went for two beats. Green usually played around seven notes between rests, and rarely played two consecutive notes followed by a rest. When comparing Green's use of rests to other performers, he was most similar to Davis and Metheny. The frequency of Green's rests was most affected by the phrase position of the preceding note, the interval size of the notes surrounding the rest, and the number of notes played per bar. Nearly all of Green's rests occurred between phrases, with Green rarely playing rests at the beginning of a phrase or in the middle. Intra-phrase notes with a large interval between them were also more likely to have a rest between the notes, with Green rarely playing playing rests between notes that were less than a third apart. Bars that contained at least one rest also had a significantly lower metrical density than those without rests.

6.5 Examples

The two examples presented within the Rhythm Domain examined a similar concept, rhythmic values, from two points of view. The first looked at the variety of rhythms (e.g. quavers, triplets) within Green's improvisations, while the second looked at the metrical density of Green's playing. The rhythmic variety section investigated rhythms at the beat level and focused solely on the variety of rhythmic patterns that were found. The metrical density section investigated rhythms at the bar level, with comparisons to how other features changed with the metrical density.

6.5.1 Rhythmic Variety

Rhythmic variety referred to the sub-beat rhythms Green played in his improvisations. Scott, in discussing a 1963 improvisation of Green's over *Blue's For Lou*, said that "the predominant rhythmic unit [of the improvisation was] the triplet" (2006, 3). While *Blue's For Lou* was not one of the selected improvisations for transcription, the following analysis investigated the predominant rhythmic units of Green's improvisations. It analysed both their general usage and how they were influenced by the tempo and the tonality mode.

Within *MeloSpy* there were two features used to describe sub-beat divisions, division and tatum. Division described the number of subdivisions for each beat (e.g. two for quavers or three for quaver triplets). Tatum was the position within the subdivision in which the note was played (e.g. three notes in a division of three would have tatums of one, two, and three). The algorithm that calculated these features was called the FlexQ algorithm, a "new specially devised algorithm" (Pfleiderer, Frieler, et al. 2017, 24) developed by the Jazzomat Research Project.[23] The FlexQ algorithm was designed to divide each beat into a number of quantised sub-beats (divisions) such that they adequately represented the actual note onsets within the beat. It aimed to represent the data in the simplest form (smallest number of divisions) possible, while minimising the quantisation errors (the time difference from the true onset to the nominal onset from the algorithm). In short, the "algorithm ... [found] the optimal subdivision for all onsets between two beats" (Pfleiderer, Frieler, et al. 2017, 319) with the algorithm preferring "smaller subdivisions ... [and] binary and ternary subdivisions." (Pfleiderer, Frieler, et al. 2017, 320)

Due to the inherent structures of swung jazz, the FlexQ algorithm did occasionally divide beats into a greater number of divisions than one would use for writing out a transcription in symbolic notation. However, the more precise nature of these divisions made them useful for analyses. Unlike rhythms in standard notation, the division and tatum did not contain note length data, and therefore were best used to describe the placement of notes within beats, and the variety of rhythmic units played by Green.

Finally, an important concept related to divisions and tatums was that of partially and fully occupied beats. As seen in Table 6.8, a fully occupied beat is defined as any beat where the number of notes matched the division of the beat. Whereas, a partially occupied beat was any beat where the number of notes was less than the division, whether that be due to a mixture of held note lengths (e.g. division four with a semiquaver, quaver, and semiquaver) or rests within the beat.[24]

[23]Full details of the algorithm can be found in *Inside The Jazzomat: New Perspectives for Jazz Research* (Pfleiderer, Frieler, et al. 2017, 319).

[24]In fully occupied beats it was not guaranteed that the last note would have a similar duration to those before it, it could be held into the next beat without changing the division or tatum of the beat in which the onset occurred.

Table 6.8: Metrical divisions: fully and partially occupied beats.

Division 4	Tatum (Sub-division)
	Fully Occupied (1, 2, 3, 4)
	Partially Occupied (1, 2, 4)

It was illustrative to see how Green's density of beats changed across tempos.[25] This was accomplished by dividing Green's tempos into 20 BPM bins, and for each bin the mean metrical density for the four most common divisions (one, two, three, and four) were calculated. This can be seen in 6.23, where the size and opacity of the points showed the relative frequency of that division within the tempo segment.

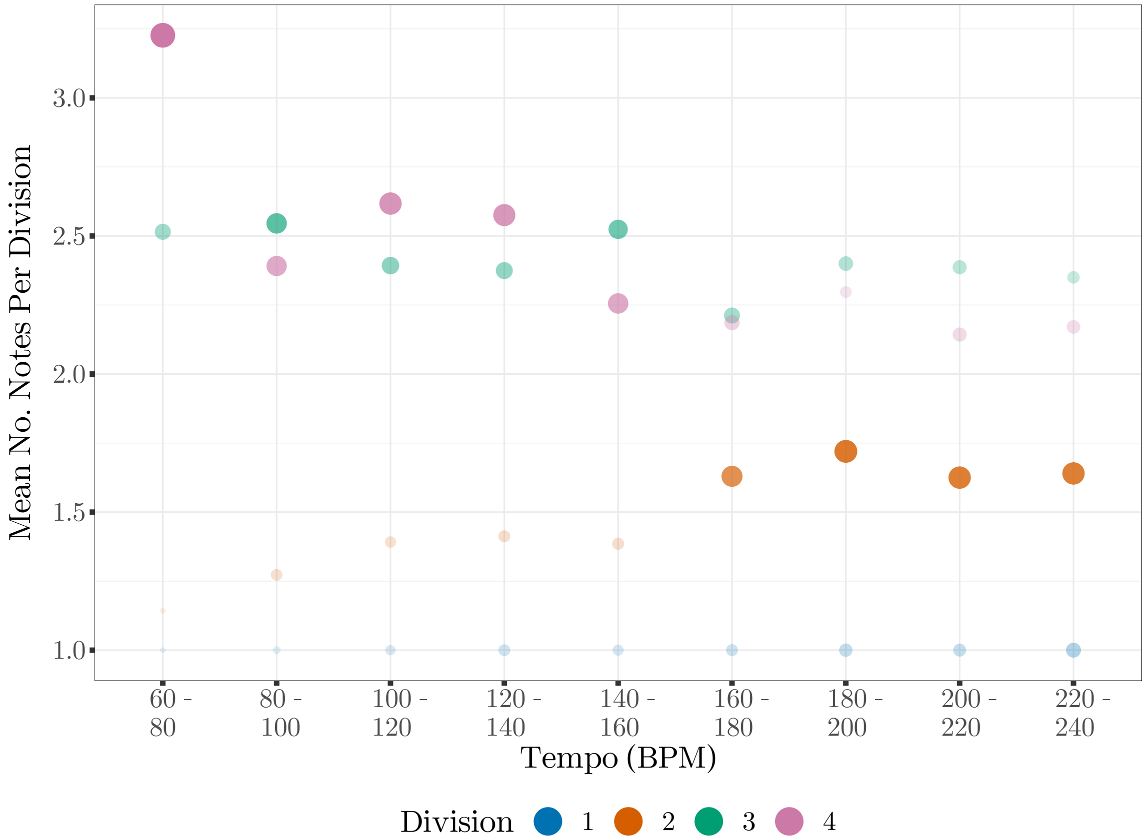

Figure 6.23: Mean metrical density of the most frequent beat divisions across the range of tempos. The size and opacity of the points showed the relative frequency of that division for each tempo segment.

While fully investigated below, this graph indicated that at slow tempos divisions of three and four were most common, with a division of two becoming most common in Green's playing when the tempo went above 160 BPM. Generally, the three and four

[25]With the caveat that certain tempos had very few data points.

note divisions tended to be equally occupied, around 2.5 notes per division, with this decreasing at higher tempos to around 2.25 notes on average. Consequently, divisions of three tended to be more fully occupied than divisions of four. These results also suggested that some beats of division three and four may have been heavily swung beats. An inverse trend was observed in divisions of two, with the mean occupancy increasing at higher tempos, with a mean occupancy rate of 1.75 notes per division. This indicated a high level of occupancy in Green's faster improvisations, suggesting the use of quaver-note lines.

The divisions generated by the FlexQ algorithm were used to assess the variety of rhythmic units Green used in his improvisations. Figure 6.24 shows the distribution of divisions found in Green's corpus. This data showed that nearly all (93.35%) beats in Green's improvisations had a divisions of one, two, three, or four. Of the higher divisions, only those with six notes were slightly frequent, trailed by beats with divisions of eight and five. The plurality (34.31%) of the beats had a division of two, unsurprising as quaver note lines form the basis of much of improvised jazz. Around one fifth had a division of either four or one, with a division of one indicating that there was only one note, played on the beat, but could have been of any duration. Of the four main divisions, beats with a division of three were least frequent.

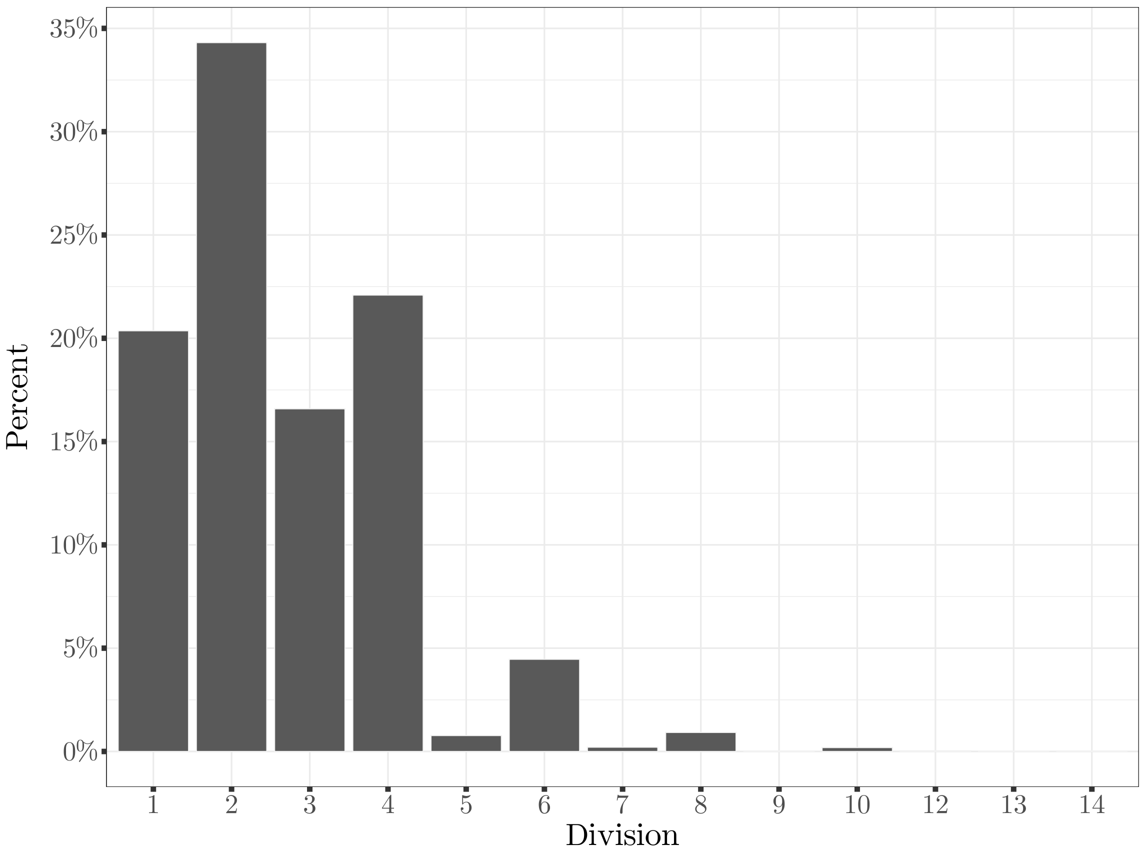

Figure 6.24: Distribution of divisions in Green's corpus.

As nearly all the beats in Green's improvisations had divisions of one, two, three, or four, the following analyses focused on only those divisions. Figure 6.25 shows the single beat MCM distributions for the four most common divisions (divisions one to four from top left to bottom right).[26,27]

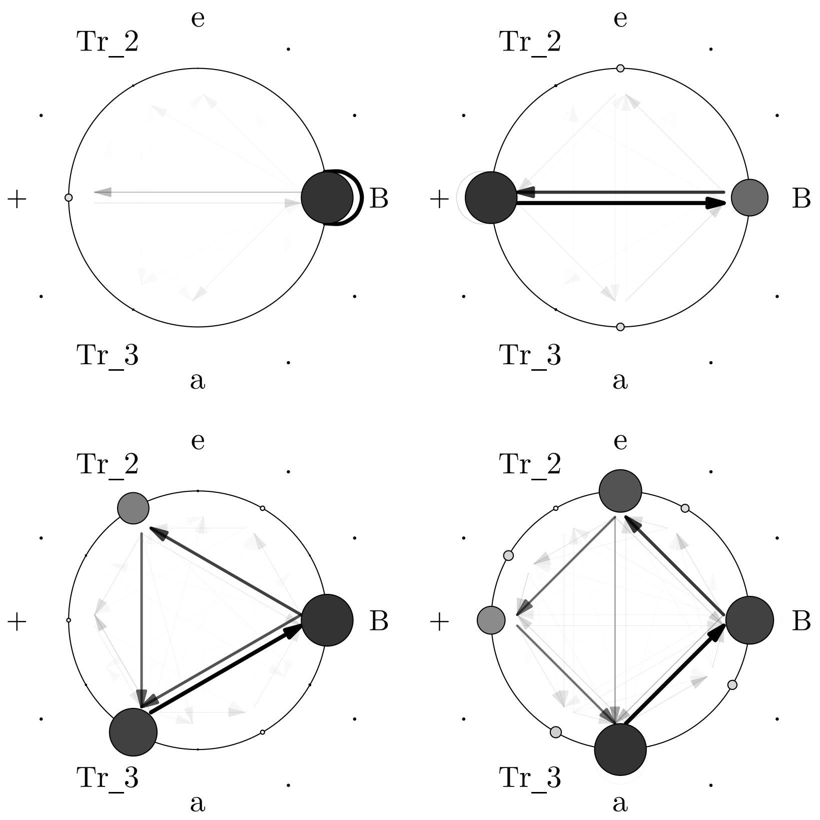

Figure 6.25: Single Beat MCM circle graphs for divisions 1, 2, 3, and 4 in Green's corpus. Top left: division 1. Top Right: division 2. Bottom Left: division 3. Bottom Right: division 4.

Notes played on the beat, with a beat division of one, were most likely to be followed by another note on the beat. Notes played in beats with division two mainly moved between the two quaver-equivalent on and off beat positions, as in a quaver note line. The repeated note line around the '+' position was indicative of off-beat notes followed by another off-beat note, as in syncopated lines. As suggested previously some quaver note pairs were assigned to higher divisions. This can be seen in the MCM graph for division three, where the down beat note went only slightly more frequently to the second tatum (854) than the third (772). Nearly all notes played by Green on the second or third tatum followed on to their next respective tatum. This showed that for beats with division three played by Green,

[26] Tr_2 and Tr_3 labels show the second and third triplet position in a beat, while the . show other unlabelled sub-beat placements.

[27] The bar distributions of the four divisions for each time signature can be found in Appendix A, Figures A.4, A.5, and A.5, for $\frac{4}{4}$, $\frac{8}{4}$, and $\frac{3}{4}$ respectively.

both true triplets and swung quavers pairs were both likely. For notes played in a beat with division four, there was no clear evidence of heavily swung quavers being incorrectly classified. Most of Green's notes played in the first tatum were followed by a note in the second tatum, with all other notes most likely to be followed by a note in the next sequential tatum. Although the majority of Green's notes moved sequentially throughout the tatums, the figures also indicated many other complex sub-beat movements.

Altogether, this data showed a high degree of rhythmic variety in Green's improvisations, with complex sub-beat transitions not uncommon common within his improvisations, even if sequential sub-beat movements were most frequent. The following sub-sections investigated how other features influenced the distribution of Green's divisions, as a representation of the rhythmic variety in his improvisations.[28]

Rhythmic Variety vs. Tempo

The first investigation focused on the relationship of Green's distribution of divisions and the tempo range, building upon the analyses in the Beat Distribution section, and the Rhythmic Variety introduction. Based on those initial analyses, and prior musical experience, a connection between the tempo and variety of divisions was expected to be observed in Green's improvisations. Specifically, it was hypothesised that the proportion of beats with a division of one or two would increase in the higher tempo range, while beats with division three and four would decrease.

Figure 6.26 shows the distribution of the four main divisions for each tempo range. This data showed that beats with divisions of one and two were more frequent at BPM $<$ 170 while divisions of three or four were less common. The greatest difference was observed for division two, increasing from 15.61% of beats at lower tempos, to 50.64% of all beats at higher tempos. The division with the second largest change was beats with a division of four, decreasing 27.74 PPs. A χ^2-test found a significant relationship with a large effect size between the tempo range and the distribution of these divisions in Green's improvisations ($\chi^2(3) = 1814.78$, $p =< .001,\ V = .43$).

[28]Except where explicitly specified the following analyses were based on a single data point for each beat in every bar.

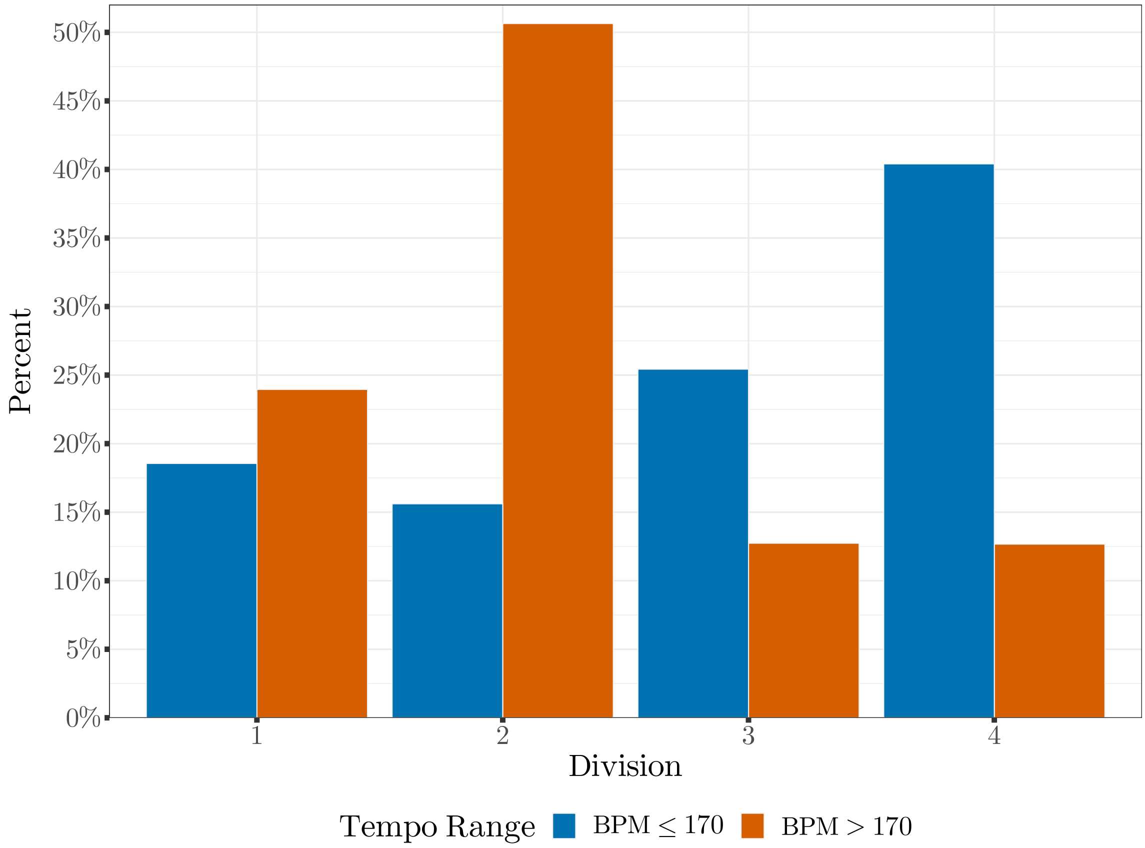

Figure 6.26: Distribution of divisions in each tempo range in Green's corpus.

This data supported the hypothesis, with Green's distribution of divisions differing significantly based on the tempo range. Figure 6.27 shows the single beat MCM graphs for tempos ≤ 170 BPM (left) and > 170 (right) BPM. These graphs showed the varied and complex rhythmic variety that occurred at lower tempos in Green's improvisations compared to more simplified and rigid structures at higher tempos.

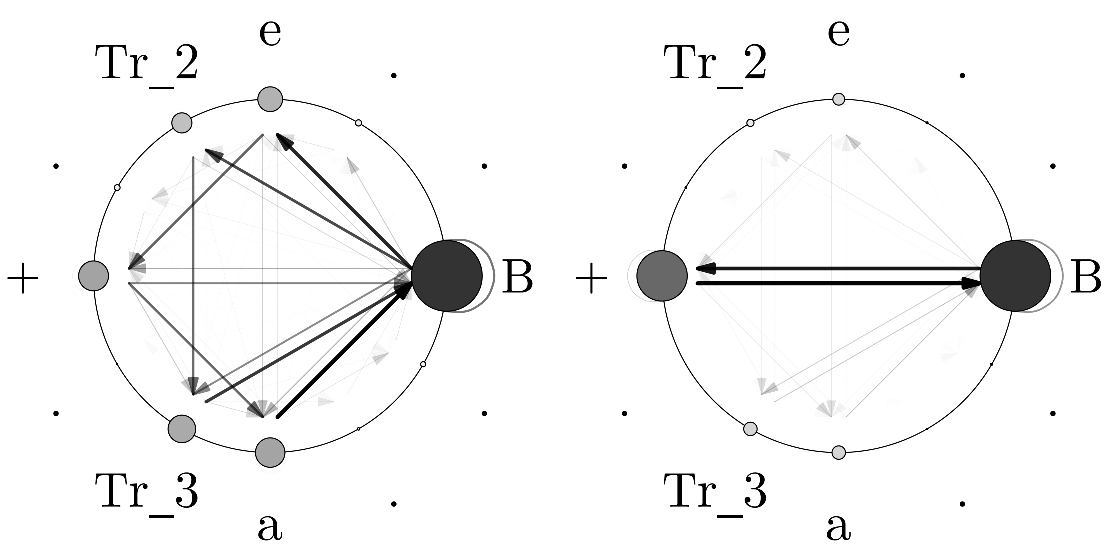

Figure 6.27: Single beat metrical circle maps showing sub-beat movements for both tempo ranges in Green's corpus. Left: BPM ≤ 170. Fight: BPM > 170.

Rhythmic Variety vs. Tonality Mode

The tonality mode analysis aimed to investigate Scott's statement that Green's predominant rhythmic unit, in a single blues improvisation, was a triplet (Scott 2006). Expanding upon Scott's statement to apply more broadly to Green's improvisational style, the hypothesis was that Green's divisions would differ significantly based on the tonality mode. Specifically, it was hypothesised that beats of division three were frequent over a blues.

Figure 6.28 shows the distribution of divisions across the tonality modes in Green's improvisations. While beats with a division of one were consistent across the tonality modes, Green played substantially fewer beats with division two when improvising over a blues compared to major and minor tonalities. Consequently, there were more beats with a division of three or four in Green's blues improvisations. Beats with division three were least frequent for all tonality modes. For the blues tonality mode, the frequency of beats with division three (20.78%) was slightly lower than beats with division one (20.90%).

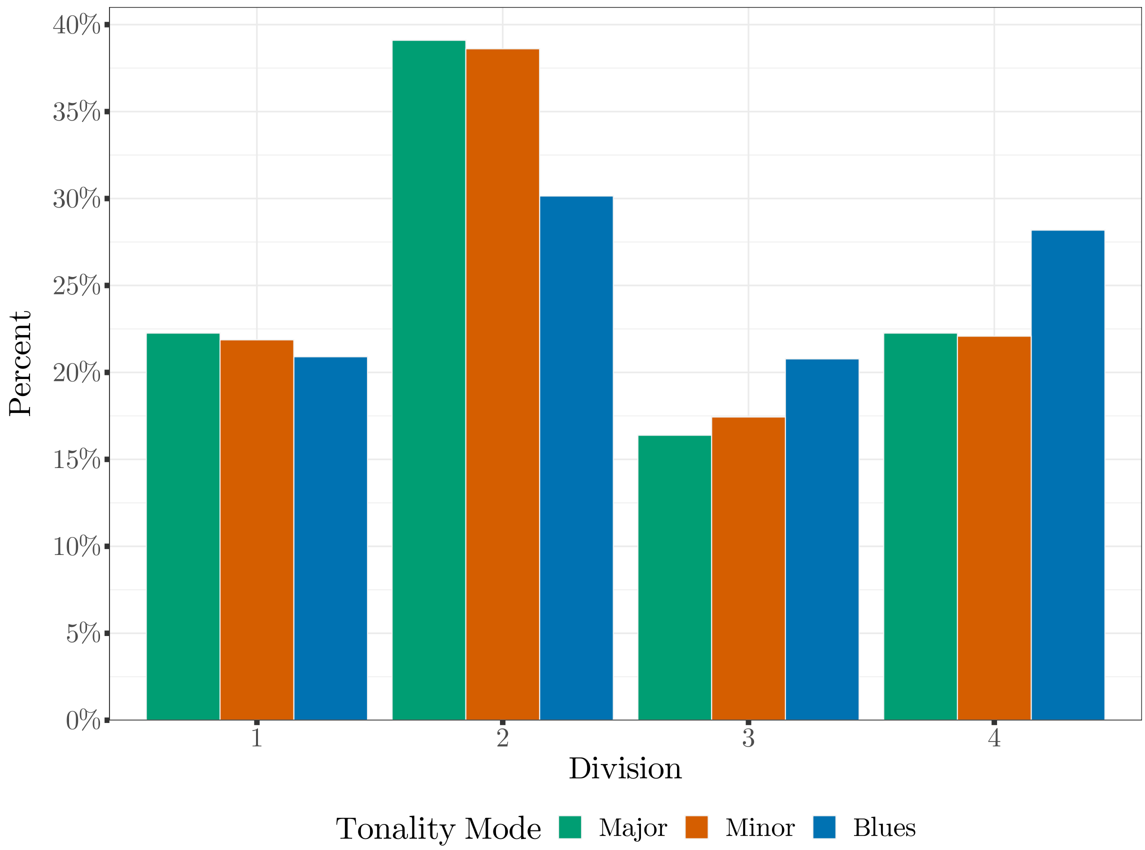

Figure 6.28: Distribution of divisions in Green's corpus dependent on the tonality mode of the improvisation.

A χ^2-test found a significant relationship between the distribution of divisions between the tonality modes ($\chi^2(6) = 85.62,\ p =< .001,\ V = .07$), with a small effect size. Subsequent post-hoc tests found no significant pairwise differences between major and minor tonalities ($p = .703$), but found significant differences in the distribution of divisions between blues and both major and minor tonalities ($p < .001$). Although the effect size was small, this data did support the hypothesis that Green's beat divisions did differ significantly depending on the tonality mode. Regarding Scott's statement, while this analysis found that beats with division three were more common over a blues when compared to major and minor tonalities, the plurality of all beats were division two.

Figure 6.29 shows an example of a predominantly quaver based line in a major tonality and a predominantly quaver triplet based line over a blues. The examples were explicitly selected to show these particular rhythmic units over the tonality modes, a similar rhythmic pattern could also be found in the opposite tonality modes. The top example, of quavers in a major tonality mode, is from Green's improvisation over *I Wish You Love* (Green 1964a). The bottom example, of quaver triplets over a blues, is from Green's improvisation over *Freedom March* (Solo 1, Green 1961d).

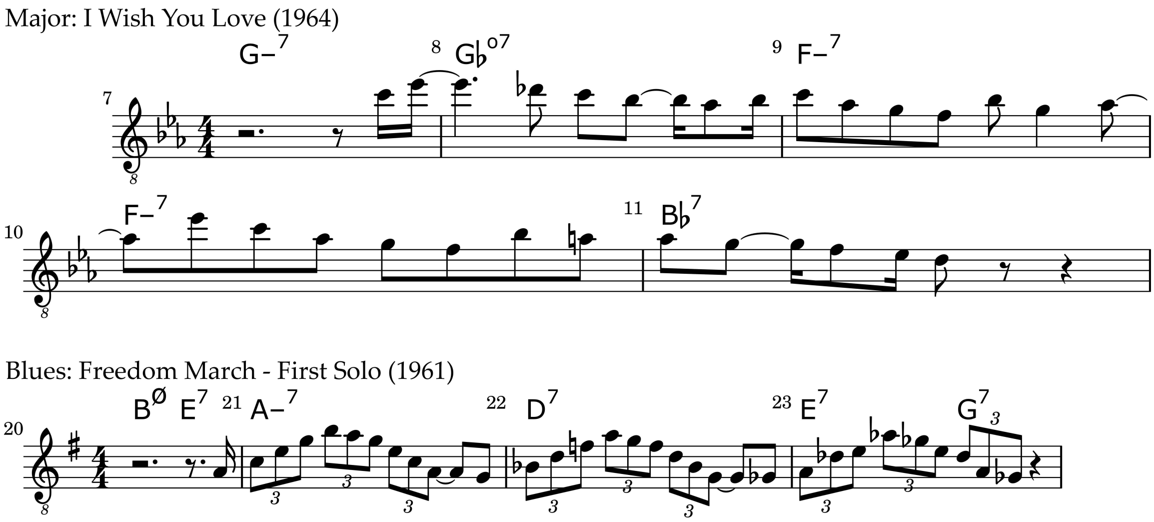

Figure 6.29: Examples of the difference in rhythmic variety of Green's improvisations in a major and blues tonality mode. Top: *I Wish You Love* (1964), bars 7–11. Bottom: *Freedom March* (Solo 1, 1961), bars 20–23.

Rhythmic Variety Summary

These analysis investigated the variety of rhythmic units played throughout Green's improvisations. They found that around one-third of all beats Green played had a division of two, with divisions of three being the least common of the main divisions. A note played by Green on a beat with division of one was most likely to be followed by another note played on the down beat. Notes played in beats with division of two likely to move between on and off-beat positions, with some evidence of syncopation in Green's playing. The tempo range of the improvisation had the largest effect on Green's distribution of divisions, with more than half of all beats at higher tempos having a division of two. The predominant rhythmic unit for all tonality modes was that of a quaver-equivalent pulse (division two), while Green was slightly more likely to play triplet rhythms (division three) when improvising over a blues.

Overall, Green had a reasonable amount of rhythmic variability within his improvisations. At slower tempos, the extra time allowed him to play more complex rhythms with a greater variety of sub-beat placements. As the tempo increased, the rhythmic variety and complexity of sub-beat placements decreased; however, a reasonable amount of variability in Green's rhythmic units remained. While other features may have had some impact on the variety of Green's rhythmic units, their effect would have been overshadowed by the strong influence of the tempo range.

6.5.2 Metrical Density

The metrical density investigated how many notes Green played in each bar of his improvisations. While the rhythmic variety looked at the sub-beat patterns played by Green, the metrical density took a broader view, extrapolating away from the specific beat patterns to their general effect on the density of notes played. Two ends of metrical density spectrum are often considered to be Coltrane, with his very dense playing referred to as "sheets of sound" (Ira Gitler in Frieler 2020, 127), and Davis who was "more economical, playing far fewer notes" (Griffin and Washington 2008, 10).[29] This section investigated both the general metrical density of Green's improvisations, as well as the relationship between the metrical density and the tempo range, the proportion of chromatic intervals in a bar, and the ratio of harmonic to non-harmonic tones in a bar.[30,31]

[29] The "sheets of sound" descriptor is "normally ... reserved for Coltrane's style from ca. 1958–1960" (Frieler 2020, 127).

[30] Due to their rarity, bars with seventeen or more notes were excluded from these analyses, this excluded nineteen bars from Green's corpus, keeping 99.43% of all data.

[31] The metrical density analyses focused only on bars that contained at least one note onset event, any bars with zero onsets were not included.

The metrical density, as notes per bar, was highly dependent on the time signature of the improvisation. As time signatures with more beats per bar could then have a proportionally higher number of notes per bar without changing the actual density of the improvisation. For Green's corpus this presented two challenges, the first for improvisations that were initially transcribed in $\frac{8}{4}$, and the second for how to deal with the differing number of beats in $\frac{4}{4}$ and $\frac{3}{4}$. The solution to $\frac{8}{4}$ was simple, the bars could be split to make two bars of $\frac{4}{4}$, with the number of notes per bar counted for each of the two new $\frac{4}{4}$ bars. For the differing number of beats between quadruple and triple time two possible solutions were considered. First, the number of notes per bar could be scaled by the time signature by dividing the number of notes per bar by the number of beats in that bar. While the advantage would have been that this would scale across a wide range of time signatures, there were also disadvantages. The main disadvantage was that the resulting division would create a mean number of notes per beat, instead of notes per bar, making it dissimilar to the rhythmic variety. The second option, and the one chosen, was to focus the analysis on only those improvisations that were in quadruple time. This option removed data from only three improvisations, which comprised 246 bars and 7.38% of Green's data.

Metrical Density General Distribution

Figure 6.30 shows the distribution of the number of notes in each bar of Green's improvisations (that were not in $\frac{3}{4}$ and had sixteen or fewer notes). There were 3068 bars from which to draw data, with the mean (6.20 ± 2.93 notes per bar) and median (6, IQR: 4–8, notes per bar) both indicating a similar general metrical density of six notes per bar. The distribution was neither heavily skewed (skewness: 0.74) nor had a long tail (kurtosis: 0.74), with the graph and these statistics indicating the data was only slightly right-skewed. Green's data showed that 64.80% of his bars had between four and eight notes per bar. This agreed with the previous findings of Green often using quaver-equivalent notes to form his lines. This was supported by the metrical density data, with notes in bars with only one or two notes having a mean $\text{duration}_{\text{BeatProp}}$ of just under a crotchet-equivalent (1: 0.83 beats; 2: 0.86 beats); while bars with between four and eight notes had mean $\text{duration}_{\text{BeatProp}}$ between a quaver triplet or quaver-equivalent (4: 0.53 beats; 8: 0.37 beats).

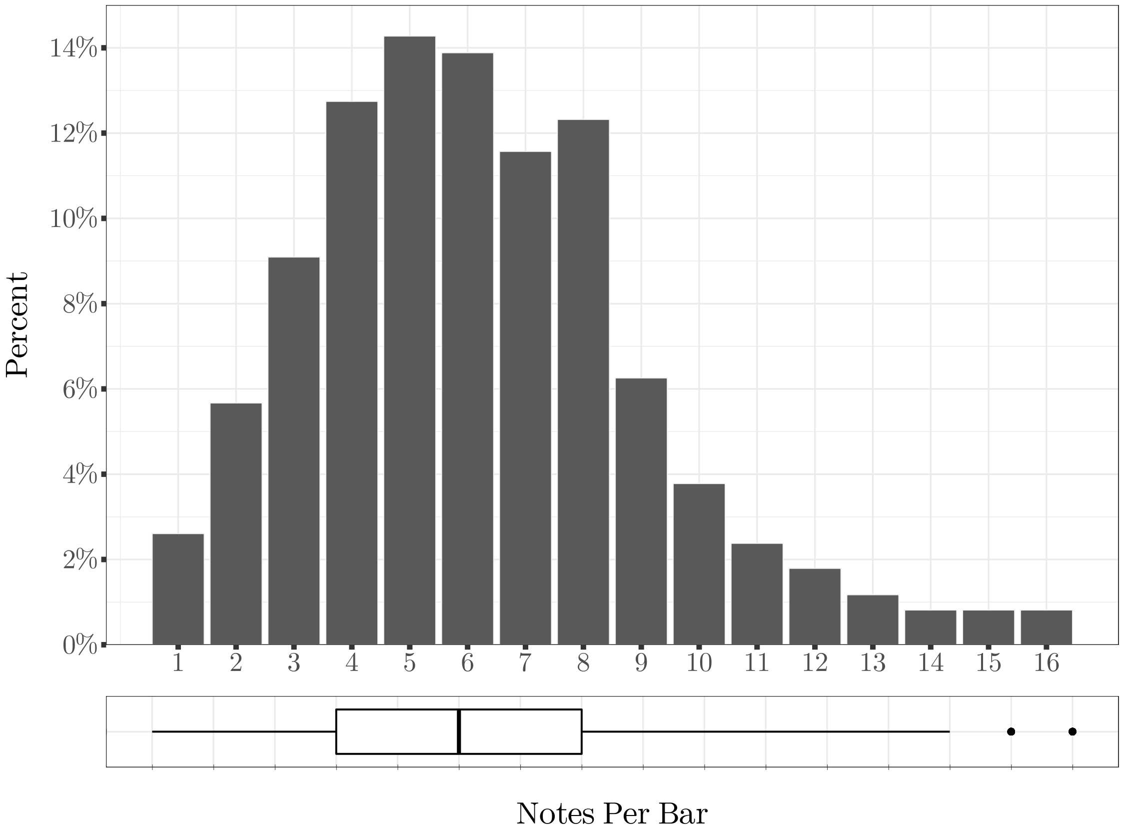

Figure 6.30: Distribution of the metrical density, as number of notes per bar, in Green's corpus.

To contextualise Green's metrical density distribution, it was compared to those of Coltrane, Davis, and Parker.[32] The distributions of the four performer's metrical density can be seen in Figure 6.31. While the data supported the concept that Davis had a lower metrical density, due to him playing a lot of space in his improvisations, it did not support Coltrane's "sheets of sound". This was due to the WJazzD having limited data from the period where this concept was usually applied to his improvisations, with his data spanning nearly a decade of improvisations (1956–1964). Green had a similar metrical density to Coltrane, tending to be more dense than Davis but less than Parker. This fit within expected distributions for Green considering the period of time under investigation, the styles of jazz he predominantly improvised over during that time, and his early influences, which included Parker. While there were differences between the performers, the median metrical density only differed by three from the lowest – Davis with a median four notes per bar – to the highest – Parker with seven notes per bar.[33]

[32]Bars with seventeen or more notes or played in non-quadruple time were excluded from all performers data.

[33]An ANOVA found a statistically significant difference in the metrical density between the performers, with a small effect size ($F(3, 8404) = 166.01$, $p < .001$; $\eta^2 = .06$). Subsequent post-hoc tests using Tukey's HSD procedure found significant pairwise difference between the performers at $p < .001$.

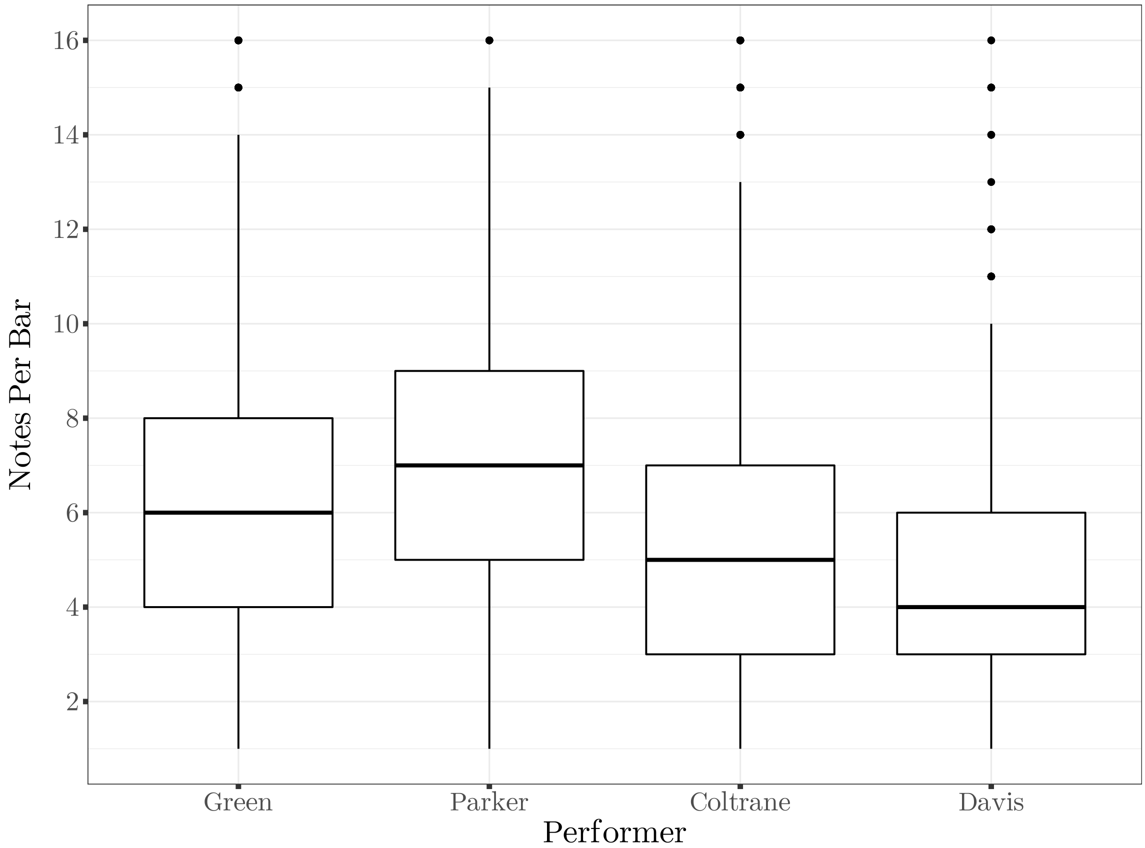

Figure 6.31: Comparison of metrical densities between Green and Parker, Coltrane, and Davis.

Metrical Density vs. Tempo

Based on the prior analyses, the feature that was expected to have the largest effect on Green's metrical density was the tempo of the improvisations. It was hypothesised that Green's metrical density would be lower at tempos $>$ 170 BPM. It was also hypothesised that there would be a smaller range of metrical densities in Green's improvisations at these faster tempos when compared to tempos $\leq$ 170 BPM.

Figure 6.32 shows the metrical density distribution for each tempo range. When Green improvised at a tempos $>$ 170 BPM the most common metrical densities were bars with four, five, six, or eight notes per bar. Bars with three or seven notes were the next most frequent, while only a few bars (forty-two) had more than nine notes, with none of Green's bars having more than twelve notes. The difference of metrical densities from seven to eight notes was almost certainly due to Green playing fully-occupied bars of quavers when improvising. When Green improvised at tempos $\leq$ 170 BPM, 70.17% contained between four and ten notes. Bars where Green played more than ten notes were not very common, with 17.06% contained between eleven and sixteen notes. This data also showed a wider variety of metrical densities for Green at the lower tempo range when compared to higher tempos.

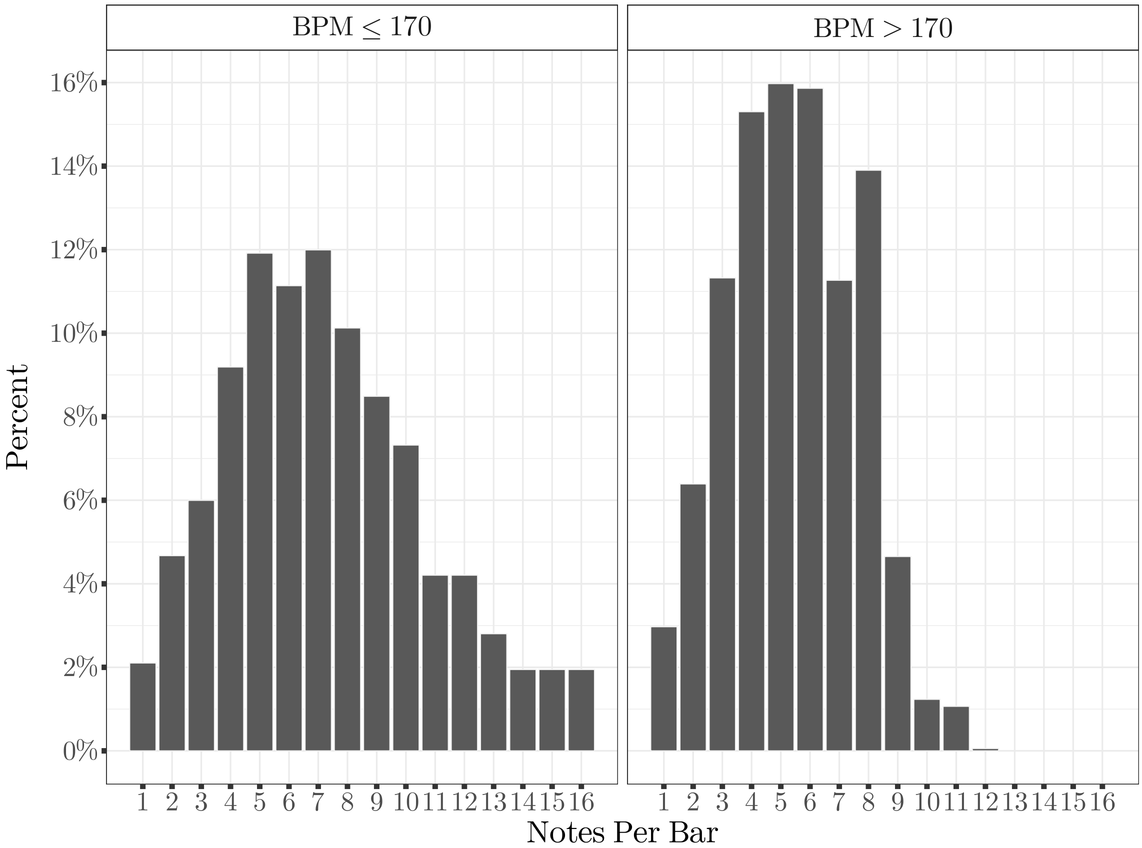

Figure 6.32: Distribution the number of notes per bar for the binary tempo ranges in Green's corpus.

When improvising at tempos $\leq$ 170 BPM Green played a mean of 7.28 $\pm$ 3.44 notes per bar. In comparison, when improvising at tempos > 170 BPM Green's metrical density dropped to a mean of 5.43 $\pm$ 2.19 notes per bar. These results showed that Green played on average two fewer notes in each bar at higher tempos when compared to lower tempos. A t-test found a significant difference in the number of notes per bar between the two tempo ranges, with a large effect size ($t(2017.43) = -16.98,\ p < .001;\ d = -0.76$). These results supported the hypothesis that Green's metrical density decreased at higher tempos.

The distributions shown in the graph, and the smaller standard deviation found at the higher tempos, also supported the second hypothesis that the range of the metrical densities played by Green was less at higher tempos. The vast majority (83.63%) of bars at tempos > 170 BPM had between three and eight notes. In comparison, less than two-thirds (60.36%) of bars at tempos $\leq$ 170 BPM had a metrical density within this range. Together these indicated that Green played a smaller variety of metrical distributions at higher tempos when compared to lower tempos, with Green's playing also tending to be more metrically dense throughout the lower tempo range.

Metrical Density vs. Chromatic Interval Proportions

The chromatic interval proportion (chromatic proportion) was the proportion of notes in a bar that moved chromatically to the following note. A chromatic proportion of 0 meant there were no chromatic intervals in the bar while a proportion of 1 meant all notes moved chromatically. The hypothesis was that the proportion of chromatic intervals would increase with the metrical density. While a chromatic interval did not necessarily imply a higher degree of NHTs or NDTs, a higher proportion of chromatic intervals could be indicative of this. The frequency of NHTs, as it related to the metrical density, was analysed separately.

Figure 6.33 shows the distribution of chromatic proportions for each of the metrical densities in Green's corpus. While an increasing trend can be observed in the data, there was also a lot of variation. There were bars throughout the metrical density distributions that had both very high and low chromatic proportions. A significant positive correlation was found between the density and chromatic proportion, with a medium effect size ($r = .29,\ t(3066) = 16.91,\ p < .001,\ r^2 = .085$).

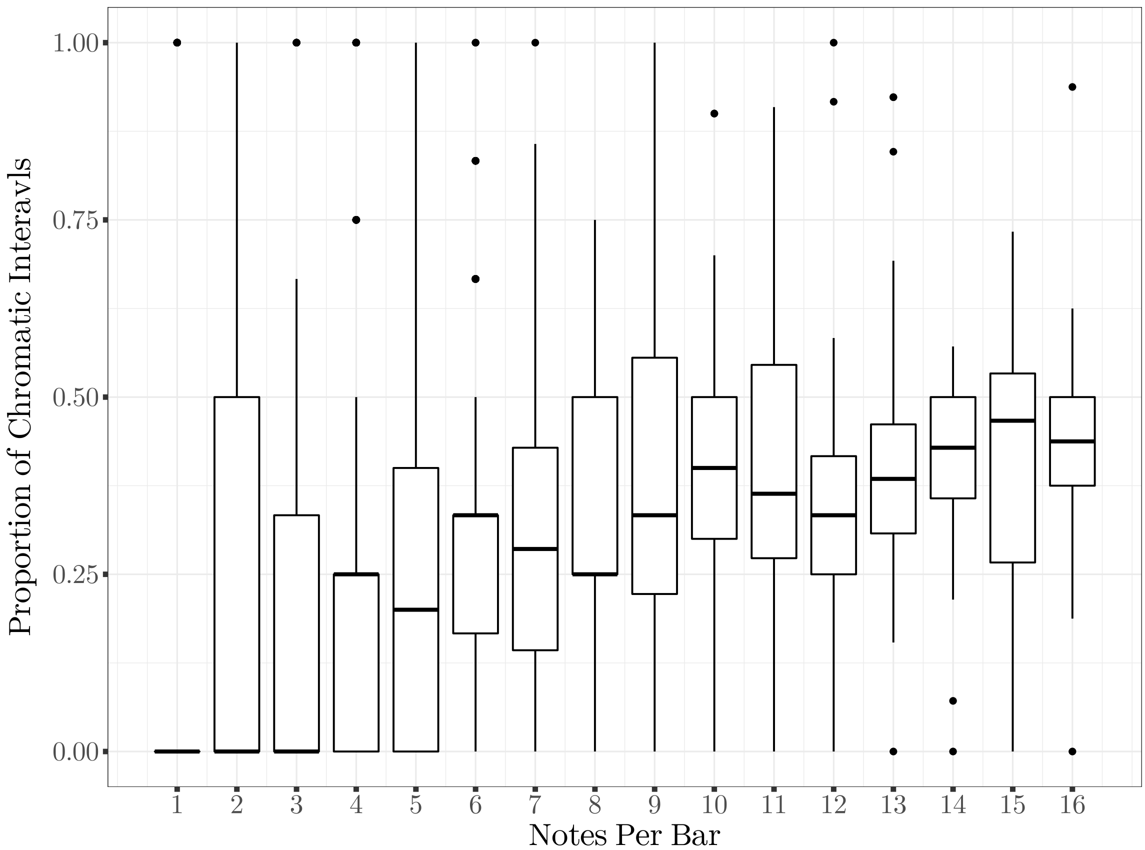

Figure 6.33: Distribution of proportion of chromatic intervals for each metrical density in Green's corpus.

As a medium correlation, it supported the hypothesis that as the metrical density increased, so did Green's level of chromaticism. It was probable that much of the correlation was related to the limited range of available chromatic intervals when there were only a small number of notes per bar. Although the data in the graph indicated a rising median chromatic proportion as the metrical density increased above six notes per bar, the median of bars with eight or more notes never extended beyond the 0.25 to 0.50 range. This indicated that, for metrically dense bars, Green tended to play between a quarter and half of his notes with chromatic movements.

Two likely explanations for the increase in the chromatic proportion as the metrical density increased were:

1) A higher metrical density resulted in more off-beat notes, allowing Green more opportunities to play chromatic passing or chromatic approach tones;
2) Patterns of alternating tones that were a semitone away from each other.

Examples of both of these can be seen in Figure 6.34. The top musical example, with more frequent chromatic passing and approach tones, was from Green's improvisation over *I'll Remember April* (Green 1961k). There was a chromatic approach tone on the second tatum of beat one and a chromatic passing tone on the second tatum of beat two. There were two other chromatic intervals in that bar, from the C in tatum two of beat three to the B in tatum one of beat four, and the last note of the bar, which moved chromatically down to a G# on the first beat of the next bar.[34] The bottom musical example showed Green alternating between two notes that were a semitone away (A and B♭), from his improvisation over *At Long Last Love* (Green 1965a). This bar had a high metrical density and proportion of chromatic intervals, but a low degree of melodic complexity or variation.

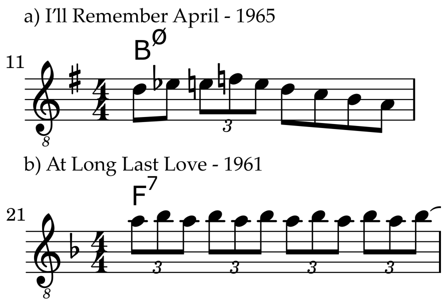

Figure 6.34: Examples of chromatic movement in metrically dense bars. a) Chromatic approach and target tones, *I'll Remember April* (1961), bar 11. b) Alternating chromatic intervals *At Long Last Love* (1965), bar 21.

This analysis found that the proportion of chromatic intervals did increase with the metrical density in Green's improvisations. There were multiple reasons why the proportion of chromatic intervals may have increased with the metrical density. Therefore, while an interesting correlation between the two features, it did not provide insight into Green's improvisational style.

[34]These were chromatic approach tones and part of a descending diatonic line.

Metrical Density vs. Frequency of Non-Harmonic Tones

The final analysis investigated how an increased metrical density influenced the proportion of HTs and NHTs in a bar. The hypothesis was that the proportion of NHTs would increase in more metrically dense bar. Figure 6.35 shows the distribution of proportions of NHTs for each metrical density in Green's improvisations. This data showed a small increase in the median proportion of NHTs as the metrical density increased. Statistically, a significant correlation between the proportion of NHTs and metrical density was found, with a small effect size ($r = .16,\ t(3066) = 9.11,\ p < .001,\ r^2 = .026$). However, these results did not show the true relationship between the features. As can be seen at the lower metrical densities, while most had no NHTs, there were also less dense bars containing only NHTs. Therefore, another approach was to investigate the frequency of bars with a non-zero number of NHTs. The updated hypothesis was that, while many of the lower density bars would most often have zero or one NHTs, bars with a higher metrical density would nearly always have at least one NHT.

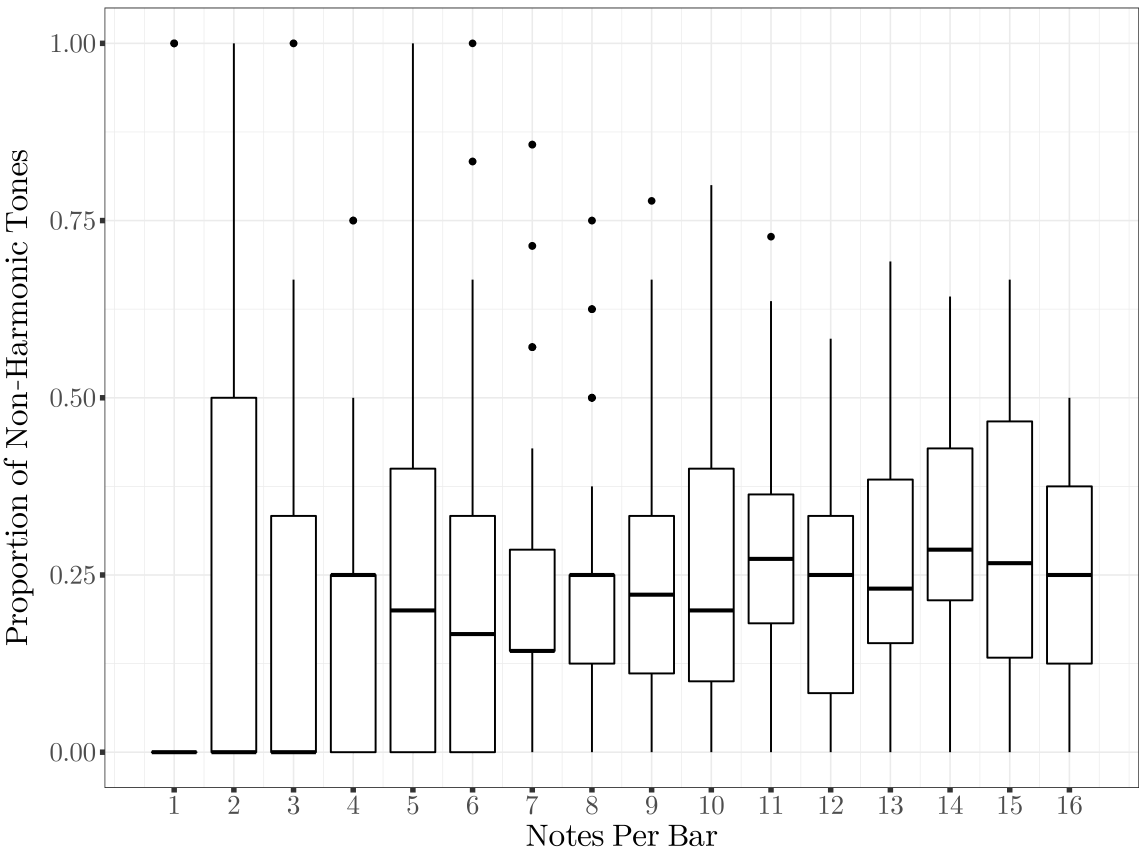

Figure 6.35: Distribution of proportions of NHTs for each metrical density in Green's corpus.

Figure 6.36 shows, for each metrical density, the proportions of bars in Green's data that contained one or more NHTs. This data indicated strong support for the stated hypothesis. For bars with only a single note, there were only five bars (6.25%) where it was a NHT. This jumped up to 32.76% of bars with two notes containing at least one NHT. The proportion of bars containing at least one NHT increased fairly linearly from bars with two to six notes, with 64.32% of six note bars having at least one NHT played in them. There was another jump in the data at a metrical density of seven, with 77.18% of bars containing one or more NHTs. For the highest density bars, between approximately 75% and 95% of bars had at least one NHT. A correlation test found a significant relationship between the metrical density and proportion of bars that contained at least one NHT, with a large effect size (r = .88, $t(14) = 6.79,\ p < .001,\ r^2 = .77$).

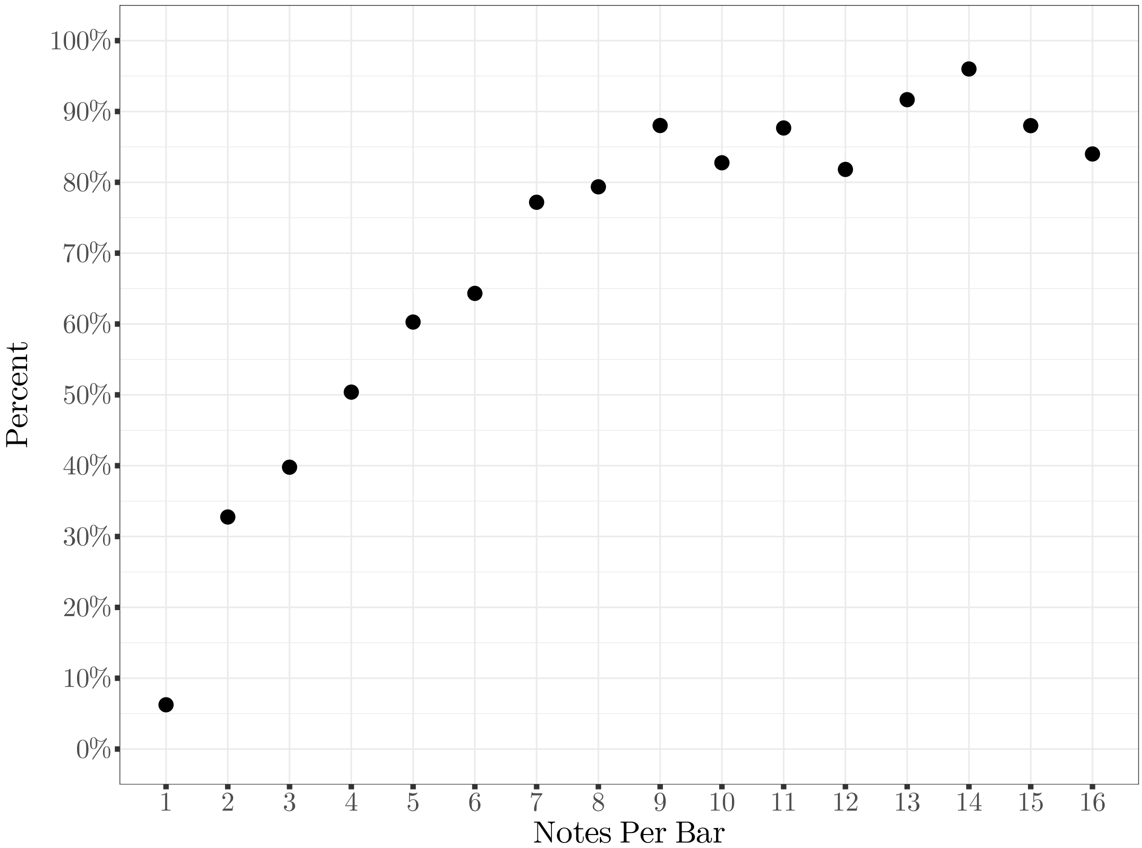

Figure 6.36: Percentage of bars for a given metrical density that had at least one non-harmonic tone.

These results supported the hypothesis that within Green's improvisations there was a significant relationship between the metrical density and the occurrence of NHTs. The analyses found that of all bars with four or more notes, the majority contained one or more NHTs. While the higher the metrical density went, the more likely it was to contain at least one NHT.

Metrical Density Summary

This analysis focused on the metrical density of bars within Green's improvisations. Green played six notes per bar on average, with two-thirds of bars having between four and eight notes. As with the variety of rhythmic units used by Green in his improvisations, while many features may have influenced Green's metrical density, the tempo range in which he was playing had the largest effect. At lower tempos Green averaged around seven notes per bar, with this dropping to five notes per bar at tempos > 170 BPM. Green's also played a tighter distribution of metrical densities at higher tempos, with few highly dense (> 10 notes per bar) bars. The analysis found some evidence that the frequency of chromatic intervals increased along with the metrical density. This was likely caused by higher density bars having more sub-beat notes for chromatic passing and approach tones, or the use of repeated note patterns that contained a chromatic movement. Investigations into the relationship between the metrical density and frequency of NHTs indicated that as the number of notes per bar increased, so did the frequency of NHTs in Green's improvisations. In bars with three or fewer notes, Green was unlikely to play any NHTs. In contrast, more than three-quarters of bars with seven or more notes had at least one NHT.

6.6 Green's Improvisational Style In The Rhythm Domain

This chapter focused on the analysis of Green's improvisational style within the rhythm domain. The chapter began with an analysis of the lengths of the notes Green played, as both duration and IOI, and the distribution of notes within the beat structures of the bars. This was followed by analyses into the metrical weights of Green's notes and an investigation into the use of rests within Green's improvisations. This chapter concluded with investigations into two specific rhythmic based features, the variety of rhythmic units and the metrical density of bars in Green's improvisations.

The note length analysis found that the majority of Green's notes had a length between a semiquaver-equivalent and quaver-equivalent note. Green's note duration was not found to be heavily dependent on the tempo range of the improvisation; although short fuzzy IOI notes were less frequent at higher tempos. Green played more medium and long notes when the note was played on-beat when compared to off-beat notes. Similarly, Green was less likely to play medium or long IOI NHTs, with arpeggio tones less likely to have a short length.

The investigation into Green's beat distribution agreed with the findings of the note length analysis, with Green predominantly placing notes on the beat and on quaver-equivalent off-beat positions. The tempo had a substantial influence on Green's beat distribution, with Green playing a wider distribution of rhythms at lower tempos. As there were only a limited number of on-beat metrical positions within each bar, the analysis into Green's metrical weight found that the majority of notes were played off-beat. The metrical weight vs. tempo investigation also agreed with the prior analyses, with the results showing that at higher tempos around half the notes were played both on and off the beat, indicative of Green playing a quaver-equivalent note line.

The analysis into Green's use of rests within his improvisations presented an entirely new feature that was absent from *MeloSpy*. The analysis found that Green's rests went for around one beat on average, with this halving when the rest was played within a phrase, and doubling when played between phrases. The phrase position of the preceding note was strongly related to the length of Green's rests, with rests, especially long rests, nearly always occurring between phrases.

The rhythmic variety and metrical density analyses were related, and built upon the findings of the previous analyses. The rhythmic variety analysis found that the plurality of all beats, and the majority at higher tempos, had a division of two. There was also evidence that Green was more likely to use quaver triplet-equivalent rhythmic units when improvising over a blues. Overall, the analysis found that Green had a reasonable amount of rhythmic variety within his improvisations, with tempo having the greatest impact. At lower tempos, the extra time allowed him to play more complex rhythms with a greater variety of of sub-beat placements, with this decreasing at higher tempos. The metrical density was also strongly influenced by the tempo, with Green playing on average seven notes per bar at lower tempos compared to five notes at higher tempos. Across all tempo ranges, two-thirds of bars had between four and eight notes. The frequency of NHTs was also related to the metrical density, with sparse bars rarely having NHTs, while they were more frequent in denser bars.

In summary, this analysis found that Green's improvisational style with relation to the rhythm domain was heavily influenced by the tempo range in which he was improvising. The majority of notes had a rhythm ranging from a semiquaver to a quaver equivalent, with long notes and rests most frequently occurring only between phrases. Green's rhythms at lower tempos had a greater variety, with the extra time allowing Green greater flexibility in sub-beat placements. This was simplified at higher tempos, with Green predominantly playing quaver-equivalent lines.

The previous two chapters have focused on those features most commonly investigated through traditional analytical methods. The following chapter, analysing features of the Micro domain, was able to fully utilise the precise timings that came from the transcription method developed by the Jazzomat Research Project.

Chapter 7

Micro Domain

The micro domain contained micro-timing features that took full advantage of the high degree of precision provided by the transcription process. Unlike many of the features from the pitch or rhythm domains, most of those in the micro domain cannot be represented in standard symbolic notation. The three features that comprised this analysis of the micro domain were: swing; micro timings (articulation and micro-gaps between notes); and the placement of notes in respect to their nominal placement. Swing was the ratio of the two note lengths in a quaver pair. Articulation and micro-gaps between notes both describe similar features, with micro-gaps being the complement to rests, while the articulation was the ratio between the $\mathrm{IOI}_{\mathrm{BeatProp}}$ and $\mathrm{duration}_{\mathrm{BeatProp}}$. The final feature investigated whether Green tended to play his notes ahead of or behind the beat, by measuring the time offset between when Green played a note and the nominal placement of the note based on the FlexQ algorithm.

7.1 Swing

Swing is the long-short pattern of notes in a quaver pair that is synonymous with many styles of jazz. The swing ratio is a measure of the proportional note lengths between the first, on-beat, note and the second, off-beat, note. Following previous literature (e.g. Benadon (2006), Butterfield (2011), Wesolowski (2012), Hernandez (2020), and Corcoran and Frieler (2021)) the term beat-upbeat ratio (BUR) was preferred to refer to the swing ratio of the quaver pairs. Due to the importance of swing in jazz, it is amongst the individual features that has undergone the most study, both within the micro domain as well as more generally (see Table 1 in Corcoran and Frieler (2021), 373).

An example of three possible BURs, ranging from no swing (1:1) to a heavy swing (3:1), can be found in Figure 7.1. The BUR could be expressed as either a ratio, e.g. 2:1 or 3:2, or transformed to the form n:1, and written as a single value, e.g. 2 or

1.5, with the :1 implied. As many of Green's BURs were not simple ratios, the latter form was used.

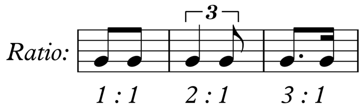

Figure 7.1: Symbolic notation equivalent of three possible BURs ranging from no swing (1:1) to heavy swing (3:1).

Due to how the swing values were encoded in the data exported from *MeloSpy*, the BURs had to be recalculated.[1] This also provided greater flexibility in calculating the BURs in Green's improvisations. Two main changes were made in the calculation of BURs. The first was that the criteria for what was considered a quaver note pair was changed. The second was a change in the note length features used in the calculation of the BURs. Additionally, although most features generated by *MeloSpy* were able to successfully be replicated from the raw data, the author was not able to replicate the BURs.[2]

In Corcoran and Frieler (2021), the BUR was calculated for "all annotated beats that contain[ed] exactly two events" (374), and where "the sum of both IOIs [were] less than the local beat duration (plus a small offset … tolerance window)" (374).[3] Their BUR was calculated by dividing the first note's IOI by the second note's IOI ("IOI: BUR = $\text{IOI}_1/\text{IOI}_2$" (Corcoran and Frieler 2021, 374).) In contrast, this research considered:

- all beats with a division of two, three, or four;
- for beats with a division of three, the notes had to be on the first and last tatum;
- for beats with a division of four, the first note had to be on the first tatum while the second note was on the third or fourth tatum.

The BUR was then calculated by taking the ratio of the first note's IOI to the second note's duration (BUR = $\text{IOI}_1/\text{Duration}_2$). The advantage of using the duration instead of the IOI for the second note can be seen in Figure 7.2. In this example, the notes in bar thirteen, beat three, of Green's improvisation over *At Long Last Love* were a swung pair (with a BUR of 2.83 when using the duration for

[1] In the exported data all the BURs were listed against the first n note events in each improvisation (where n was the number of swing events in each improvisation), instead of against the notes that were actually swung.

[2] In their paper Corcoran and Frieler provided more detail about the selection criteria and the generation of BURs (2021, 374). This appeared to apply only to the data used within that paper, and did not aid in reconstructing the data generated by *MeloSpy*.

[3] The paper's github page indicated that they only considered beats with a division of two, three, or four, https://github.com/klausfrieler/swing_ratios/blob/master/swing_ratio_analysis.R, line 212.

the second note); the Corcoran and Frieler (2021) selection criteria would exclude this note pair as the IOI of the second note was too long. Finally, only BURs between 0.98 and 3.02 were considered to be true swung notes.[4]

Figure 7.2: Example of a potential swing pair (bar thirteen, beat three) when using the duration of the second note, but not when using the IOI, *At Long Last Love* (1965), bars 13–14.

In attempting to replicate the data as presented in Corcoran and Frieler (2021), an updated version of the function to manually generate the BUR was written. In that process, 319 note pairs were found that had been incorrectly labelled as swing pairs from the original function. This would reduce the number of swung pairs in Green's data from 2971 to 2652. Although it was not clear what caused the mislabelling, as the original data was already used in many of the analyses, and in the training of the models in Part III, this chapter continued to use the original data with the extra BURs. The updated function would have resulted in only a marginally higher mean BUR of 1.77 ± 0.50, compared to the 1.75 ± 0.50. Critically, the BURs calculated were the same across both functions, with only the number of swung note pairs differing.[5] The following analyses investigated the overall frequency and distribution of BURs, how Green's distributions compared to those of Parker and Metheny's, and how BURs were affected by the tempo range, metrical density, and beat location.

7.1.1 BUR Distribution and Initial Comparisons

Within Green's corpus there were 2971 beats identified as containing a swung pair of notes with a BUR between 0.98 and 3.02. This represented 27.90% of beats in Green's corpus. There were an additional 729 beats that met the selection criteria but had a BUR outside of 0.98 and 3.02. Corcoran and Frieler (2021) labelled pairs of notes as one of three categories: snap (short-long), even, and swung (long-short). Although they set a range (1:4 to 4:1) for all BURs, they did not report a limit for each category. As the limits set for the BUR in this research were already different, a range for each of the categories was determined based upon those limits, informed

[4]A value less than one would indicate the second note was longer than the first while a value greater than three would be indicative of a pair of notes that would not sound swung. Corcoran and Frieler (2021) used the term "snap" to refer to note pairs with a short-long pattern. The values of interest were those between 1 and 3, with a small buffer to ensure all valid data was included.

[5]See Swing (Beat-Upbeat Ratio) in Appendix B for the function `manualSwing` that was written to generate the swing data.

by the prior research. Snap was defined as a BUR between 0.3 and 0.9, Even between 0.9 and 1.12, and Swing between 1.12 and 3.02. The distribution of Green's data into these classes can be found in Table 7.1, including the number of two-note beats that did not meet the BUR selection criteria, and those with a BUR outside of the limits for Snap or Swing.

Table 7.1: Distribution of Green's BUR for beats with two notes into Snap, Even, and Swing.

	No BUR		Snap	Even	Swing	
		$x < 0.3$	$0.3 \leq x < 0.9$	$0.9 \leq x < 1.12$	$1.12 \leq x < 3.02$	$3.02 \leq x$
Count	799	67	309	277	2733	314
Percent	17.76%	1.49%	6.87%	6.16%	60.75%	6.98%

This data showed that the majority of beats (66.04%) with two notes matched the selection criteria to have a BUR calculated for it (BUR between 0.98 and 3.02), with the vast majority of those (73.86%) falling within the Swing category. There were only a few beats with a very low BUR (< 0.3) while there were approximately the same number of Snap, Even, and very high BUR (≥ 3.02) beats. This distribution showed that Green's use of swing did not match the findings of Corcoran and Frieler, who found that "uneven eights [were] considerably less common than more even eights in … improvisations" (2021, 376). This may be a feature of Green's improvisational style, but was also influenced by the styles of jazz Green most commonly played. Corcoran and Frieler found that improvisations in hard bop and post-bop tended to have higher BURs (Corcoran and Frieler 2021, 379), the styles most commonly associated with Green between 1960 and 1965. Due to the relative rarity of BURs outside of the Swing class in Green's improvisations, the separation of the data into classes was dropped in favour of a single BUR feature that encompassed all values between 0.98 and 3.02. Therefore, the change between swung and even notes was identified by a higher or lower BUR.

Green's distribution of BURs can be seen in Figure 7.3. The box plot showed that half of Green's BURs fell between 1.38 and 2.01, with a median value of 1.67. The histogram showed three main spikes at BURs of 1 (1:1), 1.5 (3:2), and 2 (2:1). The box plot showed that only a quarter of Green's BURs had a value higher than 2.01, indicating that while he did occasionally play note pairs with a very heavy swing the 75% of the time his BUR was between one and two. Green's mean BUR was slightly higher than his median, at 1.75 ± 0.50. Altogether, this data indicated that Green played a wide variety of BURs, predominantly within the range of one to two, with higher frequency of BURs around standard swing values (1, 1.5, and 2).

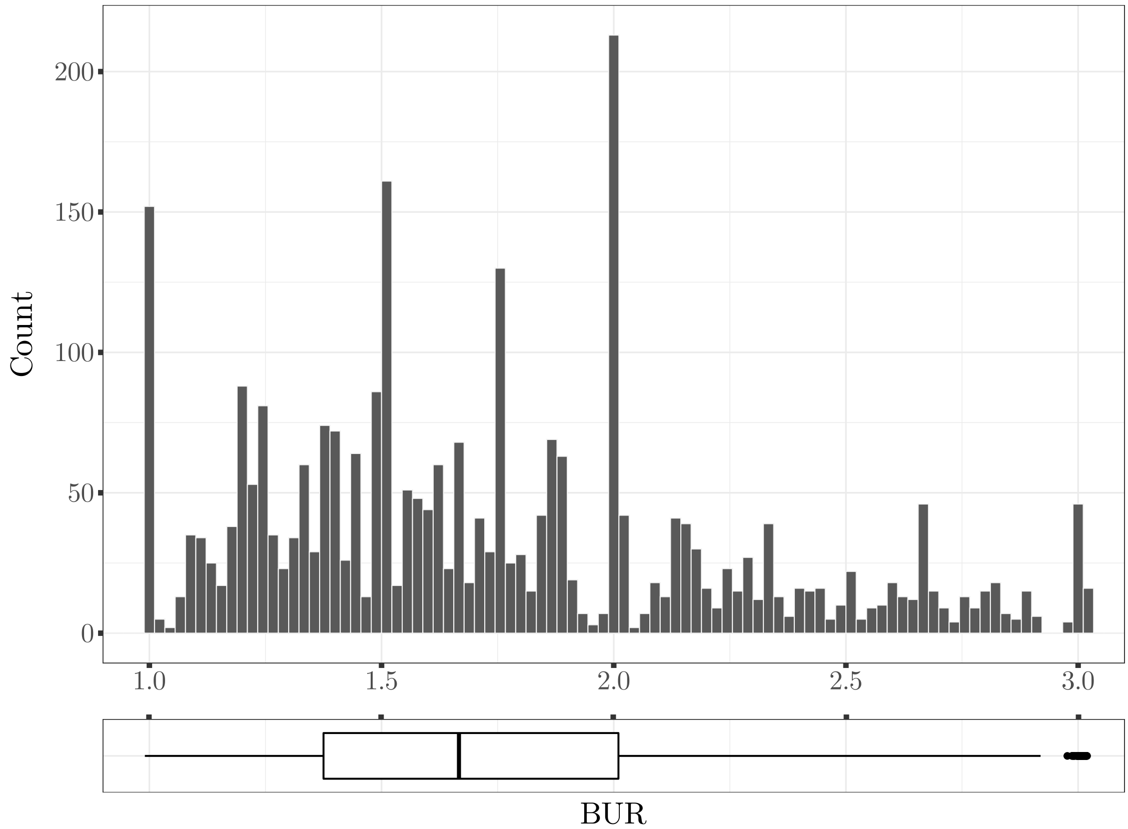

Figure 7.3: Distribution of BURs in Green's corpus.

To situate Green's use of swing, his BUR was compared to all other performers in the WJazzD. Only thirteen of the seventy-eight musicians in the WJazzD had a higher mean BUR than Green (including Coleman Hawkins, Eric Dolphy, Louis Armstrong, and Cannonball Adderley), with the remaining sixty-five (83.33%) having a lower mean BUR. The WJazzD BUR ranged from Red Garland ($\bar{x} = 1.37 \pm 0.34$) to Curtis Fuller ($\bar{x} = 2.10 \pm 0.60$). This indicated that Green tended to swing harder than most of the musicians in the WJazzD, based on their available data.

Figure 7.4 shows Green's BUR distribution in comparison to Metheny's, the guitarist with the most data in the WJazzD, and Parker's, who had a similar amount of data to Metheny. The figure shows that Green's distribution was substantially different from that of the other two performers, not only in the central tendencies of the data, but also the number of peaks in the data. An ANOVA found a significant difference between the performers and their BUR, with a small effect size ($F(2, 3988) = 48.83,\ p < .001,\ \eta^2 = .02$). Subsequent post-hoc tests using Tukey's HSD procedure found a significant pairwise difference between Green and both Metheny and Parker ($p < .001$), while no significant difference was found between Metheny and Parker ($p = .379$).

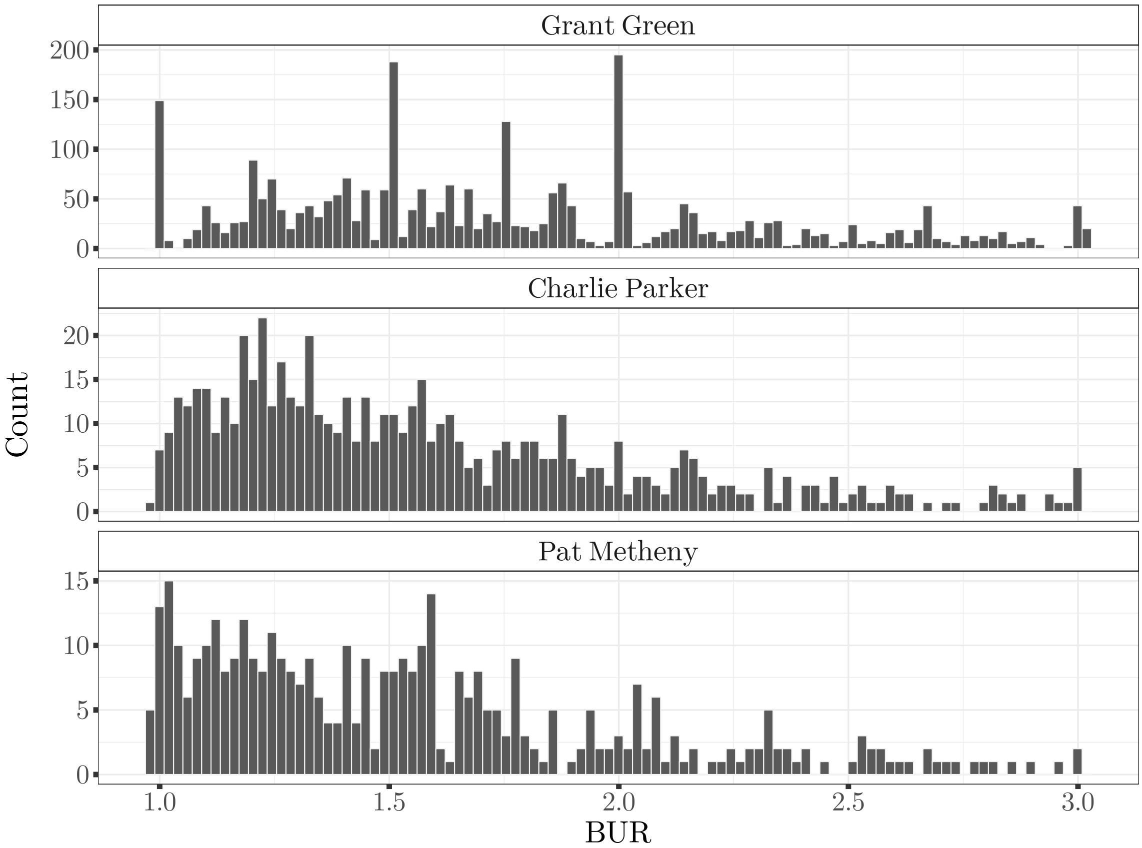

Figure 7.4: Distribution of BURs for Green, Parker, and Metheny.

Compared to Green's mean BUR of 1.75 ± 0.50, both Parker and Metheny's mean BUR were significantly lower. Parker had a mean BUR of 1.59 ± 0.46, while Metheny's was 1.55 ± 0.46. Both Parker and Metheny also had a median BUR of below 1.5 (Parker: 1.48; Metheny: 1.45), meaning that more than half of all their swung notes had a BUR between 0.98 and 1.5.

Even though there were some spikes in Metheny and Parker's data, there were not as many, or as extreme, as those in Green's. There were two possible explanations for this difference:

1) Green was highly consistent in his BUR, and played a small number of specific BURs;
2) There was – despite the best efforts of the author to match the transcription process of the Jazzomat Research Project – some difference in the transcription of the onset and offset of notes that resulted in a difference in the distribution of the BURs.

It is likely that both explanations played some part in the disparate distributions. This could have been investigated through re-transcribing some transcriptions from the WJazzD; however, this was outside the scope of this research. It was initially thought that part of the difference may have been due to the instrumentation.

Guitars have a different ADSR envelope compared to horn instruments, which comprised the vast majority of the WJazzD.[6] The guitar, with a short attack, has a clear onset time for the note, while also having a reasonably clear release with which to measure the offset. However, the same concentration of data was not seen in that of the four guitarists in the WJazzD, which suggested there may have been some other differences in the transcription process.[7] Despite the differences in the appearance of the distributions of Green's BUR to those within the WJazzD, Green's central tendencies fit within the ranges of the other data, and there was no indication that the data was incorrect. The following analyses investigated how the BUR changed depending on the tempo range and tonality mode.

7.1.2 BUR vs. Tempo Range

It is widely accepted that BUR decreases as the tempo increases, with this confirmed experimentally (e.g. Friberg and Sundström (1997), Friberg and Sundström (2002), Butterfield (2011), Corcoran and Frieler (2021)). It was hypothesised that the same changes would be observed in Green's improvisations, with higher BURs expected to be found in the lower, BPM ≤ 170 tempo range. Honing and Hass, in their investigation of BURs in jazz drummers, found that at "beat durations [greater than 350ms (BPM ≤ 170)] the swing ratio [seemed] to stabilise" (2008, 475).

Green's BUR for each tempo range can be seen in Figure 7.5, which showed a substantial difference in BUR distributions. An ANOVA was used to investigate the difference in Green's BUR between the two tempo ranges; a significant effect of the tempo range on the BUR was found, with a large effect size ($F(1, 2969) = 349.92, p < .001, \eta^2 = .11$). When improvising at tempos ≤ 170 BPM Green's mean BUR was 2.06 ± 0.58 – significantly higher than when he was improvising at tempos > 170 BPM where his mean BUR dropped to 1.66 ± 0.44. This data also indicated that Green's BUR was more consistent at tempos > 170 BPM (SD: 0.44) compared to tempos ≤ 170 BPM (SD: 0.58). Additionally, three-quarters of Green's BURs at tempos > 170 BPM had a value below the median BUR of the tempo range with a tempos ≤ 170 BPM. These results supported the preconceived wisdom, prior studies, and hypothesis that the tempo had a significant impact on the BUR of swung note pairs Green played. Specifically, the tempo had a limiting factor on how heavy the swing was at higher tempos.

[6] ADSR is Attack, Decay, Sustain, Release, and describes the amplitude of a note over time.

[7] There still could be an instrumentation difference, but as there was so little data for the four guitarists in the WJazzD, the difference may have been there but not identifiable.

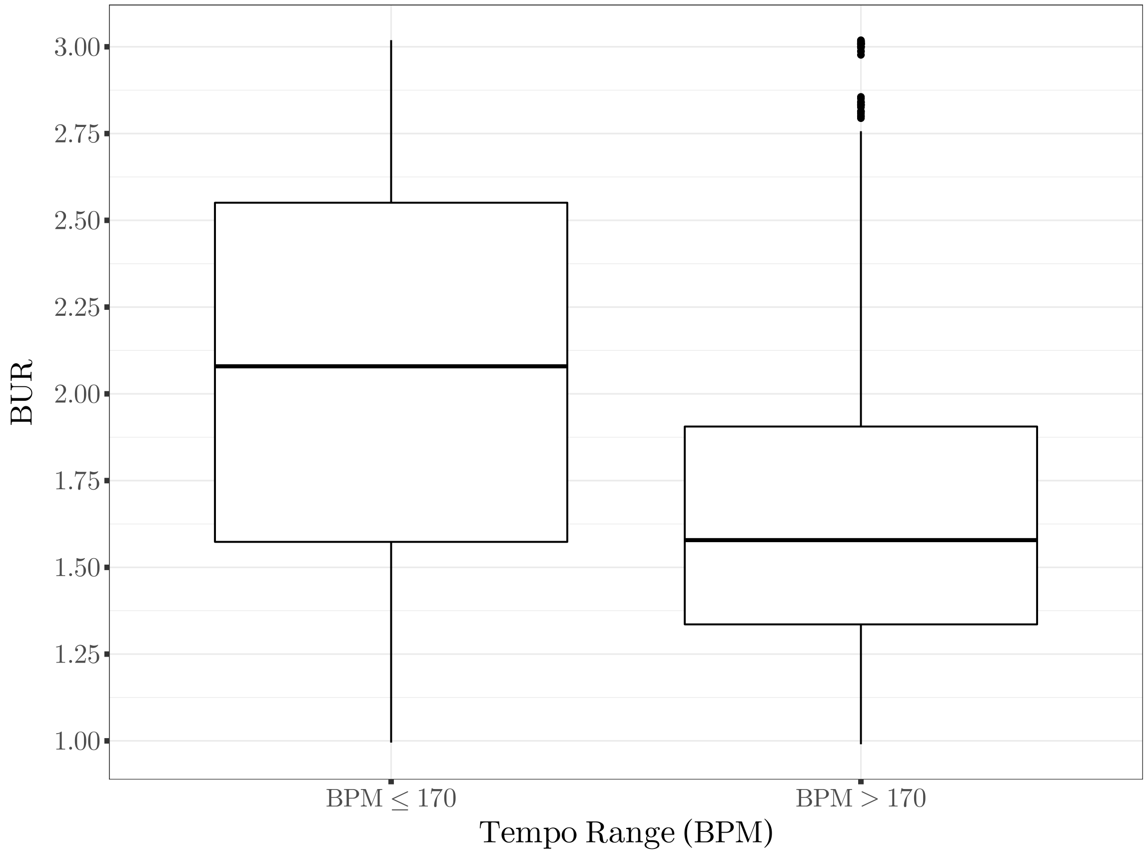

Figure 7.5: Distribution of Green's BUR for each of the binary tempo ranges.

Another factor the tempo may have had on Green's use of swing was the frequency of swung note pairs in each tempo range. There were a similar number of note events in each of the tempo ranges, around 10,000 note events each. However, there were around 1500 more beats at the higher tempos (6095 beats) compared to the lower tempo range (4552 beats). At tempos $\leq$ 170 BPM, 14.21% of the beats contained a pair of swung notes. In comparison, beats containing swung note pairs were more than twice as frequent at tempos $>$ 170 BPM, with 38.13% of beats containing swung notes. This indicated that although Green did not swing as heavily at higher tempos, he did play more swung note pairs. A χ^2-test found a significant relationship between the tempo range and the frequency of beats containing swung note pairs, with a small effect size ($\chi^2(1) = 739.69,\ p =< .001,$ $V = .26$). Taking into consideration the results of the prior analysis into rhythmic variety, it was likely that the fewer beats with swung notes at tempos $\leq$ 170 BPM was due to greater variety in Green's rhythmic units at slower tempos, playing fewer quaver-equivalent note lines with swung notes.

This analysis found the tempo had a significant impact on Green's use of swing, with Green swinging harder at lower tempos while playing swung notes with a lower and more consistent BUR at higher tempos. Likely due to the variety of rhythms played, Green also played more swung notes when improvising in the higher tempo range.

7.1.3 BUR vs. Tonality Mode

Corcoran and Frieler (2021) found that different styles of jazz could "be divided into 'soft' and 'hard' swinging" (379). They found that styles including swing, bebop, and cool had BURs with a range between 1.29 and 1.47, while hard bop and post-bop had BURs between 2.02 and 2.14 (Corcoran and Frieler 2021, 379). The relationship between blues and hard bop informed the hypothesis that Green, who often played in the hard bop style between 1960 and 1965, may have swung harder when improvising over a blues. Due to the large influence the tempo range had on Green's BUR, the investigation into the tonality mode split the data into the two ranges and investigated them separately. Figure 7.6 shows the proportion of beats for which a BUR was calculated for each tonality mode in both tempo ranges.

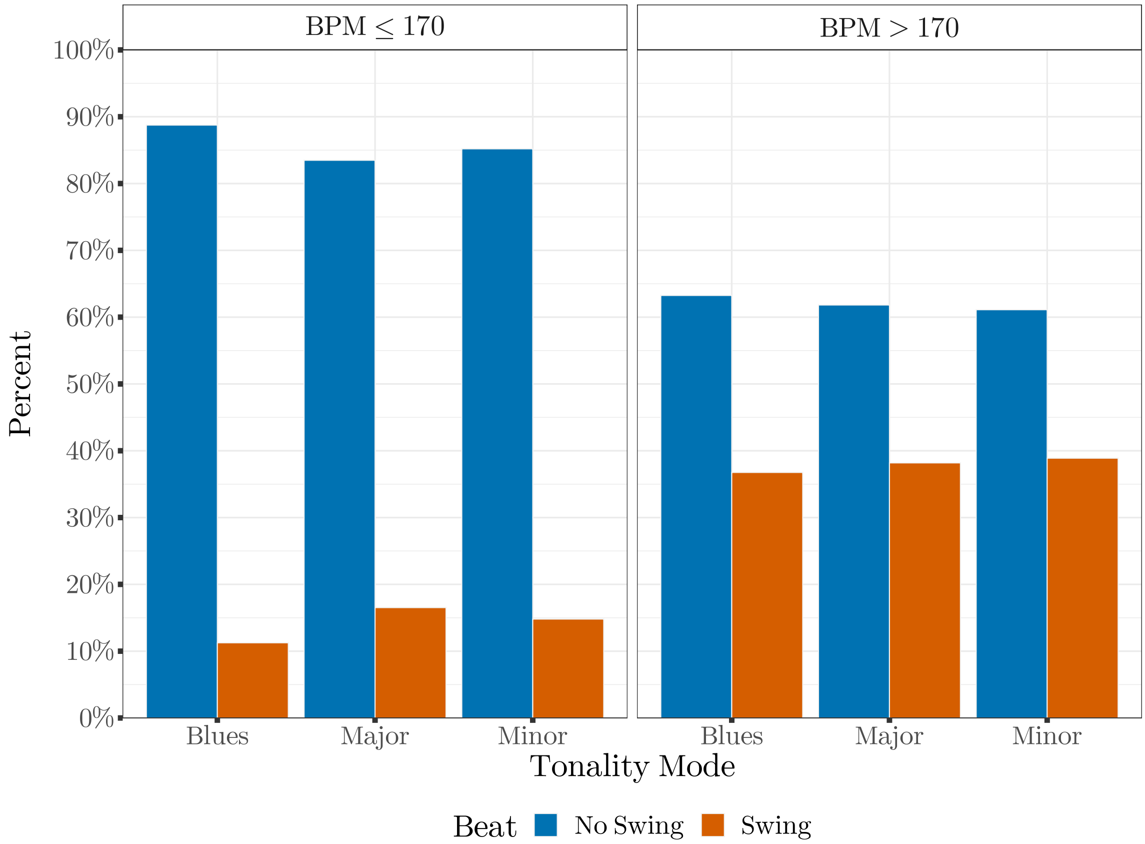

Figure 7.6: The proportion of beats for each tonality mode in each tempo range that were classified as swung beats for which a BUR was calculated.

The data showed that, in both tempo ranges, there was a smaller proportion of swung beats in Green's improvisations over a blues compared to the major or minor tonalities. For the lower tempo range, 11.25% of blues beats were swung, compared to 16.52% and 14.81% for major and minor tonalities. For the higher tempo range the proportions for each tonality mode were even closer, 36.75% for beats over a blues, while major and minor had 38.18% and 38.89% of beats swung respectively. A separate χ^2-tests for each tempo range found a statistically significant difference for BPM ≤ 170, with a very small effect size ($\chi^2(2) = 19.20,\ p =< .001,\ V = .06$). No significant difference was found between the tonality modes at BPM > 170 ($\chi^2(2) = 1.35,\ p = .510,\ V = .01$). This indicated that Green played slightly fewer swung notes over a blues in comparison to major or minor tonalities, in the lower tempo range.

The frequency of swung beats, while useful to understand how the tonality mode affected how often Green played swing pairs, did not investigate how hard Green swung over each tonality mode. Figure 7.7 shows the distribution of BURs for each tonality mode and tempo range. This figure showed that Green's median BUR was the highest at both tempo ranges for the blues. At BPM ≤ 170 the differences were smaller, with blues and minor tonalities having a similar distribution. When the BPM > 170, Green swung substantially harder over a blues when compared to a major or minor tonality.

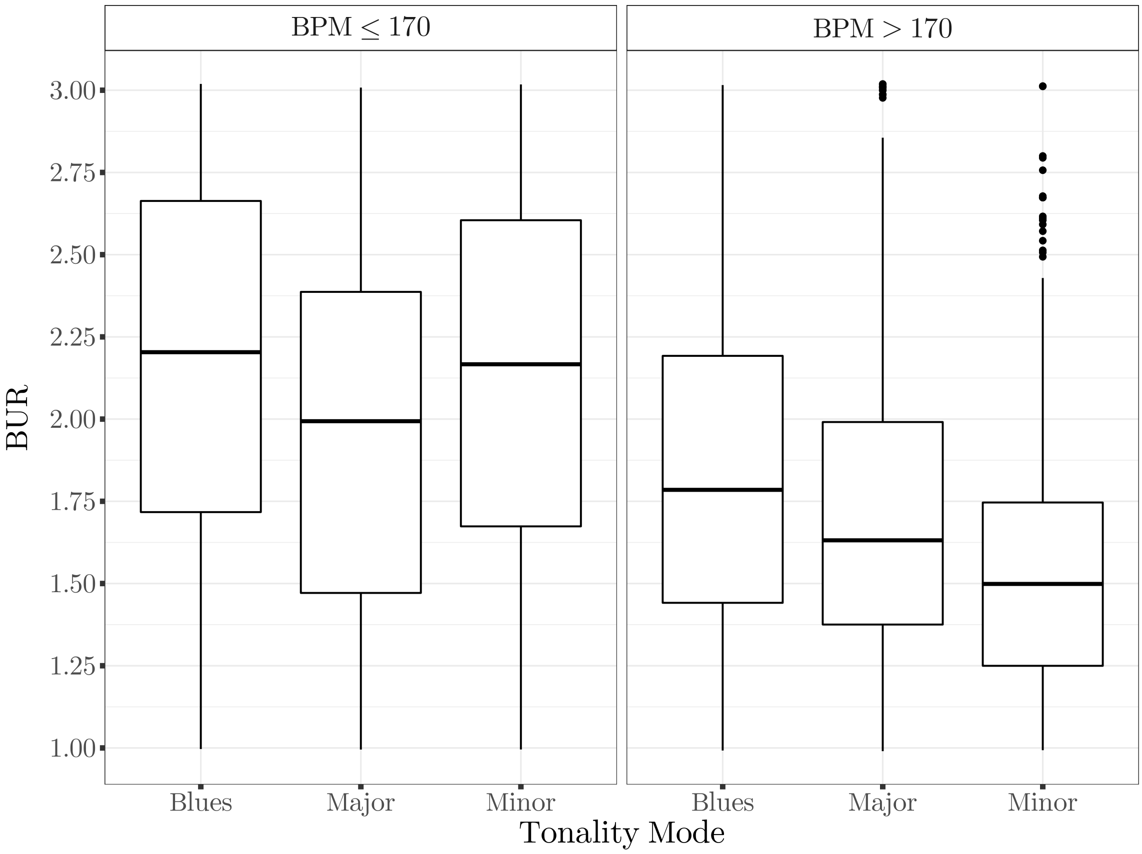

Figure 7.7: Distribution of Green's BUR for each tonality mode in both tempo ranges.

This data showed that Green swung hardest over the blues, compared to major or minor tonalities. When the tempo was ≤ 170 BPM Green's mean BUR over a blues was 2.16 ± 0.58, with the minor tonalities mean BUR also very similar 2.13 ± 0.58. The mean BUR for major tonalities was around 0.2 lower at 1.95 ± 0.56. When the tempo was > 170 BPM Green's BUR over a blues ($\bar{x} = 1.85 \pm 0.53$) was only slightly lower than that of the major tonality mode at BPM ≤ 170. This was followed by the major ($\bar{x} = 1.68 \pm 0.43$) and minor ($\bar{x} = 1.52 \pm 0.36$) tonality modes.

An ANOVA was run for each tempo range to investigate the relationship between the BUR and tonality mode. A significant effect of the tonality mode on the BUR was found for both tempo ranges. A small effect size was found for BPM ≤ 170 ($F(2, 644) = 9.13,\ p < .001;\ \eta^2 = .03$). Subsequent post-hoc tests with Tukey's HSD procedure found significant pairwise differences between the major and blues tonality ($p < .001$) and minor and major tonalities ($p = .003$), with no significant difference found between the Blues and Minor tonalities ($p = .854$). A medium effect size was found at BPM > 170 ($F(2, 2321) = 76.99,\ p < .001;\ \eta^2 = .06$). Subsequent post-hoc tests found significant pairwise differences between each tonality mode ($p < .001$).

These results supported the hypothesis that Green did have a tendency to swing harder when improvising over a blues. This was especially true when comparing the blues to the major tonality mode where it was significantly higher at both tempo ranges. Although there was no significant difference between the BUR for blues and minor tonalities when the BPM ≤ 170, the largest difference was observed between these two tonalities when the BPM > 170. This suggested that the stylistic differences Corcoran and Frieler (2021) found for post-bop and hard bop were accentuated by Green when improvising over a blues.

7.1.4 Swing Summary

The results of the analysis into Green's use of swing found that Green's BUR tended to be higher than many of the musicians in the WJazzD. Green's BUR data was also distributed substantially differently from the WJazzD, with spikes found around specific BURs of 1, 1.5, and 2. This result was likely a combination of Green's improvisational style and slight differences in transcription procedures between the WJazzD and this research. The tempo of an improvisation had a significant impact on Green's BUR. While there were more swung beats found at higher tempos, the BUR of the swung notes were lower. This supported the hypothesis, and prior studies, that an increasing tempo had a limiting effect on the BUR. The tonality mode was also found to have a significant impact on Green's swing, with swung pairs in Green's blues improvisations tending to have a higher BUR than those in major or minor tonalities. The BUR difference in tonality modes was less pronounced at the lower tempo range, where blues and minor tonality improvisations had comparable BURs. When improvising over a blues tonality, even at higher tempos, Green's BUR was close to a heavy swing of 2:1, with a mean BUR of 1.85. This analysis found that Green tended to swing harder than many of the performers in the WJazzD, with his BUR significantly affected by the tempo range and tonality mode.

7.2 Micro Timings

This micro timings section focused on two types of micro timing in Green's improvisations, articulation and micro-gaps between notes. Articulation, in the *MeloSpy* system, was defined as the ratio between the duration and IOI of a note (articulation = duration/IOI). The micro-gaps between notes was the complement to the rests section in the previous chapter, investigating all gaps with a $\text{rrest}_{\text{prop}}$ of ≤ 0.3 beats. There was overlap between these features, and between articulation and rests, as a staccato note or a note followed by a rest or micro-gap could be transcribed with the same onsets and offsets.

7.2.1 Articulation

Articulation measured how long a note's duration was compared to the length between notes (IOI). The previous chapter found that the inter-phrase IOI tended to be substantially longer than the intra-phrase IOI. Therefore, this articulation analysis focused on intra-phrase notes. There are no widely accepted standards for what constitutes a note being played with a normal articulation, compared to staccato or legato. Consequently, cut-off values were created, with the following ranges used within this research:

- short: articulation < 0.6;
- normal: $0.6 \leq$ articulation < 0.9;
- long: $0.9 \leq$ articulation.

The use of the terms short, normal, and long articulation were preferred over staccato and legato so as not to cause confusion with the musical terms, and the intent they imply.[8] This research used both the numerical descriptor and the three class categorical variable throughout the analysis of Green's articulation.

The distribution of Green's articulation can be seen in Figure 7.8. The data in this graph showed that the plurality (27.60%) of notes played by Green had an articulation ≥ 0.99, with 49.58% of all notes having an articulation ≥ 0.80. Green's mean articulation was 0.78 ± 0.20 (skewness: -0.72; kurtosis: -0.13). Regarding the articulation classes, 48.77% of the notes Green played had a normal articulation,

[8] The measured articulation calculates the values based on the duration and IOI, but can say nothing about the intent of Green's playing. For example, a note with a short articulation could either be a staccato note or a note followed by a rest.

33.71% had a long articulation, and 17.52% were played with a short articulation. This showed that Green mostly played his notes with normal to long articulation. There were likely many features that had an impact on Green's articulation; however, this research investigated the influence of the tempo range and the size of the fuzzy interval following the note.

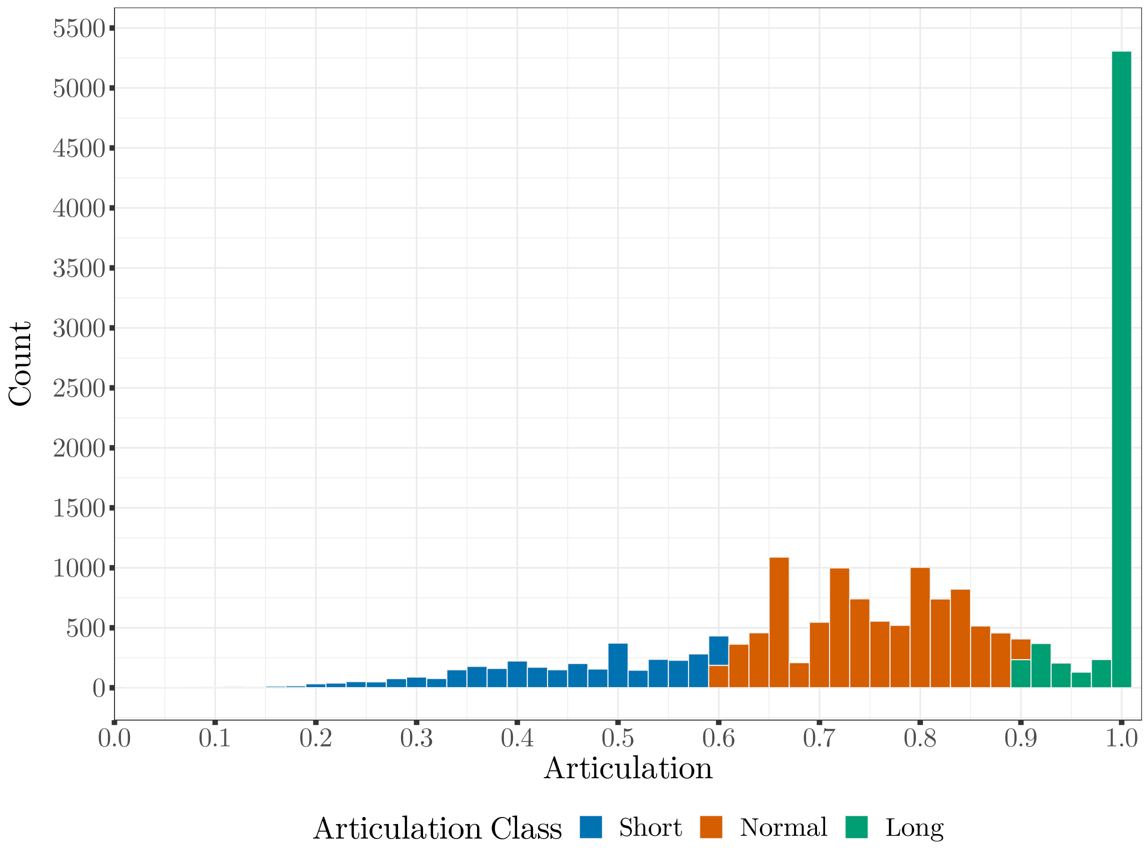

Figure 7.8: Distribution of note articulations in Green's corpus.

Articulation vs. Tempo Range

It was hypothesised that Green was likely to play more notes with a long articulation when improvising at tempos > 170 BPM. This hypothesis was based on the assumption that at higher tempos there was less rhythmic variety in Green's playing, suggesting more continuous quaver note lines with long articulation.

Figure 7.9 shows the proportion of notes in each articulation class from both tempo ranges. The data in this graph showed evidence of changes in articulation contrary to the stated hypothesis. Instead of the proportion of notes with a long articulation increasing when the tempo was > 170 BPM, it decreased by 8.60 PP. The majority of this difference was from Green playing a higher proportion of notes with a short articulation at higher tempos (20.44% vs. 14.68%). Regardless of the tempo range, approximately half of all the notes Green played had normal articulation, with there being marginally more normal notes at higher tempos (50.21% vs. 47.37%).

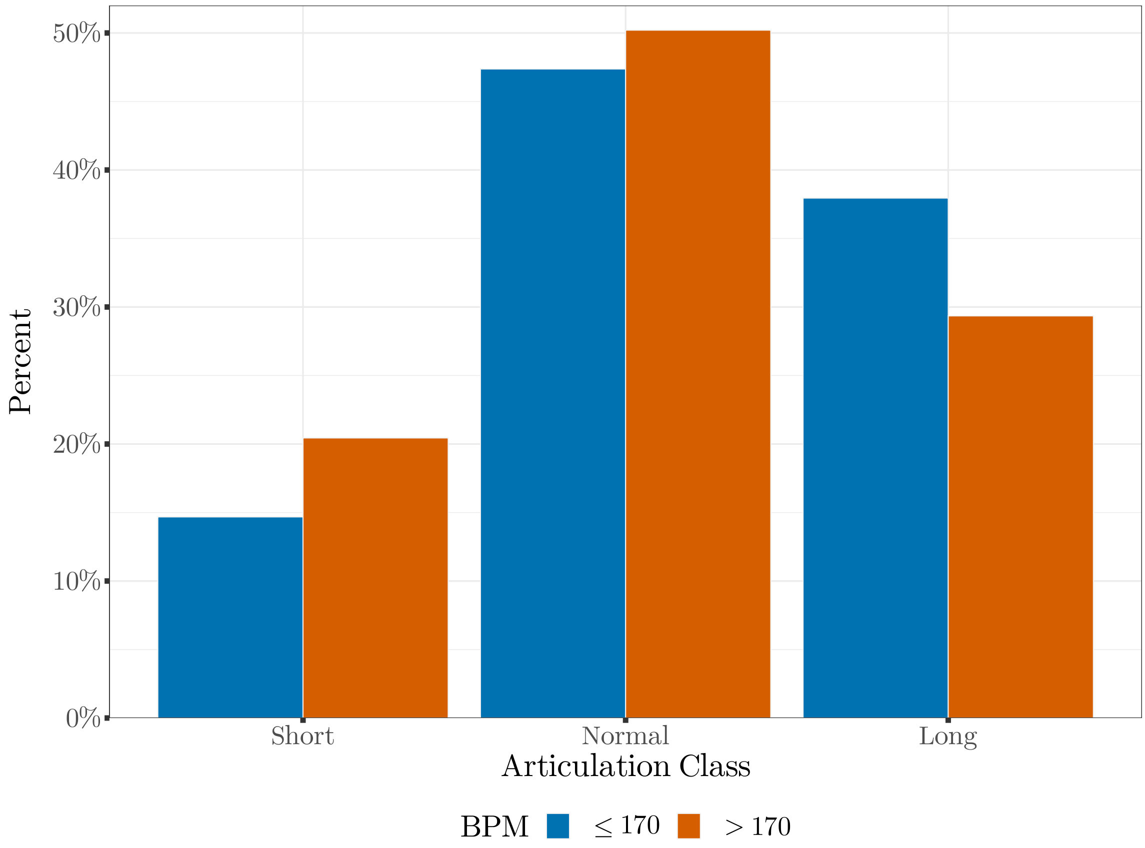

Figure 7.9: Proportion of notes in each articulation class played by Green in both tempo ranges

A χ^2-test found a significant difference in the number of notes played within each articulation class between the two tempo ranges, with a small effect size ($\chi^2(2) = 204.42,\ p =< .001,\ V = .10$). Figures 7.10 and 7.11 show two musical excerpts that exemplified the two ends of the articulation spectrum.[9] The first, Figure 7.10, from Green's improvisation over *Round About Midnight* (Green 1961r), shows a phrase in the lower tempo range (BPM = 120) where the majority of notes had a long articulation, with a median articulation of 0.996. In contrast, the second musical example in Figure 7.11, from Green's improvisation over *The Song Is You* (Green 1962l), shows a higher tempo (BPM = 240) phrase where the majority of notes had a short articulation. This phrase's median articulation was 0.523, nearly half of that of the previous phrase.

[9]The musical examples were edited to more accurately reflect the true articulation of the notes.

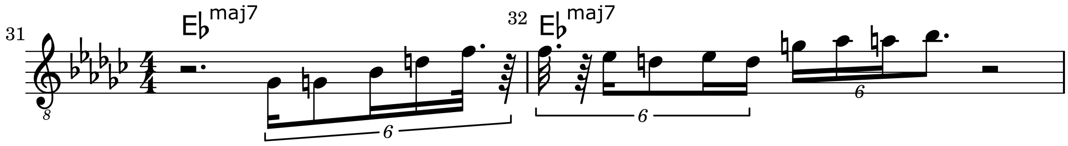

Figure 7.10: Example of a phrase at a lower tempo (BPM = 120) with most notes played with long articulation, *Round About Midnight* (1961), bars 31–32.

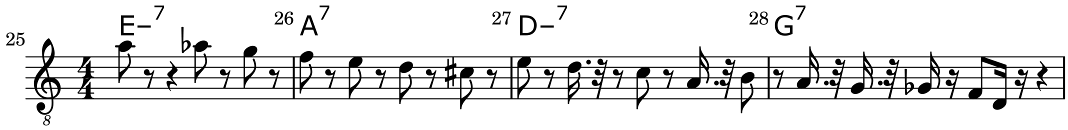

Figure 7.11: Example of a phrase at higher tempo (BPM = 240) with most notes played with short articulation, *The Song Is You* (1962), bars 25–28.

In summary, this analysis found results contrary to those hypothesised, with notes played with long articulation more frequent when the tempo was $\leq$ 170 BPM. Notes played with a short articulation occurred more often when the tempo was $>$ 170 BPM. In both tempo ranges, notes played with a normal articulation were most common.

Articulation vs. Fuzzy Interval Size

The analysis into the relationship between the articulation of notes and the fuzzy interval played after the note investigated the hypothesis that a larger interval would lower the articulation value.[10] As it was only the magnitude of the interval that was of interest the ascending and descending fuzzy interval classes were combined together to form an absolute fuzzy interval class. Figure 7.12 shows the distribution of articulation values for each of the absolute fuzzy interval classes in Green's improvisations. This showed that Green played small intervals (steps and leaps) with longer articulation than larger intervals or repeated notes. The very high articulation values for intervals of the step class could be attributed to notes played with slides, bends, hammer-ons, or pull-offs; techniques that are nearly always played with full articulation. In contrast to the small intervals, Green's articulation of repeated notes was most similar to that of the larger intervals.

[10]Unlike the interval analyses in the pitch domain, which used an expanded version of the fuzzy interval classes, this analysis used the default fuzzy interval classes from *MeloSpy* to ensure there was enough data in each of the classes. This was due to the articulation analyses excluding notes that occurred at the end of phrases, where the largest intervals were most likely to occur.

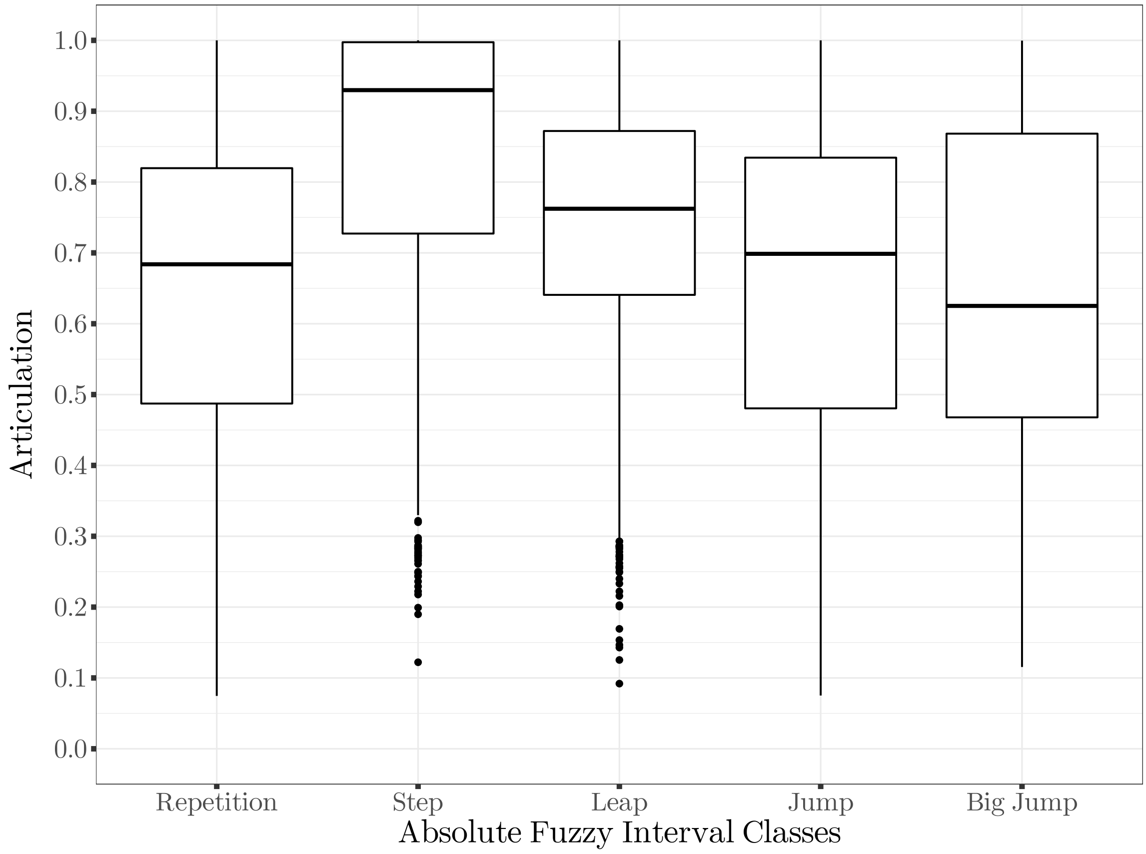

Figure 7.12: Distribution of articulation values for each absolute fuzzy interval class in Green's corpus.

An ANOVA was run to investigate the relationship between the fuzzy interval classes and articulation; finding a significant difference, with a medium effect size ($F(4, 19222) = 544.95,\ p < .001;\ \eta^2 = .10$). Subsequent post-hot tests with Tukey's HSD procedure found significant pairwise differences at $p < .001$ for all comparisons that included Step or Leap classes, with the other comparisons not found to be significant.[11] This indicated that Green's articulation of steps and leaps were found to be significantly different from all other interval classes, while the others were not found to be significantly different. These results supported the hypothesis, with larger intervals found to have significantly lower articulations than smaller intervals. However, repeated notes were also found to have lower articulations on average.

The lower articulation for repeated notes suggested that when Green repeated notes there was a variety of rhythmic and articulated components to the notes. This suggested that Green used varying articulation, instead of pitch, to add distinguishing features and interest to his improvisations. An example of this can be seen in Figure 7.13, which shows an excerpt from Green's improvisation over *Freedom March* (Solo 1, Green 1961d), where he played six E3 notes in a row. The

[11] $p < .001$: step vs. repetition, step vs. leap, step vs. jump, step vs. big jump, leap vs. repetition, leap vs. jump, leap vs. big jump. Repetition vs. jump ($p = .123$), repetition vs. big jump ($p = .998$), jump vs. big jump ($p = .746$).

figure displays two representations of the excerpt, the top is the automatically generated symbolic notation. The bottom is a screenshot from the *Sonic Visualiser* transcription file, with the notes aligned as close was possible. This combined view showed that Green played a variety of articulations to generate interest and complexity over the repeated note sequence.

Figure 7.13: Example of Green varying the articulation of notes in a repeated sequence, *Freedom March* (Solo 1, 1961), bar 55. Top: Symbolic notation. Bottom: *Sonic Visualiser* transcription.

In summary, Green tended to articulate steps, the smallest non-repeated interval size, the longest, with a median value within the long articulation class. Fuzzy intervals of a leap were found to be played with a significantly longer articulation than repeated notes, jumps, and big jumps, but shorter than steps. The largest intervals in Green's improvisations had the shortest articulation, along with repeated notes. The shorter articulation for repeated notes suggested that Green substituted pitch changes for articulation variety.

Articulation Summary

This analysis into Green's articulation found that the plurality of notes he played had an articulation > 0.99, while approximately half had an articulation ≥ 0.80. In terms of articulation classes, the vast majority of notes Green played had a normal or long articulation. The tempo had a significant impact on Green's articulation, although the effect size was small. Counter to the hypothesis, the analysis found that Green's articulation decreased at higher tempos, with Green playing more notes with a short articulation. Regardless of the tempo, around half of the notes Green played had a normal articulation. The size of the fuzzy interval following a note also had a significant impact on the articulation. Notes that moved by a step had the highest articulation values, with a median articulation within the long articulation class. Notes that moved by a leap were found to have the second longest articulation on average, followed by repetitions, jumps, and big jumps. While repetitions, jumps, and big jumps were not found to have a significantly different articulation from each other, all three were found to have a significantly lower articulations than steps or leaps. The very high articulation for notes that moved by a step may have been influenced by slurs played by Green, which are associated with an articulation of 1.[12]

[12] Although slurs could be used to transition between notes larger than a 2nd, they were frequently used for minor or major 2nd transitions.

In contrast, the trend of lower articulation for repeated notes was likely influenced by Green deliberately varying the articulation as a substitution for pitch variation. In summary, although there were instances of Green playing notes with shorter articulations, especially at higher tempos and for repeated notes or between large interval gaps, the vast majority of notes played by Green had a normal to long articulation.

7.2.2 Micro-Gaps Between Notes

Micro-gaps between notes (micro-gaps) was the complement to the rest feature, considering all gaps between notes that had a $\text{rest}_{\text{prop}} \leq 0.3$ beats. In Green's data, notes were far more likely to be followed by a micro-gap than a rest, with 82.31% (16 855) of the notes Green played followed by a micro-gap. Figure 7.14 shows the distribution of micro-gaps in Green's improvisations. The data was split into two graphs to better show the full distribution. The graph on the left shows a zoom of the range 0–0.01 beats, as there was a high concentration of notes with micro-gaps very close to 0 beats. The graph on the right shows the remaining distribution between 0.01 and 0.3 beats. Although there was a substantial concentration of micro-gaps ≤ 0.01 beats, the majority (67.36%) of non-rest notes had a micro-gap between 0.01 and 0.3 beats.

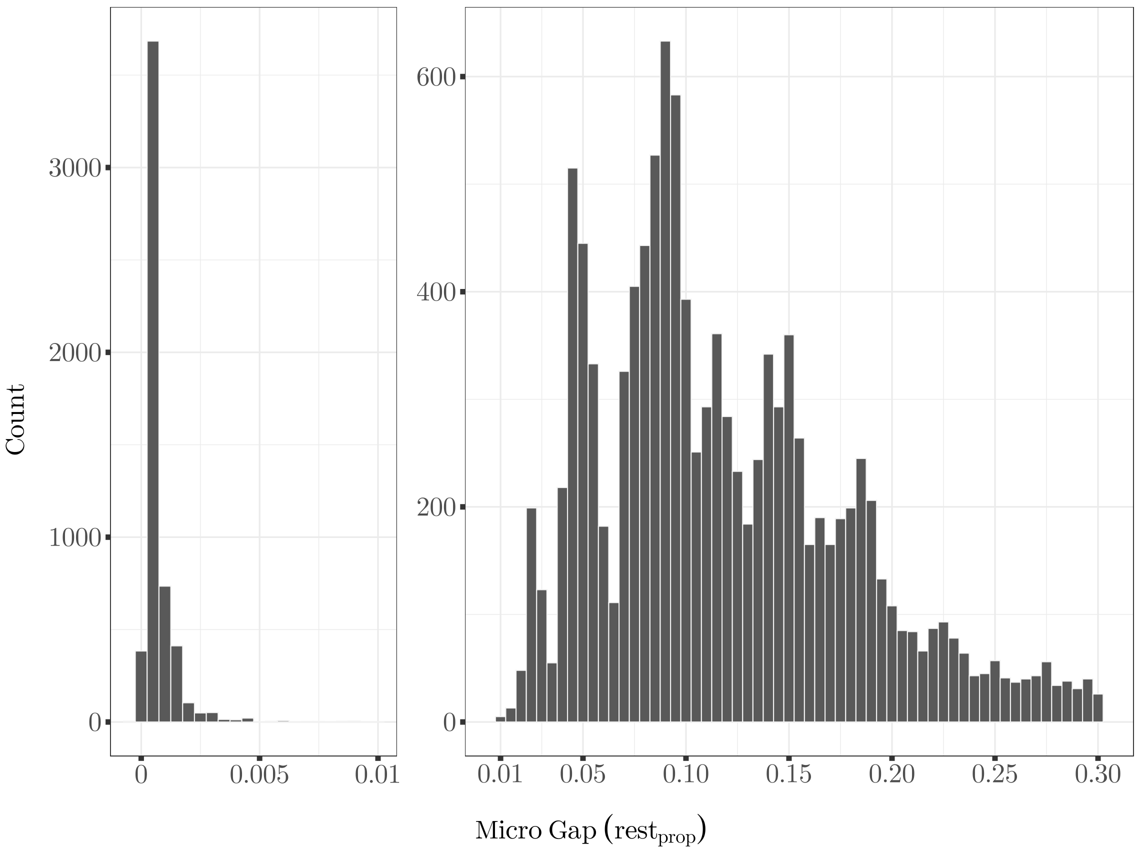

Figure 7.14: Distribution of micro-gaps between notes in Green's corpus. Left: 0–0.01 beats. Right: 0.01–0.3 beats.

The articulation of the two micro-gap ranges provided insight into their effect in Green's improvisations. The shortest micro-gaps (≤ 0.01 beats) had a mean articulation of 1.00 ± 0.004, while the longer micro-gaps had a mean articulation of 0.75 ± 0.12. These results indicated that the shortest micro-gaps were most likely artefacts from the transcription process and not true gaps between the notes. Consequently, notes with a micro-gap proportion of ≤ 0.01 beats were excluded from the analyses.

The graph on the right of Figure 7.14, which displays the valid micro-gaps, shows a distribution of micro-gaps centred around 0.1 beats ($\bar{x} = 0.12 \pm 0.06$ beats), with some noticeable spikes around 0.05, 0.10, 0.115, 0.15, and 0.1875 beats. While some of these aligned with standard rest values (e.g. 0.1875 is equal to the nominal duration of a dotted demisemiquaver rest), it was most likely that the uneven distribution of micro-gaps was related to how Green improvised, the transcription process, and the rhythms of the notes he was playing. Figure 7.15 plots the micro-gap proportion against the number of notes per beat (for the four most frequent beat densities), with the beat density acting as a proxy for the rhythms Green played.

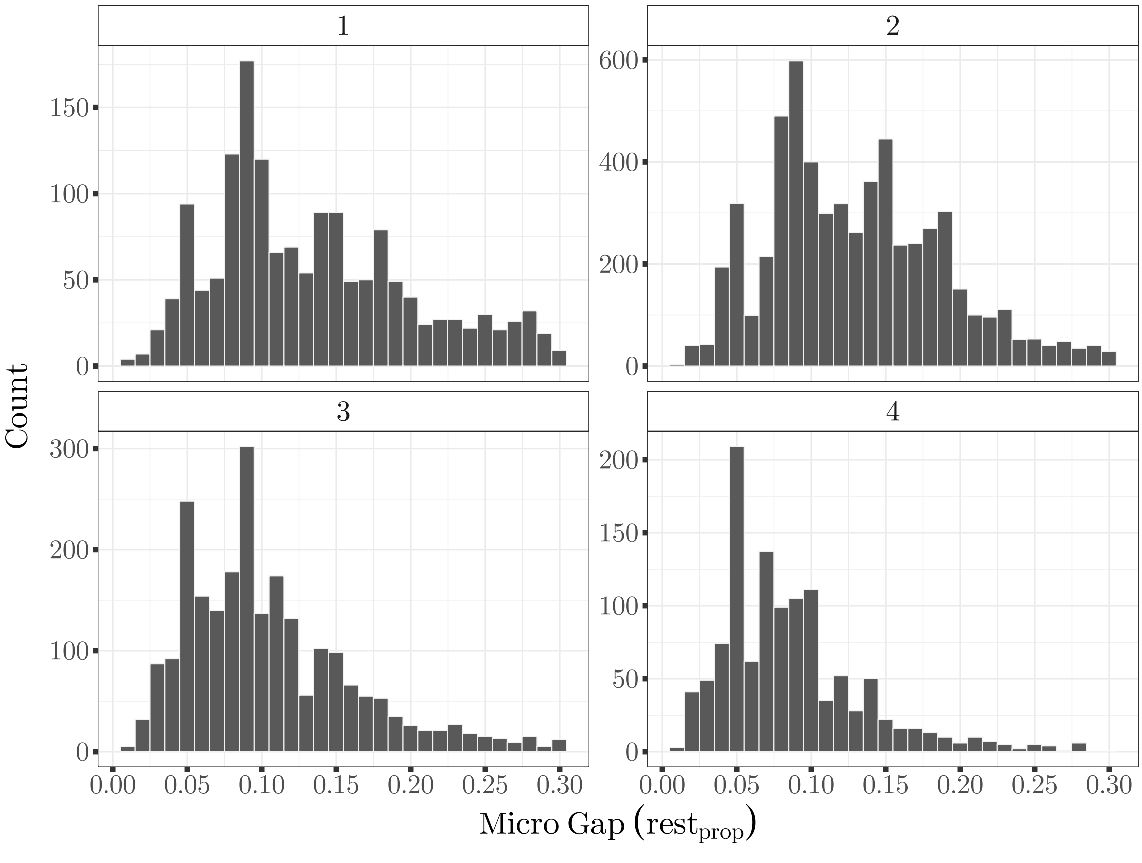

Figure 7.15: Distribution of micro-gaps for the four most frequent beat densities (notes per beat) in Green's corpus.

The graphs showed that as the beat density increased, the micro-gaps shortened.[13] Although this was expected – a beat with more notes would have fewer or shorter micro-gaps – there were similar spikes of data across the beat distributions. Specifically, the spike around 0.10 beats was frequent for all beat densities except four notes per beat. This suggested that although the beat density did naturally have a limiting effect on the micro-gaps, it did not appear to have an impact on the clustering of micro-gaps. Therefore, the clustering was most likely a combination of an aspect of Green's improvisations and an artefact of the transcription process.

The phrase position was the variable that had the largest impact on the duration of the rests Green played; however, this was not expected to be the case for micro-gaps. The importance of phrase position for $\text{rest}_{\text{prop}}$ was due to notes at the end of a phrase having a significantly longer rest than other notes. In contrast, only thirty-six (0.32%) micro-gaps occurred at the end of a phrase. Nearly all (94.44%) micro-gaps occurred in the middle of a phrase. A feature that was found to have a significant, if small, effect on Green's rests and was also expected to effect the length of his micro-gaps, was the tempo range of the improvisation.

Figure 7.16 shows the distribution of the micro-gaps for both of the binary tempo ranges. There were approximately an equal number of micro-gaps in each tempo range (BPM $\leq$ 170: 5796; BPM $>$ 170: 5558), indicating that the tempo range did not affect the frequency at which micro-gaps appeared in Green's improvisations. The data in the graph showed that Green's micro-gaps tended to be longer in the higher tempo range, with a mean micro-gap of 0.14 ± 0.06 beats. In comparison, Green's mean micro-gap when improvising at lower tempos was 0.10 ± 0.06 beats. An ANOVA was run to investigate the relationship between the tempo range and Green's micro-gaps, with a significant difference found between the two tempo ranges, with a medium effect size ($F(1, 11\ 352) = 1139.38,\ p < .001;\ \eta^2 = .09$).

The difference in mean micro-gap lengths between the tempo ranges was likely related to Green playing notes with shorter articulations at higher tempos. These results suggested that, in addition to articulation, Green may have used micro-gaps to add rhythmic punctuation. This was especially true at higher tempos, where his rhythmic variety was more limited.

[13] An ANOVA found a significant relationship between the beat density and micro-gap proportion with a medium effect size ($F(3, 10944) = 266.94,\ p < .001;\ \eta^2 = .07$). Significant differences were found between all pairwise comparisons at $p < .001$, except for beat densities of 2 vs. 1, where no significant difference was found ($p = .117$).

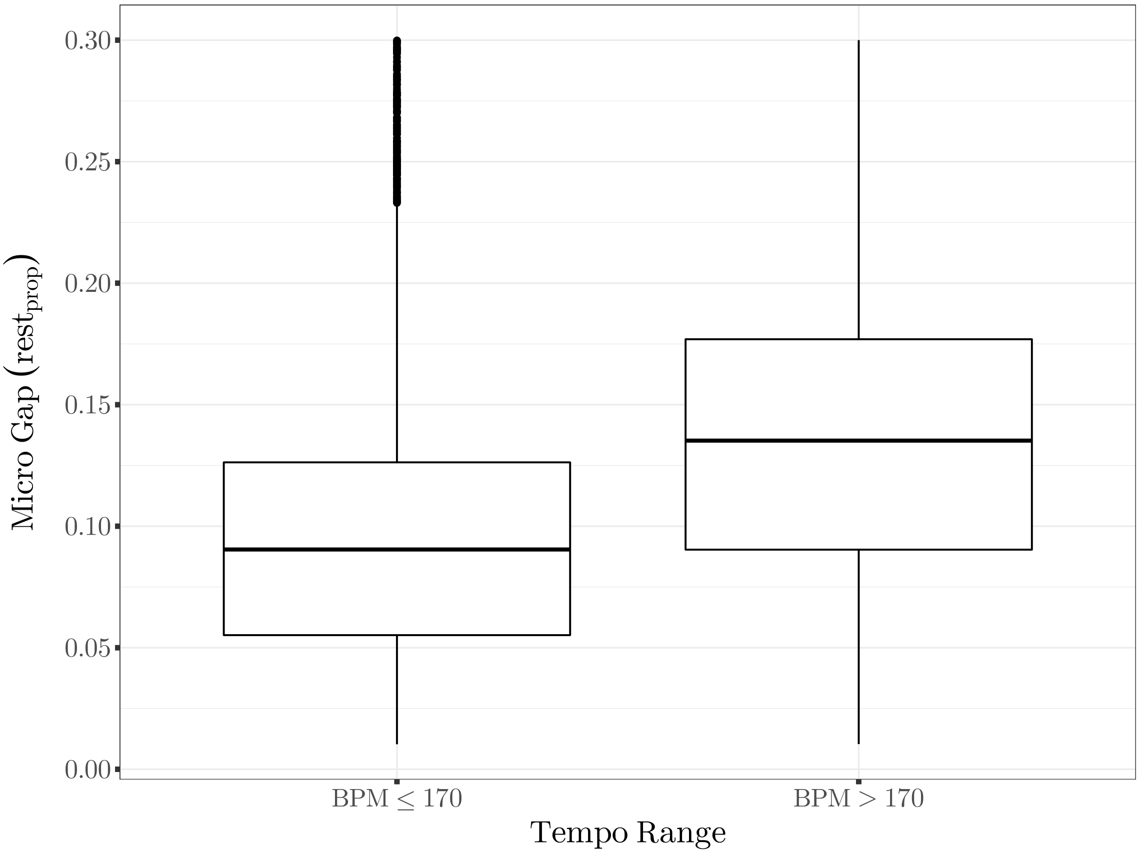

Figure 7.16: Distribution of micro-gaps for each of the binary tempo ranges in Green's corpus.

7.2.3 Micro Timings Summary

The analysis into the micro-gaps between notes did not provide great insight into Green's improvisational style. While it did find that micro-gaps were very common in Green's improvisations, with a relatively short mean length, it also found that they largely followed expected tendencies. As micro-gaps were, definitionally, very short, it was difficult to ascribe meaning and intentionality from Green's perspective. The micro-gaps results were most useful in combination with those of the articulation, supporting evidence of longer micro-gaps at higher tempos. This may have been a response to more restricted rhythmic variety at higher tempos, resulting in greater variation in short articulations and therefore longer micro-gaps. The concept of varying aspects of a note when other limitations existed was also evident in Green's articulation of repeated notes. When a note was repeated, limiting the variation from the pitch domain, Green was more likely to play his notes with a greater variety of articulations for added interest. The majority of all notes Green played had a normal articulation. This agreed with the data from the micro-gaps analysis where the majority of notes were not separated by a particularly short or long micro-gap. This analysis into micro timings found that in general the articulation of a note was a more insightful feature than the length of the micro-gaps between the notes.

7.3 Note Placement

The final feature to be analysed in the micro domain was the placement of notes. The specific investigations related to whether Green tended to play his notes ahead of or behind their nominal metrical onset. The feature that measured this was the difference to the nominal metrical onset. The nominal metrical onset was the optimal calculated position of a note based on the division and tatum from the FlexQ algorithm. There was little documentation provided about how the nominal metrical onset was calculated. Therefore, to be confident of the analyses regarding the nominal metrical onset, the author re-calculated the values. The Equations (7.1) and (7.2) successfully recreated the feature.

The equation required the beat onset, division, and tatum of the note to calculate the nominal metrical onset. Let the current beat onset time be b_0 and the onset time for the next beat be b_1, the length of the current beat was given by: $b_l = b_1 - b_0$. Let the division of the current beat be d, and the tatum of the current note within the beat be t_n, where $1 \leq n \leq d$. The nominal metrical onset in seconds from the beginning of the transcription, o_n, was then given by:

$$o_n = b_0 + ((\frac{b_l}{d}) \times (t_n - 1)) \tag{7.1}$$

The difference to the nominal metrical onset in seconds, o_d, was then given by the difference from the actual onset time of the note in seconds from the beginning of the transcription, o_a, to the nominal metrical onset time:

$$o_d = o_a - o_n \tag{7.2}$$

The difference to the nominal metrical onset could then be divided by the surrounding beat length, to get the difference from the nominal metrical onset as a proportion of the beat (onset difference). This scaled features allowed for easier comparisons across the various tempos. A note with a positive onset difference was played behind (after) the beat, while a note with a negative onset difference was played ahead of (before) the beat. The distribution of onset differences in Green's transcriptions can be found in Figure 7.17, with the graph on the left showing the raw difference in seconds while the graph on the right shows the difference as a proportion of the surrounding beat length.[14] The centre line (onset difference = 0) was highlighted to show the nominal note placement.

[14] There was a single outlier, a note with a difference from the nominal onset of 0.88 seconds, that was removed from the data.

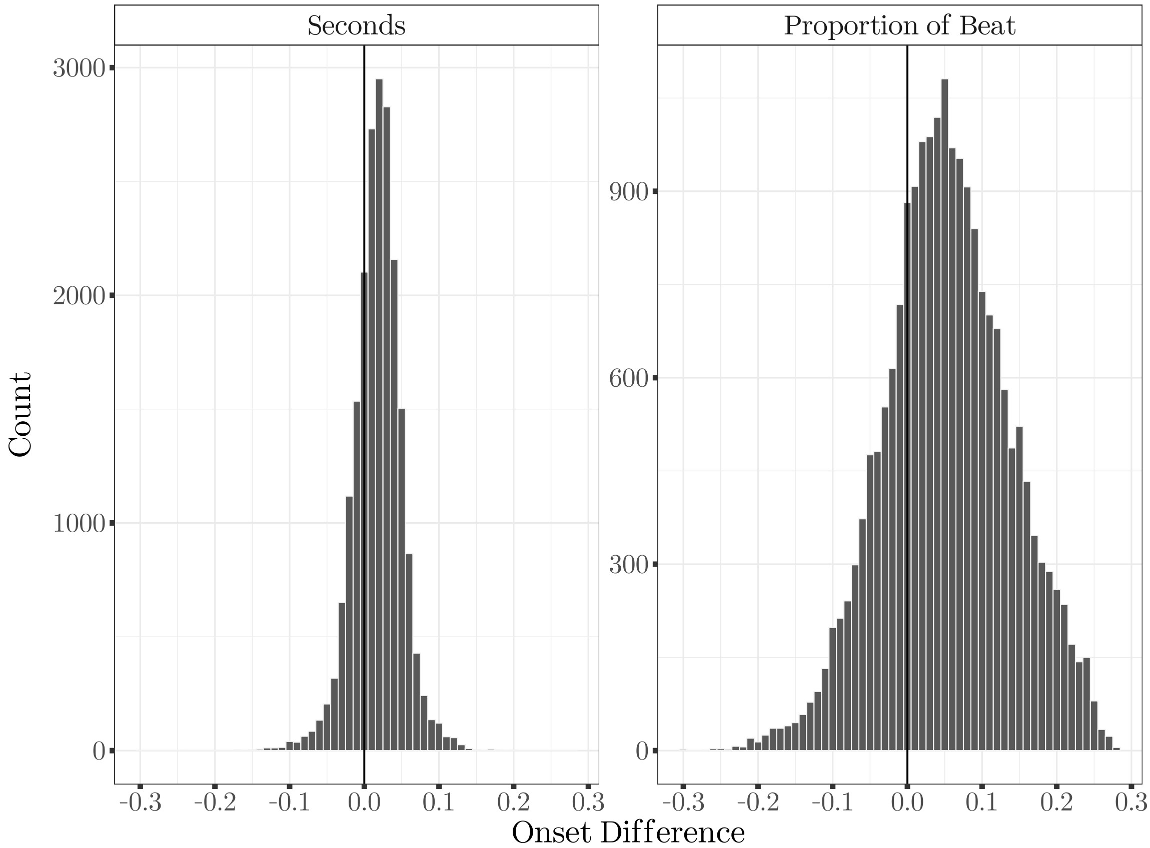

Figure 7.17: Distribution of note onset differences in Green's corpus. Left: raw difference in seconds. Right: difference as proportion of surrounding beat length.

These graphs showed that Green's onset differences followed a mostly normal distribution (seconds: skewness -0.25, kurtosis 2.47; proportion: skewness -0.06, kurtosis 6.13×10^{-4}), but shifted to the right, indicating that Green tended to play behind the beat (seconds: $\bar{x} = 0.02 \pm 0.03$; proportion: $\bar{x} = 0.05 \pm 0.08$). These results suggested that Green played behind the beat by around 5% on average. Categorising the notes into whether they were played before, on, or after the beat, showed that the vast majority of notes were played after the beat, with very few played on the beat, as seen in Table 7.2.[15] These results agreed with Scott's description of Green's playing as being "behind-the-beat" (2006, 1), with the data showing Green had a strong preference for playing behind the beat. This also agreed with a concept that while "older jazz styles had [note] placement ... ahead of the beat, ... traditional swing from the thirties [had it] on the beat, ... bebop and modern jazz [had] the placement ... behind the beat" (Solstad 2015, 103-104).

[15]The "on beat" category included all notes that had a proportional onset difference of 0 ± 0.001 beats, or 0.1%. This value was based on inspection of the *Sonic Visualiser* transcription files. While there were only fifty-eight notes with an onset difference of exactly 0, the small buffer increased this slightly to 235 note events.

Table 7.2: Distribution of note placement categorical features in Green's corpus.

	Ahead	On	Behind
Count	5096	235	15 146
Percent	24.89%	1.15%	73.97%

The following sub-sections investigated how Green's onset difference interacted with a variety of features, including: the tempo range; the division of the beat the note was played in; whether the note was played on or off the beat; and swing.

7.3.1 Note Placement vs. Tempo Range

It was hypothesised that the tempo range would have a similar limiting effect on the onset difference as it did on the swing BUR. Specifically, it was hypothesised that the the mean onset difference would decrease at the higher tempos, as Green aligned closer to the nominal note placements. Figure 7.18 shows the distribution of onset differences for each of the binary tempo ranges. This data showed results that were counter to the stated hypothesis. Instead of the onset difference decreasing at higher tempos it increased from a mean of 0.03 ± 0.076 beats at tempos ≤ 170 BPM to a mean of 0.08 ± 0.086 beats at tempo > 170 BPM. An ANOVA found a significant difference in the onset differences between the two tempo ranges, with a medium effect size ($F(1, 20475) = 1434.17,\ p < .001;\ \eta^2 = .07$).

A possible explanation was that Green played more notes ahead of the beat at lower tempos, as part of his greater rhythmic variability, lowering the overall mean onset difference. This was supported by the data, with 30.63% of notes played before the beat at lower tempos, in comparison to 19.01% of notes at higher tempos. However, investigation into the absolute values of the onset differences still found that Green played further away from the nominal placement at higher tempos compared to lower tempos. At tempos ≤ 170 BPM Green's mean absolute onset difference was 0.064 ± 0.053 beats. In comparison, at tempos > 170 BPM, Green's mean absolute onset difference was nearly 10% of the surrounding beat length, 0.096 ± 0.062 beats.

These results indicated that the tempo range that Green was improvising over did have a significant effect on his placement of notes in respect to their nominal onset. In opposition to the hypothesis, this analysis found that Green played further behind the beat at higher tempos. Although the majority of all notes were played behind the beat in both tempo ranges, Green did play a higher proportion of notes ahead of the beat at lower tempos.

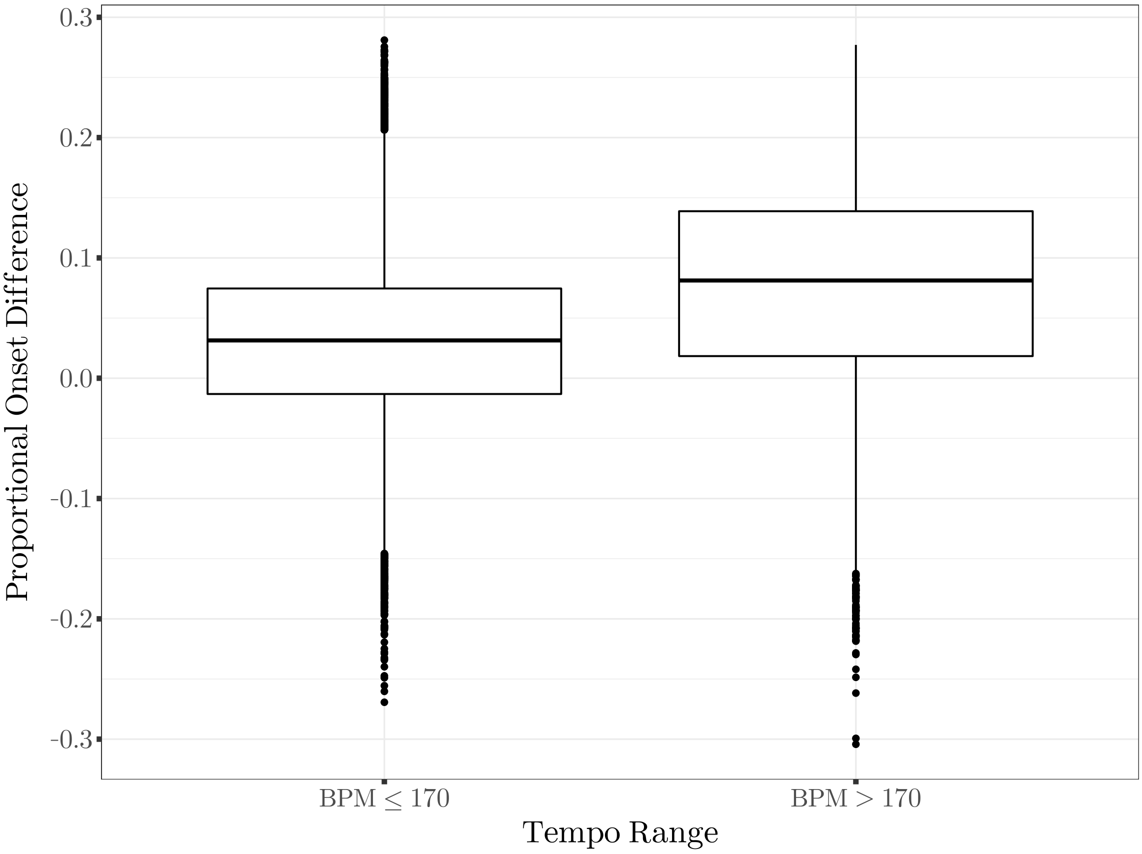

Figure 7.18: Distribution of proportional onset differences for BPM $\leq$ 170 and BPM $>$ 170 in Green's corpus.

7.3.2 Note Placement vs. Division

This note placement and division analysis focused on the five most frequent divisions (1, 2, 3, 4, 6). Each of these divisions had more than 1000 note events associated with it, and contained 95.44% of Green's data. The division was used here as a proxy for beat density and complexity, with higher divisions indicating beats that were denser or required more specific sub-beat placements. It was hypothesised that due to structural limitations of higher divisions, Green's onset difference would tend towards the nominal onset as the metrical density increased.

The distribution of Green's onset differences for each of the divisions can be seen in Figure 7.19. The graph showed that while notes played in a beat with division one were slightly closer to their nominal position than those in a beat with division two, the higher divisions had a lower onset difference. Green's mean onset difference was very close to zero for notes in beats with division four ($\bar{x} = 0.009 \pm 0.066$ beats) and six ($\bar{x} = -0.002 \pm 0.058$ beats). An ANOVA found a significant relationship between the division and Green's onset difference, with a large effect size ($F(4, 19539) = 1956.79,\ p < .001;\ \eta^2 = .29$). Subsequent post-hoc testing with Tukey's HSD procedure found significant pairwise differences for all comparisons at $p < .001$.

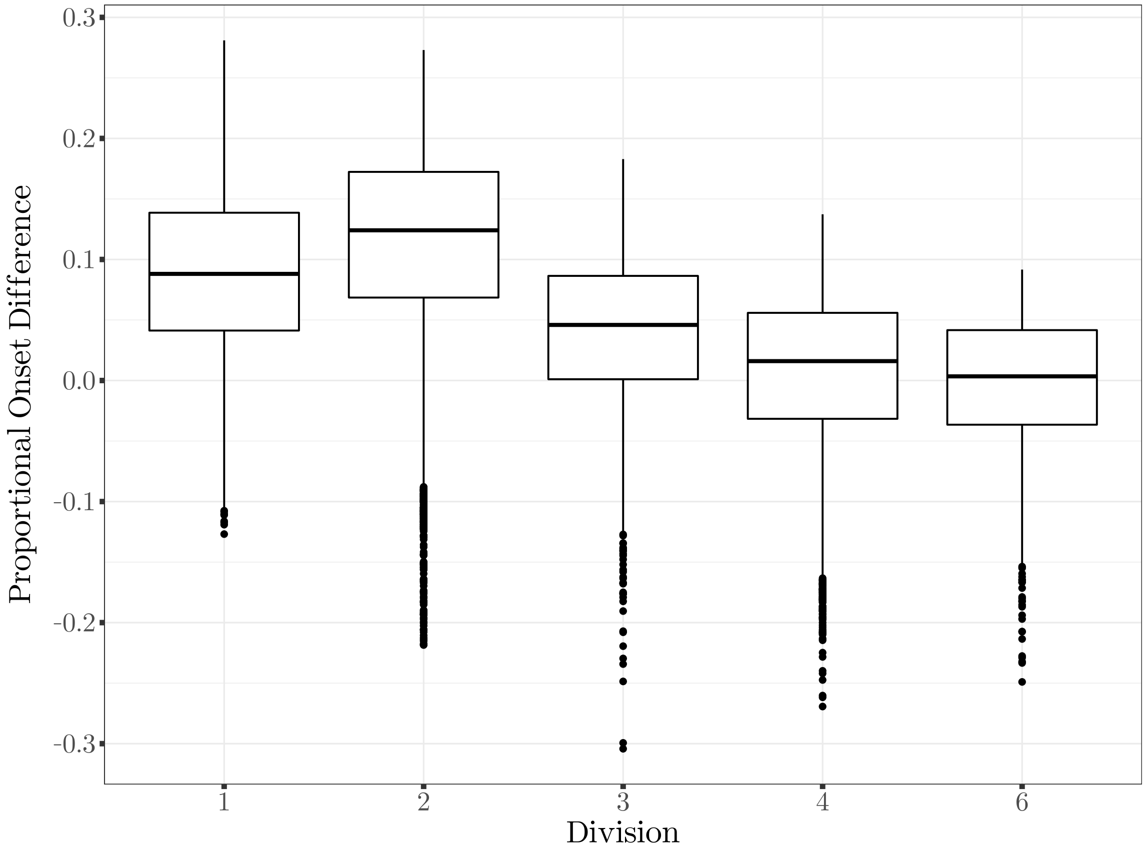

Figure 7.19: Distribution of proportional onset differences for divisions that contained more than 1000 note events in Green's corpus.

These results indicated that the division, as a proxy for beat density and complexity, did have a significant impact on Green's onset differences. In support of the hypothesis, the higher divisions had a mean onset difference close to zero, suggesting that the structural limitations of higher density beats resulted in Green playing closer to the nominal placement of the notes. The decrease was not linear, with Green's onset difference being highest for notes played in a beat of division two. Notes played in a beat of division one could only be played on the beat. Therefore, the lower onset difference may have been due to Green adhering closer to nominal onsets for on-beat notes compared to off-beat notes. It was also possible that the notes played in a beat of division two had a higher onset difference due to notes played in a swing feel, with the second note definitionally being played behind the nominal onset. Both of these hypotheses were investigated below in the following On or Off Beat and Swing sub-sections.

7.3.3 Note Placement vs. On or Off Beat

It was hypothesised that when Green played an on-beat note, it would be closer to its nominal metrical onset than those played off-beat. Running contrary to this was the previous finding that notes played in more dense divisions had a lower onset difference, which suggested that off-beat notes had a lower onset difference. The distribution of Green's note onset differences for both on-beat and off-beat notes can be seen in Figure 7.20. This graph shows that, contrary to the hypothesis but suspected from the division analysis, Green's off-beat notes were played closer to their nominal position than those played on-beat. This was supported by an ANOVA, which found a significant difference between the onset differences and the on or off beat note placement, with a small effect size ($F(1, 20475) = 252.15,\ p < .001;\ \eta^2 = .01$). The observed difference in the mean between the two categories was very small, with on-beat notes having a mean onset difference of 0.07 ± 0.08 beats while off-beat notes had a mean onset difference of 0.05 ± 0.09 beats.

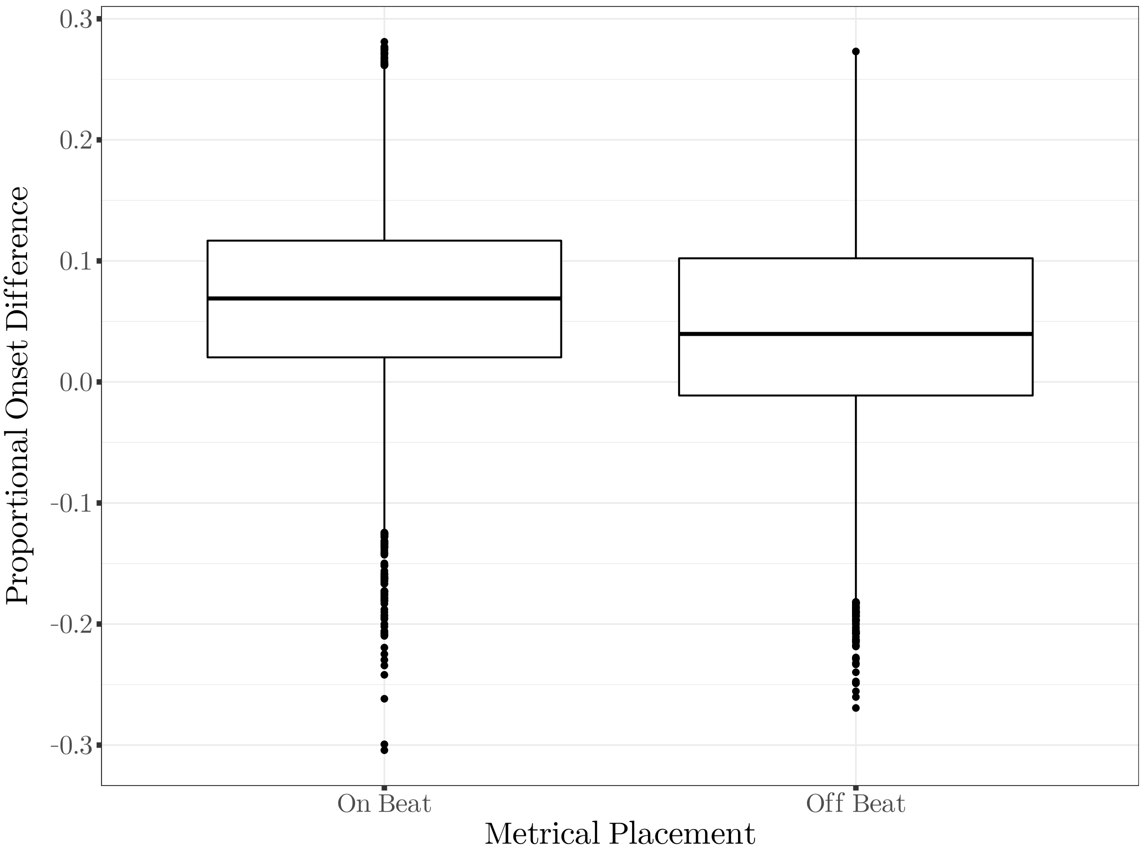

Figure 7.20: Distribution of proportional note offsets for notes played on or off the beat in Green's corpus.

These results found that while Green's onset difference did significantly differ depending on whether the note was played on or off the beat, the actual difference between the two situations was small. Both on-beat and off-beat notes were played with a mean onset difference of around 5%. This indicated that the onset difference observed for notes played in a beat of division one, being lower than those played in a beat of division two, was not due to Green playing on-beat notes closer to their nominal position.

7.3.4 Note Placement vs. Swing

Lastly, Green's note onset difference was investigated in relationship to whether or not a note was played in a swing pair, and if so whether it was the first or second note. The hypothesis was that, due to the BUR nature of swung note pairs, the second note would have been played further behind the beat than either the first note of the pair, or non-swung notes Green played. Secondly, based on the results of the division analysis, it was hypothesised that both notes of a swung quaver pair would have a larger offset than those not in a swung pair.

Green's distribution of onset differences for each of the three swing positions can be seen in Figure 7.21. The graph shows that, as hypothesised, the second note in a swung pair tended to be played further away from the nominal onset than either the first note or non-swung notes. The second note in a swung pair had a mean onset difference of 0.10 ± 0.08 beats. The first note was only played slightly closer to its nominal position, with a mean onset difference of 0.08 ± 0.07 beats. Notes that were not in a swing pair were only played behind the beat by around 5% of the surrounding beat length ($\bar{x} = 0.04 \pm 0.08$ beats). Additionally, non-swung notes were the only notes played ahead of beat by more than 15% of the beat length.

An ANOVA investigated the statistical relationship between the three swing positions (first note, second note, not swung) and Green's onset difference; a significant difference was found, with a medium effect size ($F(2, 20474) = 964.13,\ p < .001;\ \eta^2 = .09$).[16] These results supported the hypothesis that whether or not a note was part of a swung pair had a significant impact on Green's onset difference. Both notes in a swung pair were played further behind the beat than non-swung notes and, due to the nature of swung pairs, the second note was played even further behind the nominal onset. These results also partially explained why notes played in a beat of division two were played further behind the beat than notes in any other division.

[16] Subsequent post-hoc tests with Tukey's HSD procedure found significant pairwise differences for all comparisons at $p < .001$.

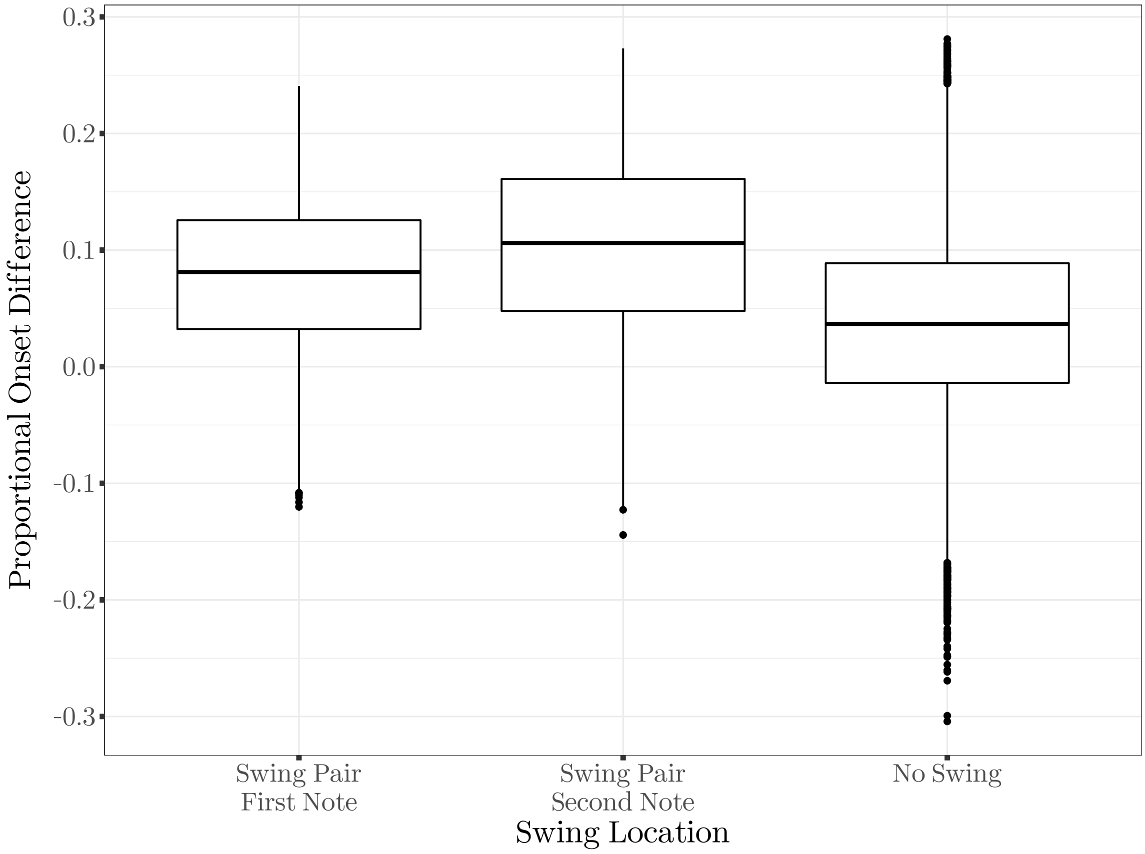

Figure 7.21: Distribution of Green's onset difference for the first note of a swing pair, the second note of a swing pair, and non-swung notes.

7.3.5 Note Placement Summary

The analysis into Green's note placement found that he tended to play behind the beat, by around 5% of the surrounding beat length. Due to the lack of available documentation, this research successfully reverse-engineered the note onset difference calculations to ensure the validity of the results. As nearly three quarters of the notes Green played were behind the beat, the results of the analysis also confirmed Scott's description of Green's playing as being "behind-the-beat" (2006, 1). The tempo range of an improvisation was found to have a significant impact on Green's note onset differences. The results indicated that Green played further behind the beat at higher tempos, while having a lower mean onset difference at lower tempos. At the lower tempo range, Green was also found to play a higher proportion of notes ahead of the beat, likely due to the increased rhythmic variety and flexibility found at this tempo range. The division of the beat a note was played in, which acted as a proxy for the beat density, was also found to have a significant effect on Green's note placement. Notes played in less dense beats, with fewer divisions, were played further behind the beat than those played in higher division beats. Green's mean onset difference approached zero when notes were played in beats with a higher division. This suggested that he both played notes closer to their nominal onset, as well as playing more notes ahead of the beat. It was hypothesised that Green's lower mean note onset difference for beats in division one may have been related to Green playing on-beat notes closer to their nominal position than off-beat notes. Analysis into the metrical placement of the notes found that Green's mean onset difference was very similar for both locations. It was also hypothesised that the higher mean onset difference for notes played in beats with a division of two was due to the effect of swung note pairs. This was supported by the analysis, which found that each note in a swung pair was played further behind the beat when compared to all other notes Green played. The second note in the pair was also found to be played significantly further behind the beat than the first note. Altogether, Green consistently played slightly behind the beat, with the tempo range and division having a significant impact on his placement of notes.

7.4 Green's Improvisational Style In The Micro Domain

This chapter focused on analysing aspects of Green's improvisational style with features from the micro domain. The analysis specifically focused on Green's use of swing, the micro timings between notes, and his placement of notes in relation to their nominal metrical onset.

The swing analysis found that Green tended to swing harder than many of the performers in the WJazzD. It also found that, consistent with prior research, an increasing tempo lowered Green's BUR, while his BUR increased when improvising over a blues. The effect of the tonality mode may also explain Green's higher mean BUR, as the predominant jazz styles he was playing between 1960 and 1965 (hard bop and post-bop) had substantial blues influences.

The micro timings analysis found the micro-gaps between notes was not useful in describing Green's improvisational style; however, the articulation was. Although the vast majority of the notes Green played had an articulation class of normal or long, both the tempo and interval size had a significant effect on Green's articulation. Green's articulation was found to be shorter at the higher tempo range, and longer at the lower tempo range. The analysis also found that the highest articulation occurred between notes a semitone or tone apart, the most frequent interval sizes for slurs. These techniques were likely more common at the lower tempo range, with the increased rhythmic variety and flexibility provided by the tempos, increasing the mean articulation values. The results of the analysis also indicated that Green's mean articulation for repeated notes was lower than expected, as there is no technical limitation in having an articulation of 1 between repeated notes on the guitar. Investigation into Green's use of repeated notes indicated that he deliberately changed the articulation of repeated notes, possibly in compensation for not changing the pitch.

The note placement analysis found that Green tended to play behind the beat. Nearly three-quarters of all note events in Green's corpus were played behind the beat, matching the findings from Scott (2006, 1). Green's mean note offset was around 5% of the surrounding beat length. The tempo of the improvisation was found to have a significant effect on Green's note placement, with Green playing further behind the beat in the higher tempo range. The lower tempo range lowered Green's mean note onset difference, with Green also more likely to play notes ahead of the beat. The division of a beat was also found to have a significant impact on Green's note placement, with division acting as a proxy for beat density. Green's mean note placement approached the nominal metrical onset as the division of the beat increased. Whether or not a note event was part of a swing pair was also found to have a significant effect. Both notes in a swing pair were played further behind the beat than non-swung notes, with the second note played furthest from the nominal placement.

In summary, the analysis of Green's improvisational style in the micro domain found that many of the features were substantially affected by other rhythmic based features, including the tempo and division. There were consistencies to Green's use of the micro domain features, with Green swinging hard, playing normal to long articulated notes, and playing behind the beat. The following chapter, Macro Domain, analysed Green's improvisational style from the opposite viewpoint of the micro domain, investigating the broad structures of his solos.

Chapter 8

Macro Domain

The macro domain referred to broad structural features of an improvisation. The analyses of this chapter focused on Green's use of phrases within his improvisations and large scale features that changed over the course of an improvisation. The micro domain analyses were aided by the precise nature of the transcriptions; in contrast, the analysis of large scale features took advantage of the corpus analytical approach. Having complete and easy access to the full corpus of Green's data in a computer-readable format allowed for the investigation of how features changed broadly across the course of an improvisation. The features under investigation included: metrical density; median interval; median pitch; proportion of NDTs; and proportion of NHTs. All of these features were calculated per bar, and then examined across the course of each improvisation.

8.1 Phrases

In his book, *How To Improvise: An Approach To Practicing Improvisation*, Crook defined a phrase as "a period of continuous, but not necessarily constant, melodic/rhythmic activity" (1991, 26). A phrase could also be thought of as "a series of notes that display a complete musical sense, and that form a natural division of the melodic line" (*Encycloédie Larousse* in Nattiez (1990), 158), with it being recognised as "one of the most ambiguous [terms] in music" (Stein in Nattiez (1990), 159). There is no widely accepted definition for the minimum or maximum length of a phrase. For the purposes of this research a phrase was any set of notes from a single phrase annotation with four or more notes. This definition removed fifty-four phrases from Green's corpus, leaving 1197 phrases, or 95.68% of the data. The following analyses into Green's use of phrases explored descriptive features, including the phrase shape (contour) and the length of phrases, followed by investigations into how Green began and ended phrases.

8.1.1 Phrase Descriptors

This section focused on the analysis of Green's phrase shapes (contours), and the lengths of his phrases. Other features that could be analysed in this section (e.g. inter vs. intra phrase rests) were not included as they had been analysed previously.

Phrase Shape (Contour)

The phrase shape, or contour, was a description of the overall trend of the pitch in a phrase. Two sets of contour codes were included in *MeloSpy*:

1) Huron contour codes, based on Huron's 1996 paper "The Melodic Arch in Western Folksongs" (1996);
2) Abesser contour codes, named after a researcher on *The Jazzomat Research Project*, with Abesser contour codes similar to the Huron codes, but based on "more stable estimators [that are] more suitable for longer sequences" (Jazzomat Research Project 2017).

The Huron contour codes were calculated by:

1) dividing each phrase into three groups, the first pitch, the last pitch, and the middle pitches;
2) the first pitch was compared to the mean value of the middle pitches (MIP) based on their MIDI note representation, and was determined to be either descending, horizontal, or ascending;
3) the MIP was compared to the last pitch, and was again determined to be descending, horizontal, or ascending.

These two comparisons were then used to determine which of the nine full contour codes the shape of the phrase best matched, seen in Table 8.1. The reduced code is a simplification of the descending, horizontal, and ascending full codes based on the First $\rightarrow$ MIP direction.

Table 8.1: Phrase shape (contour) descriptors for both reduced and full codes.

Reduced Code	Full Code	First → MIP[1]	MIP → Last[1]
Descending	Descending	Descending	Descending
	Descending-Horizontal		Horizontal
Concave	Concave		Ascending
Horizontal	Horizontal-Descending	Horizontal	Descending
	Horizontal		Horizontal
	Horizontal-Ascending		Ascending
Convex	Convex	Ascending	Descending
Ascending	Ascending-Horizontal		Horizontal
	Ascending		Ascending

[1] MIP: Mean Inner Pitch

Although the Jazzomat Research Project described the Abesser contour code as using "more stable estimators" (Jazzomat Research Project 2017), documentation was not found on what these more stable estimators were, or how they calculated the Abesser contour codes. The author was also unable to successfully reverse engineer the contour codes. Visual inspection of over fifty phrases found that the Abesser contour codes generally aligned with the shape of the phrase.[1] As the contour was a high level description of the shape of a phrase, the descriptions provided by Abesser contour code were sufficient for the analyses.

Unless specified otherwise, all graphs in this section had the pitch and onset time normalised for each phrase. The highest note in each phrase was labelled one while the lowest was labelled zero, and the first note had an onset of zero while the last had an onset of one. This allowed for all of the phrases to be compared, regardless of their length or range. Figure 8.1 shows all of Green's phrases, normalised, with a line of best fit showing the general trend to his phrases.[2] The overall trend indicated that Green began his phrases with an ascending sequence of pitches, followed by a descending line.

[1]The phrases were visually inspected by plotting the pitch MIDI numbers against the note number in the phrase as a scatterplot, with a quadratic line of best fit. The code used to generate the plots can be found in Appendix B, code block B.7.

[2]The line of best fit was generated using the `geom_smooth` function of `ggplot2`, with the formula 'y~poly(x,4)', a fourth order polynomial.

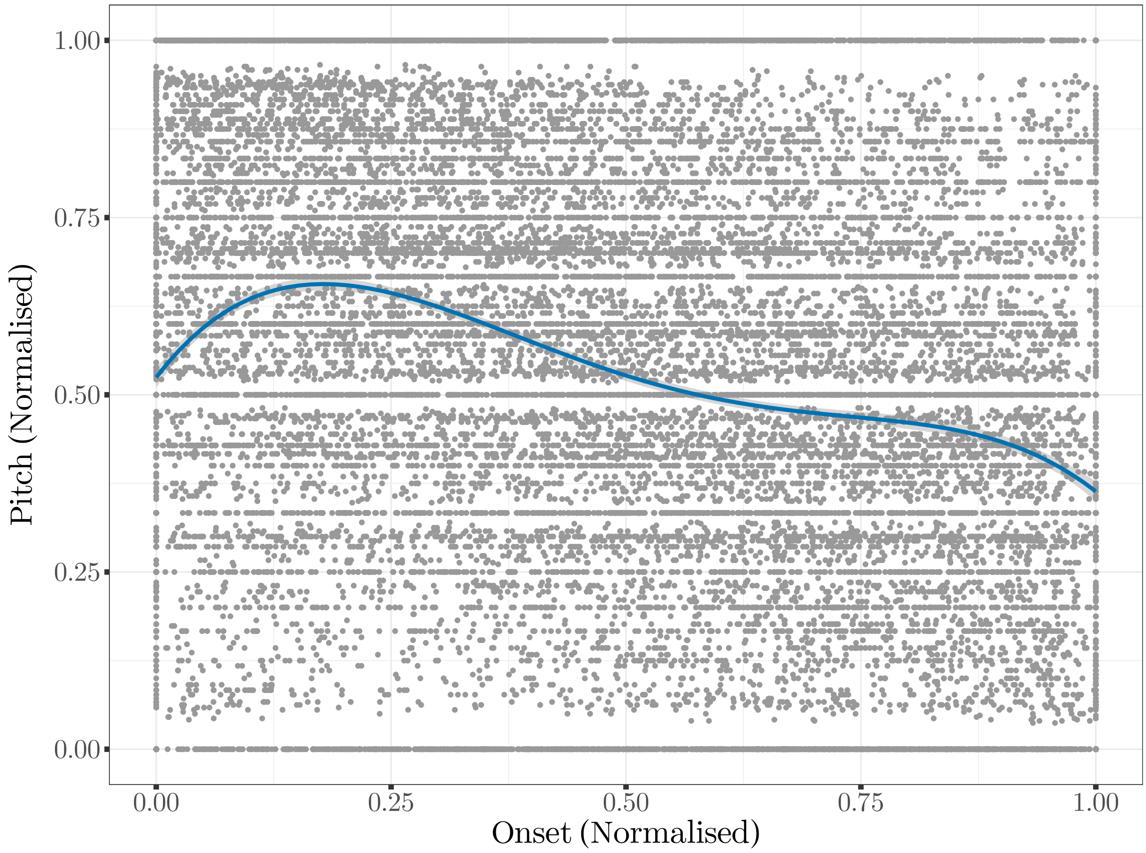

Figure 8.1: Normalised note onset vs. pitch for all phrases combined in Green's corpus with line of best fit indicating overall trend.

Table 8.2 shows how frequently each contour was found in Green's improvisations. The two most frequent Abesser codes were descending (25.56%) and convex (20.55%), together comprising nearly half of all phrases Green played. The next three most frequent contours were ascending (10.94%), concave (12.36%), and horizontal-descending (10.53%). These five contours, each with more than 100 occurrences, comprised the vast majority (79.95%) of all phrases Green played. All of the five least frequent phrase shapes played by Green involved a section of horizontal playing (horizontal-descending, descending-horizontal, ascending-horizontal, horizontal, and horizontal-ascending), suggesting that Green rarely played long stretches of notes without increasing or decreasing the pitch.

Table 8.2: Abesser code frequency of phrases played by Green.

	Asc	Asc-Hor	Hor-Asc
Count	131	60	37
Percent	10.94%	5.01%	3.09%
	Convex	Hor	Concave
Count	246	56	148
Percent	20.55%	4.68%	12.36%
	Desc	Desc-Hor	Hor-Desc
Count	306	87	126
Percent	25.56%	7.27%	10.53%

Short names of Abesser codes

Figure 8.2 took the data in Figure 8.1 but split it into each of the nine Abesser contour codes. Six observations could be made about the general shape of Green's phrases from these graphs:

1) regardless of the Abesser code, phrases tended to begin with an upward trend in pitch (horizontal-ascending phrases excepted);
2) ascending phrases tended to have a section of more horizontal playing in the middle of the phrase, in comparison to descending phrases that had a more consistent contour throughout;
3) convex phrases tended to reach their peak pitch around a third of the way through the phrase before descending again;
4) concave phrases tended to have a small increase in pitch at the start of a phrase and a small descending line at the end of a phrase;
5) horizontal phrases had an overall trend most similar to concave phrases, but with a smaller decrease in pitch at the middle of a phrase, and a larger descending pitch trend at the end of the phrase;
6) none of the Abesser codes that included a labelled horizontal section were consistently horizontal, with each having substantial pitch deviation throughout the phrase.

This showed that phrases that were predominantly descending, or contained substantial descending segments, were most common in Green's improvisations. Additionally, the contours suggested that Green's phrases tended to begin with ascending intervals and finish with descending intervals.[3]

[3] The specifics of how Green's phrases began and ended will be investigated fully in the Phrase Beginnings and Endings section.

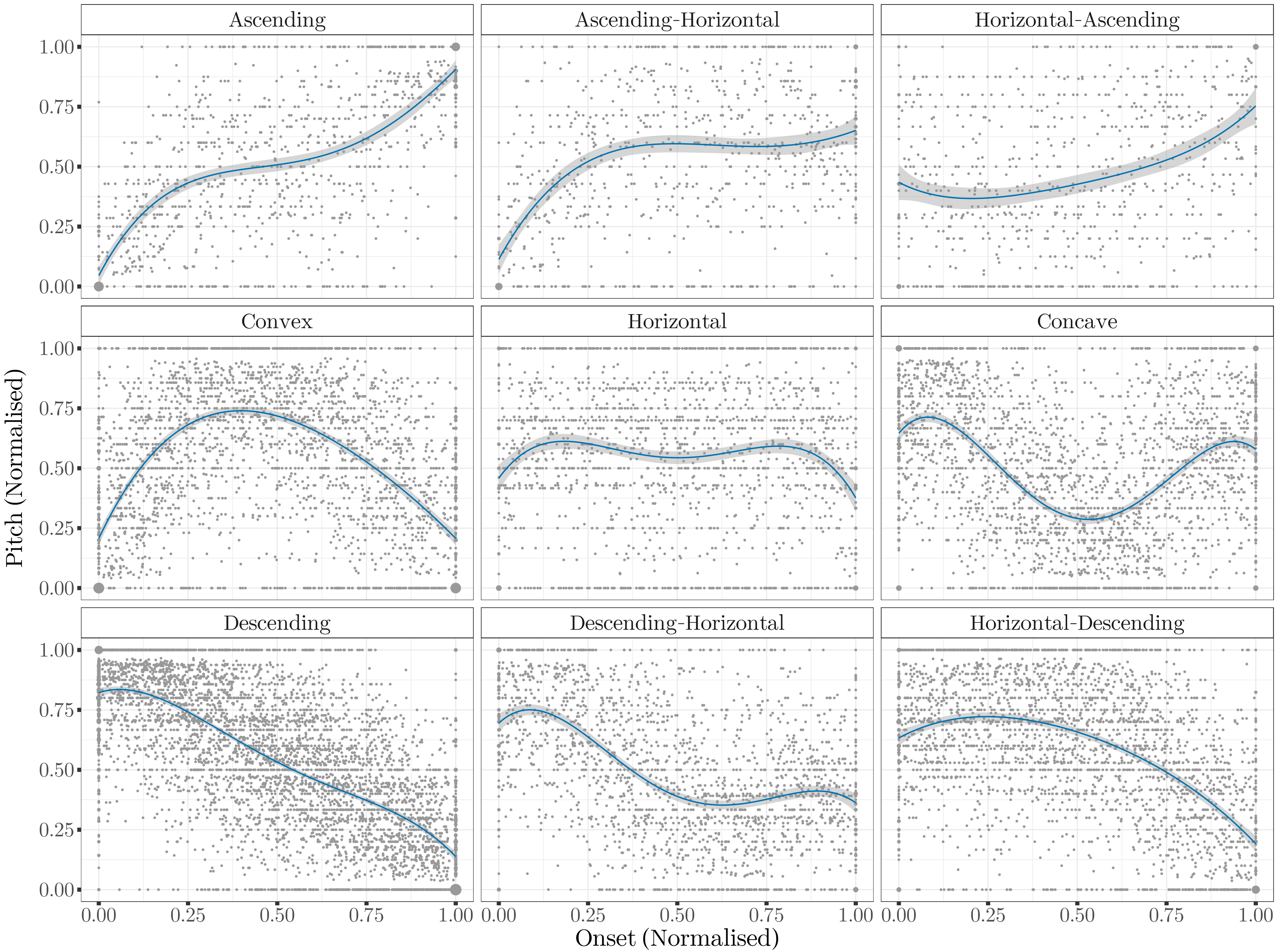

Figure 8.2: Normalised onset vs. pitch for each contour code in Green's corpus. Line of best fit from fourth-order polynomial.

Phrase Length

There were two approaches for measuring the length of a phrase, the number of notes in a phrase and the length of the phrase in beats. The phrase length in beats was calculated by summing the $\mathrm{IOI}_{\mathrm{BeatProp}}$ of all notes in a phrase with the $\mathrm{duration}_{\mathrm{BeatProp}}$ of the last note, as in Equation (8.1) for a phrase with n notes.

$$\mathrm{PhraseLength}_{\mathrm{Beats}} = \mathrm{IOI}_{\mathrm{BeatProp}_1} + \mathrm{IOI}_{\mathrm{BeatProp}_2} + ... + \mathrm{IOI}_{\mathrm{BeatProp}_{n-1}} + \mathrm{duration}_{\mathrm{BeatProp}_n} \tag{8.1}$$

Table 8.3 shows the distribution of Green's phrase lengths for both descriptors. This data showed that, on average, Green played seventeen notes per phrase, with phrases going for just under nine beats. The mean length of Green's phrases was around 40% longer than the median length, due to Green playing a small number of very long phrases. Within the corpus, Green played 150 phrases with thirty or more notes. The length of phrases in beats was approximately half of that of the note count, due to Green often playing quaver or quaver triplet equivalent rhythms. The phrase length as note count and beat sum was found to be significantly correlated, with a large effect size ($r = .88,\ t(1195) = 65.24,\ p < .001,\ r^2 = .78$).

Table 8.3: Statistical descriptors for both measures of Green's phrase lengths.

	Mean	SD	Min	Q1	Median	Q3	Max
Note Count	16.98	15.16	4	8	12	20	158
Beat Sum	8.92	7.61	1.02	4.19	6.69	11.62	95.85

Figure 8.3 shows the distribution of phrase lengths, both as note count and as beat sum, for phrases with thirty or fewer notes. The most frequent phrase lengths, based on note count, were between six and twelve notes. Although 53.55% of Green's phrases went for between four and twelve beats, the most frequent beat lengths were between three and seven beats (40.18%). An additional 12.61% of phrases went for between 12 and 16 beats, 7.29% went for between 16 and 24 beats, with only 3.93% of Green's phrases going for more than 24 beats (6 bars in $\frac{4}{4}$). Figure 8.4 shows a phrase, from Green's improvisation over *Born To Be Blue* (Green 1961c), that had the most common contour (convex) and had a length in both note count (twelve) and beats (5.08) close to the median length.

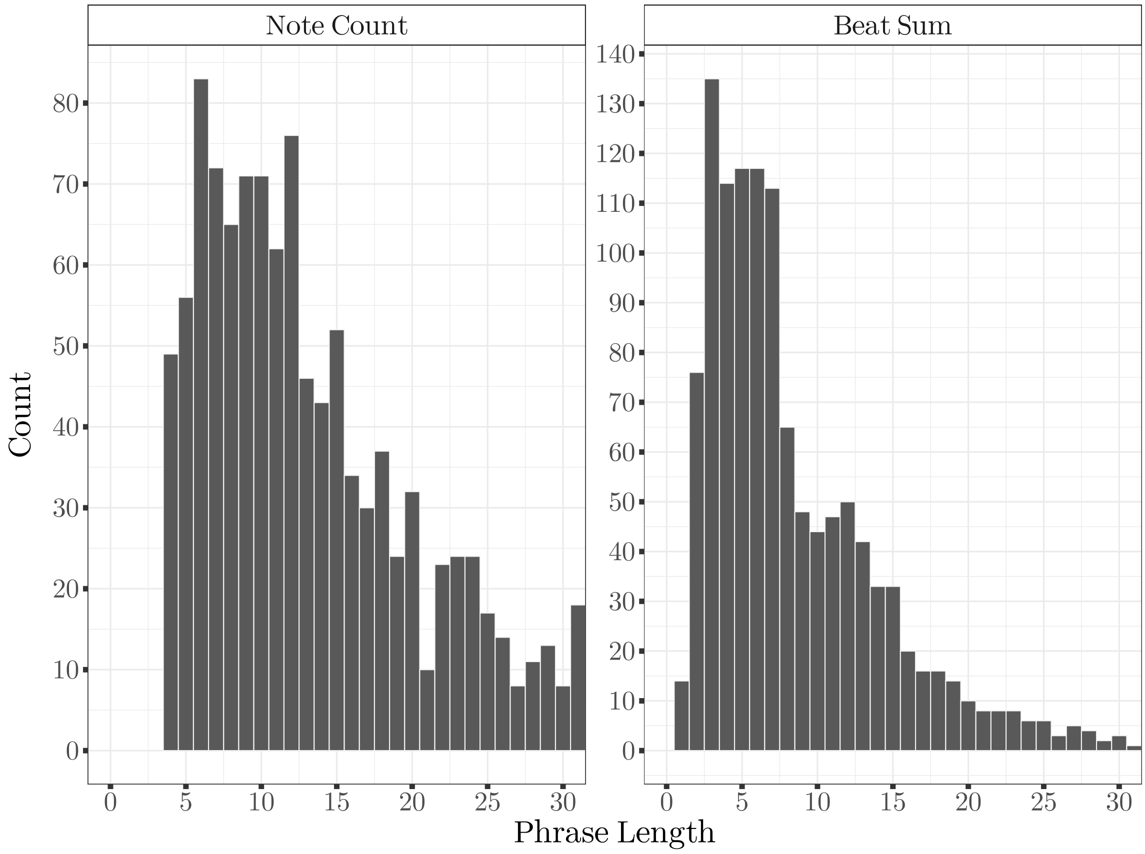

Figure 8.3: Distribution of phrase lengths, as both note count and beat sum, for phrases with thirty or fewer notes in Green's corpus.

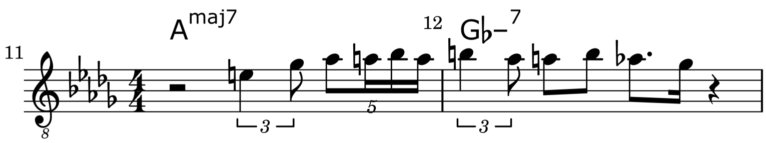

Figure 8.4: Example of a phrase with the most common contour and length, *Born To Be Blue* (1961), phrase 6, bars 11–12.

It was hypothesised that there was a relationship between the length of the phrases played by Green and their contour. Figure 8.5 shows the distribution of phrase lengths, as note count, for each of the nine Abesser contour codes. This data suggested that phrases with different contours tended to have different lengths in Green's improvisations. For example, ascending phrases had a median length of seven notes, while Green's horizontal phrases were more than three times as long, with a median length of twenty-six notes. All remaining contours had a median length between ten (convex) and nineteen (descending-horizontal) notes.

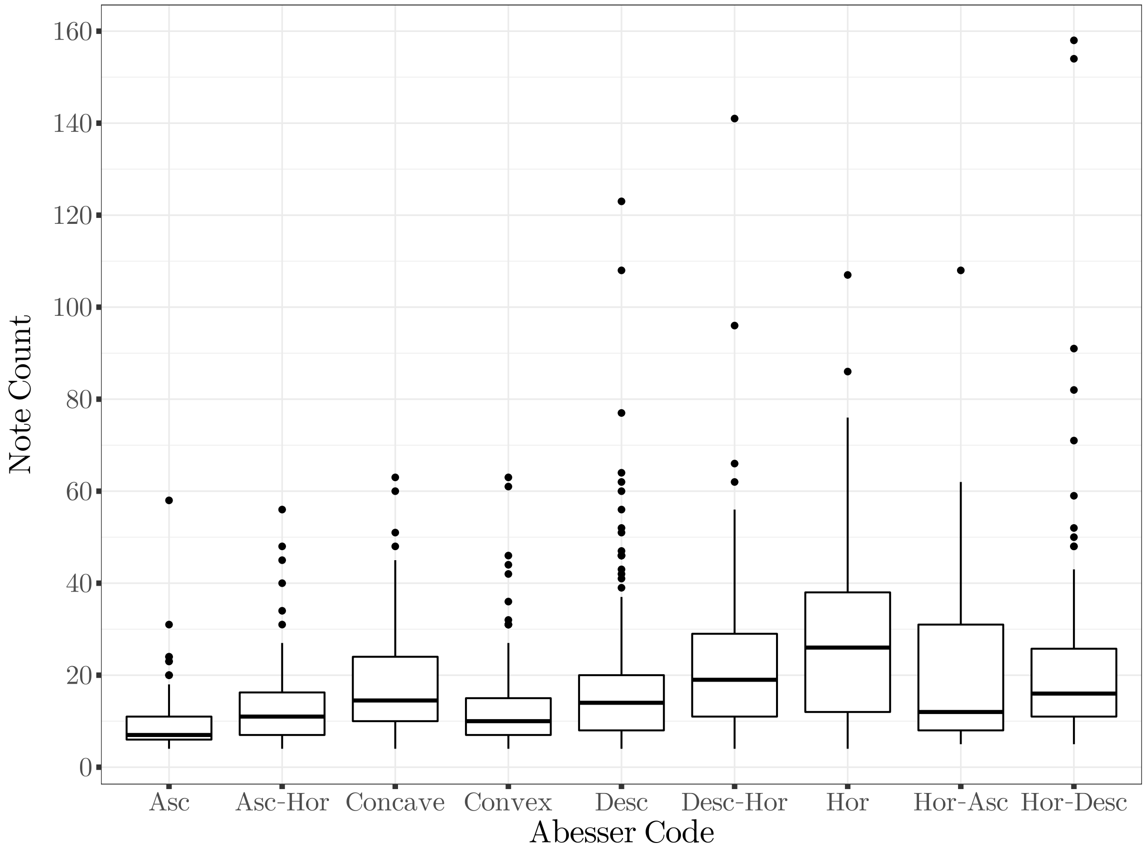

Figure 8.5: Distribution of Green's phrase lengths for each Abesser contour code.

An ANOVA found a significant relationship between the Abesser contour code and the length of the phrases, with a significant effect found with a medium effect size $(F(8, 1188) = 19.60,\ p < .001;\ \eta^2 = .12)$. The results of the post-hoc pairwise comparisons with Tukey's HSD procedure are shown in Table 8.4 due to the high number of comparisons, with significant differences shown in italics. The three contours that were most frequently found to have a significantly different length from the other contours were ascending, convex and horizontal. The data from Figure 8.5 showed that these contours were representative of short, medium, and long phrases played by Green. These results supported the hypothesis that the length and contour of Green's phrases were related.

Table 8.4: Abesser code vs. phrase length TukeyHSD *p*-values.

	Asc	Asc Hor	Hor Asc	Convex	Hor	Concave	Desc	Desc Hor	Hor Desc
Asc	-	.31	*<.001*	.53	*<.001*	*<.001*	*<.001*	*<.001*	*<.001*
Asc-Hor	.31	-	.32	.98	*<.001*	.78	.95	*.008*	*.008*
Hor-Asc	*<.001*	.32	-	*.010*	.08	.93	.68	1.00	1.00
Convex	.53	.98	*.010*	-	*<.001*	*.004*	*.006*	*<.001*	*<.001*
Hor	*<.001*	*<.001*	.08	*<.001*	-	*<.001*	*<.001*	.09	*.023*
Concave	*<.001*	.78	.93	*.004*	*<.001*	-	1.00	.15	.16
Desc	*<.001*	.95	.68	*.006*	*<.001*	1.00	-	*.009*	*.005*
Desc-Hor	*<.001*	*.008*	1.00	*<.001*	.09	.15	*.009*	-	1.00
Hor-Desc	*<.001*	*.008*	1.00	*<.001*	*.023*	.16	*.005*	1.00	-

Significant results in italics

The final analysis into Green's phrase lengths investigated how the number of unique pitches played interacted with the length of the phrase. This was based upon an observation that Green often played repeated note sequences. For example, Figure 8.6 shows a repeated sequence from Green's improvisation over *Seepin'* (Solo 1, Green 1960b). This excerpt shows a repeated two note pattern from bars 176–179, with the phrase concluding with a two note ascending chromatic resolution to E♮. This forty-five note phrase had a unique pitch count (UPC) of four notes, resulting in a unique pitch proportion (UPP) of 8.89%, with the most frequent note (D♭5) comprising 51.11% of the phrase.[4] The hypothesis was that the majority of Green's longer phrases were of the type in Figure 8.6, a phrase mainly consisting of a repeated motif.

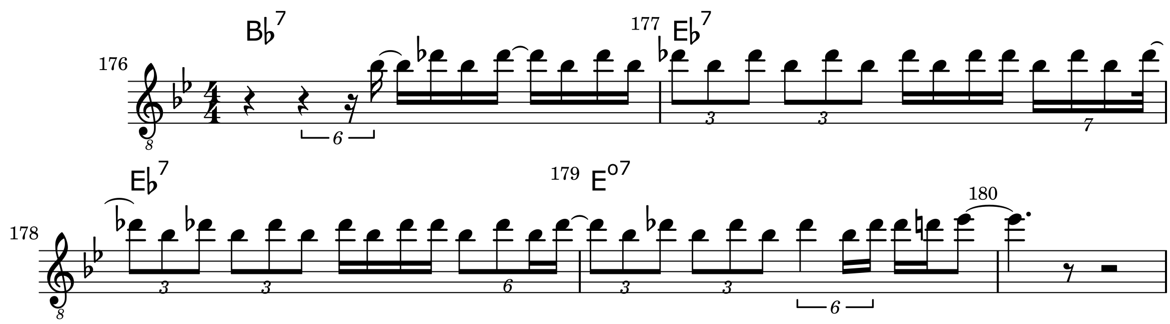

Figure 8.6: Example of a phrase with a repeated motif, *Seepin'* (Solo 1, 1960), bars 176–180.

[4] UPP = UPC $\div$ phrase length$_{\text{notes}}$

To aid in this analysis, a new categorical variable based on the phrase length beat sum was created, fuzzy phrase length, with four levels: brief (less than one bar); short (one to two bars); medium (two to four bars); and long (longer than four bars). The mean number of notes per fuzzy phrase were: brief, 6.88 ± 2.58; short, 11.69 ± 4.70; medium, 20.45 ± 8.43; and long, 43.98 ± 25.16. In considering the number of unique pitches in a phrase, it was necessary to account for a guitar having only forty-five unique pitches.[5] Therefore, for the fifty-two phrases with more than forty-five notes, the UPP was calculated by dividing the number of unique notes in the phrase by forty-five, rather than the number of notes in the phrase. Figure 8.7 shows the two methods of calculating the frequency of unique pitches for each fuzzy phrase length, with the UPC on the left and the UPP on the right.

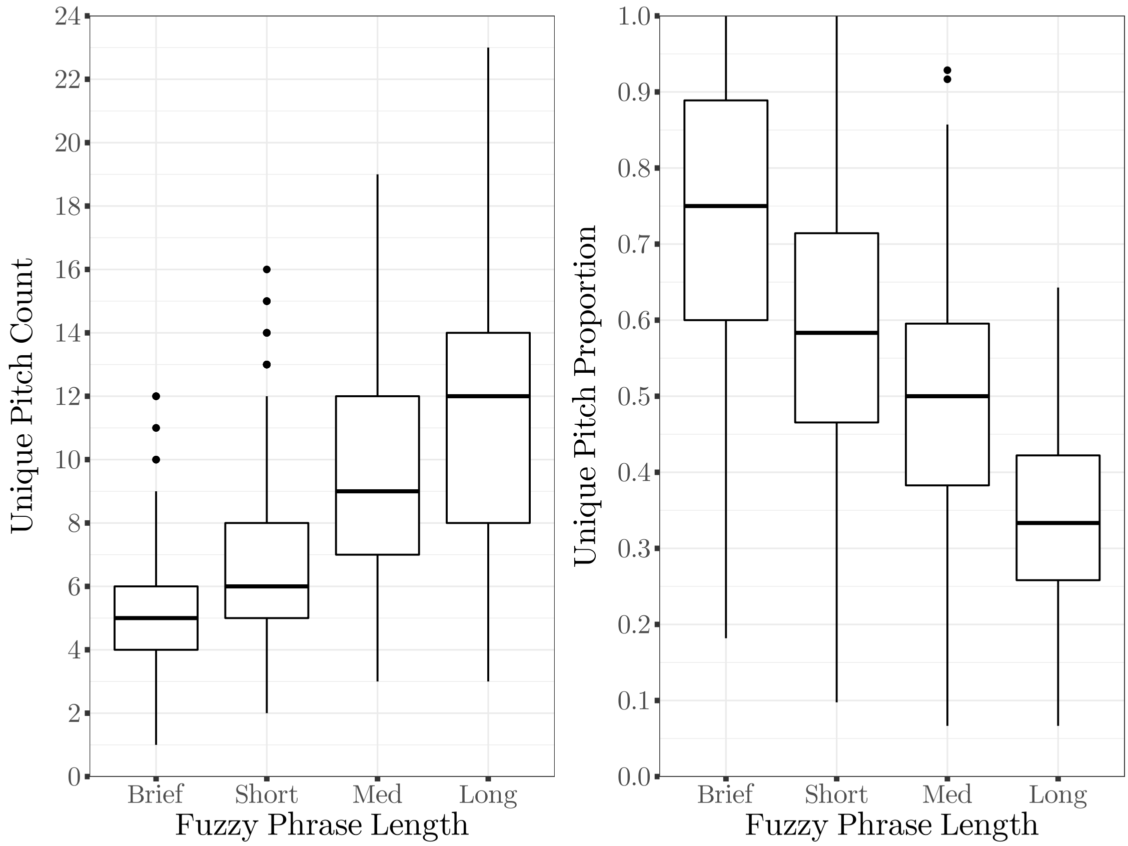

Figure 8.7: Distribution of frequency of unique pitches for each fuzzy phrase length. Left: Unique pitch count. Right: Unique pitch proportion.

These graphs showed the expected behaviour of shorter phrases having a lower UPC but a higher UPP. An ANOVA found a significant relationship between the fuzzy phrase length and the number of unique pitches in the phrase, with a large effect size ($F(3, 1193) = 207.44,\ p < .011;\ \eta^2 = .34$).[6] As the length of the phrases increased,

[5]For a guitar in standard tuning, with six strings and twenty frets. The highest pitch in Green's corpus was on the 18$^{\text{th}}$ fret of the first string, B♭5.

[6]Subsequent post-hoc comparisons with Tukey's HSD procedure found significant pairwise differences for each comparison at $p < .001$.

so did the UPC. Longer phrases also tended to have a lower UPP, indicating that comparatively fewer pitches comprised the majority of the phrase. The median UPC for the shortest phrases was four or five pitches, while medium and long phrases only had a slightly higher median at nine and twelve pitches respectively. However, if the majority of Green's longer phrases were based on repeated note motifs, they would have been expected to have lower UPC and UPP. Therefore, these results did not support the hypothesis that longer phrases played by Green were predominantly based on repeated motifs. Although the data did show that these did exist in Green's improvisations, with minimum UPC of three notes for both medium and long phrases, this was not Green's main approach for longer phrases.

Phrase Descriptor Summary

Across all of Green's phrases, regardless of the contour, he appeared to favour beginning his phrases with an ascending sequence of pitches, followed by a descending line. Abesser contour codes of convex and descending appeared most frequently in Green's transcriptions, comprising nearly half of all phrases. Ascending, concave, and horizontal-descending phrases were each played around 10% of the time. Green's phrases had an average length of seventeen notes, or around nine beats, with Green only playing 150 phrases with more than thirty notes. Green's most frequent phrase lengths had between six and twelve notes, while three-quarters (76.27%) of his phrases were between three and sixteen beats. A significant relationship was found between the Abesser contour and the length of phrase. Ascending phrases tended to be the shortest, concave phrases were of middling length, while horizontal phrases were, on average, the longest. There was also a significant relationship found between the length of the phrase and the number of unique pitches, with longer phrases having more unique pitches. Consequently, these results did not support the hypothesis that Green's longest phrases were comprised predominantly of short repeated motifs.

8.1.2 Phrase Beginnings and Endings

The other investigation into Green's use of phrases focused on how he began and ended his phrases. The features of interest were the starting and ending CPC, interval, and beat placement. Of all the notes in Green's corpus, only 2394 (11.78%) began or ended a phrase.

Chordal Pitch Class

The main pitch feature of interest was the CPC_{Weight}, with the hypothesis being that arpeggio tones were more common as the last note in the phrase, while scale tones and NHTs were more common at the start of the phrase. Beginning with the start of Green's phrases, Figure 8.8 shows the CPC_{Weight} distribution for notes that began a phrase, compared to all other notes Green played. The graph shows a very similar distribution for both situations, with Green playing slightly fewer NHTs and arpeggio tones at the beginning of the phrase, with slightly more notes coming from the scale. A χ^2-test found no significant difference in the CPC_{Weight} distribution for notes which did or did not begin a phrase ($\chi^2(2) = 2.35,\ p = .309,\ V = .01$). Therefore, it can be concluded that approximately half of Green's phrases began with an arpeggio tone, just under 30% from the scale, with only around 20% beginning with a NHT; in total, approximately 80% of all phrases began with an HT. These results did not support the hypothesis that scale and NHTs were more common at the start of a phrase, when compared with Green's overall CPC_{Weight} distribution.

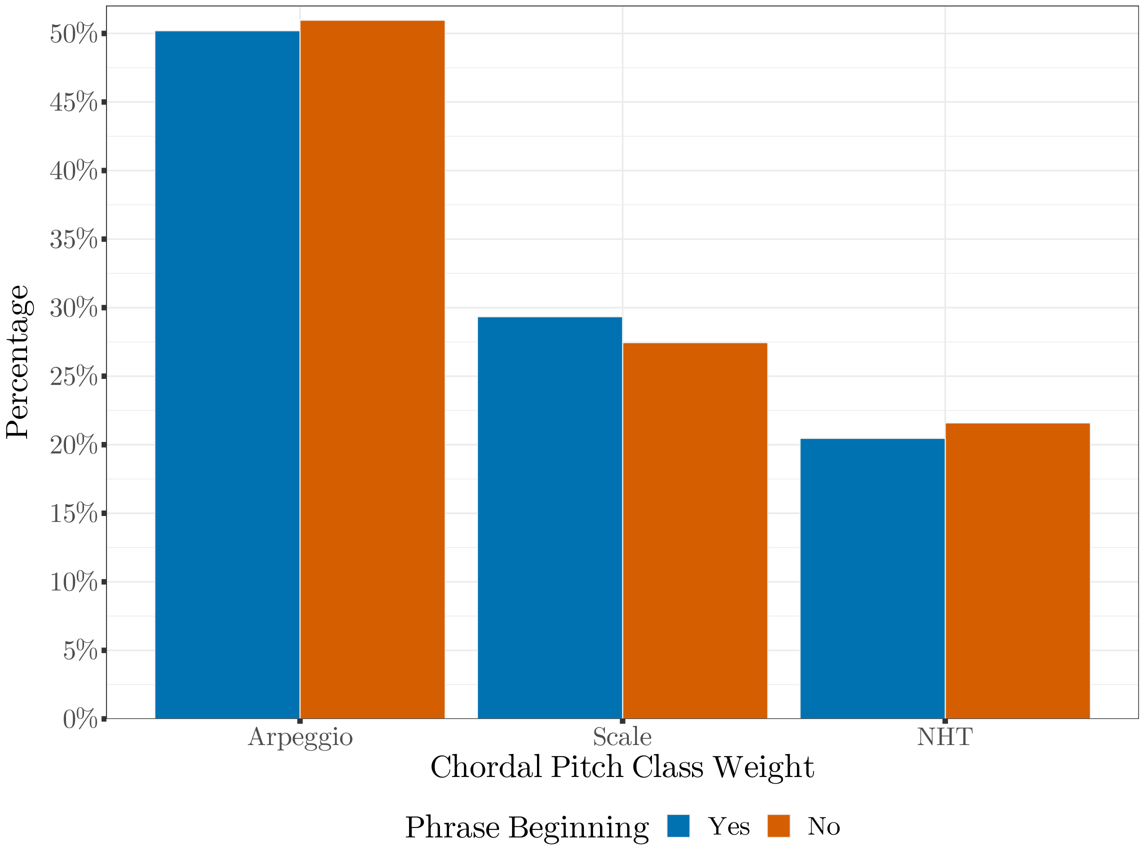

Figure 8.8: Distribution of CPC_{Weight} for notes that began a phrase compared to all other notes in Green's corpus.

Figure 8.9 shows the $\text{CPC}_{\text{Weight}}$ distribution for notes that ended phrases, compared to all other notes. This data showed that nearly two-thirds of Green's phrases ended with an arpeggio tone, while NHTs were substantially less common. Unlike notes that began a phrase, those that ended a phrase had a significantly different distribution compared to all other notes, with a small effect size ($\chi^2(2) = 103.41$, $p =< .001$, $V = .07$). These results supported the hypothesis that arpeggio tones were significantly more likely to be played at the end of Green's phrases.

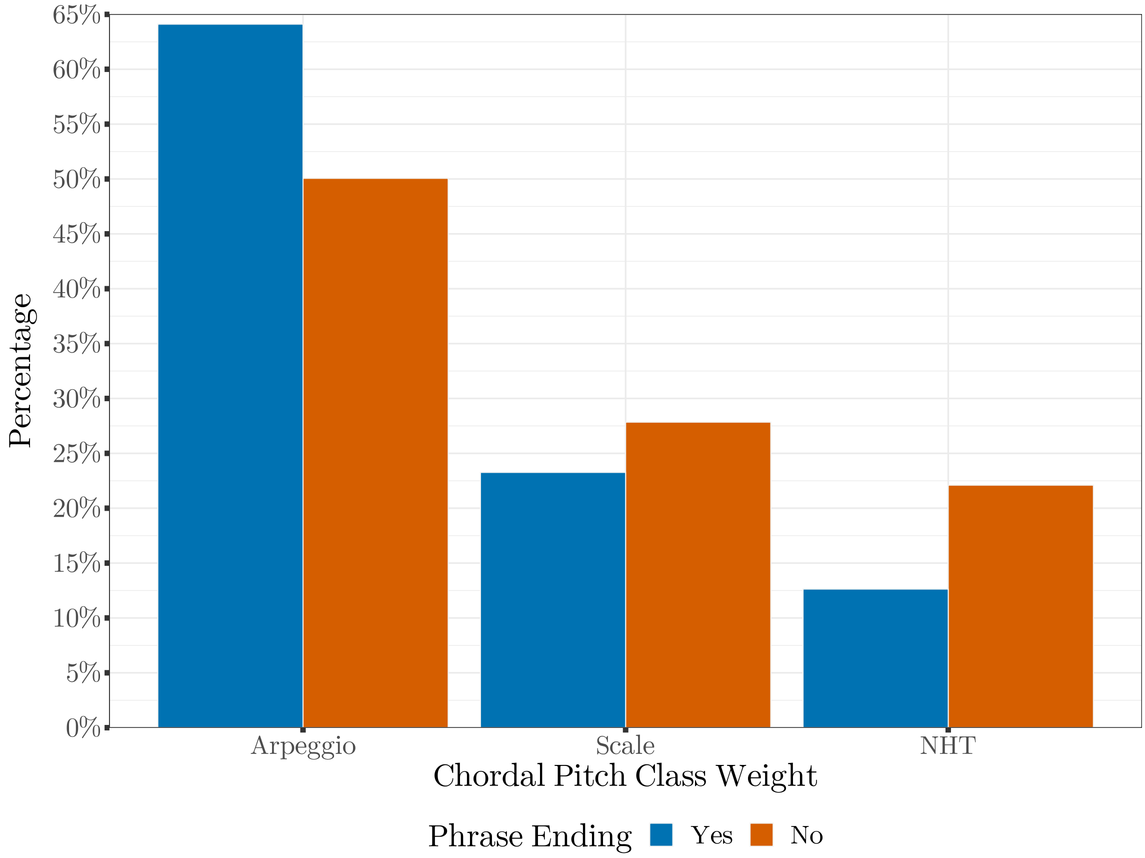

Figure 8.9: Distribution of $\text{CPC}_{\text{Weight}}$ for notes that ended a phrase compared to all other notes in Green's corpus.

Figure 8.10 separates the $\text{CPC}_{\text{Weight}}$ into the CDPCX. This graph shows that the two most common CDPCX that Green ended his phrases with are often considered the most stable tones, the tonic and 5^{th}. Following this were the third and the seventh, with each of the scale tones occurring approximately the same amount of time. This showed that Green most often ended his phrases with a stable arpeggio tone.

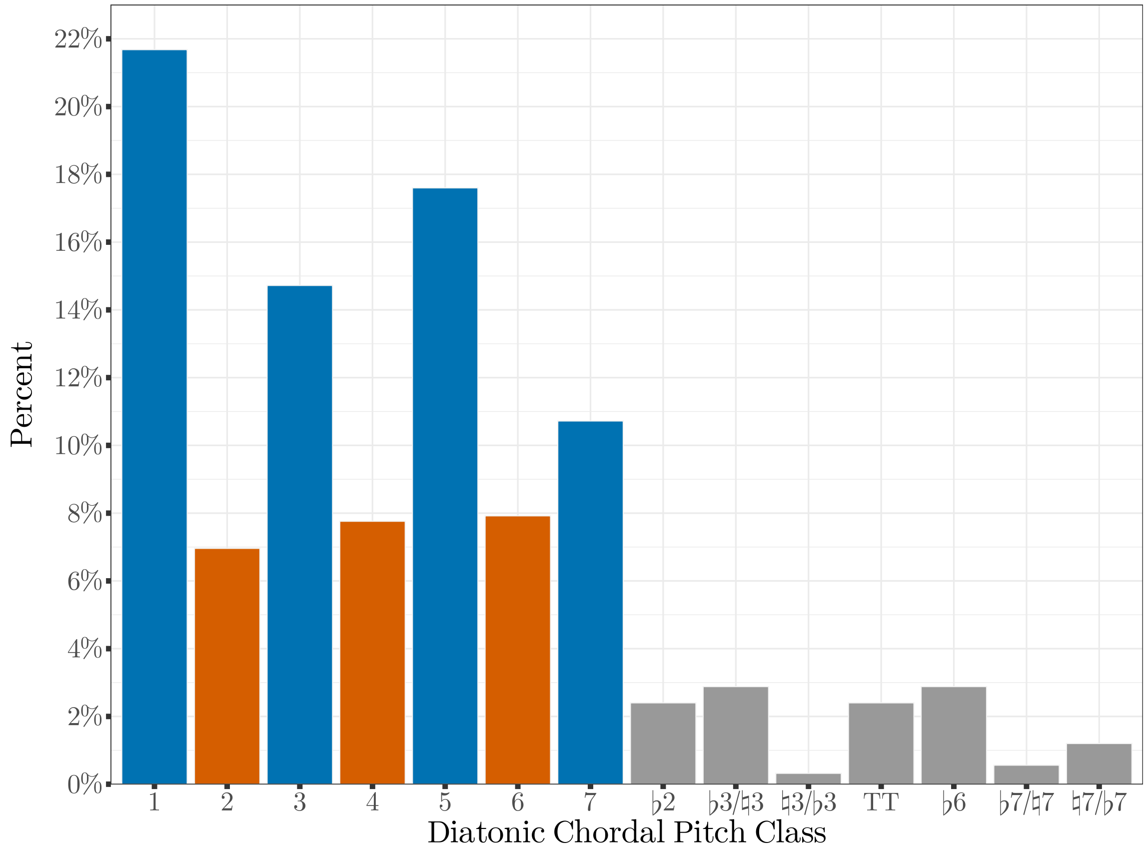

Figure 8.10: Distribution of CDPCX for notes that ended phrases in Green's corpus.

Intervals

The two interval features of interest in how Green began and ended his phrases were the Parsons and the specific intervals played. For phrase endings it was the interval into the final note that was analysed, instead of the interval between phrases. Based on the phrase contour analysis, it was hypothesised that the majority of Green's phrases would begin with an ascending interval, with Green's phrases ending with descending movement.

Green's Parsons distribution for each phrase position is shown in Figure 8.11. The graph showed that the direction in which Green started and ended his phrases was substantially different not only from each other, but to the middle notes as well. A χ^2-test found a significant difference in the distribution of the Parsons in comparison to the phrase position, with a small effect size ($\chi^2(4) = 801.14,\ p =< .001,\ V = .14$). Subsequent post-hoc tests with Tukey's HSD procedure found significant pairwise differences in all comparisons at $p < .001$. As hypothesised, 76.82% of phrases Green played started with an ascending interval. In comparison, 49.84% of Green's phrases ended with a descending interval into the final note, with a further 32.16% ending with an ascending interval. The highest proportion of repeated Parsons occurred at the end of a phrase, with 18.00% of phrases ending with a repetition.

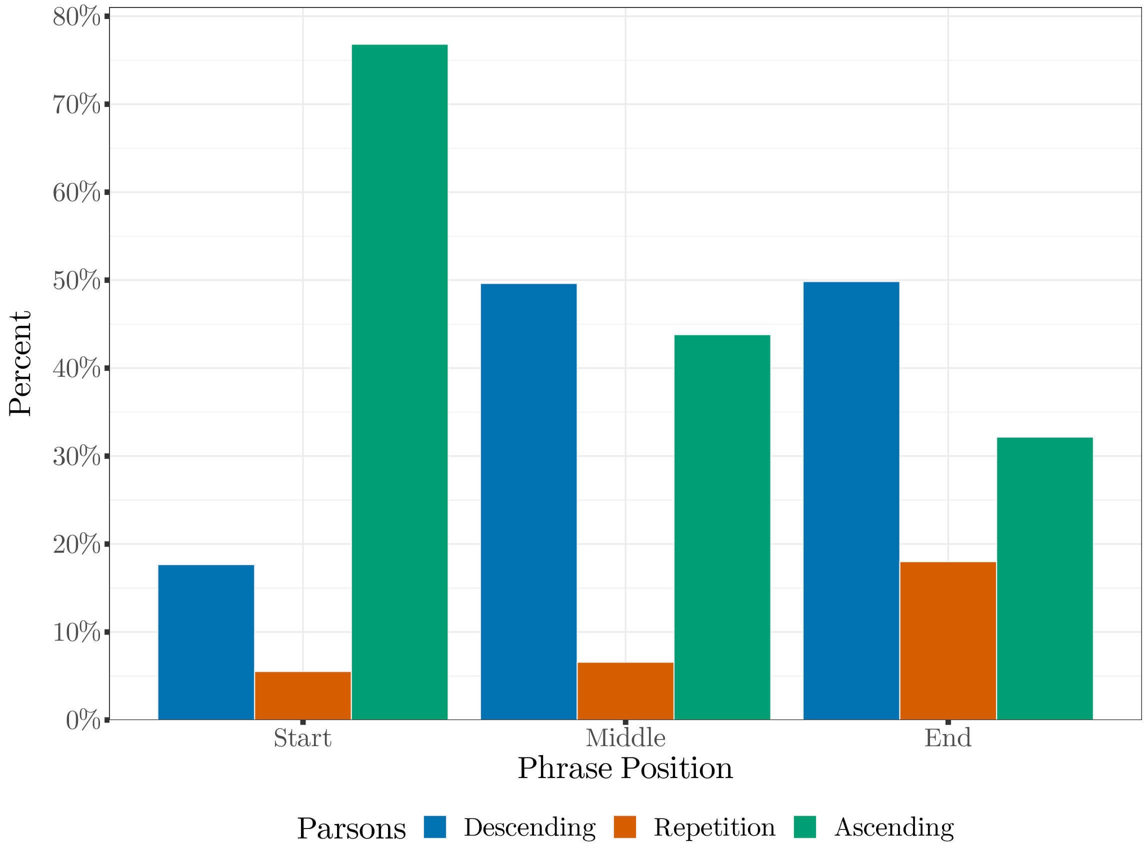

Figure 8.11: Parsons distribution for each phrase position in Green's corpus.

Figure 8.12 show the specific intervals Green played over each phrase position. The figure shows all intervals within an ascending or descending fifth, containing 98.10% of all the data.[7] This data showed that by far the most common intervals Green used to begin his phrases were an ascending semitone or minor third. For the end of Green's phrases, no individual intervals stood out as they did for the beginning of his phrases, with the most common single interval being a repetition. Overall, these results supported the hypothesis that Green was most likely to begin his phrases with an ascending interval, specifically a semitone or minor third, with descending and repeated intervals found more commonly at the end of a phrase.

[7] 99.44% of starting notes, 98.14% of middle notes, and 96.24% of ending notes.

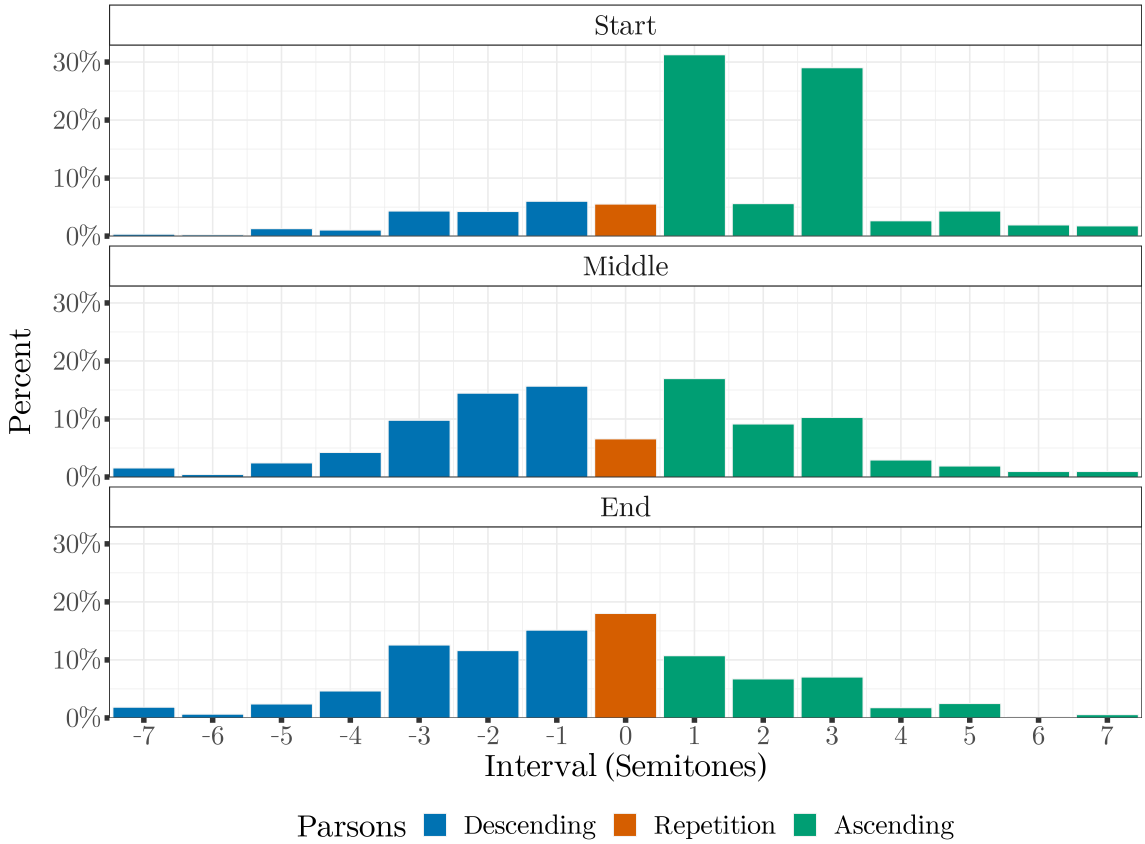

Figure 8.12: Distribution of raw intervals Green played in each phrase position.

Beat Placement

The final feature analysed was the beat placement of the first and last notes of Green's phrases. For the first note of the phrase, both the starting location and the transition to the second note was investigated; for the last note, only the location was analysed. It was hypothesised, based on standard jazz practices, that phrases were more likely to begin off the beat, followed by a note in subsequent down beat. It was also hypothesised that the last note of a phrase would be more evenly distributed across the beat and sub-beat placements in a bar.

Figure 8.13 shows, as a circle map, the frequency and transition for the start of Green's phrases (top) compared to the rest of his notes (bottom). Separate circle maps were created for improvisations played in quadruple time (left) and triple time (right). This data showed that Green started his phrases in a fairly consistent manner, and that they were substantially different to the distribution of his other notes. For both time signatures the most frequent location where Green started a phrase was on the quaver off-beat position, with the following note played on the down beat.[8]

[8]The quaver off-beat position is the halfway point between two beats, where a nominal quaver would be played in a quaver note line.

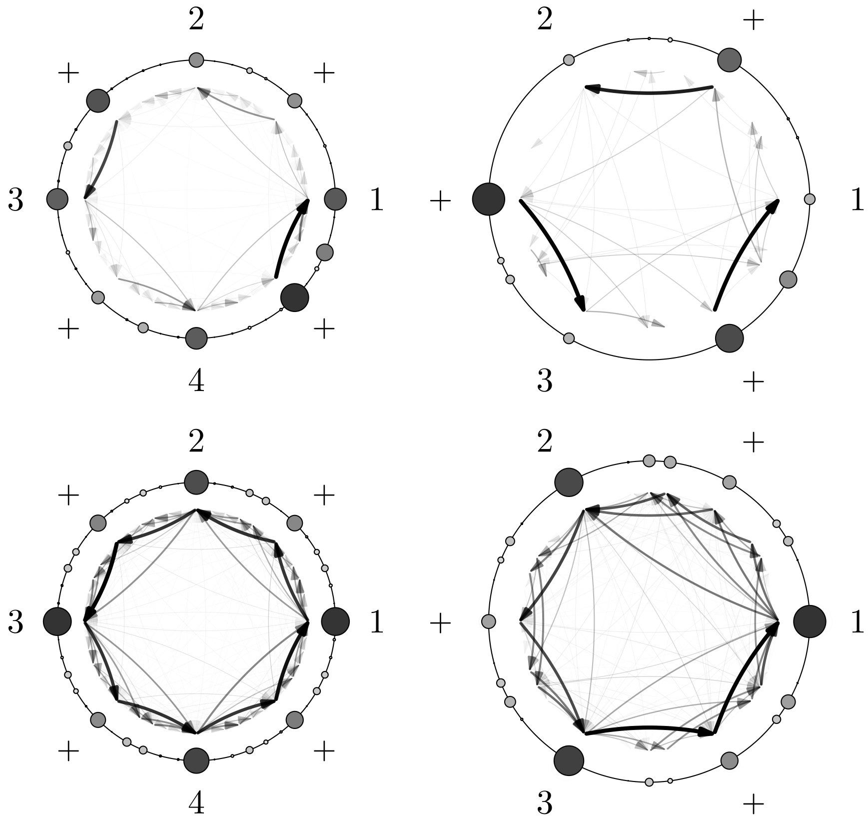

Figure 8.13: MCM beat placement and bigram transitions for notes of Green's phrases. Top: first note. Bottom: all other notes. Left: quadruple time. Right: triple time.

For improvisations in quadruple time, the most frequent transitions were from `4+` $\rightarrow$ `1` and `2+` $\rightarrow$ `3`. These were the most common places to start phrases due to their proximity to chord changes, providing forward momentum to an improvisation. The `1+` $\rightarrow$ `2` and `3+` $\rightarrow$ `4` also appeared reasonably often, with the off-beat to on-beat transition also providing some forward momentum. Phrases beginning on the beat were not uncommon, with 31.12% of phrases in quadruple time beginning on the beat. Unlike phrases beginning on the quaver off-beat position, Green did not have a strong trend of targeting a particular location after beginning a phrase on the beat, with both quaver off-beat position and the following beat being common.

For improvisations in triple time, the vast majority (80.00%) of phrases began off the beat, with the graph indicating that these were most frequently on the quaver off-beat position. Due to the infrequent use of triple time in Green's corpus, there were only seventy-five phrases to draw data from. As with the quadruple time, there was a strong tendency in Green's improvisations for these notes to be followed by a note played on the subsequent beat. The data indicated that transitions to `2+` $\rightarrow$ `3` and `3+` $\rightarrow$ `1` were the most common (ten occurrences each), while `1+` $\rightarrow$ `2` was played by Green marginally less often (nine occurrences). This suggested that in triple time Green did not prefer any specific quaver off-beat position to down beat movement. Each other bigram transition in triple time was played three times or fewer.

Figure 8.14 shows the beat distribution of the final note of Green's phrases in both quadruple time (left) and triple time (right). This data showed that in triple time Green more frequently ended phrases on the down beat; however, due to the small number of phrases in $\frac{3}{4}$ this was inconclusive. For quadruple time, the data showed a fairly similar distribution to the 'all notes' graph in Figure 8.13, with many phrases ending both on and off the beat. There was evidence of Green slightly preferring to finish a phrase with an on-beat note, with 44.56% of phrases ending on the beat compared with 37.19% of non-phrase ending notes played on-beat.[9]

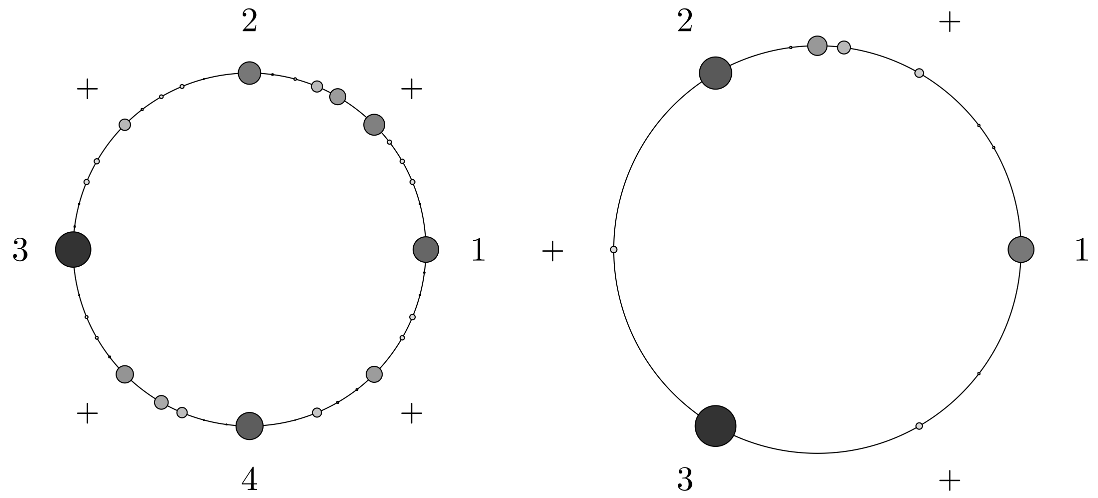

Figure 8.14: Beat placement circle map for the final note in Green's phrases. Left: quadruple time. Right: triple time.

As previously suggested, aligning the start of a phrase to a chord change could provide forward momentum in an improvisation. Therefore, it was hypothesised that more of Green's phrases would begin in the beat before a chord change, when compared to all other notes. The comparison between the phrase position and whether or not the note was played in the beat before a chord change can be found in Table 8.5. This data showed that phrases were more likely to start in the beat before a chord change, compared to notes played in the middle or end of a phrase. A χ^2-test found a significant difference in the distribution of notes played in each phrase position depending on if the next beat had a chord change, with a small effect size ($\chi^2(2) = 48.70,\ p =< .001,\ V = .05$).

[9] A significant difference was found in the on-beat vs. off-beat proportion of notes that ended a phrase compared to all other notes Green played, but with a small effect size ($\chi^2(1) = 25.20$, $p =< .001,\ V = .04$).

Table 8.5: Distribution of Green's notes in each phrase position depending on if the following beat had a chord change.

	Beat Before Chord Change	Not Beat Before Chord Change
Start		
Count	404	847
Percent	32.29%	67.71%
Middle		
Count	4769	13208
Percent	26.53%	73.47%
End		
Count	250	1000
Percent	20.00%	80.00%

Phrase Beginnings and Endings Summary

Green's phrases had common elements that occurred at the start and end of his phrases. Approximately 80% of all Green's phrases began with an HT, around 50% arpeggio tones and 30% scale tones; this did not differ significantly from his standard distribution. The vast majority of Green's phrases also began with an ascending interval, specifically ascending semitones and minor thirds. Green's phrases also tended to begin on a quaver off-beat position, followed by a note on the next down beat. The `2+` $\rightarrow$ `3` and `4+` $\rightarrow$ `1` transitions were most common in quadruple time, while all quaver off-beat position to down beat transitions were equally likely in triple time. Green was also slightly more likely to begin his phrases in the beat before a chord change. Green's phrases were most likely to end with an arpeggio tone, with just under two thirds of his phrases ending in this manner. Green was most likely to end a phrase on the tonic or 5th, with the other arpeggio tones also more likely to end a phrase than any of the scale tones. Consequently, colour tones such as the 6th or 9th were not common in Green's phrase endings. The most common single interval Green used to end a phrase was a repeated note, with descending intervals being most frequently played into the last note of a phrase. Unlike the start of Green's phrases, no obvious trend was found in the beat placements for the end of his phrases, although he was slightly more likely to play the final note on a down beat. In summary, a standard Green phrase would begin on a quaver off-beat position in the beat before a chord change with an HT, followed by an ascending interval. The end of his phrase would finish on a tonic or 5th and was slightly more likely to be played on the beat, and approached by a descending interval or repeated note.

8.1.3 Phrase Summary

This analysis found distinctive characteristics in the construction of Green's phrases. The overall trend in Green's phrases was to begin with an ascending sequence of notes, and then descend throughout the rest of the phrase, with these results supported by the interval analysis of his phrase beginnings and endings. That analysis found that Green preferred to begin his phrases with an ascending semitone or minor 3rd, with most phrases approaching the final note with a descending interval. The analysis also found that around half of Green's phrases matched a convex or descending contour. Green's shorter phrases had a high frequency of unique pitches, as the phrases became longer only around a third of the pitches were unique. Longer phrases also tended to have a horizontal contour, but were not frequently associated with a repeated motif. Green also had a tendency of ending his phrases with either the tonic or 5th. Green's phrases tended to start on a quaver off-beat position before beats 1 and 3; consequently, Green also frequently began phrases on the beat before a chord change. Combined, these created a sense of forward momentum at the start of Green's phrases. In summary, this analysis found that Green favoured certain approaches to beginning his phrases, the overall contour matched a convex or descending shape, and the majority of his phrases had a length between one and four bars.

8.2 Large Scale Feature Trends

The final section of the macro domain focused on investigating the trends of features over the course of an improvisation. The five features analysed were: the metrical density; the median interval size; the median pitch; the proportion of NDTs; and the proportion of NHTs. These analyses were largely descriptive in nature. To generate the data used in this section, each feature was calculated for every bar in Green's improvisations. Each improvisation's bars were then normalised between 0 (first bar) and 1 (last bar), to make the data comparable across the solos.

Figure 8.15 shows, for each feature, the trend across the course of the normalised improvisations, with a line of best fit added to show the overall movement.[10] None of the features showed a particularly strong or consistent trend in movement over an improvisation. While there were some small trends – for example, a slight increase in the metrical density over the course of an improvisation, or a decrease in the median pitch at the end of an improvisation – they were not substantial enough to form any strong conclusion about Green's improvisational trends.

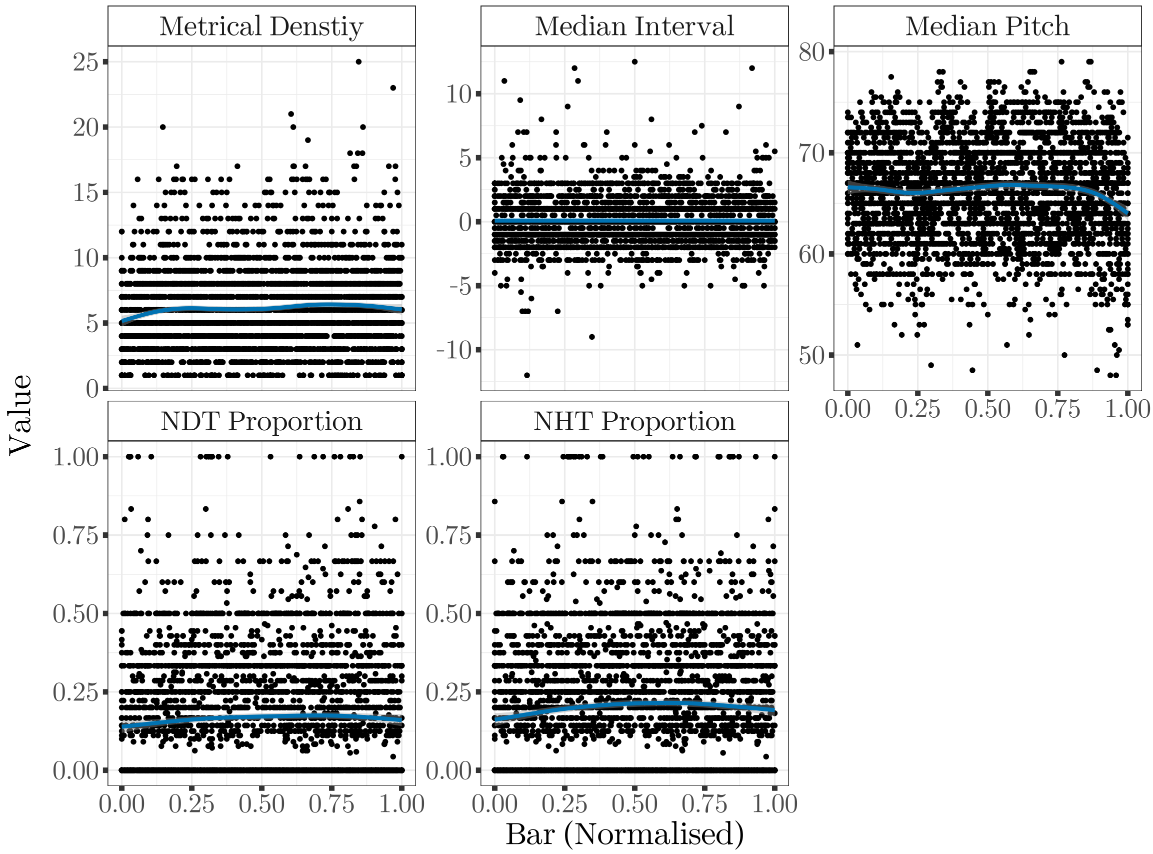

Figure 8.15: Overall bar-wise trends over the course of an improvisation for the metrical density, median interval, median pitch, NDT proportion, and NHT proportion.

Although there was not a single trend in the features across the course of Green's improvisations, there may have been a variety of different trends. For example, the phrase contour plot (Figure 8.1) which combined of all the phrases averaged out to a contour that had less movement than any individual phrase. Drawing inspiration from the phrase contour descriptors, visual inspection was undertaken for each feature trend for all forty of Green's improvisations.[11] This visual inspection determined that Green's feature trends could be categorised into five broad shapes.

[10]The line of best fit was generated with `ggplot2`'s `geom_smooth` function, with default values.

[11]As in Figure 8.15, all five features for each improvisation were plotted as a scatterplot, with a `geom_smooth` line of best fit added. It was from the automatically generated line of best fit that the shape categorises were determined.

The labels for each of these shapes were based on lower case letters and symbols, and described as modes (of the form * mode). These labels described the frequency of curves or straight sections observed in the trends of a feature.[12] The five modes, examples of which can be seen in Figure 8.16, were n mode (*Brazil*, Solo 2, Green 1962b), s mode (*Take These Chains From My Heart*, Green 1963f), r mode (*Sunday Mornin'*, Solo 1, Green 1961s), m mode (*Our Miss Brooks*, Green 1961o), and - mode (*Seepin'*, Solo 1, Green 1960b).[13] The modes were direction independent (e.g. n and u or w and m shaped features were categorised together), and also independent of the magnitude and location of the curves and peaks. The categories were deliberately broad, as it allowed for discussion of the general trend of the features, without needing to create a new mode for each improvisation.

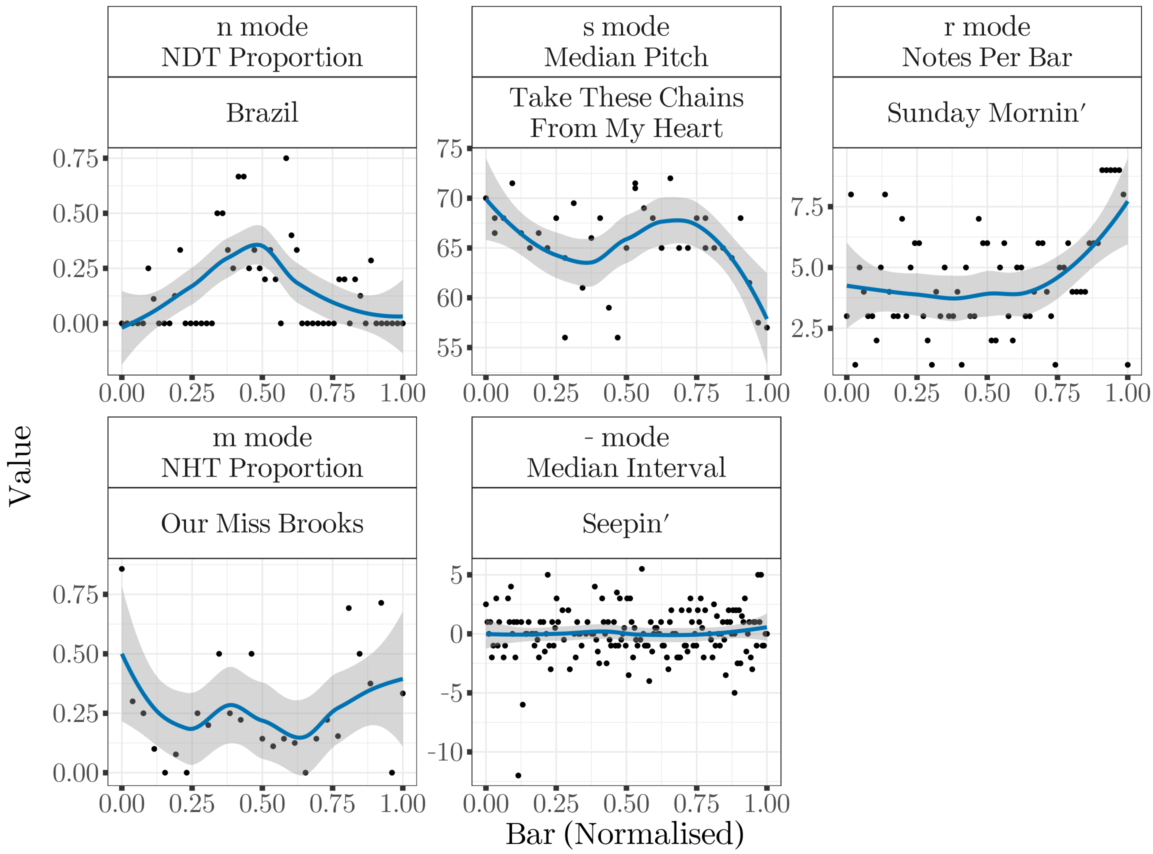

Figure 8.16: Examples of the five feature trend mode shapes. The improvisations the data were drawn from were (top left to bottom right): *Brazil* (Solo 2, 1962); *Take These Chains From My Heart* (1963); *Sunday Mornin'* (Solo 1, 1961); *Our Miss Brooks* (1961); *Seepin'* (Solo 1, 1960).

[12] Other research has used polynomials to describe complex contours; for example, Müllensiefen and Wiggins (2011). In keeping with the descriptive terms used to describe phrase contours, mode contour descriptors were used in this research.

[13] n mode improvisations could also have a horizontal section leading into or out of the trend.

The improvisations shown in Figure 8.16 were selected due to how clearly they represented each mode; however, not all were as easy to identify, or had clear modes of movement.[14] Table 8.6 shows the number of improvisations that were classified as having each mode shape for each feature. This data showed that for all features the most common shape was the s mode. The second most frequent shapes were either m mode or n mode for all features aside from median interval, where - mode was the second most common. For all other features, the - mode was the least frequent shape. This suggested that there were common trends in features across Green's improvisations, with features having one to three frequent mode shapes. The following sub-sections investigated the frequency of the different modes for each of the features. The focus was on showing examples of the most common feature trend modes to highlight how features changed over the course of specific improvisations.

Table 8.6: Frequency of feature trend mode shapes in Green's corpus.

	Mode Shape				
	n	s	r	m	-
Notes Per Bar	8	16	5	9	2
Median Interval	8	14	4	3	11
Median Pitch	8	18	4	10	0
NDT Proportion	8	21	1	10	0
NHT Proportion	10	16	2	11	1

8.2.1 Notes Per Bar

The three most common trend shapes in Green's metrical density were the s mode, m mode, and n mode, with 82.50% of improvisations having one of these modes. All of the m and n modes had convex curves; while for the s modes, half began with an increasing metrical density with the other half beginning with a decreasing metrical density. Figure 8.17 shows an example of each of the three most common mode shapes of metrical density in Green's improvisations. While the three improvisations varied substantially in their length – *Minor League* (Green 1964d, 2nd solo) twenty-four bars, *The Surrey With The Fringe On Top* (Green 1963g, 2nd solo) thirty-nine bars, and *Oleo* (Green 1962i, 1st solo) ninety-four bars – they all had a similar range of metrical densities. *Minor League* and *Oleo* ranged from one to nine notes per bar, while *The Surrey With The Fringe On Top* had two bars with ten

[14]All of the graphs, for each improvisation and each feature, can be found in Appendix E.8.

notes. These results indicated that Green frequently had multiple rises and falls in his metrical density across an improvisation. These changes in metrical density would have helped shape Green's solos, especially at the start or end of his improvisations.

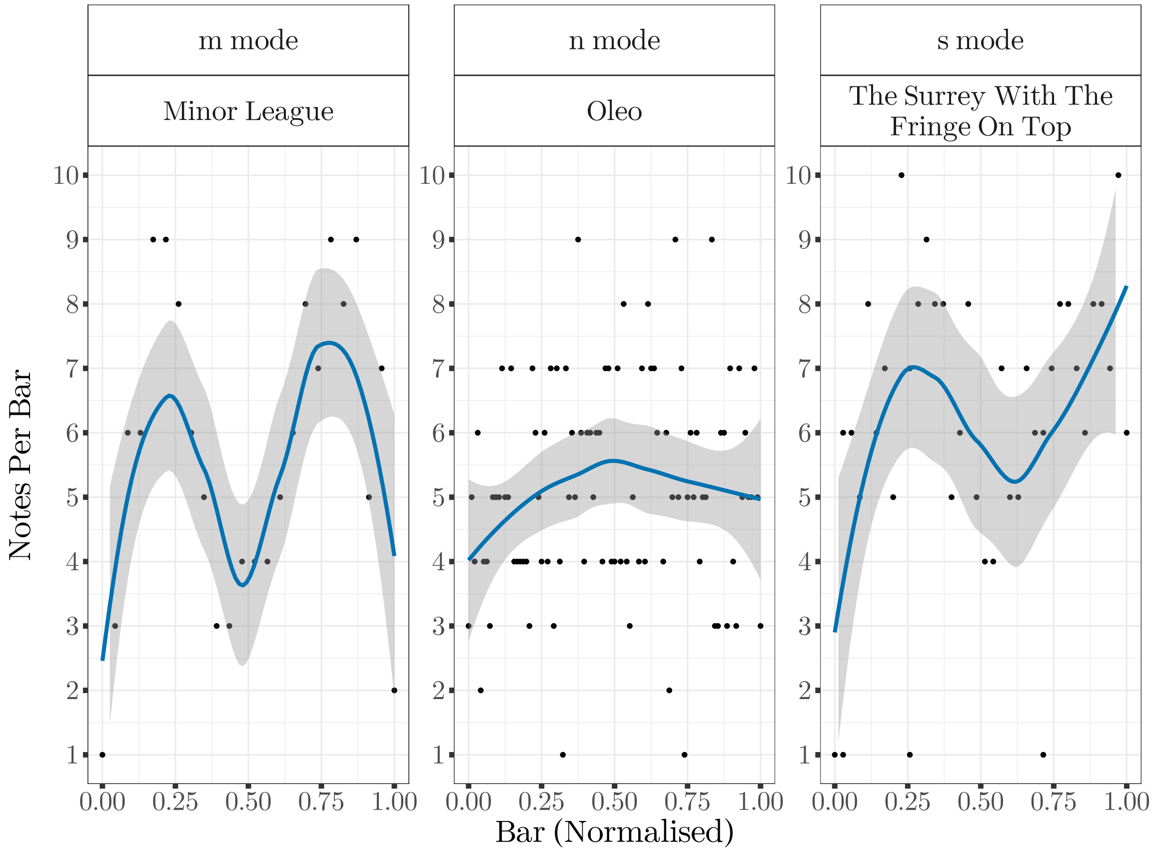

Figure 8.17: Examples of common metrical density mode shapes in Green's corpus.

8.2.2 Median Interval

The median interval was the only feature where - modes were common in Green's improvisations. This suggested that Green was less likely to shape aspects of his improvisation by changing the size of the intervals he played. There was also a relationship between the range of median intervals in an improvisation and the mode shape, as can be seen in Figure 8.18. Counter-intuitively, - mode improvisations tended to have a wider range of median intervals. This was explained by Green playing a wide variety of both ascending and descending intervals, which resulted in a flat trend of median intervals. For example, *Sonnymoon for Two* (Green 1960c) had median intervals ranging from -5 to 7.5 semitones. In comparison, *Little Girl Blue* (n mode, Green 1961l) and *Nancy (With The Laughing Face)* (s mode, Green 1962f), had smaller, and more similar ranges, from a median interval of -3 to 4 and 3 semitones respectively.

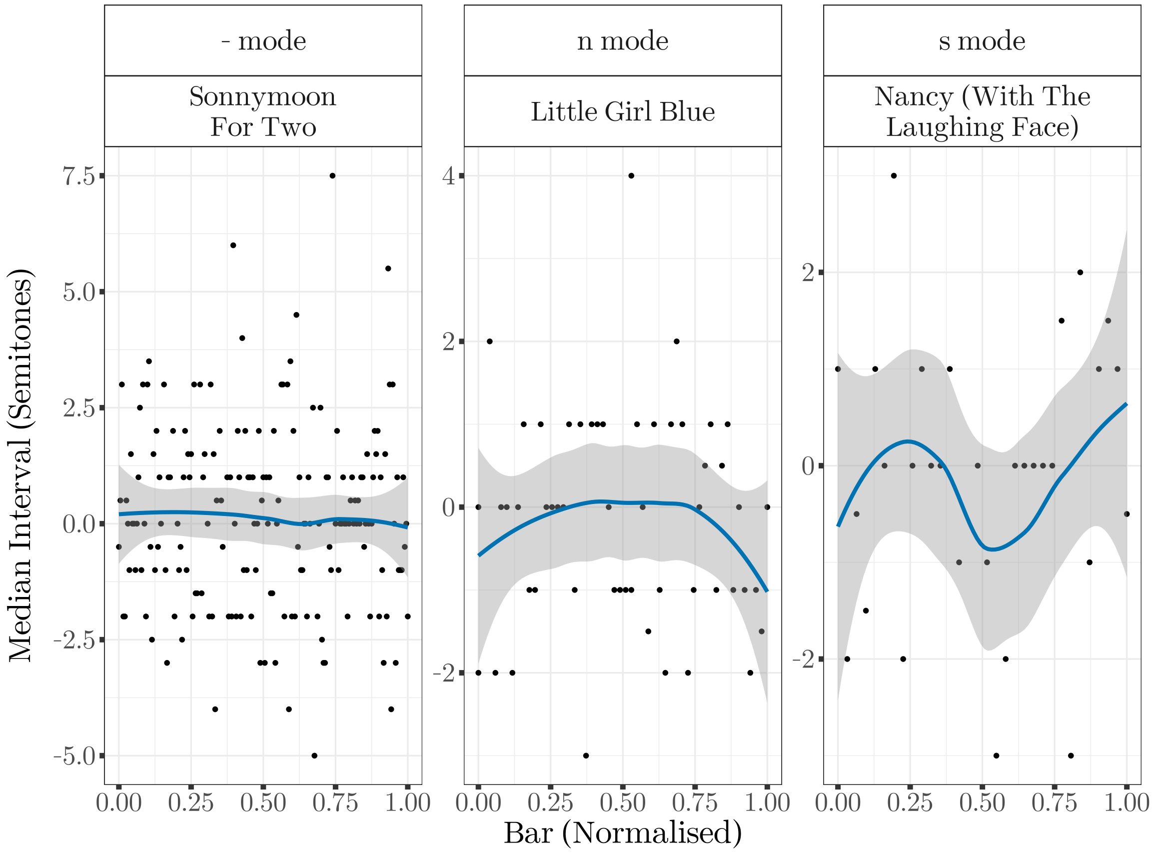

Figure 8.18: Examples of common median interval mode shapes in Green's corpus.

The infrequency of the m mode for median interval trends suggested that when Green did modulate his median interval across an improvisation, it only changed once (n mode) or twice (s mode). The median interval was the feature that Green used to shape his improvisations the least, with - modes being more frequent than in all other features combined. Although - modes indicated that Green's overall median interval size did not trend ascending or descending, they did often have a larger range than the other mode shapes. This highlighted the difference between the overall trend of a feature, and the magnitude of change within a single bar or improvisation.

8.2.3 Median Pitch

The median pitch was only one of two features, along with the proportion of NHTs, which had no - mode improvisations. As with many of the other features, the most common trend shapes were s, m, and n modes. This suggested that Green frequently used the pitch of the notes to help shape his improvisations. Of the m mode, seven began with an increasing median pitch, while the remaining three began with decreasing median pitch. For n modes, only one began with a decreasing median pitch (*Seepin'*, Solo 2, Green 1960b), while four began with a horizontal section before the n shape. For s modes, twice as many began and ended with a descending median pitch trend than those that began and ended with an ascending median pitch (twelve vs. six). Due to the wider variety of starting directions within the modes, Figure 8.19 shows two improvisations for each of the n, m, and s modes.

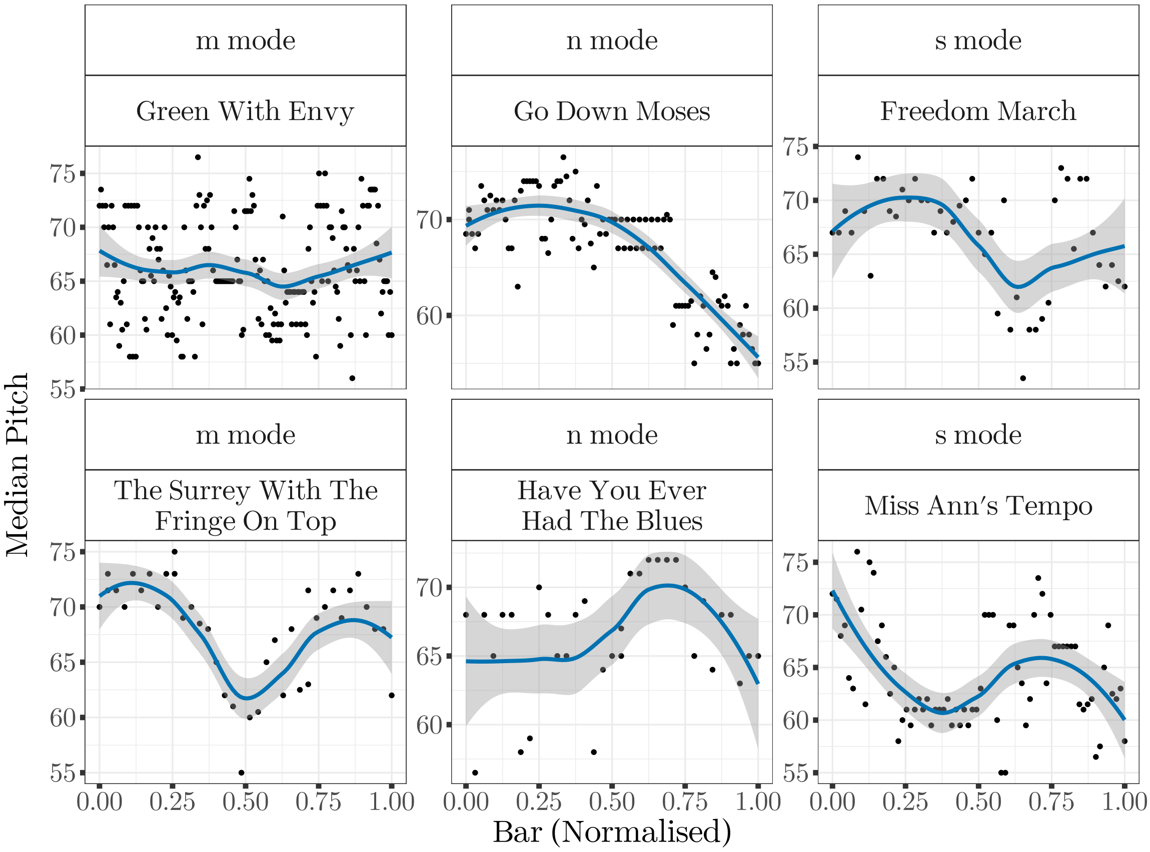

Figure 8.19: Examples of common median pitch mode shapes in Green's corpus.

Within the m mode data, *Green With Envy* (Solo 1, Green 1961j) had more noisy data than *The Surrey With The Fringe On Top* (Solo 2, Green 1963g). However, it did show an increase in the median pitch just before the middle of the improvisation. The n mode graphs showed two distinct uses of the trend shape. *Go Down Moses* (Solo 2, Green 1962d) showed Green playing the first three-quarters of the improvisation in the mid to upper registers of the guitar, with a small increase over the first quarter. The last quarter of the improvisation was then played around an octave lower. In contrast, the data from *Have You Ever Had The Blues* (Green 1963c) showed a wide variety of median pitches over the first half of the solo, resulting in a horizontal section, with the structured increase and decrease of median pitch occurring over the last half of the improvisation. The s mode graphs (*Freedom March*, Solo 2, Green 1961d) and *Miss Ann's Tempo* (Solo 2, Green 1961m) showed the two possible versions of the trend, with the bottom graph representative of the most common s mode median pitch trends. All of Green's improvisations showed evidence of Green changing his median pitch throughout the solo to aid in shaping his improvisations. The majority of Green's improvisations also tended to have multiple transitions from low to high median pitches (s and m modes).

8.2.4 Proportion of Non-Diatonic Tones

In comparison to the previous features, the NDT proportion data had more noise, caused by the fact that 45.75% of bars had zero NDTs. The majority of improvisations followed an s mode trend, with eleven of these beginning and ending with a decreasing proportion of NDTs. Of the m and n mode improvisations, only one of each began with decreasing proportion of NDTs and finished with an increasing proportion. This data indicated that many of Green's improvisations began with an increasing proportion of NDTs, with their proportion decreasing towards the end of a solo. Figure 8.20 shows an example for each of the most common modes for the proportion of NDTs. n mode *I'll Remember April* (Green 1961k) had bars that covered the entire NDT proportion range (0–1). In comparison, m mode *Red River Valley* (Green 1962j) and s mode *Sunday Mornin'* (Solo 1, Green 1961s) had a lower maximum NDT proportion, at $0.\overline{66}$ and 0.75 respectively.

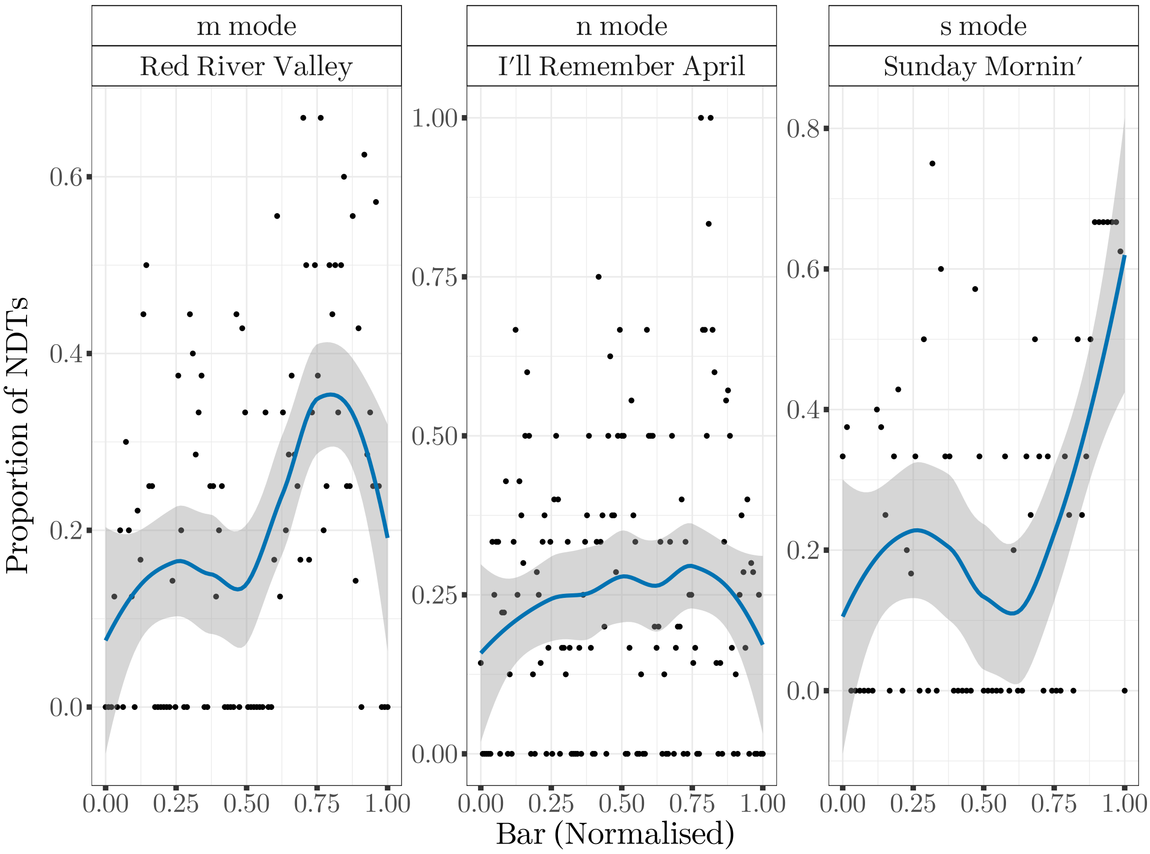

Figure 8.20: Examples of common proportion of NDT mode shapes in Green's corpus.

The n mode graph contained more deviations compared to the more consistent curves seen in previous features. The main trend was an increase in the proportion of NDTs up to around three-quarters through the improvisation, before decreasing the proportion towards the end of the solo. The m and s mode improvisations had a similar overall trend, with the main difference being that in the m mode Green decreased the proportion of NDTs towards the end of the solo, while in the s mode the proportion of NDTs increased. In summary, as many of Green's improvisations had a high number of bars without any NDTs, there was substantial noise in the trend data. However, there was still evidence of Green increasing and decreasing the proportion of NDTs across an improvisation to generate sections of tension and release.

8.2.5 Proportion of Non-Harmonic Tones

Similar to NDTs, 36.66% of Green's bars contained zero NHTs. As with the other features, s mode trends were most common, followed by m and n modes. Five of the eleven m modes started with a decreasing proportion of NHTs, while only one of the n modes did (*The Surrey With The Fringe On Top*, Solo 1, Green 1963g). Of the s modes, eight began and ended with bars with an increasing proportion of NHTs, while the other half began and ended with a decreasing proportion. Figure 8.21 shows an example of each of the three most common mode shapes. The s mode improvisation, *Idle Moments* (Green 1963e), had the smallest range, with the highest proportion of NHTs in a bar at 0.55. In comparison, the n mode improvisation, *I'm An Old Cowhand* (Green 1964b), had bars that covered the entire 0–1 range of NHT proportions, while the maximum for the m mode improvisation, *Our Miss Brooks* (Green 1961o), was 0.86.

The n mode *I'm An Old Cowhand* showed that Green used the proportion of NHTs to help shape his improvisations, with a sustained increase of NHTs around the middle of his solo. In contrast, the m mode *Our Miss Brooks* showed multiple changes in the proportions of NHTs, with fewer played at the one-quarter and three-quarter points of the improvisation. Finally, although the *Idle Moments* s mode's data had substantially more noise, the line of best fit did indicate a slight s mode trend. These graphs showed evidence of Green creating tension and release over the course of an improvisation by altering the proportion of NHTs in his lines.

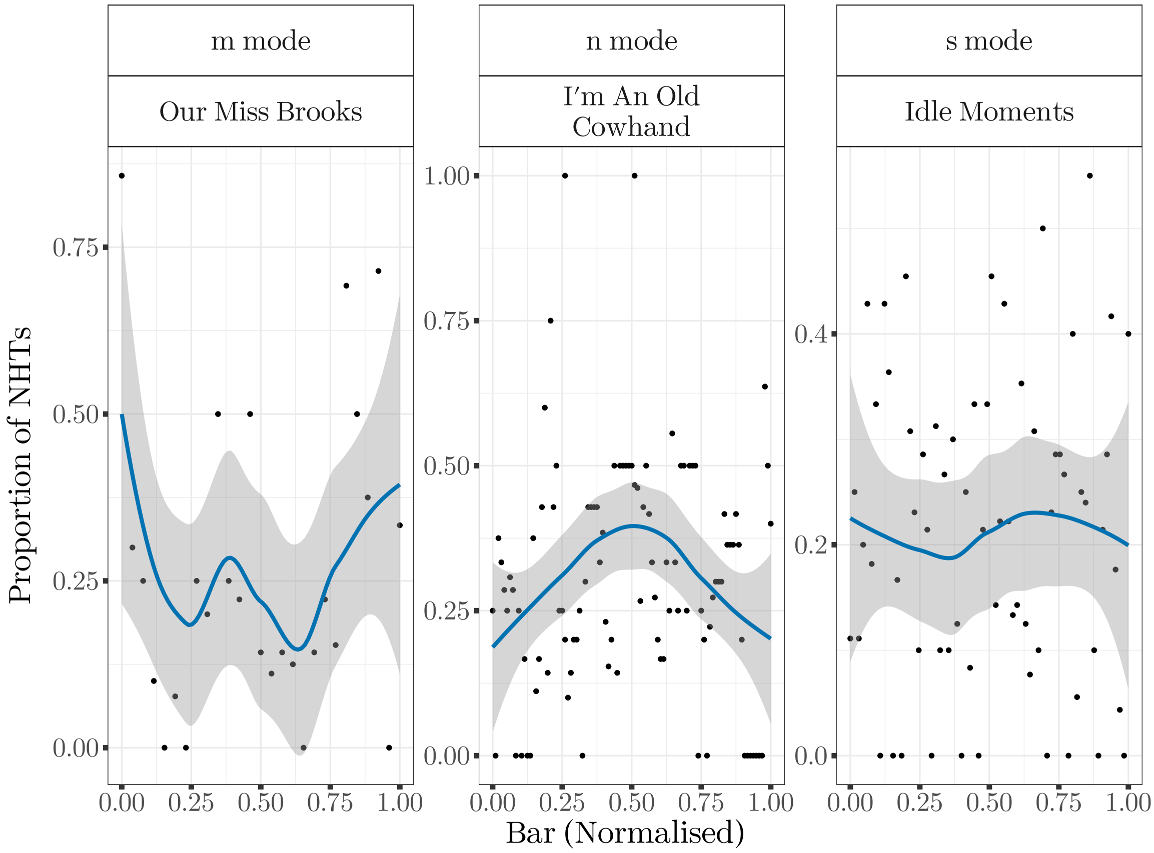

Figure 8.21: Examples of common proportion of NHT mode shapes in Green's corpus.

Large Scale Feature Trends Summary

This section focused on investigating the large scale trends of features across the course of Green's improvisations. There was not strong evidence of Green following one consistent large scale trend in features throughout his improvisations. With the values normalised across the length of a solo there were no consistent spikes or dips in the data. Instead, visual inspection of each feature over all forty improvisations found five common trend mode shapes that described the general shape of features across a solo. In the features investigated, the three most commonly found mode shapes in Green's improvisations were s, m, and n. The only feature where a horizontal, - mode, was frequently found was the median interval size. Although there was no single consistent trend in features across all of Green's improvisations, there were frequent trends found in his solos. These modes showed that Green changed features, including the metrical density, median pitch, and proportion of NDTs and NHTS, to help shape his improvisations.

8.3 Green's Improvisational Style In The Macro Domain

This chapter focused on analysing aspects of Green's improvisational features from the macro domain. The two elements that the analysis focused on were Green's phrases and larger scale feature trends.

The phrase analysis found that the majority of Green's phrases were between one and four bars long. They also tended to begin with an ascending sequence of notes, frequently starting with an ascending semitone or minor 3rd, followed by a descending trend towards the final note. Although the most common phrase contours in Green's improvisations were convex and descending, as the length of the phrase increased, the more likely it was to have a horizontal contour. Although Green had some consistent elements at the start of his phrases, including the interval and beat placement, the ends of his phrases were more varied.

The large scale feature analysis found that there was not one consistent trend that described how features changed across one of Green's improvisation. Instead, five modes were found to be able to describe the broad trend of the features. Of these, three particular modes (s, m, and n) were found to be frequently used by Green throughout his improvisations and across multiple features. These results showed that Green had a variety of approaches to using changes in features to shape his improvisations.

The analysis of Green's improvisational style in the macro domain found some consistency in his approach to phrases, especially in how he began them, while he had greater variation in how he used features to shape his improvisations.

Chapter 9

Findings of Grant Green Analysis

The chapters in Part II followed the methodology developed and presented in Chapter 3 to analyse elements of Green's improvisational style in reference to features from the pitch, rhythm, micro, and macro domains. The pitch and rhythm domains most closely followed the stated methodology, with initial analyses focused on the broad feature categories of the domains. Based on these results, more specific features, and interactions between features, were investigated. Finally, both chapters investigated two separate but interrelated examples, combining features from that domain. For the micro and macro domains, as there were fewer features, and without the same broad feature categories, these chapters were most similar to the example sections from the pitch and rhythm domain. They still followed the methodology, starting with overview analyses of each feature, followed by more specific investigations based upon these results.

The analyses in this part predominantly examined Green's solos at a corpus level, investigating the general tendencies of his improvisational style. These corpus based analyses showed the greatest advantages of the computer-aided and statistical approach to analysis over more traditional, close reading, analyses. They provided new insight into the structures of Green's improvisations, without fixating on specific examples. They also allowed for a broader investigation of Green's improvisational style to be undertaken, compared to a close reading analysis. For a close analysis, each improvisation must be manually collated for each feature, often requiring multiple passes of each transcription. In contrast, much of the time required for collating the data was only required to be completed once at the transcription stage, with the data then easily explorable. Therefore, the computer-aided and statistical approach allowed for a systematic approach to analysing Green's improvisational style between 1960 and 1965 by following the developed methodology. While the computer-aided and statistical approach most greatly benefited the corpus-style analyses, it did also allow for detailed investigations of specific feature interactions and examples. Future research could take a two-pronged approach, beginning with a corpus analysis through computer-aided and statistical methods, with the results used to inform a close reading analysis.

Throughout the analyses presented in Part II, many features were investigated. Of these, only a small number could be directly extracted from the transcriptions without any processing; for example, the pitch class or duration. Many of the features only required minimal data manipulation, with much of this already completed by *MeloSpy*; for example, comparing the pitch class to the key (TPC) or chord (CPC), or comparing the notes placement in regards to the surrounding beat structure (metrical weight). There were also features that only required the minimal data manipulation that was completed by *MeloSpy*, but had to be recreated to ensure correct results and data reproduction; for example, swing. Finally, there were features that either required little data manipulation, but not available from *MeloSpy* (rests), or required more complicated data manipulation (USTs). The analyses presented through Part II focused on those features that were most useful for investigating Green's improvisational style. As discussed in Chapter 3, Approaching A New Methodology, following the presented methodology for a different improviser or time period, would likely result in a different set of features analysed. Due to constraints on this thesis, only limited examples of feature investigations that did not return insight into Green's improvisational style could be presented, e.g. USTs. The included features were selected due to their frequent appearance in prior literature, or to serve as examples that could be applied to other performers. Although some features acted as both dependent and independent variable (e.g. $\text{CPC}_{\text{Weight}}$ or metrical weight) there were few truly independent variables that were useful for describing differences in Green's improvisational style. For pitch based features, the tonality mode and chord type were most useful, while for rhythmic features, the binary tempo range was most useful, with the phrase position also frequently influencing Green's playing. More detailed tempo features would have been more broadly useful if not for the uneven distribution of Green's tempo data. Many of the broad feature categories presented in Green's analyses would be investigated in nearly all performers, as they form part of the general musical foundation. The included features and analyses should be taken as an indication and example of what could be undertaken in future research.

The results of the analyses into Green's improvisational style found that, generally, his improvisations fit within the expectations of the era and styles in which he improvised. While Green was a highly respected improviser, between 1960 and 1965 he was not a pioneer of new jazz styles. The results of the analyses found unique elements of Green's improvisational style while also setting a baseline to compare future analyses against. Although limited, where there were specific, testable, statements regarding Green's improvisational style from previous studies. In general, this research often supported those previous statements.[1] The main results of the analysis into Green's improvisational style can be found in Figure 9.1.

In summary, the results of the analyses were able to explain many facets of Green's improvisational style between 1960 and 1965. Although the findings did not indicate that Green fundamentally changed jazz improvisation, they provided an insight into the unique elements of Green's improvisational style. The following part, Part III, built upon the results of these analyses to undertake a performer classification and comparative analysis.

[1] Due to constraints on this document – that needed to focus on facets of Green's improvisational style which were identifiable – counter examples for previous Green studies that could not be replicated or were not supported by the analyses were not presented.

Pitch Domain

- Favoured playing diatonically
- Strong evidence of blues influenced language
    - Favoured ♭3 over TT
- Majority of notes were arpeggio tones
    - Vast majority were harmonic tones
- NHTs were most often played off the beat
    - NHTs more frequent over 7 chords
    - NHTs more frequent in beats leading to a chord change
- Favoured smaller intervals
    - Step-wise and arpeggio movements most common
- Repeated notes not frequently played throughout corpus
    - Concentrated in a few improvisations
- Slight preference for descending intervals
    - Ascending intervals were larger
    - Suggested predominantly descending movement followed by ascending leaps
- SNFs were played throughout Green's lines
    - Tonic and third most commonly targeted

Micro Domain

- Tended to swing harder than many performers in the WJazzD
- Tempo had an expected limiting effect on BUR
    - Green also swung harder over blues
- Greater variation in articulation when repeating notes
    - Added interest in compensation for pitch
- Nearly three-quarters of notes were played behind the beat
    - Indicated a relaxed playing style

Rhythm Domain

- Features heavily influenced by tempo
    - Even accounting for the surrounding beat length
- Played an average of six notes per bar
    - Five notes per bar at higher tempos
    - Seven notes per bar at lower tempos
- Majority of notes had a length equivalent to semiquavers, quaver triplets, or quavers
- Lower tempos provided greater rhythmic variety and complexity of sub-beat placements
    - As the tempo increased, simplified quaver note lines were more common
- Most rests occurred between phrases
    - Average rest duration was two beats
    - Rests within a phrases were much shorter, an average of half a beat

---

Macro Domain

- Majority of phrases were between one and four bars in length
- Most common phrase contours were convex and descending
    - Longer phrases tended to be horizontal
- Note length and interval size were strongly connected to the phrase position
    - Longer notes and larger intervals tended to occur between phrases
- No single common trend found in how Green's features changed over the course of a solo
    - Five trend modes were found, with s, m, and n the most frequent modes

Figure 9.1: Summary of findings into Green's improvisational style

Part III
Performer Classification and Comparative Analysis

Chapter 10

Feature Information of Additional Performer for Classification and Comparative Analysis

Part III built upon the previous analysis of Green to undertake a classification and comparative analysis on a set of performers. Alongside Green, the other three performers whose improvisational data was used in this task were Coltrane, Parker, and Davis. These improvisers were chosen for three reasons:

1) they had the most transcriptions in the WJazzD, providing the most available data for the classification and analysis (Coltrane – 20; Davis – 18; Parker – 17);
2) they are three of the most recognisable names in jazz;
3) they are representative of at least three styles of jazz: Parker as an innovator of Bebop; Davis as an innovator of Cool jazz; and Coltrane as an innovator in avant-garde and free jazz.[1]

Although Green's data was selected to be broadly representative of his improvisations between 1960 and 1965, there was no selection criteria provided for the transcriptions in WJazzD. Therefore, any discussion of results should be considered with the caveat that the findings apply only to the data within the WJazzD, rather than to a performer's style more broadly.

The main aim for Part III was to use ML algorithms to train models that could successfully classify performers based solely on their improvisational data. The feature importance results from these models was subsequently used as the basis of the example comparative analysis. As the goal was primarily musicological in nature, only interpretable ML algorithms were used. To this end, three tree based algorithms were used for the classification task, all implemented with the `caret` package (Kuhn 2008). These were implementations of C4.5 (C4.5-like) and C5.0 decision trees (C5.0, originally developed by Ross Quinlan), and random forest (RF).

[1]There was overlap between these performers, and they were not confined to a single style.

The information in Table 10.1 shows a summary of the WJazzD data available for Coltrane, Davis, and Parker. The data in these tables presented the important data shown in the tables for Green in Chapter 4, Feature Information from Grant Green's Transcriptions.

Table 10.1: Summary Information for Coltrane, Davis, and Parker in the WJazzD.

	General (No. of)			
	Notes	Bars	Chords	Phrases
Charlie Parker	5672	715	942	311
John Coltrane	19428	3626	4142	1215
Miles Davis	6392	1348	1543	510

	Mode			Pitch Range		Time Signature	
	Major	Minor	Other	Min	Max	3/4	4/4
Charlie Parker	15	2	-	D3	Ab5	-	17
John Coltrane	15	5	-	F2	Eb6	2	18
Miles Davis	14	-	4	E3	Gb6	-	18

	Tempo Class					
	Medium Slow	Medium	Medium Up	Quick	Up	Fast
Charlie Parker	4	2	4	4	2	1
John Coltrane	1	2	7	1	3	6
Miles Davis	2	7	2	1	-	6

	Recording Year		
	1940s	1950s	1960s
Charlie Parker	13	4	-
John Coltrane	-	11	9
Miles Davis	-	10	8

The top table shows general count information about the transcriptions. Coltrane's data contained nearly as many note events as Green's, but from half as many improvisations. This suggested either that Coltrane's solos were longer on average, or that he played denser improvisations. Davis and Parker, with two and three fewer transcriptions respectively, had only approximately a quarter of the note events of Coltrane. The information related to the numbers of bars, chords, and phrases with at least one note event, were proportional to the total number of note events in each performer's data.

Most of the data presented in the other tables were excluded from the classification task. This was either because too much data was condensed in a single level (e.g. time signatures), or the features were not musical in nature (e.g. recording year). The raw pitch was not included in the comparative analysis as it could be used to identify the instrument and therefore the performer. Instead, the scaled NITP was selected as an input feature. In a task with more performers, or if the classification of the performers was the main goal, a selection of these excluded features could be used. For this research these features would have simplified the classification task to the degree where it would provide no musicological insight into the differences in improvisational style.

The mode data showed that the majority of improvisations were played in a major key, with the four 'Other' transcriptions from Davis: *Bitches Brew* (Solo 1 and 2, labelled C Chromatic in the WJazzD); and *Agitation* and *Dolores* (not given a key in the WJazzD). The time signature data showed that the vast majority of improvisations were played in $\frac{4}{4}$, with only two transcriptions from Coltrane in $\frac{3}{4}$.

The tempo class table showed that all three performers played in a wide variety of tempos, with only the Slow class having no data. As will be discussed in the following chapter, Feature Selection, despite its influence on other features, tempo based features were not able to be used as input variables for the classification task. As many features were scaled to the surrounding beat length, part of the tempo's influence was carried into the classification task. Full details on the improvisations of the three performers, including the specific years of recording, can be found on the "Database Contents" page of the Jazzomat Research Project website (Jazzomat Research Project 2017).[2]

Prior to the training and evaluation of the ML models, the entire corpus was randomly split into two datasets, a training set and a testing set. The data was split in a way that ensured that the proportions of data from each performer was the same in both datasets. The data was split 70%/30% for the training/testing datasets, with a seed of `314159`. The training data was used in the training and evaluation of the models, while the testing set was used to test how well the models performed on unseen data.

This chapter provided background details on the distribution of non-improvisational features in the WJazzD datasets of Coltrane, Parker, and Davis. The following chapter expands upon this, discussing the selection of the features used as input variables for the ML algorithms. The model results, performance metrics and variable performance, are the presented, followed by the example comparative analysis. Part III concludes with a summary of the performer classification and comparative analysis findings.

[2]https://jazzomat.hfm-weimar.de/dbformat/dbcontent.html

Chapter 11

Feature Selection

The first step in the performer classification and comparative analyses was the selection of features to use as input variables. This was informed by prior research and the completed analysis into Green's improvisational style. From this, features were included or excluded based upon early training and evaluation, and hypotheses around which could be useful for classification. The feature selection process was critical to ensuring that trained models were accurate while still being musicologically meaningful. The features were set for each abstraction level, with the same set used for every classifier and comparison. Although there were features that would have substantially increased the classification accuracy, they would have removed all possible musicological insights. These included, the instrument being played, the time signature, or the key. If classification was the main aim, or if the number of performers being classified was larger, some of these features could be useful for reducing the complexity of the classification problem. Only four abstraction levels are discussed in this chapter – solo, phrase, bar, and note – as both bar levels were trained on the same set of features.

Table 11.1 shows the number of features used in the training of the classification models, with the rows representing the four domains from Part II and an additional row for independent variables. All categorical features had to be encoded as one-hot variables (e.g. one feature per level), with the exception of the note level. There was also experimentation required for the selection of features for each abstraction level. For example, the inclusion of tempo only occurred at the solo level, as its inclusion at the other abstraction levels caused the models to overfit by using only the tempo to classify the performers.[1]

[1]Overfitting is a process where ML algorithms tuned the parameters too closely to the training data such that it was unable to successfully classify the unseen testing data.

Table 11.1: Distribution of features selected as input variables for the ML classifiers by domain.

	Raw	Distinct	Combined	Condensed	Simplified
Pitch	164	57	22	13	4
Rhythm	71	30	18	12	5
Micro	22	8	6	3	2
Macro	23	19	19	17	4
Independent	8	8	3	2	2
Total	288	122	68	47	17

The first column shows the raw number of input features used across all abstraction levels, including duplicate variables used in multiple abstractions. The distinct column removes all of the duplicate features, returning the unique number of features from each domain. The combined column shows the count when the one-hot encoded (OHE) variables from the solo, phrase, and bar abstraction levels were re-combined into a single feature. The condensed column combines the calculated and transformed features from the non-note abstraction level with the raw features used at the note level. Finally, the simplified column shows the count when the features were categorised into their base function. Although the condensed and simplified feature names were not directly used in the training of the models, they were used when discussing the features. This was due to a condensed or simplified feature being represented in myriad ways depending on the abstraction level, or to ensure the models did not overfit. The simplified feature names were taken primarily from the sub-section headings from each domain chapter, including:

- Pitch: raw pitch features, tonal pitch class, chordal pitch class, and intervals;
- Rhythm: note length, beat distribution, metrical weight, rests, and metrical density;
- Micro: swing and note placement;
- Macro: gradient, phrase descriptors, starting phrase features, and ending phrase features.

Rests were not split between true rests and micro-gaps. The raw $\text{rest}_{\text{prop}}$ values were used at the note level. At the other abstraction levels rests were represented as the proportion of the total rest time from the total beat length of each abstraction The starting and ending phrase features were separate from the phrase descriptors as they related directly to the structure of the phrases. The two independent variables were tempo (solo level) and normalised onset (note level). There were seventeen simplified features; however, two were unique to the phrase level and two were independent variables used for one abstraction level each. Consequently, there were thirteen simplified features that could have been selected for each abstraction level. The list below shows the condensed features that comprised each simplified feature, including the phrase level specific features:

- Pitch Domain:
  - Raw Pitch: normalised pitch, octave, pitch spread;
  - TPC: TPC;
  - CPC: CDPCX, CPC_{Weight}, HT or NHT;
  - Intervals: arpeggios, chromatic intervals, fuzzy interval classes, raw intervals, Parsons;
- Rhythm Domain:
  - Note Length: Duration class, fuzzy duration class, fuzzy IOI class, $duration_{BeatProp}$, $IOI_{BeatProp}$;
  - Beat Distribution: Division;
  - Metrical Weight: metrical weight, beat weight, on or off beat;
  - Rest: rest;
  - Metrical Density: notes per bar, number of empty bars;
- Micro Domain:
  - Swing: swing;
  - Note Placement: categorical note placement, onset difference;
- Macro Domain:
  - Gradient: gradient, pitch extrema;
  - Phrase Descriptors: phrase length, phrase contour, phrase position;
  - Starting Phrase Features: beat weight, CPC_{Weight}, fuzzy interval, fuzzy IOI, metrical weight, Parsons;
  - Ending Phrase Features: beat weight, CPC_{Weight}, fuzzy interval, fuzzy IOI, metrical weight, Parsons.

Some simplified features were comprised of many condensed features (e.g. intervals or note length), mainly those that had multiple representations of the same concept. Other simplified features were only comprised of one or two condensed features (e.g. metrical density or swing). Table 11.2 displays the number of distinct, condensed, and simplified features at each of the abstraction levels. This table showed that the phrase level had the most distinct, condensed, and simplified features, with the bar levels having the second most distinct and condensed features. All abstraction levels had a similar number of simplified features. The phrase level contained more features than the other abstractions due to the extra features that described the start and end of phrases. The note level did not use the metrical density simplified feature and the bar levels did not use phrase descriptors.

Table 11.2: Frequency of selected features for each abstraction level.

	Abstraction			
	Solo	Phrase	$\text{Bar}_{4\|2,2\|1}$	Note
Distinct	45	80	72	19
Condensed	19	35	25	19
Simplified	11	15	12	13

All abstraction levels had their categorical variables OHE, with the exception of the note level. The OHE created the final set of features that could be transformed to create a single row of data for each distinct abstraction (e.g. one row for each phrase at the phrase level). There were three categories of transformation:

1) measures of centre: mean, standard deviation, median, or mode;
2) proportions for OHE variables: the proportion each class contributed at that abstraction;
3) counts and descriptors: e.g. count of number of empty bars, features that began a phrase, or phrase contour.

The transformations for each condensed feature are listed below, grouped by the transformation category.[2]

Measures of Centre:

- $\text{duration}_{\text{BeatProp}}$ – mean
- $\text{IOI}_{\text{BeatProp}}$ – mean
- onset difference – mean
- division – median, mode
- raw intervals – mean, SD, median
- notes per bar – mean
- normalised pitch (NITP and *MeloSpy*) – mean
- octave – median, mode
- swing – mean, SD
- pitch – SD[3]
- tempo – mean
- gradient – mean

[2]HT or NHT and on or off beat were not listed as they were only used in the note level, where the data was not transformed.

[3]The mean and median of the pitch were excluded as they acted as a proxy for the instrument played. The standard deviation was kept as it only described the spread of pitches.

Proportions:

- chromatic intervals[4]
- arpeggio intervals[5]
- pitch extrema
- rest length
- One-hot encoded (proportions sum to 1):
    - duration class
    - fuzzy duration class
    - fuzzy IOI classes
    - fuzzy interval class
    - Parsons
    - TPC
    - CDPCX
    - $\text{CPC}_{\text{Weight}}$
    - note placement
    - metrical weight
    - beat weight

Counts and Descriptors:

- tempo class
- number of empty bars
- starting phrase features
- ending phrase features
- phrase length
- phrase shape

Below is a discussion of the simplified and condensed features that were selected for the abstraction levels. First, the features that were used consistently across the abstraction levels are discussed. This is followed by discussion of specific considerations for each of the abstraction levels.

[4]Proportion of all intervals in each abstraction that were chromatic.
[5]Proportion of all intervals in each abstraction that were thirds.

11.1 Consistent Features

There were thirteen simplified features that were used in two or more abstraction levels.[6] Phrase features were used for only the phrase and note levels. Three simplified features were used in three abstraction levels: metrical density at the solo, phrase, and bar levels; and rests and TPC at the phrase, bar, and note levels. The remaining nine simplified features (pitch, CPC, interval, note length, metrical weight, beat distribution, note placement, swing, and gradient) were used in all of the abstraction levels. These features were used in all abstraction levels because they contained representations of the building blocks of music.

There were forty-seven condensed features, of which twenty-three were used in only one abstraction level (including the two independent variables), five in two abstractions, eleven in three abstractions, and eight in all four abstraction levels. The distribution of these features across the abstractions can be seen in Figure 11.1, with features used in one or two abstractions on the left and features used in three or four on the right. Fifteen of the twenty-three condensed features that were only used in one abstraction were related to the phrase level, with twelve of these combined into the Start and End Phrase labels. The note level was the only non-phrase abstraction that included a phrase feature (phrase position). This data showed that features that were most fundamental to describing music were more frequently selected.

The similarity between some of the condensed features highlighted the myriad ways musical details could be understood and encoded. The inclusion of a particular representation at one abstraction level and not another, or the exclusion of a feature altogether, was predominantly due to overfitting during early model training and evaluation. When a feature was found to be overfitting for a particular abstraction level, a different encoding of the feature was explored. If the overfitting continued, the feature was then removed entirely. The aim was to include as many features as possible, to allow the models the best chance for classifying the performers and returning musicological information.

There were two suspected causes of overfitting for most of the features. First was a lack of available data, which affected features including the tempo. As discussed throughout Part II, there were gaps in the tempo range where there were no improvisations by Green. This was then multiplied through the four performers, providing the ML algorithms data that could classify the performers based solely on

[6] Of the seventeen total simplified features, two were specific to the phrase level and two were independent variables.

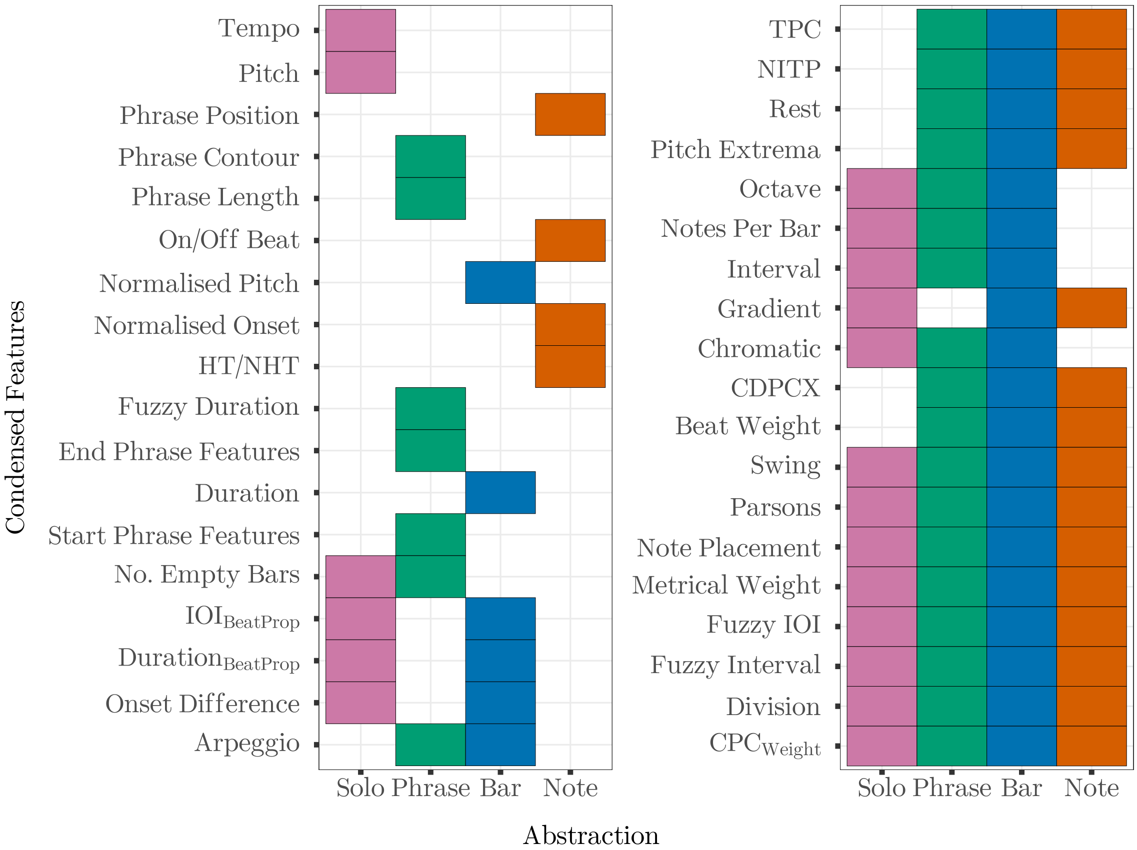

Figure 11.1: Condensed features that were selected as input variables for each abstraction. Left: one or two abstractions. Right: three or four abstractions.

the tempo of the improvisations.[7] The distribution of tempos for each performer can be found in Figure 11.2. The other suspected cause was differences in the transcription process between the data transcribed for this research, and that transcribed by the Jazzomat Research Project for the WJazzD. This applied predominantly to features in the micro domain (e.g. articulation). The distribution of raw articulation values for the four performers is shown in Figure 11.3. This data showed that although each of the performers had a different distribution of articulations, Green's data was substantially different from the others. These differences in articulation caused the models to overfit, resulting in the complete removal of the feature. Occasionally, similar representations of a feature were included at the same abstraction level. This was to allow the ML algorithms to determine which representation, or combination thereof, best classified the performers. The repetition of similar features mainly applied to those that were represented as both OHE variables and measures of centre.

[7]The binary tempo range used throughout Part II could not be used as the limit (170 BPM) was selected specifically for Green's data. For example, 77.97% of Coltrane's note events were in the BPM > 170 range.

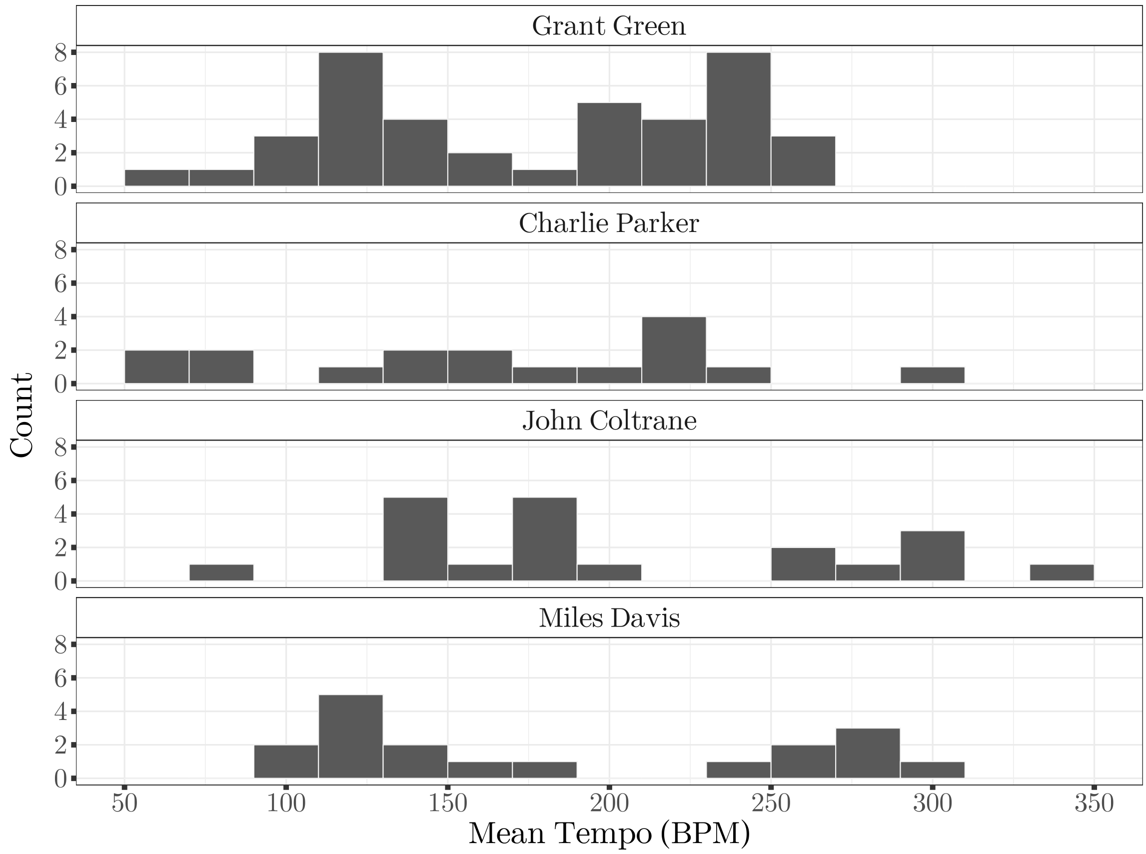

Figure 11.2: Distribution of tempos for all performers in classification task.

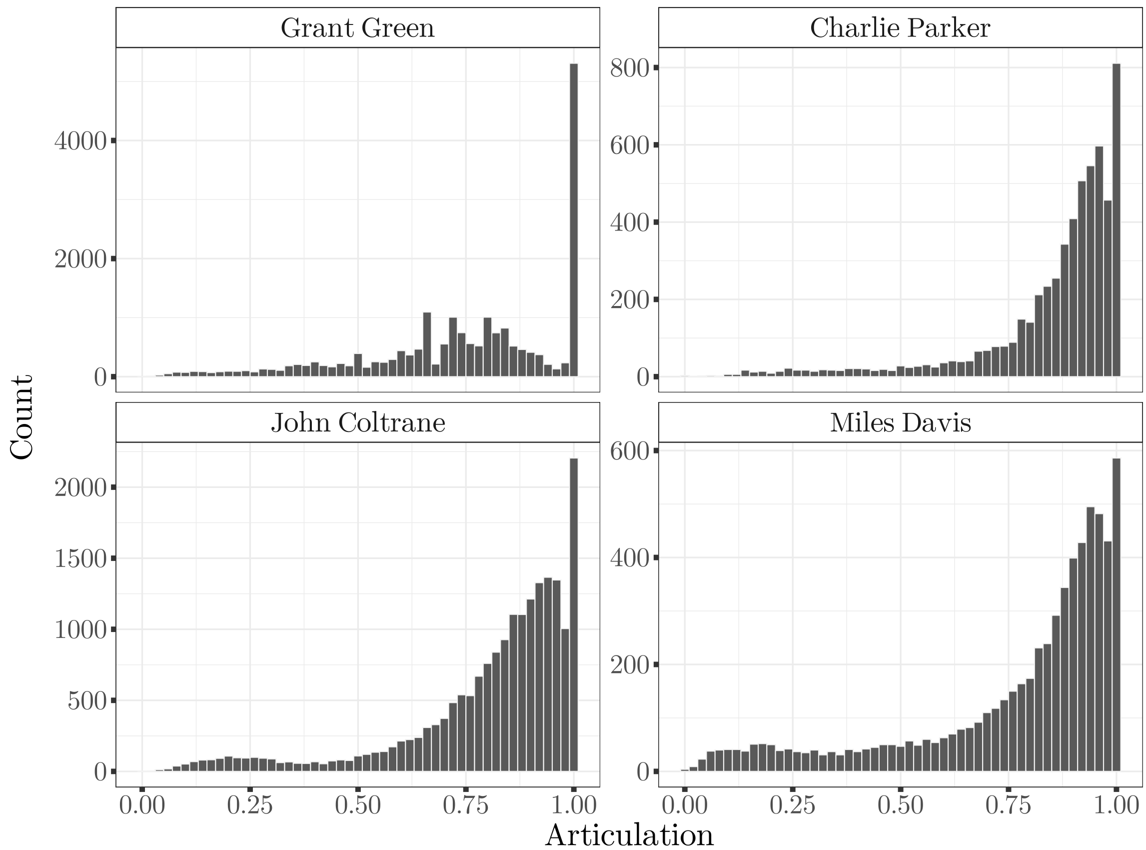

Figure 11.3: Distribution of raw articulation values for all performers in classification task.

11.2 Solo Level

The solo level was most difficult to select features for, as they required a single informative data point for an entire improvisation. Twenty features were transformed and calculated, resulting in forty-five data points for each improvisation, including: six OHE features (fuzzy intervals, fuzzy IOI, $\text{CPC}_{\text{Weight}}$, Parsons, note placement, metrical weight); and two features with measures of centre, mean and SD (interval and swing). The solo level also had many features with multiple representations, including: intervals; IOI; and note placement.

The mean and median pitch were removed after early training suggested they were acting as a proxy for the instrument played. The articulation was removed as it was overfitting. The models were trained in order of abstraction level (solo, phrase, bar, note). Features that were found to overfit at one abstraction and were expected to overfit at the other abstraction were completely excluded. Consequently, the mean pitch, median pitch, and articulation were not included in any of the other abstractions. The solo level was the only abstraction level that kept any representation of the tempo, which was likely due to the small datasets at the solo level.

11.3 Phrase Level

The phrase abstraction level considered any phrase with three or more notes, unlike the phrase analysis in the Macro Domain that only analysed phrases with four or more notes. This change was to increase the sample size for the other performers, specifically Davis, where the inclusion of shorter phrases added thirty-five phrases, an increase of 7.69%.[8]

Thirty-five features were transformed and calculated, resulting in eighty data points for each phrase. This included five descriptors for the start and end of each phrase ($\text{CPC}_{\text{Weight}}$, fuzzy IOI, fuzzy interval, metrical weight, and Parsons). The only differences between the starting and ending descriptors were that the the fuzzy interval and Parsons values were calculated heading into the final note of a phrase, instead of measuring the inter-phrase interval. The mean $\text{duration}_{\text{BeatProp}}$, $\text{IOI}_{\text{BeatProp}}$, and onset difference were all removed to stop the ML algorithms from overfitting.

A final consideration for the phrase level was how to deal with the final note of an improvisation. The final note had features that could not be calculated (e.g. IOI or

[8]Coltrane's data increased by fifty-seven phrases (+5.14%), Parker's only by three (+0.98%), and Green's by forty-five (+3.76%).

rest) as there was no following note. In other abstractions, this data was excluded when calculating means or proportions. However, as these notes were the ending note of a phrase, they could not be discarded. Therefore, the IOI feature was replaced with the equivalent duration feature for the note, and for all other features the missing values were set to `0`. This was to allow for the largest amount of data to be included in the training and evaluation process.

11.4 Bar Level

Both bar levels ($\text{bar}_{4|2}$ and $\text{bar}_{2|1}$) used the same set of features. Different sets of features could have been considered for each sliding window size. However, as they were representations of the same abstraction, it was determined that the same set of features should be used.[9] A custom function, `slideFunct`, was written to generate the data for each sliding window.[10] The full preparation code for each abstraction, including the implementation of `slideFunct`, can be found in Appendix E.3. Prior to running `slideFunct`, the necessary data was OHE, and the frequency of each feature in every bar was calculated. For continuous features (e.g. notes per bar, rests, swing) the sum of the feature was calculated, along with additional features to later calculate measures of centre. The `slideFunct` function was then applied to this data, which calculated the frequency or sum of each feature specified by the window and step size (e.g. 4 and 2 for $\text{bar}_{4|2}$). Finally, the proportions of the OHE features or measures of centre were calculated to generate the final dataset.

Twenty-five features were transformed and calculated to create seventy-two data points for each bar abstraction. This included calculating the mode of two features, the division and octave, both of which required special consideration. These considerations were required due to the smaller abstraction of the bar levels. For example, at the solo level only a single octave or division would be the mode, due to the sample size.[11] However, at the bar levels, there would often be two or more divisions or octaves that had the same frequency. For the octave feature, when there were two or more mode values, the lowest octave was selected.

The calculation for the division was more complicated. First, any abstraction where at least one mode of the division was > 8 was assigned the `9+` category. For all divisions where there were multiple mode values that had a division ≤ 8, the selection of the division mode attempted to balance the frequency of common

[9] If there were situations of overfitting at one of the bar abstraction levels and not the other, this determination would have differed.

[10] The code can be found in the Functions.R file in Appendix E.2.

[11] This would not always be the case, but was for the data used in this project.

divisions, while aiming for minimal complexity without obfuscating the true data. The selection from the multiple mode values to a single mode can be seen in the following list, in the format of 'Final Division – divisionMode1:divisionMode2, divisionMode1:divisionMode2, . . . ':

- Division 2 – 2:1, 2:3, 2:4, 2:6, 1:2:3, 1:2:4, 2:3:4, 2:4:5, 2:6:8
- Division 3 – 3:1, 3:4, 3:5, 3:6, 1:3:6, 3:4:5, 3:4:5
- Division 4 – 4:1, 4:5, 4:6, 4:8, 4:5:6
- Division 5 – 5:1, 5:2, 5:6
- Division 6 – 6:1, 6:8

There were other methods for dealing with this data and simplifying the modes. However, this approach worked for the available dataset without overfitting on these features.

11.5 Note Level

The note level was the simplest abstraction to set up, as none of the features required transformation or calculation. There were nineteen features for each note event that were passed to the ML algorithms for training. The note level was the abstraction where the most features were removed throughout the initial training and evaluation process. This was likely due to the increased size of the dataset at the note level, which caused the ML algorithms to overfit more frequently. The excluded condensed features were: phrase contour; CPC; chord type and chord weight; raw interval values; duration class and $\text{duration}_{\text{BeatProp}}$; IOI class and $\text{IOI}_{\text{BeatProp}}$; onset difference; WJazzD normalised pitch; and MCM.

Condensed features which formed part of the training data included: CDPCX; fuzzy intervals; fuzzy duration; fuzzy IOI; note placement; NITP; and division. Throughout the initial training the NITP appeared to be very important for classifying the performers. However, as the NITP allowed values outside of 0–1 for altissimo or other extended techniques, it was able to singularly identify Coltrane. This was because Coltrane was the only performer of the four with NITP outside this range. This did not affect the other abstraction levels, as the mean NITP calculated always fell within 0–1. To stop the models overfitting on this data, all notes that fell outside the 0–1 range were removed. This removed only 211 notes from Coltrane's data (1.09%), and only 0.41% of the total dataset of the four performers.

11.6 Feature Selection Summary

This chapter discussed the selection of features used as the input variables for the ML algorithms. After all categorical variables were OHE, all features were transformed to form the final training data set (except at the note level). These transformations were: measures of centre; proportions; or counts and descriptors. As many features, or transformations thereof, described similar representations, the discussion focused on the condensed and simplified versions of the features.

Features that could be considered representations of fundamental building blocks of music were selected for all abstractions. These simplified features included: raw pitch; CPC; intervals; note length; metrical weight; beat distribution; note placement; swing; and gradient. Additionally, the metrical density, rests, and TPC were all included in three of the four abstraction levels. Through initial testing and evaluation two features had to be excluded entirely as input variables: tempo based features[12]; and articulation. Both features caused the ML algorithms to overfit, resulting in models that neither performed well with new data, nor provided insight into the performer's improvisational styles.

The phrase level had the largest number of input variables due to additional features for describing the starting and ending of phrases. These features described how phrases stated and ended in relation to the $\mathrm{CPC}_{\mathrm{Weight}}$, fuzzy IOI, fuzzy interval, metrical weight, and Parsons. For the bar levels there were two features that required special consideration, octave and division. The considerations were required to find the mode of each feature, where the small sample size often resulted in multiple modes. Finally, the note level required the least amount of setup and transformations, as the raw data could be passed into the ML algorithms. However, the raw distribution of the data tended to increase overfitting with the ML algorithms. Consequently, the largest number of features were removed in the initial training and evaluation process.

In summary, the more fundamental a feature was to describing the basic building blocks of music, the more likely it was to appear throughout the abstraction levels. The exceptions to this were for features including the tempo, where the ML algorithms tended to overfit on the data. Following the selection of the features, the final versions of each of the models – one each for every combination of ML algorithm, abstraction level, and comparison – were trained.[13] The following chapter reports on the results of these models, both the performance metrics and the features found to be important in classifying the improvisers.

[12] Excluding at the solo level.
[13] Code for training the models can be found in Appendix E.3.

Chapter 12

Model Results

This chapter discusses the results of the ML models. The performance metrics of the models were reported first. Following this, the feature analysis section investigated the variable importance measures from the models. These variable importance scores aided in the identification of features that were frequently used to successfully classify the performers, and informed the example comparative analysis.

The models were trained using interpretable tree-based algorithms, and a naive approach would have been to investigate the models by visually investigating the final decision trees. A small decision tree, from the C4.5-like solo level n-way comparison, can be seen in Figure 12.1. However, the majority of models were not possible to visualise as they contained too many branches. Additionally, because C5.0 and RF classifiers used multiple trees within each model there was not a single final tree to plot. Therefore, the models were evaluated based on performance metrics extracted from the models with the aid of *R* libraries.[1]

[1] `caret` (version 6.0-86) was used for generating the confusion matrices (`confusionMatrix`), standard performance metrics from the confusion matrices, and variable importance (`varImp`). `yardstick` (version 0.0.7) was used for generating the `mcc` performance metric.

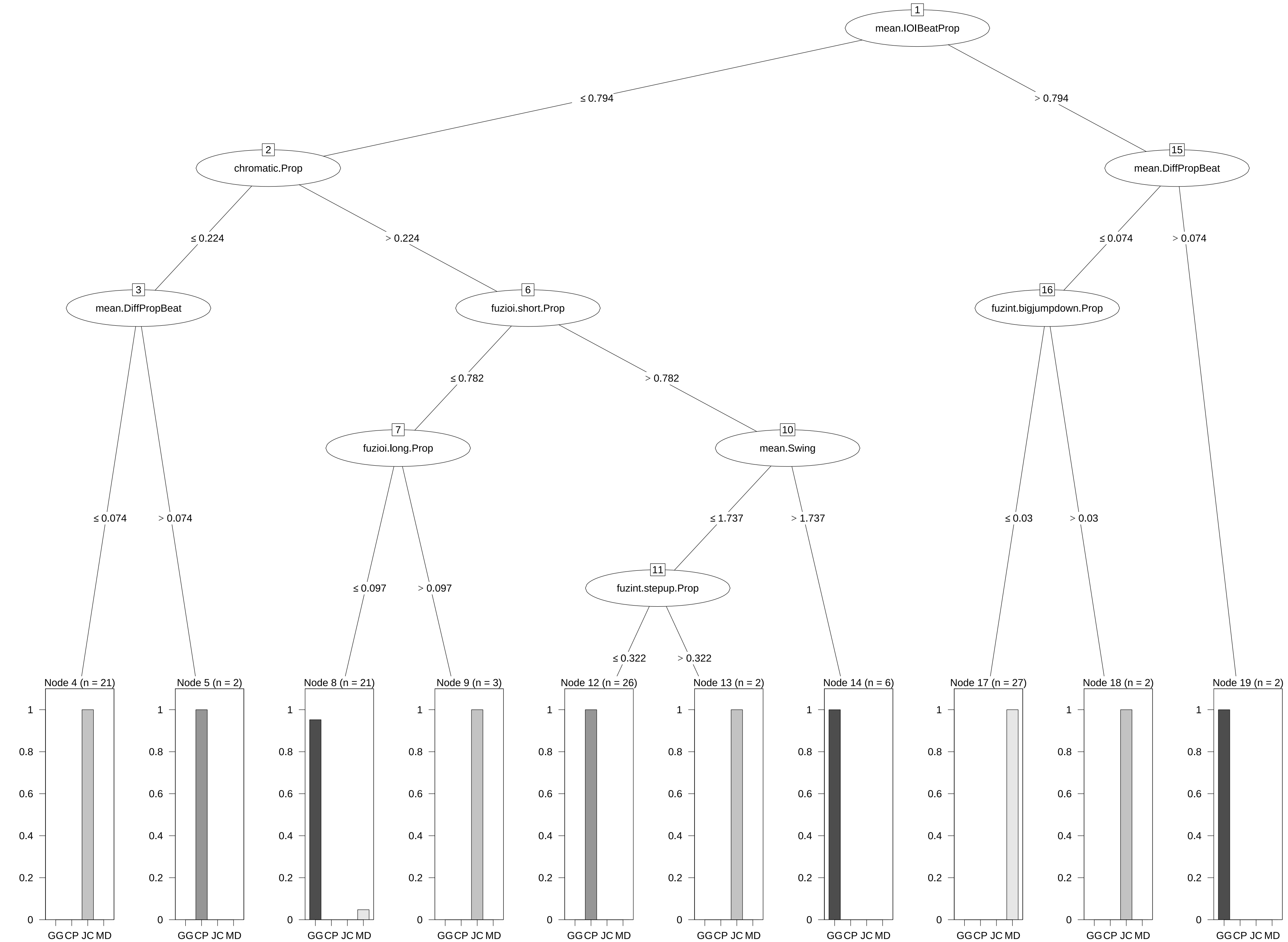

Figure 12.1: Decision tree for the n-way C4.5-like solo level model.

12.1 Model Performance Metrics

This section reported on the performance metrics of the models that had a classification accuracy significantly better ($p < .05$) than the no information rate (NIR), referred to as significant models. Within this section a summary of the performance metrics were presented. A broader set of metrics for each significant model can be found in Appendix D.[2] The NIR was the baseline accuracy if the model only predicted the majority class, and was equivalent to the proportion of the majority class in the dataset. For example, if there were ten solos to be classified, six from Green and four from Coltrane, the NIR would be $6/10 = 0.60$ or 60%. Only models with an accuracy significantly better than the NIR were investigated further with other performance metrics.

The testing data was run through each of the 165 trained models to evaluate how the models performed on unseen data. The `caret` package was then used to generate summary statistics and a confusion matrix for each model. A confusion matrix is a table that shows how each model's predicted classes matched the true reference data. For a model with only two classes (e.g. one-vs-all or one-vs-one models), the resulting 2×2 table had four segments that represented the number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN), as seen in Figure 12.2.

These four values provided the underlying data that the *R* libraries and functions used to calculated the summary statistics, with the equations shown in Table 12.1. The Matthews Correlation Coefficient (MCC), "originally developed by Matthews in 1975 … was re-proposed by Baldi and colleagues in 2000 as a standard performance metric for machine learning" (Chicco and Jurman 2020, 2). Two advantages of the MCC were that it "produce[d] a high score only if the prediction obtained good results in all of the four confusion matrix categories (true positives, false negatives, true negatives, and false positives)" (Chicco and Jurman 2020, 1) and that it was "an effective solution [for] overcoming … class imbalance issue[s]" (Chicco and Jurman 2020, 2). For these reasons the MCC was the main metric on which the performance of the significant models was evaluated. The MCC is a specific example of the *Pearson product-moment correlation coefficient*; therefore, the interpretation of the results was the same, with a value of -1 indicating complete disagreement (e.g. if every event was perfectly misclassified), a value of 0 indicating randomness, and a value of 1 indicating perfect classification. The other statistics were included for reference, and are presented in the summary performance metrics in Appendix

[2]Additionally, the confusion matrix from the testing data for all of the models can be found in Appendix E.5.

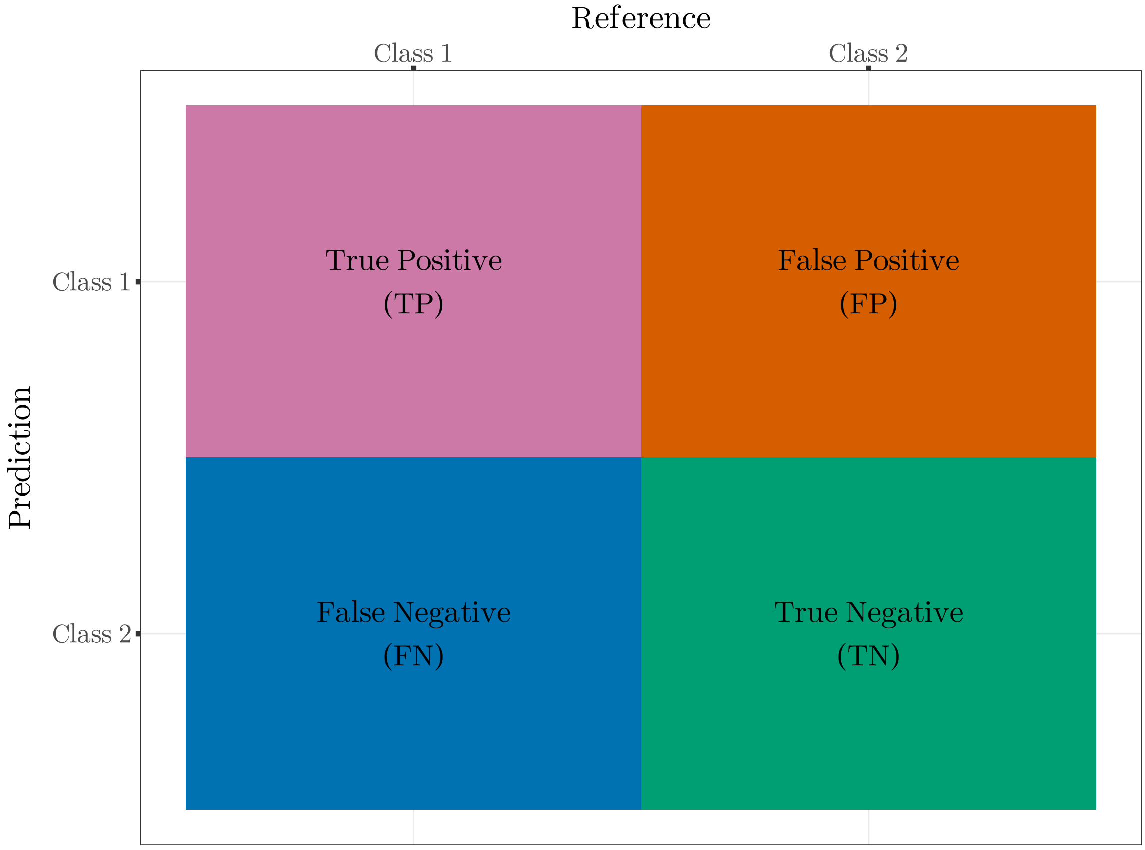

Figure 12.2: An example confusion matrix.

D. The balanced accuracy (BA) was the mean of the sensitivity and specificity, the F-score (F_1) was the geometric mean of the positive predictive value (PPV) and sensitivity, and varied depending on which class was considered the "positive" class. The scores for both classes were presented as F_1^+ for the "Positive" class and F_1^- for the "Negative" class.

Table 12.1: Equations for Summary Statistics.

	Equation
MCC	$\frac{\mathrm{TP}\times\mathrm{TN}-\mathrm{FP}\times\mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})\times(\mathrm{TP}+\mathrm{FN})\times(\mathrm{TN}+\mathrm{FP})\times(\mathrm{TN}+\mathrm{FN})}}$
Accuracy	$\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}$
Sensitivity	$\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$
Specificity	$\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}$
Balanced Accuracy	$\frac{\mathrm{Sensitivity}+\mathrm{Specificity}}{2}$
PPV	$\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$
NPV	$\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}}$
F_1^+	$2 \times \frac{\mathrm{PPV}\times\mathrm{Sensitivity}}{\mathrm{PPV}+\mathrm{Sensitivity}}$
F_1^-	$2 \times \frac{\mathrm{NPV}\times\mathrm{Specificity}}{\mathrm{NPV}+\mathrm{Specificity}}$

TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative
MCC = Matthews Correlation Coefficient
PPV = Positive Predictive Value
NPV = Negative Predictive Value
F_1 = F-score

The distribution of summary statistics for each classifier, for models which significantly outperformed the NIR, is shown in Table 12.2. This showed that the RF classifier produced the most significant models, with forty-nine of fifty-five models (89.09%) outperforming the NIR. This was followed by the C5.0 (81.82%) and C4.5-like classifiers (67.27%). The distribution of the significant models across the abstraction levels and comparisons can be seen in Figure 12.3. This indicated that the main difference between the RF and C5.0 classifiers was the number of significant one-vs-one models at the solo level. The RF classifier also produced one extra one-vs-one model (Coltrane vs. Parker) at the note level compared to the C5.0 classifier.

Table 12.2: Summary statistics for models which significantly outperformed the NIR.

	Min	Q1	Med	Q3	Max	Mean	SD
C4.5-like Tree (Count: 37)							
MCC	0.39	0.54	0.61	0.69	1	0.63	0.13
Accuracy (%)	65.05	77.84	83.23	88.60	100	82.62	7.95
Sensitivity (%)	58.90	72.40	80.99	88.06	100	79.69	11.02
Specificity (%)	67.47	78.31	83.63	90.06	100	84.19	7.59
BA (%)	69.41	77.31	81.54	85.11	100	81.94	6.73
PPV (%)	57.37	68.91	75.71	93.32	100	78.94	13.72
NPV (%)	58.66	75.86	83.58	88.28	100	81.75	10.69
F_1^+	0.58	0.70	0.77	0.91	1	0.79	0.12
F_1^-	0.63	0.78	0.83	0.88	1	0.83	0.09
C5.0 Tree (Count: 45)							
MCC	0.53	0.74	0.80	0.86	1	0.79	0.11
Accuracy (%)	70.37	89.70	92.59	95.13	100	91.77	5.22
Sensitivity (%)	38.24	81.05	89.37	97.49	100	86.57	13.11
Specificity (%)	44.00	88.33	93.57	96.78	100	89.76	11.89
BA (%)	68.72	84.71	89.44	92.90	100	88.17	6.92
PPV (%)	71.25	86.02	90.00	96.35	100	90.19	7.00
NPV (%)	79.41	88.57	92.19	94.85	100	91.74	4.59
F_1^+	0.52	0.82	0.90	0.96	1	0.88	0.10
F_1^-	0.57	0.88	0.93	0.95	1	0.90	0.09
Random Forest (Count: 49)							
MCC	0.42	0.68	0.75	0.83	1	0.75	0.13
Accuracy (%)	72.84	86.90	90.89	93.69	100	90.29	5.36
Sensitivity (%)	32.53	75.00	90.75	97.20	100	84.43	16.81
Specificity (%)	36.84	83.33	93.33	96.56	100	85.94	15.62
BA (%)	65.90	80.95	85.44	90.44	100	85.18	8.54
PPV (%)	64.74	85.84	90.47	93.47	100	89.25	6.93
NPV (%)	66.67	86.18	90.45	94.06	100	90.26	6.33
F_1^+	0.46	0.80	0.91	0.95	1	0.86	0.12
F_1^-	0.47	0.85	0.90	0.95	1	0.87	0.11

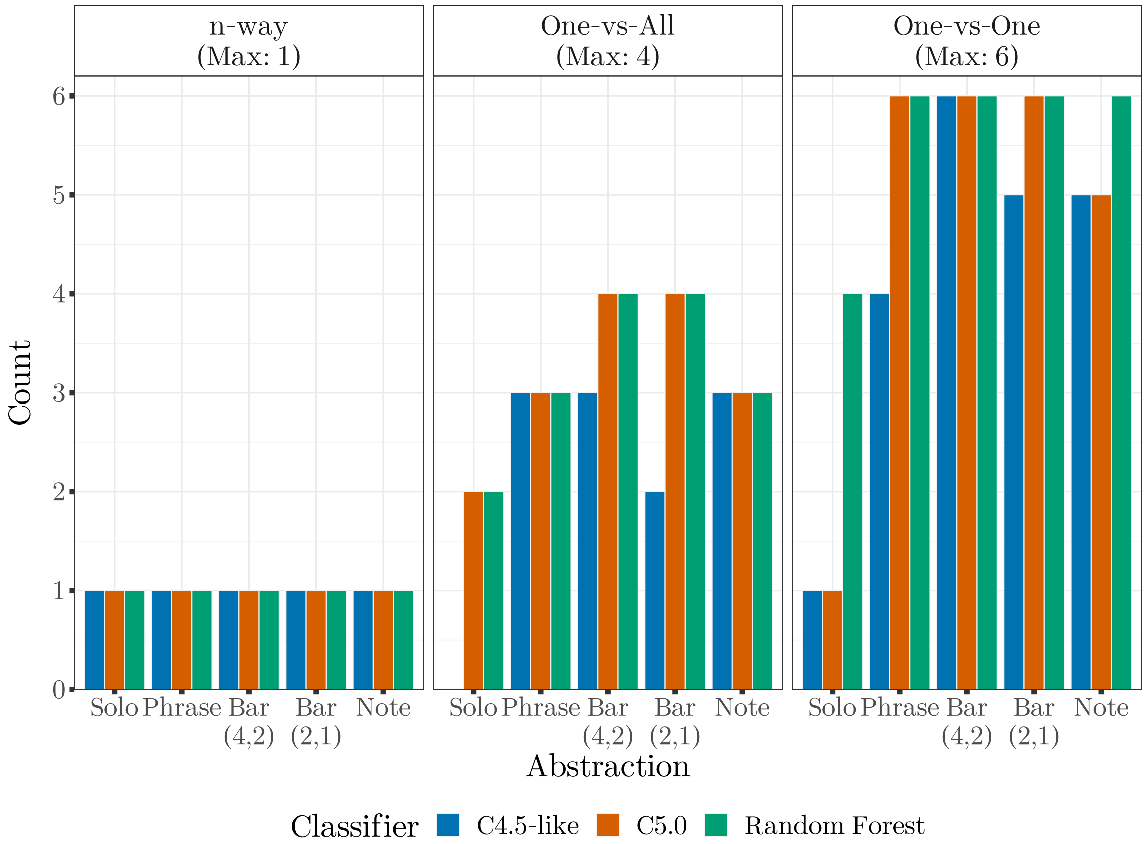

Figure 12.3: Number of models, for each comparison type, that performed significantly better than the NIR.

Although the RF classifier produced the most significant models, the summary statistics indicated that the C5.0 produced the best performing models. Across the performance metrics, the C5.0 classifier had marginally, but consistent, better results compared to the RF. These results indicated that both the C5.0 and RF classifier produced many highly performant models. In comparison, the C4.5-like classifier produced fewer models than the other classifiers and models with substantially lower performance metrics.

An additional consideration in the performance of the classifiers was the time it took to train each model. The median training time for the C5.0 algorithm was the fastest, 2 min 31 sec, followed by the C4.5-like at 16 min 29 sec, while the RF algorithm was the slowest with a median training time of 37 min 57 sec. The time tended to scale with the amount of input data and number of variables. The solo abstraction level was the quickest to train (median: 2:44), while the $\text{bar}_{2|1}$ was the longest (median: 28:42), the other abstractions had a median time around fifteen minutes (phrase: 15:18; $\text{bar}_{4|2}$: 16:30; note: 15:11). Although the C5.0 classifier's performance metrics only slightly outperformed the RF, the C5.0 classifier was able to produce better results in approximately one-fifteenth of the time.

The majority of the situations (combination of abstraction and comparison) where only the RF classifier produced a significant model were at the solo level. Due to the small sample size at the solo level, small changes in categorisation had a substantial impact on the accuracy of the models. For example, the RF Coltrane vs. Davis solo level comparison had one misclassification while the C5.0 only had two. As there were only eleven testing data points, only the result from the RF model was considered significant. The small sample size also resulted in inflated performance metrics for solo level models. The other abstractions, with around thirty significant models each (phrase: 28; $\text{bar}_{4|2}$: 32; $\text{bar}_{2|1}$: 30; and note: 28) had mean MCC values just over 0.70.[3] In contrast, the solo abstraction, with only thirteen significant models, had a mean MCC of 0.86. The small sample size of significant models, and the small testing datasets, meant that although they had the highest mean MCC, the performance metrics were inflated. Consequently, as the results of the solo level were inconclusive, they were excluded from all further analyses.

Figure 12.4 shows an example for why a certain degree of misclassification, even for well performing models, was expected. This figure shows two musical examples that were misclassified by the C5.0 classifier in the Green vs. Davis phrase level model. The top musical example shows the 14$^{\text{th}}$ phrase from Davis' improvisation over *Eighty-One* (Jazzomat Research Project 2017), which was incorrectly identified as Green. The bottom musical example is the 45$^{\text{th}}$ phrase from Green's improvisation over *Sonnymoon For Two* (Green 1960c), which was misclassified as being played by Davis. Although these two examples were played over different harmonies, there were similarities. For example, in the rhythmic structures of the lines, and that both phrases ended on the 7$^{\text{th}}$.

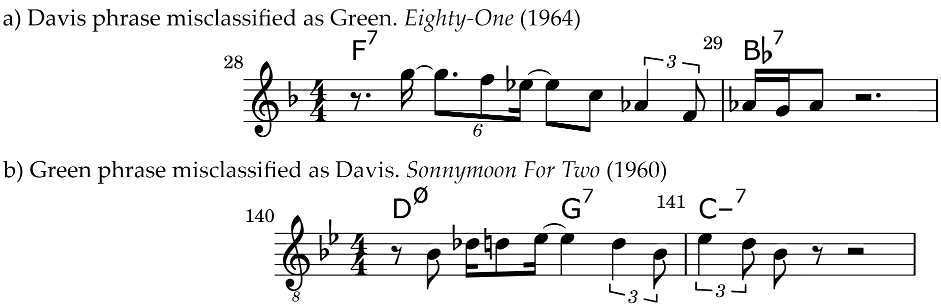

Figure 12.4: Phrase misclassifications between Green and Davis. a) Phrases from Davis misclassified as Green, *Eighty-one* (1964), bars 28–29. b) Phrase from Green misclassified as Davis, *Sonnymoon For Two* (1960), bars 140–141.

The following sub-sections discussed the MCC results for each of the three classifiers. The discussion focused on the significant models from the phrase, $\text{bar}_{4|2}$, $\text{bar}_{2|1}$, and note levels.

[3] MCC $\bar{x}$ = phrase: 0.72 ± 0.14; $\text{bar}_{4|2}$: 0.72 ± 0.11; $\text{bar}_{2|1}$: 0.71 ± 0.12; and note: 0.71 ± 0.15.

12.1.1 C4.5-like Results

Table 12.3 shows the MCC for each significant model trained by the C4.5-like algorithm. The abstraction with the highest number of significant models was the bar$_{4|2}$ level, while the phrase level had the highest mean MCC (0.63). This was influenced by the higher MCC for Coltrane vs. Davis (0.89), Green vs. Davis (0.81), and Davis vs. All (0.70). These results suggested there were identifiable elements of Davis' improvisational style at the phrase level. For both of the bar levels, the C4.5-like models tended to have the highest MCC for comparisons that included Coltrane, with the best bar models being Coltrane vs. Davis (0.75 and 0.80). Comparisons including Green tended to perform best a the note level, with Green vs. Coltrane having the highest note level MCC (0.80). The C4.5-like classifier was not able to successfully classify Parker. Parker vs. All (all abstractions), Green vs. Parker (phrase and bar$_{2|1}$ levels), and Coltrane vs. Parker (phrase and note levels) models did not classify significantly better than the NIR. The significant models that included Parker also tended to have the lowest MCC.

Table 12.3: MCC results for C4.5-like models.

	Phrase	Bar$_{4\|2}$	Bar$_{2\|1}$	Note
n-way	0.54	0.53	0.49	0.67
One-vs-All				
Green vs. All	0.47	0.53	0.52	0.59
Coltrane vs. All	0.62	0.61	0.66	0.54
Davis vs. All	0.70	0.57	-	0.61
Mean	0.60	0.57	0.59	0.58
One-vs-One				
Green vs. Coltrane	0.47	0.67	0.66	0.80
Green vs. Parker	-	0.55	-	0.61
Green vs. Davis	0.81	0.55	0.60	0.70
Coltrane vs. Parker	-	0.66	0.59	-
Coltrane vs. Davis	0.89	0.75	0.80	0.71
Parker vs. Davis	0.58	0.67	0.43	0.39
Mean	0.69	0.64	0.62	0.64
C4.5 Mean	0.63	0.61	0.59	0.62

The best performing n-way model was the note level, with the other models performing below the average for each abstraction. The n-way comparisons struggled most with the classification of Parker. The most frequent misclassifications in the n-way models occurred between Parker and Green or Coltrane.[4] At the note level, most of the misclassifications were between Parker and Coltrane. This suggested that the similarity in instrumentation between Parker and Coltrane resulted in more difficulty in differentiating note based features. At larger abstractions Coltrane and Green's improvisational styles were most similar to Parker's, with both performers drawing influence from his improvisations as one of the fathers of bebop. The C4.5-like classifier struggled to classify Parker both due to his enduring influence throughout jazz improvisation, as well as his comparatively smaller sample size in the WJazzD. In comparison, Davis' improvisations, especially at the phrase and note level, were distinct from the other performers. Overall, the C4.5-like classifier's performance metrics indicated that although there were many misclassifications it was able to frequently, and fairly accurately, classify the performers.

12.1.2 C5.0 Results

The MCC results for all significant C5.0 models is shown in Table 12.4. Following from the summary results in Table 12.2, the C5.0 classifier outperformed the C4.5-like classifier in every model. The best C5.0 model was Green vs. Coltrane at the note level (MCC: 0.98). There were only nine misclassifications from the 724 data points in the testing dataset, with the misclassifications split between both performers. There were only three C5.0 models that did not significantly outperform the NIR, and all included comparisons of Parker: Parker vs. All at the phrase and note levels; and Coltrane vs. Parker at the note level. These results suggested that, as with the C4.5-like classifier, the C5.0 classifier also had the most difficulty separating Parker from the other performers. The only models with an MCC less than 0.70 were for models that included Parker as a main point of comparison (excluding n-way comparisons).

Similar to C4.5-like models, the best C5.0 models that included Davis were at the phrase level, while those that included Green performed best at the note level. The best performing C5.0 models that included Coltrane were at the bar and note levels. Overall, the best performing C5.0 models tended to be at the note level (MCC: $\bar{x} = 0.88 \pm 0.07$), although two note level models were not significant (Parker vs. All and Coltrane vs. Parker). The second-best performing abstraction was the bar$_{4|2}$ level (MCC: $\bar{x} = 0.79 \pm 0.08$), with all trained models significantly outperforming the NIR.

[4]Performer based misclassifications in the n-way models can be seen in the confusion matrix data in Appendix E.5.

Table 12.4: MCC results for C5.0 models.

	Phrase	$Bar_{4\|2}$	$Bar_{2\|1}$	Note
n-way	0.75	0.81	0.77	0.84
One-vs-All				
Green vs. All	0.70	0.78	0.78	0.96
Coltrane vs. All	0.74	0.85	0.85	0.81
Parker vs. All	-	0.62	0.53	-
Davis vs. All	0.82	0.74	0.76	0.93
Mean	0.75	0.75	0.73	0.90
One-vs-One				
Green vs. Coltrane	0.80	0.89	0.86	0.98
Green vs. Parker	0.55	0.73	0.72	0.82
Green vs. Davis	0.94	0.82	0.81	0.88
Coltrane vs. Parker	0.59	0.78	0.77	-
Coltrane vs. Davis	0.93	0.88	0.90	0.91
Parker vs. Davis	0.80	0.77	0.66	0.79
Mean	0.77	0.81	0.78	0.88
C5.0 Mean	0.76	0.79	0.76	0.88

These results indicated that there were strongly identifiable elements of Davis' improvisational style at the phrase level, with additional note level features that aided in identification. In contrast, the classifier struggled more with Coltrane at the phrase level, suggesting his improvisational style at the phrase level was more similar to that of Green and Parker's. Compared to the C4.5-like classifier's middling performance on Green, the C5.0 classifier achieved very high MCC when classifying Green, especially at the note level. The classification of Parker improved with the C5.0 classifier, but was still the most difficult performer to classify, with the most misclassifications. Overall, the C5.0 performed substantially better than the C4.5-like classifier, and was consistently able to successfully identify the soloists across the comparisons and abstractions.

12.1.3 Random Forest Results

The RF classifier produced the most significant models. The only models that did not significantly outperform the NIR were Parker vs. All at the phrase and note levels. Table 12.5 shows the MCC results for the significant RF models. As the RF results were most similar to those of the C5.0, the discussion focused on the differences between these classifiers.

Table 12.5: MCC results for random forest models.

	Phrase	$Bar_{4\|2}$	$Bar_{2\|1}$	Note
n-way	0.75	0.81	0.74	0.58
One-vs-All				
Green vs. All	0.66	0.74	0.72	0.68
Coltrane vs. All	0.76	0.83	0.82	0.63
Parker vs. All	-	0.48	0.55	-
Davis vs. All	0.85	0.74	0.72	0.55
Mean	0.76	0.70	0.70	0.62
One-vs-One				
Green vs. Coltrane	0.81	0.84	0.81	0.72
Green vs. Parker	0.59	0.66	0.64	0.68
Green vs. Davis	0.94	0.78	0.79	0.71
Coltrane vs. Parker	0.57	0.76	0.70	0.42
Coltrane vs. Davis	0.95	0.85	0.88	0.75
Parker vs. Davis	0.71	0.77	0.68	0.58
Mean	0.76	0.78	0.75	0.64
Random Forest Mean	0.76	0.75	0.73	0.63

Although the overall MCC of the RF was lower than the C5.0 classifier, there were seven models where the RF had a higher MCC, including (difference in MCC): Coltrane vs. All at the phrase level (0.02); Davis vs. All at the phrase level (0.03); Green vs. Coltrane at the phrase level (0.01); Green vs. Parker at the phrase level (0.04); Coltrane vs. Davis at the phrase level (0.02); Parker vs. All at the $bar_{2|1}$ level (0.02); and Parker vs. Davis at the $bar_{2|1}$ level (0.02). All models where the RF outperformed the C5.0 were at the phrase level, except Parker vs. All and Parker vs. Davis at the $bar_{2|1}$ level. Although many of the models where the RF classifier outperformed the C5.0 classifier were at the phrase level, the mean MCC was the same for both classifiers (mean MCC: 0.76). Additionally, the RF classifier trained one additional significant model, Coltrane vs. Parker at the note level. However, this model had the lowest MCC of all significant models across the three classifiers.[5] All other RF models either had the same or lower MCC when compared to the C5.0 classifier. The RF models at the note level performed substantially worse than the C5.0 classifier. Considering the results of all three classifiers, this suggested that the C5.0 classifier performed particularly well at the note level.

Similar to the other classifiers, the RF also struggled with comparisons that included Parker. Comparisons that included Coltrane or Green tended to perform the best at the bar abstraction levels. The best performing abstractions for the RF classifier were the larger abstraction levels, phrase and $\text{bar}_{4|2}$, influenced by the high performing Davis models. Overall, the RF classifier only rarely, and marginally, outperformed the C5.0 classifier; however, the RF algorithm took fifteen times longer on average to train the models.

12.1.4 Model Performance Metrics Summary

This section reported the performance metrics from the significant models trained by the three classifiers. The MCC was chosen as it provided a single metric on which to evaluate the entire confusion matrix of each model, and also worked well on unbalanced datasets. The results found that the C4.5-like classifier performed worse across the board, with both the C5.0 and RF classifiers performing similarly. The C5.0 classifier nearly always outperformed the RF, and when it did not its MCC was only marginally lower. Additionally, the C5.0 classifier trained the models substantially faster than the RF classifier, fifteen times faster on average. The C5.0 classifier also massively outperformed the other classifiers at the note level.

There were issues with the small sample size of the solo level resulting in marginal classification differences having substantial impact on the performance of the models. Consequently, the solo abstraction level models were excluded from consideration. Models that included comparisons with Davis tended to perform the best at the phrase level. This suggested that distinct elements of Davis' improvisational style were well represented by the phrase level features. Green vs. All tended to perform worst amongst the one-vs-all comparisons. Considering the one-vs-one results, this was likely due to the substantial influence Parker had on Green's improvisational style (Green 1999, 6). All classifiers tended to struggle with comparisons that included Parker. This was likely due to a combination of Parker

[5]The confusion matrices of the RF and C5.0 models showed that the C5.0 model, while not significant, did achieve a higher rate of classifications for Parker (higher specificity). This suggested that the RF model may not have been strictly better than the C5.0 model.

having the smallest dataset of the four performers and his substantial and lasting influence on jazz improvisation. The note level comparisons for Coltrane vs. Parker suggested that similarities in instrumentation led to greater difficulty in separating them at the note level.

In summary, these results indicated that the C5.0 classifier was the best algorithm to use for this performer classification task, being both the fastest and best performing algorithm of the three tested. The models tended to perform very well on Davis, especially at the phrase level, while they performed worse for comparisons including Parker. Models that included Green or Coltrane tended to perform consistently well across the board, with the exclusion of comparisons with Parker. The highest MCC for any model was the C5.0 Green vs. Coltrane at the note level, with near perfect classification on the testing data (MCC: 0.98). Following the evaluation of the models' performance metrics, the features that were used to classify the improvisers were extracted and investigated.

12.2 Feature Analysis

This section investigated which features of the improvisations were most frequently used to separate the performers. This was achieved through investigation of the variable importance metrics. All models that did not significantly outperform the NIR were excluded from the feature analysis, leaving 131 of the 165 models. Additionally, as the solo level results were inconclusive, all solo level models were also excluded. Finally, it was decided that the feature analysis would focus on the best performing model of each abstraction and classification situation. This decision was based on the supposition that the best performing models would provide the most insight into the features that best separated the improvisers. This left forty-two distinct abstraction and classification situations with a significant model, out of a possible forty-four.[6] The only situations where no model significantly outperformed the NIR were Parker vs. All at the phrase and note level.

First, the MCC for each abstraction and classification was compared across the three classifiers. None of the forty-two best performing models came from the C4.5-like classifier, eight were from the RF classifier, with the remaining thirty-four from the C5.0 classifier. Consequently, the C4.5-like models were not considered further. The C5.0 classifier had MCC scores up to 0.38 higher than the RF (note level – Davis vs. All); however, the largest difference where the RF outscored the C5.0 was by 0.047 (phrase level – Green vs. Parker). There was also a single model,

[6]The total 165 trained models was the result of fifty-five models for each classifier. The removal of the solo level reduced the total to 132 models, forty-four per classifier.

Coltrane vs. Parker at the note level, where only the RF performed significantly better than the NIR; however, this model had a low MCC of 0.42.

The method by which each classifier generated the feature importance score varied, as did the reporting of the variable importance. For the C5.0 and RF classifiers, the variable importance was generated from the same package that trained and evaluated the models, `caret`, and used the built in variable importance measures of the respective model packages. The definitions for both the C5.0 and RF classifier's variable importance metrics are quoted below. Each classifier had two measures of variable importance.

C5.0 Variable Importance Metrics

> By default, C5.0 measures predictor importance by determining the percentage of training set samples that fall into all the terminal nodes after [each] split (this is used when `metric = "usage"`). For example, the predictor in the first split automatically has an importance measurement of 100 percent. Other predictors may be used frequently in splits, but if the terminal nodes cover only a handful of training set samples, the importance scores may be close to zero. The same strategy is applied to rule-based models as well as the corresponding boosted versions of the model. … When `metric = "splits"`, the percentage of splits associated with each predictor is calculated.
>
> (Kuhn 2020)

Random Forest Variable Importance Metrics

> The first measure is computed from permuting OOB [out-of-bag] data: For each tree, the prediction error on the [OOB] portion of the data is recorded (error rate for classification …). Then the same is done after permuting each predictor variable. The difference between the two are then averaged over all trees, and normalized by the standard deviation of the differences. If the standard deviation of the differences is equal to 0 for a variable, the division is not done (but the average is almost always equal to 0 in that case).
> The second measure is the total decrease in node impurities from splitting on the variable, averaged over all trees. For classification, the node impurity is measured by the Gini index. (More information can be found at https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm#varimp)
>
> (Liaw and Wiener 2017)

For the C5.0 classifier, the default `metric = "usage"` option was used. Both C5.0 and RF are ensemble methods, meaning they combine the results of multiple trees internally for the final classification, using a boosting (C5.0) or bagging (RF) approach. The result of the boosting and variable importance metric selected for the C5.0 classifier was that multiple features could be assigned the maximum importance score of 100. Assuming that the boosted trees outperformed any single tree, any feature that occurred at the top of any individual tree would be assigned an importance score of 100. Consequently, while the RF classifier only had one variable with an importance score of 100 for each model, the C5.0 classifier could, and frequently did, have multiple.

The substantially different reporting of variable importance scores made it difficult to compare the use of features across the classifiers. Therefore, it was decided that the feature analysis would focus on only one classifier to provide consistency across the models. The C5.0 models were selected as they outperformed the RF in the majority of situations, and only had marginally lower MCC when they did not. For the situation where only RF produced a significant model, due to its low MCC, its exclusion from the feature analysis was unlikely to negatively impact the findings. Therefore, the variable importance scores from forty-one C5.0 models formed the basis of the feature analysis.[7] A threshold for variable importance scores needed to be determined so that only the most important features were investigated. Due to the metric selected for the C5.0 classifier, the threshold was set at 100. Therefore, any features that were used at the top of any of the individual trees were considered.

Similar to the feature selection, the discussion on the feature analyses focused on condensed and simplified features. Consequently, additional consideration was required for features that were OHE, regarding how to combine the n classes variable importance scores to a single metric.[8] The issue was that only one class of a categorical variable may have had an importance score of 100. The solution was that if any class of a feature scored 100, that was indicative of that feature's usefulness in identifying the performers. Therefore, the maximum variable importance score for all classes of a feature was taken for the simplified or condensed version. If a OHE feature was found to be important, it was then split out again to investigate which particular classes were most useful. The threshold was applied after combining the features to their condensed form. Throughout the following discussion, features that

[7] All solo level models, Parker vs. All at the phrase and note level, and Coltrane vs. Parker at the note level were excluded.

[8] The C5.0 classifier also dummy encoded some of the output variables. If a categorical feature has n levels, OHE will result in n separate variables whereas dummy encoding will result in $n-1$ variables. This predominantly occurred at the note and phrase levels, and with division and octave features. While there is a difference between OHE and dummy encoded, the discussion will exclusively use the term OHE, as that was the process used in preparing the data for training. The solution to combining the data was the same for OHE and dummy encoded features.

are described as being found to be important refer to only those where the reported importance score was 100.[9]

12.2.1 n-way

The distribution of the simplified features found to be important for the n-way classifications across the abstractions can be seen in Figure 12.5. The colours indicated the domain to which each feature was related. The two bar levels contained the most simplified features that met the importance threshold (eleven), slightly more than the phrase (seven) and note (nine) levels. Across the four abstractions, there were thirteen important simplified features, with eleven appearing at two or more abstraction levels. The important classes in each abstraction for select OHE simplified features can be found in Table 12.6. The select features were those that had the most disparate set of individual classes, or were important in the discussion (the same applied for the one-vs-all, Table 12.7, and one-vs-one tables, Table 12.8).

[9]Other features that had high importance scores were also useful in classifying the performers, but were not considered in this research.

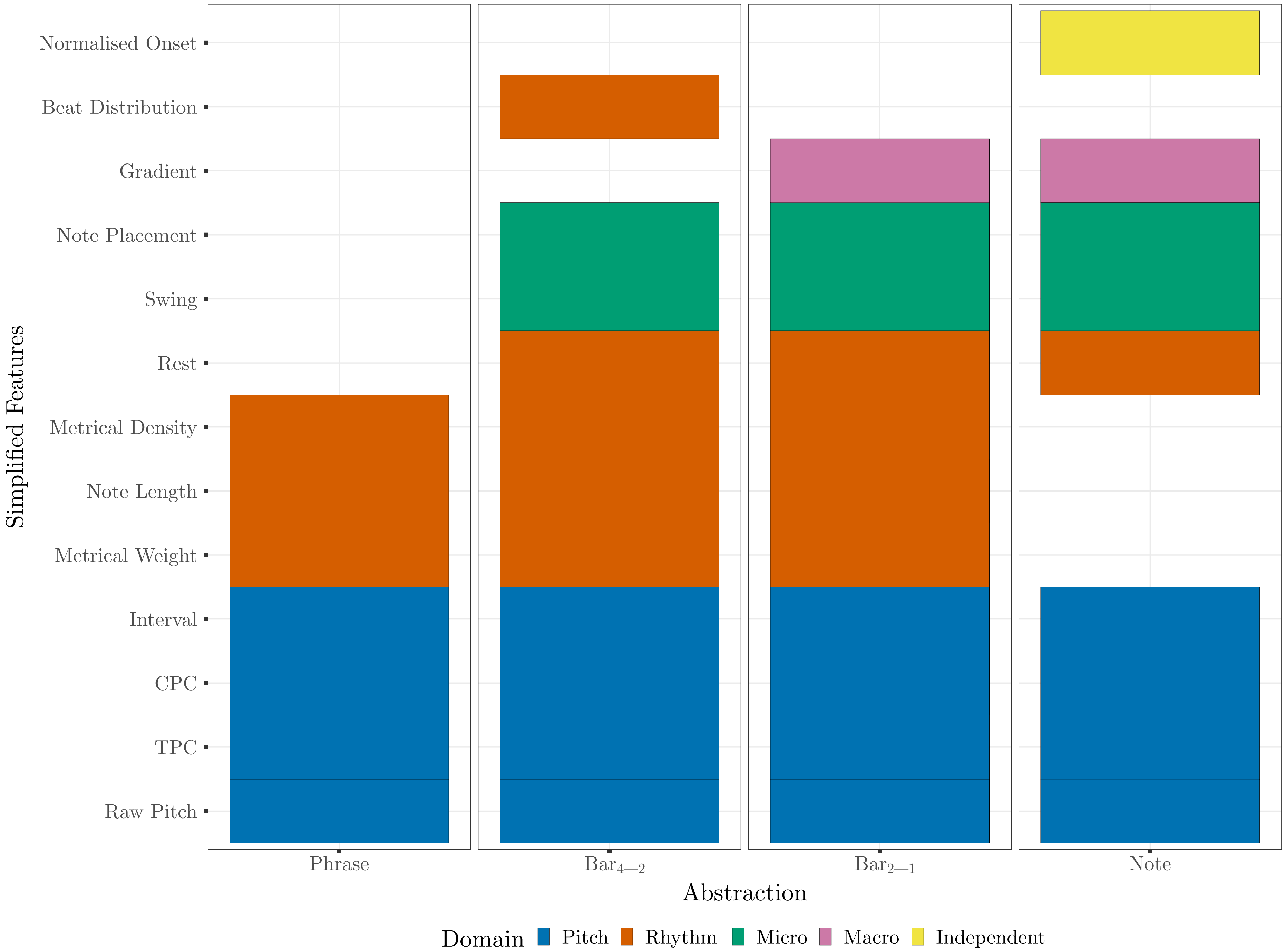

Figure 12.5: Frequency of simplified features across the abstraction levels from the C5.0 classifier in the n-way classification task.

Table 12.6: Select OHE simplified feature classes found to be important in n-way comparisons.

	Phrase	Bar$_{4\|2}$	Bar$_{2\|1}$	Note
Interval				
Fuzzy Interval	1	1	1	-2, -1, 0, 4
Parsons	-	Rep, Asc	Rep, Asc	-
Tonal Pitch Class				
TPC	Min 2$^{\text{nd}}$	Tonic, Min 3$^{\text{rd}}$, Maj 3$^{\text{rd}}$	Tonic, Maj 3$^{\text{rd}}$	Maj 2$^{\text{nd}}$, Per 4$^{\text{th}}$, TT, Min 7$^{\text{th}}$
Chordal Pitch Class				
CPC$_{\text{Weight}}$	Arpeggio	Arpeggio, NHT	Arpeggio, NHT	-
CDPCX	-	1, 2, ♭3/♮3, ♮3/♭3, TT	3, 5, ♭3/♮3, ♮3/♭3, TT, ♮7/♭7	♮3/♭3, ♭6
Note Length				
Fuzzy IOI	Long	Med, Short	Long, Med, Short	-
Fuzzy Duration	Long	-	-	-
Duration Class	-	Very Short, Med, Long, Very Long	Very Short, Med, Long, Very Long	-
Metrical Weight				
Beat Weight	First	First, Mid, Last	First, Mid	-
Metrical Weight	-	Weak	Weak, Off	-

[1] 1: Step Up, -2: Leap Down, -1: Step Down, 0: Repetition, 4: Big Jump Up

Four simplified features were found to be important at all n-way abstractions, raw pitch, TPC, CPC, and intervals. The raw pitch, including NITP, octave, and normalised pitch, described where each performer played on their instrument. The frequent importance of these features suggested that their combination may have been used as a proxy for instrument identification. The fuzzy interval and Parsons classes indicated that ascending intervals were more useful in identifying the performers. At the larger abstractions, the proportion of step ups in a given abstraction was particularly useful. In contrast, the variety of fuzzy intervals at the note level suggested that their use within the flow of an improvisation was useful in classifying the performers. Additionally, the chromatic feature (phrase and $bar_{2|1}$) and mean interval size (both bar levels) were found to be important.[10]

As the tonality of the pieces were unknown the TPC values were discussed in reference to the intervallic difference from the tonic (e.g. TPC 0 = tonic, TPC 3 = minor 3rd, TPC 6 = TT, TPC 11 = major 7th). There were some similarities in the TPC classes found to be important, especially between the bar levels. Both the phrase and note level had clear use of NDTs (minor 2nd and TT), with the bar level results suggesting that tonality may have been inferred and used as a filter. At least one measure of CPC was found to be important at each abstraction. The CPC_{Weight} was found to be important at the phrase and both bar levels, while CDPCX was found to be important at the note and both bar levels. As with the previous features, the important classes were similar for both bar levels. The results indicated that the use of arpeggios and NHTs were most useful in classifying the performers, with only one scale class found to be important (2nd for $bar_{4|2}$).

A measure of the note length was found to be important at all abstractions except the note level. Fuzzy IOI was the most common note length feature, with classes found to be important at three abstractions. The fuzzy duration was only found to be important at the phrase level, with the standard duration classes found to be important at both bar levels. The $duration_{BeatProp}$ was also found to be important for both bar abstraction levels, with the $IOI_{BeatProp}$ only found to be important at the $bar_{4|2}$ level. The final two features associated with the note lengths were the beat distribution (division: $bar_{4|2}$) and metrical density (notes per bar: phrase; $bar_{4|2}$; and $bar_{2|1}$). The notes per bar feature was not a OHE variable, while the division was dummy encoded by the C5.0 classifier. The mode division of eight was the only class found to be important.[11] Representation of note length was found to be important across the abstractions. The classes indicated a slight preference for using the proportion of medium and longer notes to identify the performers.

There were two simplified features that related to the placement of notes, the metrical weight and note placement. The metrical weight included the beat weight and metrical weight, while the note placement focused on the onset difference. The metrical and beat weights represented the proportion of notes played in those beats or metrical positions within an abstraction. Only the proportion of notes played in the first beat was important in the majority of abstractions, indicating a difference between the performers. The mean onset difference was important at both bar levels, while the categorical note placement feature was found to be important at the

[10] The simplified interval features of chromatic and arpeggio represented the proportion of intervals at a particular abstraction that moved chromatically or in thirds.

[11] The division mode indicated that the most frequent division across the abstraction was a beat divided into eight sub-beat tatums.

$bar_{4|2}$ and note level. At the $bar_{4|2}$ level, it was the proportion of notes played before the beat that was the important class; at the note level it was notes that were played after the beat. The lack of the beat and metrical weight features at the note level suggested that on a note-by-note basis the location of a note within the structure of a bar was not useful in classifying the performers. The results indicated that at higher abstractions the proportion of beats played in the first beat of a bar or on weak beats was distinct between the improvisers.

The use of rests and swing were found to be important at the note and both bar levels. The gradient was found to be important at the $bar_{2|1}$ and note level, while the normalised onset, which was only available at the note level, was also found to be important.[12] The normalised onset indicated that changes in features over the course of an improvisation may have been useful in classifying the performers.

The summary of the feature analysis for the n-way comparison found a wide variety of features important in identifying the performers. These features included those related to intervals, TPC, CPC, raw pitch, swing, note length, metrical placement, and note placement. These features covered many of the fundamental elements of music. Although there were many features, typically only a few specific classes of each feature were identified as being important. The results also indicated that while there were limited differences, both bar abstraction levels found similar features and classes important. The following one-vs-all and one-vs-one feature analyses provide further insight into specific features and classes useful for identifying each performer.

12.2.2 One-vs-All

The distribution of the sixteen simplified features found to be important across the one-vs-all comparisons can be found in Figure 12.6. The features are shown across the four comparisons, with colours indicating the abstraction. Across the abstractions, nine condensed features were found to be important for Parker vs. All, thirteen for Coltrane and Davis vs. All, and fourteen for Green vs. All. Across the comparisons the note and $bar_{2|1}$ levels had ten important features, $bar_{4|2}$ had eleven, and the phrase level had thirteen. The important classes for the select OHE simplified features are shown in Table 12.7.

[12] Gradient was the measure of the change in pitch (ΔP) over the number of intervals (ΔT) between pitch extrema ($\Delta P \div \Delta T$). Pitch extrema were notes surrounded by higher or lower pitches.

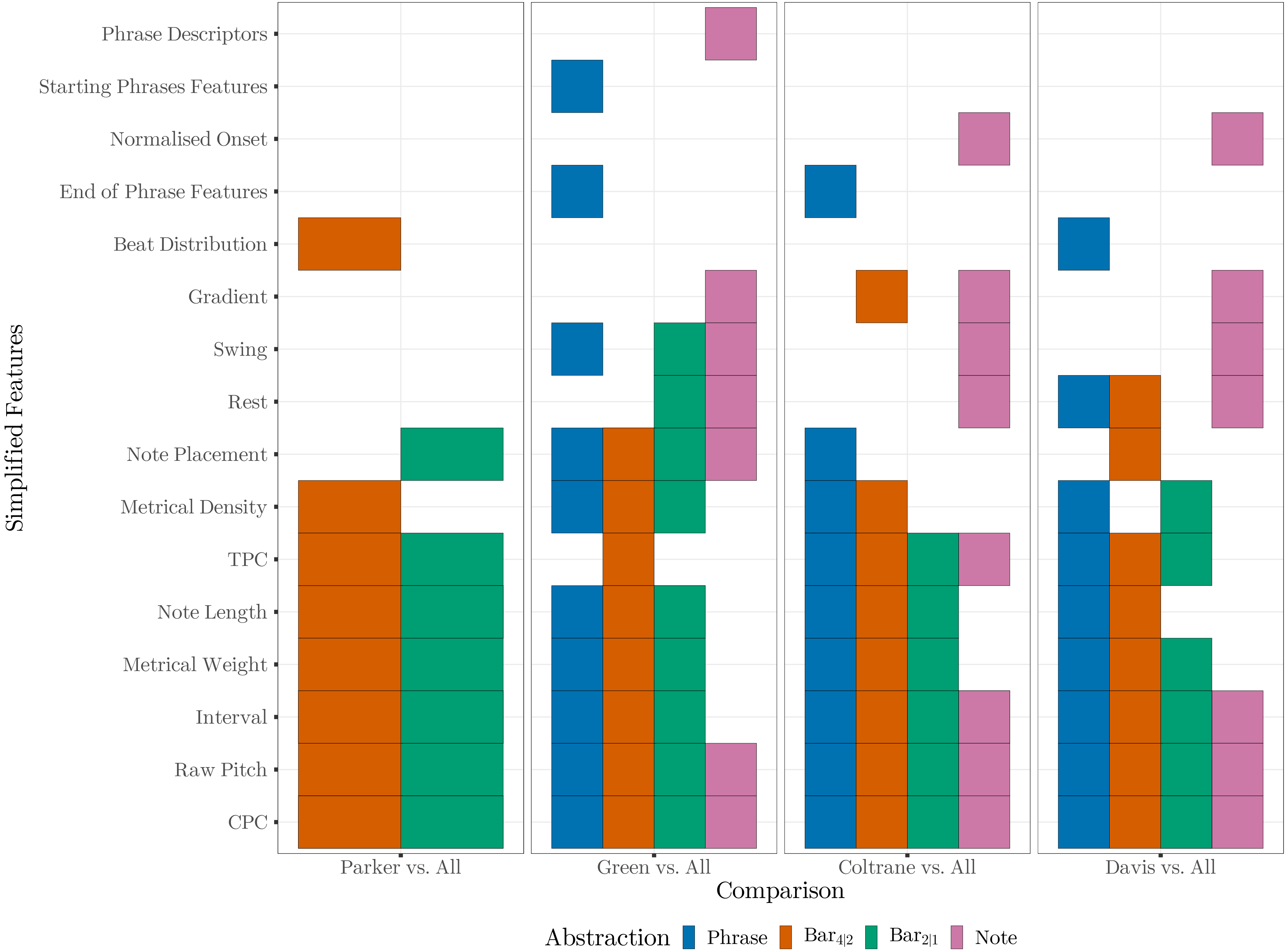

Figure 12.6: Frequency of simplified features across the abstraction levels from the C5.0 classifier in the one-vs-all classification tasks.

Table 12.7: Select OHE simplified feature classes found to be important in one-vs-all comparisons.
GG: Grant Green. JC: John Coltrane. CP: Charlie Parker. MD: Miles Davis.

	Phrase	Bar$_{4\|2}$	Bar$_{2\|1}$	Note
Interval – Fuzzy Interval				
CP vs. All	-	-	3	-
GG vs. All	-4, 1	-4, 1, 4	-4	-
JC vs. All	-	-4	-4	2, 3, 4
MD vs. All	1	1	-4, -3, 1, 2	2, 3, 4
Interval – Parsons				
CP vs. All	-	Asc	Desc, Rep	-
JC vs. All	Rep	Rep	Rep	Asc
MD vs. All	-	Asc, Rep	-	-
Tonal Pitch Class – TPC				
CP vs. All	-	Tonic, Min 2nd, Min 3rd	TT	-
GG vs. All	-	TT	-	-
JC vs. All	Tonic	Maj 3rd	Maj 3rd	Min 2nd, Per 4th
MD vs. All	Min 2nd	Maj 3rd, Per 5th Maj 7th	Min 3rd, Per 4th Per 5th, Min 7th	-
Chordal Pitch Class – CPC$_{Weight}$				
CP vs. All	-	Arpeggio, NHT	Arpeggio, Scale NHT	-
MD vs. All	-	-	Scale	-
Chordal Pitch Class – CDPCX				
CP vs. All	-	3, ♭2	5, 6, ♮3/♭3	-
GG vs. All	♮3/♭3, ♭6, ♮7/♭7	6	♭3/♮3, ♮7/♭7	♭3/♮3, ♮3/♭3, ♭7/♮7
JC vs. All	♭3/♮3, ♮7/♭7	6, ♮3/♭3	2	4, ♮3/♭3
MD vs. All	♭2	♭3/♮3, ♮3/♭3, TT	1, 4, ♭3/♮3, ♮3/♭3,	♮3/♭3, ♭6
Note Length – Fuzzy IOI				
CP vs. All	-	Med, Long	Long	-
GG vs. All	Long	Long	-	-
JC vs. All	-	-	Med	-
MD vs. All	Short	-	-	-

Note Length – Fuzzy Duration				
GG vs. All	Short	-	-	-
JC vs. All	Med	-	-	-
Note Length – Duration Class				
CP vs. All	-	-	Very Long	-
GG vs. All	-	Long	Very Long	-
MD vs. All	-	Very Short, Med	-	-
Metrical Weight – Beat Weight				
CP vs. All	-	Last	First	-
GG vs. All	First	First, Last	First	-
JC vs. All	First	First	First	-
MD vs. All	First, Mid	First, Mid, Last	First, Mid	-
Metrical Weight – Metrical Weight				
CP vs. All	-	Strong, Weak	Strong, Weak, Off	-
GG vs. All	Off	Off	-	-
JC vs. All	Off	-	-	-
MD vs. All	-	Weak	-	-
Note Placement – Note Placement				
CP vs. All	-	-	On	-
GG vs. All	Before, After	Before, After	Before	After
JC vs. All	Before	-	-	-

For the n-way comparison, features that were only found to be important in a few models could not provide much insight into the differences between the performers. In comparison, features that were only important for specific one-vs-all or one-vs-one comparisons could be indicative of specific differences for those performers. For example, the beat distribution, which included the division feature, was only found to be important for the models Parker vs. All ($\text{bar}_{4|2}$) and Davis vs. All (phrase). A division class that was found to be important for both the n-way and one-vs-all comparisons was the mode division of eight for Parker vs. All.[13] This suggested that the frequency of beats with division eight was identifiably different in Parker's playing compared to the other performers.

[13]The beat distribution feature for Davis vs. All at the phrase level was the median division across the abstraction.

The simplified features that were found to be important in the majority of the one-vs-all models were: intervals; TPC; CPC; note length; metrical weight; raw pitch; and metrical density. Of these, only the metrical density was not OHE. The mean number of notes per bar was found to be important in at least one abstraction for each of the four comparisons, with it most frequently found in the Green vs. All comparisons.

The simplified interval feature was comprised of both OHE features – fuzzy intervals and Parsons, raw interval descriptors – and the proportion of chromatic and arpeggio intervals. The mean interval value was only important for Parker vs. All at the $\text{bar}_{4|2}$ level. The proportion of arpeggio intervals was found to be important for Davis vs. All at the $\text{bar}_{4|2}$ level and Parker vs. All at the $\text{bar}_{2|1}$ level. The proportion of the chromatic interval was found to be important for Coltrane vs. All at the phrase and $\text{bar}_{4|2}$ levels, Green vs. All at the $\text{bar}_{4|2}$ level, and Davis vs. All at the $\text{bar}_{2|1}$ level. The most frequent important Parsons class was repeated notes, often found to be important in the Coltrane vs. All models. Although descending Parsons were not frequently found to be important, they did comprise a third of the important fuzzy intervals. The important descending fuzzy interval classes were also the largest intervals. This suggested that the use of large descending intervals differed between the improvisers, especially Green and Coltrane. Smaller ascending intervals were found to be important in the Davis and Green vs. All comparisons at larger abstractions. The interval results suggested that the most important classes were the proportion of repeated Parsons, big jump downs, and step ups.

Important TPC classes were most common in Davis vs. All models, followed by Coltrane and Parker vs. All. The importance of the major 3^{rd}, perfect 5^{th}, and major 7^{th} at the $\text{bar}_{4|2}$ level for Davis vs. All suggested that the presence of a major tonality was useful in separating Davis from the other performers. In contrast, at the $\text{bar}_{2|1}$ level, the TPC classes indicated the importance of a minor tonality. From the CPC simplified feature, there were many CDPCX classes found to be important across the models, with the $\text{CPC}_{\text{Weight}}$ only important for three. Only the CDPCX 7 class was not found to be important at least once. The vast majority (70.87%) of important CDPCX classes were NHTs, with ♭3/♮3 and ♭7/♮7 most frequent. The CPC simplified feature results indicated that NHTs were generally more useful than arpeggio or scale tones in classifying the performers.

The length of the notes was not found to be important in any of the note level one-vs-all models. At the $\text{bar}_{4|2}$ level, the mean $\text{IOI}_{\text{BeatProp}}$ was found to be important for both Coltrane and Davis vs. All. The $\text{duration}_{\text{BeatProp}}$ was found to be important at the $\text{bar}_{2|1}$ level for both Green and Parker vs. All. It was expected that longer notes would be more frequent in the Davis vs. All models; however, only the proportion of short fuzzy IOI note lengths were found to be important at the phrase level. This may still have been influenced by Davis playing longer notes, or notes with a larger IOI, with the shorter note classes used to identify the 'All' class. The same was also likely true for the Parker vs. All comparisons, where only long note classes were found to be important.

Similar to the n-way comparisons, none of the metrical weight simplified features were found to be important at the note level. This again suggested that the metrical or beat placement of a single note was not a useful filter for this abstraction. The important metrical weight classes suggested that the proportion of notes played off-beat was broadly useful in classifying the performers. However, the frequency of notes played on weak beats was also important for the Parker and Davis vs. All comparisons. Compared to metrical weight, beat weight classes were more commonly found to be important. All three of the beat weights were found to be important at least three times. However, the only non-note level model where the proportion of beats played in the first beat was not found to be important was Parker vs. All at the $bar_{4|2}$ level. This indicated that the density of the first beat of the bar was very useful in separating the performers.

Continuing the trend of the n-way comparisons, raw pitch features were frequently found to be important across the trained models. The NITP was found to be important in thirteen of the fourteen significant one-vs-all models, all except Davis vs. All at the $bar_{2|1}$ level. However, normalised pitch was found to be important for Davis vs. All at both bar levels. The normalised pitch was also found to be important for Green vs. All at the $bar_{2|1}$ level and Parker vs. All at the $bar_{4|2}$ level. The only octave mode found to be important for Coltrane vs. All was the 3^{rd} octave at both bar levels. The 3^{rd} and 6^{th} octaves were also an important mode for Davis vs. All at both bar levels, and the 5^{th} octave at the $bar_{2|1}$ level. The distribution of these octaves and the importance of the NITP continued to support the supposition that they were being used as a proxy for instrument identification.

The note placement simplified feature was found to be important in half of the one-vs-all models, including all of the Green vs. All models. The Parker vs. All comparison at the $bar_{2|1}$ level was the only model were the proportion of notes played on the beat was important. This suggested that Parker's proportion of notes played exactly on their nominal position was different from the other performers. Phrase classes found to be important for Green vs. All were phrases starting in the last beat of the bar and phrases ending with a medium fuzzy IOI. For Coltrane vs. All a repeated note at the end of a phrase was found to be important. The swing feature was found to be important at the note level for Coltrane, Davis, and Green vs. All, while the mean swing was found to be important for Green vs. All at both the phrase and $\text{bar}_{2|1}$ level. The frequency of the swing feature for Green vs. All comparisons suggested that Green's swing was substantially different from the other performers. The rest feature was found to be important for most Davis vs. All models. This was likely related to widely accepted assumptions about Davis' greater use of space. Rests were also found to be important for Green vs. All at both the note and $\text{bar}_{2|1}$ level, as well as at the note level for Coltrane vs. All.

In summary, many of the features found to be important for the n-way comparison were again found to be important for many of the one-vs-all classifiers. For example, the mode division of 8 was found to be important, with the results indicating that it was especially useful in separating Parker from the other performers. The metrical density and NITP was also found to be important in many of the models. The interval analysis indicated that repeated notes (as Parsons) and fuzzy intervals of a big jump down and a step up were most frequently important in identifying the performers. NHT classes were commonly found to be important for classifying the performers, with alterations to the 3^{rd} most frequent. The results of the TPC found a disparate set of classes used throughout the models. Similarly, note lengths were represented in myriad ways. This suggested that while the general features were important in classifying the performers, there were no specific classes that were consistently used. Micro domain features, swing and note placement, were most frequently used in the Green vs. All models, indicating their ability to separate Green from the other performers. The metrical weight of a note was not overly useful in classifying the performers. However, the beat weight, specifically the proportion of notes played in the first beat of a bar, was frequently useful. The following feature analysis for the one-vs-one comparisons provided further insight into the importance of the features for each performer.

12.2.3 One-vs-One

Across the six one-vs-one comparisons and abstractions, all sixteen simplified features were found to be important at least once.[14] The distribution of important simplified features across the abstractions and comparisons can be found in Figure 12.7. Only one model, Coltrane vs. Parker at the note level, did not significantly outperform the NIR and was excluded from the features analysis. The one-vs-one feature analysis focused on the remaining twenty-three significant models. The important classes for the select OHE simplified features are shown in Table 12.8.

[14]With the exclusion of the solo level, the tempo simplified feature was note used in the training of the models, reducing the available simplified features from seventeen to sixteen.

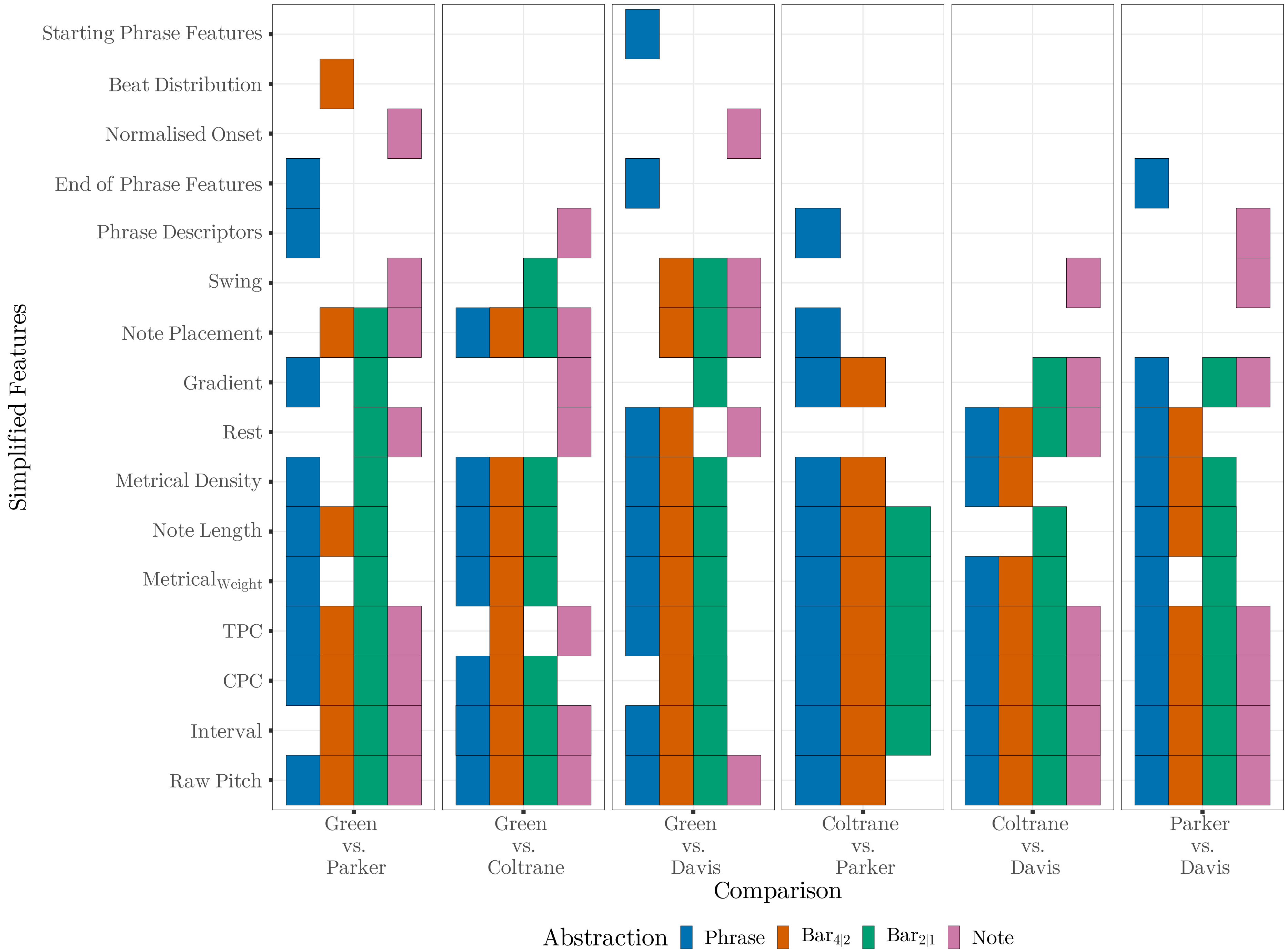

Figure 12.7: Frequency of simplified features across the abstraction levels from the C5.0 classifier in one-vs-one classification tasks.

Table 12.8: Select OHE simplified feature classes found to be important in one-vs-one comparisons.
GG: Grant Green. JC: John Coltrane. CP: Charlie Parker. MD: Miles Davis.

	Phrase	$Bar_{4\|2}$	$Bar_{2\|1}$	Note
Interval – Fuzzy Interval				
GG vs. CP	-	-	-	0
GG vs. JC	4	-4	-4	0, 4
GG vs. MD	1	-3, 1, 2	1	-
JC vs. CP	-4, 1	-4, -1, 3	3	-
JC vs. MD	-	-4, 1	-4	-1, 4
CP vs. MD	1	-	-	0, 2
Interval – Parsons				
GG vs. CP	-	-	Rep	-
GG vs. JC	Rep	-	Rep	Asc
GG vs. MD	-	Asc, Rep	Asc	-
JC vs. CP	Desc	-	-	-
JC vs. MD	-	Desc, Rep	Rep	Asc
CP vs. MD	Asc, Desc	Asc, Rep	Desc, Rep	-
Tonal Pitch Class – TPC				
GG vs. CP	Tonic, Maj 2nd, Per 4th, Per 5th	Min 3rd	Tonic, Min 3rd, Maj 3rd, TT, Min 7th	Per 4th
GG vs. JC	-	Min 2nd	-	TT
GG vs. MD	Per 4th	Min 3rd, Per 4th, TT	Tonic, Maj 3rd, Per 4th	-
JC vs. CP	Min 2nd, Maj 2nd, Per 4th	Maj 2nd, Min 3rd, Maj 6th	Min 2nd	-
JC vs. MD	Tonic, Min 2nd, Maj 3rd	Maj 3rd	Tonic, Maj 3rd, Min 7th, Maj 7th	Per 5th
CP vs. MD	Maj 7th	Tonic, Per 4th	Maj 2nd	Min 3rd, Per 4th, TT, Min 7th
Chordal Pitch Class – CPC_{Weight}				
GG vs. CP	Arpeggio, Scale	Arpeggio, NHT	Arpeggio, Scale, NHT	-
GG vs. MD	-	NHT	Scale	-
JC vs. CP	NHT	Arpeggio, NHT	Arpeggio, NHT	-
CP vs. MD	Arpeggio, NHT	Arpeggio, NHT	Arpeggio, NHT	-

Chordal Pitch Class – CDPCX				
GG vs. CP	1, ♭2, TT	2, 3, ♭2	1, 2, 5, 7	♭3/♮3, ♭7/♮7
GG vs. JC	♮3/♭3	2, 5	2	-
GG vs. MD	-	♭3/♮3, TT	1, ♭3/♮3, TT, ♭6	-
JC vs. CP	4, 5 ♮3/♭3, ♭7/♮7	3, 5, TT	5, 6, ♭2 ♮3/♭3	-
JC vs. MD	♭3/♮3	♭3/♮3, ♮3/♭3	♭3/♮3, ♮3/♭3 TT, ♭7/♮7	♮3/♭3
CP vs. MD	1, 2	1	1, 2, ♭3/♮3	3, 5, 7, TT
Note Length – Fuzzy IOI				
GG vs. CP	Med, Long	Short, Med, Long	Long	-
GG vs. JC	Med	-	-	-
GG vs. MD	Med, Long	Long	-	-
JC vs. CP	-	Long	Long	-
CP vs. MD	Short, Long	Short, Long	Short, Long	-
Note Length – Fuzzy Duration				
GG vs. JC	Short	-	-	-
JC vs. CP	Long	-	-	-
CP vs. MD	Long	-	-	-
Note Length – Duration Class				
GG vs. JC	-	Long	Very Long	-
GG vs. MD	-	Med, Long	Short, Med, Very Long	-
JC vs. MD	-	-	Med	-
Metrical Weight – Beat Weight				
GG vs. JC	First	First, Mid, Last	First	-
GG vs. MD	First	-	First	-
JC vs. CP	First, Mid	First, Mid, Last	First	-
JC vs. MD	First, Mid	First, Mid	First	-

Metrical Weight – Metrical Weight				
GG vs. CP	Strong, Weak	-	Strong, Off	-
GG vs. JC	Off	Off	-	-
GG vs. MD	Off	Weak, Off	Weak	-
JC vs. CP	-	Weak	-	-
CP vs. MD	Strong, Weak, Off	-	Off	-
Note Placement – Note Placement				
GG vs. CP	-	Before	On	After
GG vs. JC	Before, After	Before, On	Before	After
GG vs. MD	-	Before	-	After
JC vs. CP	On	-	-	-

Comparisons that included Green tended to have the highest number of important simplified features. Fifteen features were found to be important in the Green vs. Parker comparison, fourteen for Green vs. Davis, and twelve for Green vs. Coltrane. Parker vs. Davis also had twelve important features, with ten important features for Coltrane vs. Parker and Davis. Each of the abstraction levels had a similar number of important features, with thirteen for the phrase level, twelve for $\text{bar}_{4|2}$, eleven for $\text{bar}_{2|1}$, and ten for the note level.

Unlike the previous n-way and one-vs-all comparisons, there was no simplified feature that was found to be important in every model. However, the most frequent important simplified features were similar, including: intervals; TPC; CPC; note length; metrical weight; raw pitch; and metrical density. The rest and contour features were found to be important at least once in every comparison, while note placement was important in eleven models. All other features were important in seven or fewer models.

Metrical density, as mean number of notes per bar over each abstraction, was found to be important in fifteen of the eighteen non-note level models. This included all six comparisons at the phrase level, all $\text{bar}_{4|2}$ comparisons except Green vs. Parker, and all $\text{bar}_{2|1}$ comparisons except Coltrane vs. Parker and Coltrane vs. Davis. This suggested that the metrical density was widely useful in classifying the performers.

Interval features were frequently found to be important when classifying the performers. They were found to be important fifty-seven times across the models. This included the OHE fuzzy interval and Parsons features, the mean interval size, and the proportion of arpeggio or chromatic intervals in an abstraction. The mean interval size was found to be important for Green vs. Davis at both bar levels and Coltrane vs. Parker at the $\text{bar}_{4|2}$ level. The proportion of chromatic intervals was found to be important for Green vs. Coltrane and Coltrane vs. Davis at the phrase, $\text{bar}_{4|2}$, and $\text{bar}_{2|1}$ levels, and Green vs. Davis at the $\text{bar}_{2|1}$ level. The proportion of arpeggio intervals was found to be important in, Green vs. Parker at both bar levels and Green vs. Davis at the phrase and $\text{bar}_{2|1}$ levels.

Similar to the one-vs-all comparisons, the repeated Parsons was the interval class most frequently found to be important, especially at the bar abstractions. There were also similarities with the important fuzzy interval classes. Out of the twenty-five important classes, six each were big jump down and step up. All the big jump down classes were found to be important in comparisons that included Coltrane, and five of the six step up classes included Davis. This indicated that Coltrane's use of large descending intervals and Davis' use of ascending steps differed identifiably from the other performers. Overall, sixteen of the twenty-five important classes (64.00%) were in comparisons that included Coltrane. This suggested that fuzzy interval classes were most useful in separating Coltrane from the other performers.

A TPC class was found to be important forty-four times across twenty of the twenty-three one-vs-one models. The classes were most frequently found to be important were: perfect 4$^{\text{th}}$ (eight occurrences); tonic (six occurrences); minor 3$^{\text{rd}}$ (five occurrences); and major 3$^{\text{rd}}$ (five occurrences). The two most frequent NDT classes were the minor 2$^{\text{nd}}$ and TT, each with four occurrences.[15] TPC was least frequently important at the note level (seven classes), followed by the $\text{bar}_{4|2}$ (eleven), phrase (twelve), and $\text{bar}_{2|1}$ (fourteen). Comparisons that included Parker or Davis tended to contain the most important TPC classes. This suggested that Davis and Parker's TPC differed the most from each other and the other performers.

CPC classes were also found to be important in twenty of the significant one-vs-one models. Across these models a CPC class was found to be important seventy-one times. Twenty of these were $\text{CPC}_{\text{Weight}}$ classes with the remaining fifty-one CDPCX classes. Continuing the trend of similar results between the one-vs-all and one-vs-one models, scale tone $\text{CPC}_{\text{Weight}}$ were least commonly important. Although the $\text{CPC}_{\text{Weight}}$ was not important in all of the comparisons, it was important in every comparison that included Parker. These results suggested that Parker's distribution of arpeggio and NHTs was substantially different from those of the other performers.

A disparate set of CDPCX classes were found to be important in all twenty models where the simplified CPC feature was important. The only CDPCX class not found

[15] As the tonality of the pieces was unknown, either of the major or minor thirds and sevenths could be a DT or NDT.

to be important was ♮7/♭7. In contrast to the one-vs-all models, only twenty-six CDPCX classes (50.98%) were NHTs. Of these, the most frequent classes were ♭3/♮3, ♮3/♭3, and TT. From the remaining twenty-five CDPCX classes, seventeen were arpeggio tones and eight were scale tones.[16] The proportion of ♭3/♮3 notes over an abstraction was found to be important predominantly for identifying Davis. Similarly, the ♮3/♭3 class was useful in classifying Coltrane, with all six models where it was found to be important including Coltrane. As with the $\text{CPC}_{\text{Weight}}$, many of the important CDPCX classes were in models that included Parker, with thirty-three (64.71%) of the classes being from these comparisons. Altogether this suggested that the CPC features were useful in separating Parker from the other performers. The CDPCX results also aligned with the $\text{CPC}_{\text{Weight}}$ results, with most of the important classes being arpeggio or NHTs.

All thirty-four note length classes found to be important were at the phrase or bar levels. The $\text{IOI}_{\text{BeatProp}}$ was found to be important for Green vs. Davis at the $\text{bar}_{4|2}$ level and Green vs. Parker at the $\text{bar}_{2|1}$ level; the $\text{duration}_{\text{BeatProp}}$ was found to be important at both bar levels for Green vs. Davis and at the $\text{bar}_{2|1}$ level for Parker vs. Davis. The majority of the OHE note length features were fuzzy IOI classes. Across all three of the OHE features, only six classes represented the proportion of short notes played. This suggested that the proportion of medium and long classes were more commonly used in classifying the performers, with the majority of all classes representing long notes. This data could not indicate whether it was a higher or lower use of long notes that was important, or how it correlated with the use of shorter notes. The C5.0 algorithm likely favoured the use of medium and long notes over short note lengths as the vast majority of all notes had a short length.[17] These results indicated that the length of notes was widely useful in classifying the performers, with the C5.0 algorithm favouring longer classes.

Similar to the note length features, and one-vs-all comparisons, none of the metrical weight features were found to be important at the note level. Although metrical weight classes were found to be important at least once in every comparison, the majority (54.54%) were beat weight classes. All models where the beat weight was found to be important included comparisons with Coltrane. The proportion of notes played in the first beat of the bar was again the most frequent beat weight class. This supported the supposition from the one-vs-all comparison that the proportion of notes played in the first beat of a bar was useful in classifying the performers. Additionally, the one-vs-one data indicated that this class was particularly important for separating Coltrane from the other performers. The proportion of

[16]The most frequent arpeggio tones were the tonic and 5$^{\text{th}}$, while the most frequent scale tone was the 2$^{\text{nd}}$.

[17]70.70% of fuzzy IOI and 89.44% of fuzzy duration classes were short.

notes played on a metrically strong beat was the least frequent metrical weight class found to be important. This was followed by the proportion of notes played on metrically weak beats and off-beat. Although metrical weight classes were frequently found to be important in classifying the performers there was no consistent trend across the performers or abstractions. This suggested that while the beat weight was useful in identifying Coltrane, the metrical weight was used generally to separate the performers.

The most common important raw pitch classes were related to the NITP. The NITP was found to be important in all models except Parker vs. Davis at the $\text{bar}_{4|2}$ level and Coltrane vs. Parker at the $\text{bar}_{2|1}$ level. This matched the results from the n-way and one-vs-all comparisons, where the NITP was widely found to be important in classifying the performers. In comparison, the normalised pitch was only found to be important for Green vs. Coltrane and Green vs. Parker at both bar levels, and Green vs. Davis at the $\text{bar}_{2|1}$ level. This suggested that Green's normalised pitch was frequently different from that of the other three performers at the bar level. All comparisons where the octave mode was found to be important included Davis. The only octave modes found to be important were the 3^{rd}, 5^{th}, and 6^{th}. This continued to support the hypothesis that a combination of the octave mode and NITP was used as a proxy for identifying the instrument played.

The note placement feature analysis indicated that it was found to be particularly important in comparisons that included Green. The mean onset difference was only found to be important at bar levels that included Green (Green vs. Parker, Coltrane, and Davis at the $\text{bar}_{4|2}$ level; and Green vs. Coltrane and Davis at the $\text{bar}_{2|1}$ level). The note placement classes also indicated their importance in identifying Green, with only one class found to be important in a non-Green model. The proportion of notes played before the beat was the most common important class, although a note played after the beat was important at the note level. These results indicated that Green's note placement differed substantially from those of the other performers.

At least one phrase feature was found to be important for every comparison except Coltrane vs. Davis, with nearly all features found to be important at the phrase level. The starting phrase feature of phrases starting in the last beat of the bar, which was found to be important for Green vs. All, was only found to be important in Green vs. Davis. For the end of phrase only two features were found to be important: the fuzzy interval leading into the final note; and the fuzzy IOI of the final note. For Parker vs. Davis the important fuzzy interval leading into the end of the phrase was a jump down. The same fuzzy interval was found to be important for Green vs. Parker, as well as repeated notes. A final fuzzy interval of big jump up was also found to be important in the Green vs. Davis comparison. Both fuzzy IOI classes found to be important were with comparisons involving Green. The final note of a phrase having a short fuzzy IOI was important for Green vs. Parker, while an ending fuzzy IOI of medium was important for Green vs. Davis. The other phrase descriptors – phrase position, length, and contour – were predominantly found in models that included Parker or Coltrane. These results indicated that there were identifiable differences in how the performers began and ended their phrases.

The only important beat distribution feature was the same from the n-way and Parker vs. All comparisons. The important class was the mode division of 8 at the $\text{bar}_{4|2}$ level for Green vs. Parker. This consistency in results across the comparisons indicated that there was a substantial difference in Parker's distribution of divisions from the other performers, especially Green. Normalised onset was only found to be important at the note level for Green vs. Parker and Davis. The note level swing BUR was important for Parker vs. Davis, Green vs. Parker, Green vs. Davis, and Coltrane vs. Davis. The mean BUR was important for Green vs. Coltrane and Green vs. Davis at the $\text{bar}_{2|1}$ level, and Green vs. Davis at the $\text{bar}_{4|2}$ level.

The proportion of pitch extrema was commonly found to be important in comparisons that included Parker. These models included: Parker vs. Davis and Green vs. Parker at the phrase and $\text{bar}_{2|1}$ level; and Coltrane vs. Parker at the phrase and $\text{bar}_{4|2}$ level. These results suggested that Parker played an identifiably different proportion of pitch extrema compared to the other performers. The rest feature was found to be important in twelve of the models, with nine from comparisons that included Davis. These included: Coltrane vs. Davis at all four abstractions; Green vs. Davis at the phrase, $\text{bar}_{4|2}$, and note level; and Parker vs. Davis at the phrase and $\text{bar}_{4|2}$ level. The three non-Davis comparison models where rests were important were Green vs. Parker at the $\text{bar}_{2|1}$ and note level, and Green vs. Coltrane at the note level. These results indicated that Davis' use of rests was substantially different from the other performers. Additionally, Green's use of rests differed from Coltrane and Parker's.

In summary, the results of the one-vs-one feature analysis provided more insight into how the features and their classes were used to identify the performers. The results indicated that all of the simplified features were important for at least one model. These included specific features and classes for certain models, including the starting phrase beat (Green vs. Davis) or division mode (Green vs. Parker). The feature analysis also showed that although broader features (e.g. TPC, CPC, intervals) were widely used, specific classes were more frequently found to be important. There were also features (e.g. metrical weight) where the feature was important, but no specific trend in classes were observed. This suggested that they were used in collaboration with other features to aid in the identification of the performers.

12.2.4 Feature Analysis Summary

This feature analysis based on the variable importance scores of the C5.0 classifier identified features that were frequently used to classify the performers. The feature analysis also aided in identifying features to focus on in the comparative analysis. Many of the important features identified at the n-way comparison were subsequently found to be important at the one-vs-all and one-vs-one comparisons. The feature analysis with the one-vs-all and one-vs-one comparisons aided in identifying specific features and classes that were important for classifying each performer. These results suggested that a combination of the n-way feature analysis and either the one-vs-all or one-vs-one comparisons would have provided sufficient data on which to base the comparative analysis. The selection of an additional comparison could be influenced by the number of performers and the aim of the comparative analysis.[18]

12.3 Model Results Summary

This chapter reported the results of the models trained for the performer classification and comparative analysis task. The performance of the models was evaluated first through the use of the MCC metric. The most frequently used important features and classes from the C5.0 models were then investigated.

The evaluation of the models found that the C5.0 classifier was the best performing algorithm of the three tested. It nearly always trained the best performing models while also being the quickest on average. These results indicated that of the algorithms tested, the C5.0 was the best for performer classification based on improvisational data. All classifiers tended to struggle with identifying Parker, likely caused by his smaller dataset and enduring legacy on jazz improvisation. In contrast, the classifiers were able to very successfully identify Davis, especially at the phrase level. This suggested that unique elements of Davis' improvisational style were well encapsulated by the phrase level data. Models that included Green or Coltrane also tended to perform well. The C5.0 Green vs. Coltrane note level model almost perfectly classified all of the testing data. However, both Green and Coltrane comparisons with Parker tended to perform the worst; at all abstraction levels for Green vs. Parker and Coltrane vs. Parker at the note level. This was likely due to Parker's specific influence on Green's improvisational style, and similarities between saxophonists at the note level. Overall, the classifiers were able to identify the

[18] If the only aim was the identification of features that differed, the one-vs-one feature analysis provided the most detailed feature analysis.

performers with a high degree of accuracy based solely on the improvisational data of their solos.

Based on the results of the model evaluation, the feature importance analysis focused on the C5.0 models. The results of the feature analysis found that many of the important features appeared consistently across the abstraction levels and comparisons. The n-way comparison provided an excellent first pass for important features on which to base a comparative analysis. The one-vs-all and one-vs-one aided in identifying specific sets of features and classes that were specific to individual performers. A wide variety of features from across the domains were frequently found to be important, including:

- Pitch domain: Raw pitch, TPC, CPC, intervals;
- Rhythm domain: Note length, beat weight, metrical density;
- Micro domain: Swing, note placement;
- Macro domain: Phrase features, gradient.

Generally, foundational music features were found to be important more frequently. There were also features not widely used across all comparisons, but were identified as important for a single performer. For example, the beat distribution, as division, for comparisons that included Parker. Following the evaluation of the model results a subset of important features were selected to form the example comparative analysis. The comparative analysis provided insight into specific differences in improvisational style between the performers.

Chapter 13

Comparative Analysis

This chapter showed an example comparative analysis between Green, Coltrane, Parker, and Davis. The main aim of Part III was the development and exploration of the methodology for the classification and extraction of important features through the use of interpretable ML algorithms. Therefore, this comparative analysis aimed to serve as an example of one application of the results, rather than an in-depth comparison between the performers. The example comparative analysis focused primarily on the specific classes previously identified as important. The analyses investigated the identified abstractions and performers from the feature analysis. They also included other performers, as important features for one comparison were also likely important in others. The analyses also investigated the use of the feature more broadly, outside of specific abstractions, to investigate whether a trend identified at an abstraction applied generally.

Although the results of the feature analysis identified myriad features on which to base a comparative analysis, an analysis of all features was beyond the scope of this research. A subset of features were selected based on repeated specificity of a single class, or frequent appearances across abstractions or comparisons. The comparative analysis aimed to highlight specific differences or elements of the performer's improvisational styles.[1] The features selected to be analysed were:

- NITP and octave;
- TPC – Tonic, minor 2$^{\text{nd}}$, minor 3$^{\text{rd}}$, major 3$^{\text{rd}}$, perfect 4$^{\text{th}}$, and TT;
- CDPCX – ♭3/♮3 and ♮3/♭3;
- Intervals – fuzzy intervals;
- Beat distribution – division mode;
- Metrical density – mean number of notes per bar;
- Beat weight;
- Beginning phrase features – beat weight of first note;
- End of phrase features – fuzzy interval into (repetition and jump down).

[1]Features that have already been studied, either in previous literature or in comparisons in Part II, were excluded. For example, Davis' use of rests of long notes has been explored in many studies.

The results of the analyses are presented below grouped by their domain. The only domain where no feature was selected for the comparative analysis was the micro domain. This was due to the small number of input variables from the micro domain, their disparate use as an important feature for classifying the performers, and some comparisons already undertaken in Part II.

13.1 Pitch Domain

13.1.1 Raw Pitch

Amongst the most frequently identified important features were those related to the raw pitch values played by the performers, both the NITP and octave. Across all the models there was only one (Coltrane vs. Parker at the $\text{bar}_{2|1}$ level) where a raw pitch feature was not found to be important. It was hypothesised that the frequent importance of the NITP and octave was due to the C5.0 classifier using the features as a proxy for instrument identification. Figure 13.1 shows the distribution of the performer's NITP against the NITP distribution for their main instrument from the WJazzD.[2] The graph focused on the majority of the data, with outliers beyond the range shown. This graph showed that there was substantial overlap in the distribution of NITP among the performers and between the instruments. However, there were also sufficient differences such that they could contribute to instrument identification.

An ANOVA was run to investigate the relationship between the performers and their NITP distribution. A significant difference in the distribution of NITP was found, with a medium effect size ($F(3, 51966) = 1881.37,\ p < .001;\ \eta^2 = .10$). Subsequent post-hoc tests with Tukey's HSD procedure found significant pairwise differences for all comparisons at $p < .001$.[3] This suggested that there was enough difference in the NITP to act as a proxy for instrument identification.

[2]92.45% of Coltrane's note events were from improvisations where he played tenor saxophone, with the remaining 7.55% on the soprano saxophone.

[3]An ANOVA also found a significant difference in the NITP between the four main instruments in the WJazzD, with a medium effect size ($F(3, 162436) = 4542.98,\ p < .001;\ \eta^2 = 0.08$). Subsequent post-hoc tests found significant pairwise differences for all comparisons at $p < .001$.

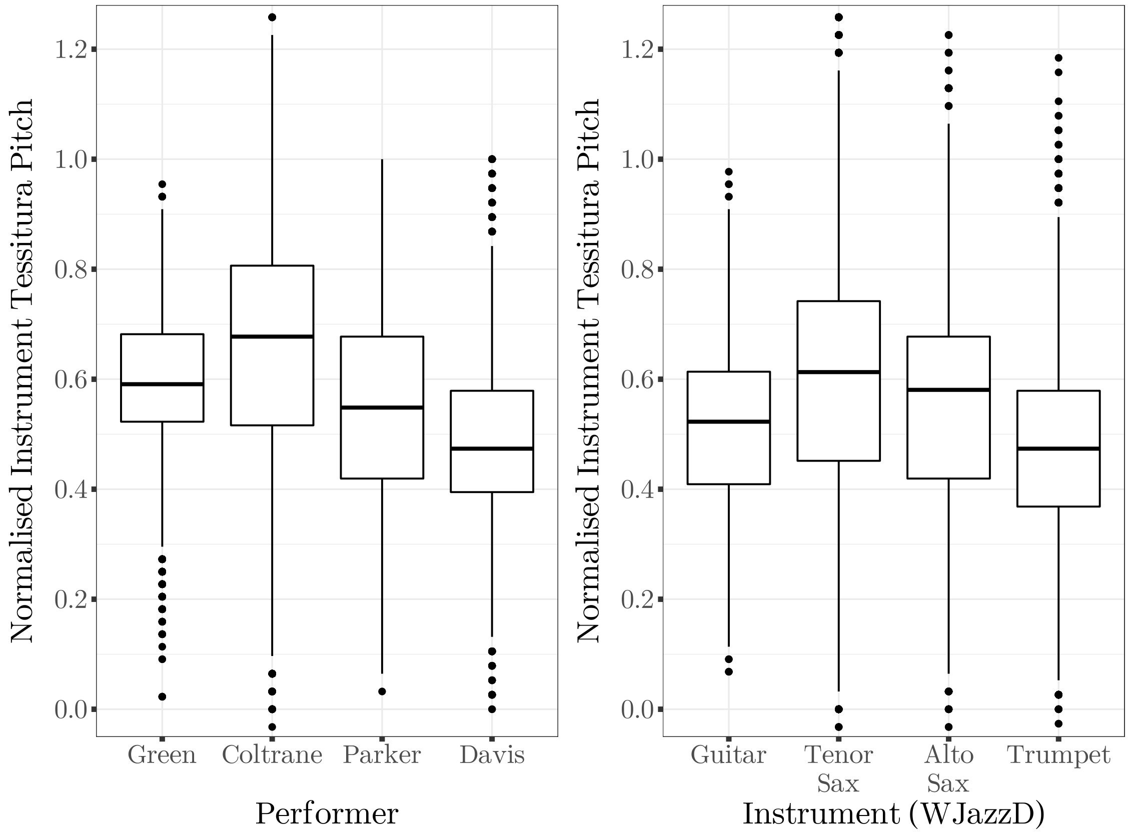

Figure 13.1: Distribution of NITP for the four performers compared to their main instrument in the WJazzD.

Figure 13.2 shows the interaction between the NITP and octave features.[4,5] A substantial relationship was observed between the NITP, octave, and performer or instrument being played. For example, the split between the 4th and 5th octave occurred at the 0.5 NITP for Davis, whereas for Coltrane all of the 4th octave was above 0.5. This data showed that the combination between the NITP and octave was unlikely to be helpful in separating Green and Parker, as the NITP thresholds for each octave were very similar. In contrast, it would have been very successful in separating Coltrane from Davis, as there was little to no overlap in the combination of these features. However, with the exception of the note level, the NITP was calculated as the mean over the abstraction, and the octave as the median or mode. This would have lowered the overall usefulness of the combination of these features. Additionally, in the one-vs-all comparisons the pooling effect for the 'All' class would have impacted the usefulness and importance of this feature combination.

[4] With the x-axis following the same limits from Figure 13.1.

[5] The stacked colours for Coltrane were due to the small number of improvisations where he played a soprano saxophone.

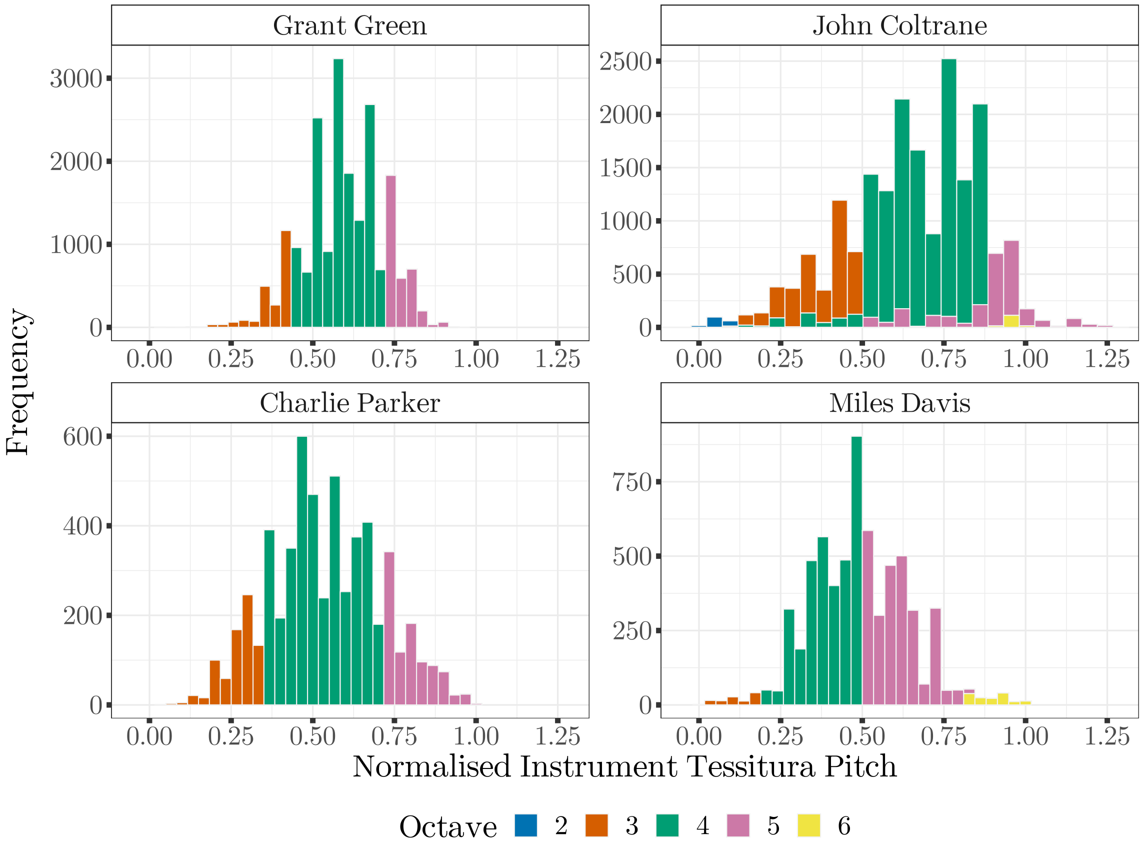

Figure 13.2: Distribution of interaction between NITP and octaves for each performer.

These results indicated that the combination of these features appeared to be used as a proxy for instrument identification. Although the NITP distribution significantly differed between the performers, the combination with the octave feature clearly separated specific performers and instruments. Future research with similar aims should avoid these features, or used a different raw pitch feature. However, the myriad features identified as important in the feature analysis showed that the models were not able to identify the performers based solely on the NITP and octave.

13.1.2 Tonal Pitch Class

The results of the feature analysis indicated there were specific TPC that were frequently used to classify the performers. Across all the models the most frequently used TPC were the tonic, minor 2nd, minor 3rd, major 3rd, perfect 4th, and TT. The distribution of those TPC are shown in Figure 13.3.[6,7]

[6] The percentages for each performer do not sum to 100% as they were calculated from the full TPC distribution.

[7] The overall TPC distribution was found to be significantly different between the performers, with a small effect size ($\chi^2(33) = 1284.94,\ p =< .001,\ V = .09$). Subsequent post-hoc tests found significant pairwise differences for all comparisons at $p < .001$.

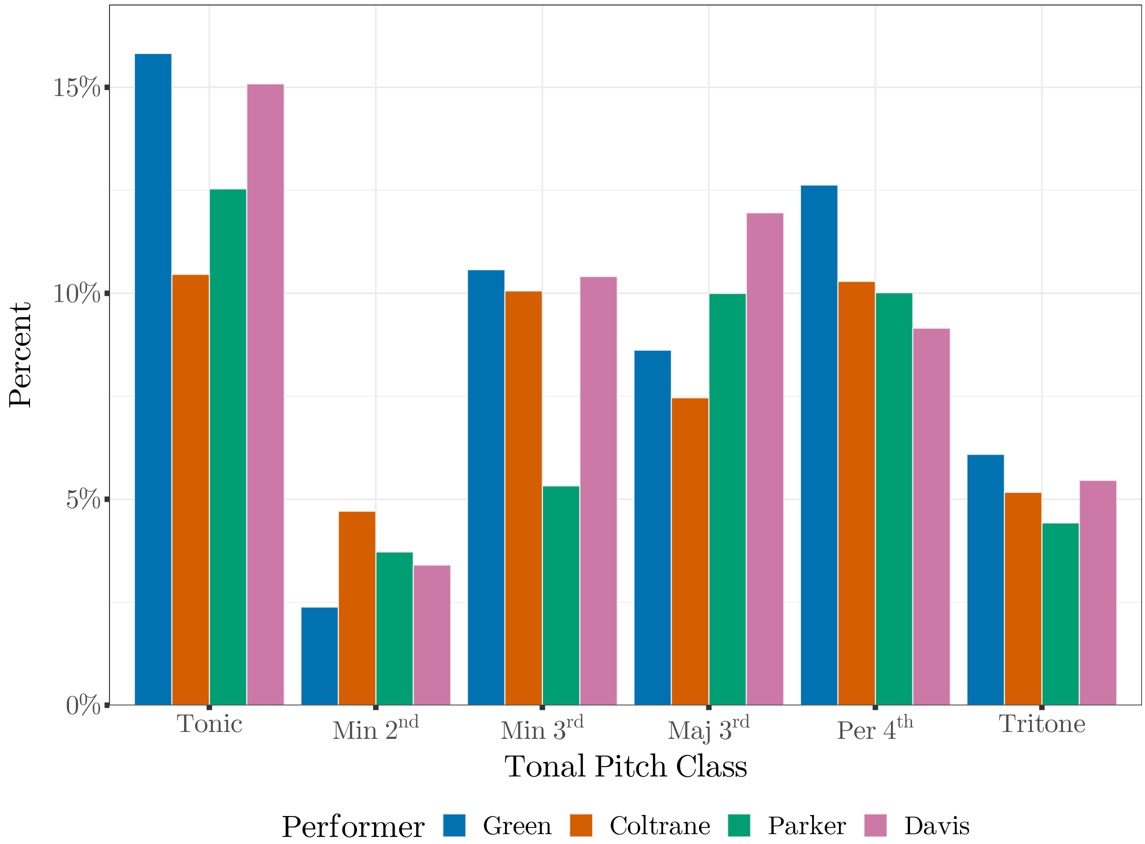

Figure 13.3: Distribution of TPC classes identified as important in the feature analysis for each performer.

Each TPC, aside from TT, had one performer who used that class substantially more or less than the other performers. For example, Parker played fewer minor thirds than the other performers. The most likely explanation for this difference was that only 12.80% of Parker's note events, 11.76% of his improvisations, were classified as being from a minor tonality, lower than the other performers.[8] As very few of his improvisations in the WJazzD were in a minor tonality, the proportion of minor thirds would be lower. Other differences, including Coltrane playing substantially fewer tonics or Green playing more perfect fourths, were likely indicative of improvisational style rather than tonality. As with the TPC analysis in Chapter 5, there was limited insight this feature could provide. This analysis suggested that while some TPC differences between the performers were related to their improvisational style, others indicated differences in tonalities.

[8] In the dataset, zero of Davis' improvisations were classified as being from a minor tonality, but there were model pieces such as *So What* that were in a minor mode.

13.1.3 Chordal Diatonic Pitch Class

There were specific CDPCX that were often found to be important in the feature analysis. This included the ♭3/♮3 for comparisons that included Davis and ♮3/♭3 for comparisons that included Coltrane. The distribution of these two classes can be seen in Figure 13.4.[9] This data supported the feature analysis findings that Davis' use of ♭3/♮3 CDPCX was substantially different from the other performers. The same was true for Coltrane's use of ♮3/♭3, although to a lesser degree. Davis played more than four times as many ♭3/♮3 compared to Coltrane (6.82% vs. 1.63%), two and half as many as Parker (2.73%), and just under twice as many as Green (3.69%). The use of ♮3/♭3 was less common in all of the performer's improvisations. Coltrane played three times as many ♮3/♭3 than Davis (2.29% vs. 0.72%), just over twice as many as Parker (1.02%), and just under twice as many as Green (1.28%).

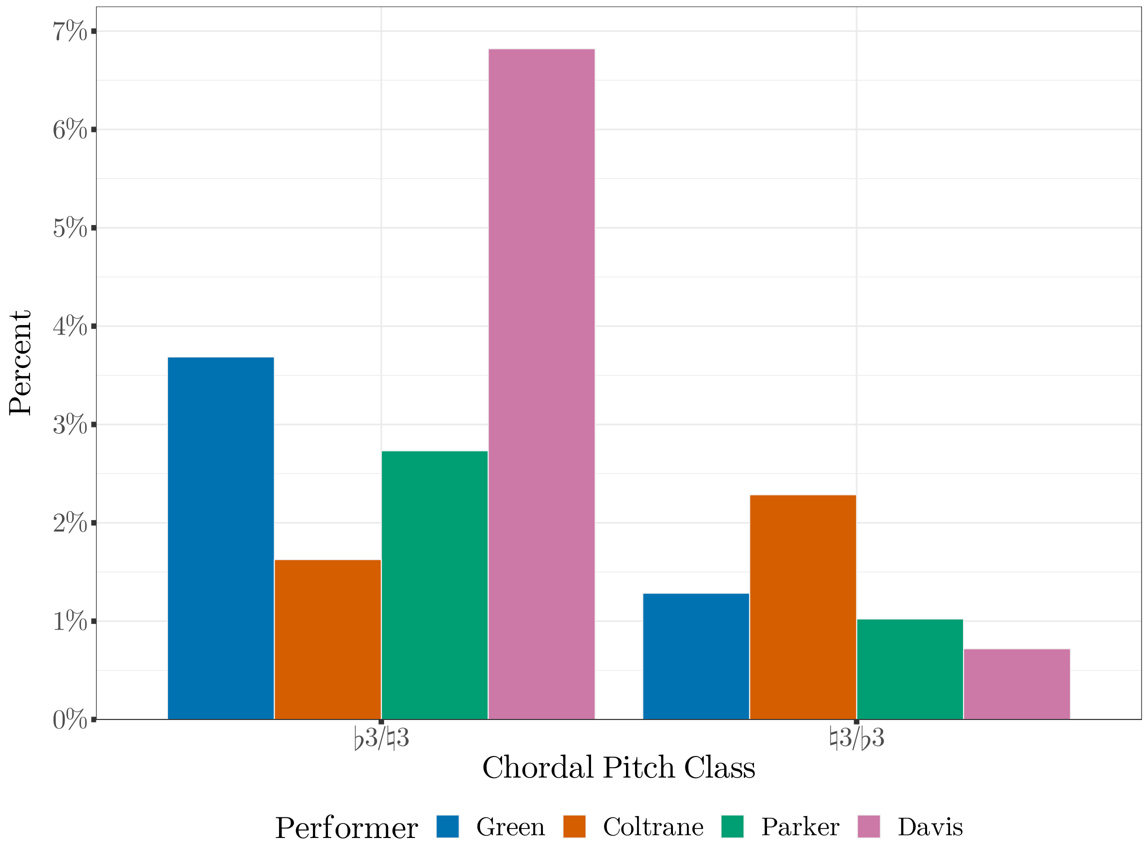

Figure 13.4: Distribution of ♭3/♮3 and ♮3/♭3 CDPCX classes for each performer.

However, there were differences in the frequency of chord types in the performer's datasets. While only 18.16% of Davis' note events were over chords with a ♭3 (m7 or ø7), 59.83% of Coltrane's note events were over a m7 chord.[10] Consequently,

[9] A χ^2-test comparing the complete CDPCX distribution found a significant difference between the four performers, with a small effect size ($\chi^2(39) = 1138.33$, $p =< .001$, $V = .09$). Subsequent post-hoc tests found significant pairwise differences for all comparisons at $p < .001$.

[10] Around 30% of Green and Parker's note events were over m7 chords.

Coltrane had many more opportunities to play ♮3/♭3 while Davis had more chances to play ♭3/♮3. Across the phrase and bar abstractions the mean proportions were similar to the overall distribution shown in Figure 13.4. This indicated that these CDPCX were evenly distributed throughout the performer's improvisations. Figure 13.5 displays a phrase from both Davis and Coltrane highlighting the use of the altered 3rd NHT more common in their data. Davis' phrase was from his improvisation over *Blues By Five* (Jazzomat Research Project 2017), with the phrase starting on the semiquaver before beat 3 in bar 85. Coltrane's phrase was from his improvisation over *Nature Boy* (Jazzomat Research Project 2017).

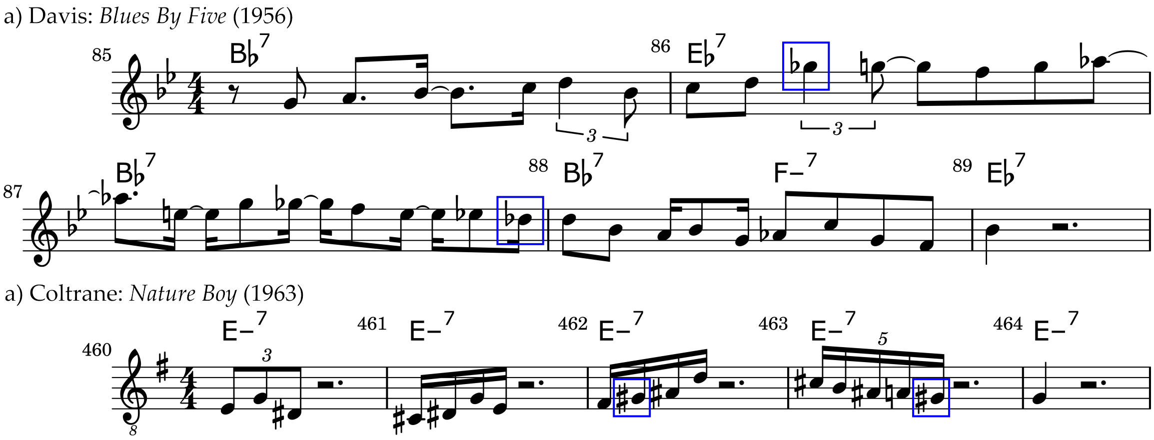

Figure 13.5: Example of phrases from Davis and Coltrane highlighting the use of the altered 3rd more common in their data. a) Davis, *Blues By Five* (1956), bars 85–89. b) Coltrane, *Nature Boy* (1963), bars 460–464.

13.1.4 Intervals

Many of the fuzzy interval classes were found to be important in identifying the performers. In the one-vs-one comparisons the big jump up class was found to be important in comparisons including Coltrane, while the step up class was important for Davis. For the one-vs-all comparisons the big jump down class was also found to be important, as was the repeated Parsons class. The classes found to be important with the n-way comparison were step down and leap down. All remaining classes were found to be important at least once. As the fuzzy intervals were found to be broadly important, the analysis focused on the raw underlying data. A χ^2-test found a significant difference in the distribution of fuzzy intervals between the performers, with a small effect size ($\chi^2(24) = 1258.59,\ p =< .001,\ V = .09$). Subsequent post-hoc tests found significant pairwise differences for all comparisons at $p < .001$. The distribution of fuzzy intervals can be seen in Figure 13.6.

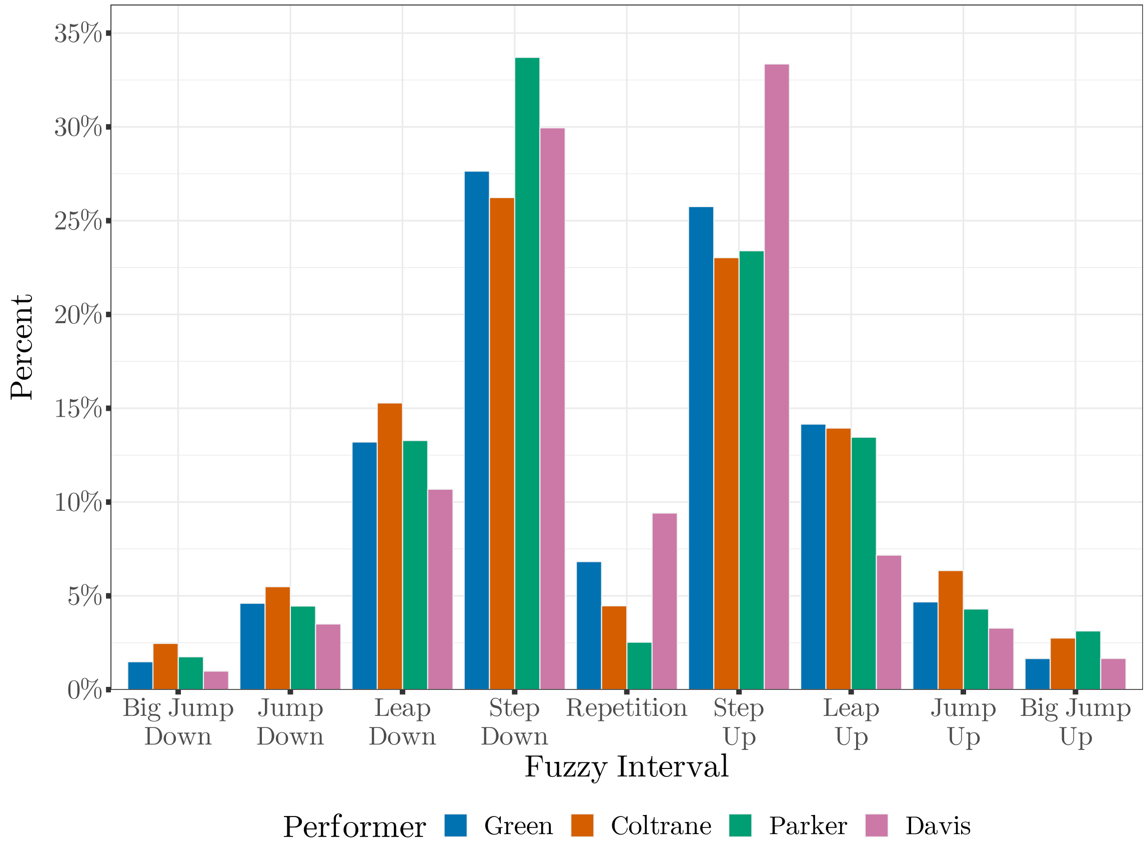

Figure 13.6: Fuzzy Interval distribution for each performer.

This data mostly aligned with the classes found to be important in the feature analysis. For example, Davis had a substantially higher proportion of step up fuzzy intervals while playing fewer leap ups. This suggested that Davis slightly favoured scalic over arpeggio movements. Davis also played a higher proportion of repeated notes, with Green playing more repeats than Coltrane or Parker. The absolute differences in the big jump fuzzy interval classes were not substantial, a spread of around 1.5PP. However, they were frequently found to be important in identifying the performers. The hypothesis was that except in classes with substantial differences (i.e. step up), the less frequently used classes were more useful in identifying the performers. These classes also tended to have greater proportional differences between the performers. For example, Coltrane played 2.49 times as many big jump downs as Davis, while Davis played 3.72 times as many repeats as Parker. These results indicated that when the majority of events fell within a minority of the classes, the classes with fewer events but larger marginal differences were more useful in classifying the performers. These results also highlighted the advantages of combining the feature analysis results with comparative or standard analyses. While the differences in the minority classes would have been noted, it was unlikely that such attention would have been drawn to them if not for the feature analysis.

13.2 Rhythm Domain

13.2.1 Beat Distribution

The beat division was identified as an important feature across comparisons that included Parker. Specifically, at the $\text{bar}_{4|2}$ level the most frequent division being one with eight tatums. Figure 13.7 shows two distributions of the division. The graph on the left shows the frequency of beats with each division across all of the performers' improvisations. The graph on the right shows the data used in the $\text{bar}_{4|2}$ abstraction, showing the division mode for each sliding window. In both graphs Parker has significantly more beats with division eight than the other three performers.[11]

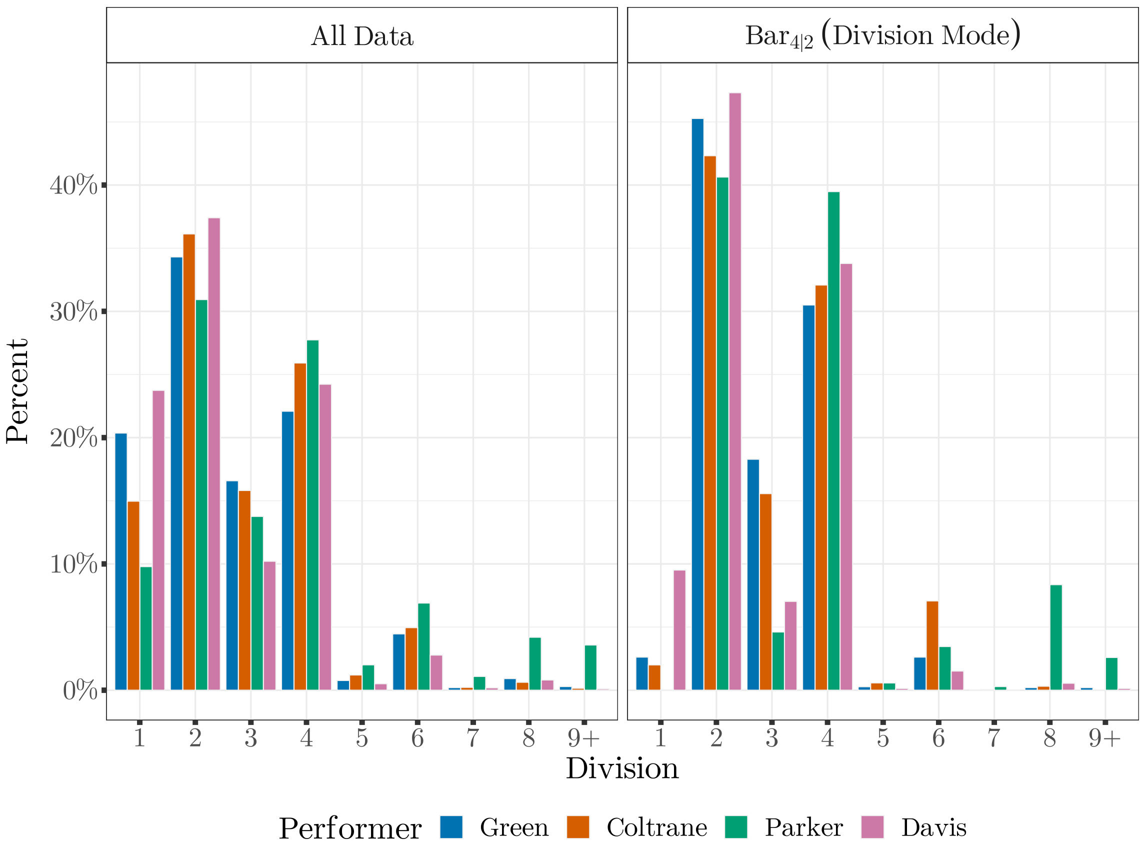

Figure 13.7: Distribution of divisions for each performer. Left: all data combined. Right: at the $\text{bar}_{4|2}$ abstraction level.

These results provided insight into elements of Parker's improvisational style. The differences in the two distributions, with Parker's division mode of eight being higher than the underlying frequency suggested that these beats were not evenly distributed in his improvisations. Investigation of Parker's data supported this hypothesis. Only

[11] A χ^2-test on the raw data found a significant difference in the distribution of divisions between the performers, with a small effect size ($\chi^2(24) = 1240.75,\ p =< .001,\ V = .13$). Subsequent post-hoc tests found significant pairwise differences between all performers at $p < .001$.

nine of the seventeen Parker transcriptions in the WJazzD had any beats with division eight, with four containing the vast majority of these beats, including:

- *Don't Blame Me* – 1947, 64 BPM, twenty-four beats;
- *Out Of Nowhere* – 1947, 68 BPM, twenty-three beats;
- *Embraceable You* – 1947, 72 BPM, nineteen beats;
- *How Deep Is The Ocean* – 1947, 72 BPM, sixteen beats.

All of these improvisations were recorded in the same year, although other Parker transcriptions from that year were present in the WJazzD. Additionally, these four improvisations were the only Parker transcriptions in the database with a tempo below 100 BPM. Therefore, these were likely double-time passages within Parker's improvisations. An example of this can be seen in Figure 13.8, which shows an excerpt Parker's improvisation over *Embraceable You* (Jazzomat Research Project 2017).

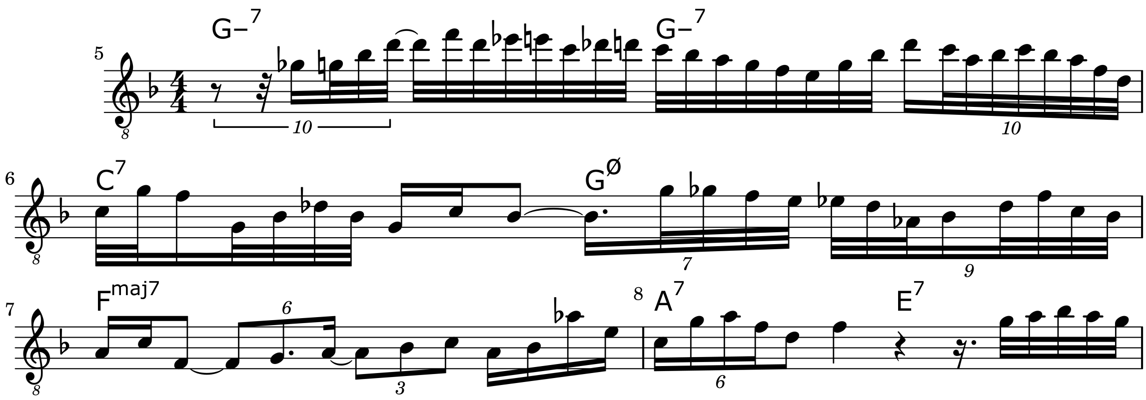

Figure 13.8: Example of Parker playing a double time phrase with beats of division 8, *Embraceable You* (1947), bars 5–8.

13.2.2 Metrical Density

The metrical density, as mean number of notes per bar, was widely found to be useful in classifying the performers. The distribution of notes per bar can be found in Figure 13.9. The raw data showed the distribution of notes per bar; the data for each abstraction showed the distribution of mean note per bar values. Across all four of the graphs there was a consistent trend of Green and Parker playing more notes per bar than either Coltrane or Davis. Parker consistently had the highest density bars while Davis had the lowest. Across all the data, 76.46% of metrical densities had a value between four and ten.

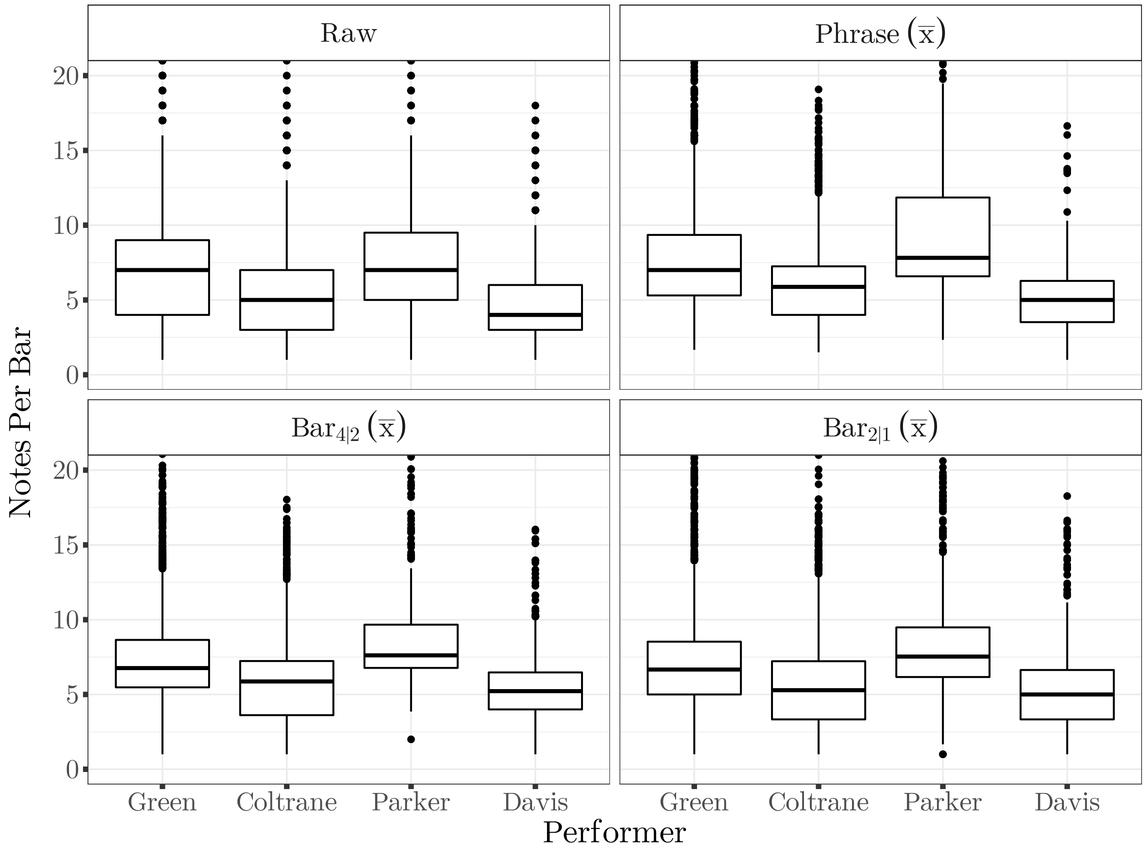

Figure 13.9: Distribution of metrical densities for each performer. Overall distributions across all data, and distributions of mean metrical densities for the phrase and bar levels.

An ANOVA was run on the raw data, and found a significant difference in the metrical density between the four performers, with a medium effect size ($F(3, 9018) = 297.95,\ p < .001;\ \eta^2 = .09$). A post-hoc test using Tukey's HSD procedure found significant pairwise differences for all comparisons at $p < .001$, except for Green vs. Parker where no significant difference was found ($p = .111$). These results indicated that both Green and Parker played significantly more notes per bar than either Coltrane or Davis. Green and Parker played on average around eight notes per bar, compared to six for Coltrane, and five for Davis.

13.2.3 Beat Weight

The most common beat weight class found to be important was the proportion of notes played in the first beat of each bar. Due to the small number of beat weight classes, and that each class was found to be important multiple times, all three classes were analysed. Figure 13.10 shows the overall distribution of beat weights for the four performers. As suggested from the one-vs-one feature analysis, Coltrane's distribution of notes played in the first beat of a bar was higher than the other performers. For the other performers, around a quarter of their notes were played in the first beat (25% $\pm$ 1.2PP). In contrast, 29.69% of Coltrane's notes were played in the first beat of a bar. All of Davis and Parker's improvisations were in $\frac{4}{4}$, while 90% of Coltrane's and 92.5% of Green's improvisations were in quadruple time.

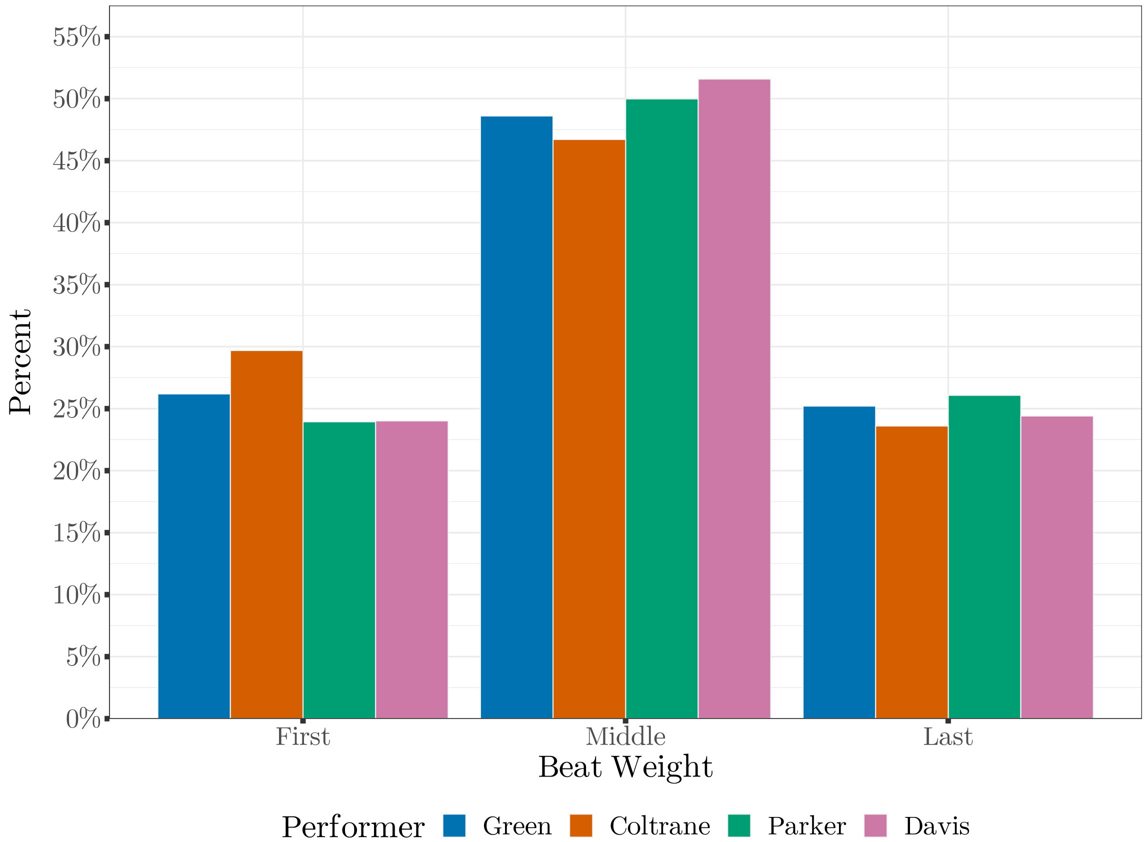

Figure 13.10: Proportion of notes played on each beat weight for all performers.

This distribution of time signatures could explain a small increase in first beat proportions. However, it would be expected that many more improvisations would need to be played in triple time to change the expected proportions substantially away from 25%/50%/25% (first/middle/last beat). Based on the number of note events Coltrane played over each time signature, the expected distribution, if all beats were equally likely, was 25.63%/48.74%/25.63%. These results indicated that Coltrane played a higher proportion of notes in the first beat in the bar.

A χ^2-test found a significant difference in the beat distribution between the performers, with a small effect size ($\chi^2(6) = 141.81,\ p =< .001,\ V = .04$). Subsequent post-hoc tests found significant pairwise differences between all comparisons at $p < .001$ except for: Green vs. Parker ($p = .003$); and Parker vs. Davis, where no significant difference was found ($p = .089$). Figure 13.11 shows a phrase from Coltrane's improvisation over *Blue Train* (Jazzomat Research Project 2017), where 61.54% of the notes were played in the first beat of the bars.

Figure 13.11: Example of Coltrane playing a high proportion of notes in the first beat of a bar, *Blue Train* (1957), bars 87–88.

13.3 Macro Domain

13.3.1 Phrase Features

There were three specific phrase features that appeared in the one-vs-all and one-vs-one comparison feature analyses. One was related to the start of phrases, the beat weight of the first note in a phrase. The other two described elements of the last note of the phrase: the fuzzy IOI of the final note; and the fuzzy interval used to move into the final note. The analysis of the phrase features focused on the data used to train the phrase level models.

The beat weight class found to be important in classifying the performers were phrases beginning in the last beat of a bar. This class was found especially useful in separating Green from the other performers. A χ^2-test found a significant difference in the distribution of starting phrase beat weights between the performers, with a small effect size ($\chi^2(6) = 158.12,\ p =< .001,\ V = .16$). Subsequent post-hoc tests found significant pairwise differences between all pairs of performers at $p < .001$ except for Parker vs. Davis and Coltrane vs. Davis where no significant difference was found ($p = .061$).

Figure 13.12 shows the distribution of beat weights for the first note of a phrase. This showed that, regardless of the performer, the plurality (Green and Coltrane) or majority (Parker and Davis) of phrases began in one of the middle beats of a bar. This data also supported the findings of the feature analysis, with Green beginning significantly more phrases in the last beat of the bar compared to Coltrane, Parker, and Davis. 32.58% of Green's phrases began in the final beat of a bar, compared to

$< 20\%$ for the other performers. The data indicated that Coltrane and Davis played more phrases that began in the first beat of the bar than the final beat.

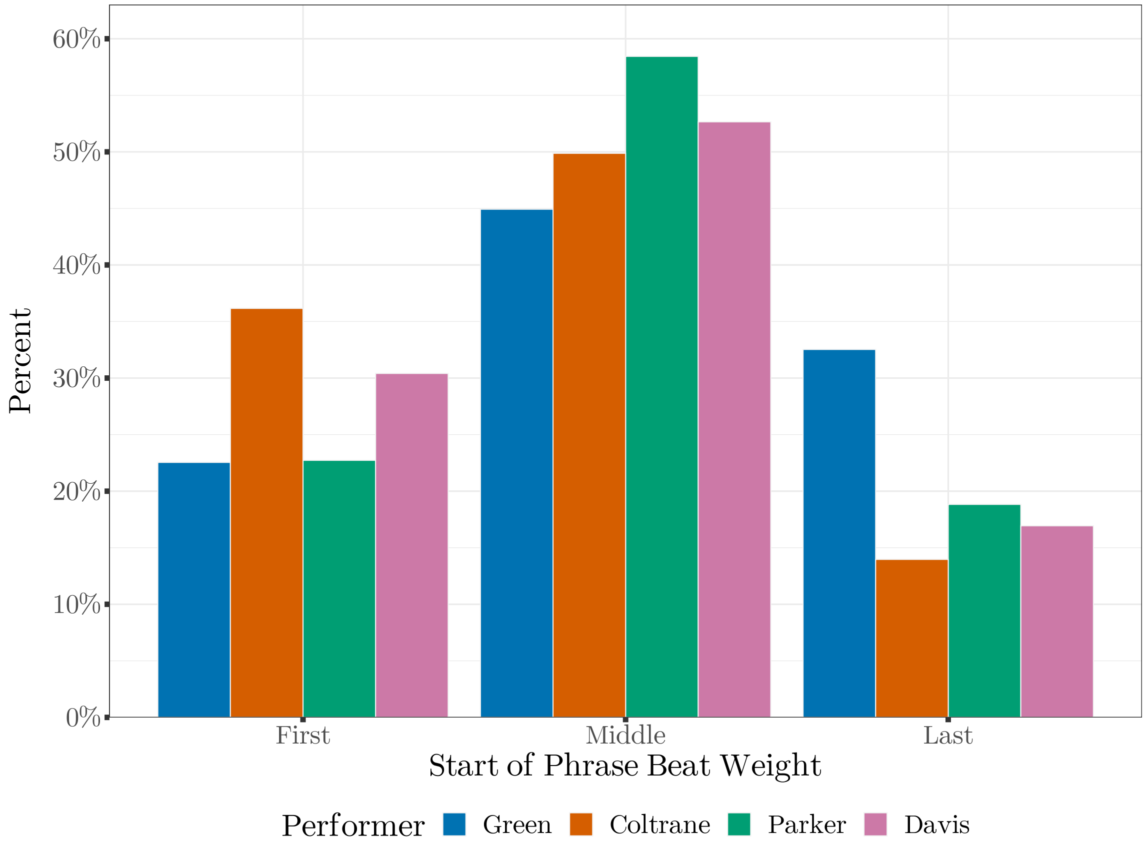

Figure 13.12: Distribution of starting notes of phrases across the beat weights for each performer.

Two fuzzy interval classes into the final note of a phrase were frequently found to be important: repeated notes and jump downs. A χ^2-test was run to analyse the entire distribution of phrase ending fuzzy intervals between the performers, finding a significant difference with a medium effect size ($\chi^2(6) = 158.12,\ p =< .001,$ $V = .16$). Subsequent post-hoc tests found significant pairwise differences between all performers, with Parker vs. Coltrane and Green vs. Davis at $p = .010$ and all others at $p < .001$. The distribution of phrase ending fuzzy intervals can be found in Figure 13.13. For all four performers, the plurality of phrases ended with a step down into the final note of the phrase. The largest differences between the performers were in the repetition and jump down classes. Specifically, Green and Davis were more likely to end a phrase with a repeated note while Coltrane and Parker were more likely to end the phrase with a jump down. Coltrane and Parker were also more likely to play a big jump down into the final note of a phrase. For Green, the use of a repeated note was the second most common phrase ending, while it was the third most common ending for Davis.

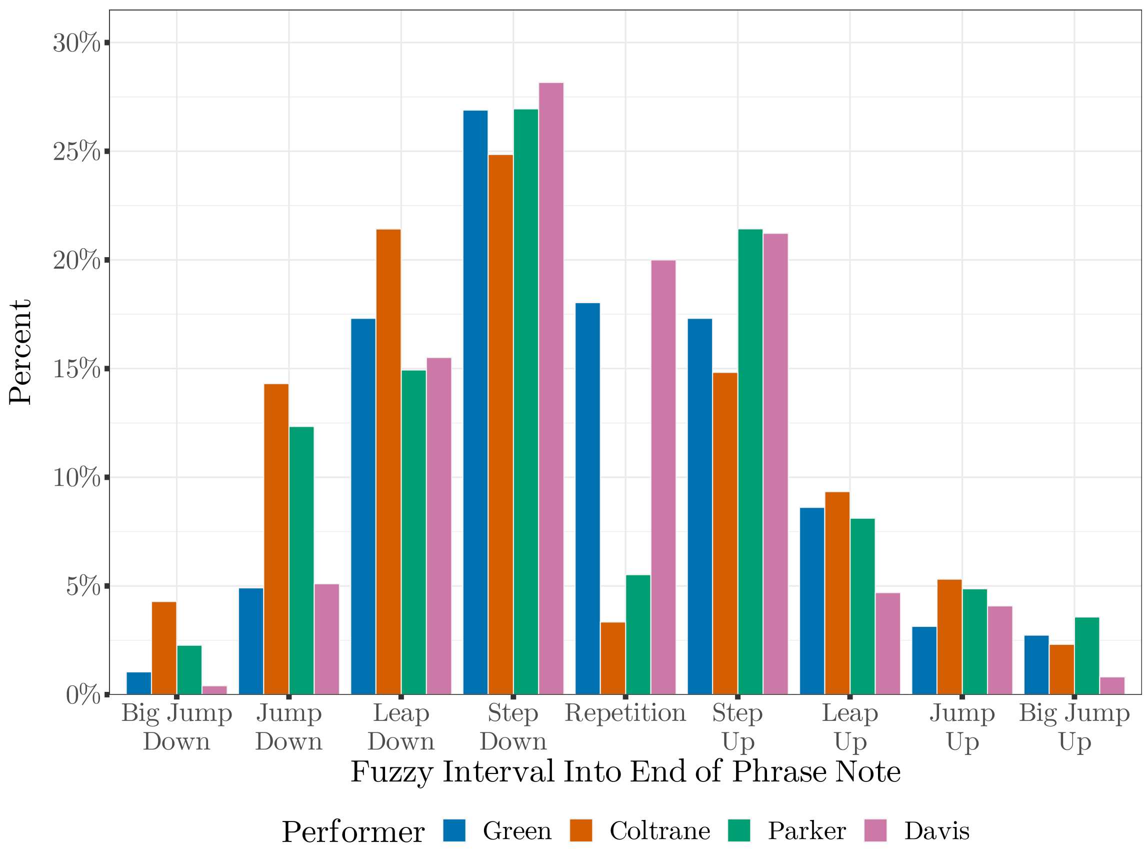

Figure 13.13: Distribution of fuzzy intervals into the final note of a phrase for each performer.

There were two end of phrase fuzzy IOI classes that were found to be important: medium length notes for Green vs. All and Green vs. Davis; and short notes for Green vs. Parker. A χ^2-test found that the distribution of end of phrase fuzzy IOI differed significantly between the performers, with a small effect size ($\chi^2(6) = 158.12,\ p =< .001,\ V = .16$). Subsequent post-hoc tests found significant pairwise differences in all comparisons at $p < .001$, except Coltrane vs. Davis ($p = 0.011$) and Green vs. Parker ($p = .634$), the latter which was not found to be significantly different. The distribution of end of phrase fuzzy IOI for the four performers can be seen in Figure 13.14. The graph showed that the distributions of Green and Parker were most similar to each other. Regardless of the performer, the majority of phrases played ended with a long fuzzy IOI note.

The reported importance of medium fuzzy IOI notes for Green vs. All and Green vs. Davis fit the data, with Green having a higher proportion of phrases ending with a medium fuzzy IOI note than Davis. However, the difference in short fuzzy IOI ending notes for Green vs. Parker was not as obvious. Green only had a slightly higher proportion of phrases that ended with a short fuzzy IOI, 3.38% compared to Parker's 2.27%. The frequent importance of the short class was likely due to the same effect of majority vs. minority classes found in other features. The results of the analysis indicated that the vast majority of phrases from all performers ended with a long note, with medium notes also more common for Green and Parker.

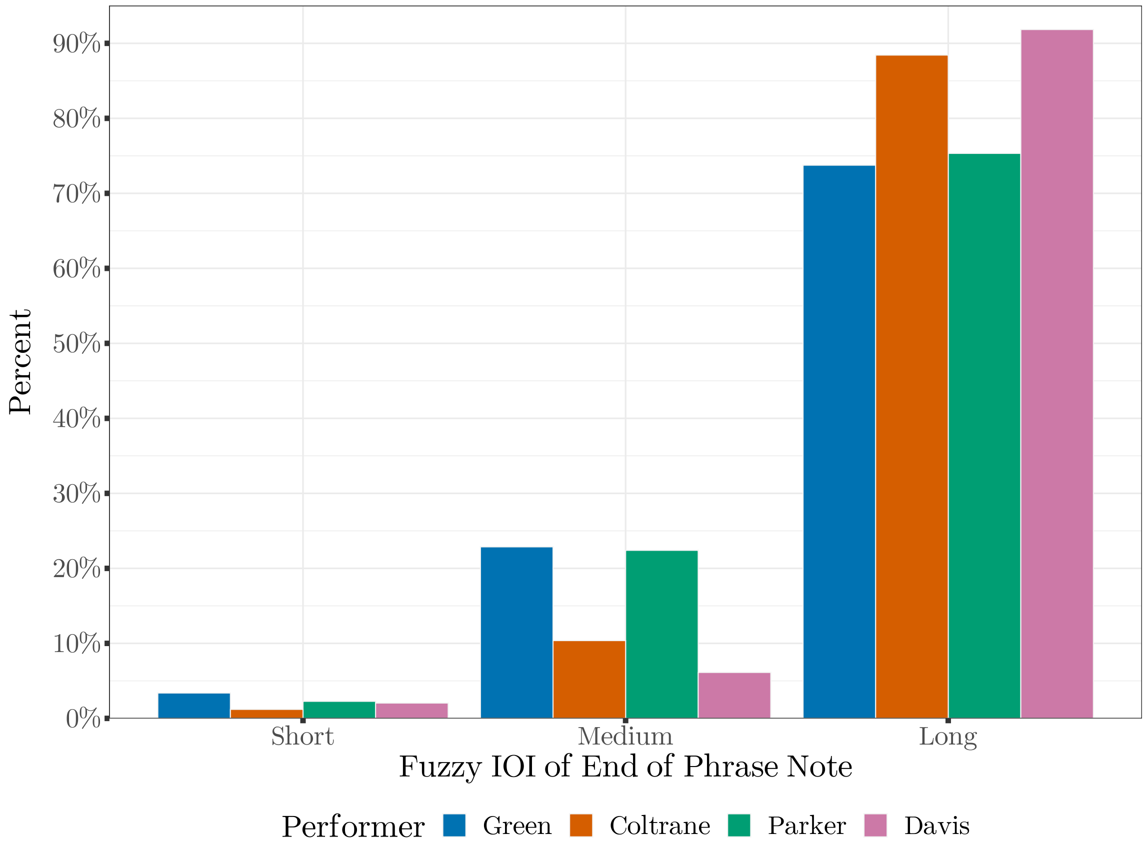

Figure 13.14: Distribution of fuzzy IOI for the last note of a phrase for each performer.

In summary, all performers were most likely to begin their phrases in the middle beats of a bar; however, Green was more likely to begin his phrases in the last beat of the bar while Coltrane and Davis were more likely to begin in the first beat. Green and Davis were more likely than Coltrane and Parker to end with a repeated note, while the opposite was true for ending phrases with a jump down fuzzy interval. The vast majority of all phrases ended with a long fuzzy IOI note. However, more than a fifth of Green and Parker's phrases also ended with a medium fuzzy IOI. An example phrase from both Green and Parker is shown in Figure 13.15, highlighting some of their specific phrase tendencies. The excerpts come from Green's improvisation over *Tico-Tico* (1962m), and Parker's improvisation over *Yardbird Suite* (Jazzomat Research Project 2017). Green's phrase started on the last beat of the bar and ended with a repeated medium fuzzy IOI note. Parker's ended with a medium fuzzy IOI with the interval into the final note being a fuzzy jump down.

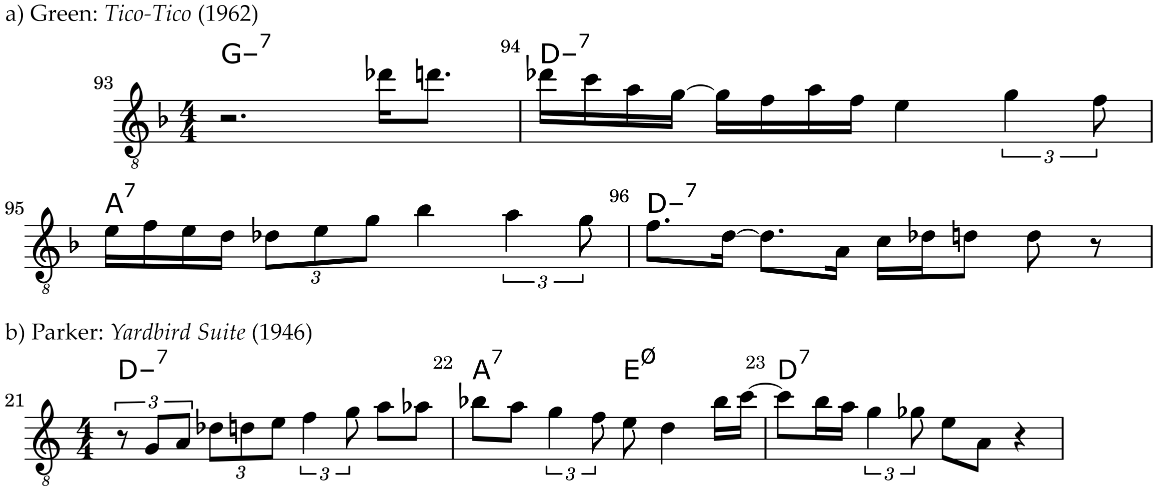

Figure 13.15: Examples showing specific beginning and ending phrase features in Green and Parker's improvisations. a) Green: *Tico-Tico* (1962), bars 93–96. b) Parker: *Yardbird Suite* (1946), bars 21–23.

13.4 Comparative Analysis Summary

This chapter explored an example comparative analysis based on the results of the feature importance of the C5.0 models. Nine features were selected for the comparative analysis, with features coming from the pitch, rhythm, and macro domains. As this comparative analysis only aimed to show one possible implementation of the results of the classification it was limited in scope.

The results of the analysis found that, as hypothesised, the combination of the raw pitch features, NITR and octave, were used partially as a proxy for instrument identification. The TPC analysis provided only limited insight into the differences in the performers improvisational style, with the largest differences more indicative of tonality differences than note choice. The CDPCX analysis focused on the two altered 3rd classes, which were most frequently found to be important. The analysis found that Coltrane played a higher proportion of ♭3/♮3, while Davis had the highest proportion of ♮3/♭3. However, this was likely influenced by the distribution of chords in their data. The interval analysis was an example of where the feature analysis aided in highlighting specific classes that might have otherwise been missed. The feature analysis indicated that some of the rarer fuzzy interval classes were more important in classifying the performers, despite smaller absolute differences in proportions. The analysis found that Green and Davis played identifiably fewer big jump fuzzy intervals than Coltrane or Parker. The analysis also found that repetitions were more common in Green and Davis' improvisations, and that Davis favoured an ascending step over a leap.

The beat distribution analysis focused on beats with a division of eight, as this was identified as being specifically important for classifying Parker. Across both the entire dataset and the $bar_{4|2}$ dataset, Parker played beats with a division of 8 significantly more frequently than the other three performers. The metrical density analysis also supported this finding, with it additionally finding that Parker and Green had the highest metrical density. This indicated that the WJazzD may not have included many of Coltrane's improvisations from his "sheets of sound" era. The beat weight analysis found that Coltrane played a higher density of notes in the first beat of a bar.

Coltrane also began his phrases in the first beat of the bar significantly more often than Green or Davis. In contrast, Green was significantly more likely to start a phrase in the last beat of the bar. However, the plurality of all phrases began in one of the middle beats of the bar. The analysis found that Coltrane and Parker were significantly more likely to end a phrase with a jump down. In comparison, Green and Davis were significantly more likely to end a phrase with a repeated note. The final phrase analysis focused on the fuzzy IOI of the last note of a phrase. This found that although the vast majority of all phrases ended with a long fuzzy IOI note, there were differences between the performers. Specifically, more than a fifth of Green and Parker's phrases ended with a medium fuzzy IOI, compared to $\leq 10\%$ of Coltrane or Davis'.

Many of these features may have be chosen for a comparative analysis without the training, evaluation, and feature extraction of the performer classification. However, that process acted as a first pass, taking the role of exploratory data analysis. Additionally, by highlighting both specific features and classes, the analysis could focus on those elements and examine their differences more closely. The specific features and classes identified as important by the C5.0 classifier allowed the comparative analysis to find distinct, and sometimes subtle, differences in the performer's improvisational styles.

Chapter 14

Findings of Performer Classification and Comparative Analysis

Part III focused on exploring the methodology developed in Chapter 3 to undertake a performer classification and comparative analysis task. First, informed by the results of the analysis of Green, the prior literature, and initial experiments of training models, the features selected for each abstraction level were presented. The models were then trained, and the performance of the models evaluated. The model results focused both on the performance metrics of the trained models, and the feature importance scores. The importance scores identified the features and classes that were most useful in classifying the performers. Finally, based on the feature analysis, an example comparative analysis was undertaken that explored the similarities and differences between Green, Coltrane, Davis, and Parker's improvisational styles.

The feature selection, informed by the analysis of Green's improvisational style, was also influenced by the features available and how fundamental they were as building blocks of music. The analysis of Green aided in the creation of features, including rests and $\text{CPC}_{\text{Weight}}$, that were then used within the performer classification task. All of the selected features were split into three broad types of transformation: measures of centre; proportions (OHE categorical variables); and counts or descriptors. Common simplified features included: raw pitch; TPC; CPC; intervals; note length; metrical weight; beat distribution; note placement; metrical density; rests; swing; and gradient. As many of the fundamental building blocks of music related to features in the pitch and rhythm domains, these domains were most represented within the input data. Other important features, including the tempo or articulation, had to be excluded as they caused the classifiers to overfit on the training data. Additionally, extra input features were required at the phrase level to represent phrase descriptors and how the improvisers started and ended their phrases.

The trained models performance metrics found that the two best performing classifiers were C5.0 and RF. However, the C5.0 classifier tended to outperform the RF classifier and trained the models fifteen times faster on average. The C5.0 classifier also massively outperformed the other classifiers at the note level. These results indicated that the C5.0 was most useful for this particular interpretable performer classification task based solely on improvisational data. The model results also found that the classifiers were able to achieve a high level of accuracy on the testing dataset. The best result from any model came from the C5.0 Green vs. Coltrane comparison at the note level, with an MCC of 0.98 and only nine misclassifications (accuracy: 98.76%). All classifiers tended to struggle the most with identifying Parker, while Davis was especially identifiable at the phrase level.

Due to the exceptional performance of the C5.0 classifier, and its efficiency in training the models, the feature analysis focused on only the significant C5.0 models. The feature analysis results indicated that although the n-way models may not have had the highest performance metrics, the features identified tended to carry throughout the one-vs-all and one-vs-one models. However, the one-vs-all and one-vs-one comparisons aided in identifying specific features and classes that were important to individual improvisers. Influenced by the feature selection, the feature analysis found that fundamental musical features were more commonly found to be important for classifying the performers. These included: raw pitch; CPC; intervals; beat distribution; beat weight; metrical density; swing; and phrase features.

Combined, these results indicated that similar future projects could train n-way models with a C5.0 classifier as a good first filter for features to investigate in a comparative analysis. Additional one-vs-all or one-vs-one models could then be trained to find more specific features and classes for individual soloists, but training both sets would often be unnecessary. The training of models at separate abstraction levels proved generally useful. However, this could be simplified down into only three or fewer levels, e.g. phrase level, one of the two bar levels, and note level.[1] The selection of the two bar levels did not appear to be very important, with both returning similar results. Future analyses could consider the average length of phrases (in bars) and select the bar level that would best fit between the phrase and note level.

The results of the performance metrics from the models informed the investigation into the feature importance, which subsequently formed the basis of the comparative

[1]More research with a larger dataset would be required to investigate the usefulness and applicability of a solo level comparison.

analysis. This comparative analysis focused on the similarities and differences of the improvisational styles of Green, Coltrane, Davis, and Parker through the examination of nine features. Although many of the features investigated would be included in many comparative analyses, the feature importance analysis aided in identifying specific classes within the features to study. For example, the feature analysis identified that for classifying Parker a beat with a division of eight was particularly useful. The comparative analysis found that Parker was significantly more likely than the other performers to play beats with divisions of eight or more. This result, combined with the metrical density analysis, indicated that few improvisations from Coltrane's "sheets of sound" era were included in the WJazzD. The interval analysis found that Green and Davis were significantly more likely to play repeated notes compared to Coltrane or Parker. Small absolute differences, including the frequency of big jump fuzzy intervals, were also identified through the feature analysis. The comparative analysis found that big jump fuzzy intervals were more common in Coltrane and Parker's improvisations.

In summary, the performer classification and comparative analysis was broadly successful. The C5.0 classifier was able to create models that performed very well on the testing data. This showed that there were identifiable differences between the four performers' improvisational styles throughout the abstraction levels. The feature importance scores from these models were useful in identifying specific features and classes on which to base the comparative analysis of the performers. This comparative analysis, while limited in scope, did identify specific similarities and differences in the improvisational styles of Green, Coltrane, Davis, and Parker.

Chapter 15

Conclusion

This research developed a new methodology for analysing improvised jazz using computer-aided and statistical methods. Part I reviewed the previous literature from four main fields of study: analysis of improvised jazz; computer-aided musical analysis; computer-aided jazz analysis; and machine learning in music. This literature then built the foundations for the methodologies presented in Chapter 3, Approaching a New Methodology. Part II explored the methodology developed for investigating a single performer's improvisational style through an exploration of Green's improvisations between 1960 and 1965. The results of this analysis then aided in the research presented in Part III, a performer classification and comparative analysis. This task used interpretable ML algorithms to classify Green, Coltrane, Parker, and Davis based solely on the improvisational content of their solos. The features and classes found to be important in classifying the performers were then used as the basis of an example comparative analysis.

The overarching aim of this research was to develop a new methodology for systematically analysing the improvisational style of a performer through the use of computer-aided and statistical methods. Despite the limitations and issues discussed in Chapter 3, the methodology developed and explored throughout this research was effective. It allowed for both a broad investigation of Green's improvisational style, while also providing an approach for deep analysis into important features. The computer-aided and statistical methods provided the flexibility and speed to undertake both broad and deep analyses of Green's improvisational style. The results of Green's analysis found that he conformed to many of the expected customs of the time and styles in which he improvised. These results could therefore be used as a baseline or point of comparison for future work. The analysis also provided insight into distinct elements of Green's improvisation style, including:

- Green's note choices were predominantly diatonic and harmonic;
- There was frequent evidence of blues influenced language;
    - Green specifically favoured the ♭3 blues note over the TT;

- The frequency of NHTs in Green's improvisations increased in the beats leading up to a chord change;
- Repeated notes, while not common throughout Green's corpus, were more likely to be played again in the 200 notes following a repeated note;
    - This indicated that there were certain improvisations when repeating notes was a critical feature of Green's improvisational style;
    - When Green played repeats he was also likely to vary the articulation of the notes, compensating for the lack of pitch change;
- Green played a median six notes per bar, with this increasing at lower tempos and decreasing at higher tempos;
- When Green improvised at lower tempos his rhythmic variety and complexity of sub-beat placements increased;
- Green rarely played rests within a phrase, although when he did they went for around half a beat;
- Green tended to swing harder than many of the performers in the WJazzD, with Green also swinging harder when improvising over a blues;
- Green played predominantly behind the beat;
- Green's most common phrase contours were convex and descending, although longer phrases tended to be horizontal.

The main focus of this research was on the development of the methodology to analyse a single performer's improvisational style. Part III then explored a potential application of these results. The performer classification task found that interpretable ML algorithms were able to successfully identify performers based solely on their improvisational data. The results found that the C5.0 classifier was both the best performing and fastest of the three classifiers tested. The results also found that the n-way comparison provided an initial filter for selecting features to investigate, with one-vs-all and one-vs-one comparisons highlighting specific features and classes that related to individual improvisers. This classification task also found that there was currently not enough solo level data to draw any conclusions at that abstraction level. A subset of three other abstractions (phrase, bar, and note), provided a broad set of features on which to base a comparative analysis. The comparative analysis was an example of a possible application of the results of the classification task. Consequently, the comparative analysis was limited in scope and focused on only a small subset of all features that were found to be important for identifying the performers. The comparative analysis found distinct differences and similarities between Green, Coltrane, Parker, and Davis' improvisational styles, including:

- Parker was more likely to play denser beats (division ≥ 8) when compared to the other performers, within the available data;
- Green and Davis played more repeated notes than Coltrane and Parker;
- Parker and Coltrane were more likely to play large intervals ($\geq$ minor 6^{th});
- Coltrane played a higher density of beats in the first beat of the bar compared to the other performers.

In summary, this research found that the methodologies developed provided an effective and efficient approach for undertaking analyses of performer improvisational style. The methodology allowed for both broad and deep investigations into a wide range of features across the four defined domains. Similarly, the application of these results to a performer classification task were also successful, with the models trained able to accurately identify the performers in a range of abstraction levels and comparison types. Finally, the results of the classification provided a set of features that were able to form the basis of a comparative analysis, finding distinct similarities and differences in improvisational style.

Future Work

Through undertaking this research there were numerous areas where additional research or work was required. This project broadened the available high quality data by creating a large dataset of guitar transcriptions. However, there is a clear and present need for high quality data with a wider representation of diversity. This includes instrumentation, gender, and nationality. Future research should take this into consideration and create more representative datasets. Larger and more diverse datasets will also aid in compensating many of the specific limitations of analysis found in this research (e.g. tempo, key signatures, or time signatures).

A continuing issue is the time investment required to complete many high quality transcriptions. Since the beginning of this project there have been developments in the field of source-separation. An example of this was the development of the open source project *Spleeter* (Hennequin et al. 2020). *Spleeter* is an open source project by deezer research that splits tracks into individual instrument stems using their pre-trained models.[1] Their pre-trained models are not purpose-trained for the separation of jazz tracks; however, they do provide the ability to train up new models. New models trained on jazz improvisations could then be used to separate improvisations into individual instruments. High quality instrument separated tracks would then allow for automatic transcription and annotation tools

[1] *Spleeter* is a Python based project using Tensorflow.

(e.g. automatic pitch or beat transcription) to have a higher success rate. This would reduce the time required for manual transcription and annotation. These tools would allow researchers to more quickly and efficiently create large datasets, while also reducing the barrier to entry. Larger and more diverse datasets would benefit the field of jazz computer-aided musicology, spurring even more transcriptions and furthering the computer-aided analysis of improvised jazz.

Two additions or extensions to *MeloSpy* would greatly improve its applicability to jazz analysis. The first is polyphony, which as discussed in Chapter 3 is a complex issue to solve. The complexity lies not only in the transcription of the data, but also the storage, feature extraction, and analysis. The other is the addition of harmonic analysis.[2] Beyond investigation of the chord type, harmonic analysis allows for greater study of the function of the chords played. Harmonic analysis would provide greater detail to the analysis of the notes played within an improvisation by considering not only how they related to the chord of the moment, but to the broader function of that chord within a piece.

This research has demonstrated the application of computer-aided and statistical methods to the analysis of a performers improvisational style. It has shown the applicability of this approach while also providing an understanding of Green's improvisational style between 1960 and 1965. Through showing the possibilities of this approach, this research forms part of the foundations for further computer-aided musicological research into jazz improvisation.

[2]See Wilding (2008), Choi (2011), and De Haas et al. (2014) for some computer-aided or automatic approaches to jazz harmonic analysis.

List of Sources

Bibliography

Abeler, Niko. 2015. "Musical Composer Identification". Accessed November 15th, 2017. http://hpac.rwth-aachen.de/teaching/sem-mus-15/reports/Abeler.pdf.

Abeßer, Jakob. 2016. "Jazz Solo Analysis Between Music Information Retrieval, Music Psychology and Jazz Research". Presentation at the 17th International Society For Music Information Retrieval Conference, New York City. Accessed March 28th, 2017. https://github.com/jazzomat/knowhow/blob/master/ISMIR_2016_Tutorial/ISMIR_2016_Jazzomat_Tutorial_Part_3.pdf.

Abeßer, Jakob, Christian Dittmar, and Holger Großmann. 2008. "Automatic Genre and Artist Classification By Analyzing Improvised Solo Parts from Musical Recordings". In *Proceedings of the Audio Mostly Conference - A Conference on Interaction With Sound*, 127–131. Piteå, Sweden. Accessed December 4th, 2017. http://www.academia.edu/download/30943522/proceedings_am08_low.pdf#page=129.

Abeßer, Jakob, Klaus Frieler, Estefanía Cano, Martin Pfleiderer, and Wolf-Georg Zaddach. 2017. "Score-Informed Analysis of Tuning, Intonation, Pitch, Modulation, and Dynamics In Jazz Solos". *IEEE/ACM Transactions on Audio, Speech and Language Processing* 25 (1): 168–177. Accessed February 23rd, 2017. http://ieeexplore.ieee.org/document/7740101/.

Abeßer, Jakob, Klaus Frieler, Martin Pfleiderer, and Wolf-Georg Zaddach. 2013. "Introducing the Jazzomat Project - Jazz Solo Analysis Using Music Information Retrieval Methods". In *Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research: Sound, Music and Motion*, 653–661. Marseille, France: Publications of L.M.A. Accessed February 23rd, 2017. http://www.mu-on.org/frieler/docs/cmmr2013_jazzomat.pdf.

Ali, Muhammad Asim and Zain Ahmed Siddiqui. 2017. "Automatic Music Genres Classification Using Machine Learning". *International Journal of Advanced Computer Science and Applications* 8 (8): 337–344. Accessed November 15th, 2017. https://dx.doi.org/10.14569/IJACSA.2017.080844.

Alton-Lee, Amity Rose. 2010. "The Improvisation of Tubby Hayes In 'The New York Sessions'". Masters Thesis, New Zealand School of Music. http://researcharchive.vuw.ac.nz/bitstream/handle/10063/1200/thesis.pdf?sequence=1.

Angus, Quentin. 2014. "Phrasing and Polyrhythm In Contemporary Jazz Guitar: A Portfolio of Recorded Performances and Exegesis". PhD Thesis, University of Adelaide. http://hdl.handle.net/2440/86585.

Babbitt, Milton. 1965. “The Use of Computers in Musicological Reserach”. *Perspectives of New Music* 3 (2 (Spring – Summer)): 74–83. Accessed March 21st, 2022. https://www.jstor.org/stable/832505.

Bakker, Arthur, Jinfa Cai, Lyn English, Gabriele Kaiser, Vilma Mesa, and Wim Van Dooren. 2019. “Beyond small, medium, or large: points of consideration when interpreting effect sizes”. *Educational Studies in Mathematics* 102: 1–8. Accessed November 22nd, 2020. https://doi.org/10.1007/s10649-019-09908-4.

Beaty, Roger E., Klaus Frieler, Martin Norgaard, Hannah Merseal, Maryellen C. MacDonald, and Daniel J. Weiss. 2022. “Expert Musical Improvisations Contain Sequencing Biases Seen In Language Production”. *Journal of Experimental Psychology: General* 151 (4): 912–920. Accessed May 18th, 2022. https://doi.org/10.1037/xge0001107.

Bechtel, Teague Stefan. 2018. “Grant Green: An Analysis Of The Blue Note Guitarist’s Musical Vocabulary”. Doctor of Arts Dissertation, University of Nothern Colorado. https://digscholarship.unco.edu/dissertations/479.

Bellaviti, Sean David. 2005. “Combining Macro and Micro Approaches to Jazz Performance Analysis”. Masters Thesis, Toronto, Ontario: York University. http://search.proquest.com.proxy.library.adelaide.edu.au/pagepdf/305367228/.

Benadon, Fernando. 2006. “Slicing The Beat: Jazz Eight-Notes As Expressive Microrhythm”. *Ethnomusicology* 50 (1): 73–98. Accessed February 21st, 2022. http://www.jstor.org/stable/20174424.

Beran, Jan. 2003. *Statistics In Musicology.* CRC Press.

Berliner, Paul F. 1994. *Thinking In Jazz: The Infinite Art of Improvisation.* Chicago: University of Chicago Press.

Brooks, Ronald M. 2008. “The Improvisational Style of Terence Blanchard”. PhD Thesis, University of Northern Colorado. http://search.proquest.com.proxy.library.adelaide.edu.au/docview/304541164/.

Burgoyne, John Ashley, Jonathan Wild, and Ichiro Fujinaga. 2013. “Compositional Data Analysis of Harmonic Structures In Popular Music”. In *Proceedings of the 4th International Conference on Mathematics and Computation In Music*, edited by Jason Yust, Jonathan Wild, and John Ashley Burgoyne, 52–63. Berlin: Springer Berlin Heidelberg. Accessed June 24th, 2017. https://link-springer-com/content/pdf/10.1007%2F978-3-642-39357-0_4.pdf.

Butterfield, Matthew W. 2011. “Why Do Jazz Musicians Swing Their Eight Notes?” *Music Theory Spectrum* 33 (1): 3–26. Accessed February 21st, 2022. https://doi.org/10.1525/mts.2011.33.1.3.

Centre for Digital Music at Queen Mary, University of London. 2017. “QM Vamp Plugins”. *Vamp Plugins.* Accessed February 26th, 2017. https://vamp-plugins.org/plugin-doc/qm-vamp-plugins.html.

Chicco, Davide and Giuseppe Jurman. 2020. “The Advantages fo the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation”. *BCM Genomics* 21 (6): 1–13. Accessed May 26th, 2021. https://doi.org/10.1186/s12864-019-6413-7.

Choi, Andrew. 2011. “Jazz Harmonic Analysis as Optimal Tonality Segmentation”. *Computer Music Journal* 35 (2): 49–66. Accessed February 17th, 2022. https://doi.org/10.1162/COMJ_a_00056.

Choi, Keunwoo, György Fazekas, Kyunghyun Cho, and Mark Sandler. 2017. “A Tutorial on Deep Learning For Music Information Retrieval”. *Corr*: 1–16. Accessed January 26th, 2018. http://arxiv.org/abs/1709.04396.

Christiansen, Corey. 2003. *Essential Jazz Lines In the Style of Grant Green For Guitar*. Pacific, MO: Mel Bay.

Cohen, Jacob. 1988. *Statistical Power Analysis For the Behavioral Sciences.* 2nd edition. Lawrence Erlbaum Associates.

Cook, Scott Alexander. 2012. “Referential Sets, Referential Tonics, and the Analysis of Contemporary Jazz”. PhD Thesis, Vancouver: the University of British Columbia. https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0072739.

Corcoran, Christopher and Klaus Frieler. 2021. “Playing It Straight: Analyzing Jazz Soloists’ Swing Eight-Note Distributions With The Weimar Jazz Database”. *Music Perception* 38 (4): 372–385. Accessed April 29th, 2021. https://doi.org/10.1525/MP.2021.38.4.372.

Corrêa, Débora C. and Francisco Ap. Rodrigues. 2016. “A Survey on Symbolic Data-Based Music Genre Classification”. *Expert Systems With Applications* 60: 190–210. Accessed November 15th, 2017. http://dx.doi.org/10.1016/j.eswa.2016.04.008.

Crook, Hal. 1991. *How To Improvise: An Approch To Practicing Improvisation.* Advance Music.

Cross, Peter and Andrew Goldman. 2021. “Interval Patterns Are Dependent On Metrical Position In Jazz Solos”. *Musicae Scientiae* OnlineFirst. Accessed August 31st, 2021. https://doi.org/10.1177/10298649211033973.

Cuthbert, Michael Scott. 2017. “What Is Music21?” *Music21.* Accessed June 27th, 2017. http://web.mit.edu/music21/doc/about/what.html.

Cuthbert, Michael Scott and Christopher Ariza. 2010. “Music21: A Toolkit For Computer-Aided Musicology and Symbolic Music Data”. In *Proceedings of the 11th International Society For Music Information Retrieval Conference*

(ISMIR 2010), edited by J Stephen Downie and Remco C Veltkamp, 637–642. Utrecht, Netherlands: International Society For Music Information Retrieval. Accessed May 9th, 2017. http://hdl.handle.net/1721.1/84963.

Cuthbert, Michael Scott, Christopher Ariza, and Lisa Friedland. 2011. "Feature Extraction and Machine Learning on Symbolic Music Using the music21 Toolkit". In *Proceedings of the 12th International Society For Music Information Retrieval Conference*, edited by Anssi Kalpuri and Colby Leider, 387–392. Miami: International Society For Music Information Retrieval. Accessed November 15th, 2017. http://ismir2011.ismir.net/ISMIR2011_complete_proceedings.pdf.

De Haas, W. Bas, José Pedro Magalhães, Frans Wiering, and Remco C. Veltkamp. 2014. "Automatic Functional Harmonic Analysis". *Computer Music Journal* 37 (4): 37–53. Accessed February 17th, 2022. https://doi.org/10.1162/COMJ_a_00209.

Dittmar, Christian, Martin Pfleiderer, and Meinard Muller. 2015. "Automated Estimation of Ride Cymbal Swing Ratios In Jazz Recordings". In *Proceedings of the 16th International Society For Music Information Retrieval Conference*, edited by Meinard Muller and Frans Wiering, 271–277. Malaga, Spain. Accessed March 2nd, 2017. http://ismir2015.uma.es/docs/ISMIR2015_Proceedings.pdf.

Dixon, Simon, Hélène Papadopoulos, Martin Pfleiderer, and Gabriel Solis. 2017. "Dig That Lick: Analysing Large-Scale Data For Melodic Patterns In Jazz Performances". *Digging into Data.* Digging into Data. Accessed March 31st, 2017. https://diggingintodata.org/awards/2016/project/dig-lick-analysing-large-scale-data-melodic-patterns-jazz-performances.

Driever, Leonhard, Mels Loe Jagt, Kuan Lon Vu, Daniel Harasim, Andrew McLeod, and Martin Rohrmeier. 2022. "Improving Chord Prediction in Jazz Music using Melody Information". In *Proceedings of the 19th Sound and Music Computing Conference*, edited by Romain Michon, Laurent Pottier, and Yann Orlarey, 146–153. SMC Network. Accessed July 16th, 2022. https://apmcleod.github.io/pdf/chord-melody.pdf.

Eppler, Arndt, Andreas Mannchen, Jakob Abeßer, Christof Weiß, and Klaus Frieler. 2014. "Automatic Style Classification of Jazz Records With Respect to Rhythm, Tempo, and Tonality". In *Proceedings of the 9th Conference on Interdisciplinary Musicology*, edited by Timour Klouche and Eduardo R. Miranda, 162–167. Berlin, Germany. Accessed February 23rd, 2017. http://www.mu-on.org/frieler/docs/CIM14_eppler_et_al_jazz_classification.pdf.

Erlwine, Michael. 2017. “Grant Green”. *Allmusic*. Accessed February 24th, 2017. http://www.allmusic.com/artist/grant-green-mn0000738855/biography.

Frederico, Gustavo Cesar De Souza. 2006. “Feature Selection and Evaluation For Genre Classification of Symbolically Encoded Classical Music With the Aid of Machine Learning”. Masters of Computer Science Thesis, University of Ottawa. https://ruor.uottawa.ca/bitstream/10393/27244/1/MR18417.PDF.

Freedy, David Rawlings. 2003. “Brecker’s Blues: Transcription and Theoretical Analysis of Six Selected Improvised Blues Solos By Jazz Saxophonist Michael Brecker”. PhD Thesis, The Ohio State University. http://search.proquest.com.proxy.library.adelaide.edu.au/docview/305317785/.

Friberg, Anders and Andreas Sundström. 1997. “Preferred Swing Ratio In Jazz As a Function of Tempo”. *TMH-QPSR (Speech, Music, and Hearing: Quarterly Progress and Status Report)* 38 (4): 19–27. Accessed June 24th, 2017. http://www.speech.kth.se/prod/publications/files/2968.pdf.

—— 2002. “Swing Ratios and Ensemble Timing In Jazz Performance: Evidence For a Common Rhythmic Pattern”. *Music Perception* 9 (3): 333–349. Accessed June 24th, 2017. https://www.speech.kth.se/prod/publications/files/822.pdf.

Frieler, Klaus. 2007. “Visualizing Music on the Metrical Circle”. In *Proceedings of the 8th International Symposium on Music Information Retrieval*, 291–292. ISMIR2007, Wein: OCG. Accessed February 26th, 2017. http://ismir2007.ismir.net/proceedings/ISMIR2007_p291_frieler.pdf.

—— 2008. “Metrical Circle Map and Metrical Markov Chains”. In *Systematic and Comparative Musicology: Concepts, Methods, Findings*, edited by Albrecht Schneider. Frankfurt Am Main; Oxford: Peter Lang. Accessed February 26th, 2017. http://www.mu-on.org/frieler/docs/frieler_mcm_2007.pdf.

—— 2014a. “Pattern Usage In Monophonic Jazz Solos”. Paper Presented at the International Jazzomat Research Workshop, University of Music Franz Liszt Weimar, Weimar, Germany, September 26-27. Powerpoint Slides. Accessed February 23rd, 2017. http://jazzomat.hfm-weimar.de/workshop2014/presentations/frieler_jazzomat_workshop_2014_pattern_pres.pdf.

—— 2014b. “Exploring Phrase Form Structures. Part I: European Folk Songs”. In *Proceedings of the Fourth International Workshop on Folk Music Analysis*, edited by Andre Holzapfel, 108–110. Istanbul: Bogazici University. Accessed February 23rd, 2017. http://www.mu-on.org/frieler/docs/frieler_phraseformI_FMA2014.pdf.

—— 2016a. “Jazz Solo Analysis Between Music Information Retrieval, Music Psychology, and Jazz Research: Part II - Data and Tools”. Presentation at the 17th International Society For Music Information Retrieval Conference, New

York City. Accessed March 28th, 2017.
https://github.com/jazzomat/knowhow/blob/master/ISMIR_2016_Tutorial/ISMIR_2016_Jazzomat_Tutorial_Part_2.pdf.

—— 2016b. "Miles Vs Trane. Statistical Comparison of John Coltrane's and Miles Davis' Improvisational Styles". Presentation at Miles Davis and John Coltrane at 90: Retrospect and Prospect. Accessed March 29th, 2017.
http://www.mu-on.org/frieler/docs/frieler_milestrane_surrey2016.pdf.

—— 2016c. "Tracing the Creative Process. A Case Study of Bob Berg's Solo on Angles". Presentation at the Second International Jazzomat Research Workshop, Hochschule Für Musik Franz Liszt Weimar, Weimar, Germany, September 23-24. Accessed February 23rd, 2017.
http://jazzomat.hfm-weimar.de/workshop2016/presentations/frieler_tracing_creative_process_2016_cleared.pdf.

—— 2018. "A Feature History of Jazz Improvisation". In *Jazz @100 Darmstadt Studies in Jazz Research Vol. 15*, edited by Wolfram Knauer, 67–90. Hofheim: Wolke Verlag.

—— 2020. "Miles Vs. Trane: Computational and Statistical Comparison of the Improvisatory Styles of Miles Davis and John Coltrane". *Jazz Perspectives* 12 (1): 123–145. Accessed September 3rd, 2021.
https://doi.org/10.1080/17494060.2020.1734053.

Frieler, Klaus, Martin Pfleiderer, Jakob Abeßer, and Wolf-Georg Zaddach. 2016. "'Telling a Story.' On the Dramaturgy of Monophonic Jazz Solos". *Empirical Musicology Review* 11 (1): 68–82. Accessed February 26th, 2017.
http://emusicology.org/article/view/4959.

Frieler, Klaus, Martin Pfleiderer, Wolf-Georg Zaddach, and Jakob Abeßer. 2013. "Introducing the Jazzomat Project and the MeloSpy Library". In *Proceedings of the Third International Workshop on Folk Music Analysis*, edited by Peter Van Kranenburg, Christina Anagnostopoulou, and Anja Volk, 76–78. Amseterdam: Meertens Institute, Utrecht: Department of Information and Computing Sciences, Utrecht University. Accessed February 23rd, 2017.
http://dspace.library.uu.nl/handle/1874/276246.

Frieler, Klaus, Wolf-Georg Zaddach, and Jakob Abeßer. 2014. "Exploring Phrase Form Structures. Part II: Monophonic Jazz Solos". In *Proceedings of the Fourth International Workshop on Folk Music Analysis*, edited by Andre Holzapfel, 48–51. Istanbul: Bogazici University. Accessed February 23rd, 2017.
http://www.mu-on.org/frieler/docs/frieler_phraseformII_FMA2014.pdf.

Gardiner, Robert Aman. 2008. "Stylistic Evolution of Jazz Saxophone Performance As Illustrated Through Selected Improvisations By Ten Master Improvisers".

PHD Thesis, University of South Carolina. http://search.proquest.com.proxy.library.adelaide.edu.au/docview/304458746/.

Giomi, Francesco and Marco Ligabue. 1991. “Computational Generation and Study of Jazz”. *Journal of New Music Research* 20 (1): 47–64. Accessed March 9th, 2017. http://www.tandfonline.com.proxy.library.adelaide.edu.au/doi/abs/10.1080/09298219108570576.

Giraldo, Sergio and Rafael Ramirez. 2017. “A Machine Learning Approach To Discover Rules for Expressive Performance Actions in Jazz Guitar Music”. *Frontiers in Psychology* 7 (1965): 1–13. Accessed December 12th, 2017. https://doi.org/10.3389/fpsyg.2016.01965.

Green, Grant. 1961i. *Green Street*. Grant Green, Ben Tucker, and Dave Bailey. Recorded April 1, 1961. Blue Note BST 84071, CD liner notes.

Green, Sharony Andrews. 1999. *Grant Green: Rediscovering the Forgotten Genius of Jazz Guitar*. San Francisco: Miller Freeman Books.

—— 2017. *The Grant Green Story*. Vimeo, Video Documentary, 58:28. Accessed July 1st, 2022. https://vimeo.com/229777815.

Griffin, Farah Jasmine and Salim Washington. 2008. *Clawing At The Limits Of Cool: Miles Davis, John Coltrane and the Greatest Jazz Collaboration Ever*. New York: Thomas Dunne Books.

Hajj, Nadine, Maurice Filo, and Mariette Awad. 2017. “Automated Composer Recognition For Multi-Voice Piano Compositions Using Rhythmic Features, N-Grams and Modified Cortical Algorithms”. *Complex and Intelligent Systems*: 1–11. Accessed November 15th, 2017. https://link.springer.com/article/10.1007/s40747-017-0052-x.

Hal Leonard Corporation. 2004a. *The Real Book*. 6th edition. Milwaukee, Wisconsin: Hal Leonard Corporation.

—— 2004b. *The Real Book, Volume II*. 2nd edition. Milwaukee, Wisconsin: Hal Leonard Corporation.

—— 2006. *The Real Book, Volume III*. 2nd edition. Milwaukee, Wisconsin: Hal Leonard Corporation.

Hedges, Thomas, Pierre Roy, and François Pachet. 2014. “Predicting the Composer and Style of Jazz Chord Progressions”. *Journal of New Music Research* 43 (3): 276–290. Accessed November 15th, 2017. https://doi.org/10.1080/09298215.2014.925477.

Hennequin, Romain, Anis Khlif, Felix Voituret, and Manuel Moussallam. 2020. “Spleeter: A Fast and Efficient Music Source Separation Tool With Pre-Trained Models”. *Journal of Open Source Software* 5 (50): 2154. Accessed October 21st, 2020. https://doi.org/10.21105/joss.02154.

Herlands, William, Ricky Der, Yoel Greenberg, and Simon Levin. 2014. "A Machine Learning Approach to Musically Meaningful Homogeneous Style Classification". In *Proceedings of the Twenty-Eight AAAI Conference on Artificial Intelligence*, edited by Carla E Brodley and Peter Stone, 276–282. Québec City: The AAAI Press. Accessed December 4th, 2017. https://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8314/8431.

Hernandez, Adam. 2020. "The Microrhythms Of Joe Pass: Analysis And Application Of Swing Ratios And Tempo". Masters of Music Thesis, University of Colorado Boulder. https://scholar.colorado.edu/concern/graduate_thesis_or_dissertations/v118rf360.

Herremans, Dorien, David Martens, and Kenneth Sörensen. 2016. "Composer Classification Models For Music-Theory Building". In *Computational Music Analysis*, edited by David Meredith, 369–392. Switzerland: Springer International Publishing.

Heyer, David. 2011. "Vocabulary, Voice Leading, and Motivic Coherence In Chet Baker's Jazz Improvisations". PhD Thesis, University of Oregon. http://search.proquest.com.proxy.library.adelaide.edu.au/docview/917243281/.

Hodges, Glen. 2007. "The Analysis of Jazz Improvisational Language and Its Use In Generating New Composition and Improvisation: A Case Study Involving Bebop Jazz Guitarist Jimmy Raney (20/8/27 - 9/5/95)". PhD Thesis, Sydney: Macquarie University. http://www.jonraney.com/wp-content/uploads/2011/12/Glen-Hodges-PhD.pdf.

Honing, Henkjan and W. Bas de Hass. 2008. "Swing Once More: Relating Timing and Tempo in Expert Jazz Drumming". *Music Perception* 25 (5): 471–476. Accessed March 3rd, 2022. https://doi.org/10.1525/mp.2008.25.5.471.

Hunt, Jonathan Robert. 2015. "On the Shoulders of Giants: Bechet, Noone, Goodman and the Efflorescence of Jazz Clarinet and the Improvised Solo". PhD Thesis, University of Adelaide. http://hdl.handle.net/2440/101783.

Huron, David. 1991. "A Score-Based Study of Musical Dynamics In 14 Piano Composers". *Psychology of Music* 19 (1): 33–45. Accessed April 7th, 2017. http://www.musiccog.ohio-state.edu/Huron/Publications/huron.dynamics.abstract.html.

——— 1994. *The Humdrum Toolkit: Reference Manual.* Center For Computer Assisted Research In the Humanities.

——— 1996. "The Melodic Arch In Western Folksongs". *Computing In Musicology* 10: 3–23. Accessed April 7th, 2017. http://csml.som.ohio-state.edu/Huron/Publications/huron.arch.text.html.

Ishwaran, Hemant. 2007. "Variable Importance In Binary Regression Trees and Forests". *Electronic Journal of Statistics* 1: 519–537. Accessed February 12th,

2018.
https://Projecteuclid.Org/Download/Pdfview_1/Euclid.Ejs/1195157166.

Janssen, Berit and Peter Van Kranenburg. 2014. "A Comparison of Similarity Measures For Musical Pattern Matching". Paper Presented at the International Jazzomat Research Workshop, University of Music Franz Liszt Weimar, Weimar, Germany, September 26-27. Powerpoint Slides. Accessed March 29th, 2017. http://jazzomat.hfm-weimar.de/workshop2014/presentations/janssen_jazzomat_workshop_2014_similarity_measures_pres.pdf.

Jazzomat Research Project. 2017. *The Jazzomat Research Project*. Accessed February 20th, 2017. http://jazzomat.hfm-weimar.de/.

Kaiser, Steve. 2013. "The Art of Jazz Guitar". Master of Arts in Music Thesis, Eastern Illinois University. https://thekeep.eiu.edu/theses/1159.

Kaliakatsos-Papakostas, Maximos A., Michael G. Epitropakis, and Michael N. Vrahatis. 2010. "Musical Composer Identification Through Probabilistic and Feedforward Neural Networks". In *Applications of Evolutionary Computation. Evoapplications*, edited by Cecilia Di Chio, Anthony Brabazon, Gianni A. Di Caro, Marc Ebner, Muddassar Farooq, Andreas Fink, Jörn Grahl, Gary Greenfield, Penousal Machado, Michael O'Neill, Ernesto Tarantino, and Neil Urquhart, 411–420. Berlin Heidelberg: Springer. Accessed November 15th, 2017. https://doi.org/10.1007/978-3-642-12242-2_42.

Kantarelis, Spyridon, Edmund Dervakos, Natalia Kotsani, and Giorgos Satmou. 2021. "Musical Harmony Analysis with Description Logics". In *Proceedings of the 34th International Workshop on Description Logics (DL 2021)*, edited by Martin Homola, Vladislav Ryzhikov, and Renate Schmidt, 1–14. Bratislava, Slovakia: Description Logics. Accessed September 30th, 2021. http://ceur-ws.org/Vol-2954/paper-20.pdf.

Kenny, Barry. 1999a. "Structure In Jazz Improvisation: A Formulaic Analysis of the Improvisations of Bill Evans". In *Annual Review of Jazz Studies 10 1999*, edited by Edward Berger, David Cayer, Henry Martin, and Dan Morgenstern, 163–194. Newark: Scarecrow Press, Inc.

——— 1999b. "Jazz Analysis As Cultural Imperative (And Other Urban Myths): A Critical Overview of Jazz Analysis and Its Relationship to Pedagogy". *Research Studies In Music Education* 13 (1): 56–80. Accessed May 25th, 2017. http://journals.sagepub.com.proxy.library.adelaide.edu.au/doi/pdf/10.1177/1321103X9901300106.

Korman, Clifford. 1999. "'Criss Cross': Motivic Construction In Composition and Improvisation". In *Annual Review of Jazz Studies 10 1999*, edited by

Edward Berger, David Cayer, Henry Martin, and Dan Morgenstern, 103–126. Newark: Scarecrow Press, Inc.

Kuhn, Max. 2008. “Building Predictive Models in R Using the caret Package”. *Journal of Statistical Software* 28 (5): 1–26. Accessed September 12th, 2019. http://doi.org/10.18637/jss.v028.i05.

—— 2020. “Variable Importance Measures For C5.0 Models”. *C5.0 Decision Trees and Rule-Based Models*. Accessed March 3rd, 2021. https://topepo.github.io/C5.0/reference/C5imp.html.

Larsen, David William. 2021. “Creating The Cool: Understanding The Music Of Gerry Mulligan”. PhD Dissertation, Washington State University. https://www.proquest.com/docview/2572599336.

Lartillot, Olivier. 2014. “Computer-Automated Motivic Analysis of the Weimar Jazz Database Through Exhaustive Pattern Mining”. Paper Presented at the International Jazzomat Research Workshop, University of Music Franz Liszt Weimar, Weimar, Germany, September 26-27. Powerpoint Slides. Accessed March 2nd, 2017. http://jazzomat.hfm-weimar.de/workshop2014/presentations/lartillot_jazzomat_workshop_2014_patternmining_pres.pdf.

Lee, Kelli Janine. 2008. “Machine Learning Classification Analysis of Music Statistics”. Masters Thesis, University of Arkansas. https://search.proquest.com/openview/c7c7220e0a02944e573f69ffddcdcddc/1?pq-origsite=gscholar&cbl=18750&diss=y.

Levine, Mark. 1995. *The Jazz Theory Book*. Petaluma, CA: Sher Music Company.

Liaw, Andy and Matthew Wiener. 2017. “CRAN, Randomforest Manual: Importance.Rd”. *Github*. Accessed March 3rd, 2021. https://github.com/cran/randomForest/blob/bd3509cdc08aa7cd23495f836bbccbc5e4c4ea1b/man/importance.Rd#L28-L42.

Lindsay, Gary. 2005. *Jazz Arranging Techniques from Quartet to Big Band*. Miami: Staff Art Publishing.

Marsden, Alan. 2016. “Music Analysis By Computer: Ontology and Epistemology”. In *Computational Music Analysis*, edited by David Meredith, 3–30. Switzerland: Springer International Publishing.

Marshall, Wolf. 2004. *Best of Grant Green: A Step-By-Step Breakdown of the Guitar Styles and Techniques of the Jazz Groove Master (Guitar Signature Licks)*. Milwaukee: Hal Leonard.

Martin, Henry. 1996. *Charlie Parker and Thematic Improvisation*. Maryland: Scarecrow Press.

Martorell, Agustin and Emilia Gomez. 2016. “Contextual Set-Class Analysis”. In *Computational Music Analysis*, edited by David Meredith, 81–110. Switzerland: Springer International Publishing.

McCann, Michelle Roehm. 2017. *More Girls Who Rocked the World: Heroines from Ada Lovelace to Misty Copeland.* New York: Aladdin.

McEvoy, David. 2014. "Aspects of Herbie Hancock's Pre-Electric Improvisational Language and Their Application In Contemporary Jazz Performance: A Portfolio of Recorded Performances and Exegesis". Masters Thesis, University of Adelaide. http://hdl.handle.net/2440/85974.

Meredith, David, ed. 2016. *Computational Music Analysis.* Switzerland: Springer International Publishing.

Moore, David Michael. 2022. "Blue: A Stylistic Analysis of Blue Mitchell". Doctor of Arts Dissertation, University of Northern Colorado. https://www.proquest.com/docview/2665564590/.

Moore, Michael D. 1999. "A Performance Analysis of Selected Works For Trumpet By Chet Baker". PhD Thesis, University of Oklahoma. https://shareok.org/handle/11244/5856.

Müllensiefen, Daniel and Geraint Wiggins. 2011. "Polynomial Functions as a Representation of Melodic Phrase Contour". In *Systematic Musicology: Empirical and Theoretical Studies*, edited by Albrecht Schneider and Arne von Ruschkowski, 63–88. Berlin, Germany: Peter Lang Verlag. https://doi.org/10.3726/978-3-653-01290-3.

Nasridinov, Aziz and Young-Ho Park. 2014. "A Study on Music Genre Recognition and Classification Techniques". *International Journal of Multimedia and Ubiquitous Engineering* 9 (4): 31–42. Accessed November 15th, 2017. http://dx.doi.org/10.14257/ijmue.2014.9.4.04.

Nattiez, Jean-Jacques. 1990. *Music and Discourse: Towards a Semiology of Music.* Trans. by Carolyn Abbate. Princeton: Princeton University Press.

Ng, Vivian W and Leo Breiman. 2005. "Bivariate Variable Selection For Classification Problem". In *Technical Report, Department of Statistics*, 1–22. University of California-Berkeley. Accessed February 12th, 2018. http://digitalassets.lib.berkeley.edu/sdtr/ucb/text/692.pdf.

Norgaard, Martin and Ute Römer. 2022. "Patterns In Music: How Linguistic Corpus Analysis Tools Can Be Used to Illuminate Central Aspects of Jazz Improvisation". *Jazz Education in Research and Practice* 3 (1 (Spring 2022)): 3–26. Accessed January 8th, 2022. https://www.jstor.org/stable/10.2979/jazzeducrese.3.1.02.

Ostercamp, Neil Lewis. 2018. "The Improvisational Style of Steve Lacy: Analyses of Selected Transcriptions (1957-1962)". Doctor of Musical Arts Dissertation, University of Northern Carolina at Greensboro. https://www.proquest.com/docview/2063407217/.

Owens, Thomas. 1974. “Charlie Parker: Techniques of Improvisation. Volume I”. PhD Thesis, University of California. https://archive.org/details/CharlieParkerDissertationVolumeIThomasOwens1974.

Pfleiderer, Martin. 2014. “Melospyin’ the Trane. Exploring Improvisations of John Coltrane With Melospysuite”. Paper Presented at the International Jazzomat Research Workshop, University of Music Franz Liszt Weimar, Weimar, Germany, September 26-27. Powerpoint Slides. Accessed February 23rd, 2017. http://jazzomat.hfm-weimar.de/workshop2014/presentations/pfleiderer_jazzomat_workshop_2014_melospyin_trane_pres.pdf.

—— 2016. “Comparative Studies of Jazz Improvisations”. Presentation at the Second International Jazzomat Research Workshop, Hochschule Für Musik Franz Liszt Weimar, Weimar, Germany, September 23-24. Accessed March 29th, 2017. http://jazzomat.hfm-weimar.de/workshop2016/presentations/Comparative_Jazz_Studies_Pfleiderer.pdf.

Pfleiderer, Martin and Klaus Frieler. 2010. “The Jazzomat Project. Issues and Methods For the Automatic Analysis of Jazz Improvisations”. In *Concepts, Experiments, and Fieldwork: Studies in Systematic Ethnomusicology*, edited by Rolf Bader, Christiane Neuhaus, and Ulrich Morgenstern, 279–295. Peter Lang. Accessed February 23rd, 2017. http://www.mu-on.org/frieler/docs/frieler_pfleiderer_azzomat.pdf.

Pfleiderer, Martin, Klaus Frieler, Jakob Abeßer, Wolf-Georg Zaddach, and Benjamin Burkhart, eds. 2017. *Inside the Jazzomat. New Perspectives For Jazz Research*. Mainz: Schott Campus.

Pfleiderer, Martin, Wolf-Georg Zaddach, and Klaus Frieler. 2016. “Pitch Class Hierarchies In Miles Davis’ ‘So What’: Reconsidering Modal Jazz Improvisation With Computer-Based Analysis Tools”. In *Beitragsarchiv Zur Jahrestagung Der Gesellschaft Für Musikforschung Halle/Saale 2015 – ‘Musikwissenschaft: Die Teildisziplinen Im Dialog’*, edited by Wolfgang Auhagen and Wolfgang Hirschmann. Mainz: Schott Campus. Accessed February 26th, 2017. http://schott-campus.com/wp-content/uploads/2016/09/pfleiderer_zaddach_frieler-so-what.pdf.

Porter, Lewis. 1985. *Lester Young*. Boston, Massachusetts: Twayne Publishers.

Potter, Gary. 1989. “The Unique Role of ♭VII7 in Bebop Harmony”. *Jazzforschung* 21: 35–48. Accessed August 12th, 2022.

—— 1992. “Analysing Improvised Jazz”. *The College Music Society* 32: 143–160. Accessed February 22nd, 2017. http://www.jstor.org/stable/40374206.

Prince, Jon B. and Mark A. Schmuckler. 2014. “The Tonal-Metric Hierarchy: A Corpus Analysis”. *Music Perception: An Interdisciplinary Journal* 31 (3):

254–270. Accessed June 21st, 2022. https://www.jstor.org/stable/10.1525/mp.2014.31.3.254.

Ramirez, Rafael, Esteban Maestre, and Xavier Serra. 2010. "Automatic Performer Identification In Commercial Monophonic Jazz Performances". *Pattern Recognition Letters* 31: 1514–1523. Accessed November 15th, 2017. https://www.sciencedirect.com/science/article/pii/S0167865509003742.

Reyman, Christopher Eric. 2011. "An Analysis of Selected Improvisations By Andrew Hill and the Development of Progressive Jazz Piano, 1959-2005". PhD Thesis, University of Illinois. https://www.ideals.illinois.edu/bitstream/handle/2142/26152/Reyman_Christopher.pdf?sequence=1.

Rolland, Pierre-Yves and Jean-Gabriel Ganascia. 2002. "Pattern Detection and Discovery: The Case of Music Data Mining". In *Pattern Detection and Discovery: ESF Exploratory Workshop London, UK, September 16–19, 2002 Proceedings*, edited by David J Hand, Niall M. Adams, and Richard J. Bolton, 190–198. Springer Berlin Heidelberg. Accessed March 5th, 2017. http://link.springer.com.proxy.library.adelaide.edu.au/chapter/10.1007/3-540-45728-3_15.

Salamon, Justin and Emilia Gómez. 2012. "Melody Extraction from Polyphonic Music Signals using Pitch Contour Characteristics". *IEEE Trans Audio Speech Lang Process* 20 (6): 1759–1770. Accessed February 26th, 2017. http://doi.org/10.1109/TASL.2012.2188515.

Salmon, Shawn. 2011. "Imitation, Assimilation, and Innovation: Charlie Christian's Influence on Wes Montgomery's Improvisational Style In His Early Recordings (1957-1960)". PhD Thesis, Indiana: Ball State University. http://search.proquest.com.proxy.library.adelaide.edu.au/pagepdf/917929962/.

Satterthwaite, Gregory. 2020. "Beyond Fourths and Pentatonics: A Critical Analysis of Selected Recordings of McCoy Tyner form 1962 to 1963". Doctor of Musical Arts Dissertation, University of North Texas. https://www.proquest.com/docview/2489621482/.

Schnabel, Brendan. 2021. "Using Solo Transcription To Develop A Personal Jazz Improvisational Style". Doctor of Musical Arts Dissertation, George Mason University. https://www.proquest.com/docview/2572610368.

Scott, Andrew. 2006. "From Grant Green to B.B. King to T-Bone Walker: Consistent Approaches to the Blues". *Journal on Media Culture* 8 (February). Accessed May 10th, 2017. http://www.icce.rug.nl/~soundscapes/VOLUME08/From_Grant_Green_to_BB_King.shtml.

——— 2009. "Exploring the Use of a Single Formula In Grant Green'S First Chorus Improvisation on 'I'll Remember April' (1961)". *Current Research In Jazz* 1.

Accessed May 10th, 2017.
http://www.crj-online.org/v1/CRJ-GrantGreenSolo.php.

—— 2012. "Green, Grant". *Grove Music Online. Oxford Music Online.* Oxford University Press. Accessed February 24th, 2017.
http://www.oxfordmusiconline.com/subscriber/article/grove/music/A2228339.

Sher, Chuck and Bob Bauer, eds. 1988. *The New Real Book.* Petaluma, California: Sher Music Co.

—— eds. 1991. *The New Real Book: Volume Two.* Petaluma, California: Sher Music Co.

Sher, Chuck, Sky Evergreen, and Larry Dunlap, eds. 1995. *The New Real Book: Volume 3.* Petaluma, California: Sher Music Co.

Smith, Gregory Eugene. 1983. "Homer, Gregory, and Bill Evans? The Theory of Formulaic Composition In the Context of Jazz Piano Improvisation". PhD Thesis, Harvard University.
http://search.proquest.com.proxy.library.adelaide.edu.au/docview/303159862/.

Solstad, Stein Helge. 2015. "Strategies In Jazz Guitar Improvisation". PhD Thesis, Norwegian Academy of Music, Oslo. http://hdl.handle.net/11250/2358576.

Stewart, Milton Lee. 1973. "Structural Development In the Jazz Improvisational Technique of Clifford Brown". PhD Thesis, University of Michigan.
http://search.proquest.com/docview/302704602/.

Stillman, Jason Patrick. 2017. "Individualism and Audibility in Jazz: The Music of Bunky Green". Doctor of Musical Arts Thesis, University of Toronto.
https://www.proquest.com/docview/1992757956/.

Sullivan, Gail M. and Richard Feinn. 2012. "Using Effect Size – or Why the *P* Value Is Note Enough". *Journal of Graduate Medical Education* 4 (3): 279–282. Accessed December 14th, 2021. https://doi.org/10.4300/JGME-D-12-00156.1.

Technimo. 2017. "Main Playlists". *iReal Pro*. Accessed February 20th, 2017.
https://www.irealpro.com/main-playlists/.

Temperley, David and Trevor De Clercq. 2013. "Statistical Analysis of Harmony and Melody In Rock Music". *Journal of New Music Research* 42 (3): 187–204. Accessed May 2nd, 2017. http://www.tandfonline.com.proxy.library.adelaide.edu.au/doi/abs/10.1080/09298215.2013.788039.

University of Cambridge MRC Cognition and Brain Sciences Unit. 2009. "Rules of thumb on magnitudes of effect sizes". *University of Cambridge MRC CBU Wiki.* Accessed November 25th, 2017.
https://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize.

Van der Bliek, Robert. 1987. "Wes Montgomery a Study of Coherence In Jazz Improvisation". Masters Thesis, Ottawa, Canada: York University.
http://search.proquest.com.proxy.library.adelaide.edu.au/docview/303651130/.

Waggoner, Andrew. 2002. “Grant Green”. In *The New Grove Dictionary of Jazz*, edited by Barry Kernfield. 2nd edition, 88–89. London: Macmillan Publishers Limited.

Weir, Timothy Malcolm. 2006. “An Analysis of the Development of Kenny Dorham’s Jazz Improvisational Vocabulary”. DMA, University of Miami. http://search.proquest.com/docview/305312016/.

Wesolowski, Brian C. 2012. “Testing A Model Of Jazz Rhythm: Validating A Microstructural Swing Paradigm”. PhD Thesis, University of Miami. https://scholarship.miami.edu/esploro/outputs/doctoral/Testing-a-Model-of-Jazz-Rhythm/991031447346402976?institution=01UOML_INST.

Widmer, Gerhard. 2000. “On the Potential of Machine Learning For Music Research”. In *Readings In Music and Artificial Intelligence*, edited by Eduardo Reck Miranda, 69–84. New York: Routledge. Accessed December 4th, 2017. http://www.ofai.at/cgi-bin/tr-online?number+97-28.

Wild, Justin Michael. 2002. “Grant Green’s Approach to Improvisation: an Analytical Study”. Masters of Music Thesis, University of South Carolina.

Wilding, Mark. 2008. “Automatic Harmonic Analysis of Jazz Chord Progressions Using a Musical Categorical Grammar”. Master of Science, University of Edinburgh. https://mark.granroth-wilding.co.uk/files/msc_dissertation.pdf.

Young, Robert W. 1939. “Terminology for Logarithmic Frequency Units”. *The Journal of the Acoustical Society of America* 11: 134–139. Accessed August 14th, 2019. https://doi.org/10.1121/1.1916017.

Zaddach, Wolf-Georg. 2016. “Saxophone Colossus Rhythmical and Harmonic Aspects of Michael’s and Chris Potter’s Improvisations”. Presentation at the Second International Jazzomat Research Workshop, Hochschule Für Musik Franz Liszt Weimar, Weimar, Germany, September 23-24. Accessed March 29th, 2017. http://jazzomat.hfm-weimar.de/workshop2016/presentations/Zaddach_JazzomatWorkshop2016_Potter.pdf.

Zaddach, Wolf-Georg and Martin Pfleiderer. 2016. “Playin’ Outside. New Melodic and Harmonic Strategies In Jazz Improvisation After John Coltrane”. Presentation at Miles Davis, John Coltrane at 90: Retrospect, and Prospect. Accessed March 29th, 2017. http://jazzomat.hfm-weimar.de/workshop2016/presentations/Zaddach_PresentationSurrey2016.pdf.

Zanon, Patrick and Gerhard Widmer. 2003. “Recognition of Famous Pianists Using Machine Learning Algorithms: First Experimental Results”. In *Proceedings of the 14th Colloquium on Musical Informatics*, 1–6. Accessed December 4th, 2017. https://pdfs.semanticscholar.org/5381/95eac5a14e5ac8790d7c3a981213a7a13539.pdf.

Zimmer, Russell. 2016. “A Study of the Improvisational Style of Theodore ‘Fats’ Navarro: 1949-1950”. PhD Thesis, University of Nebraska. http://digitalcommons.unl.edu/musicstudent/98.

Further Reading

Aitchison, John. 1986. *The Statistical Analysis of Compositional Data*. London: Chapman and Hall.

Bilder, Christopher R and Thomas M Loughin. 2015. *Analysis of Categorical Data With R*. Boca Raton: CRC Press.

Blackwell, David. 2016. “The Jazzomat Research Project: An Evaluation Through the Analysis of Selected Improvisations By Joe Pass”. Honours Project, University of Adelaide.

Chollet, François and J.J. Allaire. 2018. *Deep Learning With R*. Shelter Island: Manning Publications Co.

Cook, Nicholas. 1987. “Traditional Methods of Analysis”. In *A Guide to Musical Analysis*, 7–26. New York: George Braziller.

Cope, David. 1989. *New Directions In Music*. 5th edition. Dubuque, Iowa: W.C. Brown.

Cuthbert, Michael and Kyle Polich. 2016. *Music21*. Data Skeptic Podcast, Podcast, 34 minutes. Accessed March 22nd, 2017. https://dataskeptic.com/blog/episodes/2016/music21.

Donato, Stephanie M. 2014. “Statistical Analysis and Graphic Representation of the Correlation of Bach and Chopin Preludes”. Honours Thesis, University of Vermont. http://scholarworks.uvm.edu/cgi/viewcontent.cgi?article=1025&context=hcoltheses.

Doyle, Frank Vincent. 2018. “Horace Silver: Transcription and Analysis of Ten Blues Compositions Including The Piano Improvisations 1964-1969”. Doctorate of Musical Arts Dissertation, Five Towns College. https://www.proquest.com/docview/2280683120/.

Frieler, Klaus, Martin Pfleiderer, Wolf-Georg Zaddach, and Jakob Abeßer. 2016. “Midlevel Analysis of Monophonic Jazz Solos: A New Approach to the Study of Improvisation”. *Musicae Scientiae*: 1–20. Accessed February 23rd, 2017. http://msx.sagepub.com/content/early/2016/03/08/1029864916636440.

Gabric, Lukas. 2020. “Coltrane Plays The Blues: Multi-Level Coherence and Stylistic Tendencies”. PhD Dissertation, The City University of New York. https://www.proquest.com/docview/2451146638/.

Ghatak, Abhijit. 2017. *Machine Learning With R*. Singapore: Springer Nature.

Gillick, Jon, Kevin Tang, and Robert M. Keller. 2009. “Learning Jazz Grammars”. In *Proceedings of the 6th Sound and Music Computing Conference*, edited by Fabien Gouyon, Álvaro Barbosa, and Xavier Serra, 125–130. Porto, Portugal: Sound and Music Computing Conference. Accessed November 18th, 2017. http://smc2009.smcnetwork.org/programme/pdfs/214.pdf.

Granville, Kyle M. 2020. “A Musical Analysis of the Improvisational Bebop Style of Nat Adderley (1955-1964)”. Doctor of Musical Arts Dissertation, The University of Nebraska - Lincoln. https://www.proquest.com/docview/2476109600.

Heo, Gayeon. 2021. “An Analytical Approach To Selecting Scat Syllables: An Application To The Solos Of Two Contemporary Singer-Horn Players”. Master of Music Thesis, William Paterson UNiversity. https://www.proquest.com/docview/2638067389/.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. *An Introduction to Statistical Learning: With Applications In R*. New York: Springer.

Kelly, Geordie F O. 1997. “The Idiom of Joe Pass: Elements In the Construction of Selected Jazz Guitar Solo”. In *International Association of Jazz Educators: Jazz Research Proceedings Yearbook: a Compilation of Research Papers Presented at the Annual International Conference. Chicago, Illinois, January 1997*, edited by Larry Fisher, 76–88. Manhattan, Kansas: IAJE Publications.

Kuhn, Max. 2019. “Variable Importance”. *The Caret Package*. Accessed March 3rd, 2021. https://topepo.github.io/caret/variable-importance.html#model-independent-metrics.

Lakens, Daniel. 2017. “Equivalence Tests: A Practical Primer For T Tests, Correlations, and Meta-Analyses”. *Social Psychological and Personality Science* 8 (4): 355–362. Accessed June 24th, 2017. http://journals.sagepub.com/doi/pdf/10.1177/1948550617697177.

Lee, Kyungho, Michael J. Junokas, Mohammad Amanzadeh, and Guy E. Garnett. 2015. “Exploratory Analysis on Expressions In Two Different 4/4 Beat Patterns”. In *Proceedings of the 2015 International Computer Music Conference*, 198–201. International Computer Music Association. Accessed July 21st, 2017. http://hdl.handle.net/2027/spo.bbp2372.2015.040.

Leevy, Joffrey L., Taghi M. Khoshgoftaar, Richard A. Bauder, and Naeem Seliya. 2018. “A Survey on Adressing High-Class Imbalance In Big Data”. *Journal of Big Data* 5 (42): 1–30. Accessed May 18th, 2021. https://doi.org/10.1186/s40537-018-0151-6.

Liu, Lu, Jianrong Wei, Huishu Zhang, Jianhong Xin, and Jiping Huang. 2013. “A Statistical Physics View of Pitch Fluctuations In the Classical Music from

Bach to Chopin: Evidence For Scaling”. *PLOS ONE* 8 (3): 1–6. Accessed July 21st, 2017. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0058710.

Marcus, Joshua C. 2022. “A Historical Re-Contextualization of the Soul Jazz of Julian ‘Cannonball’ Adderley”. Master of Music Thesis, William Paterson University. https://www.proquest.com/docview/2637985827/.

McKay, John Z. 2014. “Oxford Handbooks Online”. *Statistical Reasoning In Music Online*. Oxford University Press. Accessed July 21st, 2017. http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199935321.001.0001/oxfordhb-9780199935321-e-50.

Murari, Maddalena, Antonio Roda, Osvaldo Da Pos, Sergio Canazza, Giovanni De Poli, and Marta Sandri. 2014. “How Blue Is Mozart? Non Verbal Sensory Scales For Describing Music Qualities”. In *Proceedings of the 2014 International Computer Music Conference*, 209–216. International Computer Music Association. Accessed July 21st, 2017. http://hdl.handle.net/2027/spo.bbp2372.2014.033.

Pilzer, Leigh. 2020. “Gerry Mulligan, Soloist: Transcription and Analysis of His Small-Group Solos On ‘Line For Lyons,’ ‘Curtains,’ and ‘The Flying Scotsman’”. Doctor of Musical Arts Dissertation, George Mason University. https://www.proquest.com/docview/2427302568/.

Pressing, Jeff. 1987. “The Micro- and Macrostructural Design of Improvised Music”. *Music Perception* 5 (2): 133–172. Accessed June 19th, 2022. https://www.jstor.org/stable/40285390.

Sample, T. W. 2019. “An Analysis of the Improvisational Style of Jazz Organist Jimmy Smith through Selected Right Hand Transcriptions”. Doctorate of Musical Arts Dissertation, Five Towns College. https://www.proquest.com/docview/2384890416/.

Steinberg, Dan. 2009. “CART: Classification and Regression Trees”. In *The Top Ten Algorithms in Data Mining*, edited by Xindong Wu and Vipin Kumar. New York: Chapman and Hall/CRC. Accessed April 29th, 2022. https://doi.org/10.1201/9781420089653.

Tan, Willie. 2018. *Research Methods: A Practical Guide For Students and Researchers*. Singapore: World Scientific.

Waggoner, Andrew and Barry Kernfeld. 2003. “Grant Green”. *The New Grove Dictionary of Jazz, 2nd Ed. Grove Music Online. Oxford Music Online*. Oxford University Press. Accessed February 24th, 2017. http://www.oxfordmusiconline.com/subscriber/article/grove/music/J177000.

Weihs, Claus, Dietmar Jannach, Igor Vatolkin, and Guenter Rudolph. 2016. *Music Data Analysis: Foundations and Applications*. CRC Press.

Discography

Green, Grant. 1960a. *First Session*. Grant Green, Wynton Kelly, Paul Chambers, Philly Joe Jones, Sonny Clark, Butch Warren, and Billy Higgins. Recorded November 16th, 1960 and October 27th, 1961. Blue Note 7243 5 27548 2, CD.

—— 1961a. *Ballads*. Grant Green, Sam Jones, Louis Hayes, Sonny Clark, Ike Quebec, Ben Tucker, Ben Dixon, Kenny Drew, Al Harewood, Yusef Lateef, Jack McDuff, and Art Blakey. Recorded April 1, 1961. Blue Note 7243 5 37560 2 4, CD.

—— 1961e. *Gooden's Corner*. Grant Green, Sonny Clark, Same Jones, and Louis Hayes. Recorded December 23rd, 1961. Blue Note GXF 3058, CD.

—— 1961f. *Grant's First Stand*. Grant Green, Roosevelt Willette, and Ben Dixon. Recorded January 28th, 1961. Blue Note BST 84064, CD.

—— 1961g. *Grantstand*. Grant Green, Yusef Lateef, Jack Mcduff, and Al Harewood. Recorded August 1st, 1961. Blue Note BST 84086, CD.

—— 1961h. *Green Street*. Grant Green, Ben Tucker, and Dave Bailey. Recorded April 1, 1961. Blue Note BST 84071, CD.

—— 1961p. *Reaching Out*. Grant Green, Dave Bailey, Frank Haynes, Billy Gardner, and Ben Tucker. Recorded March 15th, 1961. Black Lion BLCD 760129, CD.

—— 1961q. *Remembering*. Grant Green, Wilbur Ware, and Al Harewood. Recorded August 29th, 1961. Blue Note GXF 3071, CD.

—— 1961t. *Sunday Morning'*. Grant Green, Kenny Drew, Ben Tucker, and Ben Dixon. Recorded June 4th, 1961. Blue Note BST 804099, CD.

—— 1962a. *Born to Be Blue*. Grant Green, Ike Quebec, Sonny Clark, Sam Jones, and Louis Hayes. Recorded March 1st, 1962. Blue Note BST 84432, CD.

—— 1962c. *Feelin' the Spirit*. Grant Green, Herbie Hancock, Butch Warren, Billy Higgins, and Garvin Masseaux. Recorded December 21st, 1962. Blue Note BST 84132, CD.

—— 1962e. *Goin' West*. Grant Green, Herbie Hancock, Reggie Workman, and Billy Higgins. Recorded November 30th, 1962. Blue Note BST 84310, CD.

—— 1962g. *Nigeria*. Grant Green, Sonny Clark, Sam Jones, and Art Blakey. Recorded January 13th, 1962. Blue Note LT 1032, CD.

—— 1962h. *Oleo*. Grant Green, Sonny Clark, Sam Jones, and Louis Hayes. Recorded January 31st, 1962. Blue Note GXF 3065, CD.

—— 1962k. *The Latin Bit*. Grant Green, John Acea, Wendell Marshall, Willie Bobo, Carlos Valdes, and Garvin Masseaux. Recorded April 26th and September 7th, 1962. Blue Note BST 84111, CD.

—— 1963a. *Am I Blue*. Grant Green, Johnny Coles, Joe Henderson, John Patton, and Ben Dixon. Recorded May 16th, 1963. Blue Note BST 84139, CD.

—— 1963b. *Blues For Lou*. Grant Green, John Patton, and Ben Dixon. Recorded February 20th and June 7th, 1963. Blue Note 21438, CD.

—— 1963d. *Idle Moments*. Grant Green, Joe Henderson, Duke Pearson, Boddy Hutcherson, Bob Cranshaw, and Al Harewood. Recorded November 4th and 15th, 1963. Blue Note BST 84154, CD.

—— 1964c. *Matador*. Grant Green, Mccoy Tyner, Bob Cranshaw, and Elvin Jones. Recorded May 20th, 1964. Blue Note GXF 3053, CD.

—— 1964e. *Solid*. Grant Green, James Spaulding, Joe Henderson, Mccoy Tyner, Bob Cranshaw, and Elivn Jones. Recorded June 12th, 1964. Blue Note LT 990, CD.

—— 1964f. *Street of Dreams*. Grant Green, Bobby Hutcherson, Larry Young, and Elvin Jones. Recorded November 16th, 1964. Blue Note BST 84253, CD.

—— 1964g. *Talkin' About!* Grant Green, Larry Young, and Elvin Jones. Recorded September 11th, 1964. Blue Note BST 84183, CD.

—— 1965b. *His Majesty King Funk*. Grant Green, Harold Vick, Larry Young, Ben Dixon, and Candido Camero. Recorded May 26th, 1965. Verve V6-8627, CD.

—— 1965c. *I Want to Hold Your Hand*. Grant Green, Hank Mobley, Larry Young, and Elvin Jones. Recorded March 31st, 1965. Blue Note BLP 4202, CD.

Additional Online Sources

Green, Grant. 1960b. *Seepin*. Grant Green, Wynton Kelly, Paul Chambers, Philly Joe Jones, Sonny Clark, Butch Warren, and Billy Higgins. YouTube Video, 11:46. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=mOn4RqyA8Wo.

—— 1960c. *Sonnymon For Two*. Grant Green, Wynton Kelly, Paul Chambers, Philly Joe Jones, Sonny Clark, Butch Warren, and Billy Higgins. YouTube Video, 6:16. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=K6ZUl2K0UzA.

—— 1961b. *Blues in Maude's Flat*. Grant Green, Yusef Lateef, Jack Mcduff, and Al Harewood. YouTube Video, 15:04. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=-n3VJQIKduU.

—— 1961c. *Born To Be Blue*. Grant Green, Ike Quebec, Sonny Clark, Sam Jones, and Louis Hayes. YouTube Video, 4:54. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=bVX_FTC2yEA.

—— 1961d. *Freedom March*. Grant Green, Kenny Drew, Ben Tucker, and Ben Dixon. YouTube Video, 8:42. Accessed November 9th, 2021. https://www.youtube.com/watch?v=amZlEovr9QE.

Green, Grant. 1961j. *Green With Envy*. Grant Green, Ben Tucker, and Dave Bailey. YouTube Video, 9:46. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=J2d2OFu9SEo.

—— 1961k. *I'll Remember April*. Grant Green, Wilbur Ware, and Al Harewood. YouTube Video, 9:09. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=xdyWcfdZTXc.

—— 1961l. *Little Girl Blue*. Grant Green, Sonny Clark, Sam Jones, and Louis Hayes. YouTube Video, 7:15. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=qCd_O39rUJc.

—— 1961m. *Miss Ann's Tempo*. Grant Green, Roosevelt Willette, and Ben Dixon. YouTube Video, 5:40. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=wcs9PdZEJwM.

—— 1961n. *Moon River*. Grant Green, Sonny Clark, Same Jones, and Louis Hayes. YouTube Video, 5:38. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=iGHCjq6G6h4.

—— 1961o. *Our Miss Brooks*. Grant Green, Dave Bailey, Frank Haynes, Billy Gardner, and Ben Tucker. YouTube Video, 6:48. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=YaV4hcHt2Gs.

—— 1961r. *Round About Midnight*. Grant Green, Ben Tucker, and Dave Bailey. YouTube Video, 7:03. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=m7vhW6Pz4mg.

—— 1961s. *Sunday Mornin'*. Grant Green, Kenny Drew, Ben Tucker, and Ben Dixon. YouTube Video, 4:01. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=MhFX3LpueLQ.

—— 1962b. *Brazil*. Grant Green, John Acea, Wendell Marshall, Willie Bobo, Carlos Valdes, and Garvin Masseaux. YouTube Video, 5:01. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=vzp8QqfUqjw.

—— 1962d. *Go Down Moses*. Grant Green, Herbie Hancock, Butch Warren, Billy Higgins, and Garvin Masseaux. YouTube Video, 7:30. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=jRyz7aWcxUc.

—— 1962f. *Nancy (With The Laughing Face)*. Grant Green, Sonny Clark, Sam Jones, and Art Blakey. YouTube Video, 6:25. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=AqXCd76WvyQ.

—— 1962i. *Oleo*. Grant Green, Sonny Clark, Sam Jones, and Louis Hayes. YouTube Video, 5:38. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=KWAmZZLJlSg.

—— 1962j. *Red River Valley*. Grant Green, Herbie Hancock, Reggie Workman, and Billy Higgins. YouTube Video, 6:04. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=6EVnc7D37wI.

—— 1962l. *The Song Is You*. Grant Green, Sonny Clark, Sam Jones, and Art Blakey. YouTube Video, 7:47. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=_TIRxwCisVo.

—— 1962m. *Tico-Tico*. Grant Green, John Acea, Wendell Marshall, Willie Bobo, Carlos Valdes, and Garvin Masseaux. YouTube Video, 7:46. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=-pcISrbblgw.

—— 1963c. *Have You Ever Had The Blues*. Grant Green, John Patton, and Ben Dixon. YouTube Video, 4:29. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=wFVzjDKAvCo.

—— 1963e. *Idle Moments*. Grant Green, Joe Henderson, Duke Pearson, Boddy Hutcherson, Bob Cranshaw, and Al Harewood. YouTube Video, 14:56. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=aq0m0hbCjFQ.

—— 1963f. *Take These Chains From My Heart*. Grant Green, Johnny Coles, Joe Henderson, John Patton, and Ben Dixon. YouTube Video, 6:11. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=rXpUjoIQu38.

—— 1963g. *The Surrey With The Fringe On Top*. Grant Green, John Patton, and Ben Dixon. YouTube Video, 5:05. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=G_qCcg2tWms.

—— 1964a. *I Wish You Love*. Grant Green, Bobby Hutcherson, Larry Young, and Elvin Jones. YouTube Video, 8:46. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=VIa9jzjmoxA.

—— 1964b. *I'm An Old Cowhand*. Grant Green, Larry Young, and Elvin Jones. YouTube Video, 6:31. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=4SyOopRWjFM.

—— 1964d. *Minor League*. Grant Green, James Spaulding, Joe Henderson, Mccoy Tyner, Bob Cranshaw, and Elivn Jones. YouTube Video, 7:05. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=znCG8eqjnXo.

—— 1964h. *Wives and Lovers*. Grant Green, Mccoy Tyner, Bob Cranshaw, and Elvin Jones. YouTube Video, 9:01. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=wHfCVbBRvRo.

—— 1965a. *At Long Last Love*. Grant Green, Hank Mobley, Larry Young, and Elvin Jones. Recorded March 31st, 1965. YouTube Video, 7:17. Accessed November 9th, 2021. https://www.youtube.com/watch?v=LIkYQt-bzHo.

—— 1965d. *Stella By Starlight*. Grant Green, Hank Mobley, Larry Young, and Elvin Jones. YouTube Video, 6:29. Accessed Novermber 9th, 2021. https://www.youtube.com/watch?v=_Xo7K39Gk_k.

Software

Abeßer, Jakob, Klaus Frieler, Martin Pfleiderer, and Wolf-Georg Zaddach. 2017a. *MelospyGUI*. Jazzomat Research Project: Hochschule für Musik Franz Liszt Weimar. Computer software, current version 1.4. Accessed December 7th, 2017. http://jazzomat.hfm-weimar.de/download/download.html.

——— 2017b. *MeloSpySuite*. Jazzomat Research Project: Hochschule für Musik Franz Liszt Weimar. Computer software, current version 1.6. Accessed December 7th, 2017. http://jazzomat.hfm-weimar.de/download/download.html.

Audacity Team. 2015. *Audacity(R): Free Audio Editor and Recorder*. Computer software, version 2.1.2. Accessed February 20th, 2017. https://audacityteam.org/.

Biolcati, Massimo. 2017. *iReal Pro*. Technimo. Android software, version 8.0. Accessed February 20th, 2017. https://www.irealpro.com/.

Cannam, Chris, Christian Landone, and Mark Sandler. 2010. "Sonic Visualiser: An Open Source Application For Viewing, Analysing, and Annotating Music Audio Files". In *Proceedings of the ACM Multimedia 2010 International Conference, Firenze, Italy*, 1467–1468. Accessed February 20th, 2017. http://www.sonicvisualiser.org/sv2010.pdf.

Coral Creek Software. 2017. *SQLite Expert Professional*. Computer software, version 5.2.3.314 (x64). Accessed February 20th, 2017. https://www.sqliteexpert.com/.

Dittmar, Christian, Estefania Cano, and Sascho Grollmisch. 2017. *Songs2see Editor*. Ilmenau, Germany: Fraunhofer Institute For Digital Media Technology. Computer software, version 1.61. Accessed February 20th, 2017. http://www.songs2see.com/en/songs2see_produkte/editor.

Nienhuys, Han-Wen and Jan Nieuwenhuizen. 2003. "LilyPond, A System For Automated Music Engraving". In *Proceedings of the XIV Colloquyum on Musical Informatics (XIV CIM 2003), Firenze, Italy, May 8-9-10*, edited by Nicola Bernardini, Francesco Giomi, and Nicola Giosmin, 167–172. Accessed March 16th, 2022. https://lilypond.gitlab.io/static-files/media/xivcim.pdf.

R Core Team. 2017. *R: A Language and Environment For Statistical*. Vienna, Austria: R Foundation For Statistical Computing. https://www.R-project.org.

Robinson, Andy. 2014. *Transcribe!* Seventh String Software. Computer software, version 8.40.0. Accessed February 20th, 2017. https://www.seventhstring.com/.

Appendices

Appendix A

Additional Graphs

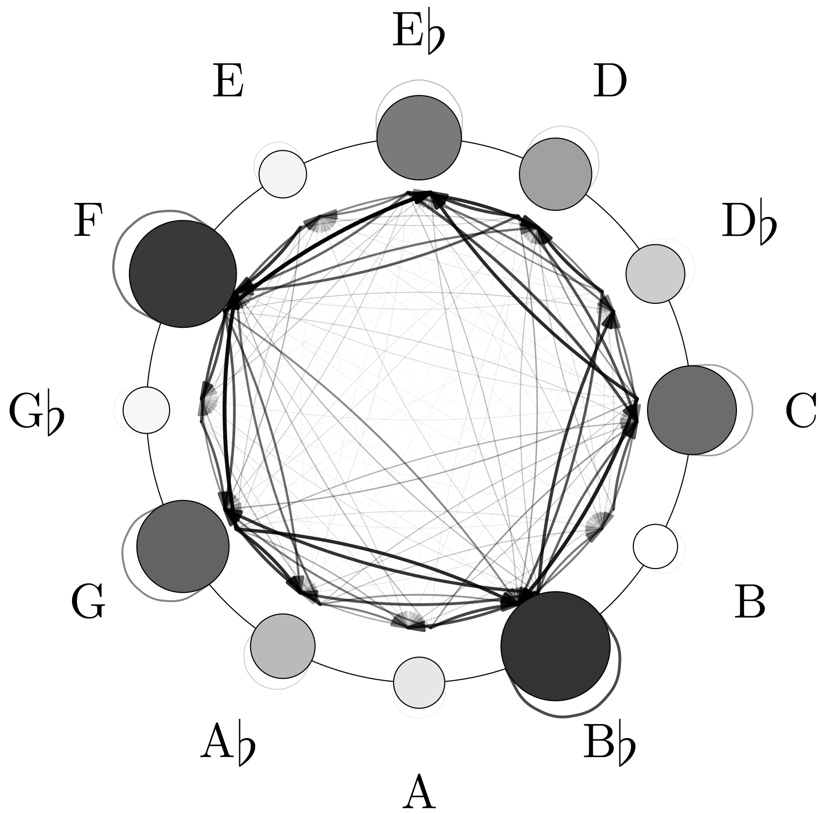

Figure A.1: Bigram Pitch Class Distribution for entire corpus of Green's improvisations.

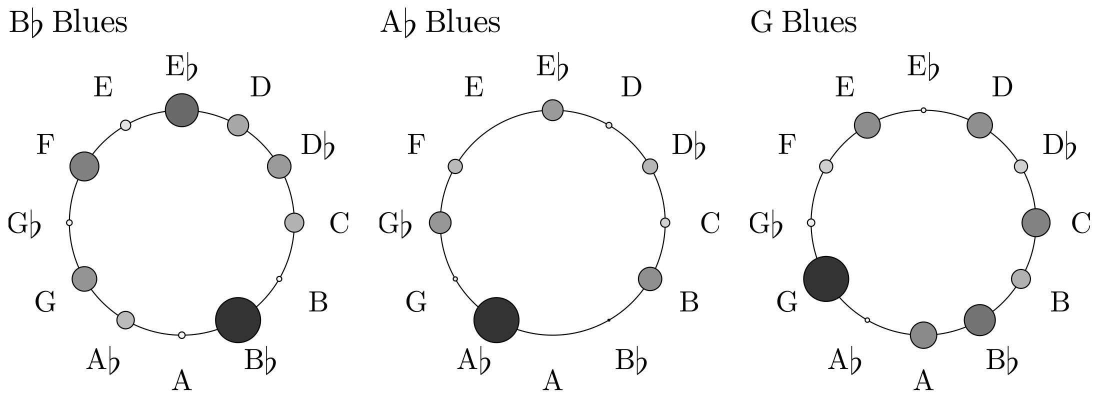

Figure A.2: Pitch class circle maps for each blues key signature in Green's corpus.

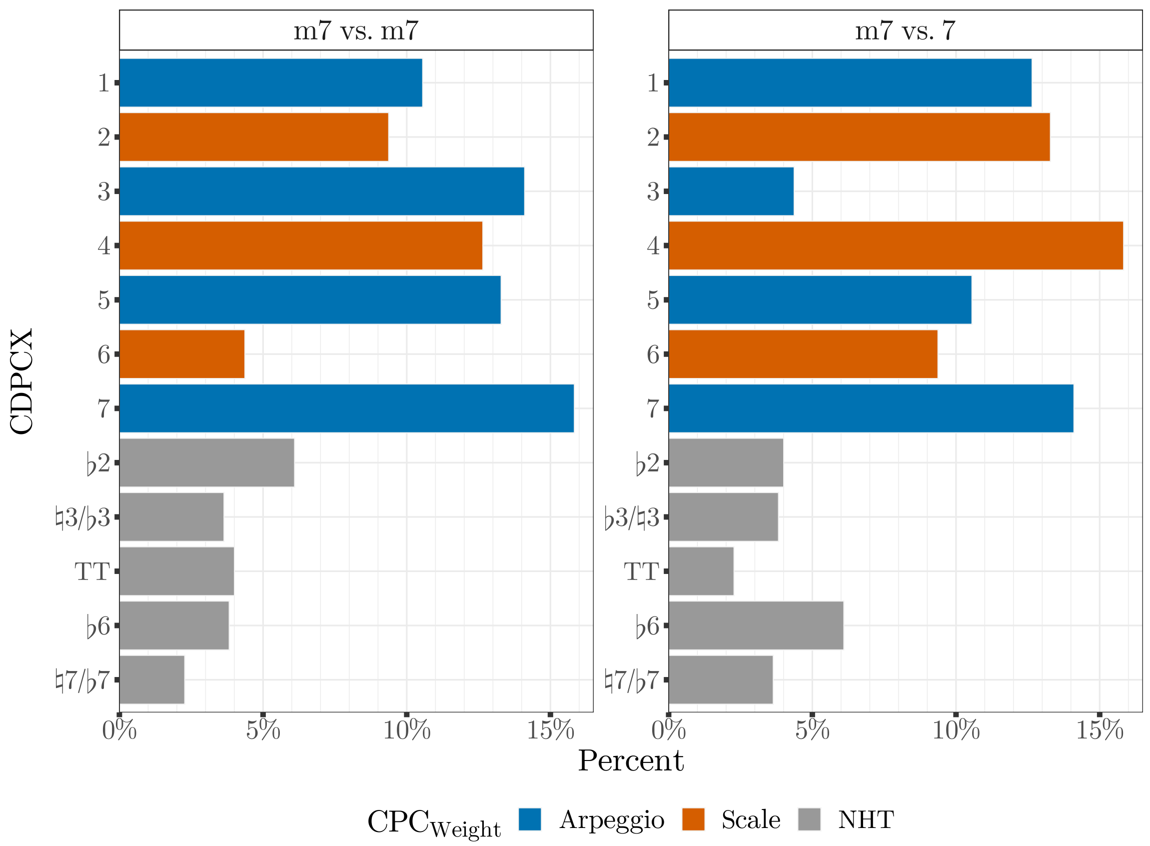

Figure A.3: CDPCX distribution for m7 in the beat before a dominant resolution to 7, compared to both chord types, in Green's corpus.

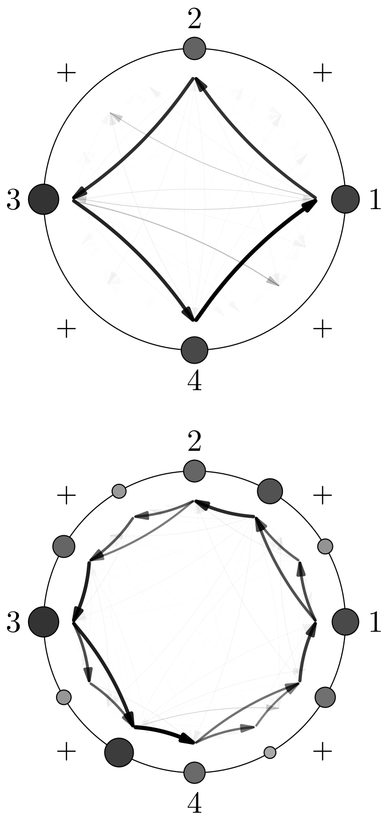

Figure A.4: Most common divisions for full bars in $\frac{4}{4}$, in Green's corpus.

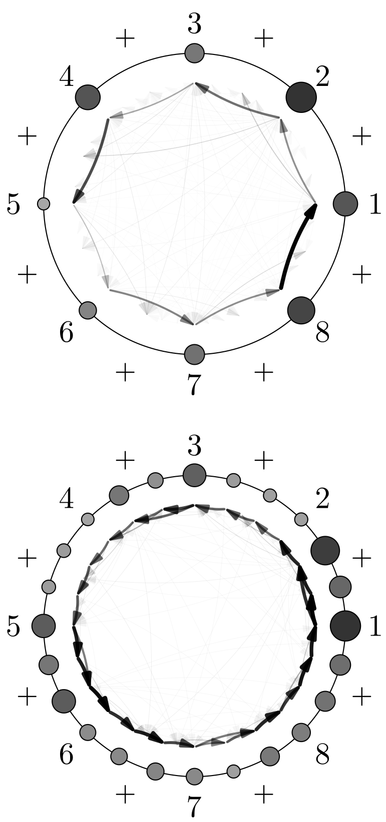

Figure A.5: Most common divisions for full bars in $\frac{8}{4}$, in Green's corpus.

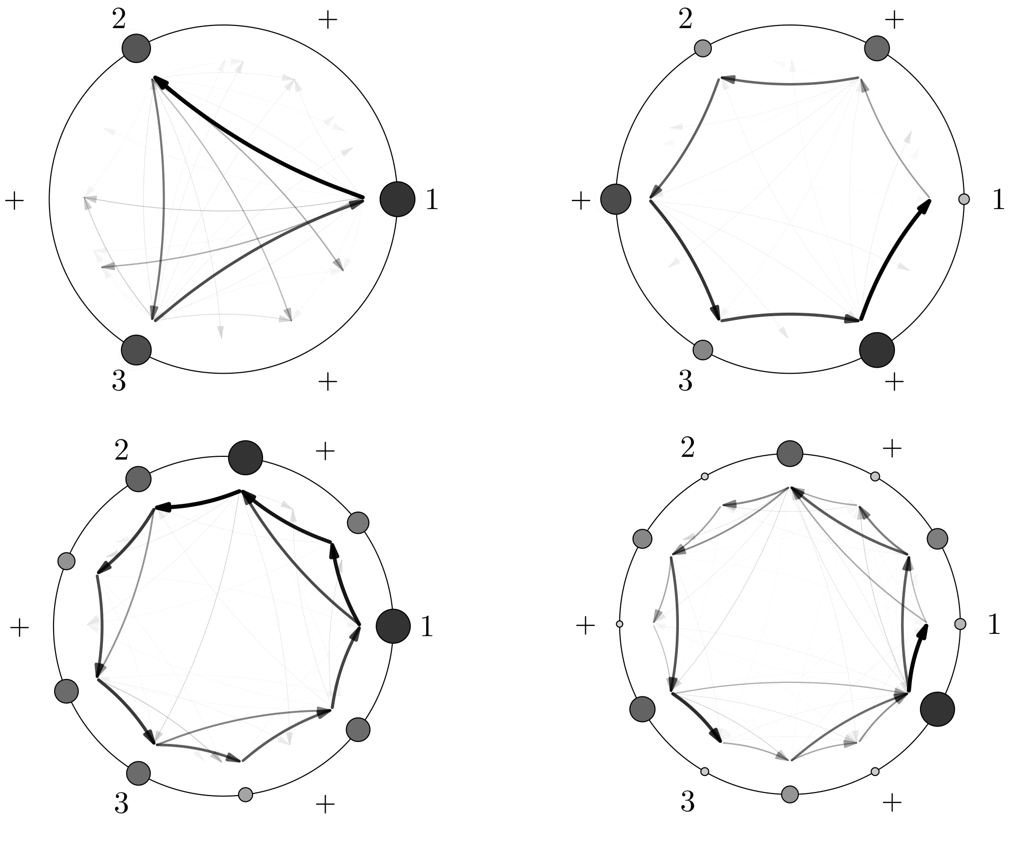

Figure A.6: Most common divisions for full bars in $\frac{3}{4}$, in Green's corpus.

Appendix B

Code Examples

Transforming $\frac{8}{4}$ to $\frac{4}{4}$

These two functions transformed the bars (Code block B.1) and beats (Code block B.2) of improvisations in $\frac{8}{4}$ to $\frac{4}{4}$. Examples of how these were applied to the data is shown in Code block B.3

```
transform84Bar <- function(bar, beat, timeSignature){
  if(timeSignature == "8/4"){
    if(beat < 5){
      n <- (bar * 2) - 1
    }
    else{
      n <- bar * 2
    }
  }
  else{
    n <- bar
  }
  return(n)
}
```

Code B.1: Code for converting bars of $\frac{8}{4}$ to $\frac{4}{4}$.

```
transform84Beat <- function(beat, timeSignature){
  if(timeSignature == "8/4"){
    if(beat < 5){
      n <- beat
    }
    else{
      n <- beat - 4
    }
  }
  else{
    n <- beat
  }
  return(n)
}
```

Code B.2: Code for converting beats of $\frac{8}{4}$ to $\frac{4}{4}$.

```r
data$bar_84 <- mapply(transform84Bar, data$bar, data$beat, data$signature)
data$beat_84 <- mapply(transform84Beat, data$beat, data$signature)
```

Code B.3: Use of functions for converting bars and beats of $\frac{8}{4}$ to $\frac{4}{4}$.

Circle Map Code

The inspiration for the circle maps came from the Jazzomat Research Project's website gallery (Jazzomat Research Project 2017), and two papers by Frieler (Frieler 2007, 2008). The plots as described on the Jazzomat Research Project's website required sequence features extracted from *MeloSpy* and a series of programs, "awk, ps-tricks, latex, and gimp" (Jazzomat Research Project 2017). As this research was based on the raw underlying data, it was decided to write a custom function to create circle maps from the raw data, and for a range of features. Circle maps were a useful visual aid for investigating features within music, as many elements of music are cyclical in nature (e.g. metrical frameworks or pitches). To generate the circle maps eight separate functions were written, with their code available in Appendix E.2, Functions.R. The first seven were helper functions, while the eighth was the function called to generate the plots. The helper functions were:

- `normaliseCircleMap`: scales unigram circles so that the most frequent class has the same size and opacity (if normalise_unigram == TRUE);
- `translatePoint`: calculates position of data around the main circle;
- `normaliseNotePlacementInBeats`: if the metre_raw feature was plotted, calculates position of the raw metrical data around the main circle;
- `generateUnigramData`: generates the data for the unigram circles;
- `generateClassLabels`: sets the default class labels, or custom labels, and calculates their position around the main circle;
- `generateBigramData`: generates the bigram data;
- `plotCircleGraph`: code for plotting the graphs;
- `circleMapGraphFromRaw`: function called to generate the graphs, calls the previous helper functions, and set the overall options for the graph e.g. the feature, whether the graph is unigram or bigram, the size of the labels, the bigram offset, or the curvature of the bigram line.

These functions allowed for great flexibility in the plotting of the circle map figures.

Swing (Beat-Upbeat Ratio)

The following functions were used to generate the BUR data used in the analysis of Green and the machine learning classification and comparative analysis (Code blocks B.4 and B.5). As discussed in the swing section of the Micro Domain, there was an error in the code of `manualSwing` that incorrectly labelled some beats as being a swung pair when they should have been ignored. Specifically, this applied to some beats of division four, with notes on the 2nd or 3rd tatum and the 4th tatum. As a result, these functions should not be used in any future research, and the updated `manualSwingMarkerAndBUR` should be used instead (code block B.6). Slight changes have been made to the line breaks within the code to fit within the paper margins.

```
# Function to manually calculate swing based on
# the beat, division, tatums, duration, and ioi
manualSwing <- function(df){
  # Marks swing beats as 1, rest as 0
  # (similar to swing markers from MeloSpy)
  df <- as.data.frame(df %>%
            dplyr::rename() %>%
            group_by(id, bar, beat) %>%
            mutate(swing = ifelse((division==2 & sum(tatum)==3), 1,
              ifelse((division==3 & sum(tatum)==4), 1,
              ifelse(((division==4 & sum(tatum)==5 & min(tatum)!=2) |
                    (division==4 & length(division)==2 & sum(tatum==4))),
                    1,0)))))
  # Calculates the swing ratio for eligible binary notes based on
  # the ioi of the first note and duration of the second note
  df$swingRatio <- (df %>%
                      dplyr::rename() %>%
                      group_by(id, bar, beat, division) %>%
                      mutate(duration1 = dplyr::lead(duration),
                             swingRatio = ifelse((swing==1),
                                          (ioi_raw/duration1), NA )) %>%
                      pull(swingRatio))
  # Removes the swing beat marker
  df <- subset(df, select = -(swing))
  # Puts swing ratios between 0.98 and 3.02 into a new column named swing
  df <- as.data.frame(df %>%
                        dplyr::rename() %>%
                        mutate(swing = ifelse((swingRatio < 0.98 |
                                              swingRatio > 3.02),
                                              NA, swingRatio)))

  data.frame(df)
}
```

Code B.4: Code for manually generating the BUR of swung note pairs.

```r
# This function expands the markers for notes in a swing pair, with:
# 0 -> no swing
# 1-> first note of swing pair
# 2 -> second note of swing pair
swingMarker <- function(df){
  df$swingMarker <- NA

  for(i in 1:length(df$swing)){
    # If there is a swing value,
    # mark current note as the first note in a swing pair
    if(!is.na(df$swing[i])){
      k <- 1
      # Else, if the previous note had a swing value,
      # mark current note as second note in a swing pair
    } else if(i>1 && !is.na(df$swing[i-1])){
      k <- 2
      # Else, mark the note as not being part of a swing pair
    } else{
      k <- 0
    }
    df$swingMarker[i] <- k
  }

  # Converts the swing markers into a factor
  df$swingMarker <- factor(df$swingMarker, levels = c("0","1","2"))

  # Returns the dataframe
  data.frame(df)
}
```

Code B.5: Code for manually labelling the swung note pairs.

```r
manualSwingMarkerAndBUR <- function(df,
                                    lower_limit = 0.98,
                                    upper_limit = 3.02){

  df <- df %>%
  group_by(id,bar,beat) %>%
  # creates columns for comparing the current notes against
  mutate(divisionNext = dplyr::lead(division),
         tatumNext = dplyr::lead(tatum),
         durationNext = dplyr::lead(duration)) %>%
  # labels each beat uniquely, and counts the number of notes in each beat
  mutate(beat_id = cur_group_id(), notesPerBeat = n()) %>%
  group_by(beat_id) %>%
  # initial binary marker for swing values
  mutate(swingMarker = case_when(
    notesPerBeat == 2 & division == 2 & tatum == 1 & tatumNext == 2 ~ 1,
    notesPerBeat == 2 & division == 3 & tatum == 1 & tatumNext == 3 ~ 1,
    notesPerBeat == 2 & division == 4 & tatum == 1 & tatumNext == 3 ~ 1,
    notesPerBeat == 2 & division == 4 & tatum == 1 & tatumNext == 4 ~ 1,
    TRUE ~ 0
  )) %>%
  # where there was a swung pair, calculates the BUR
  mutate(swingRatio = ifelse(swingMarker == 1, ioi_raw/durationNext,NA)) %>%
  # excludes BUR outside the set limits and sets to new column
  mutate(swing = case_when(
    is.na(swingRatio) ~ NA_real_,
    swingRatio < lower_limit ~ NA_real_,
    swingRatio > upper_limit ~ NA_real_,
    TRUE ~ swingRatio
  )) %>%
  # updates swing markers so that the first note of the pair is labelled 1,
  # the second note 2, and all others 0
  mutate(swingMarkerPrev = dplyr::lag(swingMarker),
         swingMarker = case_when(
           !is.na(swing) ~ 1,
           swingMarker == 0 & swingMarkerPrev == 1 ~ 2,
           TRUE ~ 0
         )) %>%
  # ungroups and removes the extra columns created that are no longer needed
  ungroup() %>%
  select(-c(divisionNext, tatumNext, durationNext,
            beat_id, notesPerBeat, swingMarkerPrev))

  return(df)

}
```

Code B.6: New code for manually marking swung note pairs and calculating the BUR.

Phrase Contour Examination

The code in Code block B.7 was used to examine the phrase contours to assess the Abesser contour codes.

```
# Randomly selects improvisation id and phrase, assigns them to variables
# assigns the abesser contour code for that id and phrase to a variable
testID <- ((df %>% sample_n(1))$id)
testPhrase <- ((df %>%
                  filter(id == testID) %>%
                  sample_n(1))$phrase_id_raw)
testContour <- ((df %>%
                  filter(id == testID & phrase_id_raw == testPhrase) %>%
                  sample_n(1))$abesserCode)

# Plots the pitch values against the note number in the phrase as scatterplot
# y-axis is extended to the nominal tessitura range of the guitar
# line of best fit (quadratic) plotted over the top
df %>%
  filter(id == testID & phrase_id_raw == testPhrase) %>%
  mutate(n=row_number()) %>%
  ggplot(aes(x=n,y=pitch_raw))+
  geom_point()+
  scale_y_continuous(limits = c(40,84))+
  geom_smooth(formula = y ~ poly(x, 2),method="lm")+
  labs(title=testContour)

}
```

Code B.7: Code for examining phrase contours.

Appendix C

Transcription Details

Sonic Visualiser layers and plugins

The final SV transcription files had ten layers, including the three default layers in each new SV file (Pane, Ruler, and Waveform), with the other seven being generated throughout the transcription process. All extra plugins were download from the vamp-plugins website (https://www.vamp-plugins.org/download.html). These plugins were[1]:

Adaptive Spectrogram

Centre for Digital Music at Queen Mary, University of London (2017), from 'QM Vamp Plugins'.

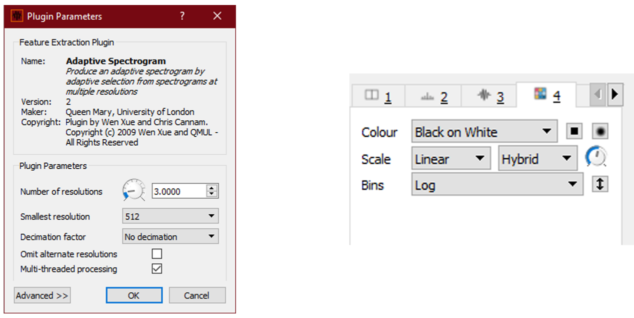

Figure C.1: Adaptive Spectrogram Settings.

[1] Screenshots were included where appropriate of the plugin settings (left) and SV view settings (right)

MELODIA - Melody Extraction

Salamon and Gómez (2012), from 'MELODIA - Melody Extraction'.

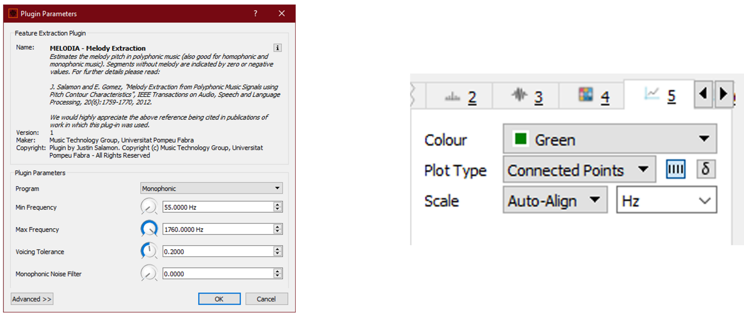

Figure C.2: Melodia - Melody Extractions Settings.

Notes - Automated MIDI Transcription

The MIDI track that came from *Songs2See*. Each of the two MIDI tracks were assigned a different colour, with the scale set to 'MIDI Notes'.

Time Instants (Beats)

Centre for Digital Music at Queen Mary, University of London (2017) from 'QM Vamp Plugins'.

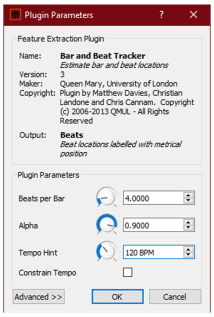

Figure C.3: Bar and Beat Tracker Settings.

Notes - Manual MIDI Transcription

The layer where the manual transcription was completed.

Regions - Phrases

The layer where phrases were annotated. The scale was set to 'Equal Spaced' and the plot type to 'Segmentation'.

Text

Although the Text layer had to be included it could be, and was, left empty in all the transcriptions. It could be used to add additional annotations such as slides, dead-notes, or vibrato (full list available on the Jazzomat Research Project website, Jazzomat Research Project 2017).

The final SV transcription files included in Appendix E.1 do not include the adaptive spectrogram, MELODIA, or automated MIDI transcription layers. The automatically generated time instants (beats) layer is also replaced by the manually calculated and annotated layer.

Creating the *Sonic Visualiser* files

To consistently initiate the file in *Sonic Visualier* another program with three functions, `automateSV`, was written in C#. Two of these functions set up the file to be transcribed, while the last checked the file after the transcription. The first function asked for the audio file to be transcribed, and opened that file in *Sonic Visualiser*. The second function created nearly all of the layers needed to accurately transcribe the improvisations, initialising the plugins listed above, with the correct settings. This also prompted the user to open the related automatic transcription file from *Songs2See*. The `automateSV` program can be found in Appendix E.8.

Transcribing the Solos

The pitch transcriptions were based primarily on the *Songs2See* MIDI transcription, the Melodia melody extraction, and the adaptive spectrogram. The adaptive spectrogram also informed the onset and offset of the notes. To ensure a high level of accuracy in the onset and offset times of each note, instead of using a mouse to draw the notes a Wacom Intuos Pro tablet and pen was used.[2]

Once the transcription was completed, the notes layer was exported as a csv file. This file was then imported into the 'check notes' function of the `automateSV` program. This checked each note in the transcription to see if there was any overlap between any of the notes, ensuring a completely monophonic transcription. If the program found any overlapping notes, the onset time of the first note would be reported, and this would then be fixed within the *Sonic Visualiser* file.[3] Once fixed, the note layer would again be exported as a csv and the function would check for any overlaps. Once no overlaps were found, the next stage of the transcription could begin.

[2]In the settings for the tablet, the zoom function was employed so that the whole area of the tablet was focused on only a small section around where the notes appeared in *Sonic Visualiser*, this again increased the accuracy of the onset and offset of each note.

[3]These errors would nearly always be caused by the pen drawing an overlap of only a few pixels, but enough to cause the resulting file to be polyphonic.

The accompaniment track, or the combined track for the three songs that could not be split, were also opened using the `automateSV` program. The prompt to open the *Songs2See* MIDI file was dismissed, and the Melodia and Adaptive Spectrogram layers were deleted. In its place a basic spectrogram layer is added,[4] with the settings: Colour - White on Black; Scale - dBVˆ2, None; Window - 256 (although this could range from 64–512 depending on the quality of the recording), 50%; Bins - All Bins, Linear. Finally, a new beats layer was added to transcribe the new beat track onto. The beat transcription was then completed as described in the main text. Once the beat transcription was completed, the beat track was exported as a CSV for annotation.

The form labels were based on standard jazz labelling, being independent of the chorus. For example, if an improvisation went for two choruses over a 32-bar AABA form the labels would be – on bars 1, 9, 17, 25, 33, 41, 49, and 57 respectively – A1 A2 B1 A3 A1 A2 B1 A3. The only form label that did not follow the standard labelling was 'I1', for introduction. This was used when a solo began at the end of the head, or in the last few bars of the previous solo, but before a new form had started. While this could have been labelled consistently with the nominal form location that it started in, this system was set in place by the Jazzomat Research Project and was followed for consistency. These sections would also be labelled with negative or zero bar numbers, so that bar one would match with the first bar of the form (e.g. if there were two bars in $\frac{3}{4}$ time before the start of the form, they would be labelled -1.1, -1.2, -1.3, 0.1, 0.2, 0.3, 1.1 . . .). On occasion, the first label needed to be annotated not onto the first transcribed beat, but what would be the first beat, which then would have a negative time value. Negative times for beats occurred when the extracted transcription file started with the first note occurring before the first beat, in these cases the nominal first beat was extrapolated backwards based on the timing of the first few beats that were transcribed.

For phrase annotations, issues occurred not for obvious phrases. For example, a series of eleven continuous quavers, with a bar of rests before and after, would be classified as a phrase by all transcribers. Instead, it was the edge cases that caused issues. For example, when there two fairly continuous streams of notes, with a small rest between, to one transcriber it may have sounded like two separate musical ideas, and annotated as two phrases, whereas another may have heard it as one continuous phrase with multiple parts. Other issues occurred when one or two notes were played close, but not definitely connected, to any other phrases that came before or after. It was again up to the transcriber to decide, based upon their musical experiences, whether the notes were connected to any other phrase, or if

[4]Built into the *Sonic Visualiser* software, not from a plugin.

they constituted a musical phrase by themselves. For the purposes of this research the basic principle was:

- small or singular notes should be grouped with other phrases if there was any doubt;
- where one phrase could be split into two, with only a small rest between phrase but with distinct musical ideas, they were annotated as two phrases.

However, the caveat that these decisions were still largely dependent on the authors prior experiences stands.

Data Manipulation

The `chordSetup` Function

The `chordSetup` function provided many of the features for analysing Green's pitches in relation to the chords. This function can be found in Functions.R in Appendix E.2. As discussed in the Section 3.1.2 – dealing with the issues of transcription – chord annotation, and comparing the notes played to the chords of a piece presented a number of issues. Beyond the issues raised there, there were other situations in which the chord changes need to be considered, specifically related to the previous or following chords. These issues were because it is not uncommon to either anticipate the change, or delay the resolution, of a chord progression, with anticipations being most common in the beat before a chord change (especially if the chord changes in the first beat of the bar), and the delays to resolution also occurring within a beat or two after the chord change. To be able to fully investigate these phenomena, two sets of features were created, one for the previous chord and one for the following chord, with features including: chord type; chord tonic; and CDPCX. It was necessary for this data to be drawn from the beat tracks generated and exported from the SQLite3 database, instead of from the data exported by *MeloSpy*, as *MeloSpy* did not generate any data for a chord (or bar), if there were no note events. For example, in bar 4 of Figure C.4 (Green 1962j), the G on beat 3 played over E♭7, the actual previous chord was B♭m7, while the data would report the previous chord as A♭Δ7.

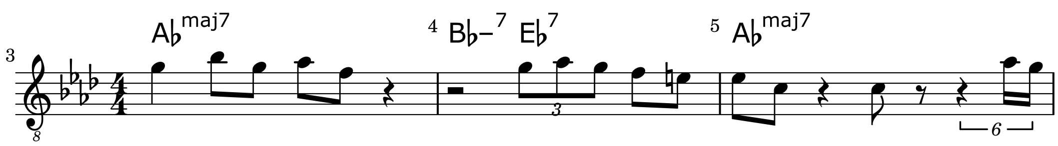

Figure C.4: Example of potential chord annotation issues, *Red River Valley* (1963), bars 3–5.

Additionally, there were 355 note events where the chord of the moment according to *MeloSpy* was in disagreement with the raw beat track chord data. This was due to how *MeloSpy* determined the chord of the moment. If Green anticipated or delayed a note at the point of a chord change, even if only slightly, *MeloSpy* would classify the note as being played over whatever the current chord was, even if the note was nominally played over a different chord. Therefore, in addition to the two set of data mentioned above, the `chordSetup` function also generated a set of nominal chord data, which compared the notes to the bar, beat, division, and tatum features from the FlexQ algorithm. This extra data allowed for the greatest flexibility when analysing Green's note choice in relation to the nominal chords of the piece.

Appendix D

Summary Performance Metrics

The model summary performance metrics are separated by classifier and abstraction. The summary performance metrics from each model that performed significantly better than the NIR were presented within a table. The performance metrics were split by comparison (n-way, one-vs-all, one-vs-one), with the n-way class results listed at the bottom. The summary metrics included were: Matthews Correlation Coefficient (MCC); Accuracy; Balanced Accuracy (BA); F-score for the "positive" (F_1^+) and "negative" (F_1^-) classes. The mathematical formulas for each of these metrics can be found in Table 12.1. The results for the solo level comparisons were included, but due to the small testing dataset – only twenty-seven data points: twelve Green improvisations, four Parker, six Coltrane, and five Davis – their results were not meaningful.

D.1 Tree: C4.5 Like

Solo Level

Table D.1: C4.5-like solo level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.69	77.78%	84.47%	74.39%	92.28%	0.75	0.92
Parker vs Davis	1.00	100.00%	100.00%	100.00%	100.00%	1.00	1.00
n-way Class Results							
	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.78	88.89%	88.33%	90.91%	87.50%	0.87	0.90
Coltrane	0.66	88.89%	80.95%	80.00%	90.91%	0.73	0.93
Parker	0.61	88.89%	83.15%	60.00%	95.45%	0.67	0.93
Davis	0.66	88.89%	85.45%	66.67%	95.24%	0.73	0.93

Phrase Level

Table D.2: C4.5-like phrase level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.54	68.91%	74.88%	60.28%	88.62%	0.61	0.89
Green vs All	0.47	74.01%	73.48%	67.31%	79.02%	0.69	0.78
Coltrane vs All	0.62	82.76%	81.14%	75.35%	86.80%	0.76	0.87
Davis vs All	0.70	91.30%	87.21%	69.03%	96.37%	0.75	0.95
Green vs Coltrane	0.47	73.81%	73.68%	75.71%	71.63%	0.76	0.72
Green vs Davis	0.81	91.86%	91.48%	96.22%	82.08%	0.94	0.86
Coltrane vs Davis	0.89	94.84%	95.69%	99.01%	87.04%	0.96	0.92
Parker vs Davis	0.58	81.51%	77.31%	78.05%	82.86%	0.70	0.87

n-way Class Results

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.48	74.84%	73.54%	70.21%	77.75%	0.68	0.79
Coltrane	0.60	81.58%	80.56%	72.37%	87.11%	0.75	0.86
Parker	0.14	88.32%	56.31%	23.08%	92.79%	0.20	0.94
Davis	0.75	93.09%	89.13%	75.47%	96.81%	0.79	0.96

Bar – Sliding Window (4, 2)

Table D.3: C4.5-like bar (4, 2) level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.53	67.30%	75.43%	60.28%	88.28%	0.61	0.88
Green vs All	0.53	77.84%	76.70%	68.81%	83.51%	0.71	0.82
Coltrane vs All	0.61	81.28%	80.99%	74.20%	86.37%	0.77	0.84
Davis vs All	0.57	86.72%	81.01%	57.41%	94.34%	0.64	0.92
Green vs Coltrane	0.67	83.33%	83.43%	80.58%	86.32%	0.83	0.83
Green vs Parker	0.55	86.08%	78.79%	92.76%	59.57%	0.91	0.63
Green vs Davis	0.55	79.46%	78.51%	88.35%	64.04%	0.85	0.70
Coltrane vs Parker	0.66	89.61%	85.11%	95.42%	66.33%	0.94	0.72
Coltrane vs Davis	0.75	88.60%	89.02%	95.48%	75.86%	0.92	0.82
Parker vs Davis	0.67	84.25%	84.58%	71.72%	92.26%	0.78	0.88

n-way Class Results

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.48	76.10%	73.46%	68.93%	79.77%	0.66	0.82
Coltrane	0.64	82.98%	81.90%	78.89%	85.49%	0.78	0.86
Parker	0.35	89.20%	69.36%	35.85%	95.21%	0.40	0.94
Davis	0.52	86.33%	77.01%	57.45%	92.66%	0.60	0.92

Bar – Sliding Window (2, 1) Level

Table D.4: C4.5-like bar (2, 1) level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.49	65.05%	73.29%	57.37%	87.45%	0.58	0.88
Green vs All	0.52	77.04%	75.76%	70.50%	81.13%	0.70	0.81
Coltrane vs All	0.66	83.84%	83.27%	77.23%	88.15%	0.79	0.87
Green vs Coltrane	0.66	83.23%	83.21%	82.90%	83.58%	0.84	0.83
Green vs Davis	0.60	82.45%	81.97%	91.70%	65.27%	0.87	0.72
Coltrane vs Parker	0.59	86.11%	82.64%	94.63%	58.66%	0.91	0.67
Coltrane vs Davis	0.80	91.39%	91.11%	95.95%	81.65%	0.94	0.86
Parker vs Davis	0.43	74.68%	71.21%	65.32%	79.09%	0.62	0.81
n-way Class Results							
	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.51	76.90%	74.87%	72.16%	79.49%	0.69	0.82
Coltrane	0.57	80.00%	78.53%	73.89%	83.59%	0.73	0.84
Parker	0.31	89.00%	65.90%	34.90%	94.39%	0.36	0.94
Davis	0.44	84.19%	73.85%	48.52%	92.31%	0.53	0.90

Note Level

Table D.5: C4.5-like note level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.67	78.31%	80.55%	70.40%	92.27%	0.70	0.92
Green vs All	0.59	79.32%	79.71%	71.79%	85.98%	0.77	0.82
Coltrane vs All	0.54	77.90%	77.62%	68.91%	84.47%	0.72	0.82
Davis vs All	0.61	90.26%	84.18%	58.11%	96.48%	0.66	0.94
Green vs Coltrane	0.80	89.78%	89.73%	89.53%	90.06%	0.90	0.89
Green vs Parker	0.61	88.50%	81.54%	94.02%	64.29%	0.93	0.67
Green vs Davis	0.70	89.37%	85.35%	93.32%	76.52%	0.93	0.77
Coltrane vs Davis	0.71	88.07%	88.21%	95.92%	70.42%	0.92	0.78
Parker vs Davis	0.39	70.37%	69.40%	62.82%	75.68%	0.64	0.75

n-way Class Results

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.79	89.70%	89.41%	87.30%	91.40%	0.88	0.91
Coltrane	0.67	84.56%	83.82%	79.15%	87.99%	0.80	0.87
Parker	0.37	89.59%	69.80%	39.33%	95.02%	0.42	0.94
Davis	0.64	92.77%	79.16%	75.82%	94.65%	0.68	0.96

D.2 Tree: C5.0

Solo Level

Table D.6: C5.0 solo level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.58	70.37%	79.29%	71.25%	89.31%	0.70	0.89
Green vs All	0.70	85.19%	85.00%	83.33%	86.67%	0.83	0.87
Coltrane vs All	0.78	92.59%	83.33%	100.00%	91.30%	0.80	0.95
Parker vs Davis	1.00	100.00%	100.00%	100.00%	100.00%	1.00	1.00
n-way Class Results							
	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.55	77.78%	77.50%	75.00%	80.00%	0.75	0.80
Coltrane	0.78	92.59%	83.33%	100.00%	91.30%	0.80	0.95
Parker	0.61	88.89%	83.15%	60.00%	95.45%	0.67	0.93
Davis	0.43	81.48%	73.18%	50.00%	90.48%	0.55	0.88

Phrase Level

Table D.7: C5.0 phrase level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.75	83.22%	84.08%	76.24%	94.11%	0.75	0.94
Green vs All	0.70	85.36%	84.69%	82.72%	87.12%	0.82	0.88
Coltrane vs All	0.74	88.34%	85.34%	89.94%	87.67%	0.82	0.91
Davis vs All	0.82	95.24%	90.40%	86.02%	96.90%	0.85	0.97
Green vs Coltrane	0.80	90.04%	89.80%	88.85%	91.58%	0.91	0.89
Green vs Parker	0.55	88.93%	70.99%	89.67%	81.48%	0.94	0.57
Green vs Davis	0.94	97.38%	96.91%	98.38%	94.85%	0.98	0.95
Coltrane vs Parker	0.59	88.64%	75.36%	90.00%	79.41%	0.93	0.64
Coltrane vs Davis	0.93	97.10%	97.61%	99.52%	92.23%	0.98	0.95
Parker vs Davis	0.80	91.10%	89.88%	87.76%	92.78%	0.87	0.93

n-way Class Results

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.76	87.99%	88.17%	82.77%	92.08%	0.86	0.90
Coltrane	0.80	90.95%	89.50%	89.16%	91.85%	0.87	0.93
Parker	0.36	91.78%	64.57%	50.00%	94.10%	0.39	0.96
Davis	0.85	95.72%	94.08%	83.02%	98.41%	0.87	0.97

Bar – Sliding Window (4, 2) Level

Table D.8: C5.0 bar (4, 2) level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.81	87.00%	88.21%	82.39%	95.49%	0.82	0.95
Green vs All	0.78	89.88%	88.28%	89.27%	90.19%	0.86	0.92
Coltrane vs All	0.85	92.74%	91.79%	93.46%	92.33%	0.90	0.94
Parker vs All	0.62	95.13%	75.88%	78.57%	96.06%	0.63	0.97
Davis vs All	0.74	93.31%	84.71%	84.83%	94.68%	0.78	0.96
Green vs Coltrane	0.89	94.57%	94.61%	93.16%	95.97%	0.94	0.95
Green vs Parker	0.73	92.51%	83.17%	93.52%	86.36%	0.96	0.77
Green vs Davis	0.82	92.07%	91.35%	95.21%	85.47%	0.94	0.87
Coltrane vs Parker	0.78	94.09%	86.37%	95.01%	88.57%	0.97	0.81
Coltrane vs Davis	0.88	94.82%	94.29%	97.01%	89.83%	0.96	0.91
Parker vs Davis	0.77	90.16%	88.35%	86.25%	91.95%	0.85	0.93

n-way Class Results

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.83	91.87%	91.45%	88.24%	94.05%	0.89	0.94
Coltrane	0.87	93.69%	93.41%	91.71%	94.97%	0.92	0.95
Parker	0.62	94.74%	79.53%	68.92%	96.71%	0.65	0.97
Davis	0.77	93.69%	88.47%	80.70%	96.23%	0.81	0.96

Bar – Sliding Window (2, 1) Level

Table D.9: C5.0 bar (2, 1) level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.77	84.19%	86.06%	80.09%	94.42%	0.79	0.94
Green vs All	0.78	89.43%	88.41%	88.28%	90.10%	0.86	0.92
Coltrane vs All	0.85	92.83%	91.54%	94.19%	92.12%	0.90	0.94
Parker vs All	0.53	94.17%	68.72%	81.25%	94.69%	0.52	0.97
Davis vs All	0.76	94.23%	85.49%	87.14%	95.26%	0.79	0.97
Green vs Coltrane	0.86	92.93%	92.90%	91.53%	94.47%	0.93	0.93
Green vs Parker	0.72	92.25%	82.57%	93.39%	85.19%	0.95	0.75
Green vs Davis	0.81	92.24%	90.24%	94.38%	86.69%	0.95	0.86
Coltrane vs Parker	0.77	93.52%	85.14%	94.13%	89.91%	0.96	0.80
Coltrane vs Davis	0.90	95.98%	95.40%	97.56%	92.19%	0.97	0.93
Parker vs Davis	0.66	84.50%	82.15%	80.16%	86.59%	0.77	0.88

n-way Class Results

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.77	89.18%	88.58%	86.16%	91.08%	0.86	0.91
Coltrane	0.84	92.28%	91.89%	89.31%	94.11%	0.90	0.94
Parker	0.59	94.47%	75.91%	72.28%	95.92%	0.62	0.97
Davis	0.72	92.46%	87.88%	72.60%	96.55%	0.77	0.96

Note Level

Table D.10: C5.0 note level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.84	89.70%	89.44%	84.59%	96.51%	0.84	0.96
Green vs All	0.96	98.03%	98.05%	97.11%	98.69%	0.98	0.98
Coltrane vs All	0.81	91.14%	90.80%	87.61%	93.38%	0.88	0.93
Davis vs All	0.93	98.47%	94.57%	98.06%	98.52%	0.94	0.99
Green vs Coltrane	0.98	98.76%	98.75%	98.67%	98.85%	0.99	0.99
Green vs Parker	0.82	95.13%	90.25%	96.58%	87.50%	0.97	0.85
Green vs Davis	0.88	95.91%	93.01%	96.35%	94.29%	0.97	0.91
Coltrane vs Davis	0.91	96.75%	94.86%	97.17%	95.37%	0.98	0.93
Parker vs Davis	0.79	89.95%	89.44%	88.00%	91.23%	0.87	0.92

n-way Class Results

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.93	96.71%	96.81%	94.82%	98.10%	0.96	0.97
Coltrane	0.84	92.33%	91.98%	89.49%	94.12%	0.90	0.94
Parker	0.57	93.43%	79.07%	60.26%	96.53%	0.61	0.96
Davis	0.85	96.93%	89.89%	93.81%	97.30%	0.87	0.98

D.3 Random Forest

Solo Level

Table D.11: Random forest solo level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.80	85.19%	90.24%	87.20%	95.21%	0.86	0.95
Green vs All	0.93	96.30%	96.67%	92.31%	100.00%	0.96	0.97
Davis vs All	1.00	100.00%	100.00%	100.00%	100.00%	1.00	1.00
Green vs Coltrane	0.88	94.44%	91.67%	92.31%	100.00%	0.96	0.91
Green vs Davis	1.00	100.00%	100.00%	100.00%	100.00%	1.00	1.00
Coltrane vs Davis	0.83	90.91%	90.00%	85.71%	100.00%	0.92	0.89
Parker vs Davis	1.00	100.00%	100.00%	100.00%	100.00%	1.00	1.00
n-way Class Results							
	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.85	92.59%	92.50%	91.67%	93.33%	0.92	0.93
Coltrane	0.66	88.89%	75.00%	100.00%	87.50%	0.67	0.93
Parker	0.70	88.89%	93.48%	57.14%	100.00%	0.73	0.93
Davis	1.00	100.00%	100.00%	100.00%	100.00%	1.00	1.00

Phrase Level

Table D.12: Random forest phrase level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	$F_1{}^+$	$F_1{}^-$
n-way	0.75	83.22%	84.89%	80.36%	94.01%	0.78	0.94
Green vs All	0.66	83.88%	82.50%	83.78%	83.94%	0.79	0.87
Coltrane vs All	0.76	89.00%	85.52%	93.49%	87.27%	0.83	0.92
Davis vs All	0.85	96.06%	90.89%	90.91%	96.93%	0.87	0.98
Green vs Coltrane	0.81	90.26%	89.93%	88.30%	92.89%	0.91	0.89
Green vs Parker	0.59	89.93%	72.40%	90.07%	88.46%	0.94	0.61
Green vs Davis	0.94	97.38%	96.91%	98.38%	94.85%	0.98	0.95
Coltrane vs Parker	0.57	88.26%	73.60%	89.27%	80.65%	0.93	0.62
Coltrane vs Davis	0.95	97.74%	98.08%	99.52%	94.06%	0.98	0.96
Parker vs Davis	0.71	86.99%	83.88%	86.05%	87.38%	0.80	0.90

n-way Class Results

	MCC	Accuracy	BA	PPV	NPV	$F_1{}^+$	$F_1{}^-$
Green	0.75	87.50%	87.75%	81.85%	92.01%	0.85	0.89
Coltrane	0.77	89.47%	88.25%	85.71%	91.46%	0.85	0.92
Parker	0.52	93.59%	71.92%	65.71%	95.29%	0.54	0.97
Davis	0.84	95.89%	91.63%	88.17%	97.28%	0.87	0.98

Bar – Sliding Window (4, 2) Level

Table D.13: Random forest bar (4, 2) level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.81	86.90%	86.94%	83.68%	95.53%	0.81	0.95
Green vs All	0.74	87.87%	85.44%	89.06%	87.33%	0.82	0.91
Coltrane vs All	0.83	91.79%	90.44%	93.99%	90.60%	0.89	0.93
Parker vs All	0.48	93.98%	65.90%	79.41%	94.47%	0.46	0.97
Davis vs All	0.74	93.31%	82.83%	89.15%	93.90%	0.77	0.96
Green vs Coltrane	0.84	92.05%	92.06%	91.26%	92.80%	0.92	0.92
Green vs Parker	0.66	91.01%	77.06%	91.11%	90.20%	0.95	0.69
Green vs Davis	0.78	90.63%	88.69%	92.78%	85.63%	0.93	0.85
Coltrane vs Parker	0.76	93.69%	84.20%	94.15%	90.63%	0.96	0.79
Coltrane vs Davis	0.85	93.96%	92.31%	95.16%	90.96%	0.96	0.90
Parker vs Davis	0.77	90.16%	88.04%	87.18%	91.48%	0.84	0.93

n-way Class Results

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.83	91.87%	91.88%	86.73%	95.15%	0.89	0.93
Coltrane	0.85	92.73%	92.54%	89.90%	94.60%	0.91	0.94
Parker	0.57	94.55%	73.37%	74.07%	95.67%	0.58	0.97
Davis	0.80	94.65%	89.98%	84.02%	96.69%	0.84	0.97

Bar – Sliding Window (2, 1) Level

Table D.14: Random forest bar (2, 1) level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.74	82.49%	83.66%	80.78%	93.87%	0.77	0.94
Green vs All	0.72	86.63%	84.60%	88.28%	85.82%	0.81	0.90
Coltrane vs All	0.82	91.68%	89.97%	94.15%	90.45%	0.88	0.94
Parker vs All	0.55	94.41%	66.85%	95.83%	94.37%	0.50	0.97
Davis vs All	0.72	93.32%	81.52%	88.52%	93.92%	0.75	0.96
Green vs Coltrane	0.81	90.46%	90.46%	90.47%	90.45%	0.91	0.90
Green vs Parker	0.64	90.57%	76.92%	91.24%	85.39%	0.94	0.68
Green vs Davis	0.79	91.45%	88.24%	92.71%	87.88%	0.94	0.84
Coltrane vs Parker	0.70	91.80%	80.65%	92.40%	87.76%	0.95	0.74
Coltrane vs Davis	0.88	95.06%	94.64%	97.37%	89.69%	0.97	0.92
Parker vs Davis	0.68	85.79%	82.47%	85.84%	85.77%	0.78	0.90

n-way Class Results

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.74	87.23%	87.53%	80.31%	92.44%	0.84	0.89
Coltrane	0.80	90.82%	90.12%	88.25%	92.34%	0.88	0.93
Parker	0.56	94.41%	71.53%	78.95%	95.16%	0.57	0.97
Davis	0.71	92.52%	85.46%	75.60%	95.56%	0.75	0.96

Note Level

Table D.15: Random forest note level summary performance metrics.

	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
n-way	0.58	72.84%	74.17%	64.74%	90.25%	0.62	0.90
Green vs All	0.68	84.57%	84.45%	79.75%	88.25%	0.82	0.87
Coltrane vs All	0.63	82.71%	80.95%	79.50%	84.46%	0.76	0.86
Davis vs All	0.55	91.58%	70.87%	79.03%	92.49%	0.56	0.95
Green vs Coltrane	0.72	85.91%	85.85%	85.68%	86.18%	0.87	0.85
Green vs Parker	0.68	91.59%	80.77%	93.11%	81.67%	0.95	0.72
Green vs Davis	0.71	90.18%	83.71%	91.84%	83.51%	0.94	0.77
Coltrane vs Parker	0.42	85.38%	66.41%	87.43%	66.67%	0.92	0.47
Coltrane vs Davis	0.75	90.89%	87.09%	93.47%	82.57%	0.94	0.81
Parker vs Davis	0.58	79.89%	78.66%	76.39%	82.05%	0.74	0.83
n-way Class Results							
	MCC	Accuracy	BA	PPV	NPV	F_1^+	F_1^-
Green	0.65	82.80%	82.75%	77.31%	87.11%	0.80	0.85
Coltrane	0.62	81.49%	81.46%	73.13%	87.64%	0.77	0.85
Parker	0.28	91.24%	60.53%	45.00%	93.36%	0.31	0.95
Davis	0.50	90.14%	71.96%	63.53%	92.87%	0.55	0.94

Appendix E

Data

The contents of the folders in the attached dataset of Appendix E are listed below.

E.1 Transcription Files

- *Sonic Visualiser* transcription files (audio files removed due to copyright)[1]
- PDF of automatically generated symbolic notation transcriptions
- Green's SQLite3 Database

E.2 Data Extraction and Setup Files

- bat and YAML files used with `MelConvert` to extract features from the database of Green's transcriptions and the WJazzD
- CSV files generated by `MelConvert` from the bat and YAML files
- CSV files of the `beats` table extracted from the SQLite3 databases
- Green_DataPreparation.R, *R* script for the initial data cleaning and processing for the analysis of Green's improvisational style (which formed part of the all_raw_df.Rds data)[2]
- Libraries.R and Functions.R, *R* scripts for loading the libraries and custom functions used throughout the research

[1] *Sonic Visualiser* requires an audio track to be present in each file. To allow for the SV transcription files to work empty tracks with the same length as the transcription are provided.

[2] Basic processing for WJazzD data used within the analysis of Green's improvisational style was completed in index.Rmd and when required in the chapter Rmd files.

E.3 Machine Learning Setup Files

- MachineLearning_DataPreparation.R, *R* script for initial data cleaning and processing for the performer classification and comparative analysis (result was the mlRaw.Rds data)
- *R* scripts for additional data cleaning and processing required for each abstraction level
- RDS files containing the cleaned and processed data for each abstraction level
- RMarkdown (Rmd) file containing the code used to train the models at each abstraction level

E.4 Machine Learning Models

- RDS file containing the final model for each of the 165 models trained[3]

E.5 Machine Learning Confusion Matrices

- *R* script containing a Shiny app for viewing all the testing data confusion matrices
- Windows program electron wrapper for the Shiny app
- PDFs containing the confusion matrices and statistics for all training and testing models

E.6 RDS Data Files

- all_raw_df.Rds contains the combined data for Green, Coltrane, Parker, and Davis. This data was split into two datasets (raw for Green and wjd for Coltrane, Parker, and Davis) in index.Rmd
- mlRaw.Rds contains the data used for the performer classification and comparative analysis
- wj_raw_df.Rds contains the slightly processed data from the entire WJazzD
- beatTrackGreen.Rds and contain beat track data extracted from Green's SQLite3 database
- confusionMatrixData.Rds contains the testing data confusion matrix performance metrics for each trained model

[3]For each classifier and abstraction level the one-vs-all and one-vs-one models were combined into single RDS files.

E.7 Thesis Files

- Rmd files containing the text of the thesis and accompanying code for generating the figures, tables, and analyses
- Supporting files for compiling the thesis or running code in the Rmd files (KnitLibraries.R, KnitFunctions.R, RDS data files, pre-rendered figures, and bibliography files)

E.8 Supplementary Files

- `randomSongSelector` program and CSV file for randomly selecting the improvisations
- PDF of all feature trend graphs
- `automateSV` program for generating the default *Sonic Visualiser* layers and checking for polyphony in the exported note layer