TOWARD A GENERAL THEORY OF POLYMETER:

POLYMETRIC POTENTIAL AND REALIZATION IN ELLIOTT CARTER'S SOLO AND CHAMBER INSTRUMENTAL WORKS AFTER 1980

by

ÈVE POUDRIER

A dissertation submitted to the Graduate Faculty in Music

in partial fulfilment of the requirements for the degree of Doctor of Philosophy,

The City University of New York

2008

UMI Number: 3340212

Copyright 2008 by
Poudrier, Eve

All rights reserved

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.
In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.



UMI Microform 3340212
Copyright 2009 by ProQuest LLC
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

© 2008

ÈVE POUDRIER

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Music in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

_____

Norman Carey

_____ _____

Date Chair of the Examining Committee

_____

David Olan

_____ _____

Date Executive Officer

Joseph N. Straus

_____

Stephen Blum

_____

John F. Link

_____

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

TOWARD A GENERAL THEORY OF POLYMETER:
POLYMETRIC POTENTIAL AND REALIZATION IN ELLIOTT CARTER'S
SOLO AND CHAMBER INSTRUMENTAL WORKS AFTER 1980

by

ÈVE POUDRIER

Advisor: Professor Joseph N. Straus

This dissertation proposes a conceptual framework and tools for the analysis of polymeter in contemporary musical practices applied to the short instrumental works of Elliott Carter, where multiple temporalities have been identified as an essential aesthetic component. Three works are discussed in more detail: *Changes* for guitar (1983), *Esprit rude/Esprit doux I* for flute and clarinet (1985), and *90+* for piano (1994). The composer's use of metrical modulations and structural polyrhythms is re-evaluated in the context of a theory of polymeter. Structural polyrhythms are represented by means of proximity charts, and Carter's compositional strategies are explored, including rhetorical use of background pulsations on the musical surface.

In contrast to current theories of meter, this study does not assume that hierarchical rhythmic structures necessarily imply recursiveness or rely on accentual structure. In Carter's music, *polymetric potential* stems from the systematic partitioning of pulses of different speeds. Textural streaming is contrasted with multileveled polyrhythmic design, and

polymetric structures are defined as contrasting rhythmic strands that are built on competing pulse streams or non-isochronous beat structures and give rise to non-synchronous metrical projections. In a *balanced* polymetric structure, the rhythmic strands are nearly equal in metrical strength; *polarized* structures either involve shifting dominance or one pulse stream that is dominant while the other(s) provide periodic syncopation; when the two streams are fused, the polymetric structure is *integrated*. The realization of polymetric potential is an interactive process involving the composer, performer(s), and the listener. From a music-analytical perspective, polymetric realization can be assessed by an investigation of the metrical properties of each pulse stream as well as the interaction of pulse streams at different structural levels (*foreground*, *middleground*, *background*).

The first part introduces various analytical tools that measure the relative mensural determinacy of competing rhythmic strands, including pulse stream reductions and density and nestedness analyses. Issues of perception and performance are briefly discussed in relation to specific interpretations. The second part explores the topic of polymeter from a psychological perspective. The results of a listening experiment based on an example from *90+* for piano are presented and possible avenues for future research are discussed.

ACKNOWLEDGEMENTS

Throughout my studies at The Graduate Center and during my doctoral work in the Music Department, I received encouragement and support from many people. I have found this place to be filled with interesting, caring, and stimulating individuals. Upon completing this dissertation, my first thought goes to all my colleagues, professors, and the staff who have made my stay so enriching and productive. In particular, I would like to thank my advisor, Joe Straus, for his outstanding mentoring in so many different areas of my professional life; my first reader, Stephen Blum, for his patience and attention in reading through the text as well as for many useful comments in making this dissertation more reader-friendly; and the dissertation committee members Norman Carey and John Link for sharing their expertise. Finally, I am much grateful to the wonderful professors who have contributed to my development as a musician and theorist: Ruth DeFord, Poundie Burstein, David Gagné, Allan Atlas, Anne Stone, Shaugn O'Donnell, and Philip Rupprecht, in whose seminar I was introduced to the music of Elliott Carter. Special thanks go to William Rothstein and Stephen Blum for their insights into the fascinating topic of rhythm and meter.

This dissertation involves elements of experimental psychology and statistics. In preparing the materials for the experiment I conducted, I was fortunate to receive a grant from the Graduate Research Grants Program at The Graduate Center (2005-2006). I am especially indebted to Paul Riker for his technical assistance in sound engineering and programming. Many people have also guided my learning of statistical methods and exploration of the field of music psychology. In the early stages of this work, I was fortunate to receive mentoring from Prahba Betne from the Mathematics Department at LaGuardia Community College and to participate in David's Rindskopf's seminar on statistical methods

in the Educational Psychology Program at The Graduate Center; Rita Aiello from the Psychology Department at New York University welcomed me in her seminar on Emotion and Music; Bruno Repp at the Haskins Laboratory reviewed an early draft of the results of the listening experiment and pointed out various areas that needed improvement; and Tina Zottoli was especially helpful in guiding me through the steps of the statistical analysis. Finally, I want to express my gratitude to Panayotis Mavromatis at New York University's Steinhardt School of Culture, Education, and Human Development for his enthusiastic interest in the cognition aspect of this project and for generously giving many hours of his time for advising and help with data processing issues.

I have received additional financial support for various aspects of this work: a second grant from The Graduate Research Grants Program (2007-2008) allowed me to go study Carter's sketches at the Paul Sacher Foundation in Basel, Switzerland. I am thankful for John Link's advise and tips in preparing for this trip and to Felix Meyer, Henrike Hoffmann, and other staff at the Foundation for helping me while I as there. The work presented in this dissertation was also supported by an Elebash Dissertation Award for research on music in New York. This grant has provided additional funds for my stay at the Paul Sacher Foundation, purchasing the recordings and scores I needed, and various expenses related to interviews, travel, and technical assistance.

Excerpts from *90+*, *Changes*, and *Esprit rude / Esprit doux I* are reprinted by permission of Hendon Music, Inc. Transcriptions of sketches from the Elliott Carter Collection are reproduced by permission of the Paul Sacher Foundation, Basel.

My final thoughts go to my friends and family for their understanding, kindness and encouragements, and to my husband, Rémi Castonguay, for his unflinching love and support.

To my mother,

for her enthusiasm, courage,

and deep confidence in all of my undertakings.

TABLE OF CONTENTS

List of Illustrations *xiv*

Abbreviations *xxvi*

CHAPTER 1 **Introduction: Elliott Carter and Multiple Temporalities** ***1***

1.1 Composing with Simultaneous Speeds *1*

1.1.1 Structural Polyrhythms and Metrical Modulation 5

1.1.2 Character-Patterns and Referential Pitch Collections 15

1.2 Purpose of Study *19*

1.2.1 Interdisciplinary Considerations 21

1.2.2 Repertoire 23

1.2.3 Dissertation Plan 29

CHAPTER 2 **Structural Levels and the Mapping of Polyrhythmic Organization** ***30***

2.1 Preliminary Remarks *30*

2.1.1 Defining Meter 32

2.1.2 Carter's Notational Practice 35

2.2 Structural Levels *38*

2.2.1 Background 39

2.2.2 Foreground 45

2.2.3 Middleground 47

CHAPTER 2 **2.3** Compositional Strategies and Interaction of Structural Levels *51*

2.3.1 Metrical Modulation Re-visited 52

2.3.1.1 Background Continuity 54

2.3.1.2 Surface Discontinuities 62

2.3.1.3 Perceptual Issues 67

2.3.2 Long-Range Polyrhythms in Action 70

2.3.2.1 Background Irregularities: Polycyclic Spectrum and Proximity Charts 76

2.3.2.2 Analytical Details 81

CHAPTER 3 **Polymetric Potential and Realization** ***92***

3.1 Streaming and Multileveled Polyrhythmic Design: *Esprit rude/Esprit doux* I (1985) *93*

3.1.1 Polymetric Structures Defined 104

3.1.2 Abstract Models 107

3.2 Local Conditions: Mensural Determinacy and Pulse Stream Reductions *113*

3.2.1 Middleground Level: Quarter Pulse Stream 114

3.2.2 Foreground Level: Elementary Pulse Streams 118

3.3 Global Conditions: Density and Nestedness Analyses *129*

3.3.1 Case Study No. 1: The Opening of Mozart's Symphony in G minor, K. 550 131

CHAPTER 3 **3.3.2** Case Study No. 2: *Changes* for Guitar (1983) 139

3.3.2.1 Musical Surface and Pulse Activity: Sectional Density Analyses 140

3.3.2.2 Polarity and Integration at Work: Nestedness Analyses 146

3.4 Polymetric Potential and Realization in *90+* for Piano (1994) *154*

3.4.1 Integrated Polymeter 155

3.4.2 Polarized Polymeter 159

3.4.3 Balanced Polymeter 174

CHAPTER 4 **The Realization of Polymetric Potential from a Psychological Perspective** ***182***

4.1 Methodological Issues *185*

4.1.1 Previous Research 187

4.1.2 Experimental Design 191

4.1.2.1 Participants 192

4.1.2.2 Source Materials 194

4.1.2.3 Stimuli 195

4.1.2.4 Apparatus and Experimental Procedure 200

4.1.2.5 Data Collection and Measures 202

4.1.2.6 Data Analysis 205

CHAPTER 4 **4.2** Results *207*

4.1.1 Overall Tapping Consistency and Tapping Contour Graphs 207

4.22 Block 1 215

4.2.2.1 Preliminary Analysis: Order of Presentation and Phase Discrimination 215

4.2.2.2 Main Analysis: The Effect of Layer and Accent 219

4.2.2.3 Exploratory Analyses: Selected Between-Subjects Factors 226

4.2.3 Block 2 234

4.2.3.1 Preliminary Analysis: Order of Presentation and Phase Discrimination 234

4.2.3.2 Main Analysis: The Effect of Layer and Articulation 238

4.2.3.3 Exploratory Analyses: Selected Between-Subjects Factors 243

4.2.3.4 Across-Blocks Analysis: Blocks 1&2 (Timbre) 259

4.2.4 Block 3 264

4.2.4.1 Preliminary Analysis: Order of Presentation and Phase Discrimination 265

4.2.4.2 Main Analysis: Layer-Specific Accentuation (Laccent and Raccent) 269

4.2.4.3 Exploratory Analyses: Selected Between-Subjects Factors 276

CHAPTER 4 **4.2.4.4** Across-Blocks Analysis: Blocks 1&3 (Version) 281

4.2.5 Block 4 290

4.2.5.1 Preliminary Analysis: Order of Presentation and Phase Discrimination 292

4.2.5.2 Main Analysis: The Effect of Attention and Condition 297

4.2.5.3 Exploratory Analyses: Selected Between-Subjects Factors 307

4.2.5.4 Across-Blocks Analyses: Blocks 2&4 (Version) and 3&4 (Timbre) 313

4.3 Final Discussion: Assessing Polymetric Realization *325*

CHAPTER 5 **Concluding Remarks: Toward a General Theory of Polymeter** ***331***

Bibliography *337*

I. General Theoretical and Analytical Approaches 337

II. Psychological Studies 340

III. Carter: Scores, Writings, Interviews, and Research Guides 344

IV. Analytical Studies of Carter's Music 345

V. Audio and Video Recordings 349

LIST OF ILLUSTRATIONS

CHAPTER 1 *1*

Example 1.1: Out-of-phase polyrhythm with a speed ratio of 23:21:18, *90+* for piano (1994), mm. 1-3 3

1.2: Transcription of Carter's polyrhythmic sketch for *Esprit rude/Esprit doux I* for flute and clarinet (1985) 7

1.3: Partial transcription of a page from Carter's sketches for the Oboe Quartet (2001) 9

1.4: Transcription of Carter's rhythmic plan for *90+* 11

Table 1.1: Structural chart with tempi, pulse streams, notational equivalences, and resultant speeds 13

1.2: Beat subdivisions and resultant speeds for *90+* 15

1.3: Elliott Carter's solo instrumental works 25

1.4: Evidence of structural polyrhythms in Elliott Carter's solo and chamber instrumental works after 1980 27

CHAPTER 2 *30*

Example 2.1: Interaction of background pulse and foreground rhythmic activity, *90+*, pulse-measures 39-41 (mm. 51-53) 43

2.2: Mapping of surface events based on the two elementary pulse streams, *Changes* for guitar (1983), mm. 1-11 44

2.3: Metrical modulation from T_1 to T_2, *Changes*, mm. 108-111 56

2.4: Metrical modulation from T_2 to T_1, *Changes*, mm. 112-115 56

Example **2.5:** Metrical modulation from T_1 to T_3, *Changes*, mm. 128-133 57

2.6: Expressive tempo fluctuation, *Changes*, mm. 122-125 58

2.7: Pivot pulse-measures O46 and X62, *Changes*, mm. 128-131 60

2.8: Written-out "grace" note, *Changes*, m. 17 63

2.9: Surface rhythmic activity in Epilogue, *Changes*, mm. 134-141 65

2.10: Annotated score, Epilogue, *Changes*, mm. 131-150 75

2.11: Displacement of background pulsation X10 and integration in competing elementary pulse stream, *Changes*, mm. 17-20 82

2.12: Displacement of background pulsation X18, *Changes*, mm. 34-38 82

2.13: Displacement of background pulsation X37 and integration in competing elementary pulse stream, *Changes*, mm. 75-79 83

2.14: Near-coincidence X21→O16 articulates the phrase structure, *Changes*, mm. 42-43 84

2.15: Near-coincidence X29→O22 participates in a contrapuntal duet, *Changes* for guitar, mm. 58-60 84

2.16: Interaction between background pulsation and formal structure, X43→O33, *Changes*, mm. 88-91 85

2.17: Annotated score, *Changes*, mm. 102-115 88

2.18: Displacement of background pulsation X58 and integration in competing elementary pulse stream, *Changes*, mm. 119-123 89

Figure **2.1:** Formula for the resultant speed of a background pulse stream 51

2.2: Final convergence phase, *Changes*, O48→X64 to X76/O57 (mm. 133-150) 74

2.3: Abstract polycyle, *Changes*, mm. 1-158 78

2.4: Realized polycyle, *Changes*, mm. 1-150 78

2.5: Abstract polycycle, *Changes*, mm. 1-102 80

2.6: Realized polycycle, *Changes*, mm. 1-102 80

2.7: Abstract polycycle, *Changes*, mm. 103-129 86

2.8: Realized polycycle, *Changes*, mm. 103-130 86

2.9: Abstract polycycle, *Changes*, mm. 130-158 90

2.10: Realized polycycle, *Changes*, mm. 131-150 90

Table **2.1:** Properties of the long-range polyrhythm for *Changes* 48

2.2: Notational equivalences in *Changes* 49

2.3: Formal Chart of *Changes* 55

2.4: "X" stream notational equivalences and resultant metronomic speeds 59

2.5: "O" stream notational equivalences and resultant metronomic speeds 59

2.6: Tempo Field of *Changes* 63

CHAPTER 3 *92*

Example **3.1:** Opening, *Esprit rude/Esprit doux I* for flute and clarinet (1985), mm. 1-5 94

3.2: Transcription and metrical interpretation of *Esprit rude/Esprit doux I*, mm. 1-5 94

3.3: Multileveled polyrhythmic design, *Esprit rude/Esprit doux I*, mm. 6-9 97

3.4: Stream segregation, *Esprit rude/Esprit doux* I, mm. 10-17 99

3.5: Groupings leading to stream segregation, *Esprit rude/Esprit doux* I, mm. 1-5 100

3.6: Background anacrusis, *Esprit rude/Esprit doux* I, mm. 1-5 102

3.7: Pulse stream reduction, middleground level, *Changes* 112

3.8: Musical surface, *Changes*, bar-measures 8-9 115

3.9: Musical surface, *Changes*, bar-measures 21-26 115

3.10: Musical surface, *Changes*, bar-measures 1-7 117

3.11: Pulse stream reduction, "X" stream, *Changes* 120

3.12: Pulse stream reduction, "O" stream, *Changes* 122

3.13: Musical surface, *Changes*, bar-measures 36-41 127

3.14: Musical surface, *Changes*, bar-measures 25-26 127

3.15: Musical surface, *Changes*, bar-measures 122-125 129

3.16: Mozart Symphony in G minor, K, 550, I, mm. 1-2 (melody only) 131

3.17: Mozart, K. 550, I, mm. 1-21 (full texture) 135

Example **3.18:** Integrated polymetric structure, *90+*, pulse-measures 41-43 (mm. 54-56) 157

3.19: Structural polyrhythm, *90+*, pulse-measures 1-11 (mm. 1-14) 161

3.20: Polarized polymetric structure with nestedness details, *90+*, mm. 1-8 166

3.21: Balanced polymetric structure, *90+*, pulse-measures 56-58 (mm. 69-71) 175

3.22: Background pulsations, *90+*, pulse-measures 55-69 (mm. 68-82) 176

3.23: Eighth- and sixteenth-level reductions of left-hand accompaniment, *90+*, pulse-measures 56-69 (mm. 69-83) 179

3.24: Quarter- and eighth-level reductions of right-hand melody, *90+*, mm. 68-82 180

Figure **3.1:** Details from a 4:3 structural polyrhythm 108

3.2: Polarity and integration variance, phrase level 150

3.3: Proximity chart, *90+*, all structural pulse streams, bar-measures 0-14 162

3.4: Contour graph of the structural polyrhythm as performed by Charles Rosen, *90+*, mm. 0-14 168

3.5: Contour graph of the structural polyrhythm as performed by Ed Spanjaard, *90+*, mm. 0-14 168

3.6: Cyclic spectrum for the 90+ pulse 171

3.7: Cyclic spectrum for the Chord pulse 171

Figure **3.8:** Polycyclic spectrum, *90+*, mm. 0-14 173

3.9: Contour graph of the right-hand melody, *90+*, mm. 71-76 181

Table **3.1:** Simple polymetric models for in-phase structural polyrhythms 106

3.2: Synchronicity as a measure of integration 110

3.3: Isochronicity as a measure of mensural determinacy 110

3.4: Elementary pulse streams' notational equivalences for pulse stream reductions 124

3.5: Notational equivalences in elementary pulse units, Mozart, K. 550, I 133

3.6: Density and nestedness data for Mozart, K. 550, I, mm. 1-2 133

3.7: Density and nestedness analyses data of Mozart, K. 550, I, mm. 1-20, downbeat 138

3.8: Sectional density analysis data for *Changes*, middleground level 141

3.9: Sectional pulse density analysis data for *Changes*, foreground level 144

3.10: Nestedness analysis data for *Changes*, section level 148

3.11: Nestedness analysis data for *Changes*, phrase level 151

3.12: Density and nestedness analysis data, pulse-measures 41-43 (mm. 54.2.2–56.2.2) 158

3.13: Density and nestedness analysis data, pulse-measures 1-6 (mm. 0-8) 163

CHAPTER 4 *182*

Figure **4.1:** Participants' experience with twentieth-century music 208

4.2: Correlation of tapping consistency and musical experience 209

4.3: Tapping contour graph of Participant 30 for Block 1 stimuli 211

4.4: Tapping contour graph of Participant 32 for Block 1 stimuli 213

4.5: Layer × Accent interaction for the coefficient of variability in Block 1 224

4.6: Layer × Degree.2grp interaction for upbeat tapping synchronization in Block 1 227

4.7: Main effect of area of specialization on tapping consistency in Block 1 229

4.8: Accent × Piano.3grp interaction for tapping consistency in Block 1 230

4.9: Accent × Gender interaction for overall downbeat synchronization in Block 1 232

4.10: Main effect of area of specialization on tapping consistency in Block 2 246

4.11: Layer × Specialization.4grp interaction for overall downbeat synchronization in Block 2 248

4.12: Layer × Instrument.3grp interaction for tapping consistency in Block 2 250

Figure **4.13:** Articulation × Instrument.3grp interaction for upbeat tapping synchronization in Block 2 252

4.14: Layer × Gender interaction for upbeat tapping synchronization in Block 2 253

4.15: Articulation × Gender interaction for upbeat tapping synchronization in Block 2 255

4.16: Accent × Degree.2grp interaction for downbeat tapping synchronization in Block 2 257

4.17: Accent × Degree.2grp interaction for overall tapping synchronization in Block 2 258

4.18: Layer × Timbre interaction for upbeat tapping synchronization across Blocks 1 and 2 262

4.19: Layer × Timbre interaction for overall upbeat synchronization across Blocks 1 and 2 263

4.20: Laccent × Raccent interaction for tapping and beat synchronization to the right hand downbeat in Block 3 275

4.21: Raccent × Degree.2grp interaction for tapping consistency in Block 3 277

4.22: Main effect of area of specialization on overall beat synchronization to the right hand phase in Block 3 279

4.23: Laccent × Gender interaction for overall synchronization with the left hand phase in Block 3 280

Figure **4.24:** Main effect of condition on tapping consistency in Block 4, with significant mean difference between conditions 1 and 4 301

4.25: Attention × Condition interaction for downbeat tapping synchronization in Block 4 303

4.26: Significant main effect of condition on upbeat tapping synchronization in Block 4 304

4.27: Attention × Condition interaction for overall upbeat synchronization in Block 4 306

4.28: Significant main effect of area of specialization on tapping consistency in Block 4 310

4.29: Significant main effect of area of specialization on overall downbeat synchronization in Block 4 311

4.30: Interaction between area of specialization and condition for downbeat tapping synchronization in Block 4 313

4.31: Tapping contour graph of Participant 8 for Stimulus 9 in Block 4 326

Table **4.1:** Stimuli description with factors and conditions 199

4.2: Estimated marginal means and mean differences for downbeat and upbeat synchronization based on order of presentation in Block 1 216

4.3: Paired-samples t tests resuls for tapping synchronization to the downbeat and upbeat phases in Block 1 218

4.4: Descriptive statistics for Block 1 220

Table

4.5: Effect of Layer in Block 1 222

4.6: Effect of Accent in Block 1 222

4.7: Estimated marginal means and mean differences for upbeat synchronization based on order of presentation in Block 2 235

4.8: Paired samples t-tests results for downbeat and upbeat beat synchronization in Block 2 237

4.9: Effect of Layer on all dependent variables in Block 2 239

4.10: Paired-samples t tests results based on streaming in the left hand layer in Block 2 241

4.11: Paired samples t-tests results based on accentuation in the right hand layer in Block 2 241

4.12: Estimated marginal means and mean differences for all dependent variables based on musical degree in Block 2 245

4.13: Order of presentation in Block 3 266

4.14: Interference levels based on accentuation and streaming 267

4.15: Paired-samples t tests results for tapping synchronization to left hand and right hand downbeat in Block 3 269

4.16: Estimated means and mean differences based on left hand accents 270

4.17: Estimated means and mean differences based on right hand accents 272

4.18: Phase dominance and metrical profiles in Blocks 1 and 3 282

Table **4.19:** Superposition operation, interference levels, and resulting metrical profiles 284

4.20: Paired-samples t tests results for downbeat tapping synchronization across Blocks 1 and 3 288

4.21: Paired-samples t tests results for overall downbeat synchronization across Blocks 1 and 3 288

4.22: Paired-samples t tests results for upbeat tapping synchronization across Blocks 1 and 3 289

4.23: Paired-samples t tests results for overall upbeat synchronization across Blocks 1 and 3 289

4.24: Summary description of Block 4 stimuli 291

4.25: Paired-samples t tests results for phase differentiation in tapping and overall beat synchronization by condition for attended left hand layer in Block 4 293

4.26: Paired-samples t tests results for phase differentiation in tapping and overall beat synchronization by condition for right hand layer in Block 4 294

4.27: Estimated marginal means of each condition for all synchronization measures across attended layers in Block 4 296

4.28: Estimated means and mean differences based on attended layer in Block 4 298

4.29: Paired-samples t tests results for the effect of version on tapping consistency across Blocks 2 and 4 314

Table **4.30:** Paired-samples t tests results for the effect of version on tapping and beat synchronization across Blocks 2 and 4 317

4.31: Paired-samples t tests results for the effect of timbre on tapping 317

4.32: Paired-samples t tests results for the effect of timbre on downbeat synchronization across Blocks 3 and 4 322

4.33: Paired-samples t tests results for the effect of timbre on upbeat synchronization across Blocks 3 and 4 323

CHAPTER 5 *330*

Excerpts from *90+*: pp. 3, 43, 157, 161, 166, 175, 176-177, 179, and 180.

© 1994 HENDON MUSIC, INC.

Excerpts from *Changes*: pp. 44, 56, 57, 58, 60, 63, 65, 75, 82, 83, 84, 85, 88, 89, 115, 117, 127, and 128.

© 1983 HENDON MUSIC, INC.

Excerpts from *Esprit rude / Esprit doux I*: pp. 94, 97, 99, and 102.

© 1985 HENDON MUSIC, INC.

LIST OF ABBREVIATIONS

A	Articulated time-points total
bpm	Beats per minute
BPU/bpu	Background pulse unit
Bsyn/Tsyn	Beat/tapping synchronization
CV	Coefficient of variability
EPU/epu	Elementary pulse unit
I/N	Independent/nested events
IOI	Inter-onset interval
ITI	Inter-tap interval
LH/RH	Left/right hand layer
LRP	John Link's "Long-Range Polyrhythms"
MEC	David Schiff's *The Music of Elliott Carter*
MM	Metronomic speed
MPU/mpu	Middleground pulse unit
P	Pulsation total of a given stream
PSS	Holdings at the *Paul Sacher Stiftung*
QD/QU	Quarter downbeat/upbeat
QL/R	Left/right hand phase quarter beat
SD/SE	Standard deviation/error
SPU/spu	Subpulse unit
T_S	Total of synchronous events

CHAPTER 1

Introduction: Elliott Carter and Multiple Temporalities

There has been much experimentation with various types of rhythmic organization in the music of the twentieth century, often leading to complex rhythmic textures. In particular temporal stratification of the kind provided by the use of *polymeter* as structural device, that is, *the layering of rhythmic strands based on two or more concurrent metric frameworks*, emerged as an organizational strategy of choice in the practice of modernist composers such as Igor Stravinsky, Charles Ives, Conlon Nancarrow, and György Ligeti. Nevertheless, despite its widespread use, this technique is still under-theorized and to date there is still no general theory of polymeter. *This dissertation proposes a conceptual framework for the study of polymetric structures applied to music by Elliott Carter*, where multiple temporalities have been identified as an essential element of the musical discourse. For the time being, a *polymetric structure* can be loosely defined as a musical texture in which rhythmic components can be traced back to at least two simultaneous meters, each of which involves at least two beat levels. While polymetric textures are necessarily polyrhythmic, *polyrhythm* is more specifically understood as the superposition of two or more rhythmic sequences in which all events are perceived as evenly spaced (e.g., 2 against 3). These terms will be defined more formally at a later point.

1.1 Composing with Simultaneous Speeds

In the music of the American composer Elliott Carter, rhythmic layering is associated with the use of simultaneous speeds that are generated by the subdivision of the temporal

fabric into a grid with two or more *pulse streams*, that is, series of equally spaced time-points also known as pulses or pulsations.[1] In this dissertation, the term *pulse* will be used to refer to the perceptible regularity of such a succession, while *pulsation* will designate a specific instance within a pulse stream. Given that each pulse stream can be associated with a specific *speed*, i.e., the rate of motion associated with the periodic return of a pulsation, the term *tempo* will be used here in relation to the specific pulse level of the *beat*, the structural timespan unit conventionally associated with the time signature (as representative of meter).[2] Thus, speed is understood as a product (or function) of pulse and tempo as an emerging property of meter. The opening measures of *90+*, a short piece for solo piano composed in 1994 as a birthday offering for the Italian composer Goffredo Petrassi, will serve to illustrate Carter's compositional strategies in dealing with multiple temporalities (see **example 1.1**).

---

[1] For a representative analysis focusing on Carter's use of simultaneous speeds, see Brenda Ravenscroft, "Setting the Pace: the Role of Speeds in Elliott Carter's *A Mirror on Which to Dwell*," *Music Analysis*, 22/3 (2003): 253–82. John Roeder introduces the concept of "pulse stream" to refer to distinct continuities of "perceptibly equal timespans, marked off by accented timepoints"; see John Roeder, "Interacting Pulse Streams in Schoenberg's Atonal Polyphony," *Music Theory Spectrum* 16/2 (1994): 234. My use of the term is consistent with Roeder's except for the latter's reliance on the concept of "accent" (unless time-points are conceptualized as accented only to the extent that they are "marked for consciousness in some way"). The term "stream" is related to the concept of "auditory streaming," i.e., the partitioning of auditory events based on aspects of pitch, dynamics, timbre, and tempo, which is explored by Albert S. Bregman in his seminal work *Auditory Scene Analysis: the Perceptual Organization of Sound* (Cambridge, MA: MIT Press, 1990); for a theory of texture based on these concepts, see Brenda Ravenscroft, "Texture in Elliott Carter's *A Mirror on Which to Dwell*" (Ph.D. dissertation, University of British Columbia, 1992).

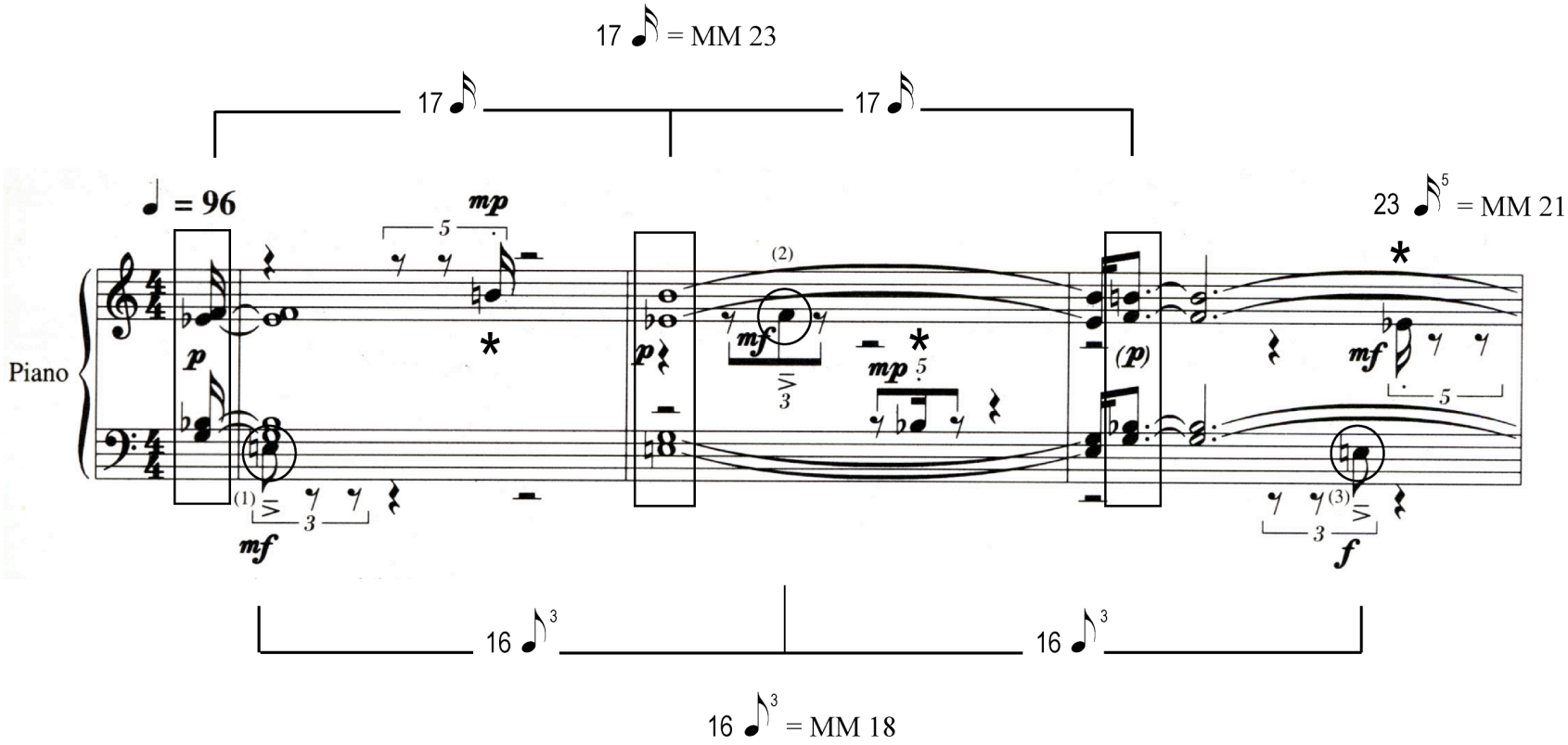
[2] The term *tactus* will be discussed later in relationship with psychological theories of meter (see chapter 4, section 4.1).

Example 1.1: Out-of-phase polyrhythm with a speed ratio of 23:21:18, *90+* for piano (1994), mm. 1-3

◯ = 90+ pulse

▯ = Chord pulse

* = Rogue pulse



As described by Carter in his program note, the piece is built around ninety slow pulsations that travel through a series of contrasting textures.[3] I will refer to this explicit structural pulse stream as the "90+ pulse" and each pulsation belonging to it will be designated as a "90+ pulsation." In **example 1.1**, each 90+ pulsation is expressed as a single triplet eighth note with a tenuto accent. The inter-onset interval (IOI) between each pulsation corresponds to 16 triplet eighths.[4] At the notated tempo of quarter = 96, this pulse gives rise to a speed of 18 pulsations per minute, or MM 18.[5] A second pulse stream, the "Chord pulse," is generated by a series of legato chords, each of which is separated by 17 sixteenths, yielding a speed of about MM 23.[6] The remaining pitch events, a series of staccato quintuplet sixteenth notes, are the product of a third intervening pulse stream, which I will refer to as the "Rogue pulse," with an associated timespan of 23 quintuplet sixteenths and a resultant speed of MM 21. The overall texture can thus be described as an *out-of-phase polyrhythm*,

---

[3] Elliott Carter, *90+ for piano* (New York: Hendon Music, Boosey & Hawkes, 1994). The program note states:

> '90+' for piano is built around ninety short, accented notes played in a slow regular beat. Against these the context changes character continually. It was composed in March 1994 to celebrate the ninetieth birthday of my dear and much admired friend, Goffredo Petrassi, Italy's leading composer of his generation. Its first performance was given by Giuseppe Scotese on June 11, 1994, at the Pontino Music Festival dedicated to Petrassi's birthday.

[4] In all annotated examples, figures, and tables, timespans will be given using musical notation; the use of a superscript Arabic numeral to denote subdivision is a standard format in the Carter literature.

[5] The resultant speed of a given pulse stream is calculated as follows: notated tempo multiplied by the number of subdivisions per notated beat divided by the number of subdivision for each pulsation. In this case, 96 × 3/16 = 18. In Carter's sketches, this formula is reversed, e.g., 3/16 × 96.

[6] More precisely, the resultant speed is MM 22 10/17 (96 × 4/17). For the sake of simplicity, all resultant speeds will be rounded up to the closest integer.

which is to say, overlapping pulse streams in which the number of pulsation in each stream is related by a non-integer ratio to the others, with a speed ratio of 23:21:18.[7]

1.1.1 Structural Polyrhythms and Metrical Modulation

Carter's aesthetic preference for a musical discourse that exhibits coherence and underlying continuity leads him to make use of *giant polyrhythms*, that is, slow structural polyrhythms such as the 216:175 polyrhythmic cycle that encompasses the entire timespan of *Night Fantasies* for piano (1980), to the extent that "in his later music the calculation of a structural polyrhythm on graph paper is usually the first step in composition, a clear indication, if one were needed, that for Carter rhythm always comes first, albeit in a highly abstract form."[8] In *Night Fantasies*, the two structural pulse streams progress with a speed of MM 10.8 and MM 8.75, respectively, and the polyrhythmic cycle has a total duration of 20 minutes.[9] On these graphs, Carter typically assigns a different color to each instrumental part or group.[10]

---

[7] Out-of-phase polyrhythms are discussed by John F. Link, in "Long-Range Polyrhythms in Elliott Carter's Recent Works" (Ph. D. dissertation, City University of New York, 1994), pp. 26-27.

[8] David Schiff, *The Music of Elliott Carter*, 2nd ed. (Ithaca: Cornell University Press, 1998), p. 46. The use of "giant polyrhythms" is discussed by Carter as early as 1968–70; see Allen Edwards, *Flawed Words and Stubborn Sounds: a Conversation with Elliott Carter* (New York: W. W. Norton, 1972), pp. 111–12. For a detailed discussion of the development of Carter's harmonic vocabulary as it relates to the composition of *Night Fantasies*, see John F. Link, "The Composition of Elliott Carter's *Night Fantasies*," *Sonus* 14/2 (1994): 67-89. In this dissertation, I will use the term "long-range polyrhythm" as formalized by Link.

[9] Link, "Long-Range Polyrhythms," p. 117.

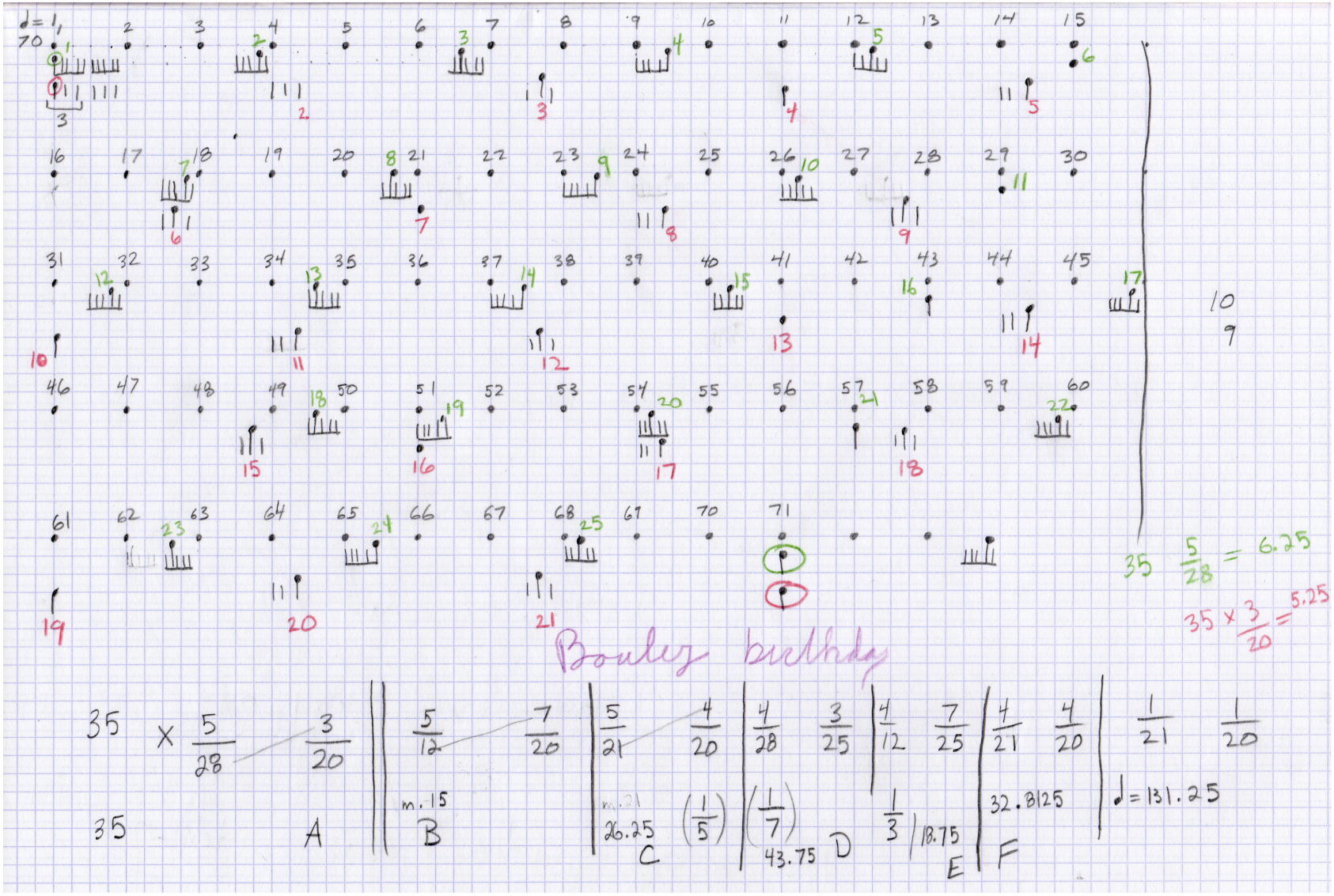
[10] I am indebted to John Link for sharing his experience of studying Carter's sketches and pointing out this particular aspect prior to my stay at the Paul Sacher Foundation in Basel, Switzerland in March 2008. The following articles also provide interesting insights into

Example 1.2 presents a transcription by the author of Carter's sketch for the long-range polyrhythm that serves as background rhythmic structure for *Esprit rude/Esprit doux I* for flute and clarinet (1985), which is based on a long-range polyrhythm of 25:21.[11] As can be observed on this sketch, Carter divides the temporal fabric into two contrasting pulse streams, each of which is associated with a specific subdivision. Carter also provides the specific location of the recurring background pulsations. Here, the flute occupies the lower stream and is associated with triplet quarter notes and the color red. The clarinet is in the upper stream and is associated with quintuplet eighth notes and the color green. As with many of the works from this period, the piece ends with the return of the coincidence point at pulsations 26 and 22 for the clarinet and flute. The graph also provides the grid that coordinates the two parts: each square represents a half-note beat in the final score, and Carter initially sets the tempo at half note = MM 35. Although the length of the piece in measures appears to be pre-determined, Carter often modifies the way the rhythm is notated at the surface, especially by using more or less frequent metrical modulations (to be discussed below).

Carter's compositional process: David Schiff, "A Paper Mountain: Elliott Carter's Sketches," in *Settling New Scores: Music Manuscripts from the Paul Sacher Foundation (*Switzerland: Schott Mainz, Germany, and Paul Sacher Stiftung, 1998), pp. 115-118; Jonathan W. Bernard, "Elliott Carter: Concerto for orchestra (1965-69)," in *Settling New Scores: Music Manuscripts from the Paul Sacher Foundation* (Switzerland: Schott Mainz, Germany, and Paul Sacher Stiftung, 1998), pp. 121-126; and Denis Vermaelen, "Elliott Carter's Sketches: Spiritual Exercises and Craftmanship," in *A Handbook to 20th-Century Musical Sketches*, eds. Patricia Hall and Friedemann Sallis (New York: Cambridge University Press, 2004), pp. 161-175.

[11] Elliott Carter Collection, Paul Sacher Foundation, Basel (accessed in March 2008). For a published version of the score, see Elliott Carter, *Esprit rude/Esprit doux I for Flute and Clarinet* (New York: Hendon Music, Boosey & Hawkes, 1985).

Example 1.2: Transcription of Carter's polyrhythmic sketch for *Esprit rude/Esprit doux I* for flute and clarinet (1985)



Long-range polyrhythms are not the only structural polyrhythms used by Carter. In many of his works, the composer uses contrasting pulse streams or implied polyrhythms as underlying grids for passages varying in scope from a phrase to an entire section.[12] For example, in *Night Fantasies* Link identifies two "local polyrhythms" of 81:80 and 16:15, which are superimposed onto the 216:175 long-range polyrhythm.[13] In other cases, Carter appears to subdivide slower pulsations into a series of faster isochronous pulsations.[14] These reveal the fluid boundary that exists in Carter's compositional thinking between underlying regularity and surface irregularities. Thus, while in many of his sketches he painstakingly lays out the pulse skeleton that gives rise to the speed of a specific instrumental part, the realization of this part into an actual rhythmic sequence is rarely regular enough for this speed to be perceived for more than a fleeting moment. Nonetheless, the importance of background pulsations as well as more local periodicities for the expression of each instrument's voice (or group of instrument) is undeniable.[15] Carter's colored marking of the location of background pulsations and his sketching of the beat subdivisions that articulate a specific speed as he works out the unfolding of pitch events in time serves as witness to the importance of this multileveled polyrhythmic grid in the building of the musical work.

---

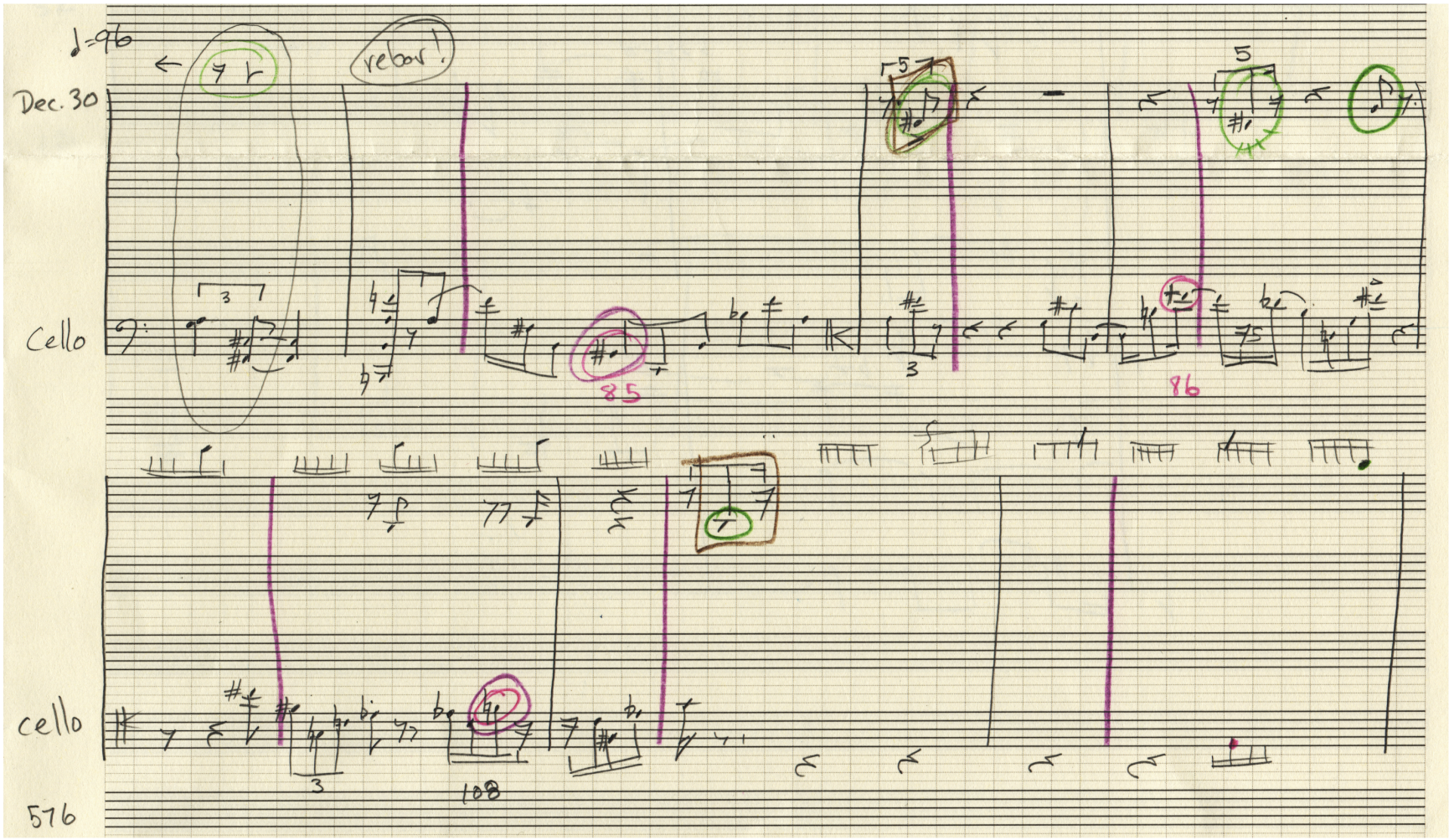
[12] In his book, Schiff appears to limit his definition of structural polyrhythms to the slow background rhythmic structures just described; see *The Music*, pp. 46-47.

[13] Link, "Long-Range Polyrhythms," pp. 68-74.

[14] This is also discussed by Link who illustrates Carter's subdivision of the slow polyrhythms of *Enchanted Preludes* (1988) and the Oboe Concerto (1986-1987), the former resulting in "a kind of metric layering." Ibid., pp. 75-89.

[15] This particular aspect of Carter's rhythmic language will be re-visited in chapter 2, subsection 2.3.2.

Example 1.3: Partial transcription of a page from Carter's sketches for the Oboe Quartet (2001)



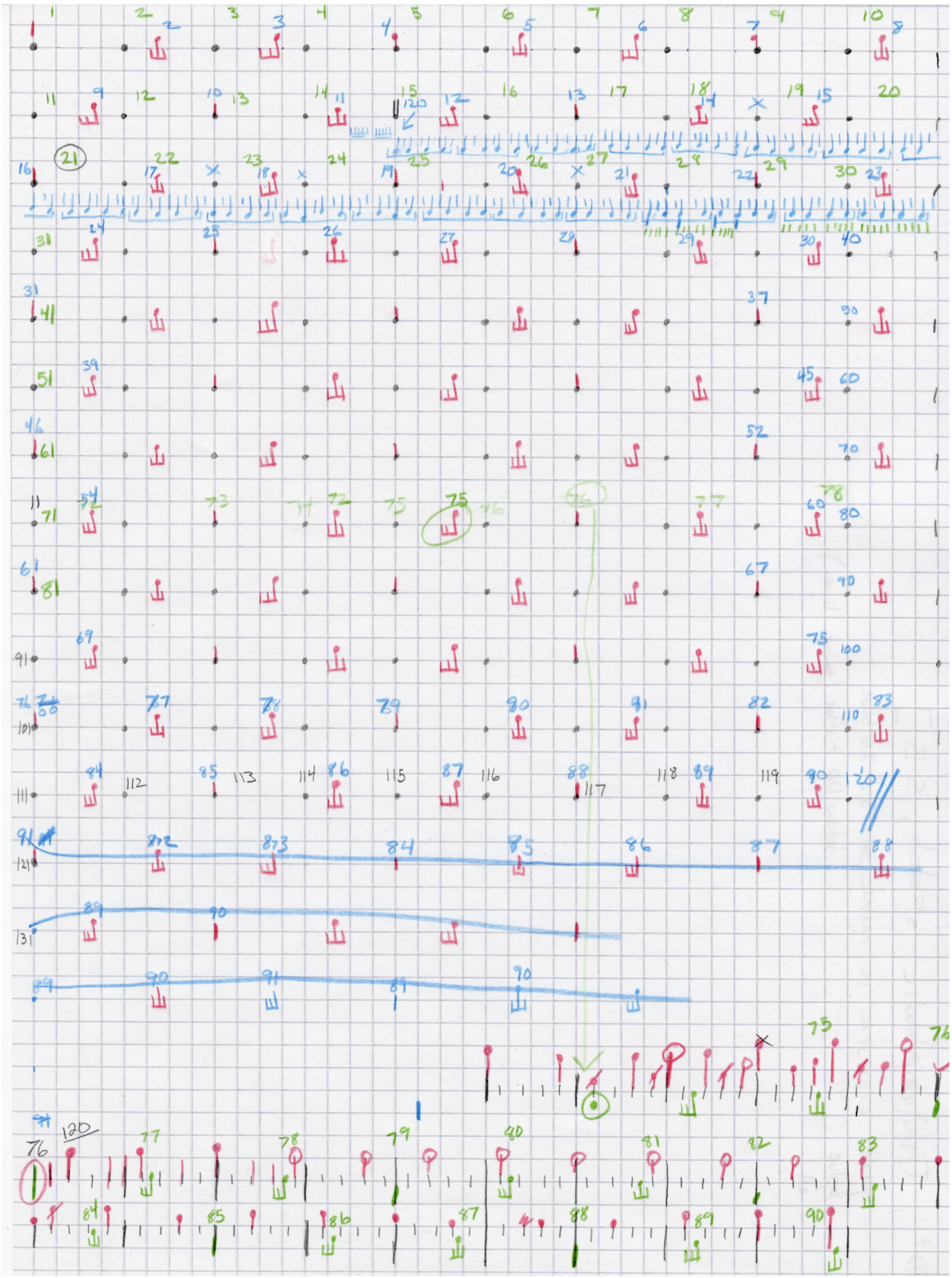
Example 1.3 presents a partial transcription of a page of the sketches for the Oboe Quartet (2001) that is representative of this aspect of Carter's sketching technique.[16] Here, the cello appears to be relatively worked out, and background pulsations are numbered (pulsations 85 and 86) and/or circled in purple. In contrast, the oboe part is only fragmentary, but is explicitly related to a timespan of 8 quintuplet sixteenths, as exhibited by the attack patterns above the staff (stems and beams represent implied pulsations and noteheads or strokes are used for implied beats). Several pitches in the staff are also boxed in brown and/or circled in green, suggesting that these are background pulsations from a strand that contrasts with the cello's. This is not to say that the polyrhythmic framework is strictly followed or made obvious throughout a given piece. On the contrary, as we will see later, Carter seems to delight in setting up expectations and denying their realization. In **example 1.3**, Carter is evidently in the process of re-barring the passage and the timespan associated with the cello's background pulsations is not entirely clear (possibly due to a metrical modulation), especially since it uses contrasting subdivisions (triplet eighths and sixteenths). In fact, the polyrhythmic graph for this piece bears the following comment (in the composer's hand): "oboe quartet discarded rhythmic plan."[17] In any case, the mapping out of the temporal space provides a framework with which the composer can play to build textures that are characterized by a wealth of contrasting and suggestive interconnections, temporal as well as tonal.

---

[16] This transcription includes only the top two systems of a 24-stave music manuscript page; Elliott Carter Collection, Paul Sacher Foundation, Basel (accessed March 2008).

[17] Elliott Carter Collection, Paul Sacher Foundation, Basel (accessed March 2008); a transcription is in possession of the author. Although Carter later abandons his polyrhythmic plan for this piece, it is not clear whether the final composition includes traces of the initial work represented by this sketch.

Example 1.4: Transcription of Carter's rhythmic plan for *90+*



Although *90+* is built on a single background pulse, as opposed to a long-range polyrhythm, it also exhibits other structural pulses (or "cross-pulses").[18] **Example 1.4** provides a transcription of Carter's rhythmic plan on graph paper for this piece.[19] Except for a few disruptions, the specific location of the 90+ pulsations is clearly marked by the use of a subdivided quarter beat within the grid provided by 4/4 measures. However, structural polyrhythms are not apparent. The Chord and Rogue pulses discussed previously (**ex. 1.1**) do not appear at the beginning of the graph, and the potentially "competing" pulse that appears in the passage re-worked at the bottom of the graph (mm. 90-113) is actually the notated measure. The re-alignment of the 90+ pulse with the bar line at m. 90 and concurrent metrical modulation suggest that Carter might have intended this time-point (90+ pulsation 76) as a clearly marked emphatic arrival. The use of a 2/4 measure also ensures that the 90+ pulsations will not align with the bar line at any point for the remainder of the piece; instead they periodically return at the mid-point (see 90+ pulsations 79, 82, 85, and 88 in **ex. 1.4**). Thus, it would appear that at least visually (and possibly conceptually) the 90+ pulse and the notated measure do combine to form a local polyrhythm, with a coincidence point every three 90+ pulsations against every five notated bar (i.e., a 5:3 structural polyrhythm).

A more detailed examination of the rhythmic landscape of *90+* for piano as realized in the score does reveal that in this piece, polyrhythms are operative at many structural levels, from the background to the foreground. **Table 1.1** provides sectional divisions of the piece based on tempo and polyrhythmic activity along with notational equivalences and

---

[18] The expression "cross-pulse" is also sometimes used to refer to the resulting polyrhythm. To avoid confusion, this term will not be used formally in this dissertation.

[19] Elliott Carter Collection, Paul Sacher Foundation, Basel (accessed March 2008).

resultant speeds for all structural pulse streams. As shown by this table, each section of *90+* is characterized by at least two structural pulse streams, each of which is associated with a particular timespan and beat subdivision.

Table 1.1: Structural chart with tempi, pulse streams, notational equivalences, and resultant speeds[20]

Section	Bar-Measures	Tempo	90+ Pulse*	Cross Pulses & Resultant Speeds**
A	1–14	♩ = 96	16 ♪ 3	17 𝅘𝅥𝅯 = MM 23 23 𝅘𝅥𝅯 5 = MM 21
B[1]	15–36	♩ = 120	20 ♪ 3	15 𝅘𝅥𝅯 = MM 32***
C[1]	37–48	♩ = 72	12 ♪ 3	9 𝅘𝅥𝅯 = MM 32
C[2]	49–89	♩ = 72	8 ♪	𝅗𝅥. = MM 24
B[2]	90–113	♩ = 120	20 ♪ 3	17 𝅘𝅥𝅯 = MM 28 14 𝅘𝅥𝅯 = MM 34 (7 𝅘𝅥𝅯 = MM 69)
Coda	114–118	*a piacere*	11 𝅘𝅥𝅯 3 to 1 𝅘𝅥𝅯 6	-

*The 90+ pulse moves at the constant speed of MM 18.

**Resultant speeds are rounded up to the closest integer.

***This speed is preceded by an accelerando from 14 𝅘𝅥𝅯 to 5 𝅘𝅥𝅯 (about MM 34 to MM 96).

The re-notation of the 90+ pulse (fourth column) is a by-product of Carter's use of "metrical modulation" as formal device, that is, "a proportional change in tempo effected by

[20] Schiff also provides a detailed description of the re-notation of the slow structural pulse as well as of the timespan associated with most cross-pulses; see Schiff, *The Music*, pp. 218 and 220. The information contained in this chart complements Schiff's description and presents a contrasting sectional segmentation based solely on metrical modulation (and thus, the use of the term "structural" rather than "formal" to qualify this chart).

the re-notation of a metronomic speed."[21] In the rhythmic plan for *90+*, two metrical modulations were already apparent, both of which prepared a modulation to a notated tempo of quarter = 120 (see m. 15 and 90+ pulsation 76 in **ex. 1.4**). In the graph for *Esprit doux/Esprit rude I* (**ex. 1.2**), the composer had also worked-out the different tempo sections below the polyrhythmic plan, specifying the value and number of subdivisions for each background pulse as it migrates from one tempo section to another: for example, as the clarinet moves from tempo section A to B, it goes from a notated value of 28 quintuplets (5/28) to 12 quintuplets (5/12).[22] In *90+*, the resulting speed of MM 18 (see **table 1.1**) also confirms that this pulse has structural significance throughout the work and that it provides a background for other rhythmic processes. Because of the metrical modulations, notational equivalences and resultant speeds will be helpful to determine the structural status of emerging pulse streams. The various beat subdivisions and resultant speeds for *90+* are illustrated in **table 1.2**.[23] Those notational values that result in equivalent speeds across tempo sections could be characterized as "pivot speeds," given that they allow the

---

[21] Schiff, *The Music*, p. 41. The expressions "metric modulation" and "tempo modulation" are also used in the Carter scholarship. The significance of metrical modulations in the context of structural polyrhythms will be explored in more details in Chapter 2, subsection 2.3.1. In the present chapter, I will use the expression "metrical modulation."

[22] Despite appearances, these equivalences are not dependent on the planned tempo of half note = MM 35. On the other hand, tempo sections will have to preserve the same speed ratio in relation to each other. I will not discuss the actual realization of these metrical modulations here.

[23] This table includes only those values that are actually used in this work. Carter sketches show that he systematically goes through all possible subdivisions and resultant speeds, e.g., 96 × 3 = 288, × 4 = 384, × 5 = 480, × 6 = 576, × 7 = 672 (Carter rarely uses more that seven subdivisions, except for tremolos or unless the tempo is slow enough to allow for clearly articulated attacks). These annotations are sometimes found on plain sheets of paper, but most often appear in the margins of the sketches on music manuscript paper.

performer(s) to establish a new tempo with an equivalent time unit. Here, the values made available by the three tempo areas and contrasting subdivisions of the beat unit insure a variegated rhythmic landscape as well as the technical means for Carter's metrical modulations.

Table 1.2: Beat subdivisions and resultant speeds for *90+* (pivot speeds are marked in bold)

	♩ = 96	♩ = 120	♩ = 72
𝅘𝅥𝅯 7	-	-	MM 504
𝅘𝅥𝅯 6	-	MM 720	MM 432
𝅘𝅥𝅯 5	**MM 480**	MM 600	MM 360
𝅘𝅥𝅯	MM 384	**MM 480**	**MM 288**
𝅘𝅥𝅮 5	-	MM 300	-
𝅘𝅥𝅮 3	**MM 288**	MM 360	MM 216
𝅘𝅥𝅮	MM 192	MM 240	MM 144

1.1.2 Character-Patterns and Referential Pitch Collections

The composer's concern for the perceptibility of contrasting layers becomes manifest in his use of *character-patterns*, which are described by Schiff as the "association of *intervals*, *metronomic speeds*, *polyrhythms* and *rhythmic characters* used to dramatize the musical personalities of instruments and instrumental groups and to make clear the stratification of texture."[24] In the present example, each pulse is clearly defined by a different

[24] David Schiff, *The Music of Elliott Carter*, rev. 2nd ed. (Ithaca: Cornell University Press, 1998), p. 36. Carter himself uses the expressions "character-continuities" and "character-structures"; see Allen Edwards, *Flawed Words and Stubborn Sounds: A Conversation with*

type of articulation (90+ pulse = tenuto accent, Chord pulse = legato, and Rogue pulse = staccato) as well as a specific combination of dynamic, registral, and textural cues. For example, in the opening 90+ (**ex. 1.1**), the 90+ pulsations are always relatively louder than the Rogue pulsations while the Chord pulse is softer than the two other pulses and is static both in dynamic and register.

In his attempt to insure intelligibility, Carter also develops his own vocabulary of pitch materials adapted to the expression of contrasting characters characterized by associated intervals and speeds.[25] The foreground discontinuities created by the juxtaposition and superimposition of contrasting gestures, which are so characteristic of modernist composition, are mediated by the composer's use of "signature" set-classes, e.g., the all-interval tetrachord (AIT) pair [0137]/[0146] and the all-trichord hexachord (ATH) [012478], as unifying harmonic devices.[26] In **example 1.1**, the pitch material used for each set of three

---

Elliott Carter (New York: W. W. Norton, 1971), pp. 67–122. For a specific investigation of such formal processes and their implication for performance, see Leonard L. Garrison, "Elliott Carter's *Scrivo in vento*," *The Flutist Quarterly: The Official Magazine of the National Flute Association*, 19/4 and 20/1 (Summer and Fall 1994): 86–92 and 75–80.

[25] For an overview, see David Schiff, *The Music of Elliott Carter*, 1st ed. (London and New York: Eulenburg Books and DaCapo Press, 1983), pp. 23-70; Jonathan W. Bernard, "The Evolution of Elliott Carter's Rhythmic Practice," in *Perspectives of New Music*, 26/2 (1998): 164–203; and Max Noubel, *Elliott Carter, ou Le temps fertile* (Genève, Switzerland: Contrechamps, 2000).

[26] This feature has been widely acknowledged in the scholarly literature. For representative studies dealing with pitch processes, see Jonathan Bernard, "Spatial Sets in Recent Music of Elliott Carter," *Music Analysis* 2/1 (1983): 5-34, and "Problems of Pitch Structure in Elliott Carter's First and Second String Quartets," *Journal of Music Theory* 37/2 (1993): 231-266; David I. H. Harvey, *The Later Music of Elliott Carter: A Study in Music Theory and Analysis* (New York: Garland Publishers, 1989); Andrew Mead, "Twelve-Tone Composition and the Music of Elliott Carter" in *Concert Music, Rock, and Jazz since 1945: Essays and Analytical Studies* (Rochester, NY: University of Rochester Press, 1995), pp. 67-102; and Marguerite Boland, "'Linking' and 'Morphing': Harmonic Flow in Elliott Carter's *Con leggerezza pensosa*," *Tempo* 237 (2006): 33–43. For in-depth analyses of representative works, see Guy

pulsations corresponds to permutations of the same registrally ordered hexachord, {E3, G3, Bb3, Eb4, F4, B4}, a representative of ATH [012478].

Carter's use of techniques emphasizing unity of sound and rhythm and the simplification of texture associated with the composer's predilection for chamber and solo instrumental media in recent years have been interpreted by some authors as representing a return to some form of classicism.[27] However, Carter's interest in background continuity and organic connection is not new, and much like his use of long-range polyrhythms his particular treatment of AIT and ATH is the result of many years of experimentation with referential collections. For example, in discussing the dramatic possibilities offered by the recourse to 'idiomatic' gestures with Allen Edwards (1968–70), Carter comments:

> And of course in any "dialogue," musical or otherwise, there must be areas of overlap and interchange as well as points of divergence. Thus in my music there is a kind of ongoing dialectic of affirming and contradicting the character of the instruments involved, *which nonetheless have an organic relation to the*

Capuzzo, "Variety Within Unity: Expressive Ends and their Technical Means in the Music of Elliott Carter, 1983-1994" (Ph. D. dissertation, University of Rochester, Eastman School of Music, 1999), and Mark Sallmen, "A Flexible Approach to Ordering and Grouping in Atonal Music in General: Text-Music Relations in Elliott Carter's *In Sleep, In Thunder* in Particular," (Ph.D. dissertation, University of Rochester, 1998).

[27] In discussing Carter's late works, Schiff emphasizes the "almost classical" design of works from the 1980s – such as *Triple Duo* (1983), *Penthode* (1985), and the Quarter no. 4 (1986) – which "balance Carter's habitual opposition of instrumental forces with a desire to bring them all together." Capuzzo argues that the composer's predilection for "variety within unity" such as that resulting from the use of "all-partition presentation" of a reference collection can be related to the 18th-century practice of "ars combinatoria." On the other hand, Bye disputes an assessment that "the fluency and simplification of texture and polyrhythm" exemplified by *Gra* for clarinet (1993) would represent a form of classicism, emphasizing instead the "rupture between moment-to-moment Dionysiac flux… and any subliminal coherence existing at the background level." See David Schiff, "Elliott Carter," in *Grove Music Online*, ed. L. Macy (Accessed 30 May 2007), <http://www.grovemusic.com>; Guy Capuzzo, "Letters to the Editor: *From Guy Capuzzo*," *Tempo* 204 (1998): 53; and Anthony Bye, "Carter's 'Classic' Modernism," *Tempo* 189 (1994): 3.

> *character of the musical ideas and to the formal-dramatic conception of the whole work* in each case [emphasis added].[28]

It appears that Carter's tendency to mitigate surface discontinuity with underlying continuity is very closely related to the value he attaches to individual character and the need for meaningful exchanges between individuals, which sometimes leads to dramatic interactions or conflicts (as when characters talk without listening to each other) but can also be a transformative experience (as when characters take on the characteristics of each other). And this has been an enduring trait in the composer's aesthetics.

Perhaps, part of the perceived classicism in Carter's music comes from the composer's adherence to a traditional model of music production and reception. Carter states:

> I realize that what I am trying to do in music always remain in a "restricted frame," in that I try to write music that will appeal to an intelligent listener's ear and will be a strong enough expression so that the listener will be drawn to hear and grasp this music when it is presented by a performer who finds it gratifying enough to play effectively. It may take many years for the listener to be convinced, but I believe that my training and experience as a composer enable me to prejudge a future listener.[29]

Thus, for Carter the art of the composer rests in his ability to effectively communicate his musical ideas through the score. His criticisms of various techniques employed by other modernist composers – serial, textural or aleatory – are especially revealing:

> It should be obvious that the idea of allowing four (as in the Quartet) or two (as in the Double Concerto) different streams of music to be heard together in any one of the possible uncoordinated ways that have been used by Ives and others in recent years will, from the point of view I am describing here, produce a form of entropy, a degrading of the possibilities of communication, which to me have ever to be revitalized and sharpened.[30]

---

[28] Edwards, *Flawed Words*, p. 69.

[29] Ibid., p.77.

[30] Elliott Carter, *Elliott Carter: Collected Essays and Lectures, 1937–1995*, Ed. Jonathan W. Bernard (Rochester, NY: University of Rochester Press, 1997), p. 271.

Considering the paramount importance given to the dramatic interaction of contrasting characters that are set apart by temporal means, the organic unfolding of interrelated structural levels becomes a key to effective communication. In this light, the fluidity and clarity of expression achieved in the late works can be understood as the latest stage of refinement in the oeuvre of a composer whose vision is deeply rooted in a modernist aesthetic.

1.2 Purpose of Study

In his work on long-range polyrhythms, Link identifies 11 works, from the 1980 *Night Fantasies* for solo piano to the 1989 *Anniversary* for orchestra, in which a single polyrhythmic cycle generates the background structure.[31] In many of his compositions, Carter also uses competing pulse streams at different structural levels. In the excerpt from *90+* discussed above, two of the concurrent pulse streams serve as middleground for the elaboration of textures with varying degrees of foreground activity. As of now, very little work has been done on the interaction between foreground polyrhythmic activity and structural pulse streams, more specifically, on the metric implications of these multileveled polyrhythmic structures.[32] The reduction of simultaneous rhythmic layers to mathematical ratios, while imbued with elegance and useful for generalizations, is poorly expressive of the various shades of effects thus created. Furthermore, despite the recognition of the primordial

---

[31] Link, "Long-Range Polyrhythms," pp. 117-120.

[32] An exception is Rebecca Jemian, "Rhythmic Organization in Works by Elliott Carter, George Crumb, and John Adams: Rhythmic Frameworks, Timepoints, Periodicity, and Alignment" (Ph.D. dissertation, Indiana University, 2001); Jemian examines various types of rhythmic processes in Carter's *Enchanted Preludes* for flute and cello (1988), but her work does not focus on the metrical implications of the structures she describes.

importance of multileveled polyrhythmic structures in the compositional process of Carter, most analyses of his works do not provide the reader with meaningful interpretations of the interaction of pitch and temporal structures beyond the identification of character-patterns and their associated speeds and intervals, if at all.[33]

This dissertation presents a theoretical framework and methodology for the study of polymetric structures in the solo and chamber instrumental music of Elliott Carter after 1980, a subset of works where previous analytical studies have shown that the use of contrasting speeds or asynchronous pulse streams is an important structural principle (Link 1994; Garrison 1994; Ravenscroft 1992 and 2003; Jemian 2001; Roeder 2006).[34] This study is also conceived as the first step in a larger project dealing with polymeter in contemporary musical practices. The usefulness of the concept of polymeter for analysis is not limited to twentieth-century music; it has also been used in analyzing various repertoires ranging from European vocal music of the late 14th century (Günther 1990 and 1995), American jazz (Folio 1995; Pressing 2002), and African music (Locke 1982; Arom 1991), although the legitimacy of this term in the latter context has recently been questioned (Agawu 2003). However, the

---

[33] There are a few exceptions to this, although none of these investigate the interaction of harmonic and rhythmic processes in a systematic way. In his dissertation, Link gives a few examples of coordination of harmony with polyrhythmic design; see Link, "Long-Range Polyrhythms," pp. 101-109. More recently, Theisen also explores various processes by which pitch collections are unfolded in time; see Alan Jude Theisen, "Emancipated Discourse, Temporal Unfolding, and the Articulation of Harmonic Regions in Selected Chamber Compositions by Elliott Carter" (M.M. Thesis, University of Southern Mississippi, 2006). Finally, Roeder proposes a narrative interpretation of *Enchanted Preludes* based on both rhythmic and pitch structures, although it is not clear exactly how rhythm and pitch processes influence one another; see John Roeder, "Autonomy and Dialog in Elliott Carter's *Enchanted Preludes*," in *Analytical Studies in World Music*, ed. Michael Tenzer (New York: Oxford University Press, 2006), pp. 377-414.

[34] For an overview, see Schiff, *The Music*, which includes short descriptions of Carter's works up to 1998.

abundance of terminology in the music-theoretical literature aimed at describing the products of the phenomena resulting from the use of polymeter serves witness not only to its relevance but also to the absence of a general theory of polymeter.[35] Given the continued relevance of polymetric structures in contemporary music, it is surprising that there have been relatively few attempts at addressing the topic of polymeter in a more formal way (Smither 1960 and Arom 1991 are two examples of such attempts).

1.2.1 Interdisciplinary Considerations

According to Carter, the rise of a new "musical psychology" is the main force behind the musical discourse represented by early modernists (e.g., Debussy and Stravinsky), and for American composers, this has meant a continued research in the "experience of time."[36] From Carter's writings and interviews, it is also clear that he composes with a listener in mind and that his understanding of the psychology of the listener has contributed to shape his musical discourse (see subsection 1.1.2). While psychological studies of music have

---

[35] The term that is most commonly used in reference to the structures resulting from concurrent rhythmic strands is "polyrhythm." In *African Polyphony and Polyrhythm* (Cambridge: Cambridge University Press, 1991), Simha Arom offers a systematic exploration of rhythmic terminology without exhausting the needs for further formalization. Important studies that offer methodological models for the analysis of polyrhythmic structures include: Brian Robison's "*Carmen Arcadiae Mechanicae Perpetuum*: Towards a Methodology for Analyzing Harrison Birtwistle's Music Since 1977" (D.M.A. diss., Cornell University, 1999), in which Robison proposes the concept of "polycycle" to refer to concurrent cycles of similar or contrasting musical elements; in his "Inversion, Subversion, and Metaphor: Music and Text in Elliott Carter's *A Mirror on Which to Dwell*" (D.M.A. diss., University of Washington, 1992), Craig Weston offers the concept of "phase cycles," which is formalized into "proximity cycles" in John F. Link's "Long-Range Polyrhythms in Elliott Carter's Recent Music" (Ph.D. diss., City University of New York, 1994).

[36] Elliott Carter, "A Further Step" (1958), in *Elliott Carter: Collected Essays and Lectures, 1937-1995*, ed. Jonathan W. Bernard (Rochester, NY: University of Rochester Press, 1997), pp. 5–11.

proposed interesting models for the perception and cognition of musical meter, polymeter remains a marginal topic.[37] In one of the more music-theoretically informed discourses on musical meter, Justin London even goes as far as to assert that from a psychological perspective, "there is no such thing as poly*meter*."[38] In addition, the interpretation of the score by a performer and his or her role in the interpretation of the musical score call for a methodical investigation of expressive variations, especially in regard to timing patterns used by different performers and their effect on the perception of polymetric structures.[39] Thus, while there are perceptual and conceptual challenges due to the complexity of these textures and given the definition of meter as currently accepted in mainstream North-American music theory, Carter's own words call for a consideration of the metrical properties of multileveled polyrhythmic structures. In developing the concepts and analytical tools presented in this dissertation, I have kept this statement in mind:

> A work that does not take into account the listener's ability to distinguish sounds, to grasp, remember, and compare in some way their combinations, both sequential and simultaneous, in small durations, intermediate lengths, as well as over the whole composition, is very unlikely to hold a permanent interest for the listener.[40]

---

[37] For reference, see Stephen Handel, *Listening: an Introduction to the Perception of Auditory Events* (Cambridge, Massachusetts: MIT Press, 1989), and Eric F. Clarke, "Rhythm and Timing in Music," in *The Psychology of Music*, ed. Diana Deutsch (San Diego: Academic Press, 1999), pp. 473-500.

[38] Justin London's *Hearing in Time: Psychological Aspects of Musical Time* (New York: Oxford University Press, 2004), p. 50.

[39] This research is summarized and discussed in London's *Hearing in Time*, pp. 144-152. Although there have been several studies on this topic since the groundbreaking work Shaffer, these are usually limited to common-practice repertoire; see L.H. Shaffer, "Performances of Chopin, Bach, and Bartók: Studies in Motor Programming," *Cognitive Psychology* 13 (1981): 326-376. For a more recent and representative example, see Bruno H. Repp, "Pattern Typicality and Dimensional Interactions in Pianists' Imitation of Expressive Timing and Dynamics," *Music Perception* 18/2 (2000): 173-211.

[40] Carter, *Collected Essays*, p. 224.

The contributions of composers, performers, and listeners shape the musical object in a variety of ways, which are not always concordant with each other. In addition, the socio-cultural context within which the musical creation takes place also has a determining effect on the meaning derived from it. The framework proposed embraces the diversity of agents involved in the creation of (and participation in) a musical entity. It was also developed with the assumption that since the composer's work and the performers and listeners involved with his music result in a particular practice of polymeter, any investigation of the use of this technique must be conducted in way that is sensitive to the context from which this practice has emerged and the modes according to which it operates. While the scope of this dissertation is necessarily too limited to engage in an in-depth study of all of these aspects, my work has been informed by the experience of expert performers of Carter's work. Finally, the work presented here also includes a more formal investigation of the realization of polymetric potential from a psychological perspective.[41]

1.2.2 Repertoire

The repertoire for this study is the solo and chamber instrumental works after 1980, although discussions and demonstrations will focus on three representative works within this repertoire, i.e., *Changes* for guitar (1983), *Esprit rude/Esprit doux I* for flute and clarinet (1985) and *90+* for piano (1994). The choice of repertoire was driven by three main factors: (1) the explicit aesthetic goals of the composer, which find a particularly refined expression

---

[41] I am especially grateful to Ursula Oppens, Virgil Blackwell, Fred Sherry, Anahid Ajemian, and Rolf Schulte for their interest in this project and for generously taking the time to meet with me and share their experience.

in these works; (2) practical considerations relating to the relatively limited length and complexity; (3) the special nature of the musical narrative suggested by the composer for many of these works.

Before 1980, Carter had only written two multi-movement works for solo instruments (see **table 1.3**), the Piano Sonata (1945-6) and the Eight Pieces for Timpani (1950-66). Although both works exhibit Carter's growing interest in experimenting with rhythmic elements—the use of proportional tempo changes in *Canaries* being the most systematic example of this—it is not until the composition of *Night Fantasies* for piano that the implications of Carter's rhythmic language in works for solo instruments can be observed. On the other hand, there is a larger number of chamber instrumental works before 1980, several of which exhibit techniques similar to those employed in the later works, including the use of contrasting (and often simultaneous) musical characters set apart by temporal devices and multi-leveled polyrhythmic textures. Although the Sonata for Cello and Piano (1948) is certainly notable for being the first chamber work to make use of contrasting rhythmic layers, the first three string quartets are probably more characteristic of Carter's writing for chamber instrumental groups. In the literature and Carter's own comments, the quartets are often taken as representative of successive stages in the composer's creative research. Thus, while the first string quartet (1950-1951) pushes the limits of ensemble playing by capitalizing on polyrhythmic layering and metrical modulation, the second (1959) crystallizes Carter's distinctive use of character-patterns. The formal opposition of two duos in the third string quartet (1971) points to one of Carter's favored procedure in ensemble works after 1980, that is, the division of a larger ensemble into more or less formal musical partnerships, a technique that has implications for the kind of large-scale organization that

can be found in later works. Thus, in the *Triple Duo* (1982-1983), each duo proceeds according to a different temporal layer in a long-range polyrhythm.[42]

Table 1.3: Elliott Carter's solo instrumental works[43]

Title	**Instrument**	**Date**
Piano Sonata	Piano	1945-1946
Eight Pieces	Timpani	1950-1966
Night Fantasies	Piano	1980
Changes	Guitar	1983
4 Lauds: I. *Statement – Remembering Aaron* II. *Riconoscenza per Goffredo Petrassi* III. *Rhapsodic Musings* IV. *Fantasy – Remembering Roger*	Violin	1984-2001 (1999) (1984) (2001) (2001)
Scrivo in Vento	Flute	1991
From Trilogy: I. *Bariolage* II. *Inner Song*	 Harp Oboe	1992 (1992) (1992)
Gra	Clarinet	1994
90+	Piano	1994
4 Figments: *Figment* [I] *Figment II* *Figment III* *Figment IV*	- Cello Cello Double bass Viola	1995-2007 (1995) (2001) (2007) (2007)
A 6 Letter Letter	English Horn	1996
Shard	Guitar	1997
Two Diversions	Piano	2000
Retrouvailles	Piano	2000
Steep Steps	Bass Clarinet	2001
Retracing	Bassoon	2002
Two Thoughts About the Piano: I. *Intermittences* II. *Caténaires*	Piano	2005-2006 (2005) (2006)
HHBH	Oboe	2007
Matribute	Piano	2007

[42] Link, "Long-Range Polyrhythms," p. 117; Schiff, *The Music*, pp. 120-124.

[43] This table does not include early works that have not been published, but does include recent works that have not yet been published.

In subsection 1.1.1, we have seen that Carter uses structural polyrhythms that vary in their scope, i.e., long-range polyrhythms that encompass the entire duration of a work, more local polyrhythms that may be superimposed on a long-range polyrhythm, and contrasting pulse streams just below the surface that allow for irregular rhythms as well as perceptible simultaneous speeds. A survey of Carter's sketches for the focus repertoire confirmed the presence of structural polyrhythms in many of these works; generally, only solo instruments that had extended polyphonic capabilities (e.g., piano, guitar, harp) showed evidence of structural polyrhythms.[44] A summary of these findings is provided in **table 1.4** with references to sources that support the findings, including the sketches. In the last column ("Structural Polyrhythms"), "Yes" means that for this work, there is a polyrhythmic plan or evidence of polyrhythmic layering from the sketches, as represented by contrasting attacks patterns accompanying the working-out of pitch events. Works marked as "Maybe" either have a polyrhythmic plan that was discarded or lost, or the evidence for polyrhythmic layering was weak, and a "No" means that there is no polyrhythmic plan and no conclusive evidence of contrasting and simultaneous attack patterns. Finally, works for which sketches were not available are marked with a question mark ("?"). This survey is certainly not definitive, but it provides a helpful point of departure for anybody interested in exploring the use of structural polyrhythms in Carter's sketches. They also suggest some of the works that may be most interesting for an in-depth analysis of polymetric potential and realization.

---

[44] It should be said, however, that even in solo works Carter often plans on implied simultaneous speeds, such as result from the alternation of two or more contrasting characters, e.g., *Riconoscenza per Goffredo Petrassi* for violin (1984), *Scrivo in Vento* for flute (1991), and *Gra* for clarinet (1994). This technique would also be interesting to explore from the perspective of polymetric potential and realization, possibly in tandem with a psychological study on the role of memory.

Table 1.4: Evidence of structural polyrhythms in Elliott Carter's solo and chamber instrumental works after 1980 (LRP = Link's "Long-Range Polyrhythms"; MEC = Schiff's *The Music of Elliott Carter*; PSS = Paul Sacher Stiftung)

Title & Date[45]	**Instruments**	**Sources**[46]	**Structural Polyrhythms**
Triple Duo (1983)	Flute, Clarinet Piano, Percussion Violin, Cello	LRP, MEC, PSS	Yes
Changes (1983)	Guitar	LRP, MEC, PSS	Yes
Canon for 4 – Homage to William Glock (1984)	Flute, Bass clarinet Violin, Cello	-	No
Riconoscenza per Goffredo Petrassi (1984)	Violin	-	No
Esprit rude/Esprit doux I (1985)	Flute, Clarinet	LRP, MEC, PSS	Yes
String Quartet No. 4 (1986)	Violin (2), Viola, Cello	LRP, MEC, PSS	Yes
Birthday Flourish (1988)	Trumpet (2), Horn, Trombone (2)	-	No
Enchanted Preludes (1988)	Flute, Cello	LRP, MEC, PSS	Yes
Con leggerezza pensosa (1990)	Clarinet, Violin, Cello	MEC, PSS	Maybe
Quintet for Piano and Winds (1991)	Oboe, Clarinet, Bassoon, Horn, Piano	MEC, PSS	Yes
Scrivo in Vento (1991)	Flute	-	No
Trilogy (1992) I. *Bariolage* II. *Inner Song* III. *Immer Neu*	Oboe, Harp	 PSS - PSS	 Maybe No Maybe
90+ (1994)	Piano	PSS	Yes
Fragment (1994)	Violin (2), Viola, Cello	PSS	Yes
Gra (1994)	Clarinet	-	No
Esprit rude/Esprit doux II (1995)	Flute, Clarinet, Marimba	MEC, PSS	Yes[47]

[45] The first date of publication is used unless the work has not been published yet.

[46] Sources are given only for cases where there is at least some evidence of structural polyrhythms.

[47] It should be noted that the information provided by Schiff (*The Music*, pp. 148-149) about the timespans associated with the flute and clarinet pulses is erroneous (flute = 51 quintuplet sixteenths, not 68; clarinet = 40 triplet eighths, not 48).

Title & Date[48]	**Instruments**	**Sources**[49]	**Structural Polyrhythms**
Figment (1995)	Cello	-	No
String Quartet No. 5 (1995)	Violin (2), Viola, Cello	PSS	Maybe
A 6 Letter Letter (1996)	English horn	-	No
Luimen (1997)	Trumpet, Trombone, Vibraphone, Mandolin, Guitar, Harp	PSS	Maybe
Shard (1997)	Guitar	PSS	Yes
Quintet for Piano and Strings (1997)	Violin (2), Viola, Cello, Piano	PSS	Maybe
Fragment No. 2 (1999)	Violin (2), Viola, Cello	-	No
Statement – Remembering Aaron (1999)	Violin	-	No
Two Diversions (1999)	Piano	-	No
Retrouvailles (2000)	Piano	PSS	Maybe
Fantasy – Remembering Roger (2001)	Violin	-	No
Figment II (2001)	Cello	-	No
Hiyoku (2001)	Clarinet (2)	-	No
Oboe Quartet (2001)	Oboe, Violin (2), Viola, Cello	PSS	Maybe
Rhapsodic Musings (2001)	Violin	-	No
Steep Steps (2001)	Bass clarinet	-	?
Au Quai! (2002)	Bassoon, Viola	PSS	No
Call (2003)	Trumpet (2), Horn	PSS	No
Intermittences (2005)	Piano	-	?
Caténaires (2006)	Piano	-	?
Clarinet Quintet (2007)	Clarinet, Violin (2), Viola, Cello	-	?
Figment III (2007)	Double bass	-	?
Figment IV (2007)	Viola	-	?
HBHH (2007)	Oboe	-	?
Matribute (2007)	Piano	-	?

[48] The first date of publication is used unless the work has not been published yet.

[49] Sources are given only for cases where there is at least some evidence of structural polyrhythms.

1.2.3 Dissertation Plan

The dissertation divides into two parts. Chapters 2 and 3 present the conceptual framework and analytical tools developed for the investigation of polymetric potential and realization in the repertoire at hand, while chapter 4 engages the question of polymeter from a psychological perspective. More specifically, chapter 2 focuses on the mapping of the polyrhythmic organization and the identification of structural levels; it also explores Carter's compositional strategies and interaction between structural levels through a re-assessment of the technique of metrical modulation as well as a detailed analysis of the background structure of *Changes* for guitar (1983), with special emphasis on background irregularities and musical narrative. After considering the issue of stream segregation and fusion in *Esprit rude/esprit doux I* or flute and clarinet (1985), chapter 3 exposes three types of polymetric structure, and returning once more to *Changes*, it investigates local as well as global conditions for polymetric potential by means of polarized reductions as well as density and nestedness analysis. While density measures allow for a consideration of the relative salience of concurrent pulse streams, nestedness analysis helps in determining the relative metrical strength of concurrent rhythmic strands. The chapter ends with a detailed discussion of three representative examples from *90+* for piano (1994), including brief discussions of the influence of listening mode and performance. In chapter 4, the issue of polymetric experience is delved into through a survey of pertinent literature from music psychology and the presentation of a listening experiment based on a passage from *90+* previously analyzed in chapter 3. Finally, chapter 5 presents a brief conclusion that puts the results of the experimental study back into the context of the proposed conceptual framework.

CHAPTER 2

Structural Levels and the Mapping of Polyrhythmic Organization

This chapter will focus on the mapping of polyrhythmic textures into structural levels; this discussion will serve as an introduction to basic terminology and concepts as well as an exploration of Carter's compositional strategies in dealing with multileveled polyrhythmic structures or *structural polyrhythms*. I will propose a partitioning of the resulting textures into three structural levels (foreground, middleground, and background) and explore the nature of the relationships between these levels in general, as well as their potential for narrative. In the course of this exposition we will first re-visit *90+* briefly, but the focus will be on one representative work in particular, that is, *Changes* for guitar (1983).

2.1 Preliminary Remarks

While it is beyond the scope of this dissertation to fully address the issues that its conceptual framework raises about the nature of musical meter, a few points should be made; these are offered in guise of premises:

(1) Although accent theory is currently predominant in Western theories of meter, it is not at all clear that meter must be dependent on accent (whether implied or actual). Other types of metrical organization are conceivable and may be encouraged by the music at hand. In some cases, some form of *durational meter* might be more appropriate, that is, *the organization of rhythmic events by recurring timespans, without implied accentual patterns* such as "strong-weak" (duple accentuation) or "strong-weak-weak" (triple accentuation). Furthermore, it

seems reasonable to posit that meter does not necessarily require isochronous beats, as exemplified by the *bell* patterns used in West and Central African drumming.[1] To some extent, the methodology proposed here moves toward a reconciliation of accent theory and durational meter.

(2) It is currently fashionable to define meter based on psychological models, and thus to regard meter as the cognitive product of the listener's mode of attention, regardless of musical style.[2] However, it is obvious that meter must be situated in the music (as a hybrid product of a score and its performance) as well as in the mind. Meter is an important organizational element for music's temporal fabric and as such it conditions the experience of all the participants involved in the creation of a musical object (i.e., the composers, performers, and listeners). Thus, a generalized theory of polymeter necessarily presupposes a theory of meter that is flexible enough to accommodate various musical styles and practices; it can neither be ethnocentric nor universalistic.

(3) The ecological validity of conclusions drawn from empirical research and applied to actual music must be regarded with skepticism. Although it is undeniable that

---

[1] For an example of a theory of rhythmic organization adapted to such structures, see Willi Anku, "Circles and Time: a Theory of Structural Organization of Rhythm in African Music," *Music Theory Online* 6/1 (2000), http://www.societymusictheory.org/mto/issues/mto.00.6.1/mto.00.6.1.anku_frames.html (accessed August 6, 2008). The author conceptualizes these cyclical patterns as "time lines" that are described as "n-time point sets" and generalized using a modified form of set theory. One advantage of this approach is its ability to relate different modes of performance, and thus to address cultural specificity.

[2] London proposes a formalized model for the graphic representation of musical meter in which a "metrical cycle" is understood as a "coordinated set of periodicities (…) that make up a particular attentional state"; see London, *Hearing in Time: Psychological Aspects of Musical Meter* (Oxford, UK: Oxford University Press, 2004), p. 76.

> there must be physical limits to sound perception, there might be insurmountable difficulties in testing the potential for perceptual learning involved in highly specialized skills such as those associated with different musical styles and practices. Because of the great number of variables in actual music and the methodological need of controlling these variables, most experimental studies in music perception do not use stimuli derived from actual music. Furthermore, it is not clear at all that any number of specific tasks conducted in a lab environment necessarily add up to the whole musical experience.

2.1.1 Defining Meter

The theory of rhythmic organization proposed by Lerdahl and Jackendoff can be taken as representative of the challenges facing a generalized theory of meter, especially because of the impact it has had on the field of music psychology, where it has become the principal source for the understanding of musical meter.[3] As introduced by its authors, this theory is based on an understanding of Western classical common-practice music. However, recent development in research involving national styles of meter strongly suggests that this scope might be even narrower than the authors believed it to be.[4] In addition, although the

---

[3] Fred Lerdahl and Ray Jackendoff, *A Generative Theory of Tonal Music* (Cambridge, MA: MIT Press, 1983).

[4] In a recent study, William Rothstein demonstrates that differences in notational styles of barring in music of the common-practice period are actually related to contrasting conceptualizations of meter with regard to beginning- and end-accented phrases; see "National Metrical Types in Music of the Eighteenth and Early Nineteenth Centuries," in *Communication in Eighteenth-Century Music*, ed. Kofi Agawu and Danuta Mirka (Cambridge: Cambridge University Press, 2008 [in press]), pp. 112–59. Thus, given its "downbeat early" preference rule, the theory presented by Lerdahl and Jackendoff could be characterized as biased toward beginning-accented or "German" meter.

authors claim a cognitive basis to their assertions (as represented by the appeal to the low-level Gestalt principles of similarity and proximity), it is not at all clear how the mind processes accentual cues, and whether it can actually differentiate between metric and phenomenological accents. Furthermore, the reliance on well-formedness rules results in an understanding of metric structure that presupposes recursiveness at all levels.[5] Given its limited scope in terms of referential musical materials, and the issues raised by assumptions about the relationship of accentual structure and musical meter, such a theory can hardly be accepted as basis for a generalized theory of meter.

Carter's use of complex polyrhythms as a basis for a composition, from the deep levels to the surface, calls for a better understanding not only of what *can* be perceived by the average listener but what can *become* perceivable by a listener/performer/analyst with specialized skills. It is my belief that a broader definition of meter would be a good starting point and avoid unnecessary biases that may deform our examination of Carter's polyrhythmic practices. Thus, we might define *meter* most generally as the *organizational patterns of time for musical events to which the composers/performers/listeners attune, either individually or collectively*. These patterns might be "consonant" or "dissonant" with each other and it can be posited that polymetric structures would tend to support conflicting interpretations.[6] The definition proposed may impress the reader as being too broad but it is

---

[5] A notable exception is provided by triple meter as well as some compound meters, which are characterized by a switch from duple to triple organization from at least one level to the next (down or up). It is suggestive that the majority of the musical examples discussed by the authors exhibit some form of duple meter.

[6] The concept of metric dissonance was introduced by Krebs to analyze metrically ambiguous passages in nineteenth-century music; see Harald Krebs, "Some Extensions of the Concept of Metrical Consonance and Dissonance," *Journal of Music Theory* 31 (1): 99-120.

my opinion that one cannot be too cautious in generalizing from empirical data that is far from complete (and here, I am referring to experimental studies in music perception that do not use actual music as source materials). Perceived patterns of duration and accentuation are predetermined to a lesser or greater extent by the participants and the cultural context; they may be limited by some innate constraints but it must be emphasized that these constraints are not well enough known to make musically useful generalizations at this point. The advantage of a broad definition is that several musical practices can be compared without the bias of a falsely universal theory.

From the point of view of the music scholar (and here I use this general label to include historical, theoretical, and ethnographical considerations), the definition can be further refined and adapted to the particular object of study as information on performance and compositional practices and their context become more specifically determined. From the point of view of the music psychologist, it can also be further refined by data gathered from scientific observation and experimental studies. Music theory and music psychology can (and should) influence one another, but advances in one should not be limited by the state of research in the other. This chapter will not discuss the perceptual and cognitive issues raised by a theory of polymeter in detail. Some data from empirical studies will be considered where appropriate. The issue of perception will be discussed in more detail in

---

Krebs methodology is extended by Folio to include three different types of polyrhythmic phenomena; see Cynthia Folio, "An Analysis of Polyrhythm in Selected Improvised Jazz Solos," in *Concert Music, Rock, and Jazz Since 1945: Essays and Analytical Studies*, eds. Elizabeth West Marvin and Richard Hermann (Rochester, NY: University of Rochester, 1995), pp. 103-134. The concept of "temporal dissonance" has also been used in relation to Nancarrow's music; see Margaret E. Thomas, "Conlon Nancarrow's "Temporal Dissonance": Rhythmic and Textural Stratification in the Studies for Player Piano" (Ph.D. dissertation, Yale University, 1996).

chapter 4 and issues related to performance practice are briefly explored in the discussion of *90+* in chapter 3. For the moment, it will be sufficient to say that the concept of polymeter and its relationship to those of meter and polyrhythm have been under-theorized with the result that perception and cognition studies that have attempted to deal with it have often concluded that it is psychologically impossible.[7] Thus, until the music itself is better understood, experiments will lack the guidance of a conceptual framework that would not only help researchers to come up with relevant questions but also allow human participants to respond in a meaningful way.

2.1.2 Carter's Notational Practice

In Carter's music, the status of the quarter note as foreground structural unit (i.e., the usual status reserved to the beat) is not a given. Rather, its primary function is to coordinate the contrasting pulse streams and events they support, since these events' duration is often measured according to only one of the pulse streams. However, while the notated measure (as delineated by the bar lines) allows the performer(s) to conceptualize how the parts work together, it is not meant to translate automatically into a performed accentual structure.[8]

That the specific details of Carter's notational practice are not to be taken lightly is made manifest by his comments about the communicative power of his scores:

---

[7] London, *Hearing in Time*, p. 50.

[8] For an informed discussion of the meaning of the bar line in Carter's notational practice and related performance issues, see Charles Rosen, *The Musical Languages of Elliott Carter* (Washington, D.C.: Music Division, Research Services, Library of Congress, 1984), pp. 7–8, and Leonard L. Garrison, "Elliott Carter's *Scrivo in vento*. I: Historical and Analytical Notes," *The Flutist Quarterly: The Official Magazine of the National Flute Association* 19/4 (Summer 1994): 90.

> [W]hile such works as mine do not always receive performances that present clearly all the materials, their relationships and expressive intentions, still, *these are there in the score*, and performers and listeners can gradually come to recognize them after successive performances [italics added]. If they were to be played quite differently every time, as is the intention of most aleatory scores, there would be little possibility of learning to hear and interpret more and more of what is in the scores.[9]

Carter uses various compositional strategies to express his conception of musical time as a dramatic narrative. As we have seen, this conception is often expressed through the interaction of contrasting character-patterns, each of which may be associated with a particular speed. In solo works, it would seem that this interaction can be characterized by an inner dialogue in which simultaneous and adjacent events come to be understood as belonging to conflicting streams (see Garrison 1994a, p. 88). Because of the technical means by which these contrasting speeds are achieved, that is, the grouping of beat subdivisions that contrast with each other, it would also seem fruitful for the analyst to explore the association of notational values and character-patterns. In *Changes* for guitar (1983), for example, there is much evidence of character-patterns. *Changes* is dedicated to David Starobin who premiered the work in New York on December 11, 1983 for the 75th birthday of the composer. Carter's own description of the work reads as follows:

> CHANGES, for guitar solo, is music of mercurial contrasts of character and mood, unified by its harmonic and rhythmic structure. Various aspects of the basic harmony are brought out in the course of the work, somewhat like the patterns used by bell-ringers in ringing changes.[10]

In *Changes,* the character-patterns or contrasting expressive styles appear to be based on the opposition of two main characters, the lyrical and the flamboyant flamenco strumming, each

---

[9] Carter, *Collected Essays*, p. 271.

[10] Elliott Carter, *Changes: for Guitar* (New York: Hendon Music, Boosey & Hawkes, 1983), p. 2.

of which is set apart by polyrhythmic processes, as testified by the use of sixteenths and quintuplet sixteenths on the musical surface, and these two characters are unified by the recourse to a long-range polyrhythm (for a preview, see **example 2.2**).[11]

Apart from the brief comments made by Schiff and Link, the only extended analysis of the work that has been made available to date focuses exclusively on pitch structures (Capuzzo 1999).[12] My own analysis deals with pitch structures only to the extent that they articulate the rhythmic processes on which I focus. My analysis of *Changes* followed four stages: (1) the mapping out of the background structure and extraction of the corresponding layers in the form of an annotated score and three polarized reductions; (2) production of a segmentation of the piece from component gestures to sections; (3) analysis of density and nestedness to explore the degree of integration and polarity (or dominance) of each structural pulse stream at the sub-phrase, phrase, and section level, including an exploration of polymetric potential; (4) a detailed observation of characteristic gestures and rhythmic processes at foreground and middleground levels, with an exploration of polymetric realization based on isochronicity and synchronicity. Because the purpose of this chapter is to introduce the reader to the conceptual framework and analytical tools, the analytical results will not be presented in this order. All unfamiliar terms will be defined as they arise in the discussion.

---

[11] Link, "Long-Range Polyrhythms," p. 118, and Schiff, *The Music*, 1998, p. 137.

[12] In Carter's *Harmony Book* (New York: Carl Fisher, 2002), pp. 10-14, Link also presents a harmonic analysis of the coda (mm. 131-150), where the bell-like quality of the "ringing changes" is made most transparent.

2.2 Structural Levels

The use of the notion of structural levels requires some cautious clarifications. Having received much attention from their use for the analysis of tonal structures by the theorist Heinrich Schenker (*Der freie Satz,* 1935; published in an English translation by Ernst Oster as *Free Composition*, 1977), the terms *background*, *middleground*, and *foreground* are often mistakenly understood in the field of music theory as expressing a view of embedded structures that are necessarily generative and recursive. That such an understanding of structural levels also applies to rhythm and meter has been greatly emphasized by the popularity of the methodology advanced by Fred Lerdahl and Ray Jackendoff (*A Generative Theory of Tonal Music*, 1983), who regard meter itself as dependent on the recurrence of a specific (and usually binary) pattern at successive levels (see **example 3.16** for an example inspired by their approach). However, research in the role of memory in rhythmic perception would seem to suggest that from the point of view of the cognitive processes involved, structural levels could not be so readily reduced to isomorphic patterns.[13]

As used in this dissertation, the terms background, middleground, and foreground simply suggest gradation in the scope of the structures discussed. These three levels are not completely separate; they constantly interact (albeit not always in the same manner) and even though each level will be associated with a representative timespan, each structural level results from the interaction of pulse levels. The existence of distinct structural levels is suggestive of a network of metric functions and the exploration of such structures is meant to stimulate a better understanding of the various processes at play in the temporal fabric, not

[13] The implications of the different types of memory on the conceptualization of structural levels of musical meter are discussed by Candace Brower in "Memory and the Perception of Rhythm," *Music Theory Spectrum* 15/1 (1993): 19-35.

only from the perspective of the listener/analyst but also from that of the composer. Whether these structural levels can be equated with the types of memory at work in the cognition of music's temporality (as suggested by Brower 1993), calls for further investigation.

2.2.1 Background

The uninterrupted series of slow pulsations, such as those used as the structural pulse of Carter's *90+* for piano, can be understood as giving rise to a *background pulse stream*; thus, the timespan associated with each pulsation will be designated as the *background pulse unit* (BPU or bpu).[14] The frequent subdivision of the background pulse into a series of synchronized shorter values supports the interpretation of many of these timespans as implicit *measures* (or *hyper-measures*).[15] Hence, each such timespan will be referred to as a *pulse-measure* while notated measures will be referred to as *bar-measures*.[16] The metrical properties of the pulse-measure stem from its systematic parsing into a stream of smaller timespans that are themselves subject to form larger patterns with more or less *mensural determinacy*, i.e., the ability of past durational events to give rise to specific metric

[14] The lower-case format will be used throughout as measurement unit; the upper-case format will be used in the tables and figures to refer to the type of pulse unit represented.

[15] The term "measure" is used here in a technical sense to denote a perceivable group of beats that does not necessarily correspond to the "bar," that is, the notated measure as delineated by the bar lines. These pulse-measures could also be described as "hyper-measures," and in the case where these are subjected to larger groupings, "hyper-beats."

[16] The latter term is borrowed from Christopher Hasty, *Meter as Rhythm* (New York: Oxford University Press, 1997). As will become apparent, my conception of meter is partly influenced by Hasty's theory.

expectations.[17] The following short example from *90+* for piano (1994) will clarify this point.

In **example 2.1**, each 90+ pulsation corresponds to a timespan of 8 eighth notes that is articulated by eighth, sixteenth, and triplet-eighth subdivisions.[18] The transcription and successive metrical reductions of pulse-measure 39 (provided below the musical excerpt) clarify the role of pulse subdivisions in defining metrical projections. In each successive reduction, the initial rhythmic value is understood as providing a "measuring stick" for the grouping of subsequent events. The conception of meter represented by the pulse-measure and the reductive technique employed here rest on Yeston's definition of meter as "an outgrowth of the interaction of two levels—two differently-rated strata, the faster of which provides the elements and the slower of which groups them."[19]

> This recurrent act of grouping, whether it is conceptual or whether it is represented by something in the music, then becomes a pulse itself (B), having a rate of recurrence that is necessarily slower than the rate of level A. The fundamental requirement for meter is therefore that there be a constant rate

[17] Hasty, *Meter as Rhythm*, pp. 84–95. In Hasty's theory, meter is understood as the result of a projective process experienced by the listener. "Mensural determinacy" is characterized by the ability of a determinate durational quantity to "exert future relevancy" for the becoming of a successive event. Thus, a "measure" consists of a completed event that has become durationally determinate (a "beginning") and its associated projected potential (the "continuation").

[18] Each pulsation has been labeled with an Arabic numeral that corresponds to its temporal ordering. These labels will be used to designate the background pulsations as well as the timespans (i.e., the pulse-measures) with which they are associated.

[19] Maury Yeston, *The Stratification of Musical Rhythm* (New Haven, Connecticut: Yale University Press, 1976), p. 76.

> within a constant rate—at least two rates of events of which one is faster and another is slower.[20]

The fact that in Yeston's conception of meter the recurrence of a timespan as the manifestation of a pulse is not dependent on sounded events is especially significant for the repertoire under examination, where the periodic return of slow (background) pulsations is insured by a much faster (foreground) pulse stream that also serves as basis for highly irregular patterns, the design of which can undermine the salience of, or at least often contrasts with, background pulsations.

According to Yeston, for any rhythmic foreground there may be one or more middleground levels, each of which may be used to render this foreground "meter-defined," and with each of which it will generate a particular meter.[21] In **example 2.1**, the surface activity in pulse-measure 39 initially suggests the interpretation of the eighth note as the first middleground unit (as shown by eighth-note level reduction below the musical excerpt). The following two reductions (based on the quarter- and half-note middleground levels) further elucidate the metrical organization of the whole-note background pulse unit (BPU) that corresponds to the determinate duration projected by the 90+ pulse. However, the three triplet eighths (m. 53, downbeat) point to an underlying timespan of a quarter note that is syncopated (compare the transcription below the score with the eighth-level reduction).[22]

---

[20] Ibid., p. 66. As discussed in chapter 1, the recurrence of adjacent isochronous timespans results in a pulse stream that is associated with a specific speed. It follows that each metric level will be associated with a set of speeds.

[21] Ibid., pp. 66-68. I will use the expression "mensurally determinate" (borrowed from Hasty, *Meter as Rhythm*, cited above).

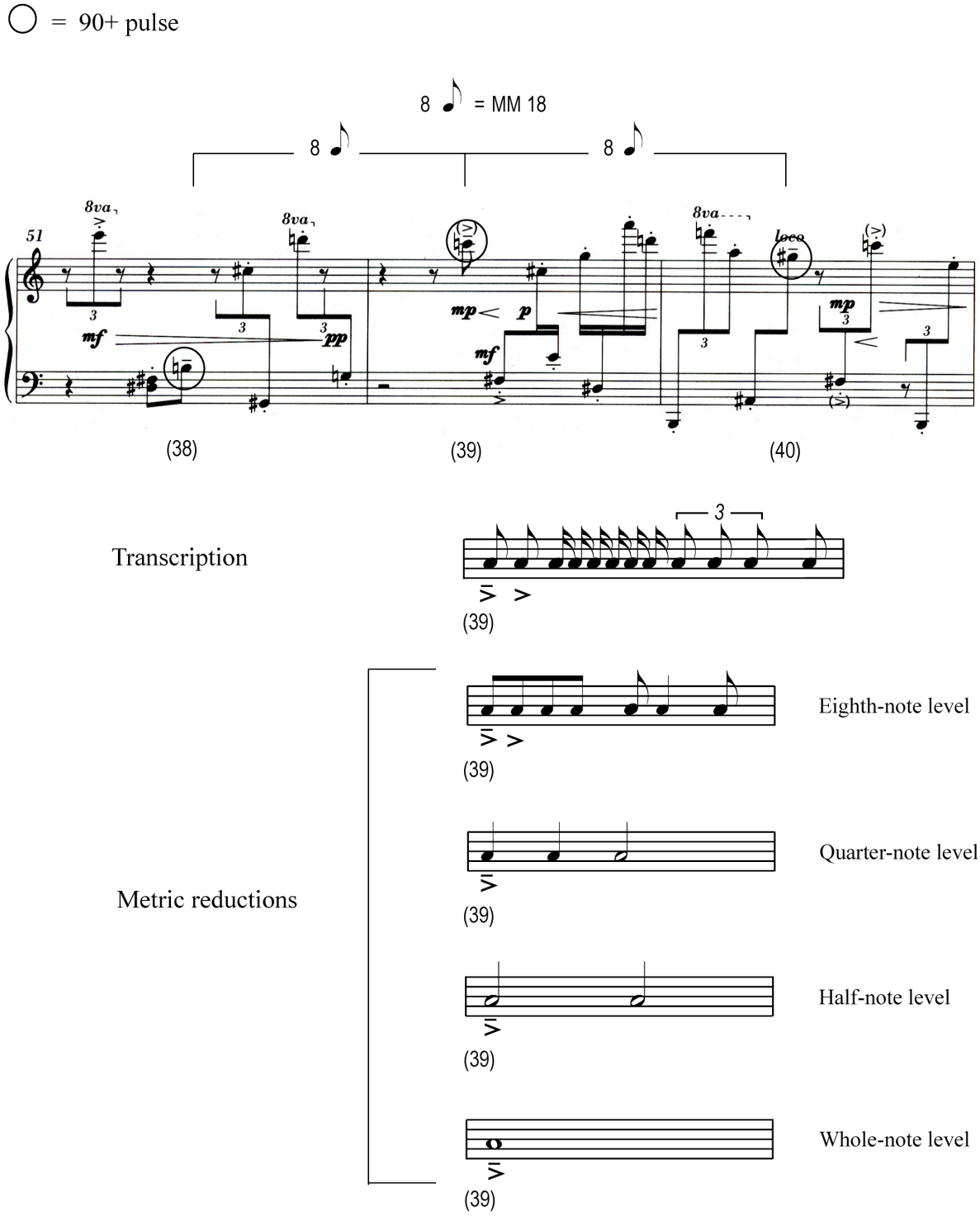
[22] If one is not relying on isochronous beats, this measure could also be interpreted as 5+3 eighths (not syncopated); this case, the three triplet eighths would correspond to an irregular subdivision of the first part of the second beat (LS pattern). Similarly, the first beat could be

Thus, although pulse-measure 39 exhibits a relatively low degree of metrical ambiguity, the presence of triplet eighths throughout the passage from which this excerpt is taken suggests an alternative middleground unit that is non-isochronous with the eighth-note level pulse unit and would result in a non-synchronous metric interpretation (the resulting syncopated quarter-note middleground pulse actually corresponds to the notated quarter beat). Depending on various factors, including pitch context, streaming, salience and the structural levels involved, these expectations will be realized in different ways. As we will see in chapter 3, in Carter's music, polymetric potential can be understood to stem from the systematic parsing of concurrent non-isochronous pulsations into non-synchronized streams of smaller timespans that are themselves subject to form larger patterns that give rise to competing metrical projections.

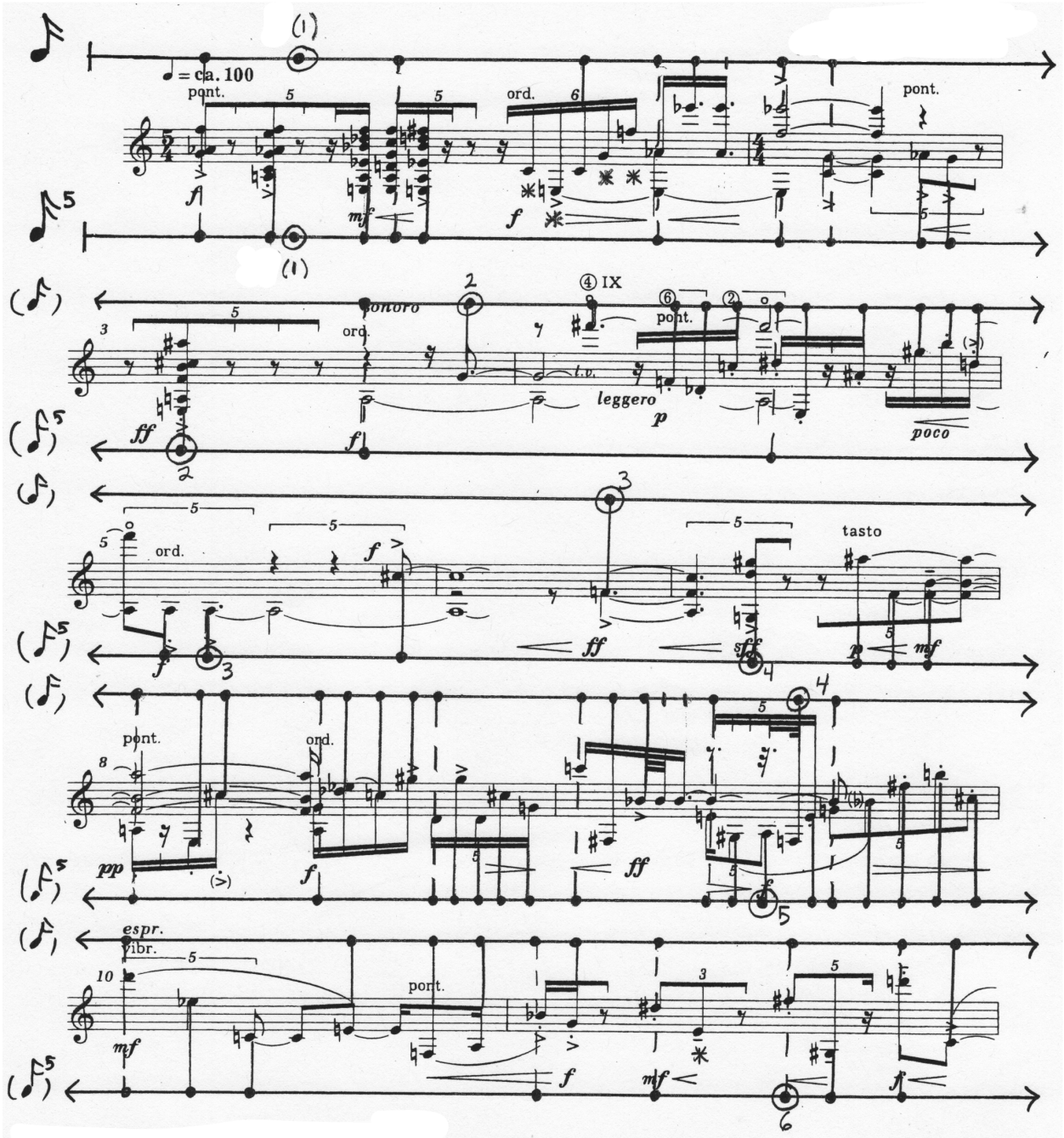
---

further subdivided into 2+3 eighths, the second part of which is further subdivided by 6 sixteenths.

Example 2.1: Interaction of background pulse and foreground rhythmic activity, *90+*, pulse-measures 39-41 (mm. 51-53)



Example 2.2: Mapping of surface events based on the two elementary pulse streams, *Changes* for guitar (1983), mm. 1-11 (upper-time line EPU = sixteenth; lower time-line EPU = quintuplet sixteenth)



2.2.2 Foreground

The lower level (and faster) pulse stream that results from the subdivision of each background (slow) pulsation into a series of isochronous timespans can be understood to form the *elementary pulse stream*; its related durational unit, the *elementary pulse unit* (EPU) acts as a referential timespan unit at the foreground level and the duration of pulse-measures will thus be expressed in number of elementary pulse units (epu).[23] The association of each stream with a specific durational unit (and notational value) will also become useful in the investigation of the musical surface. In **example 2.2**, the beginning of most surface events has been traced back to a corresponding elementary pulse stream: the upper time-line represents the sixteenth-based pulse stream, the lower time-line represents the quintuplet sixteenth-based pulse stream, and background pulsations have been circled and numbered.[24] Non-synchronous events, that is, events whose initiating time-points do not belong to either of the elementary pulse streams and thus whose inter-onset intervals cannot be measured in terms of elementary pulse units, have been marked with an asterisk (for example, see mm. 1 and 11).[25] It should be noted that the consideration of a meter arising from the interaction of

---

[23] As with the background pulse unit, the lower-case format (epu) will be used throughout as measurement unit; the upper-case format (EPU) will be used in the tables and figures to refer to the type of pulse unit represented.

[24] The location of the background pulsations are based on the data provided by Link in "Long-Range Polyrhythm," p. 118 (presented here as **table 2.1**). Once the notated duration of the background pulsations has been determined for a given tempo, the specific location of these pulsations in relation to the musical surface can be easily determined by trial and error; the irregularity of the surface rhythmic patterns usually ensures that there will be only one possible match for any extended passage.

[25] Although the initiating time-point of the fifth Ab-Eb that occurs at the end of m. 1 corresponds to an "offbeat" articulation, it is unlikely that it is meant to be experienced as syncopation. Rather, it yields a *ritenuto* effect as compared with the preceding series of

background and foreground stands in opposition to Yeston, who contends that "it would seem that a meter arising from the interaction of these two levels would not be highly significant, since background events are so widely spaced."[26] However, it would appear that in Carter's later works, this interaction is in fact significant, and that the specific ways in which background and foreground interact play an important part in shaping the expectations of the performer/listener. This particular point will be demonstrated in chapter 3 where the interaction of the flute and clarinet parts in *Esprit rude/Esprit doux I* (1985) is examined.

While the elementary pulse unit (EPU) is the smallest referential unit within one slow structural pulse stream, the presence of two contrasting elementary pulse streams calls for an additional, albeit abstract, durational unit, the *subpulse unit* (SPU), which allows surface events that are generated by contrasting elementary pulse streams to be measured according to a common unit (i.e., spu).[27] The number of subdivisions necessary to fully coordinate contrasting elementary pulse streams in a polyrhythmic cycle, or its *subpulse density*, corresponds to the smallest common denominator of the pulsation total for each pulse stream.[28] Given a foreground polyrhythm of 5:4 with a coincidence point every quarter beat,

---

sextuplet sixteenths, i.e., a temporary and highly localized surface deceleration from MM 600 to about MM 267, which is mediated by a sixteenth note (MM 400).

[26] Ibid., p. 67.

[27] As with background and elementary pulse units, the lower-case format (spu) will be used throughout as measurement unit, i.e., the upper-case format (SPU) will be used in the tables and figures to refer to the type of pulse unit represented; the same applies to middleground pulse units.

[28] My "subpulse density" is related to Link's "cyclic duration," which he defines as the total amount of time required for the completion of a polyrhythmic cycle ("Long-Range Polyrhythms," pp. 8-9). "Subpulse density" is given as a number of units, while "cyclic duration" is given as a chronometric time value (measured in minutes).

there will be 20spu per quarter beat, with each quintuplet sixteenth and sixteenth corresponding to 4spu and 5spu, respectively. Subpulse units are especially helpful because hybrid events, i.e., durations that result from the joining of contrasting elementary pulse units, frequently occur on the musical surface. For example, in **example 2.2**, the relative duration of the tones that join the two background pulses based on the quintuplet sixteenth elementary pulse unit in mm. 5-6, that is, segment <A, C#, F>, can only be ascertained using subpulse units (spu). While the A corresponds to a time-span of 14 quintuplet sixteenths, both the C# and F are made up of a combination of sixteenths and quintuplet sixteenths (C# = 2 quintuplet sixteenths + 10 sixteenths, and F = 6 sixteenths + 6 quintuplet sixteenths), resulting in the subpulse durational series <56, 58, 54>. Thus, although these events are generated from contrasting elementary pulse streams, they are nearly equivalent and are most likely to be perceived isochronously, and will give rise to a speed of about MM 36.[29]

2.2.3 Middleground

While the background and foreground levels can be, at least conceptually, easily identified, the question of the middleground is less straightforward. In order to define the middleground meaningfully for the task at hand, we will need to turn our attention to the

[29] This speed is calculated based on the average inter-onset interval of 56spu and notated tempo of quarter = 100 (100 × 20spu ÷ 56spu = MM 36). Given that each subpulse unit corresponds to a duration of 30ms (60s ÷ 2,000spu per quarter beat), the difference between successive events (2spu and 4spu, respectively) does not exceed 120ms, which at this speed is within the 10% margin usually tolerated in beat-finding experiments; see Joel S. Snyder and Carol L. Krumhansl, "Tapping to Ragtime: Cues to Pulse Finding," *Music Perception* 18/4 (2001): 457. Thus, although a listener could hear the successive rhythmic events as slightly longer or shorter, these differences are hypothetically not large enough to induce a perception of each event as categorically different from the other (i.e., as representative of a different beat value).

specific work under study. In *Changes*, a single structural polyrhythm serves as the backbone structure for the entire piece.[30] This structure results in the presence of two slow structural pulse streams, each with its corresponding elementary pulse stream. Let us now examine the properties of this long-range polyrhythm, as formalized by Link.

Table 2.1: Properties of the long-range polyrhythm for *Changes*[31]

Data	**Property**
P = 75:56	Ratio of the pulsation totals for the polyrhythmic stream
T = 100	Initial notated tempo (in ♩ beats per minute)
C = 6.3 minutes	Cyclic duration of the polyrhythm
S = 11 19/21:8 8/9	Speed ratio for the polyrhythmic stream (P/C)
B = 8.4 ♩.:11.25 ♩	Ratio of the number of notated beats between pulsations (T • C/P)

In **table 2.1**, various properties of the long-range polyrhythm and its relation to the notated score are given, including the number of pulsations in each stream, the notated beat and tempo, and the resulting total duration required for the two pulse streams to return to their initial position. Each pulse stream is also associated with a particular speed, and the specific proportional relationship of these two pulse streams could be understood as corresponding to an approximate ratio of 4:3. The use of decimals in the specific number of notated beats between pulsations betrays the internal make-up of the background pulse

[30] Link and Schiff both provide basic information on the structural polyrhythm for *Changes*; see Link, "Long-Range Polyrhythms," p. 118, and Schiff, *The Music*, p. 137. This topic will be discussed in more detail later.

[31] The information in this table is adapted from Link, "Long-Range Polyrhythms," pp. 117-118.

streams. The decimals point to the smallest unit (or elementary pulse unit) with which each background pulse stream is associated (see **table 2.2**).

Table 2.2: Notational Equivalences in *Changes*

Pulse Stream	Notated Tempo	Pulse Duration	Beat Equivalence	Pulse Speed
"X" stream	♩ = 100	42 $\overset{5}{\text{𝅘𝅥𝅯}}$	8 ♩ + 2 $\overset{5}{\text{𝅘𝅥𝅯}}$	MM 11 19/21
"O" stream	♩ = 100	45 𝅘𝅥𝅯	11 ♩ + 1 𝅘𝅥𝅯	MM 8 8/9

Here, the faster of the two pulse streams has been labeled the "X" stream while the slower one will be referred to as the "O" stream.[32] Each "X" pulse-measure corresponds to 42 quintuplet-sixteenth notes.[33] The quintuplet sixteenth divides each quarter into five equal parts and two quintuplet sixteenths correspond to 2/5 of a quarter note. Likewise, each "O" pulse-measure corresponds to 45 sixteenth notes, the sixteenth being the durational value that corresponds to 1/4 of a quarter note. Furthermore, each background pulsation will be labeled according to its generating pulse stream and with an Arabic numeral that corresponds to its temporal ordering. These labels will be used to designate the background pulsations as well as the time-spans (or pulse-measures) with which they are associated. For example, in bar-measure 9 (**ex. 2.2**), there are two background pulsations, each of which corresponds to the initiation of a pulse-measure, and which would be designated as X5 (lower time-line) and O4 (upper time-line).

---

[32] The practice of labeling two concurrent slow pulses with "X" and "O" is borrowed from Link, "Long-Range Polyrhythms."

[33] The use of a superscript Arabic numeral for irregular subdivisions of a basic durational value is a standard shorthand in the analytical literature on Carter's music; this type of notation will be used in all the tables, figures, and musical examples.

The internal make-up of the long-range polyrhythm is particularly suggestive of the meaning of the notated beat. In *Changes*, the quarter beat is the implied *coincidence point* of the 5:4 polyrhythm that results from the superposition of the two elementary pulse streams. A coincidence point can be formally defined as the co-occurrence of two pulsations from contrasting pulse streams; usually, coincidence points are considered only when they are realized (i.e., they correspond to actual sounded events), but in this dissertation I will show that implied coincidence points are also relevant, especially when considering polymetric potential. Thus, events that begin on a notated quarter beat, whether they are measured in terms of sixteenths, quintuplet sixteenths, or both, are nested in the concurrent pulse stream, and the quarter beat can be understood as a manifestation of their *synchronicity*, i.e., the property by which event onsets are conceptualized as operating within the same temporal layer. (In **example 2.2**, the onset of nested events is joined by a vertical dotted line as opposed to a solid line, which is used for independent events.) Consequently, the quarter beat functions as a low-level *middleground pulse unit* (MPU or mpu) and the bar-measure as a higher-level one, although other temporary middleground units are likely to emerge based on specific durational and accentual patterns exhibited by local events, especially given the fact that the quarter beat is still quite removed from the background. The formula for determining the resultant speed of a background pulse stream can thus be reformulated as shown in **figure 2.1**.

Figure 2.1: Formula for the resultant speed of a background pulse stream

Resultant speed (S) = notated tempo (T) × middleground pulse unit (MPU) ÷ background pulse unit (BPU), where each pulse unit is expressed as a number of elementary pulse units (epu).

Thus, $S = T \bullet MPU_{epu} / BPU_{epu}$

Since the actual metrical significance of the quarter-note beat (and the whole-note hyperbeat associated with the bar line) depends on the specific design of the musical surface, and given the specific relationship of the two elementary pulse streams (the fact that there is an implied coincidence between them every four sixteenths and five quintuplet sixteenths), a high degree of *nestedness*, i.e. the property by which an event onset belongs to two or more pulse streams, can be understood as a factor leading to quarter-beat (and possibly bar-measure) mensural determinacy, although such determination will depend on the specific placement and grouping of the nested events. This particular property will be further explored in chapter 3.

2.3 Compositional Strategies and Interaction of Structural Levels

One aspect of Carter's rhythmic language, that is, his use of metric (or tempo) modulation, needs to be addressed before our attention may be turned to a more detailed assessment of Carter's compositional strategies in dealing with structural levels and their interaction, especially given the conventional association of beat and tempo, as specified by tempo markings that give the metronomic speed of the beat prescribed by the time signature. From a psychological point of view, "[a] sense of tempo and motion is a hierarchically emergent property of the musical surface, and not simply a product of note-to-note

transitions."[34] Whether the prescribed tempo will be perceived as *the* tempo in a specific performance of the piece is a question that will not be formally addressed in this dissertation. Nevertheless, it is an issue that has some bearing on the topic at hand, especially given the fact that the meaning of the notated beat is not to be assumed.

2.3.1 Metrical Modulation Re-visited

The technique of metrical modulation is Carter's most widely recognized rhythmic idiom, although its implications and effects are yet to be fully understood.[35] The interest of Carter in building a musical discourse that presents a series of contrasting textures that are nonetheless related in some consistent way appears to be the basic aesthetic motivation behind his use of this technique. In his essay "Music and the Time Screen" (1976), Carter's describes "metric" modulation:

> as a mode of proceeding smoothly or abruptly from one speed to another and as a formal device to isolate one section from another. Generally, these work together, for very often a new section with a different speed and character starts while another layer continues at the same speed.[36]

That formal design does not prevent the continued relevance of some underlying relationship is attested by the second part of the above quote. Thus, in Carter's later works, while the notation of the durational values associated with a structural pulse stream is at times affected

---

[34] Justin London, "Rhythm," in *Grove Music Online*, ed. L. Macy (Accessed November 19, 2006), <http://www.grovemusic.com>, §II: 5.

[35] For a brief but instructive discussion, see Schiff, *The Music*, pp. 41-43; for an overview of the use of this technique by Carter see Noubel, *Elliott Carter*, pp. 107–10; for a more detailed discussion and analytical applications, see George Peter Tingley, "Metric Modulation and Elliott Carter's First String Quartet," *Indiana Theory Review* 4/3 (1982): 3–11.

[36] Carter, *Collected Essays*, p. 273. Carter uses "metric" rather than "metrical."

by a metrical modulation, its actual speed is usually unchanged, and if so, it changes only imperceptibly.[37] This would seem to hold true even in works based on a long-range polyrhythm, where two structural pulse streams remain constant.

The characterization of these modulations as "metrical" (or "metric") stands from its frequent correlation with a change of time signature and/or re-notation of an equivalent beat or sub-beat unit, as in [←𝅗𝅥 = ♩→]. On the other hand, the expression "tempo modulation" de-emphasizes the technical aspect of metrical modulations and underlines the effect by which the musical surface gives way to the perception of a proportionally related speed for the prevalent pulse level. Although the two processes are related, *one may occur without the other*. For example, if the notational equivalence given above is correlated with a time signature change from 2/2 to 2/4, no change of tempo will follow (based on beat level notational equivalence). Alternatively, if it is associated with a change from 2/2 to 6/4, the new tempo will be slower than the old one in a ratio of 3:1 (given the beat to sub-beat notational equivalence and the ratio of the number of subdivisions in each beat). While both of these perspectives may be relevant, the most significant effect of metrical modulations in *90+*, *Changes*, and many other works by Carter is not the use of different time signatures or metronomic speeds associated with the notated beat, but rather a change in the interaction between structural pulse stream(s) and musical surface. As Schiff notes (in relation to Carter's First Quartet, composed in 1950 and revised in 1994), "tempo modulations are often used to continue one strand of the polyphony *while enabling its combination with ideas at*

[37] In chapter 1, we have seen how the structural pulse in *90+* keeps a constant speed of MM 18 despite the surface tempo changes suggested by the metric modulations. Schiff also demonstrates this principle using the oboe solo in *Sandpiper* as example; see Schiff, *The Music*, p. 43.

different speeds [italics added]."[38] In other words, because of the newly available durational values, the texture can change and the general perception of tempo relations will be influenced by those metronomic speeds favored by the newly available durational units.

In order to demonstrate these different viewpoints, let us take a closer look at the overall structure of *Changes* from the perspective of metrical modulation.

2.3.1.1 Background Continuity

A formal chart of *Changes* will help identify and contextualize the use of metric modulation (see **table 2.3**). This chart is offered as the product of a potential segmentation rather than a definitive one.[39] There are three explicit metrical modulations in *Changes*. Successive tempi will be designated by the upper-case letter "T" followed by the subscript Arabic number corresponding to its temporally ordered first appearance. The modulation from T_1 to T_2 (quarter = ca. MM 100 → eighth = ca. MM 125) corresponds to the "outburst," the dramatic climax of the piece, which is situated at the end of the Scherzando (bar-measure 110; see **example 2.3**).[40] The second (T_2 → T_1) marks the return of the original tempo at the beginning of Episode 4 and also corresponds to background pulsation O41 (bar-measure 114, beat 3; see **example 2.4**).[41] Finally, the third modulation from T_1 to T_3 (quarter = ca. MM 100 → quarter = MM 67) initiates the Epilogue (bar-measure 131; see **example**

---

[38] Schiff, *The Music*, p. 41.

[39] For an alternative segmentation, see Capuzzo, "Variety Within Unity," pp. 43-44.

[40] The label "outburst" is borrowed from Capuzzo; note that despite its brevity, Capuzzo presents this passage as a separate section.

[41] The actual beginning of Episode 4 is coordinated with the background pulsation X54 for reasons that will be discussed in subsection 2.3.2.2. The *scherzando* markings at bar-

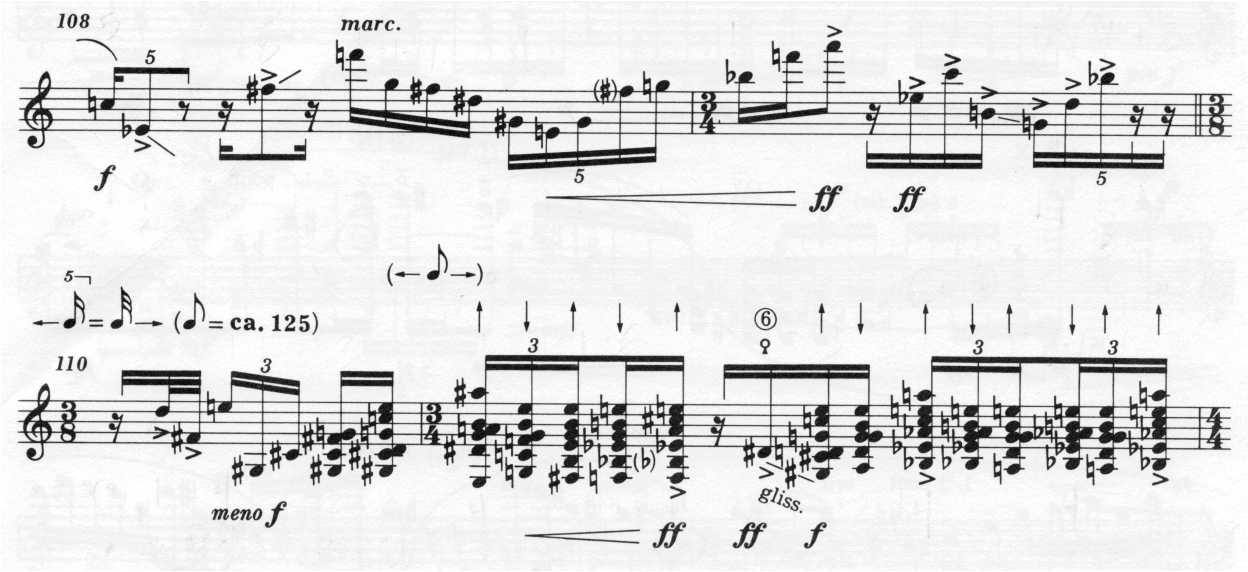
2.5). Of these, only the first and last are accompanied by an explicit sub-beat or beat equivalence, i.e., [← ♪5 = 𝅘𝅥𝅯 →] in the transition from T_1 to T_2, and [← ♩. = ♩→] in the transition from T_1 to T_3.

Table 2.3: Formal Chart of *Changes*

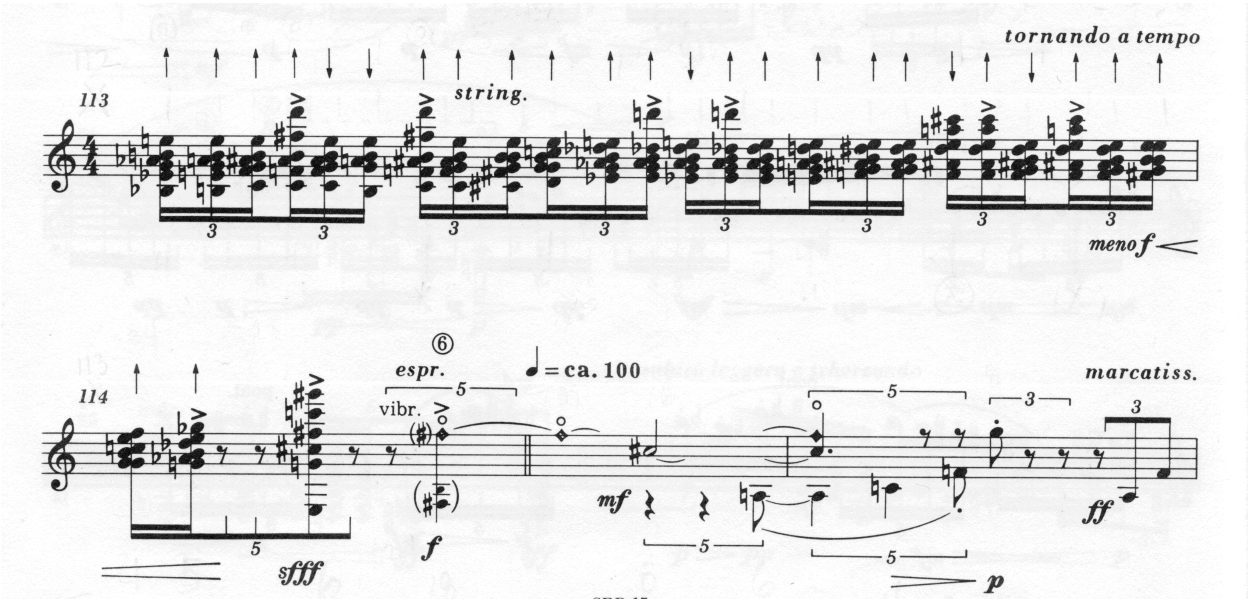
Section	Notated Tempo	Bar-Measures	Pulse-Measures "X"	Pulse-Measures "O"
Opening	T_1: ♩ = ca. 100	1-14	1-7	1-5
Episode 1	-	14-27	7-13	5-10
Episode 2	-	27-53	13-25	10-19
Episode 3	-	53-89	26-42	19-32
Scherzando	*subito leggero e scherzando*	89-113	43-54	32-40
-	(*scherzando*)	(91)	-	(33)
("Outburst")	(T_2: ♪ = ca. 125)	(110-113)	(52-54)	(39-40)
-	*(stringendo… tornado a tempo)*	(112)	-	-
Episode 4	T_1: ♩ = ca. 100	114-130	54-62	40-46
-	*(allargando… a tempo)*	(124-125)	-	-
Epilogue	Lento tranquillo T_3: ♩ = 67	131-150	62-76	46-57

measures 89 and 91 (included in table 2.1, second column), although not directly related to tempo, do relate to questions of segmentation, more specifically to the relation between the formal design and background pulsations X43 and O33; this interesting correlation will also be discussed in subsection 2.3.2.2.

Example 2.3: Metrical modulation from T_1 to T_2, *Changes*, mm. 108-111

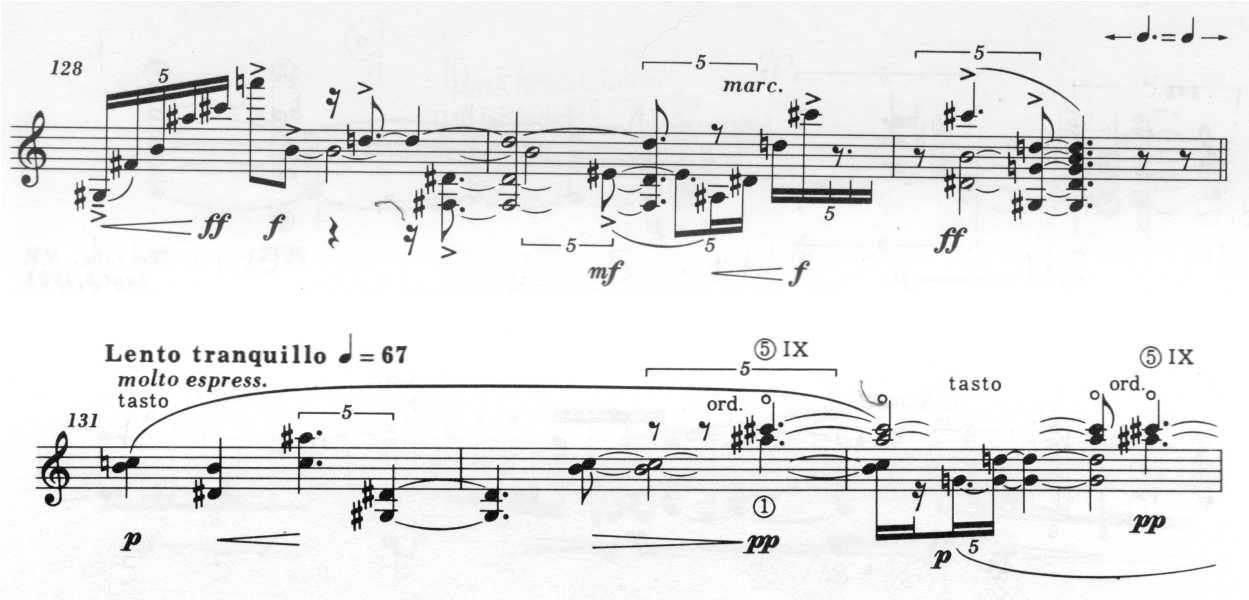


Example 2.4: Metrical modulation from T_2 to T_1, *Changes*, mm. 112-115 (corrected)[42]



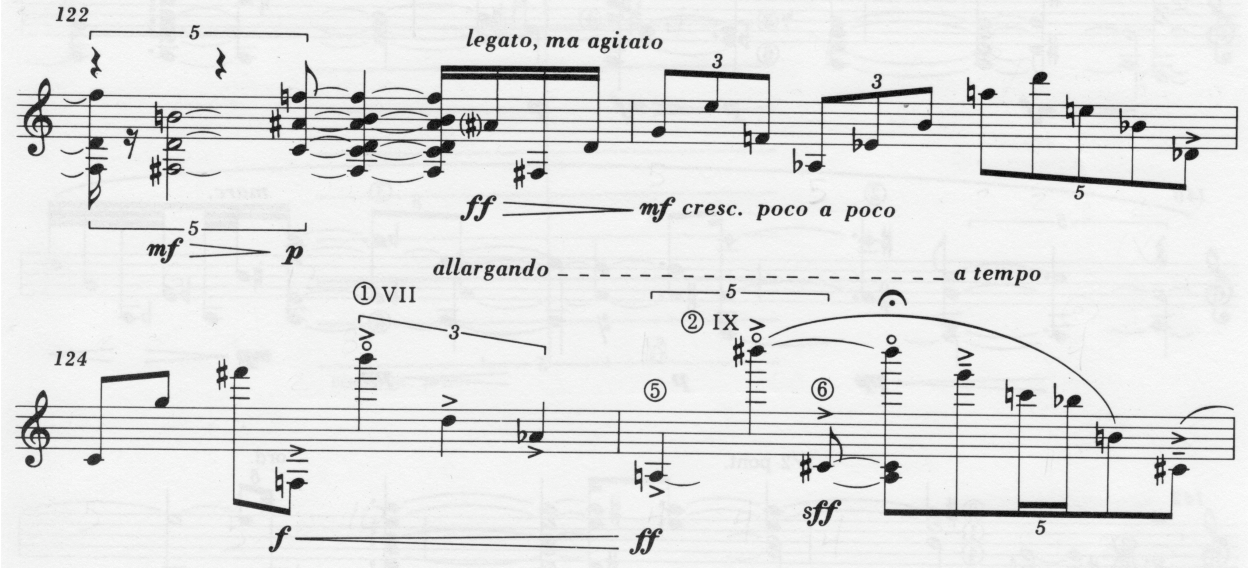
[42] It should be noted that in the published score (Carter, *Changes*, p. 10), there are two erroneously marked measure numbers, that is, bar-measures 113 and 114 should read as 112 and 113, respectively; these "real" measures are those used in this dissertation.

Example 2.5: Metrical modulation from T_1 to T_3, *Changes*, mm. 128-133



While the first two tempo changes are clearly "metric," i.e., they both involve a change of time signature and a re-notation of the sub-beat level (explicit or implicit), the third one (T_1 → T_3) would seem to be only a matter of tempo: the time signature is unchanged and despite the equivalence given, i.e., [← ♩. = ♩→], the dotted quarter has no perceptible influence in the measures preceding the modulation. In addition to the metrical modulations, there are two other more temporary tempo fluctuations that occur; these are included in the formal chart (see **table 2.3**, bar-measures 112-114 and 124-125). The first of these achieves a smooth transition from the flamboyant "outburst" to the more subdued episode that follows (**ex. 2.4**) while the second enhances a lyrical moment (see **example 2.6**).

Example 2.6: Expressive tempo fluctuation, *Changes*, mm. 122-125



Despite the metrical modulations, the background pulse streams and their associated speeds remain mostly unchanged, with the exception of the "outburst." For each notated tempo, background pulse-measure equivalences are provided (see **tables 2.4** and **2.5**). Interestingly, although the sub-beat equivalence [← $\overset{5}{\text{♪}}$ = ♬ →] that accompanies the metrical modulation from T_1 and T_2 theoretically allows for the "X" pulse stream to remain constant (**table 2.4**), this is not what happens in the music, where each background pulsation is actually extended, as if subjected to a *ritardando*.[43] The "O" stream regularity (**table 2.5**) is also theoretically insured by an implicit sub-beat near equivalence, i.e., [←♪ = $\overset{3}{\text{♪}}$ →],

[43] In his brief discussion of anomalies in the long-range polyrhythm of *Changes*, Link notes:

> In an earlier version the dramatic chords beginning in m. 110 were much faster, but Carter rewrote the passage because Starobin felt the chords could not be negotiated effectively at the original tempo.

These comments are based on letters exchanged between Carter and the piece's dedicatee, David Starobin (see "Long-Range Polyrhythms," p. 112). It may be inferred from that comment that it is at least possible that in earlier versions the background pulse streams were not altered, but sufficient evidence is lacking at the moment. (A more detailed discussion of

in which the clock time value of the elementary pulse unit is modified from 150 milliseconds to 160 milliseconds (a potentially imperceptible difference).[44]

Table 2.4: "X" stream notational equivalences and resultant metronomic speeds

Tempo Area	Notated Tempo	Pulse Duration	Beat Equivalence	Pulse Speed
T_1	♩ = 100	42 ♬ 5	8 ♩ + 2 ♬ 5	MM 11 19/21
T_2	♪ = 125	42 𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯	10 ♪ + 2 𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯	MM 11 19/21
T_3	♩ = 67	28 ♬ 5	5 ♩ + 3 ♬ 5	MM 11 27/28

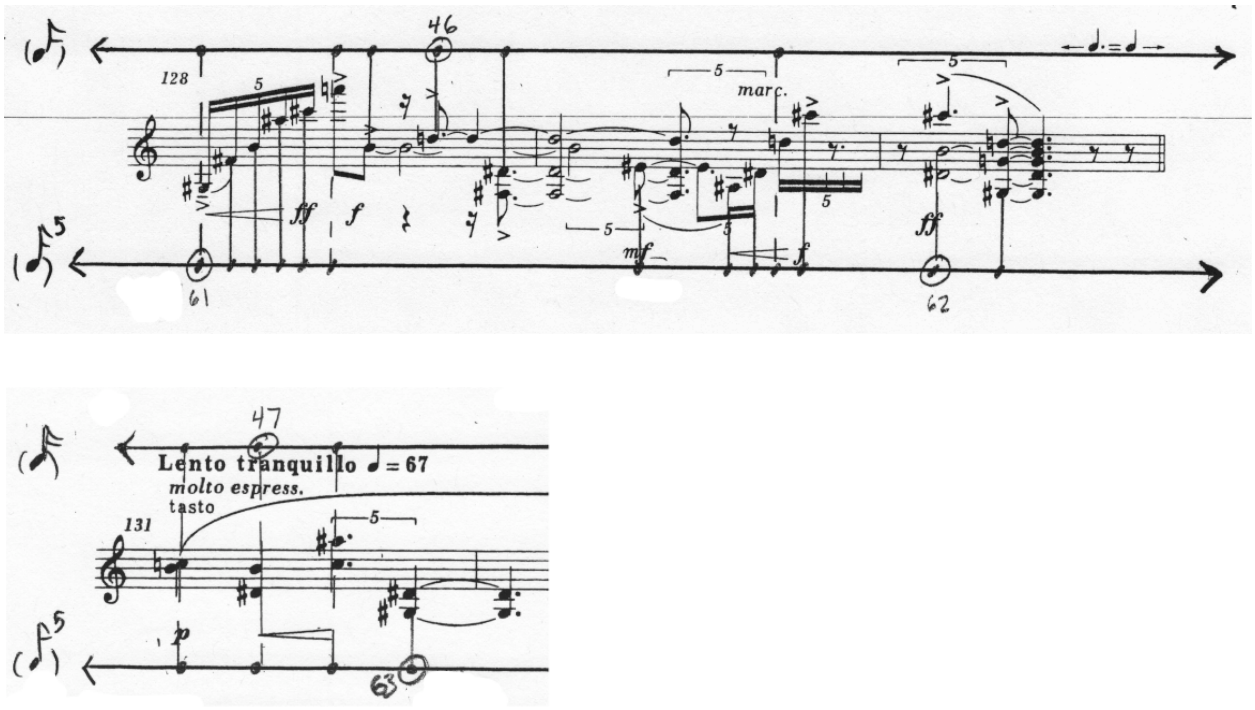
Table 2.5: "O" stream notational equivalences and resultant metronomic speeds

Tempo Area	Notated Tempo	Pulse Duration	Beat Equivalence	Pulse Speed
T_1	♩ = 100	45 ♬	11 ♩ + 1 ♬	MM 8 8/9
T_2	♪ = 125	42 ♬ 3	14 ♪	MM 8 13/14
T_3	♩ = 67	30 ♬	7 ♩ + 2 ♬	MM 8 14/15

background irregularities, including those preceding this section, is provided in subsection 2.3.2.2; the internal make-up of the corresponding background pulse-measures can be observed in the polarized reductions in chapter 3, section 3.2).

[44] At T_1, the resultant speed of a stream of sixteenths is MM 400 and the duration of one sixteenth is 60s/400bpm (0.15s or 150ms). At T_2, the resultant speed of a stream of triplet sixteenths is MM 375 and the duration a one triplet sixteenth is 60s/375bpm (0.16s or 160ms).

Example 2.7: Pivot pulse-measures O46 and X62, *Changes*, mm. 128-131



All metrical modulations in this piece occur in the middle of pulse-measures, requiring an adjustment of the notational value of the *pivot* pulse-measures (i.e., the pulse-measures in which the metrical modulation takes place) if the actual speed of the background pulse streams is to be preserved.[45] This is exactly what happens in bar-measures 128-131, which contains the pivot pulse-measures O46 and X62 (see **example 2.7**). Based on the 3:2 ratio between notated tempi, the new tempo requires only 28 quintuplet sixteenths and 30 sixteenths to yield the almost equivalent background speeds. Consequently, pulse-measure O46 has a combined notational value of 39 sixteenths at quarter = MM 100 plus 4 sixteenths at quarter = MM 67 (6epu @ T_1 = 4epu @ T_3) and pulse-measure X62 is notated as 18

[45] The plural is used here because *Changes* is based on a background polyrhythm and thus a metric modulation at the surface will usually affect the notated timespan associated with both background pulse streams, unless it occurs in synchrony with one of the background pulsations.

quintuplet sixteenths at quarter = MM 100 plus 16 quintuplet sixteenths at quarter = MM 67 (24epu @ T_1 = 16epu @ T_3). Hence, in this case the "metric" modulation itself pertains only to the background level, while the "tempo" modulation affects the middleground and foreground levels.

The example just discussed shows that the composer went to some pains to preserve the speed of the background pulse streams while achieving a metrical modulation. This is not to say that there are no changes in at least some of the perceived speeds, or that these are not part of the intended effect. The "outburst" does sound faster, especially if one's attention is placed on the strumming chords, and the build-up in activity to it strongly suggests that it functions as a climactic moment, a turning point in the narrative after which the characters' interaction, as represented by the two elementary pulse streams, is transformed (**ex. 2.4**). Similarly, the overall effect of the Epilogue is one of a slowing down of the beat, the perception of which is encouraged by the series of quarter notes at the beginning, thus bringing the piece to a peaceful conclusion (**ex. 2.5**). The point to be made here is that when considering the texture and emerging speeds, the focus cannot be put only on the middleground or foreground levels (here, the beat and sub-beat). Not only should the background be considered, but the interaction of these levels as well. To do so is to acknowledge Carter's compositional process as well as his aesthetic goals, i.e., the careful crafting of a discourse that joins together surface discontinuities and background continuity. The next subsection will further illustrate this point.

2.3.1.2 Surface Discontinuities

In discussing rhythmic organization, a consideration of the *tempo field* of a given piece is useful, which is to say, the collection of related metronomic speeds available in a given piece. Let us now take a look at Carter's tempo field for *Changes*, which includes the two elementary pulse levels (sixteenth and quintuplet sixteenth) as well as all other possible subdivisions (see **table 2.6**).[46] Most of the surface events not included here could be described as written-out "grace" notes, although this is not the only use of such notational values; for example, see the quintuplet thirty-second Eb in bar-measure 17 (**example 2.8**).[47] In **table 2.6**, speed equivalences are in shaded boxes. It should be noted that although metronomic speeds are provided for all notational values in each of the three notated tempi, not all of the available values are used in a given tempo; speeds that are not used have been put in parentheses. As we will see, some of the specific compositional choices made by Carter are particularly foretelling of the unfolding of this musical drama.

---

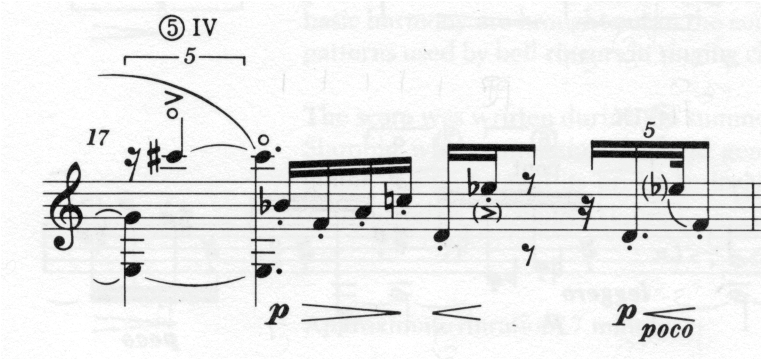
[46] Schiff, *The Music*, p. 43. It should be noted that the speed prescribed by the equivalence at T_3 is actually quarter = MM 66 2/3. In table 2.6, the notated tempo of quarter = MM 67 is given, but the resultant speeds have been calculated using the mathematically correct tempo. It should be noted that the actual differences in clock time between the pulse units represented in this table range from about 0.6ms to 1.5ms, with a difference between the dotted quarter at T_1 and the quarter at T_3 is about 4.5ms. These durations are well below observed perceptual thresholds, even if one considers larger differences resulting from the compounding of smaller values, e.g., the differences in duration for the background pulsations in the "X" and "O" pulse streams are about 59.5ms and 44.4ms. Furthermore, when considering the background pulse, the question of perceptible difference is complicated by the fact that the pulsations are subdivided into series of shorter and mostly non-isochronous durations. Perceptual issues are discussed in more detail in subsection 2.3.1.3.

[47] This issue will be clarified in the context of the interaction between background pulse streams and surface activity in subsection 2.3.2.2.

Table 2.6: Tempo Field of *Changes*

Notational Values	**Notated Tempi & Resultant Metronomic Speeds**		
	T_1: ♩ = 100	T_2: ♪ = 125	T_3: ♩ = 67
𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯	(MM 800)	MM 500	MM 533
𝅘𝅥𝅯3 or 𝅘𝅥𝅯6	MM 600	MM 375	(MM 400)
𝅘𝅥𝅯5	MM 500	MM 312.5	MM 333
𝅘𝅥𝅯	MM 400	MM 250	MM 267
♪5	MM 250	(MM 100)	MM 167
♪	MM 200	MM 125	MM 133
♪3	MM 300	(MM 187.5)	MM 200

Example 2.8: Written-out "grace" note, *Changes*, m. 17



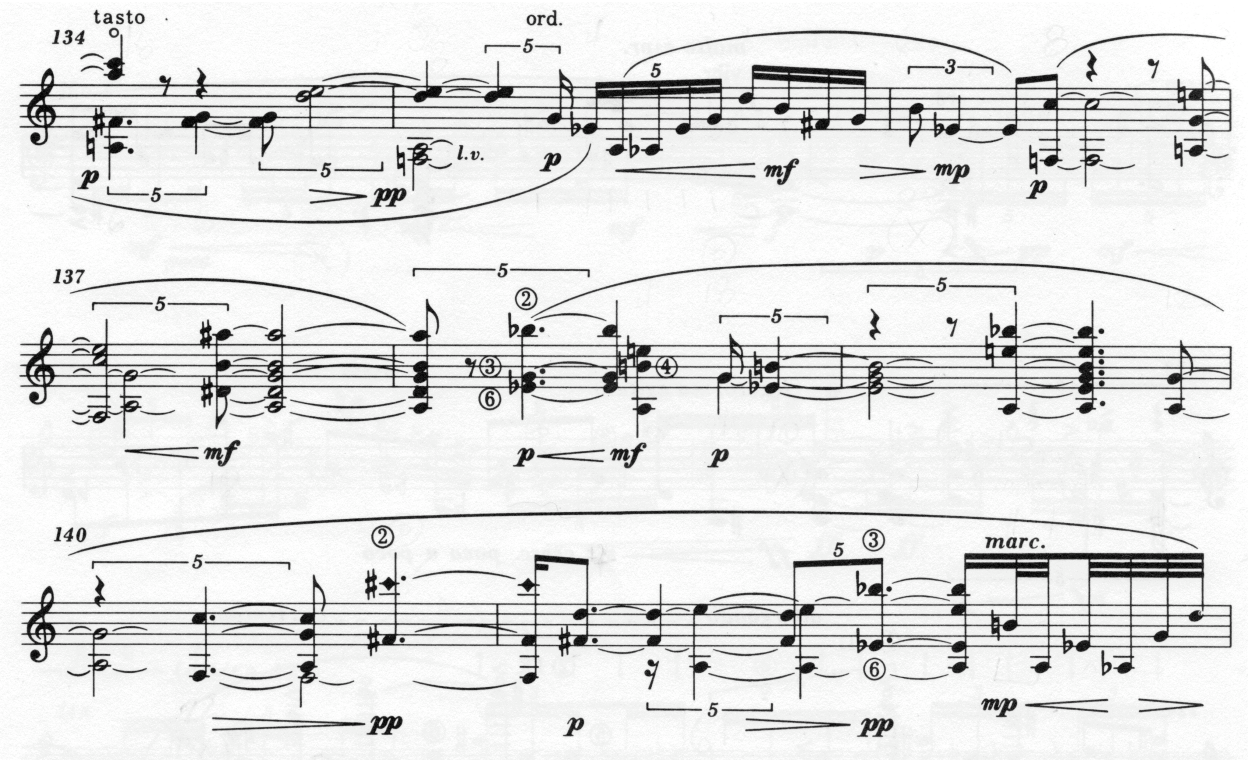
From a purely technical point of view, the tempo field makes evident the choice of elementary pulse units across tempi. From T_1 to T_2, there are two constant durational values (or speeds), i.e., [𝅘𝅥𝅯5 = 𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯] (chosen by Carter as the notated sub-beat equivalence) and [♪5 = 𝅘𝅥𝅯]. However, both of the new values are used only towards the beginning of the T_2 passage (**ex. 2.3**, bar-measures 110-111) and they represent only about 19% of the total number of events, i.e., nine out of the 48 events. The majority of the remaining events

articulate a speed of MM 375, which is most closely related to the sixteenth note at T_1 (MM 400), with an actual difference of merely 10 milliseconds per elementary pulse unit (epu). With the exception of the eighth note, which is present just below the surface (as a grouping pulse or low-level middleground beat), the influence of the remaining durational values on tempo perception is negligible, especially given the "*stringendo… tornado a tempo*" indication in bar-measures 112-113 (**ex. 2.4**).

In the transition from T_1 to T_3 (**ex. 2.5**), equivalence is achieved by [←♪ = ♪^{3}→] and [←♪ = ♪^{3}→]. Again, the presence of the new durational values on the surface is minimal, and the first equivalent durational value, i.e., the triplet or sextuplet-sixteenth note at MM 400, is not present at all. There are two gestures highlighting some of the smaller durational values, i.e., a series of 6 quintuplet sixteenths followed by 4 sixteenths in bar-measure 135 (MM 335 → MM 267) and a series of 6 thirty-seconds in bar-measure 141 (MM 533), but these flourishes seem to have little to do with the overall rhythmic landscape offered by this section (see **example 2.9**). Most of the texture is composed of longer values that are built on the two original elementary pulse streams (or their multiples), which now exhibit somewhat foreign tempi. In fact, the new quintuplet sixteenth is closest in speed to the triplet eighth of T_1 (a difference of 20 milliseconds), which was only infrequently present at the surface, and the new sixteenth is even closer to the quintuplet eighth of T_1 (a difference of 15 milliseconds), which functioned there as a low-level middleground pulse for the quintuplet sixteenth elementary pulse unit (EPU) associated with the "X" stream. That the new sixteenth notational value is closer in speed to the previous quintuplet sixteenth is an interesting reversal in speed identity, which suggests a reconciliation of the two competing

elementary streams, a narrative that is consistent with the dramatic trajectory set by the long-range polyrhythm (discussed in subsections 2.3.2.1 and 2.3.2.2).

Example 2.9: Surface rhythmic activity in Epilogue, *Changes*, mm. 134-141



These observations confirm that in his use of the tempo field for *Changes*, Carter maximizes surface contrast by giving precedence to those values that are less nearly equivalent. However, it would appear that beyond the constancy of the speeds generated by the background pulse streams, he also creates a link between contrasting textures by insuring that most of the available durational values are used, even if only infrequently. The passage in T_2 is the most restricted in that sense and thus the most contrasting. It uses only three out of the seven available speeds and the notated tempo and its triple subdivision is greatly emphasized. The resulting texture in the T_2 passage is one where levels have collapsed on

one another (the background pulsations are irregular at best), and metaphorically speaking, it offers reduced metrical "depth." Conversely, in the final section at T_3, there is a sense of space, the texture is less dense (even the chord tones are more widely spaced), and the background pulsations can interact more perceptibly with the middleground, while occasional flourishes recall the gestures of the first half of the work in a more relaxed manner.

Ultimately then, it would seem that Carter's use of metric (and tempo) modulation results in a kind of large-scale rubato of the type that involves "one part of the texture [that] remains constant," usually the one that is associated most closely with the expression of the metric structure, "while the other part pushes ahead and/or pulls behind."[48] This is also the view suggested by Hudson when discussing metric modulations in Carter.[49] By making a distinction between metric and tempo modulations, and by examining their use on separate levels, I have shown that there are many possible combinations in the modulation from one texture and its associated speeds to another. Furthermore, I believe that by examining the interaction of speeds rather than focusing only on the most perceptible tempo or the underlying constancy, we can come at a more refined understanding of formal and textural elements as well as their contribution to musical narrative.

---

[48] London, "Rhythm," §III: 7.

[49] Hudson associates this type of rubato with C.P.E. Bach and Chopin; see Richard Hudson, *Stolen Time: the History of Tempo Rubato* (Oxford: Clarendon Press, 1994), p. 419.

2.3.1.3 Perceptual Issues

Finally, the relation of these tempi with current perception and cognition studies should be discussed. As summarized by London, it would appear that *metrical entrainment*, i.e., the phenomenon by which a listener attunes to a regular pattern of time-points, can be found to occur when a listener is presented with a pulse with inter-onset intervals from 100 milliseconds up to about 5 or 6 seconds (i.e., about 10 to 600 pulsations per minute).[50] Furthermore, studies have shown that "100ms seems to be the threshold for reliable judgments of length."[51] The preferred *tactus*, i.e., the pulse level most likely to be acknowledged by sympathetic bodily responses, resides between MM 60 and 120, which corresponds to inter-onset intervals between 500 milliseconds and 1 second.[52] On the other hand, Clarke and Parncutt report that "[m]usical pulses are confined to a tempo range of roughly 30 to 300 beats per minute, or an inter-onset interval range of 200 milliseconds to 2 seconds."[53] While the latter's lower limit may correspond to a "metrical" pulse (i.e., a beat), and thus be based on a definition of meter as the sum of at least two isochronous time-spans, the upper limit of 2 seconds seem rather conservative. Clarke and Parncutt also note that when asked to tap "at equally spaced intervals" participants seem to prefer pulse streams with inter-onset intervals between 400 and 900 milliseconds (i.e., 150 to about 67 pulsations per

---

[50] London, *Hearing in Time*, p. 46.

[51] London, "Rhythm," §I: 2.

[52] London, *Hearing in Time*, p. 46.

[53] Carolyn Clarke and Richard Parncutt, "Psychology of Music," in *Grove Music Online*, ed. L. Macy (Accessed November 19, 2006), <http://www.grovemusic.com>, §II. 2.

minute), but that "[s]ensitivity to small changes in tempo is most acute in the range 300 to 800 milliseconds" (i.e., 200 to 75 pulsations per minute).

Given the tempo field provided above, we can see that in *Changes* most surface events are within the perceptible limits of a pulse, although many of these speeds fall outside of Clarke and Parncutt's most sensitive range for detecting changes in tempo, including the elementary pulse streams.[54] Thus, there would seem to be no perceptual issue with considering any of these notational (and durational) values as building blocks for meter, but whether fluctuations in durational values can be perceived across tempi is much more problematic. On one hand, the lack of reliability in judging lengths that are less than 100 milliseconds seems in favor of near-equivalences. On the other hand, the experimental evidence also suggests that a listener might not be able to differentiate between the two elementary pulse units since these yield an inter-onset difference of only 30 milliseconds. It may be that in the case of small subdivisions, the listener (and performer) uses a top-down approach, using a slower pulse level as ruler and a counting strategy to identify the subdivision being used. While it might be difficult (or even impossible) to identify smaller durational values in isolation, there would seem to be more hope for differentiating between faster and slower series of them, especially if rhythmic events are associated to each other by some other musical parameters such as timbre, dynamics, accentuation, and/or articulation.[55]

---

[54] This is an interesting issue that connects perception and performance, and raises questions on whether (and how) these values or speeds—which are often meaningful in Carter—can be made perceivable by the performers (through aural and/or visual cues). This topic and other similar issues will be addressed in chapter 6.

[55] London, "Rhythm," §I: 3.

The situation is also uncertain when considering the background pulse streams. The faster of the two background pulse streams ("X") appears to be already approaching the upper limit provided by London (MM 12 = 5 seconds) and the slower one ("O") exceeds it by about 700 milliseconds (MM 9 = about 6.7 seconds). And, both tempi are beyond the upper limit provided by Clarke and Parncutt. Commenting on experiments that involve performing series of beats, London notes:

> Interestingly, they reveal little effect of musical training on basic beat-keeping abilities. In more complex contexts, however, including the perception and performance of polyrhythms […] and in passages that require subdivision, musically skilled subjects perform significantly better than non-musicians.[56]

Whether the same principle would apply to tasks involving keeping track of larger recurring time-spans and polymeter is not clear.[57] Although the speeds observed in *Changes* are significantly slower than the preferred tactus revealed by perception studies (which could be understood as a conservative middleground), it is not unreasonable to suggest that with appropriate training, a listener could follow such a pulse. In any case, a definition of musical meter cannot be completely reliant on perception studies. Musical practices such as the counting strategies used for Indian *tãl* cycles can serve as witness to the many forms that meter can take.[58]

---

[56] Ibid.

[57] The preliminary results from the listening experiment presented in chapter 4 suggest that it does.

[58] When listening to a performance of this music, informed participants track the beats of these long cycles with patterned movements involving fingers, hands, and/or head. Martin Clayton provides an interesting discussion of this practice and its implications for a generalized theory of rhythm and meter in "Le mètre et le *tãl* dans la musique de l'Inde du Nord" in *Cahiers de Musiques Traditionelles* 10 (1997): 169-189 (an English version of this article is provided at http://www.open.ac.uk/Arts/music/mclayton/metre.htm [accessed August 5, 2008]). Clayton also explores this topic from a music-psychological perspective

Finally, it is important to note that in Carter's music the subpulse stream presents a challenge for direct perception and it is not clear at which point differences calculated in subpulse units (spu) can become perceptible; for example, at a tempo of quarter = MM 100 (as in *Changes)*, the speed resulting from a stream of subpulse units (quarter = 20spu) would correspond to 2,000 pulsations per minute, or $33^{1/3}$ pulsations per second, an inter-onset interval of 30ms, a periodicity that is currently recognized as below the threshold of rhythmic performance and perception.[59]

2.3.2 Long-Range Polyrhythms in Action

As noted in the introduction, the most notable development in Carter's later works is his use of a "giant polyrhythm" as the background structure of a piece.[60] These structures have been alternately referred to as "slow polyrhythms" and "long-range polyrhythms"; in this dissertation, I use the latter term, but I will refer to the individual streams of a long-range polyrhythm as *background pulse streams*.[61] Because each background pulse stream is associated with a faster and notationally distinct elementary pulse stream, a given long-range

---

using empirical methods and audio/video recordings of an expert performer and his audience; see "Time, Gesture and Attention in a *Khyãl* Performance," in *Asian Music* 38/2 (2007): 71-96 (this article is also available online at http://www.open.ac.uk/Arts/music/mclayton/khyal-annexe.htm [accessed August 5, 2008]). The performance studied in the latter article was based on a *tãl* cycle (*ektãl*) comprising 12 *mãtrãs* (time unit similar to a measure) subdivided into 4 pulses or "beats"; in the studied performance by Viswanath Shirodkar, the *tãl* cycle was associated with a chronometric duration of 35-48 seconds or 15-20 *mãtrãs* per minute (i.e., MM 15-20), speeds similar to Carter's background pulse units in *90+*, but faster than those for *Changes*.

[59] London, *Hearing in Time*, p. 27.

[60] Edwards, *Flawed Words*, p. 101.

[61] The expression "slow polyrhythm" is the one preferred by Schiff (1983 & 1998); as noted previously, the expression "long-range polyrhythm" is borrowed from Link.

polyrhythm will have implications for the kinds of surface proportions (or polyrhythms) available at higher levels. As a result, in Carter's later music, although the musical surface exhibits a high degree of irregularity, there is a sense in which discontinuous events relate to one another, a feature that is emphasized by the composer's reliance on metric modulations for the seamless juxtaposition of contrasting materials. Thus, Schiff states:

> Carter has said that he wanted to develop a more systematic way of approaching all the tempo relations in a piece, and also wanted to develop the structural equivalent of the slow phrase rhythms of tonal music. [62]

The most attractive and palpable effect of a long-range polyrhythm is the patterned movement of its competing pulsations, that is, its *proximity cycle*.[63] While a long-range polyrhythm might not be perceivable as a numerical entity, it is hypothesized that the listener will hear (or feel) the tension produced by the out-of-phase cycles and will become increasingly aware of the cycles moving toward and away from each other, and that these *convergence* and *divergence* phases will provide a sense of expectation.[64] The relative degree of perceptibility of such a cycle would seem to depend on a variety of factors, including speed and surface activity. Carter himself posits:

> It is *only* at more moderate and slow tempi that a large-ratio polyrhythm begins to have clearly tension-value, such that a polyrhythm like 71:49 starts off sounding like 3:2 but then clearly begins to get out of phase and then to get back into phase in a perceptually transparent way.[65]

---

[62] Schiff, *The Music of Elliott Carter*, p. 47. According to Schiff, his first use of this technique is found in the Introduction and Coda of the Double Concerto (see p. 46).

[63] Link, "Long-Range Polyrhythms," pp. 11-14; this phenomenon received its first formal treatment under the term "phase cycle" in Craig Alan Weston, "Inversion, Subversion, and Metaphor: Music and Text in Elliott Carter's *A Mirror on Which to Dwell*" (D.M.A. dissertation, University of Washington, 1992).

[64] The expressions "convergence phase" and "divergence phase" are also from Weston (1992).

[65] Edwards, *Flawed Words*, p. 112.

From Carter's comments and writings, it becomes evident that the composer not only believes that these structures are perceptible at least on some level, but also that their perception plays an important formal and expressive role in his music. Since the long-range polyrhythm comes with its own time-line associated with patterns of tension and release, it also provides a dramatic trajectory for the work it unfolds. For instance, in *Changes*, the end of the piece corresponds with the first coincidence point of the long-range polyrhythm, the initial one having been 'deleted' by a rest (**ex. 2.2**; this interpretation will be discussed in more detail in subsection 2.3.2.1). The proximity cycle is most readily observable in the Epilogue, where as noted previously, the texture is made thinner by the use of longer (or slower) values, and the re-notation of the pulse-measures into fewer elementary pulse units (through metric modulation) makes these background pulsations closer to the surface (refer back to **tables 2.4** and **2.5**). Given the relatively low number of events at the surface, not only are the near-coincidence points emphasized, but the relation of the complex polyrhythm (75:56) to its simpler model (4:3) also becomes more perceptible (see **figure 2.2** and **example 2.10**). It should also be noted that in the polyrhythmic sketch on graph paper for this piece, Carter marked the near-coincidences, providing evidence of their use by the composer.[66]

In **figure 2.2**, background pulsations (or pulse-measures) have been coordinated using a subpulse stream to allow for the consideration of various subdivisions (at T_3, "O" = 150spu and "X" = 112spu). Asterisks precede each of the near-coincidence points so that the progression of the convergence phase of the polyrhythm can be readily observed (note that

---

[66] This polyrhythmic sketch is also help at the Paul Sacher Foundation (accessed March 2008). The author is in possession of a partial transcription of this graph.

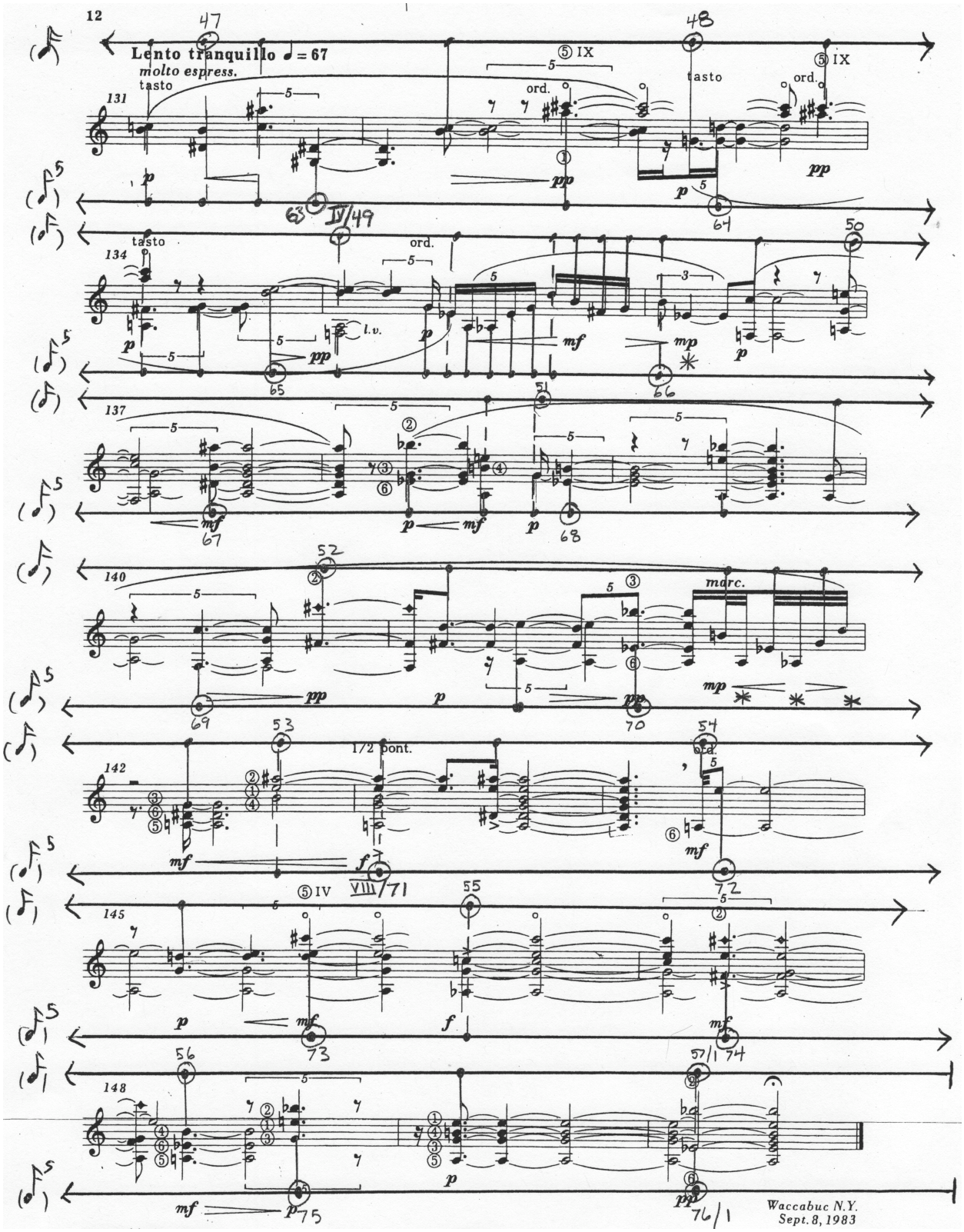
near-coincidence points are increasingly closer by 2spu, a progression of -2spu); in order to simplify the representation, events that do not belong to the background polyrhythm have not been included in the subpulse durational series (SPU). If each of these near-coincidence points is understood as a particularly accented event, it seems reasonable to posit that the two background pulsations can be perceived as one event, leading to the conceptualization of this passage as a slightly out-of-phase 4:3 polyrhythm leading to the final coincidence point; in this representation of the proximity cycle, alternative groups of corresponding pulsations are placed in shaded boxes to clarify this relationship. The subpulse unit value attached to each convergence point represents a timespan imbued with tension onto which the listener's attention can be directed, which holds potential for expectation building (and thus, tension resolution), and thus, this structure could also be thought of as a 4:3 hypermetrical polyrhythm.

When looking at the annotated score (**ex. 2.10**), this passage also illustrates the reason behind Carter's use of the very small subdivisions previously described as written-out "grace" notes. However, unlike grace notes, these values are exact subdivisions, carefully calculated so that the background pulsations can be located precisely where they would be expected. In the Epilogue, this is made especially clear if one observes the near-coincidence point O54→X72, which corresponds to a quintuplet thirty-second note (**ex. 2.10**, m. 144). A similar example is found in the Opening section at near-coincidence point X5→O4, which exhibits the use of a quintuplet sixty-fourth note (**ex. 2.2**, m. 9). Such notational oddities provide an instructive glimpse of Carter's compositional process. Thus, in Carter's notational practice, what appears like overly precise notation is in fact the result of an interaction between the surface and some deeper process.

Figure 2.2: Final convergence phase, *Changes*, O48→X64 to X76/O57 (mm. 133-150)

"O"	**48**			**49**		**50**		**51**			**52**		**53**		**54**			**55**		**56**		**57**
SPUs	***6**	112	32	80	70	42	108	***4**	112	34	78	72	40	110	***2**	112	36	76	74	38	112	***0**
"X"		**64**	**65**		**66**		**67**	**68**	**69**		**70**		**71**		**72**	**73**		**74**		**75**		**76**

Example 2.10: Epilogue, *Changes*, mm. 131-150 ("O" stream EPU = sixteenth; "X" stream EPU = quintuplet sixteenth)



The next subsections will focus on the treatment of the background polyrhythm as a whole, where traces of the composer's hand can also be found. More examples of interaction between the background and other structural levels will also be presented.

2.3.2.1 Background Irregularities: Proximity Charts and Polycyclic Spectra

In his study of long-range polyrhythms, Link discusses irregularities in Carter's treatment of abstract models.[67] He notes that it is not unusual for the composer to take liberties if aesthetic or practical reasons demand it (as was the case with the reworking of the "outburst" section discussed previously in subsection 2.3.1.1). In discussing the *Triple Duo* (also discussed by Link), Schiff comments:

> In practice Carter feels free to make the structural polyrhythm explicit at some points in a piece while not articulating it at others and he has also felt free to omit parts of the polyrhythm, as he did in the last section of the *Triple Duo*.[68]

In the case of *Changes*, both Link and Schiff refer to a cut in the long-range polyrhythm and a subsequent expansion at the request of the dedicatee. Link notes:

> In the finished score, pulsations in both streams disappear after m. 102, and are taken up again in m. 115, about five beats behind schedule. A second anomaly is the absence of an initial coincidence point, which should occur on the second beat of the first measure. The initial 5/4 meter signature—which affects only the first bar—hints at the reason. Had the first measure been 4/4, like most of the rest of the piece, the coincidence point would have been on the downbeat.[69]

Link's interpretation of the opening measure (previously shown in **example 2.2**) is convincing, especially in light of the fact of the dramatic trajectory, i.e., its stated dual

---

[67] Link, "Long-Range Polyrhythms," pp. 111-113.

[68] Schiff, *The Music*, p. 47.

[69] Link, "Long-Range Polyrhythms," p. 112.

character, intensifying conflict, the climactic "outburst," and the subsequent reconciliation and directed motion toward the harmony of a final coincidence expressed by the serene interval of a perfect twelfth (**ex. 2.10**). Putting the question of the genesis of the work aside, my analysis shows that the final result is even closer to the abstract model than the comments about the passage preceding the "outburst" might suggest. The use of *proximity charts*, i.e., graphic representations of the location of background pulsations within bar-measures, will help demonstrate this point.

In using the expression *polycyclic spectrum* to refer to simultaneous background pulse streams, I wish to emphasize the sense of directed motion that long-range polyrhythms exhibit.[70] The *abstract polycycle* (**figure 2.3**) presents us with an abstract realization of the full polyrhythmic cycle of 75:56 in an unchanged 4/4 meter, beginning and ending with a coincidence point. The *realized polycycle* (**figure 2.4**) is the representation of the background polyrhythm as found in the finished score. In all these representations, the vertical axis corresponds to the timespan of the bar-measure expressed as a number of subpulse units, i.e., one bar-measure (four quarters) is equivalent to 80spu. The two dimensional format is not entirely representative, as in reality time-point 80 in the subpulse stream is equal to time-point 0 (a three-dimensional representation would involve transforming this graph into a cylinder). The horizontal axis corresponds to the bar-measure unfolding of the piece in time.

---

[70] The term *polycycle* has also been used by Robison in his study of Birtwistle's *Carmen Arcadiae*, where these structures are compared to the large-scale polyrhythmic structures of isorhythmic motets; see Brian Robison, "*Carmen Arcadiae Mechanicae Perpetuum*: Towards a Methodology for Analyzing Harrison Birtwistle's Music Since 1977" (D.M.A. dissertation, Cornell University, 1999). Robison extends the definition of a *cycle* to include other types of recurring elements such as specific pitch, timbre, dynamic, etc.

Figure 2.3: Abstract polycycle, *Changes*, mm. 1-158

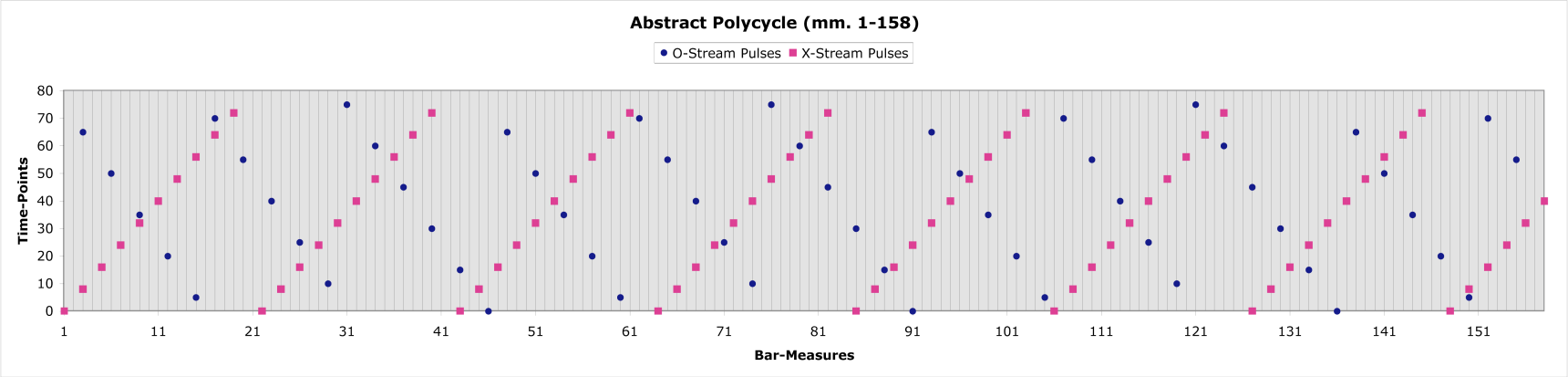
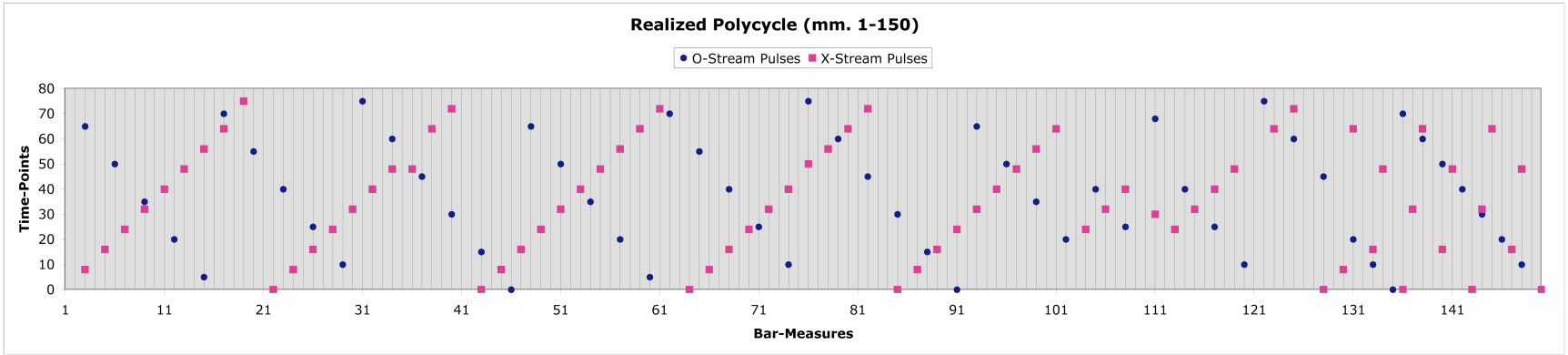


Figure 2.4: Realized polycycle, *Changes*, mm. 1-150



These representations of the polycyclic spectrum allow us to see that the realized version is very close to the abstract one. As proposed by Link, the main region of disruption begins after bar-measure 102. However, it seems to resume as early as bar-measure 113.[71] This disruption results in a displacement of one bar-measure (and not five measures "behind schedule" as suggested by Link). The length of the polycycle in terms of bar-measures also appears to be modified, but the "early" ending at bar-measure 150 instead of 158, where it would have occurred in mid-bar, is in fact due to the background metric modulation at the beginning of the Epilogue. Although in general the decision to produce a slowing down and a thinning-out of the texture through metric modulation can precede the choice of specific beat equivalences, it is interesting to consider the meaning that can be derived from the location of the unique coincidence point between the background pulse streams on the notated downbeat of the last bar-measure in this work. Did Carter want to end on a notated downbeat? Is this an example of rhetorical gesture signified by notation? Given the great care the composer takes in precisely notating various rhythmic relationships, it would seem fruitful to consider this element as more than a mere by-product of metric modulation. The sketches and earlier drafts, which indicate that the final metric modulation came rather late in the compositional process, support this interpretation.[72]

[71] As opposed to m. 115 as suggested by Link; it is possible that Link got led astray by the erroneous measure number (see **example 2.4** in subsection 2.3.1.1).

[72] These materials are held at the Paul Sacher Foundation and were accessed in March 2008.

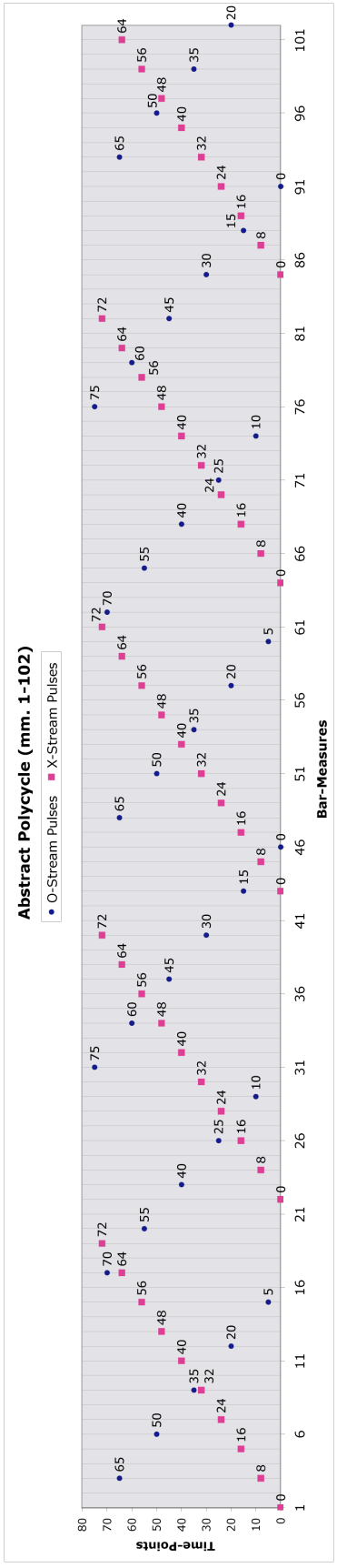


Figure 2.5: Abstract polycycle, *Changes*, mm. 1-102

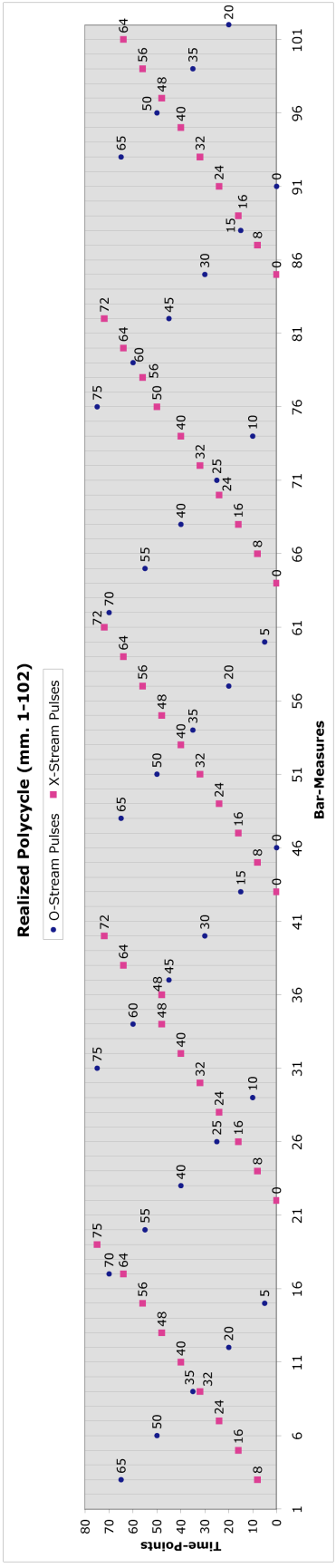


Figure 2.6: Realized polycycle, *Changes*, mm. 1-102

2.3.2.2 Analytical Details

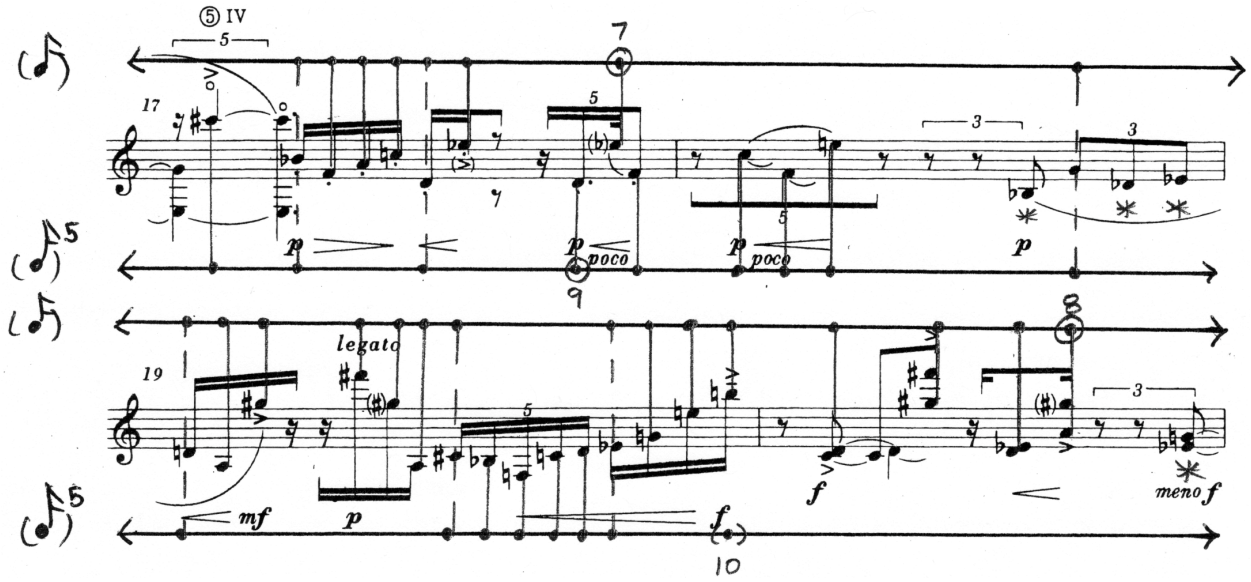
Let us now observe the differences between abstract and realized polycycles more closely. The realized polycycle for bar-measures 1-102 shows indeed very few irregularities (see **figures 2.5** and **2.6**). In the proximity charts, each background pulsation is labeled with the time-point where it appears. Apart from the initial disappearance of the original coincidence point mentioned above (annotated score shown as **ex. 2.2**), there are three irregularities in this section. In bar-measure 19, X10 is displaced by +3spu and finds itself integrated within the competing pulse stream (see **example 2.11**; in all the annotated score examples, irregular background pulsations are shown in parentheses instead of being circled).[73] Although the specific choice of durational value (i.e., sixteenth instead of quintuplet sixteenth) could be interpreted to result from practical considerations relating to ease of reading, the corresponding sub-phrase also marks a moment where the "O" stream appears to dominate.[74] In addition, the background pulsation is still accented and functions as a segment (or phrase) boundary, adding to its salience. Given the context, the sequence of sixteenths in m. 19 could be interpreted as a rhetorical gesture toward the next sub-phase, which is based on the sixteenth-note elementary pulse stream. In bar-measure 36, X18 is displaced by –8spu (-2epu), but remains within the quintuplet sixteenth pulse stream, a kind of rushing forward that occurs at a moment of increasing intensity in the musical discourse of the "flamboyant" character associated with the quintuplet sixteenth pulse stream and the "strumming" chords. Finally, in bar-measure 76, X37 is displaced by +2spu and also

[73] The positive sign denotes a displacement to the right on the time-line, i.e., it is "late" (3spu = 3/20epu = quintuplet dotted thirty-second).

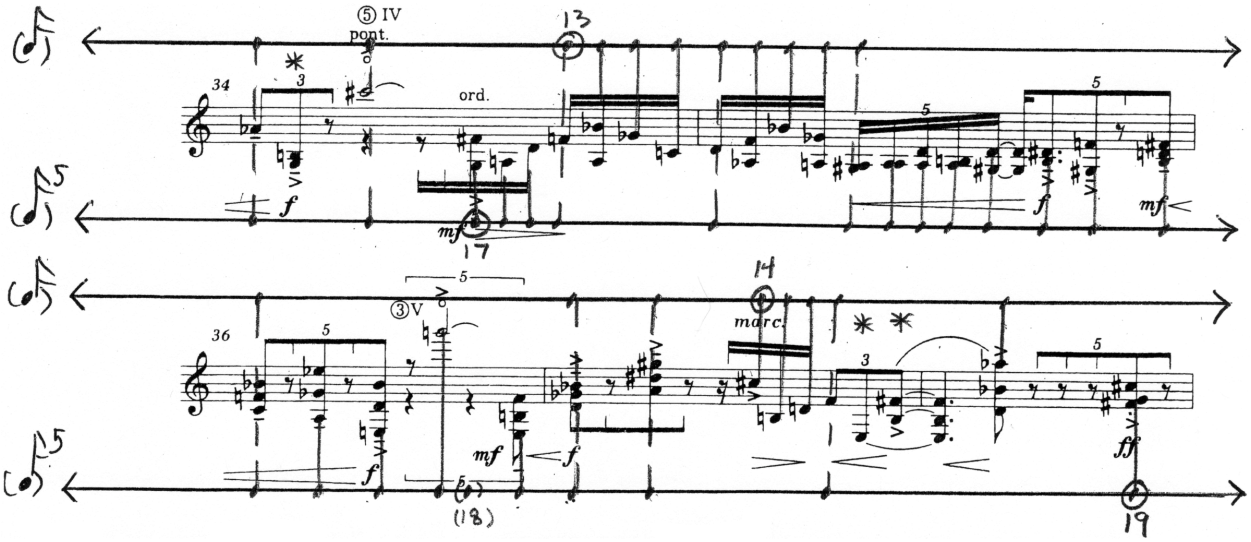
[74] Stream dominance, or *polarity*, will be discussed at more length in chapter 3, along with a presentation of the segmentation.

integrated into the competing stream (see **example 2.12**). This irregular background pulsation is a dynamic and accented goal as well as a sub-phrase boundary, although this time (as opposed to **ex. 2.11**) it ends a segment that is dominated by the "X" stream and leads to a segment dominated by "O" stream activity, an emphatic delay that could be interpreted as a kind of gesture toward the competing stream or its associated "lyrical" character.

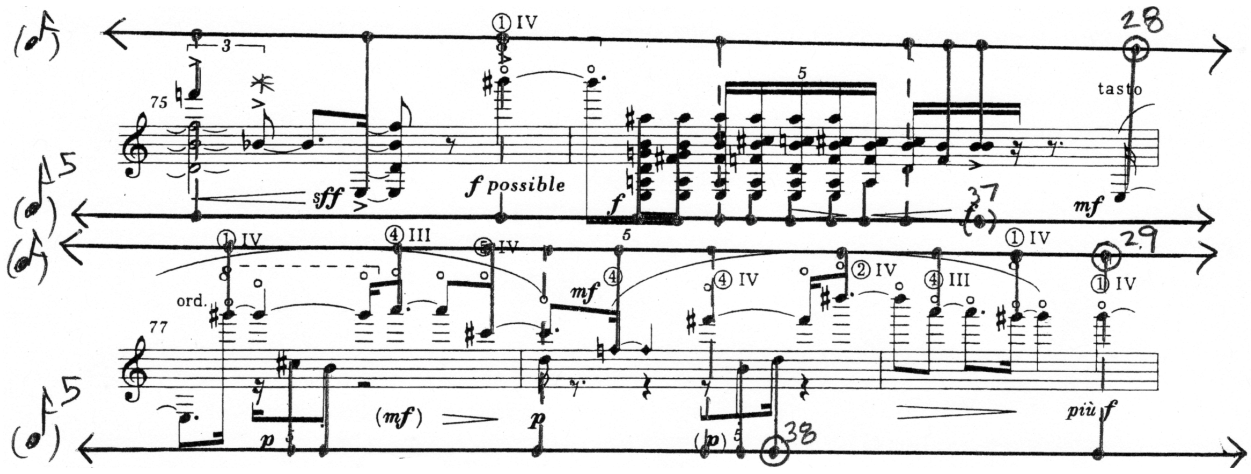
Example 2.11: Displacement of background pulsation X10 and integration within competing elementary pulse stream, *Changes*, mm. 17-20



Example 2.12: Displacement of background pulsation X18, *Changes*, mm. 34-38



Example 2.13: Displacement of background pulsation X37 and integration within competing elementary pulse stream, *Changes*, mm. 75-79

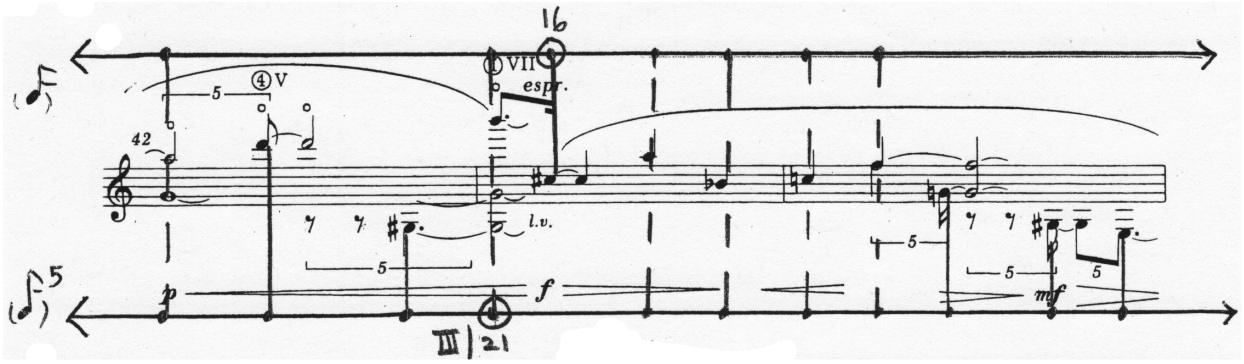


Another advantage of the proximity charts is that most near-coincidence points, and their associated convergence and divergence phases, can be located easily on the vertical axis. The original divergence phase (see bar-measures 9, 17, 26, 34, 43, 51, 59-60, and 68 on the graph) shows a rate of divergence of +3spu.[75] The passage from bar-measures 76-82 marks a turning point, with two successive near-coincidences of 27spu acting as the point of maximal divergence in the abstract polycycle (**fig. 2.5**). However, the first of these two timespans is realized as 25spu as a result of the reaching gesture of X37 noted above (**ex. 2.12**). This moment also marks the beginning of a convergence phase (-3spu), the two first near-coincidence points of which are visible in bar-measures 91 and 99 (**figs. 2.5** and **2.6**). Finally, when observing the musical surface from the annotated score, one becomes aware that many of these near-coincidence points are used to articulate the phrase structure or are clearly demarcated as two voices in a contrapuntal duet. In the section under examination, the

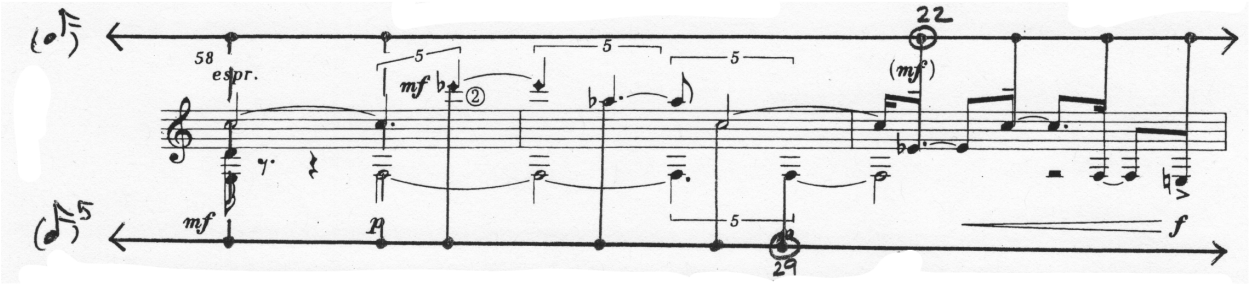
[75] A positive sign is used in divergence phases to denote a gradual widening of the distance between the background pulsations of non-synchronous pulse streams; conversely, a negative sign would signify a gradual narrowing in a convergence phase.

former is illustrated with X21→O16 in bar-measure 43 (**example 2.14**) and the latter with X29→O22 in bar-measures 59-60 (**example 2.15**).

Example 2.14: Near-coincidence X21→O16 articulates the phrase structure, *Changes*, mm. 42-43



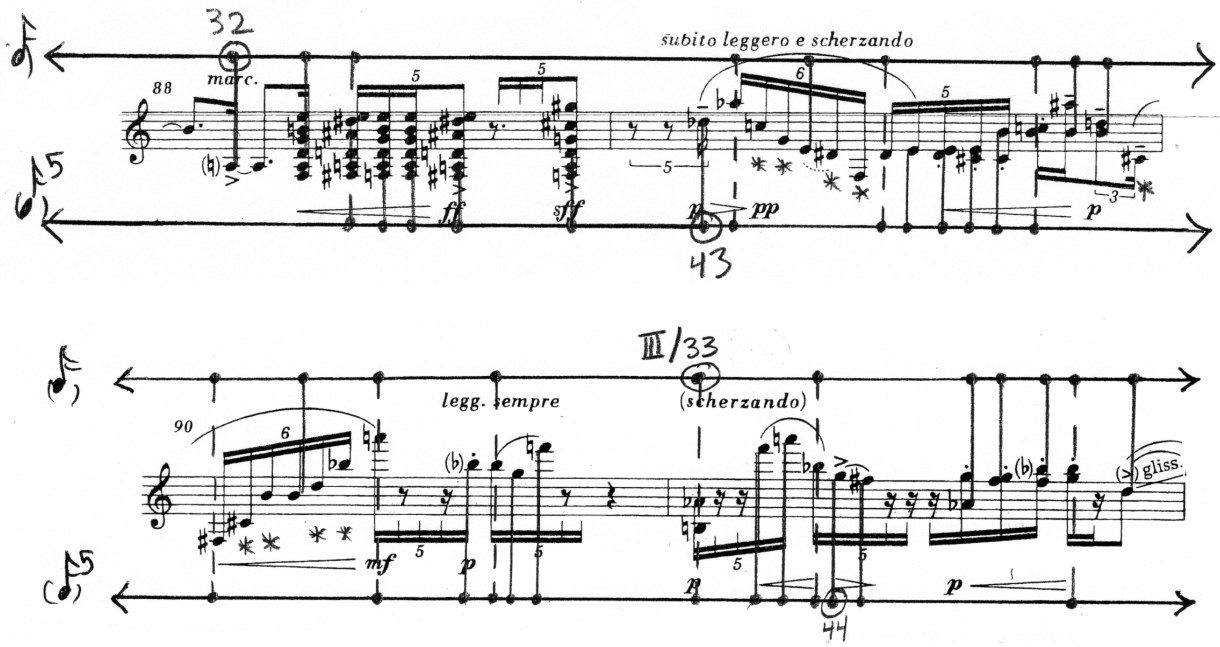
Example 2.15: Near-coincidence X29→O22 participates in a contrapuntal duet, *Changes* for guitar, mm. 58-60



Background pulsations that are used as sectional boundaries are especially salient and charged with narrative potential. The juncture from Episode 3 to the *Scherzando* section (see **table 2.3**; this passage was mentioned previously in subsection 2.3.1.1) makes an interesting case study (see **example 2.16**). Here, background pulsation X43 serves as the head of the sub-phrase that bears the expressive marking *subito leggero e scherzando* and O33 heads the following sub-phrase marked with the re-iterated *scherzando* character. It is as if each pulse stream has been given its own sectional beginning at the strategic moment that leads to the

most dramatic moment of the piece (the "outburst"). Although an innocent performance of this passage would still convey some of the significance of this moment, it becomes more and more clear that a knowledge of the background structure, an awareness of its surface instantiation as part of specific character-patterns, and an understanding of its interaction with foreground and middleground levels have much to offer to the performer.[76]

Example 2.16: Interaction between background pulsation and formal structure, X43→O33, *Changes*, mm. 88-91



[76] Another example of this kind of procedure can be found on the second page of *Esprit rude/Esprit doux*; this passage will be discussed in chapter 3.

Figure 2.7: Abstract polycycle, *Changes*, mm. 103-129

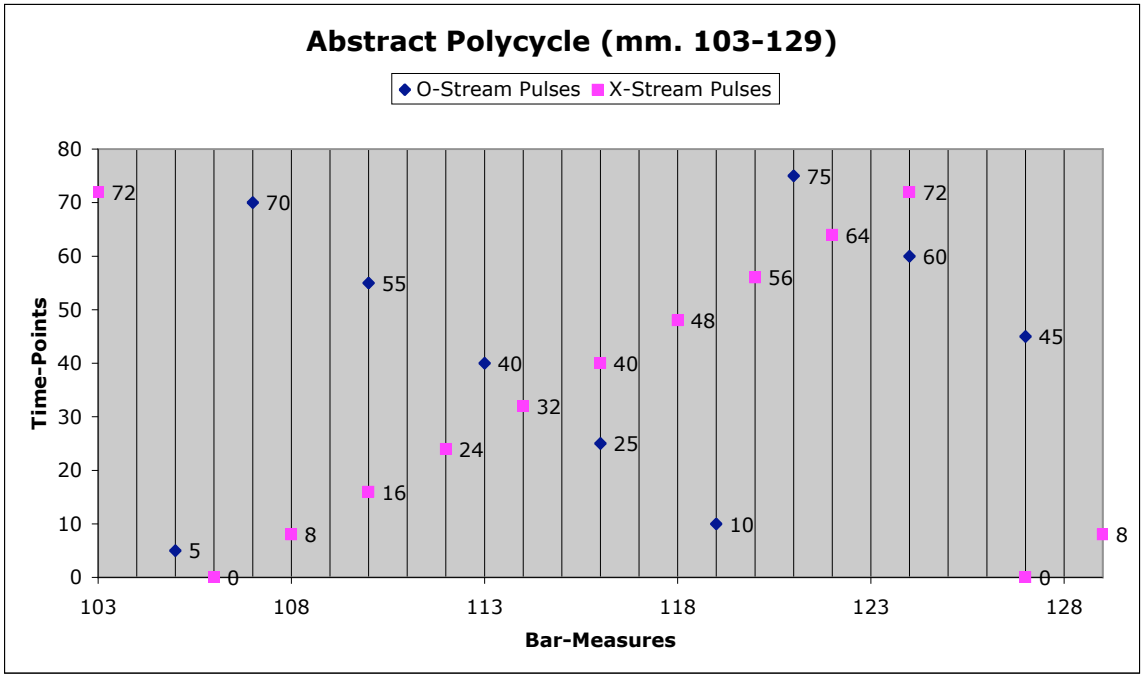
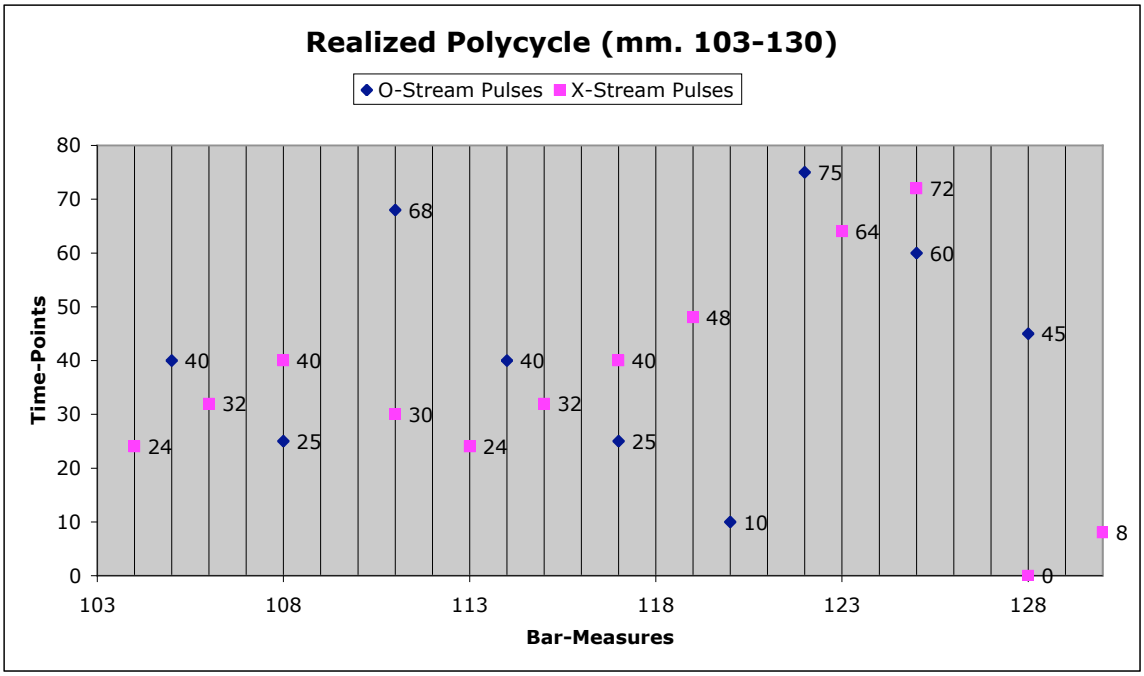


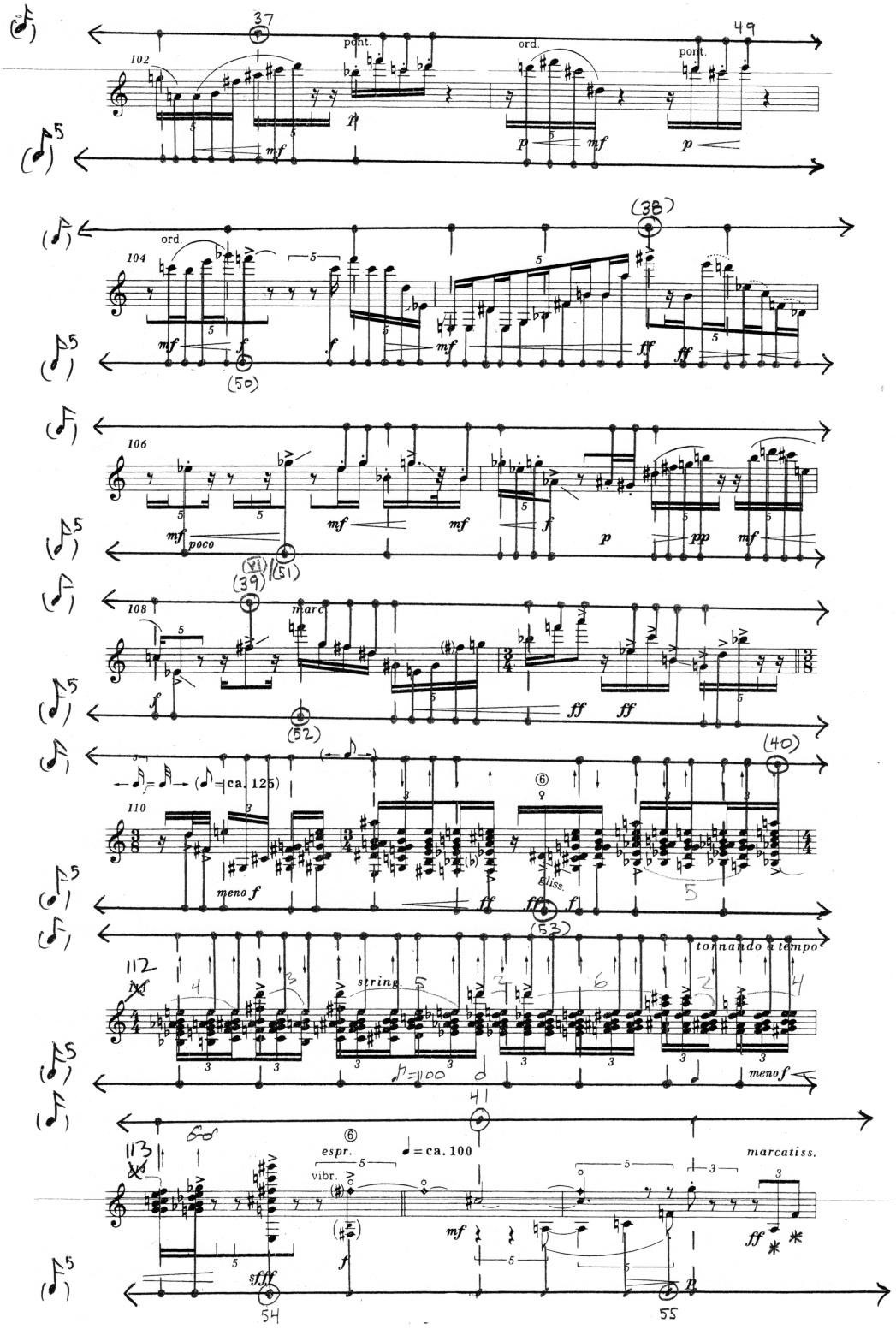
Figure 2.8: Realized polycycle, *Changes*, mm. 103-130



Let us now take a look at the realization of the rest of the polycycle (**figure 2.7** to **2.10**). A comparison of **figures 2.7** and **2.8** reveals that the main region of disruption is located at bar-measures 103-111 (**fig. 2.7**, abstract polycycle). At bar-measure 113, time-point 24 (**fig. 2.8**, realized polycycle), the long-range polyrhythm resumes its normal course.[77] However, although it is true that the expected background pulsation X50 fails to materialize in bar-measure 103, time-point 72, pulse-measures could be located until at least X52 if we consider a displacement of +8 quintuplet sixteenths for X50 (+32spu; see **example 2.17**, m. 104, lower time-line). The resulting pattern, which includes the hypothetical background pulsations X53 and O40 belonging to the very irregular "outburst," shows an interesting case of structural parallelism in its background pulse streams (see **fig. 2.8**, mm. 104-117). Thus, it would appear that for a brief moment, not only has the foreground collapsed onto the middleground (as noted previously in subsection 2.3.1.2), but the background has also been put in a metrical "corset." Finally, it should be noted that X54 is an exception of regularity within the "outburst": located at its very edge, it functions an accented phrase ending and its location can be precisely traced back from X55 based on the number of elementary pulse units (bar-measures 114-115, lower time-line, **ex. 2.17**). Overall, these observations suggest that despite the disruption created by Carter's expansion of this section, there might still be relevant periodicities worth investigating.

[77] As noted previously, because of the irregularities in the passage leading up to this point, bar-measure 113 in the realized polycycle (**fig. 2.8**) corresponds to bar-measure 112 in the abstract polycycle (**fig. 2.7**).

Example 2.17: Annotated score, *Changes*, mm. 102-115



Once the long-range polyrhythm resumes its course in bar-measure 113, it is very regular except for one displaced background pulsation and the final metric modulation discussed previously (mm. 131-150). In bar-measure 121, X58 appears to be integrated within the sixteenth-based melody as a result of a displacement of +19spu (**example 2.18**), the corresponding sub-phrase exhibiting a relatively high degree of activity in the "O" stream. Finally, although the two graphs of the final section (**figures 2.9** and **2.10**) appear very different (a problem intrinsic to the two-dimensional representation, which emphasizes time-point location at the expense of pulse-measure duration in clock time), the realized polycycle (**fig. 2.10**) is actually extremely regular. The near-coincidence points with a convergence rate of –2spu can be easily located in bar-measures 133, 136, and 144 (compare with **ex. 2.10**). As noted previously and due to the metric modulation, the final coincidence point is now located on the downbeat of bar-measure 150 (as opposed to bar-measure 158, quarter beat 3; see **fig. 2.9**, abstract polycycle).

Example 2.18: Displacement of background pulsation X58 and integration within competing elementary pulse stream, *Changes*, mm. 119-123

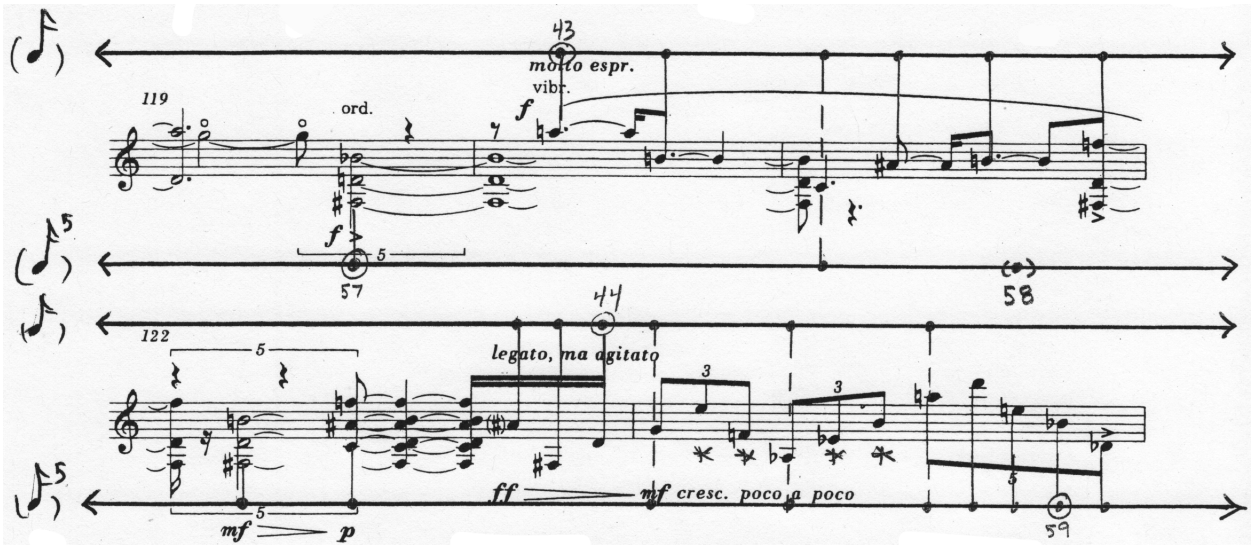


Figure 2.9: Abstract polycycle, *Changes*, mm. 130-158

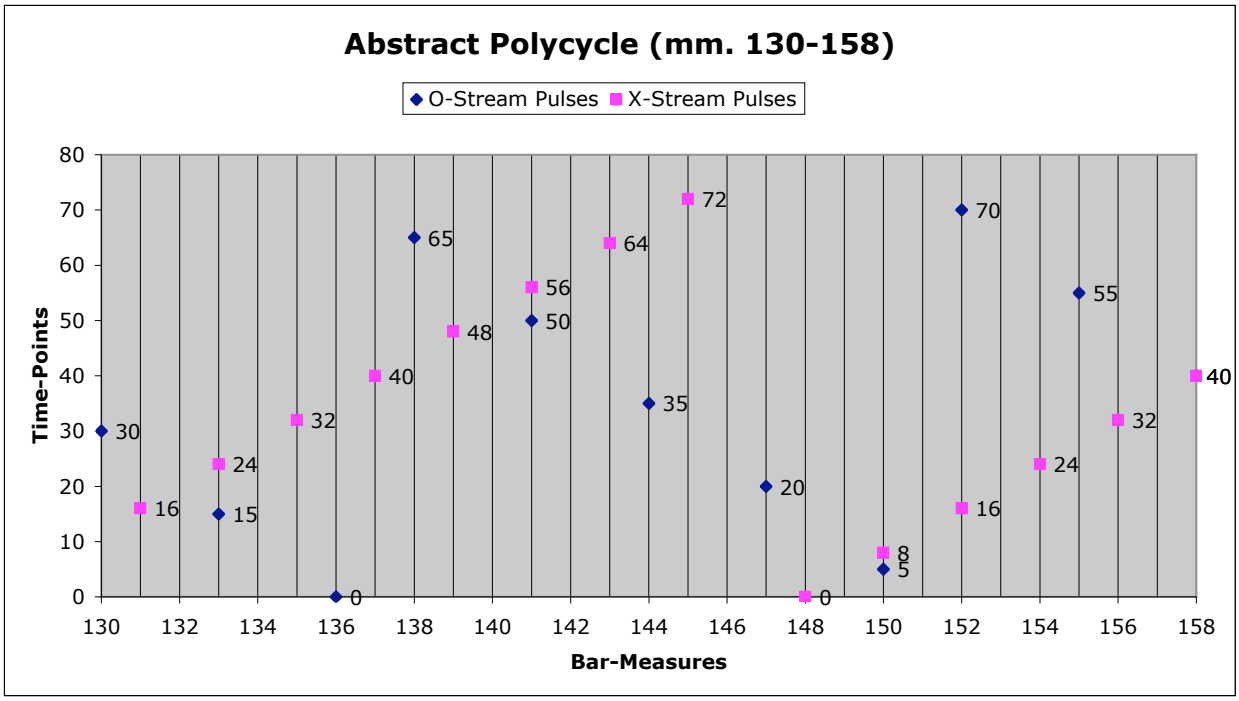
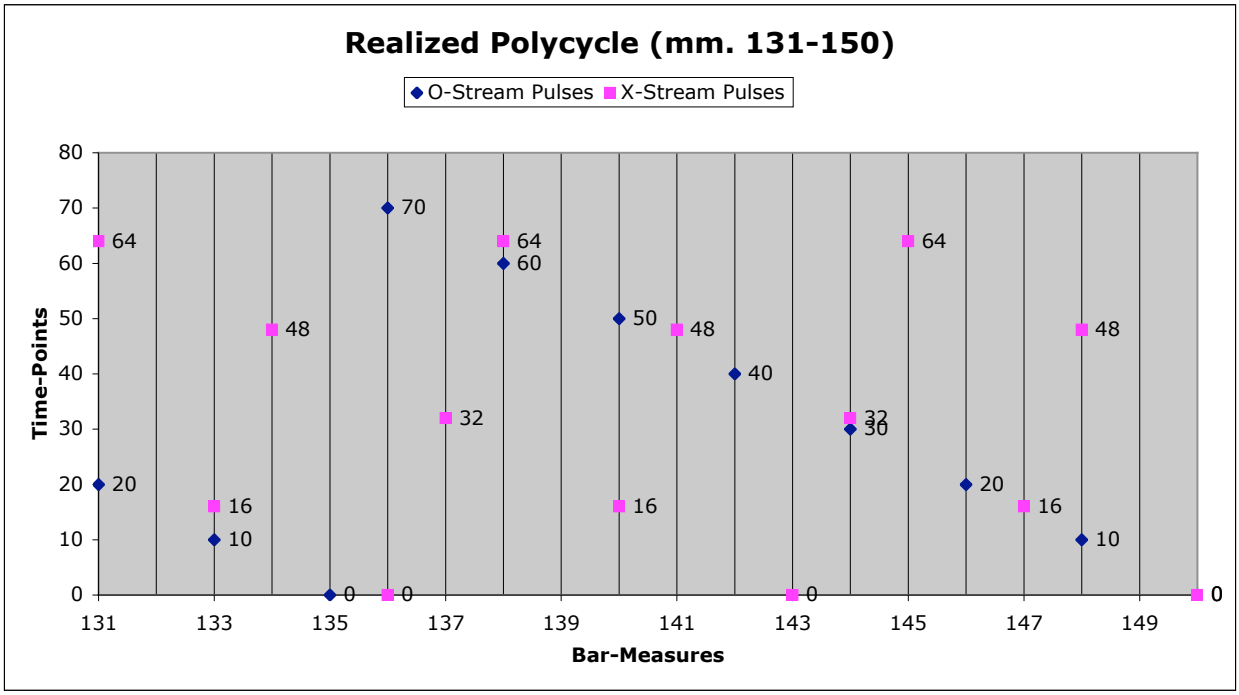


Figure 2.10: Realized polycycle, *Changes*, mm. 131-150



As I have shown, much insight can be gained from the mapping out of the background structure (by means of a polycyclic spectrum and proximity charts) and a careful examination of its interaction with the musical surface as well as the foreground and middleground levels. In this section, I have focused on two aspects of Carter's compositional technique: (1) the specific use of background pulsations and near-coincidences, and (2) background irregularities and their manifestation on the musical surface. First, we have seen that Carter often uses background pulsations as structural markers. Given this structural function and the relative salience of these background pulsations, near-coincidences at the background level can also come to be understood as marked events that regulate the musical flow. At other times, they are used to delineate different voices in the musical texture. Second, there are only a few background irregularities overall. Background irregularities can be understood to result from a flexible treatment of the long-range polyrhythm, as when a section is expanded (i.e., the "outburst" passage), or could be construed as alternative spellings chosen for ease of reading (e.g., a background pulsation based on the quintuplet-sixteenth elementary pulse stream that is surrounded by sixteenth notes is notated as a sixteenth note). However, the particular care with which the composer notates the music and the resulting structures suggest that there may be deeper connections to uncover. In this section, we have seen that in *Changes* such irregularities also appear to function as rhetorical gestures or expressive devices within a richly constructed musical discourse that relies on an awareness of contrasting temporal layers. Thus, the consideration of the association of background pulsations with one or another elementary pulse stream allows us to posit a dynamic interpretation of the events on the musical surface, and these interpretations can inform a particular listening or performance of the work.

CHAPTER 3

Polymetric Potential and Realization

It may already have become apparent to the reader that in some ways, the distance between background and foreground, which was seen as problematic by Yeston, might be a necessary characteristic of polymetric structures since a strong middleground, such as that present in common-practice dance-derived music, is much more likely to neutralize any polymetric potential. In Carter's textures, sometimes the bar-measure integrates the competing pulse streams and the texture is perceived as irregular or the quarter beat is understood as having fluctuating subdivisions (an *integrated* polymetric structure); much more often, its influence appears to be neutralized, giving rise to competing metrical projections that may result in more or less regular cross-rhythms (a *polarized* polymetric structure). When two or more metrical projections are found to be equally viable, the polymeter is potentially *balanced*, although a systematic way of measuring the realization of such potential for a given listener or group of listeners has not yet been devised.[1] In all cases, the use of the quarter beat (and the bar-measure) provides a temporal template on which the composer weaves his musical discourse, and from which performers and analysts may look for cues for the underlying narrative. However, before embarking on the exploration of polymetric potential (and realization), it will be necessary to further differentiate between textural streams and rhythmic/metric strands, as the two are not always equivalent. The following sub-section will help to illustrate this point.

---

[1] Chapter 4 presents the results of an exploratory music-psychological study based on an example of *balanced* polymetric structure.

3.1 Streaming and Multileveled Polyrhythmic Design: *Esprit rude/Esprit doux* I (1985)

The first piece bearing the title *Esprit rude/Esprit doux*, a duet for flute and clarinet composed for Pierre Boulez's 60th birthday celebration, is built on a 21:25 long-range polyrhythm.[2] This temporal plan, which rests on the use of two contrasting background pulse units (in Carter's polyrhythmic sketch, these correspond to 20 triplet quarters for the flute and 28 quintuplet eighths for the clarinet; see **ex. 1.2**) suggests a basic division of the texture into two streams, each of which is represented and made more perceptible by its association with a specific instrumental timbre and elementary pulse unit (and its resultant surface speed). While this is certainly true from the point of view of pre-compositional planning, the musical surface actually calls for a more dynamic interpretation of the texture, in which stream fusion and segregation interact with polyrhythmic design (and by extension, polymetric structure) in various ways.

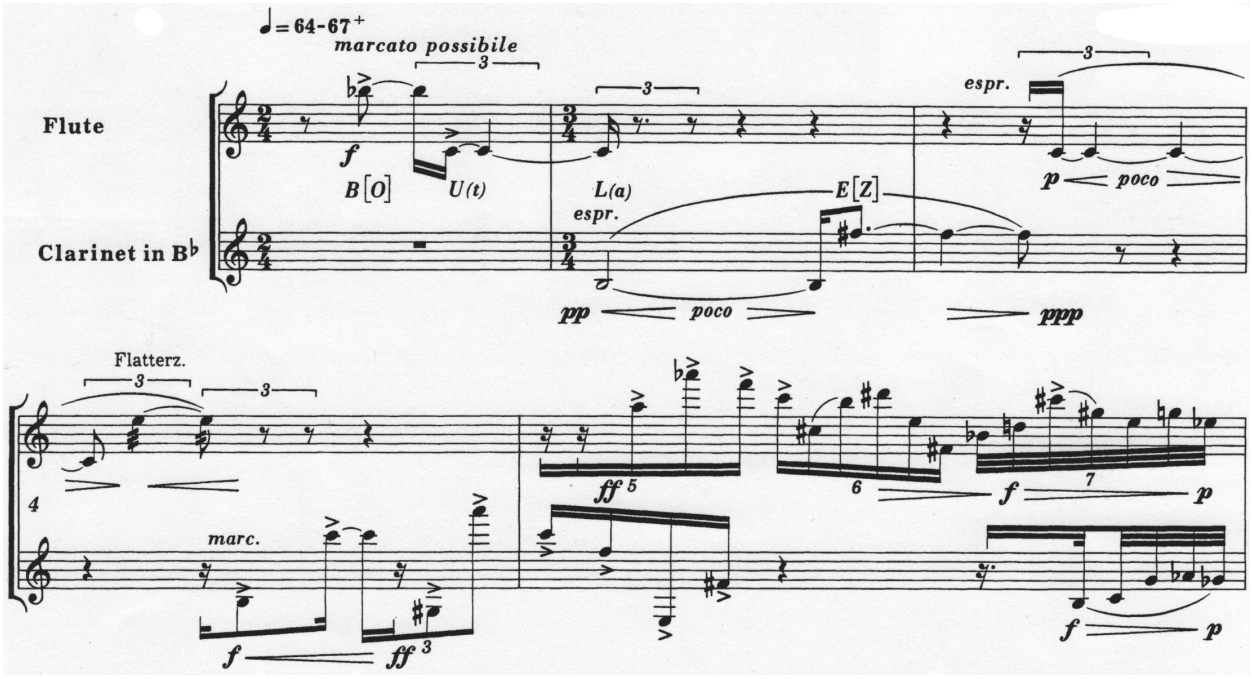
Example 3.1 presents the opening of *Esprit rude/Esprit doux* I (mm. 1-5). Here, the overall texture could be described as a *Klangfarbenmelodie* that results from the slow alternation of the flute and clarinet, an interaction that strongly favors stream integration. This interpretation is also supported by Carter's use of pitch recurrence. Thus, after the first two bar-measures state the motive derived from the dedicatee's name, i.e., <Bb, C, A, E>, a representative of the all-interval tetrachord (AIT) 4-Z29 (0137) and Carter's tetrachord no. 23, the initial recycling of these pitches in mm. 3-4 displays register invariance. Furthermore,

---

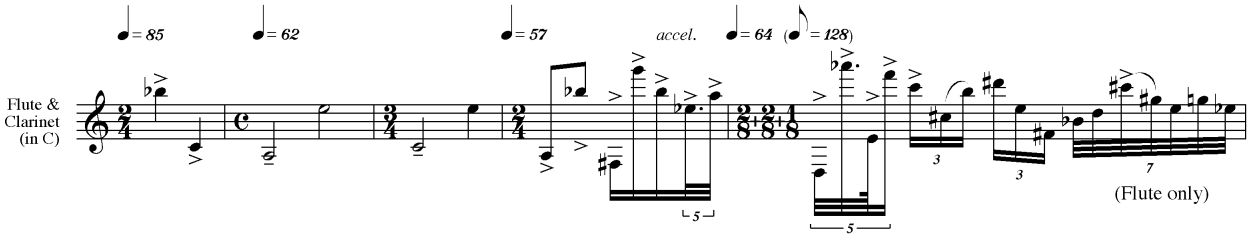
[2] Both Link and Schiff provide basic information for the background polyrhythm; see Link, "Long-Range Polyrhythms," pp. 41-42 and 118, and Schiff, *The Music*, p. 140. Schiff's speed ratio of 6.25:5.25 is based on Carter's rhythmic sketches, which set the initial tempo at quarter = MM 70 (Schiff reports the tempo used in the tempo calculations for metrical modulations, i.e., half = MM 35). The rhythmic, harmonic, and formal elements of this piece are also discussed in more detail by Matthias Truniger, "Elliott Carter: *Esprit Rude/Esprit Doux*," in *Sonus* 19/1 (1998): 26-52.

until the 12-tone aggregate is completed in m. 5, second beat, third subdivision (B5), the only pitches that are repeated are those from the opening tetrachord.

Example 3.1: Opening, *Esprit rude/Esprit doux* I for flute and clarinet (1985), mm. 1-5



Example 3.2: Transcription and metrical interpretation of *Esprit rude/Esprit doux* I, mm. 1-5



Example 3.2 presents a transcription and metrical interpretation of this passage. As shown in this transcription, despite the use of non-isochronous elementary pulse streams (mostly triplet sixteenths for the flute against duple sixteenths in the clarinet part), the temporal interaction of successive pitch events suggests a single rhythmic/metric strand that is subjected to tempo manipulations. The transcription also makes apparent the metrical

processes at work. The inter-onset intervals associated with each of the first two flute tones encourage duple grouping that is perceived as leading to the clarinet's A3, which acts as a local downbeat (the tenuto markings denote perceptibly elongated pitch events that can nonetheless be understood as equivalent).[3] The following two quasi-isochronous durations (represented by half notes), which are shared between the clarinet and the flute, not only establish a speed of about 31 beats per minute, but also suggest a larger duple measure that is supported by contour. At the same time, each of these rather slow pulsations is likely to be perceived as being subdivided into two parts, an inner pulse of quarter = MM 62, which will be perceived as a slowing down of the tactus as it was established by the opening two pitch events.[4] However, this measure quickly finds itself challenged by the durational series initiated by the flute's E5 (m. 3, beat 3), which presents two overlapping LSS rhythms that give way to more irregular articulations shared between the parts before the flute settles momentarily on a pulse of 256 pulsations per minute (m. 5, beats 2-3). The sudden acceleration felt at the surface and the dynamic contrast offered by the clarinet's entrance

[3] Here, the inter-onset interval (IOI) is used instead of the full duration because the focus is on the interaction of successive pitch events that are fused into a single strand. In the case of two or more perceptible layers, it might become more meaningful to take into consideration the total duration of each event, independently of overlapping pitch events.

[4] There are, of course, other possible scenarios, especially given the 4:3 ratio between the two tempi associated with the quarter note (as opposed to a ratio of 2:1 or 3:2, which would be more strongly suggestive). Thus, one could imagine the opening two pitches as the second and third triplet subdivision of a larger beat closer in speed to the half note, resulting in a smaller tempo fluctuation (from half = MM 28 to half = MM 31). The chosen interpretation is only the most straightforward one for a context where the listener has no specific expectation as to the grouping of first two pitches/durations.

support the interpretation of the second A3 as a new downbeat with duple subdivisions (m. 4).[5]

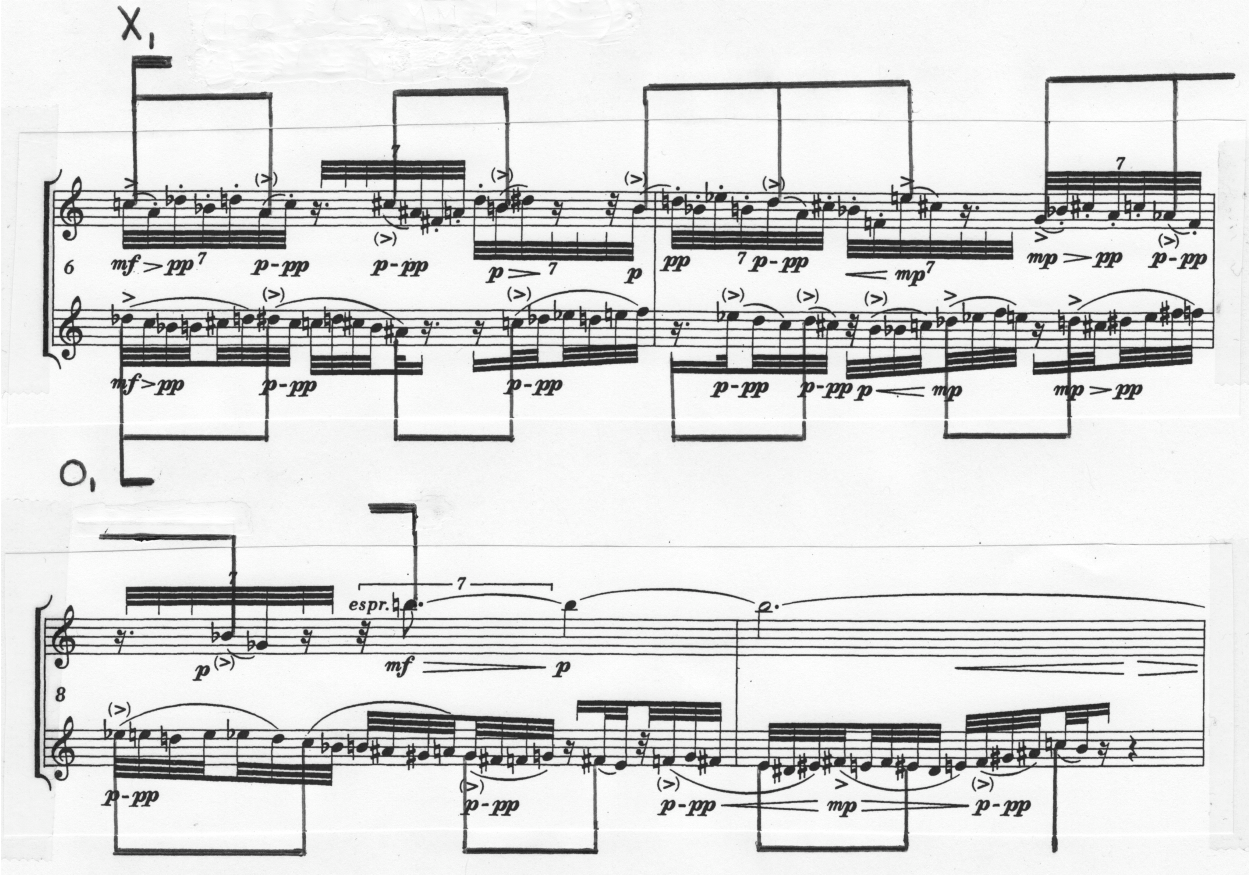
The metrical interpretation just proposed rests on the principle that the repetition of a durational unit will lead the listener to expect its continued relevance and will result in particular grouping expectations, a phenomenon that Hasty has referred to as metrical projection.[6] It also assumes that the listener's attention is engaged simultaneously with the interpretation of current events as well as the prediction of future and re-interpretation of past events. The resulting metrical groupings also demonstrate an important characteristic of Carter's measures, that is, their limited *projective field*. Thus, as soon as the musical surface puts forward a particular measure, this grouping finds itself challenged or negated by new durational events. This transcription also emphasizes the likely interpretation of the partially articulated 5:4 surface polyrhythm of the flute and clarinet (ex. 3.1, m. 5, beat 1) as fused or integrated. As we shall see however, the final beat of the same measure (another partially articulated 8:7 surface polyrhythm, represented here by the flute alone) will call for a different interpretation.

---

[5] The placement of the downbeat of m. 5 is much more ambiguous and was determined based on the underlying pulse of quarter = MM 64 as well as the grouping suggested by the contour accents (see example 3.3). An alternative interpretation is the continuation of a triple meter in m. 4, with a downbeat provided by the flute's C5.

[6] Hasty, *Meter as Rhythm*. The notable difference between Hasty's and other theoretical approaches that make use of the periodic return of a duration as a means of interpreting rhythmic events is that Hasty's approach does not require the occurrence of a third isochronous duration in order to posit metrical grouping. In other words, a duple measure does not need to be confirmed by the recurrence of a second downbeat in order to be psychologically relevant.

Example 3.3: Multileveled polyrhythmic design, *Esprit rude/Esprit doux* I, mm. 6-9



While the passage just discussed could scarcely be qualified as polyrhythmic (based on stream fusion and the integration of pitch events into a single rhythmic/metric strand), the next passage (see **example 3.3**) presents a contrasting texture where stream segregation is complemented by multileveled polyrhythmic design. Bar-measure 6 marks the beginning of the polyrhythmic cycle.[7] At the notated tempo of quarter = MM 64-67^{+}, the flute's background pulsations (the "X" stream) give rise to an actual speed of MM 4.48-4.96^{+} while the clarinet's resultant speed (the "O" stream) is MM 5.33-5.58^{+}.[8] The annotated example

[7] Schiff (1998) and Truniger (1998) also make the same assumption.

[8] The piece does have three additional notated tempi, i.e., quarter = MM 90-94+, quarter = MM 60-63, and half = MM 45-47+, but the background speeds are very nearly equivalent

(**ex. 3.3**) shows that pulse-measure X1 exhibits at least three levels of grouping (or subdivision). At the most foreground level, Carter's accents signal a series of pulsations (or beats) with an associated timespan of 5 septuplet thirty-seconds (shown with stems) that are themselves grouped into duple and then triple measures at the middleground level (shown by beams).[9] The presence of the long B5, which divides pulse-measure X1 into two equal timespans, suggests yet another (slower) middleground unit (or hyperbeat) of 50 septuplet thirty-seconds that defines the pulse-measure as duple (bar-measures 8-10; see **example 3.4** for the last 15 elementary pulse units). While somewhat more irregular, the clarinet part is nonetheless strongly metrical. Thus, pulse-measure O1, which corresponds to 96 thirty-seconds, subdivides itself into eight middleground-level duple measures of 12epu each (two beats of 6 thirty-seconds that are themselves subdivided into 2×3 based on rhythmic design; e.g., m. 7, beats 1-2). The beats of the first two of these measures are clearly demarcated by accents and contour. However, while the placement of a rest on the downbeat of m. 7 creates a momentary disturbance that is quickly resolved by the subsequent regularity of rhythmic design and accentuation, the displaced accents and slurs from m. 8, beat 3 to m. 9, beat 2, generate much more syncopation.[10]

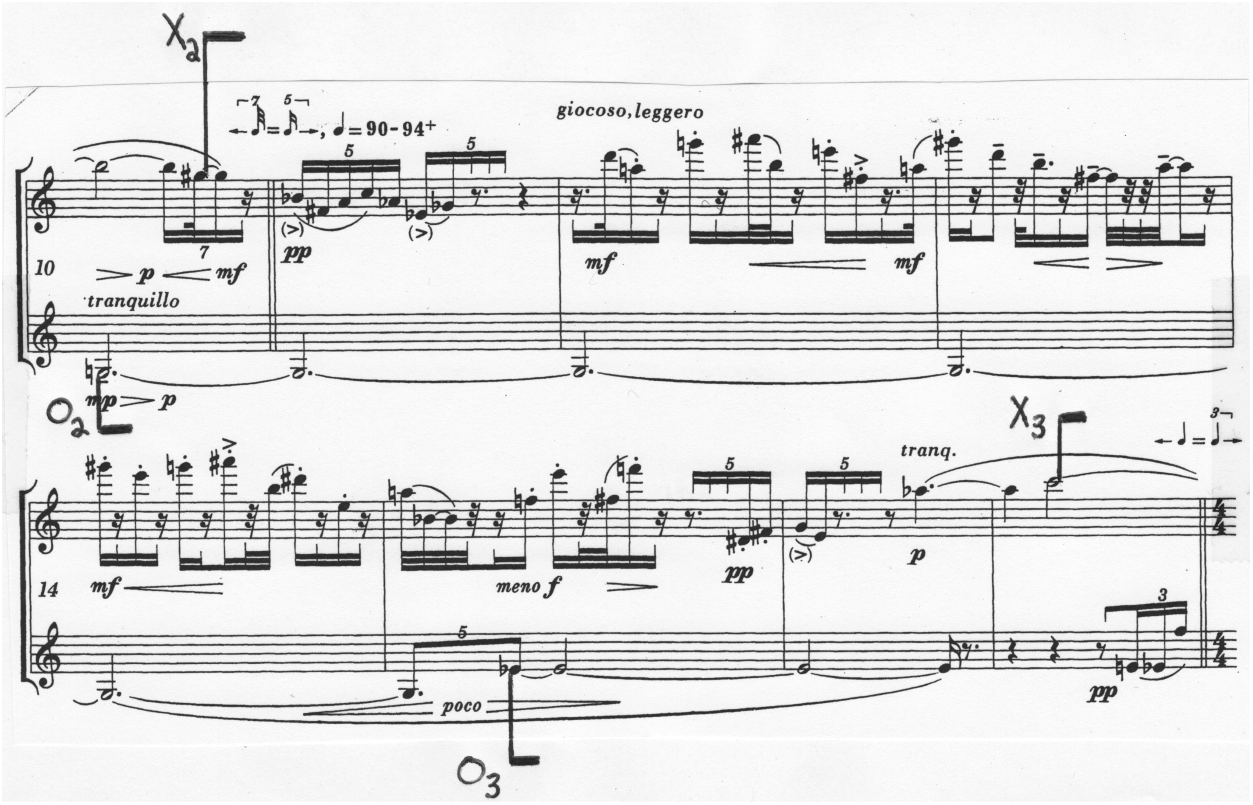
---

throughout. Based on the lower limit of quarter = MM 60, the resulting speed ratio at the background is about 4.5:5.4.

[9] This specific grouping is suggested by rhythmic design (including parallelism) and event density (interpreting more dense beats as relatively more accented). One could also consider either duple or triple measures throughout the first hyperbeat. The purpose of this example is not to present a definite metrical interpretation but to demonstrate the presence of several levels of pulse, which confirm the metrical properties of the passage.

[10] Again, there are other metrical interpretations, e.g., triple measures at the beginning would be supported by rhythmic design and event density (in a manner similar to the flute part) and dynamic shaping (in particular, the crescendo to ***mp***, m. 7, beat 2), although this measure

Example 3.4: Stream segregation, *Esprit rude/Esprit doux* I, mm. 10-17



Returning to the issue of textural versus temporal (and metrical) streaming, we can observe that while the passage from mm. 6 to 9 offers strong evidence of polymetric properties, i.e., it presents two superimposed rhythmic/metric strands with competing pulses at several structural levels, the independence of the two layers doesn't remain unchallenged throughout. When listening to this passage, it becomes clear that at the beginning of the passage (mm. 6-8), the shared register, overall speed, and dynamic level actually promote stream fusion rather than segregation.[11] However, once the flute reaches its B5 in m. 8, the

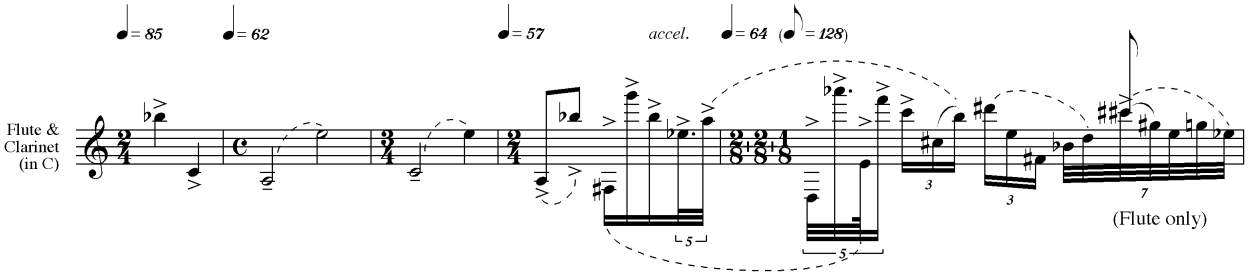
---

would be somewhat more contradicted by the accent on m. 8, downbeat, as well as the subsequent disturbances discussed above.

[11] This effect is noticeable in all of the recordings cited in the discography; see Nouvel Ensemble Moderne (ATMA Classique: ACD2 2280, 2001), P. Gallois and M. Lethiec

two streams become clearly audible again, although metrical definition begins to slip away. Thus, in mm. 10-17 (**ex. 3.4**), while the two textural streams are clearly set apart, the lack of activity, first in the flute (second hyperbeat of pulse-measure X1 initiated in m. 8; see **ex. 3.3**) and then in the clarinet (undivided pulse-measure O2 and first hyperbeat of pulse-measure O3), poses a real challenge to polymetric potential.[12]

Example 3.5: Groupings leading to stream segregation, *Esprit rude/Esprit doux* I, mm. 1-5



If this entire passage (mm. 6-17) is put back into the context of what preceded it, we can see how Carter has skillfully created a play of expectations that takes full advantage of both textural and temporal elements. **Example 3.5** presents the transcription and metrical interpretation of the opening passage previously discussed, which has been marked with dotted slurs to show how melodic grouping sets up the scene for both textural and metrical streaming. In mm. 2-4, a moment of undeniable integration, the listener is presented with three successive measures in which pitch events of contrasting registers are grouped from

(ARION: ARN 68495, 2000), and Ensemble Contrasts (Classic Produktion Osnabrück & Westdeutscher Rundfunk Köln: cpo 999 453-2, 1998).

[12] Interestingly, as both X1 and O3 form large duple measures, the first hyperbeat of X1 and the second hyperbeat of O3 stand in symmetrical relationship to each other. Note that given the local tempo of quarter = MM 90-94+ (or T2), the 21 quintuplet eighths that correspond to the inter-onset interval initiated by the clarinet's Db in m. 15 are equivalent to 14 quintuplet eighths at T3 (quarter = MM 60-63), i.e., half of the background pulse unit's duration.

low to high. This contour motive undergoes further development starting in m. 4, beat 2, where the higher pitch event in the clarinet part (G6) is followed by four additional pitches, resulting in the contour segment <1, 5, 4, 3, 0, 2>.[13] The flute's imitation of this contour motive in bar-measure 5 (starting with the final note of m. 5, **ex. 3.5**, i.e., pitch segment <A5, Ab6, F6, C6, C#5, B5>) contributes to a growing sense of contrapuntal interaction.[14] On the other hand, the quickening pace resulting from the acceleration of tactus and increased event density gives added salience to the melodic high points leading to the perception of the flute's D#6 and C#6 as accented, resulting in melodic groups of contrasting contour (i.e., <4, 3, 0, 1, 2>) and cardinality (five pitch events as opposed to six). Thus, while register contrast had been previously neutralized by the integration of pitch events into a single rhythmic/metric strand, it now becomes a driving force for metrical expectations.[15] This interpretation is emphasized in the transcription by the addition of a flag to the C#6, which by virtue of its timespan (5 septuplet thirty-seconds) acts as a metrically defined anacrusis to the polyrhythmic cycle's downbeat.

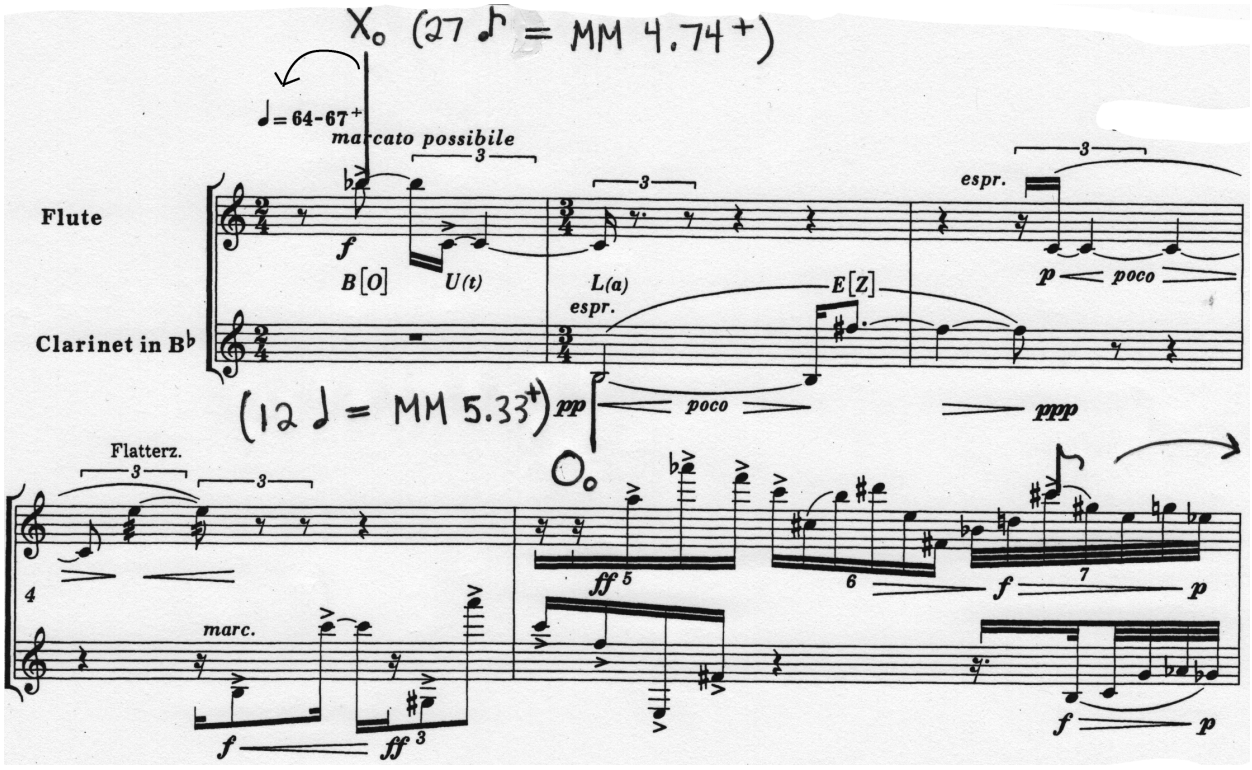
---

[13] This contour motive could be understood as the joining of two smaller motives, each of which is characterized by an ascending motion from the lowest to highest pitch, i.e., <0, 3, 2, 1> and <0, 1>.

[14] These two representatives of hexachord 6-Z37 (012348) (or Carter's hexachord no. 23) are actually even more closely related as they share the unordered pitch class interval series <1, 3, 5, 1, 2>, i.e., the clarinet's ordered pitch intervals <+37, -9, -7, -25, +14> become the flute's <+11, -3, -5, -11, +10>. Another way to conceptualize this relationship is that the ordered pitch class intervals of the two instruments are inversions of each other, i.e., <1, 3, 5, 11, 2> becomes <11, 9, 7, 1, 10>. Carter derives these two hexachords from a twelve-note, eleven-interval chord.

[15] Traditionally, register is a strong agent of stream segregation, as demonstrated by *compound melodies*, i.e., single melodic lines that are understood as representing two or more voices.

Example 3.6: Background anacrusis, *Esprit rude/Esprit doux* I, mm. 1-5



In fact, the entire opening passage (mm. 1-5) could be interpreted as an anacrusis leading to the structural downbeat provided by the beginning of the background polyrhythmic cycle. Carter's sketches, which show that the opening five measures were a later addition, support this interpretation.[16] Interestingly, the actual duration of this addition is also consistent with the background polyrhythm. As shown in **example 3.6**, the clarinet's entrance actually corresponds with a background pulsation that precedes the beginning of the background polyrhythmic cycle's downbeat (12 quarters = 96 thirty-seconds). Conversely, the timespan associated with the flute entrance results in a background pulsation that is

[16] These sketches are held at the *Paul Sacher Stiftung* in Basel, Switzerland (accessed March 2008).

slightly too short (27 eights = 94.5 septuplet thirty-seconds).[17] However, it is possible to conceive of the opening rest as the actual location of the background pulsation of the flute. This interpretation is supported by the program of the piece, which associates *esprit rude* (to which, by virtue of their dynamic level and accentuation, the opening flute's tones belong) with "rough breathing," which is produced by a sounded H preceding the initial vowel.[18] Given that the performer would most probably accompany his or her breath with some sort of motion and/or visual cue, the actual location of the initial background pulsation is very likely to be more than conceptual. The anacrustic character of the opening five bar-measures is also underscored by the surface acceleration in mm. 4-5, from eighth notes to triplet eighths, through overlapping sixteenths, quintuplet and sextuplet sixteenths, and finally leading to the septuplet thirty-seconds by means of which the flute's pulse stream is subsequently defined (a "terraced" acceleration of MM 128-134^{+} to MM 448-469^{+}). This gesture of increasing speed and density provides much energy to the structural downbeat at m. 6.

In conclusion, despite the clear streaming suggested by Carter's use of structural polyrhythms as pre-compositional devices, the composer's musical surface is characterized by rapid changes from one type of texture to another as well as by a tendency to negate or play with the metrical expectations he has set up, resulting in limited projective field. In addition, even when the surface groupings suggest strong polymetric potential, these are often manipulated in such a way as to make them perceptually ambiguous. In the example

---

[17] Truniger (1998) also notes the timespan equivalence of the clarinet's music in mm. 1-5 and the "somewhat shorter" duration of the flute's, but he does not interpret this passage as a background anacrusis; see Truniger, "Elliott Carter," p. 41, fn. 8. In order to be exact, the flute's timespan leading to m. 6 would need to be a little more than 28 eighth notes.

[18] See the program note; Elliott Carter, *Esprit rude / Esprit doux* (New York: Hendon Music, Boosey & Hawkes, 1985).

just examined, the clearest polymetric structure (mm. 6-9) was obscured by the shared register, overall speed and dynamic, as well as syncopation in one of the parts. However, limited projective field and perceptual ambiguity do not preclude the relevance of metrical processes, quite the contrary. Metrical projections resulting from the use of isochronous or quasi-isochronous durational series are frequently used by Carter to direct the attention of the listener toward a specific pitch event and enrich the interaction of instrumental parts.

3.1.1 Polymetric Structures Defined

We are now in a better position to define a *polymetric structure* as a *multileveled polyrhythmic compound with at least two concurrent rhythmic strands (as opposed to pulse "trains" or "strings") in which durational events can be traced back to contrasting pulse streams or non-isochronous beat structures that give rise to non-synchronous metrical projections*, *the latter case implying the presence of conflicting groupings on at least one structural level* (i.e., sub-beat, beat, measure, etc.).[19] The presence of two concurrent metrical projections supposes that while there might be one principal generating pulse for each rhythmic strand (e.g., Carter's elementary pulses), these strands are designed in such a way that their interaction with their associated pulse stream gives rise to *meter* defined more specifically here as *the grouping of one pulse level by another*. Furthermore, although each pulse stream is theoretically associated with a particular speed, the actual perceived tempo

[19] "Pulse train" is one of the expressions favored by music psychologists; the term "string" has been used recently by Andrew Mead in "On Tempo Relations," in *Perspectives of New Music* 45/1 (2007): 64-108. These two terms are equivalent and usually refer to a series of events of equal (or nearly equal) duration. My use of the expression "rhythmic strand" emphasizes the potential presence of non-isochronous events that can nonetheless be traced back to a discrete series of implied isochronous pulsations, which I refer to as a "pulse stream."

will depend on the specific type of polymetric structure and whether its polymetric potential is fully realized.[20]

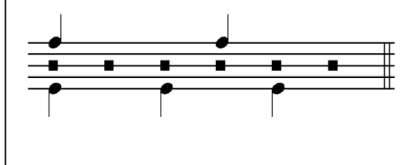
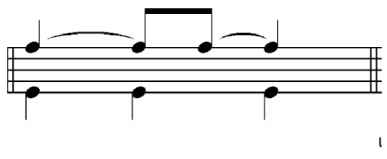
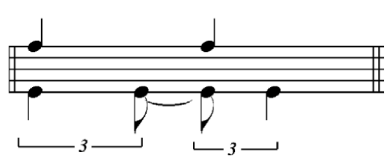
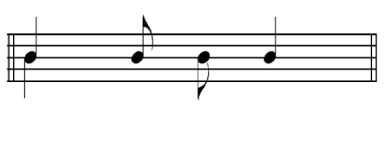
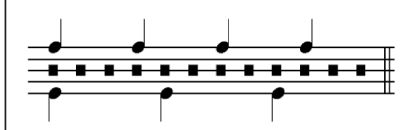
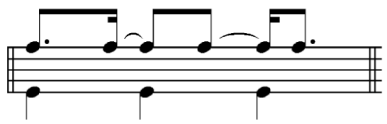
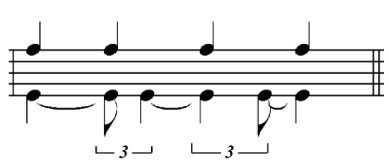
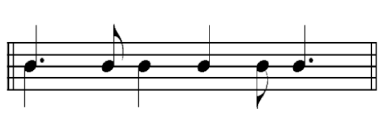
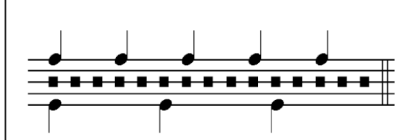
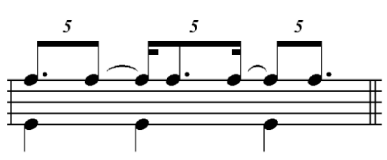
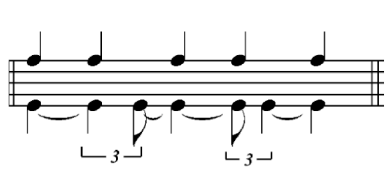
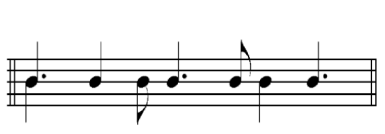
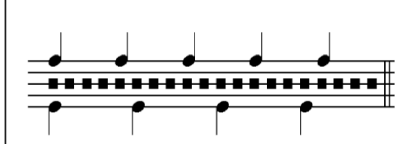
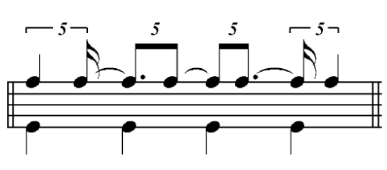
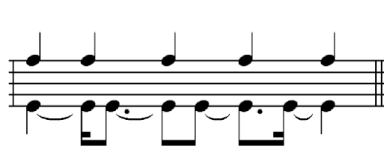
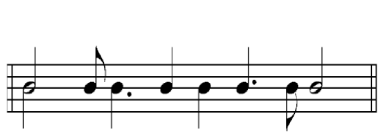
By investigating various aspects of the interaction of these rhythmic layers (mensural determinacy, subpulse and event density, nestedness, proximity, and polycyclic spectrum), I will demonstrate that emerging polymetric structures can be construed as belonging to three categories that are representative of reasonably predictable behaviors: (1) *integrated*, in which competing pulse streams are assimilated by a single metrical projection, (2) *polarized*, in which one of the pulse streams is dominant and the other provides cross-accentuation, and (3) *balanced*, in which two or more concurrent pulse streams achieve metrical significance. As will be demonstrated in this dissertation, these idealized categories offer a conceptual model that is expressive of a listener's experience and, as such, contribute to render the polyrhythmic processes more intelligible, both as compositional and perceptual products.[21]

---

[20] The issue of perceived tempo is an important one but is beyond the purview of this dissertation. It should be emphasized however that while I present tempo as an emergent property of meter, in some cases *polytempo*, that is, the presence of simultaneous speeds, might be more easily cognized than polymeter. I am indebted to John Covach for raising this issue upon hearing a presentation of my work.

[21] The 'listener' referred to here is an idealized listener. Since an in-depth discussion of the perceptual implications of the structural use of polyrhythms is beyond the scope of this chapter, the observations related to listening experience are the result of an introspective process. Issues of perception will be addressed in chapter 4.

Table 3.1: Simple polymetric models for in-phase structural polyrhythms

Ratio	Polymetric Type			
	Balanced	Polarized		Integrated
		Minor	Major	
3:2				
4:3				
5:3				
5:4				

3.1.2 Abstract Models

Table 3.1 presents in-phase structural polyrhythms with speed ratios of 3:2, 4:3, 5:3, and 5:4 that have been subjected to three polymetrical interpretations. In these simple models, the actual structural level has not been specified, that is, each compound could be taken to represent a beat, measure, hypermeasure, or even an entire composition; also, no other participating events (or subdivisions) have been taken into account. In a context where both pulse streams are given equal structural weight, a *balanced* polymetric structure may be inferred (first column). Such weight may be provided by the subdivision of the timespans associated with both pulse streams into more or less defined metrical groups as a result of the participation of shared and/or independent events. In addition to allowing for the synchronization of competing pulse trains (i.e., fully articulated pulse streams), subpulse density is also expressive of the degree of complexity involved in the interpretation of simultaneous streams as a metrical unit. Thus, the full synchronization a 4:3 polyrhythm (second row) will necessitate 12 pulse subdivisions and will generate a coincidence point every 12spu, a relatively low subpulse density (see **figure 3.1a**). Given a quick tempo and a lack of actual subpulse articulations, the ambiguity generated by the two competing pulse trains is likely to be neutralized, and the repeated coincidence point will emerge as the most measure-defining event.[22]

[22] In the case of an out-of-phase polyrhythm, the completion of the polyrhythmic cycle would correspond with the return of the two pulse streams to their initial position rather than to a coincidence point. If the IOI of the initial offset of a 4:3 polyrhythm is not synchronous with the subpulse stream defined by 1/12 of the polyrhythmic cycle, the subpulse density will be greatly increased and there will be no coincidence point. In such a case, the absence of coincidence points may be an additional factor that will encourage an interpretation of the polymetric structure as "balanced."

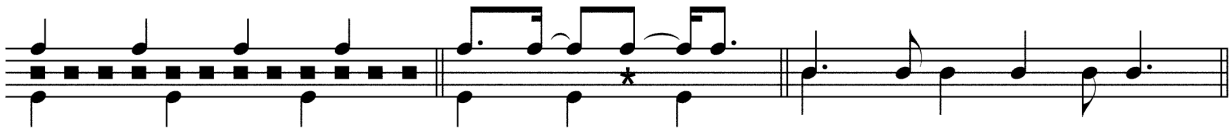
Figure 3.1: Details from a 4:3 structural polyrhythm (a. Subpulse density 12; b. Minor polarity; c. <312213> durational pattern)

Legend:

▪ = Pulse subdivision

* = Nested pulsation

a. Subpulse density 12 b. Minor polarity c. <312213> durational pattern



This situation is represented by models where the two streams have been fused, resulting in an *integrated* polymetric structure. In this example, opposite stem directions illustrate the unrealized polymetric potential of the two competing pulse trains and the double-stemmed note head corresponds to the coincidence point, which comes to function as the downbeat. For example, in the 4:3 structural polyrhythm (second row), the resulting stream projects the durational pattern <312213>, also known as the *resultant* or *composite rhythm* (see **figure 3.1c**).[23] Here, there are two possible metrical interpretations: a triple meter with a half note beat (3/2) or a quadruple meter with a dotted-quarter note beat (12/8). The perceived meter will depend on the specific characteristics of the musical surface (durational patterns, harmonic rhythm, prescribed accentuation and/or melodic contour), the manner of performance (especially applied accentuation and timing patterns), and a listener's biases.

Finally, competing metrical projections may result in *polarized* structures. A specific

[23] One of the properties of these composite rhythms is that they are palindromic (or non-retrogradable). This feature is discussed by Folio, "Analysis of Polyrhythms," pp. 106-107; Folio cites Yeston, *Stratification of Musical Rhythm*, p. 123.

polyrhythmic cycle, or portion thereof, could be understood as dominated by either its minor (**table 3.1**, second column) or major (**table 3.1**, third column) pulse stream; in both cases the competing pulse stream will be experienced as syncopated (or as providing cross-accentuation).[24] Structural factors related to the specific interaction of two given pulse streams may also be considered. For example, based on the possible interpretation of the third pulsation of the upper stream as a 2:1 subdivision of the timespan associated with the second pulsation of the lower stream, a 4:3 polyrhythmic compound may be qualified as showing a tendency toward minor *polarity*, or triple meter in this case (see **figure 3.1b**). While *nested* (or, in this case, *maximally synchronized*) pulsations may enhance the metrical potential of the concurrent pulse stream without jeopardizing the realization of a given polymetric potential, a high degree of *nestedness*, i.e., the property by which an event onset (or its associated timespan) belongs to two or more pulse streams, may lead to the integration of the two strands into a single perceptible stream.

[24] In longer passages, polarity could also be determined as shifting from one stream to the other; I am referring here to analytical determination and not the general process by which listeners might shift their attention from one stream to another.

Table 3.2: Synchronicity as a measure of integration

Factor	Definition	Effect
Frequency of Coincidence Points	Number and/or recurrence of synchronous events that belong to concurrent pulse streams	• Frequent coincidence points favor integration, either in terms of total number of occurrence or relatively short time-span separating each point • The absence of coincidence points encourages balanced or polarized polymetrical interpretations
Convergence and Divergence	Specific interaction of pulse streams with no coincidence point; the pattern of convergence and divergence is referred to as a proximity cycle	• Out-of-phase cycles are more likely to give rise to a polarized polymeter • The proximity cycle may regulate the polarity of a given pulse stream
Subpulse Density	Product of the total number of pulses for each polyrhythmic cycle	• Low degree of subpulse density favors integration • High subpulse density may favor the interpretation of non-synchronous events as expressive variations

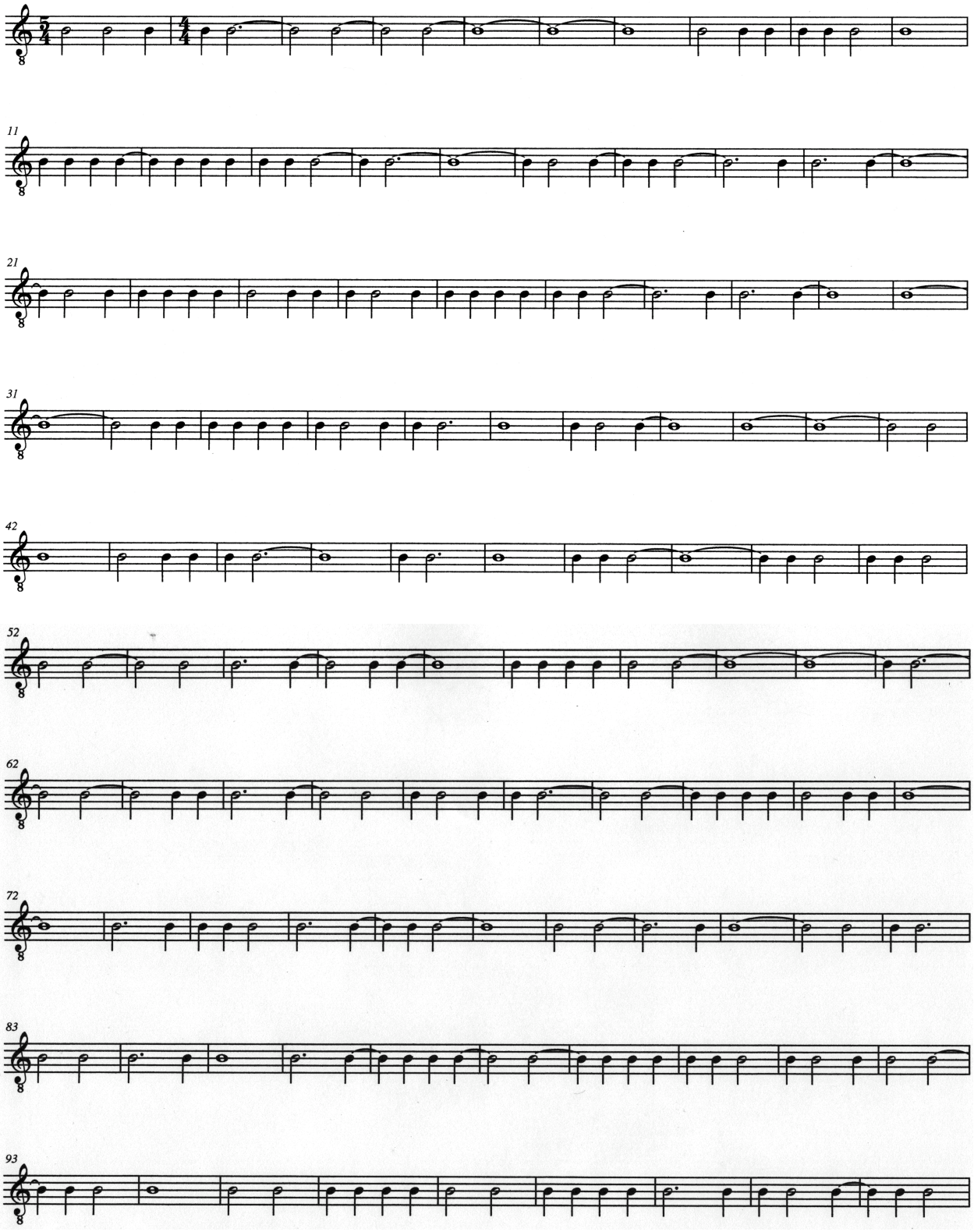
Table 3.3: Isochronicity as a measure of mensural determinacy

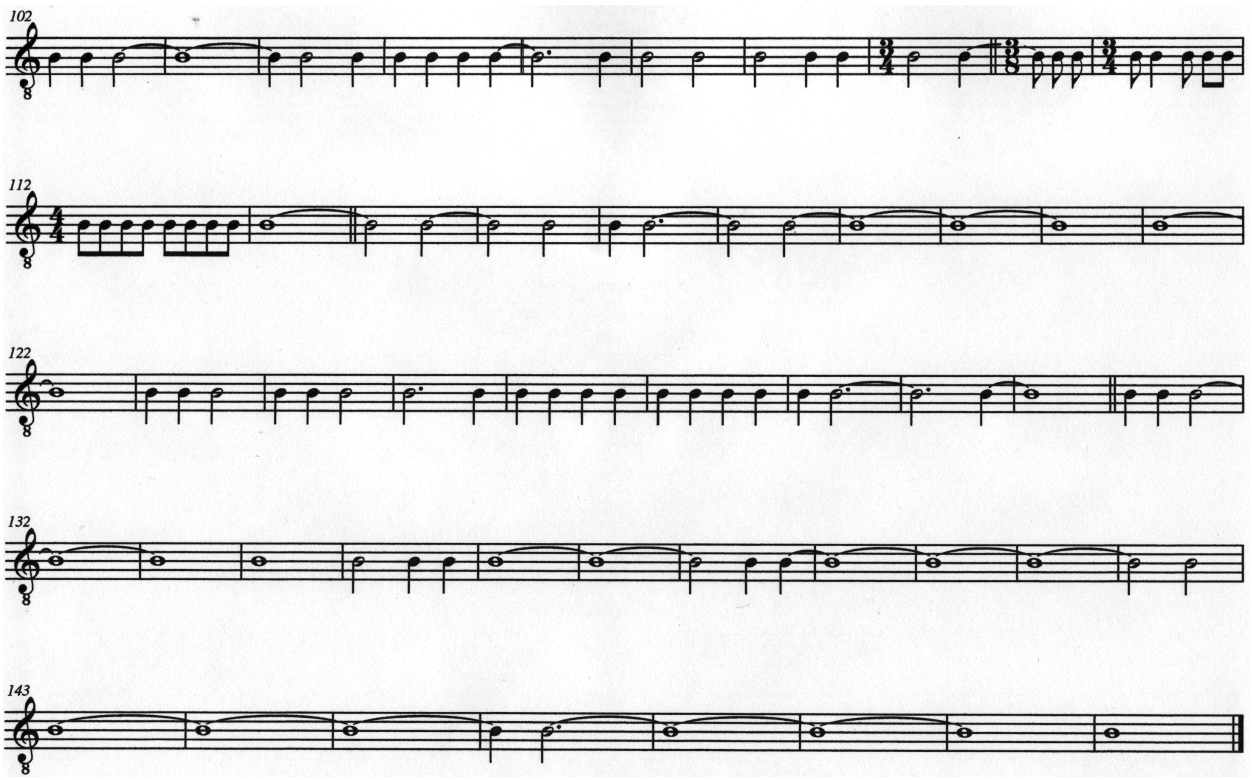
Factor	Definition	Effect
Durational Determinacy	Property of a timespan defined by two adjacent event onsets	• Recurrence of isochronous events is likely to contribute to the emergence of specific metrical projections
Periodicity	Adjacent recurrence of equivalent timespans that may or may not be subdivided	• Timespans governed by simple ratios are more likely to produce a hierarchical structure
Projective Field	Timespan associated with a given metrical projection	• Lack of activity on a specific level may lead to mensural indeterminacy • Recurrence of participative events on a specific level will contribute to the endurance of a given projection
Nestedness	Participation of relevant timespans in concurrent pulse stream	• Shared events may enhance the mensural determinacy of one or both layers • An excessive number of shared events might contribute to the integration of pulse streams into a single perceptible stream

A preliminary guide to the analysis of polymetric potential is offered as **tables 3.2** and **3.3**, where two complementary properties are associated with various structural factors that may influence the polymetric potential of a given texture. *Synchronicity* (**table 3.2**) refers to the property by which concurrent events are conceptualized as operating within the same temporal layer, while *isochronicity* (**table 3.3**) is the property by which successive events are related by simple ratios (1:1, 2:1, 3:1 etc.).[25] While both isochronous and synchronous events are defined as events whose duration can be determined according to a shared pulse stream, synchronous events will also necessarily be in-phase with each other. In exploring polymetric potential and realization in Carter's textures, special consideration must be given to his notational practice. As we have seen in chapter 2 (subsection 2.1.2), the bar-measure allows the performer(s) to conceptualize how the parts work together and is not meant to translate automatically into a performed accentual structure. However, events that begin on a notated beat – whether they are measured in terms of triplet eighths, sixteenths or quintuplet sixteenths – are usually nested in the concurrent elementary pulse stream, and the notated beat can be understood as a manifestation of their synchronicity. In the following section, I will demonstrate how reductive tools can be used to tackle the metrical potential of concurrent rhythmic strands whose events are determined by contrasting pulse streams.

[25] Although "isochronicity" usually refers to equal (or equivalent) durations, i.e., a 1:1 ratio, the concept is extended here to include groupings and subdivisions that are based on the perception of a recurrent durational unit and result in whole number multiples.

Example 3.7: Pulse stream reduction, middleground level (MPU = quarter), *Changes*





3.2 Local Conditions: Mensural Determinacy and Pulse Stream Reductions

A *pulse stream reduction* is a temporally ordered representation of the inter-onset intervals between synchronous events according to a given pulse stream; it makes use of musical notation and is re-barred according to the governing pulse stream, with each measure number corresponding to a specific bar-measure or pulse-measure.[26] As opposed to standard reductions, which are usually based on all surface events in a passage, segregated textural layer or rhythmic strand, pulse stream reductions operate from the perspective of a dominant pulse that integrates all synchronous events; events that are not synchronous are not included and conceptualized as syncopated. Since this dissertation deals almost exclusively with rhythmic organization, both standard and pulse stream reductions will only show durational

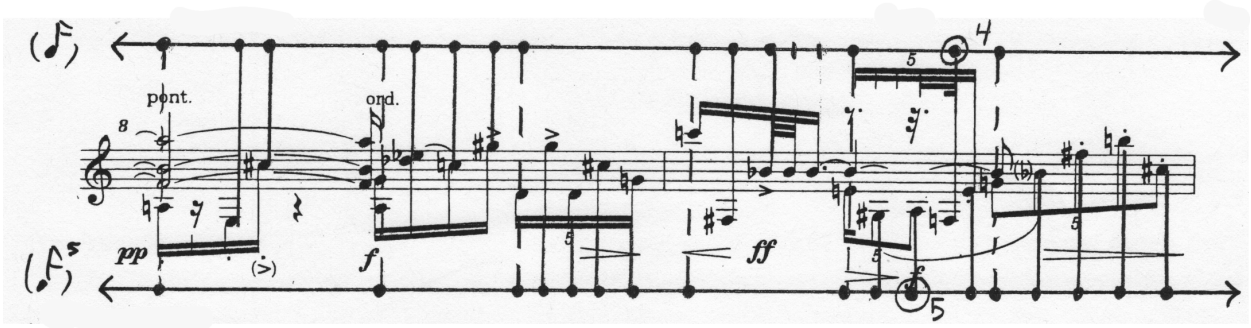
[26] In am indebted to Brian Moseley for suggesting the expression "pulse stream reduction" to describe this type of reduction.

values (unspecified pitch). I will now make a few observations that will clarify the advantages and limitations offered by the use of pulse stream reductions. The specific examples taken from *Changes* for guitar will also serve to illustrate further some of Carter's compositional strategies and their potential for narrative.

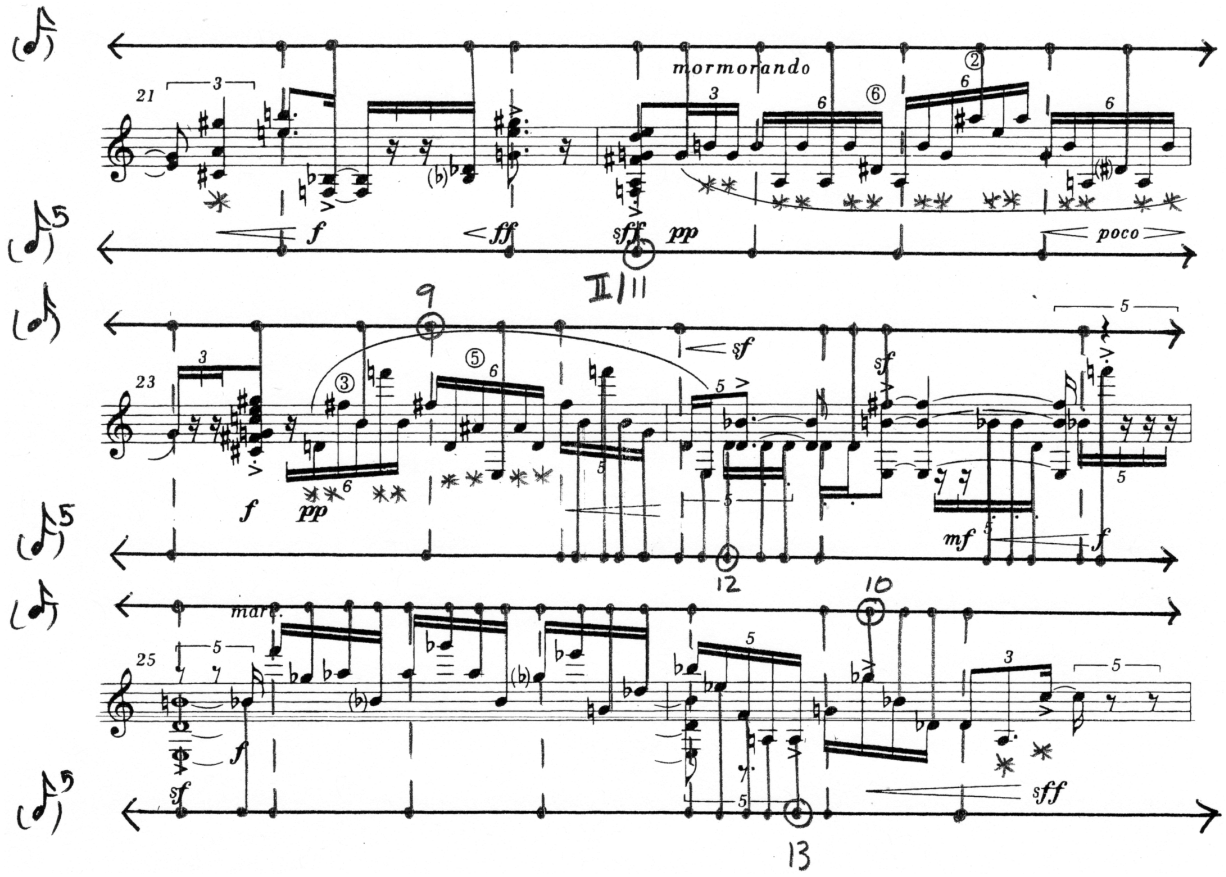
3.2.1 Middleground Level: Quarter Pulse Stream

In a pulse stream reduction, the particular make-up of measures can point to passages that show evidence either of *mensural determinacy* (or lack thereof), the latter referring to the likelihood that a given rhythmic strand will give rise to specific metrical projections. The reduction based on the bar-measure and the quarter beat pulse presents us with an alternation of simple rhythmic patterns and long held notes (see **example 3.7**). Bar-measures with quarter and half notes are indicative of a high degree of nestedness, which may result in a perceived predominance of the quarter beat and notated tempo of MM 100, although the actual perceived meter and tempo may depend on a variety of other factors, including accentuation, performance, and listening habits. For example, bar-measures 8-9 are strongly suggestive of the influence of the quarter-beat and bar-measure on metrical perception. However, when confronted with the musical surface, a given listener may be prompted to lock onto the tones marked with accents, which in this passage (**ex. 3.8**) would result in no regular sense of pulse, and consequently, no mensural determinacy, provided that upon each new event onset, the listener would be called to project a new duration.

Example 3.8: Musical surface, *Changes*, bar-measures 8-9



Example 3.9: Musical surface, *Changes*, bar-measures 21-26



Given the high degree of irregularity on the surface in general, it seems reasonable to posit that passages that exhibit quarter-beat regularity for a more significant amount of time may be more conducive to an interpretation of the quarter-beat as mensurally determinate,

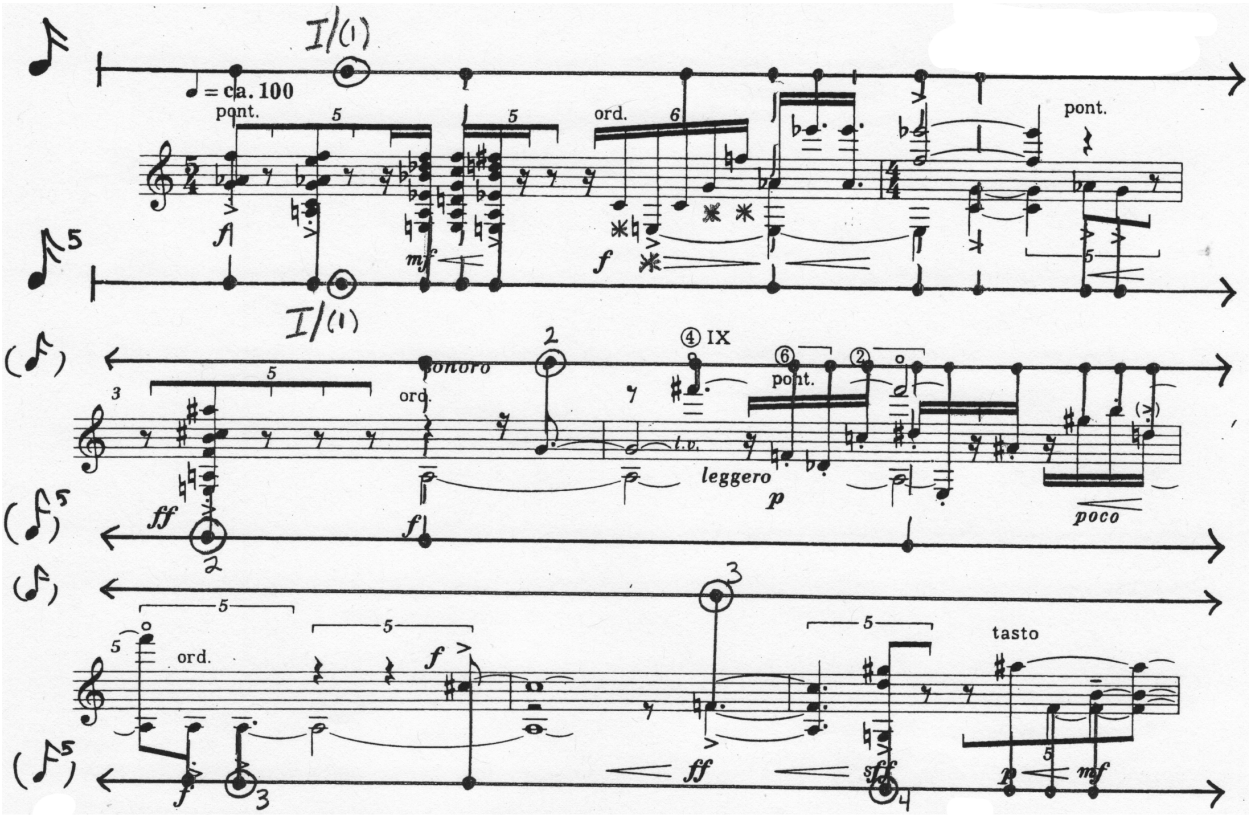
especially if the performer is influenced by the accentual pattern suggested by the notation. For example, compare **example 3.8** with bar-measures 22-26 (**example 3.9**); the latter are much more likely to be perceived as based on a quarter pulse with changing subdivisions and some syncopation. Conversely, passages with long notes in the reduction, especially when they involve several bar-measures, point to a lack of articulation of the quarter beat that is likely to give way to the influence of the two implicit structural pulse streams. For example, in bar-measures 5-7 (compare **ex. 3.7** and **ex. 3.10**), it is very unlikely that a listener will perceive the quarter-beat. In this passage (mm. 5-7), the series of accented tones {A, C#, F} may be heard as forming a triple measure followed by the "downbeat" chord {G, D, G#}, especially since these tones are almost equidistant and give rise to a tempo of about MM 36.[27] Although this triple measure includes hybrid timespans, it may be associated with the pulse stream based on quintuplet sixteenths, especially since it corresponds to a quasi-isochronous subdivision of pulse-measure X3 (lower time-line).[28] This hearing is strongly supported by the accentuation of the boundary background pulsations X3 (durational or agogic accent) and X4 (dynamic and attack density accent), which come to function as local "downbeats."

---

[27] The durational series for these three tones in subpulse units (spu) is {56, 58, 54}. In the present temporal context, i.e., quarter = MM 100 = 20spu, differences of 2spu and 4spu correspond to absolute timing differences of 60ms and 120ms, respectively. Given a basic beat of 1,680ms (56spu), these fluctuations represent less than +/-5%, which is within the set limits for successful synchronization in beat-finding experiments (e.g., Drake et al. 2000). In other words, supposing that this temporal sequence was performed with mechanical precision, the beats will be interpreted by the listener as equivalent. In an actual human performance, it is highly unlikely that the performer would achieve such precision and these durations would most probably be subjected to at least slightly greater temporal fluctuations articulating the performer's understanding of the temporal structure (i.e., *expressive timing*).

[28] Thus, 42 quintuplet sixteenths = MM 12 and MM 12 × 3 = MM 36 (multiplication here denotes a subdivision of the timespan associated with a given speed).

Example 3.10: Musical surface, *Changes*, bar-measures 1-7



In the opening bar-measures, the situation is not as clear-cut. Although the reduction (**ex. 3.7**) exhibits irregular subdivisions (bar-measures 1-2) and syncopated downbeats (bar-measures 3-4), the musical surface (**ex. 3.10**) suggests that some of these beats may be perceived as structural. Apart from the initial attack, other emphasized quarter beats include: the arrival on Ab in bar-measure 1, beat 4 (reinforced by dynamic shaping); the downbeat of bar-measure 2 (agogic and dynamic accent, dynamic shaping); and the polyphonic melody beginning in bar-measure 3, beat 3). Thus, while bar-measures 1-2 may be perceived as written at the bar level, the lack of a downbeat at bar-measure 3 and the mid-bar placement of the following phrase beginning suggest that the bar-measure loses some of its significance as the piece progresses (and thus prepares the terrain for the dominance of the quintuplet

sixteenths pulse stream in bar-measures 5-7). This passage also offers an interesting case study for one of Carter's compositional strategies, which involves the use of isochronous durations leading to a relatively accented event. Two dotted-sixteenth notes precede the downbeat of bar-measure 2. I will propose that such series direct the attention of the listener forward and mark the non-isochronous event that follows them in a special way.[29] Another example is offered in this passage (**ex. 3.10**) by the phrase beginning in bar-measure 3, where two isochronous events, the legato tones {A, G} exhibiting a duration of five sixteenths each, lead to the local apex F#5 in bar-measure 4. As can be seen throughout his works, Carter often uses this rhythmic/metric process to give added weight to events that are otherwise structurally significant, such as background pulsations and phrase beginnings.

3.2.2 Foreground Level: Elementary Pulse Streams

The pulse stream reductions of the two elementary pulse streams (**examples 3.11** and **3.12**) provide a very different picture of the piece. For the purpose of clarity and comparison, in these reductions both elementary pulse streams are represented in sixteenth notes and each notational value can be understood to correspond to a specific number of elementary pulse units (see **table 3.4**). Thus, in the pulse stream reduction of the X stream (**ex. 3.11**), the beginning time signature is 42/16 and the two associated tempi are MM 12 for the pulse-measure and MM 500 for the quintuplet sixteenth pulse, while in that of the "O" stream (**ex. 3.12**), the beginning time signature is 45/16 and the two associated tempi are MM 9 for the pulse-measure and MM 400 for the sixteenth pulse. It should be noted that in these

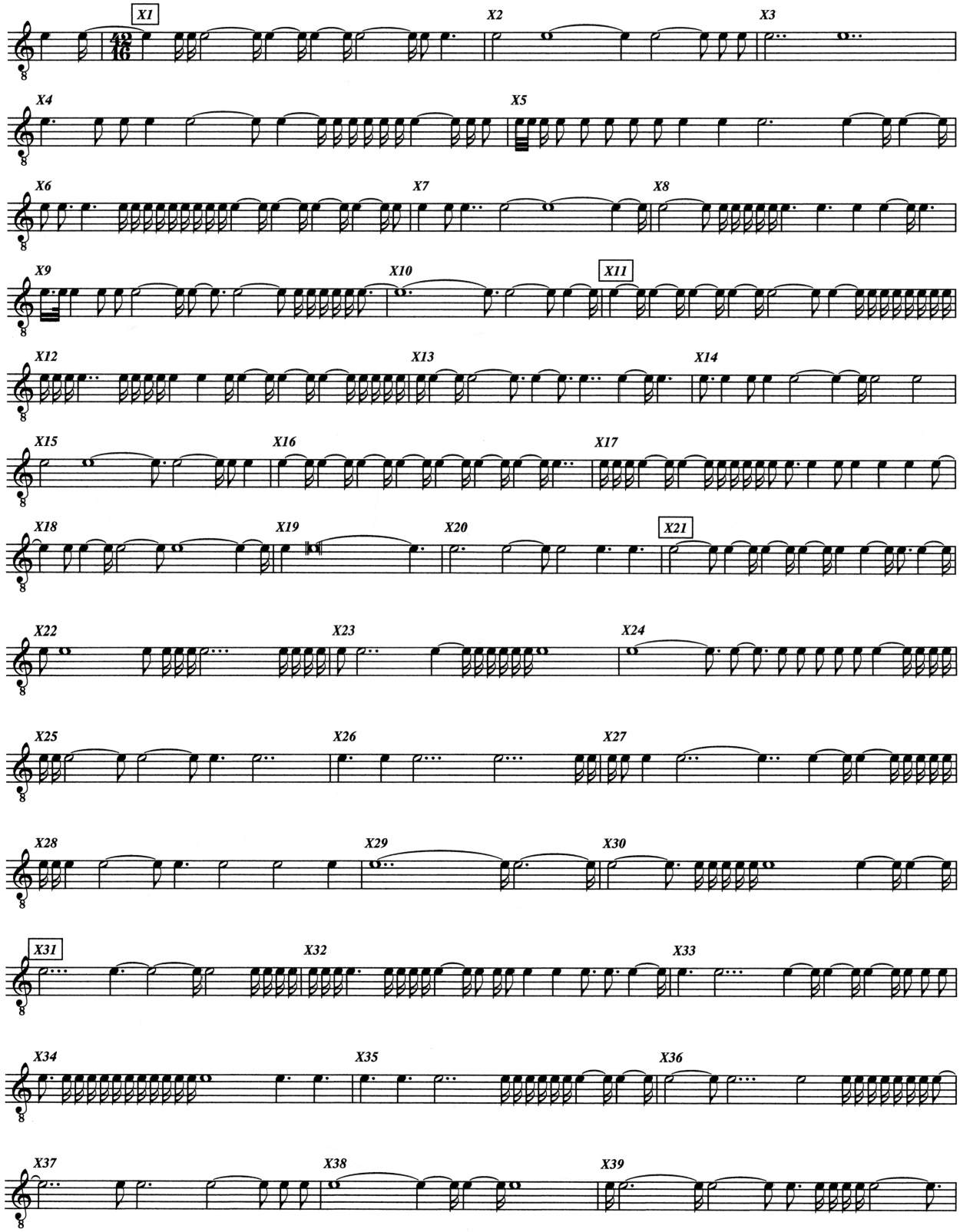
[29] This proposition is partly influenced by Hasty's theory of metrical projection and Carter's general interest in anacrustic gestures, which he traces back to the teachings of Nadia Boulanger; see Elliott Carter, *Collected Essays*, pp. 285-286.

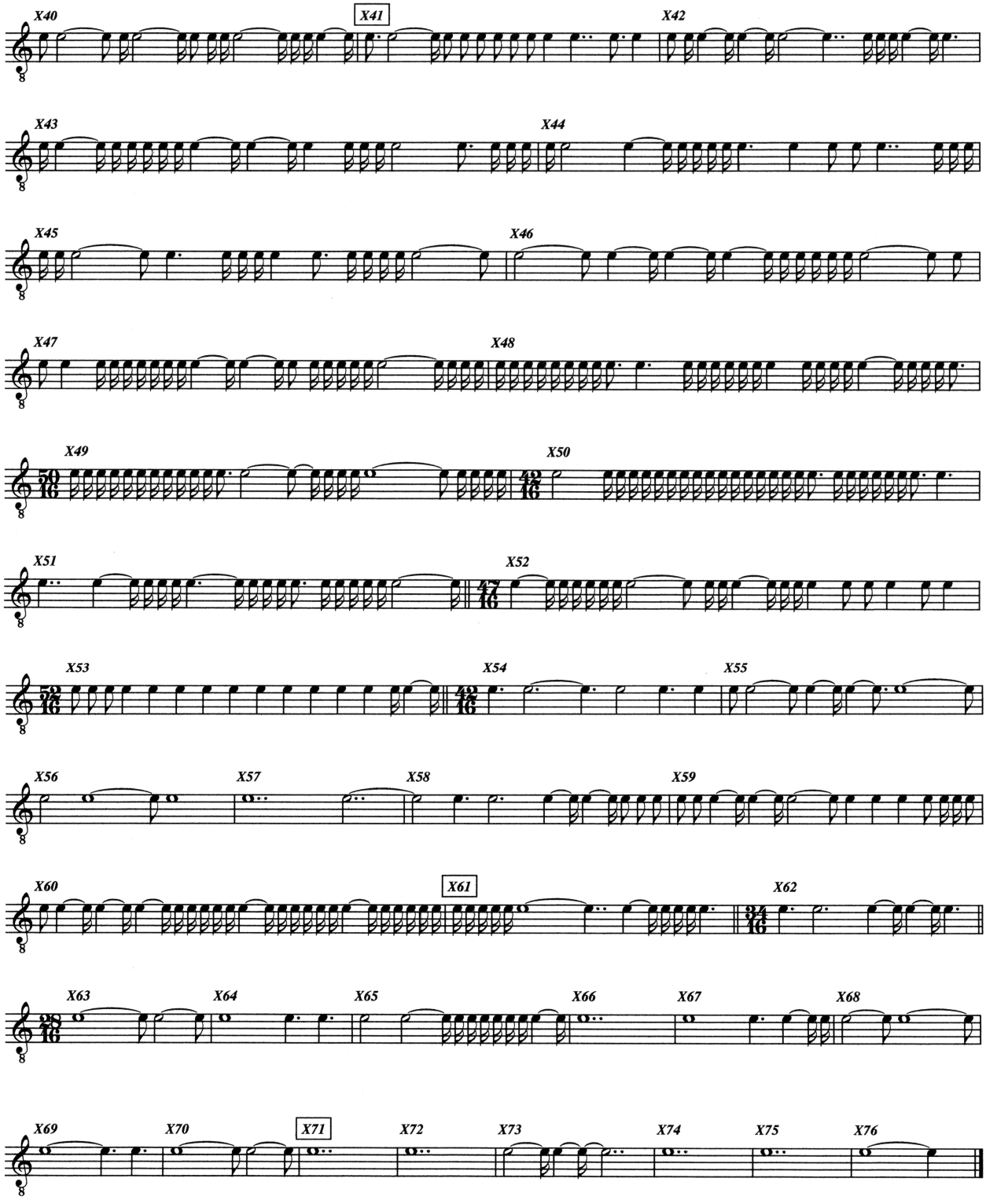
reductions, irregularities in pulse-measure duration become manifest (see X49, X52-X53, O37, and O40; the specific details of these irregularities were discussed in subsection 2.3.2.1).[30] The most noticeable difference between these two reductions and the previous one is that background pulsations are almost always articulated, as opposed to only 79 out of 150 articulated downbeats for the bar-measure (about 53%).[31] This fact attests to the structural ascendancy of the long-range polyrhythm over the notated metric grid. (A somewhat more detailed discussion of the background polyrhythm and its interaction with foreground and middleground was provided as subsection 2.3.2.)

[30] The pitch level of these pulse stream reductions ("X" stream = E5 and "O" stream = F4) was chosen arbitrarily.

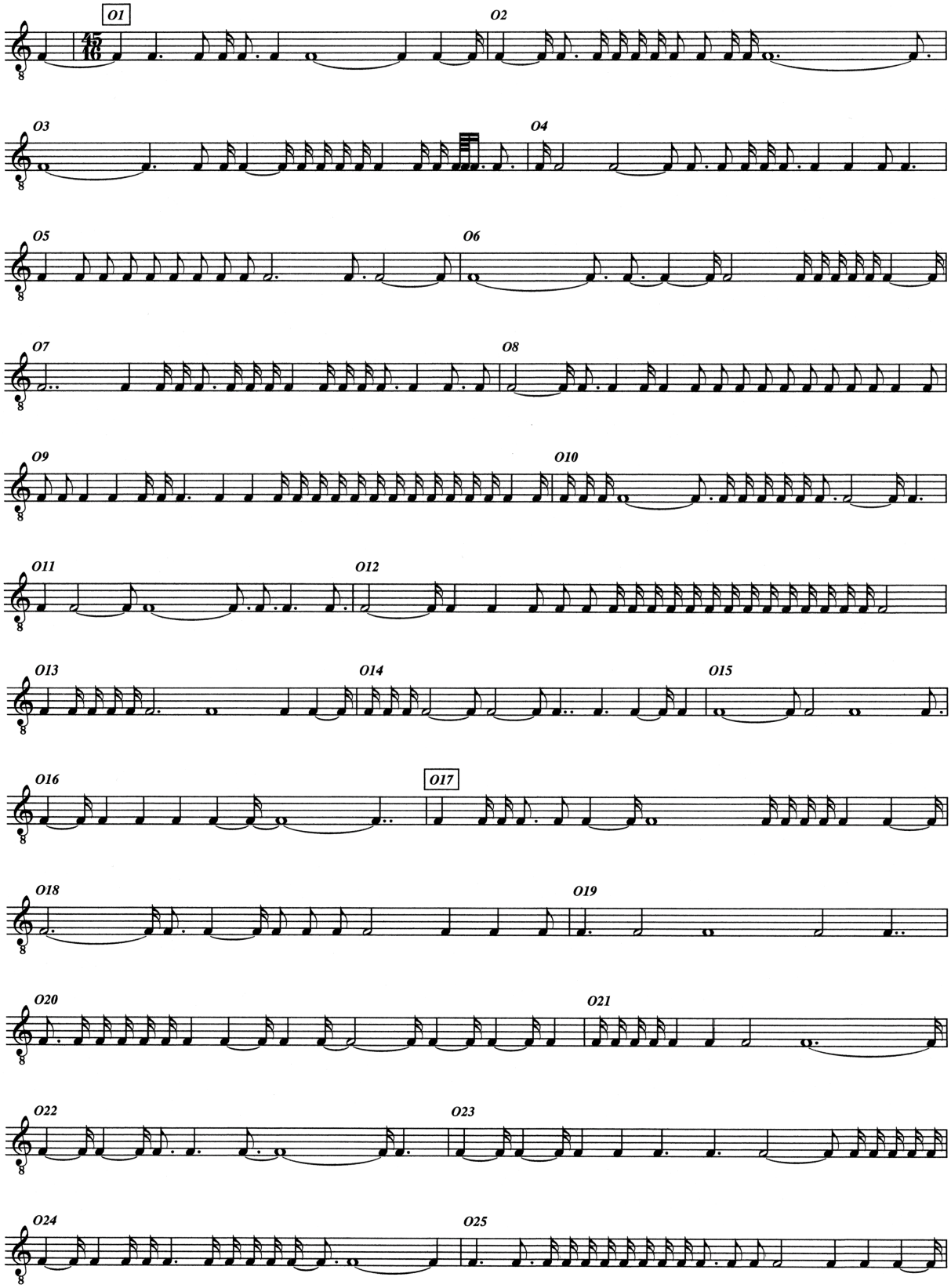
[31] The exact number of articulated background pulsations is difficult to establish because of the irregularities discussed in subsection 2.3.2.1; out of a total of 133 pulsations, 120 are regular (120/133 = 90%), 6 have been displaced by up to one quarter beat (including the initial coincidence point 'deleted' by the insertion of a quarter rest), and 7 belong to the irregular section leading to and including the "outbust."

Example 3.11: Pulse stream reduction, "X" stream (EPU = quintuplet sixteenth), *Changes*




X40
X41
X42
X43
X44
X45
X46
X47
X48
X49
X50
X51
X52
X53
X54
X55
X56
X57
X58
X59
X60
X61
X62
X63
X64
X65
X66
X67
X68
X69
X70
X71
X72
X73
X74
X75
X76

Example 3.12: Pulse stream reduction, "O" stream (EPU = sixteenth), *Changes*



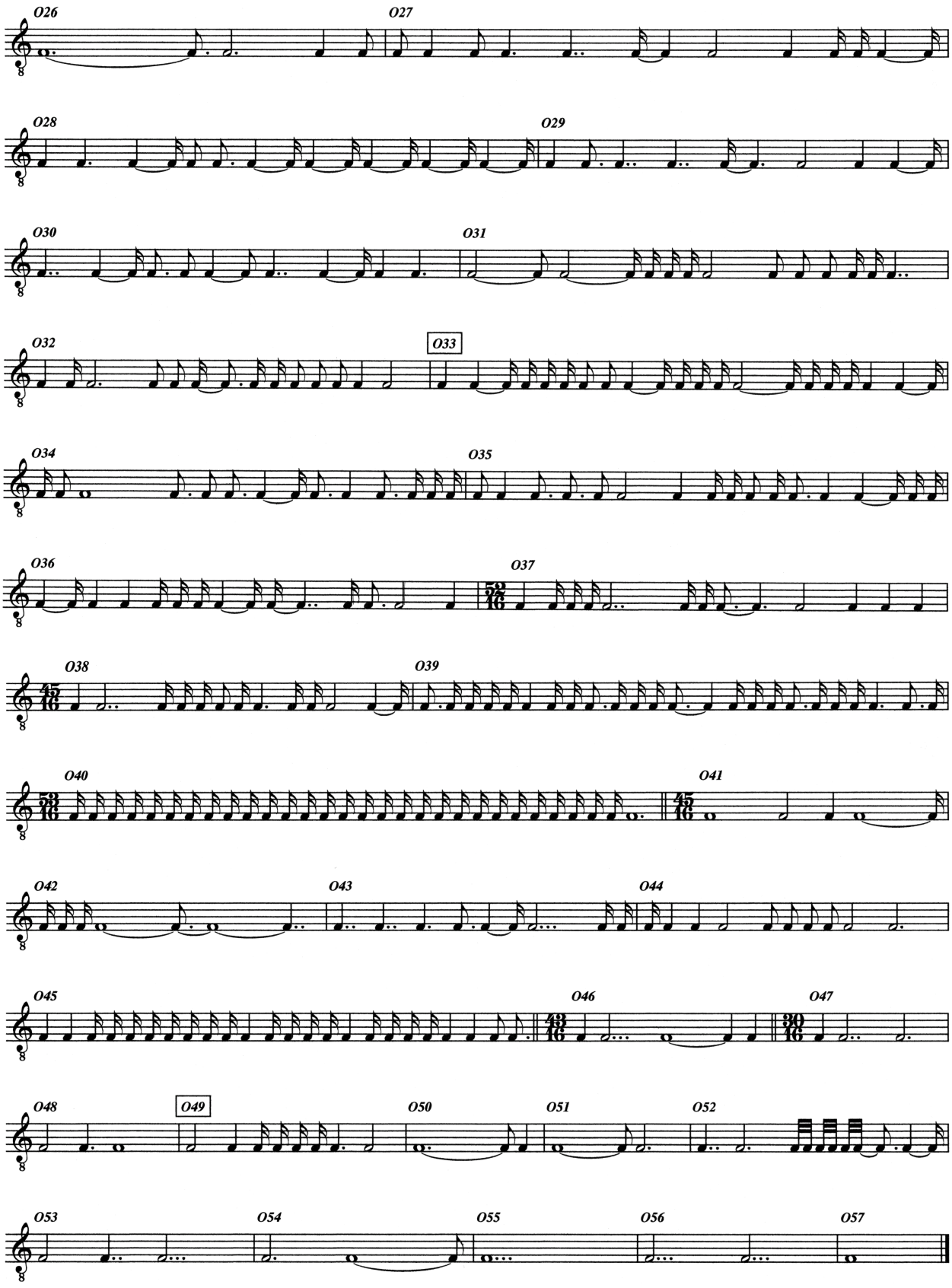

O26
O27
O28
O29
O30
O31
O32
O33
O34
O35
O36
O37
O38
O39
O40
O41
O42
O43
O44
O45
O46
O47
O48
O49
O50
O51
O52
O53
O54
O55
O56
O57

Table 3.4: Elementary pulse streams' notational equivalences for pulse stream reductions

Notational Value*	EPU Value
Sixteenth	1
Eighth	2
Dotted-eighth	3
Quarter	4
Quarter + sixteenth	5
Dotted quarter	6
Double-dotted quarter	7
Half	8
Half + sixteenth	9
Half + eighth	10
Half + dotted-eighth	11
Dotted-half	12
Dotted-half + sixteenth	13
Double-dotted half	14
Triple-dotted half	15
Whole	16
Whole + sixteenth	17
Whole + eighth	18
Whole + dotted-eighth	19
Whole + quarter	20

* The "+" sign is used to signify a tie.

The advantage of these two pulse stream reductions is that they allow the analyst to gauge the metrical potential of a single layer as manifested by the use of successive isochronous time-spans or durations that are related by simple proportions within a given pulse-measure, although a definitive interpretation cannot be made until the musical surface has been more thoroughly examined. The disadvantage is that the layers cannot be easily coordinated, a problem that is remedied by the use of these reductions in conjunction with the annotated score. For example, pulse-measure X3 (the previously discussed bar-measures 5-7; compare **ex. 3.10**, lower time-line, and **ex. 3.11**) is characterized by a SL pattern in a 2:1 ratio; in this case, a closer look at the annotated score confirms the relevance of 14 quintuplet sixteenths but also reveals that the second duration of 28 quintuplet sixteenths is subdivided

into two almost imperceptibly unequal events. Carter's notational choice here, i.e., the use of two streams to express an almost equivalent duration, is noteworthy: on one hand, it mimics the expressive timing of a performer; on the other, since it is the practice of Carter to associate contrasting pulse streams with different character-patterns, it suggests that this measure is in fact a dialogue between two dramatic personae.

Another example of the use of isochronous durations and simple proportions is offered by pulse-measures O2-O3 (compare **ex. 3.10**, upper time-line, and **ex. 3.12**). Following the directed motion to the F#5 in bar-measure 4 (the second rhythmic event in pulse-measure O2), the pulse stream reduction reveals a durational series of {3, 1, 1, 1, 1, 2, 2, 1, 1} (sixteenth = 1epu). This particular rhythmic design suggests a continued metrical grouping of three sixteenths, i.e. {[1+1+1], [1+2], [2+1], [1||]}, followed by *hiatus*, i.e. a break in the projective process.[32] Given the change of articulation and the continued presence of the F# (at least mentally, since the guitar cannot literally hold this pitch for the notated duration), the musical surface encourages a disjunction between the previous projection of five sixteenths and that of three sixteenths, thus creating a succession of irregular middleground measures.

Another phenomenon that can be traced is the use of written-out acceleration or deceleration patterns, and written-out rubato. Hudson identifies four types of rhythmic treatment in Carter's music: strict, accelerating, retarding, and free, the latter three being

---

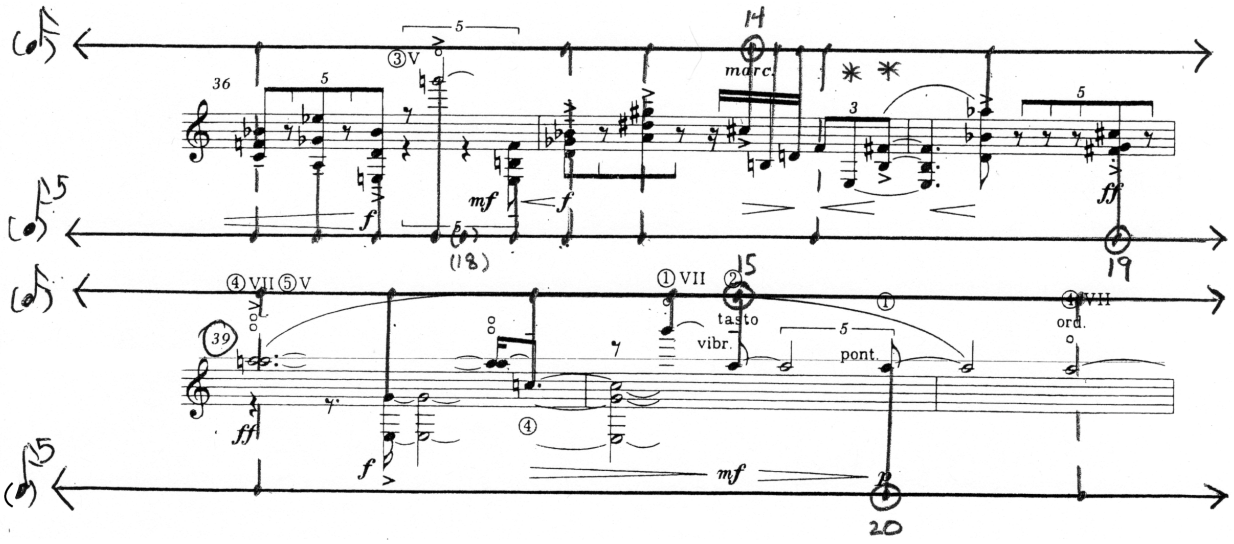
[32] The use of the term hiatus in this context is borrowed from Hasty, *Meter as Rhythm*. It should be noted that based on various elements of musical structure (e.g., contour) and/or performance manner, there are several interpretations possible, a fact that attests to the elusive and often ambiguous character of Carter's rhythmic/metric practice. The chosen interpretation is based on using the duration of the F#5 in bar-measure 4 as metrically determining.

linked to the practice of *tempo rubato*.[33] In this dissertation, written-out accelerations (or decelerations) are defined as series of durations that are increasingly shorter (or longer) either by some constant ratio or by the successive subtraction or addition of referential time units, while a written-out rubato is understood as a succession of closely related durations that could be equalized, i.e. reduced to isochronous durations. In addition, these concepts can be expanded to include groups of tones, the total timespan of which is increasingly smaller or larger, thus resulting in grouping acceleration or deceleration. The simpler form of acceleration is exemplified by pulse-measure O14 (**ex. 3.12**), which presents us with the durational series {1, 1, 1, 10, 10, 7, 6, 5, 4}, the second half of which exhibits an acceleration by the successive subtraction of 1epu (EPU = sixteenth; quarter beat = 5epu). A closer look at the musical surface (**ex. 3.13**) reveals that the first duration of 10 sixteenths (initiated by the F3 in m. 37, b. 4) is subdivided by non-synchronous triplet-eighth notes, which results in a slight deceleration at the surface from the initial series of sixteenth notes. The second timespan of 10epu is subdivided by a chord linked to the quintuplet sixteenth pulse stream (34spu + 16spu), which results in a LS pattern (or triple measure with unequal subdivision). The subsequent acceleration (beginning on the downbeat of bar-measure 39) is easily perceivable, even though the texture suggests two layers, i.e., a melody with accompaniment. Interestingly, this written-out acceleration leads to a hyper-downbeat, the onset of pulse-measure O15 in bar-measure 40 (upper time-line, A4), which is followed by X20 (lower time-line, a repetition of A4), thus setting up a point of near-convergence in the long-range polyrhythm. The resulting durational series (quarter beat = 20spu) from the last event in

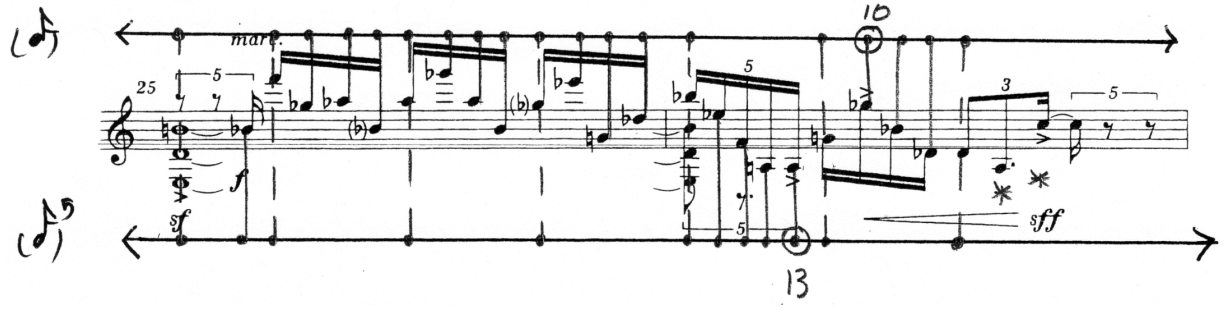
[33] Hudson, *Stolen Time*, p. 420. "Tempo rubato" refers to the practice of systematically inducing tempo fluctuations in the performance of a musical work.

pulse-measure O14 (upper time-line, B5) to the onset of X20 is {20, 42, 48}; given the ratio of about 1:2 from the end of the acceleration to the first background pulsation, it seems reasonable to conceive that the accelerating series will create a forward momentum that will diffuse its energy onto the onset of pulse-measure O15, thus giving added accentual weight to it (note that Carter has also marked the first A with a tenuto accent as well as *vibrato* and *tasto* markings), a compelling example of interaction between structural levels (in this case, surface and background events).

Example 3.13: Musical surface, *Changes*, bar-measures 36-41

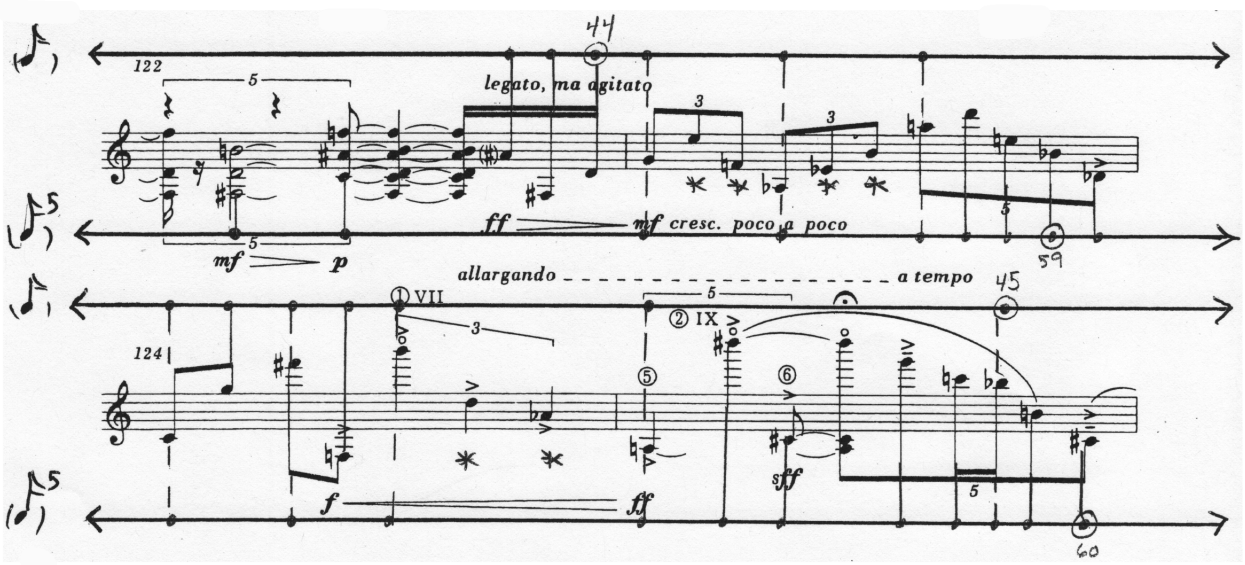


Example 3.14: Musical surface, *Changes*, bar-measures 25-26



Such accelerating patterns are also found on the musical surface resulting from the progressive change of subdivision values and are thus less apparent in the pulse stream reductions. For example, in bar-measure 26 (**ex. 3.14**), five quintuplet sixteenths give way to four sixteenths that are followed by a triplet eighth, dotted triplet-eighth, and a hybrid duration followed by a rest (an additional sixteenth rest in bar-measure 27 is not shown in the example), resulting in a written-out deceleration with a durational series of {4, 4, 4, 4, 4, 5, 5, 5, 5, 6 2/3, 10, 23 1/3} (quarter note = 20spu). The phrase beginning at the end of bar-measure 122 (*legato ma agitato*) provides a more extended example of this kind of process. In **example 3.15**, the deceleration is created by the successive subdivision of quarter beats and half notes into increasingly larger durations, from sixteenths to quintuplet quarters (compare with **ex. 3.11**, X58-X59, and **ex. 3.12**, O43-O44). The introduction of triplet quarters in bar-measure 124 is also accompanied by a performance marking that amplifies this deceleration (*allargando*). These two examples, where some or all of the deceleration is created by series of isochronous durations, each of which sets a temporary tempo, offer a rhythmic counterpart to the technique of terraced dynamics. Finally, although some melodies exhibit written-out rubato, in many cases an overall written-out rubato effect is created by the juxtaposition of short gestures exhibiting acceleration and deceleration patterns (e.g., pulse-measure O14, **ex. 3.13**).

Example 3.15: Musical surface, *Changes*, bar-measures 122-125



3.3 Global Conditions: Density and Nestedness Analyses

Density and nestedness analyses can be used to tackle *polymetric potential*, i.e., the likelihood that a given texture may give rise to simultaneous metrical interpretations or projections that are non-synchronous. Two basic types of *event density* measures will be used in this dissertation; *event* density, as used here, is to be distinguished from *attack* density in which one may wish to account for the number of tones that are sounded simultaneously. *Pulse density* (A/P or A/MPU) is a measure of the degree of activity within a particular pulse level that compares the number of articulated time-points (A) over the total number of implied pulse units (P) in a given timespan (or segment).[34] *Surface density* (T/P) will be calculated by accounting for the number of total events (T) over the total number of implied pulsations at a given pulse level (usually a middleground pulse unit). The assumption guiding the interpretation of data from a pulse density analysis is that a pulse level with a high

[34] Pulsation totals will be given in terms of elementary pulse units or middleground pulse units (P_{EPU} or P_{MPU}). The necessity for these two different ways of measuring pulse density will become apparent in the following subsections.

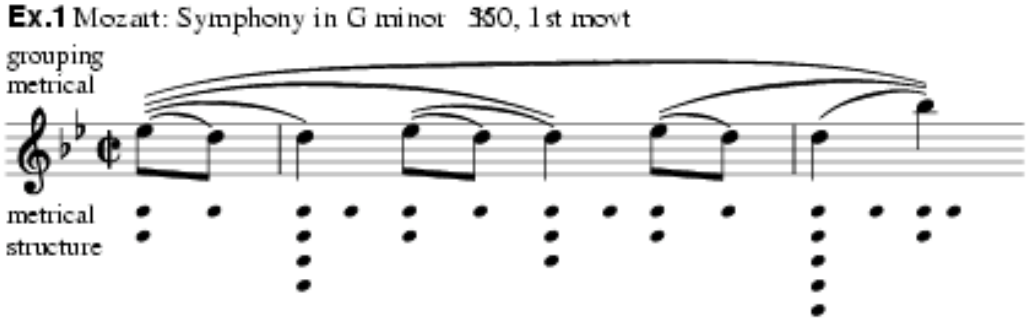
density will attract more attention and thus density may be used as a factor in the determination of which pulse level(s) may emerge as the prevalent beat(s).

A *nestedness analysis* is similar to a density analysis in that it measures the activity associated within pulse streams, but it also allows for the weighing of the relative influence of pulse levels in relation to each other. While we may speak of a degree of *potential nestedness* (i.e., two pulse streams may have a certain pre-determined nestedness value), my analysis focuses on *actual nestedness* (i.e., the actual number of articulated time-points that are shared by structural pulse streams). Relative levels of *polarity* and *integration* assess the extent to which a given pulse has potential for mensural determinacy, and are measured by comparing the number of *independent* (I) events and shared or *nested* (N) events. Here, the assumption is that a pulse level with higher polarity is more likely to support metrical projections and that a high degree of integration may prevent competing pulse streams from gaining ascendancy over each other, although such judgment will also depend on the mensural determinacy as suggested by the particular durational and accentual patterns present on the musical surface.

In order for the reader to become familiar with these analytical tools and associated terminology, and thus to be in a better position to assess the results of such analyses in the context of Carter's polymetric practice, this section will be divided into two sub-sections: (1) a case study of an excerpt of common-practice repertoire to help highlight the differences between that repertoire (and the ideas of meter it carries) and Carter's compositional practice; (2) a presentation of the segmentation and a brief discussion of the results of the density and nestedness analyses of *Changes* for guitar.

3.3.1 Case Study No. 1: The Opening of Mozart's Symphony in G minor, K. 550

Example 3.16: Mozart Symphony in G minor, K, 550, I, mm. 1-2 (melody only)[35]



In discussing the above example, Clarke and Parncutt note:

> Grouping and metrical structures are intertwined at all levels of rhythmic organization. For example, the first ten notes of the main melody of Mozart's G minor Symphony k550 (ex.1) imply as many as four different hierarchical levels of grouping (from two-note phrases to all ten notes) and five metrical levels (quavers, crotchets, minims, semibreves, double-semibreves). The second-last note (d") bears the greatest metrical accent because it belongs to all five metrical levels.[36]

Although only two levels are theoretically necessary to generate meter, in many common practice works there are several more pulse levels, each of which showing a high degree of potential nestedness, and thus leading to a high capacity for integration. The dot representation in the example above makes the time-points explicit and helps clarify this concept. Thus, the eighth-note pulse level is potentially nested within the quarter-note pulse

---

[35] Example reproduced from Clarke and Parncutt, "Psychology of Music," §II: 2 (permission for use requested and pending).

[36] Ibid. It is outside the scope of this dissertation to embark on a dispute with the authors' metrical interpretation. It should be noted, however, that the argument in support of such an interpretation rests with the practice of associating poetic meter and barline, in this case, an alignment of the last accented syllable and the downbeat in Italian Opera, which obviously represent only one type of alignment (i.e., "end-accented" as opposed to "beginning-accented" meter; see Rothstein, "National Styles").

level every two time-points; if all the eighth notes were articulated, this would represent a polarity of 50% and an integration of 50%. This relationship is repeated from the quarter-note pulse level to the half-note pulse level, and so forth. The result is a highly synchronous texture in which at least four different (but closely related) pulses could potentially emerge as the prevalent pulse (or beat), which is to say, all levels except the eighth-note level, which functions here as the elementary pulse unit (EPU).[37] Consequently, each durational value can be associated with a specific number of eighth notes (see **table 3.5**); these durational values readily express the simple ratio between levels. Here, the coordination of all levels does not require the use of a subpulse level. In such a situation, if pulse streams are called to compete, it is most likely to be across levels. Although alternate placement of the half- and whole-note pulse could be located (i.e., second and fourth quarters of each bar-measure for the half-note level, and second, third, or fourth quarter for the whole-note level), for the sake of conciseness, the time signature will be understood as metrically significant. The middleground pulse unit (MPU) could be defined either as the quarter- (MPU 1) or half-note (MPU 2) pulse level, while the background pulse unit (BPU) could be associated with the whole note (or bar-measure). In this passage, the two-bar level will not be considered; given the proper context (high speed, regularity, etc.), it too could be interpreted as a background pulse unit.

[37] The most prevalent beat is of course related to, though not exclusively representative of, the perceived tempo, which as mentioned earlier, also depends on pacing, that is, the special interaction between event density (measuring the number of events is a given timespan) and speed (defined as the rate of motion associated with recurring pulsations).

Table 3.5: Notational equivalences in elementary pulse units (EPU), Mozart, K. 550, I

Note value	♪	♩	𝅗𝅥	𝅝
EPU value	1	2	4	8

Table 3.6: Density and nestedness data for Mozart, K. 550, I, mm. 1-2 (melody only)

Pulse Level	**Pulsation Total (P)**	**Articulated Time-Points (A)**	**Pulse Density (A/P)**	**Polarity (I/A)**	**Integration (N/A)**
♪	14	10	0.71	30%	70%
♩	7	7	1	57%	43%
𝅗𝅥	3	3	1	33%	67%
𝅝	2	2	1	N/A	

Legend:
T = Total number of events
A = Number of articulated time-points at a specific pulse level
P = Number of implied pulsations (or time-points) at a specific pulse level
I = Number of independent events
N = Number of nested events

Let us now consider this brief excerpt from the point of view of density and actual nestedness. The results of this analysis are summarized below (see **table 3.6**). The eighth-note level has the highest number of time-point articulations (highest degree of pulse stream activity) and the quarter-note, half-note, and whole-note levels have the highest density, which can be expressed as a number between 0 and 1 (100% = 1). As noted previously, while pulse density is calculated over the total number of *implied* pulsations in a given pulse stream, nestedness is measured against the total number of *articulated* time-points in the

corresponding pulse stream: independent events yield a measure of polarity and nested events yield a measure of integration.[38] In the present excerpt, the eighth level has only 3 independent events, which yields a polarity of 30% (3/10 = 30%) as compared with a quarter-note integration of 70% (7/10 = 70%). Similarly, the quarter-note level has 4 independent events and 3 nested events within the half-note pulse, while the half-note level exhibits only 1 independent event as compared with 2 nested events.[39] Since the fifth level (the 2-bar level) is not being considered here, polarity and integration measures are not provided for the whole-note pulse level (polarity and integration are comparative measures). It is important to note that this particular example does not include competing pulse streams; rather, each level is partially nested within the next level in a 2:1 ratio. Thus, although polymetric potential is theoretically present, the total texture is much more likely to be perceived as a nested hierarchy. This is to be differentiated from an integrated texture, in which the beat level will be subjected to a high degree of syncopation and/or fluctuating sub-beat levels.

What conclusions can be drawn from this data? Although the eighth level has the lowest density, the fact that it has the highest total number of articulated time-points suggests that it will be influential on the perception of activity (and character) in a general sense. On the other hand, the quarter-note, half-note, and whole-note levels are all fully articulated. Of these, the quarter level has the highest polarity and the higher number of articulated time points and thus, one may be led to hypothesize that this is the level that is most likely to

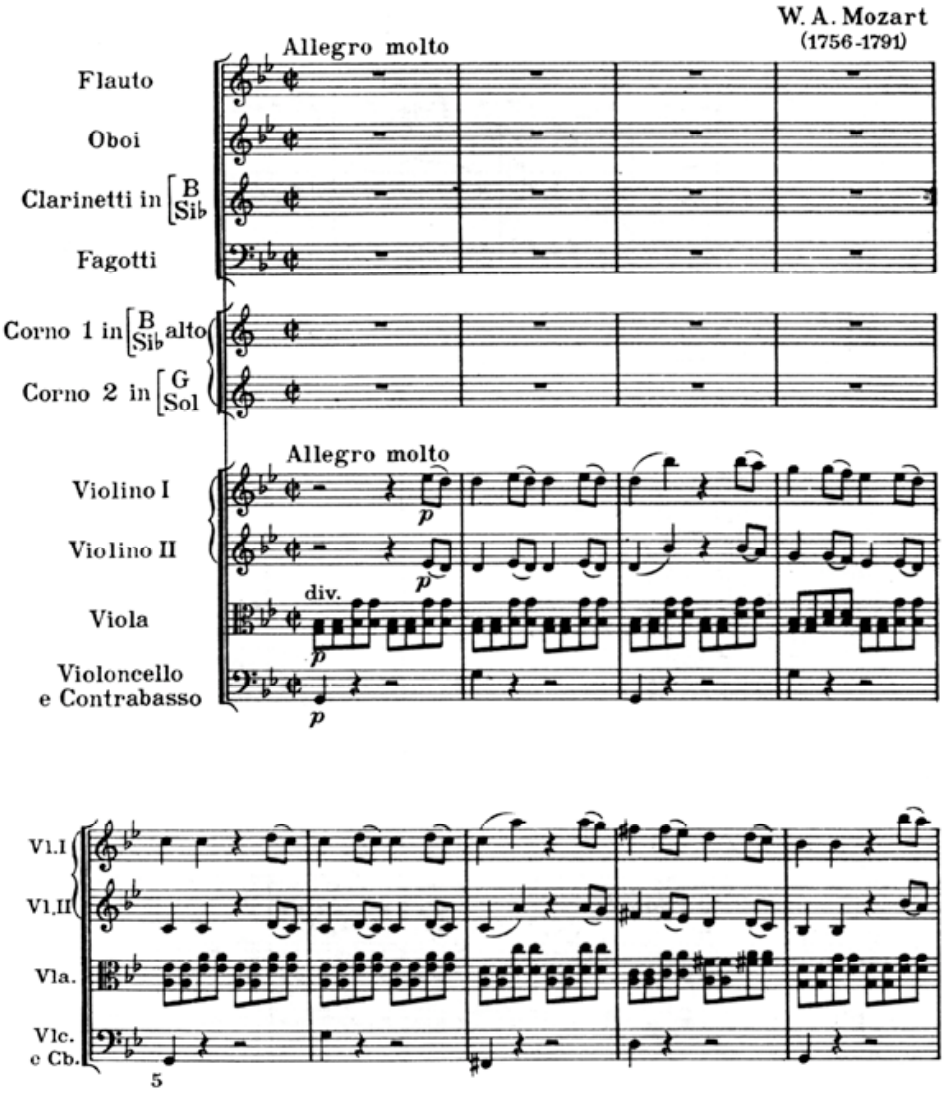
---

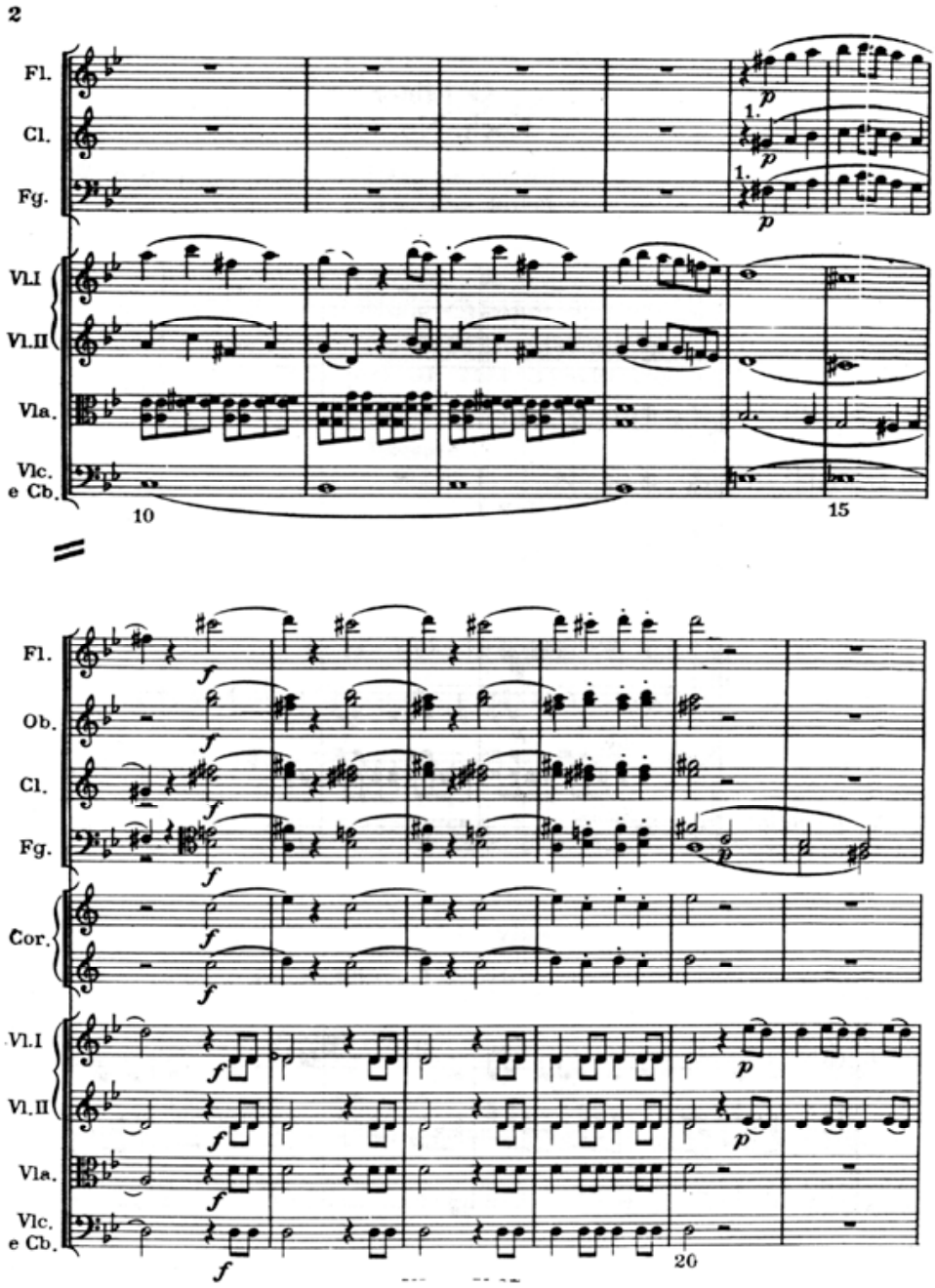
[38] Because polarity and integration measures are relative to each other, they will always add up to 100%. This interdependence is emphasized by the use of percentages (as opposed to density measures, which are expressed as numbers between 0 and 1).

[39] Since all pulse levels are synchronous, the number of independent events at a given pulse level is the difference between the number of articulated time-points between this level and the next higher level.

emerge as the dominant pulse stream (i.e., the prevalent beat). Of course, these are only two factors. We have not considered the suggested tempo (*Allegro molto*), which strongly suggest that the half note pulse might offer a more comfortable tactus for the listener to synchronize with. Nonetheless, the elements observed are part of the rhythmic (and metric) landscape of this theme.

Example 3.17: Mozart, K. 550, I, mm. 1-21 (full texture)





This symphony's melody incipit was considered in isolation. It seems reasonable to assume that the texture as a whole would give us a more complete picture of the interaction of pulse levels, and it might offer more support for the notated meter. Let us now consider the overall texture of the complete theme (bar-measures 1-20, downbeat; see **example 3.17**). Since

there are two potential middleground pulse units (MPU), surface density is calculated for each pulse level using the next pulse level as middleground unit: (1) if MPU = quarter, T/P_{MPU} = 130/77 = 1.69; (2) if MPU = half, T/P_{MPU} = 130/39 = 3.33 (see **table 3.7**).[40] As with the melodic opening (**ex. 3.16**), the eighth-note level has the highest degree of articulated time-points. However, if we take the entire phrase into consideration, the half note and whole note levels have the highest pulse density, adding more weight to these levels as compared with the quarter and eighth levels. Furthermore, the higher number of articulated time-point at the half-note level (as compared with the whole-note level) may be taken as additional support for the emergence of that level as the most stable beat. In term of polarity, we can also notice some differences.[41] The eighth pulse exhibits an increase in polarity (from 30% to 43%), while the quarter pulse shows a decrease (from 57% to 47%), although it is still slightly higher than that of the eighth. The half-note level, which also shows an increase in polarity (from 33% to 49%), is now the pulse level with the highest polarity (although it is only slightly higher than that of the quarter). In addition, it might be posited that the corresponding surface density is not so high as to make it too complex to be useful as the prevalent beat.[42]

---

[40] Surface density is not calculated based on the eighth-note and whole-note pulses since these represent the elementary and background pulse levels, respectively.

[41] As with the previous example, the number of independent events at a given pulse level is the difference between the number of articulated time-points between this level and the next higher level.

[42] This point is based on speculations concerning the number of subdivisions (including unequal subdivisions) that may be perceptually grouped as a beat. While the number of subdivisions likely to be accommodated by any given beat is largely dependent on the tempo associated with a given beat, there would appear to be certain limits that are more common, i.e., groups of between 2 and 7 elements.

Table 3.7: Density and nestedness analyses data of Mozart, K. 550, I, mm. 1-20, downbeat

Pulse Level	Pulsation Total (P)	Articulated Time-Points (A)	Surface Density (T/P_{MPU})	Pulse Density (A/P)	Polarity (I/A)	Integration (N/A)
♪	153	130*	N/A	0.85	43%	57%
♩	77	74	1.76	0.96	47%	53%
𝅗𝅥	39	39	3.33	1	49%	51%
𝅝	20	20	N/A	1	N/A	

* Since this texture is maximally synchronous, the number of articulated time-points (A) at the lowest pulse level (eighth note) also corresponds to the total number of events (T = 130).

Legend:
T = Total number of events
A = Number of articulated time-points at a specific pulse level
P = Number of implied pulsations (or time-points) at a specific pulse level
I = Number of independent events
N = Number of nested events
MPU = Middleground pulse unit

Not only can these observations provide us with a concrete way of discussing relative structural weight contextually (as opposed to unilaterally as when psychological tempo preferences are appealed to), but it also allows us to compare a single layer and the whole texture as well as smaller and larger segments. While the melody might have certain characteristics, the consideration of the context surrounding it and the potential conflict between these two perspectives may provide substance for the building of a temporal narrative, even in a texture as synchronous as this one. The normalizing tendency of event density towards the end of the phrase, i.e., its "evening-out" of differences in favor of the notated beat, provides an efficient vehicle for the conflict-resolution metaphor. Polarity could also be used as a measure of tension or relaxation, and we can imagine how easily successive segments of opposing character could be conceptualized as a dialogue (or an argument). Let

us now turn to the density and nestedness analyses of an example from the repertoire under examination.

3.3.2 Case Study No. 2: *Changes* for Guitar (1983)

It is beyond the scope of this paper to discuss segmentation in detail as such discussion would generate an entire dissertation in itself, especially if considering the interaction of temporal and textural elements of musical narrative. As mentioned previously, the segmentation proposed is not meant as definitive. Nonetheless, a few remarks will be made to clarify the segmentation process as it relates to Carter's style.

In producing any kind of segmentation, proximity and similarity are paramount.[43] Events that are close in time tend to group together; those that are separated by relatively long rests tend to be understood as separate; contrast will tend to produce a sense of new beginning and repetition, return of the same (thus grouping events that are perceived as "in between" the similar material). Salience was also an important factor in my analysis. *Marked events*, i.e., events that were perceived in contrast to surrounding events as imbued with a special meaning, were more likely to be heard as boundary points.[44] This particular phenomenon is especially significant in relation to Carter's background pulsations, which are often placed at or near the end and beginnings of segments and usually marked with a stress

---

[43] For representative studies that focus on segmentation for twentieth-century music, see Dora A. Hanninen, "A General Theory for Context-Sensitive Music Analysis: Applications to Four Works for Piano by Contemporary American Composers" (Ph.D. dissertation, University of Rochester, 1996), and Christopher F. Hasty, "Segmentation and Process in Post-Tonal Music," *Music Theory Spectrum* 3/1 (1981): 54-73.

[44] For a detailed exposition of the concept of "markedness" and its application to musical analysis, see Robert Hatten, *Musical Meaning in Beethoven: Markedness, Correlation, and Interpretation* (Bloomington, IN: Indiana University Press, 1994).

accent when this is the case. Thus, an important factor at work here is that the correlation of background pulsations with other markers was sometimes used as support for a given segmentation.

The size of segments that are based on the proximity principle also tends to vary considerably. While we may tend to think of categories of unit as having a typical size, there is no hard rule for actual size (even though it seems that in principle a gesture should be smaller than a sub-phrase, a sub-phrase smaller than a phrase, etc.). In segmenting a musical surface with a "bottom-up" approach, i.e., from gestures to sub-phrases, phrases, and sections, this might create some problems, especially in Carter's style, which privileges surface discontinuity and contrast, especially in solo works. Also, some phrases are more continuous or homogeneous than others and are thus not readily divisible into smaller segments, which means that a given unit of larger dimension might become apparent as early as the sub-phrase level. Perception of segments may also vary based on the type of (simulated) listening (i.e., prospective versus retrospective), and it tends to change as one become more familiar with the piece. Finally, as encouraged by Carter's comments, metrical modulations were also considered as potential section markers, although the exact points of juncture are not always clear.

3.3.2.1 Musical Surface and Pulse Activity: Sectional Density Analyses

Density analyses were performed by section and only for the surface, foreground, and middleground levels (see **tables 3.8** and **3.9**). In contrast to the Mozart example discussed in the previous sub-section, which could be characterized as over-determined, from the point of view of pulse synchronization, Carter's mature works usually involve at least two competing

elementary pulse streams. Thus in *Changes*, the two elementary pulse streams express a ratio of 5:4 (i.e., as sixteenths subdivisions of the quarter middleground pulse unit), a ratio that is only occasional reproduced at the next higher level (i.e., eighth subdivisions of the half note), as opposed to the Mozart, where each pulse level relates to the next in a 2:1 ratio. Furthermore, the grouping of a different number of elementary pulse units from each of these pulse streams generates two background pulse streams that are even more non-synchronous since they form a ratio of 75:56.

Table 3.8: Sectional density analysis data for *Changes*, middleground level

Section	Tempo	Bar-Measures	Pulsation Total (P_{MPU})	Surface Density		Pulse Density	
				T	T/ P_{MPU}	A	A/P_{MPU}
Opening	T_1	1-14	53	100	1.89	23	0.43
Episode 1	T_1	14-27	54	134	2.48	28	0.52
Episode 2	T_1	27-53	103	153	1.49	41	0.40
Episode 3	T_1	53-89	135	233	1.73	58	0.43
Scherzando	$T_1 \rightarrow T_2$	89-113	105	298	2.84	66	0.63
Episode 4	T_1	113-130	69	105	1.52	25	0.36
Epilogue	T_3	131-150	80	59	0.74	14	0.18

Legend: T = Total number of events
A = Number of articulated time-points at a specific pulse level
P = Number of implied pulsations (or time-points) at a specific pulse level
MPU = Middleground pulse unit (quarter pulse at T_1 and T_3; eighth pulse at T_2)

The density analysis presented in **table 3.8** includes all surface events. Here, *surface density* is measured against the pulsation total of the first middleground pulse unit for each tempo section, i.e., the quarter pulse level at T_1 and T_3, and the eighth pulse level at T_2 (T_2 is used only for the "outburst" passage, bar-measures 110-113).[45] The progression of surface

[45] It should be noted that the last event in bar-measure 113 belongs to the following section.

density provides us with a kind of contour for the rhythmic activity over the timespan of the piece, which could be expressed as <4, 5, 1, 3, 6, 2, 0>.[46] Given its cardinality ("4") and the overall surface density (T/P_{MPU} = 1082/599 = 1.81), the piece opens with relatively high level of density. We can see that from Episode 2 to the *Scherzando* section, there is a gradual density increase from relatively low density ("1") to the highest density ("6"); prior to that climax, there was a similar increase (albeit to a slightly lesser degree) from the Opening section to the Episode 1 ("4" to "5"). Given the character of the "outburst" within the *Scherzando* and Carter's aesthetic leaning toward building a musical discourse that is expressive of a dramatic confrontation between characters, this increase in surface event density could be interpreted as representative of the rising tension between the two characters as represented by the two competing elementary pulse levels. This musical narrative is especially interesting because of what happens following the *Scherzando*. Thus, after the *Scherzando*, there is a steep decline in density, with the Epilogue reaching the lowest degree of density ("0" = 0.74 events per quarter beat). Given the proposed argument scenario and the fact that the piece ends with the first coincidence point between the two background pulse streams, I think that this metaphor is particularly attractive. There is also a certain magnification effect at work here when considering the intensity as measured by surface density and sectional length. Thus, sections in *Changes* range in size from 53mpu to 135mpu, with the longest being the Scherzando (closely followed by the section leading to it, i.e.,

---

[46] In this notation, density contour is presented as a temporally ordered series in which "0" represents the section with the lowest density and that with the highest is represented by the cardinality of the set minus 1; for a related application of rhythmic contour theory (based on relative duration) to twentieth-century music, see Elizabeth West Marvin, "The Perception of Rhythm in Non-Tonal Music: Rhythmic Contours in the Music of Edgard Varèse," *Music Theory Spectrum* 13/1 (1991): 61-78.

Episode 3) and the shortest the Epilogue (closely followed by the Opening, when taking the 3:2 quarter beat speed ratio of the metrical modulation into consideration). The change of tempo of the quarter beat also has an effect on perceived density as a slower middleground pulse unit means less density (and in this case, perceived tempo is closely related to pacing).

Although it is measured against the quarter beat (or its equivalent in the case of the metrical modulation to T_2), surface density is only a tool for describing the rise and fall of activity level and does not provide information about the potential significance of the quarter pulse. Thus, when compared with an average surface density of 1.81, the quarter pulse density is rather low, i.e., only 0.43 articulated time-points per notated quarter beat (A/P_{MPU}) = 255/599 = 0.43). Consequently, the range of variation in density is also significantly lower, a difference of 0.45 for quarter density (A/P) as opposed to 2.1 for all surface events (T/P_{MPU}). The progression of pulse density for the quarter pulse level can also be expressed as a pulse density contour segment of <3, 4, 2, 3, 5, 1, 0>. Apart from the fact that there are two sections that share the density index 0.43 (Opening and Episode 3) and that the opening begins at its average density level, the two density contours are identical.[47]

Let us now examine pulse density at the foreground level. Because the presence of two contrasting elementary pulse streams necessarily includes articulations at the quarter pulse level, the examination of pulse activity at the foreground level will be most meaningful in the context of a nestedness analysis. Thus, I will presently limit my observations to a comparison of the two elementary pulse streams with each other. Here, it will prove useful to

[47] Had these two contours been different, this would have provided evidence for increased care in the avoidance of the notated quarter, which corresponds to a coincidence point for the contrasting elementary pulse streams. As it stands, the composer's treatment of the notated quarter beat is consistent throughout.

measure pulse density in two different ways, that is, against the total number of elementary pulse units (A/P_{EPU}) and against the total number of middleground pulse units (A/P_{MPU}). The results of these analyses are summarized in **table 3.9**. Again, when comparing the contour segments of density variance (A/P_{MPU}), we are faced with a very similar pattern of rising and falling figures (with Scherzando and Epilogue at the two extremes), the sixteenth level yielding a contour segment of <4, 5, 2, 3, 6, 1, 0> and the quintuplet sixteenth level resulting in <4, 4, 1, 3, 5, 2, 0>, with the quintuplet sixteenth pulse stream showing a significantly greater range of variation in density, i.e., 1.48 as compared to 0.95 for the sixteenth pulse stream. This last point could be used to characterize the quintuplet sixteenth stream as more unstable, again an element that offers narrative potential.

Table 3.9: Sectional pulse density analysis data for *Changes*, foreground level

Section	Bar-Measures	Pulsation Total (P_{MPU})	Sixteenth "O" Stream			Quintuplet Sixteenth "X" Stream		
			A	A/P_{MPU}	A/P_{EPU}	A	A/P_{MPU}	A/P_{EPU}
Opening	1-14	53	53	1	0.25	62	1.17	0.23
Episode 1	14-27	54	70	1.30	0.32	63	1.17	0.23
Episode 2	27-53	103	88	0.85	0.21	95	0.92	0.18
Episode 3	53-89	135	126	0.93	0.23	154	1.14	0.23
Scherzando	89-113	105	145	1.38	0.35	202	1.92	0.38
Episode 4	113-130	69	53	0.77	0.19	67	0.97	0.19
Epilogue	131-150	80	34	0.43	0.11	35	0.44	0.09

Legend:
A = Number of articulated time-points at a specific pulse level
P = Number of implied pulsations (or time-points) at a specific pulse level
MPU = Middleground pulse units (quarters at T_1 and T_3; eighths at T_2)
EPU = Elementary pulse units (sixteenth or quintuplet sixteenth)

As a higher degree of activity increases the probability that articulated time-points will be grouped at the next higher pulse level, the density indexes can be used to compare the

potential for mensural determinacy of the two competing pulse streams within sections. Here, the density of the quintuplet sixteenth is higher than that of the sixteenth in all the sections except for the Opening, although the two levels are nearly equal in the Epilogue (0.44 as compared with 0.43). Conversely, the point of maximum divergence in density indexes occurs in the *Scherzando*. Overall, the average for the quintuplet sixteenth is higher than that of the sixteenth density, 1.13 (A/P_{MPU} = 678/599) as opposed to 0.95 (A/P_{MPU} = 569/599). The number of time-points per quarter beat is pre-determined, i.e., there are always four implied sixteenths and five implied quintuplet sixteenths per notated quarter beat. In order to assess whether the higher density corresponds to the built-in potential or deviates from it, I have provided a second set of density indexes (A/P_{EPU}). For each foreground pulse level, the foreground event density A/P_{EPU} is directly related to the A/P_{MPU}, i.e., $A/P_{EPU} = (A/P_{MPU}) \div Pn$, where Pn corresponds to the number of elementary pulse units in one middleground pulse unit. For example, let us consider the pulse density results for the quintuplet sixteenth level in mm. 1-14. Given that the contextual middleground pulse unit corresponds to the notated quarter and that for each quarter P = 5 quintuplet sixteenths, and given a pulse density A/P_{MPU} of 1.17, the pulse density (A/P_{EPU}) = 1.17/5 = 0.23.

While pulse density (A/P_{MPU}) may be more representative of the perceived density of a given elementary pulse stream, the pulse density (A/P_{EPU}) is more representative of its "realized" potential, and thus of the composer's treatment of this built-in potential. The fact that the average pulse density for the two pulse streams in the Opening is very close, i.e., 0.23 for the quintuplet sixteenth as compared to 0.25 for the sixteenth, confirms that the higher density A/P_{MPU} corresponds to the built-in potential of the quintuplet sixteenth for higher density within the quarter beat and that overall, the composer has not privileged one

stream over the other beyond its abstract potential (although such decisions might have been made at the pre-compositional stage). In conclusion, the overall assessment of metrical potential based on density seems to favor the quintuplet sixteenth level slightly over the sixteenth level. The density contour segments also show that the pattern of variance is very similar from one stream to another, and across structural levels. Thus, given the relatively low density of the notated quarter beat, especially when compared to the Mozart example, the nestedness analysis will be most instructive in its assessment of the tendencies for metrical dominance (or ambiguity), and by extension, of polymetric potential.

3.3.2.2 Polarity and Integration at Work: Nestedness Analyses

While density analyses function as descriptors of the "background" or "wall paper" of perception, a kind of "evening-out" of the prospective and retrospective accounts, nestedness analyses make it possible to engage the question of meter in both ways. We may speculate that the prospective approach emphasizes the order of events and short-term comparisons of equivalent events; it is geared toward the continuous. The retrospective approach emphasizes memory and salience; it functions non-linearly, relating discontinuous events. The various segmentation levels may also be understood to correspond to different perspectives, from a moment-to-moment approach (sub-phrase level) to a more general outlook (section level), with the phrase level acting as a mean between these two extremes. Thus, at the phrase level the units are small enough to preserve a sense of the "dialogue" aspect generated by changes in polarity as well as by the context in which these changes operate, an aspect that can be lost if the texture is observed too closely. Another way of understanding the differences between these levels might be to make a parallel with listening experience. The smallest level may

correspond to a first or second hearing, the middle level to the changing perception of the piece as one gets to know it better, and the larger level to the general impression one is left with after extended exposure to a work. Thus, while the smaller segments keep track of changing perception, a phenomenon that we could relate to Max Noubel's evocative image of the "*changement dramatique par le biais de la pulsation*," the larger segments show a tendency toward overall assessment, and the latter is possibly the more likely interpretation after multiple listenings.[48]

The scope of this dissertation prevents me from diving into a detailed account of any of these perspectives. At this time, I will limit my discussion to a few observations. First, I will consider the sectional nestedness analysis and emphasize those elements it provides in addition to the density analysis. Then, I will identify some general trends that are found in the two lower levels. My aim at the moment is to introduce this tool and suggest some of the ways in which it can be used. It is hoped that the reader will thus be tempted to make her or his own reading of the data in tandem with the score and/or the recorded performances.

[48] This translates as "the dramatic change brought about by means of the pulsation"; Noubel, *Elliott Carter*, p. 68.

Table 3.10: Nestedness analysis data for *Changes*, section level

Bar-Measures	Total Events* (T_S)	Pulse Density (A/P_{MPU})	Sixteenths (O Stream)			Quintuplet Sixteenths (X Stream)			Polarity (I/T_S)		Integration (N/T_S)
			A	N	I	A	N	I	"O"	"X"	MPU
1-14	92	24/53 = 45%	54	24	30 = 56%	62	24	38 = 61%	33%	41%	26%
14-27	105	28/54 = 52%	70	28	42 = 60%	63	28	35 = 56%	40%	33%	27%
27-53	142	41/103 = 40%	88	41	47 = 53%	95	41	54 = 57%	33%	38%	29%
53-89	222	58/135 = 43%	126	58	68 = 54%	154	58	96 = 62%	31%	43%	26%
89-113	281	66/105 = 63%	145	66	79 = 54%	202	66	136 = 67%	28%	48%	23%
113-130	95	25/69 = 36%	53	25	28 = 53%	67	25	42 = 63%	30%	44%	26%
131-150	55	14/80 = 18%	34	14	20 = 59%	35	14	21 = 60%	36%	38%	25%

* This is the total of *synchronous* events, i.e., the attacks that belong to the two structural pulse streams. It does not include events that are non-synchronous or *syncopated*, i.e., attacks that correspond to displaced subdivisions.

Legend:

I = Number of independent events
N = Number of nested (or shared) events
T_s = Total number of synchronous events in a given timespan
A = Number of articulated time-points at a specific pulse level
P = Total number of implied pulsations at a specific pulse level
MPU = Middleground pulse unit (quarters at T_1 and T_3; eighths at T_2)
"O" = O stream (based on sixteenths elementary pulse stream)
"X" = X stream (based on quintuplet sixteenths elementary pulse stream)

While the surface density takes into account all events, including non-synchronous subdivisions, such as triplet sixteenths and other subdivisions of the quarter beat that are not synchronous with the foreground level, the nestedness analysis will take only into account those events that are synchronous with the three pulse streams under examination, i.e., the quarter, sixteenth, and quintuplet sixteenth pulse streams. Let us be reminded that isochronicity refers to durations that are equivalent and synchronicity has to do with the particular way in which events can be coordinated at the most basic level. Synchronous events are defined as events whose duration can be determined according to a common pulse stream and they usually involve simple ratios such as 2:1, 3:1, etc. This also means that synchronous events are necessarily in-phase with each other. Thus, nestedness analyses exclude events that are *syncopated*, i.e., articulated time-points that correspond to displaced subdivisions of a particular pulse stream. The overall integration factor at the section level ranges from 23% to 29% (see **table 3.10**). This is a relatively low nestedness when compared with the Mozart example, where the lowest integration value was 43% (quarter note integration within half pulse for the melody only, mm. 1-2). This kind of texture provides a favorable environment for the polarity of the two elementary pulse streams. Thus, the "O" stream polarity ranges from 28% to 40%, while the "X" stream polarity ranges from 33% to 48%. When comparing the polarity values from section to section, it can be observed that the "O" stream is clearly dominant only for Episode 1 (bar-measures 14-27).

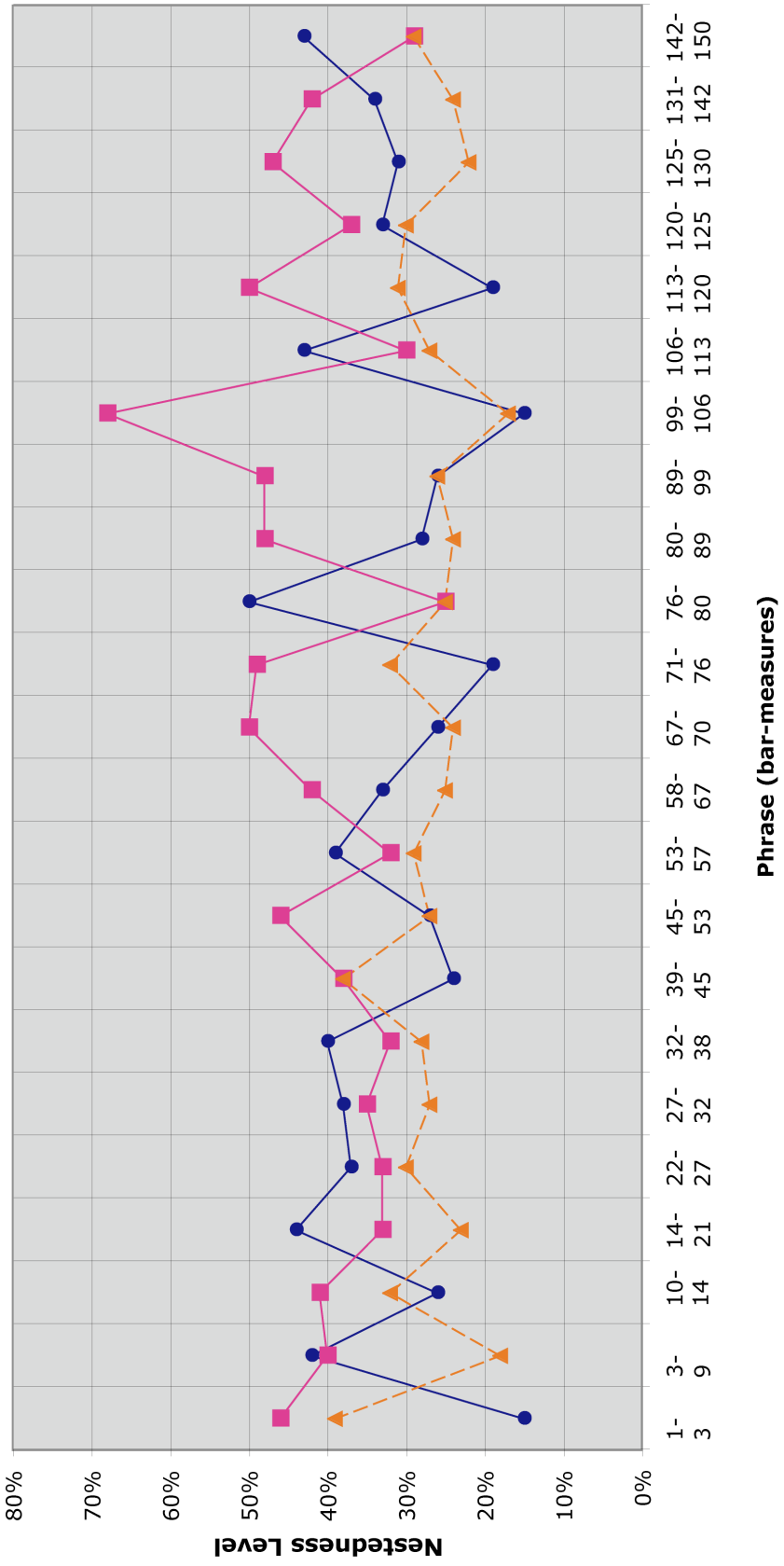


Figure 3.2: Polarity and integration variance, phrase level

Table 3.11: Nestedness analysis data, *Changes*, phrase level*

Section	Bar-Measures	Total Events (T_S)	Pulse Density (A/P_{MPU})	Sixteenths (O Stream)			Quintuplet Sixteenths (X Stream)			Polarity (I/T_S)		Nestedness (N/T_S)
				A	N	I	A	N	I	"O"	"X"	MPU
Opening	1-3	13	5/10 = 50%	7	5	2 = 29%	11	5	6 = 55%	15%	46%	39%
	3-9	45	8/26 = 31%	27	8	19 = 70%	26	8	18 = 69%	42%	40%	18%
	10-14	34	11/17 = 65%	20	11	9 = 45%	25	11	14 = 56%	26%	41%	32%
Episode 1	14-21	48	11/31 = 35%	32	11	21 = 66%	27	11	16 = 59%	44%	33%	23%
	22-27	57	17/23 = 74%	38	17	21 = 55%	36	17	19 = 53%	37%	33%	30%
Episode 2	27-32	26	7/21 = 33%	17	7	10 = 59%	16	7	9 = 56%	38%	35%	27%
	32-38	47	13/24 = 54%	32	13	19 = 59%	28	13	15 = 54%	40%	32%	28%
	39-45	21	8/28 = 29%	13	8	5 = 38%	16	8	8 = 50%	24%	38%	38%
	45-53	48	13/30 = 43%	26	13	13 = 50%	35	13	22 = 63%	27%	46%	27%
Episode 3	53-57	31	9/14 = 64%	21	9	12 = 57%	19	9	10 = 53%	39%	32%	29%
	58-67	52	13/37 = 35%	30	13	17 = 57%	35	13	22 = 63%	33%	42%	25%
	67-70	34	8/15 = 53%	17	8	9 = 53%	25	8	17 = 68%	26%	50%	24%
	71-76	31	10/24 = 42%	16	10	6 = 38%	25	10	15 = 60%	19%	49%	32%
	76-80	16	4/12 = 33%	12	4	8 = 67%	8	4	4 = 50%	50%	25%	25%
	80-89	58	14/37 = 38%	30	14	16 = 53%	42	14	28 = 67%	28%	48%	24%
Scherzando	89-99	103	27/41 = 66%	54	27	27 = 50%	76	27	49 = 64%	26%	48%	26%
	99-106	88	15/27 = 56%	28	15	13 = 46%	75	15	60 = 80%	15%	68%	17%
	106-113	90	24/37 = 65%	63	24	39 = 62%	51	24	27 = 53%	43%	30%	27%
Episode 4	113-120	16	5/26 = 19%	8	5	3 = 38%	13	5	8 = 62%	19%	50%	31%
	120-125	30	9/23 = 39%	19	9	10 = 53%	20	9	11 = 55%	33%	37%	30%
	125-130	49	11/20 = 55%	26	11	15 = 58%	34	11	23 = 68%	31%	47%	22%
Epilogue	131-142	41	10/45 = 22%	24	10	14 = 58%	27	10	17 = 63%	34%	42%	24%
	142-150	14	4/35 = 11%	10	4	6 = 60%	8	4	4 = 50%	43%	29%	29%

* See **table 3.10** for the corresponding legend.

An exploration of the polarity variance at the phrase level reveals a much more variegated landscape, one that is more expressive of the actual conflict between character-patterns. In **figure 3.2**, each stream is represented by its nestedness degree in a given phrase; the nestedness degree corresponds to the polarity results for the sixteenth ("O" stream) and quintuplet sixteenth ("X" stream), and to the integration value within the quarter beat for the quarter stream (compare with **table 3.11**). It can be observed that the "X" stream still dominates in the majority of phrases (13 out of 23), but the "O" stream is offering more resistance (9 out of 23). Also, in one instance the sixteenth polarity and quarter integration are equal and predominant over the quintuplet sixteenth. This representation also allows us to observe more enduring tendencies; for example, the "O" stream has the highest polarity from bar-measures 14-38, two phrases beyond Episode 1, which was noted previously as the only section that was clearly dominated by the sixteenth pulse stream. Also, while the section level data for the Epilogue showed the two elementary pulse levels as nearly equivalent in terms of polarity (i.e., sixteenth = 36% and quintuplet sixteenth = 38%), the representation at the phrase level (**table 3.11**) actually shows that it is first dominated by the quintuplet sixteenth pulse (bar-measures 131-142) and then by the sixteenth pulse (bar-measures 142-150; refer back to **ex. 2.9** for details of the musical surface). This particular interaction of the two elementary pulse streams could be significant, especially in the light of the opening phrase, where the relationship was reversed, as well as in the context of the "conflict-resolution" metaphor previously explored and the transformative role of the "outburst."

Similar representations of the sub-phrase level would also illuminate the interaction of pulses in a more moment-to-moment manner. Such data could be a useful tool for performance preparation, especially given the importance of the concept of "argument" as a

shaping force for musical discourse in Carter's music (as recognized by Schiff, Garrison, Noubel, and others). This analysis has also shown that there might be grounds to question Charles Rosen's assessment that "these contrasting temporal perspectives are not simply opposed but are combined in ways that rarely allow one of them to become dominant."[49] It would seem that while this statement might hold true when considering the nestedness factor over larger spans of music, it is not so in the smaller segments. Finally, I have not offered a background nestedness analysis. Although such an analysis can be a predictive tool of background pulse dominance, its most promising use may be in the identification of different types of background structures when comparing different pieces. The concept of *structural predisposition*, i.e., the dominance of one pulse stream over another based on the nature of their abstract relationship has already been touched on in comparing A/P_{MPU} and A/P_{EPU} density and would be an interesting topic to pursue, especially in comparing different realizations of the same models. Another important (and still unclear) issue is the relation and potential combined influence of quarter density and integration. As noted previously, a high quarter density would seem conducive to a structural quarter pulse (i.e., a pulse with potential for mensural determinacy), but how high would it need to be, especially if the integration value is low? And, what other factors would need to be present to realize this potential? More analyses are required before such questions can be answered, and possibly perception studies aimed at evaluating the influence of specific factors.

[49] Quoted in Hudson, *Stolen Time*, p. 423.

3.4 Polymetric Potential and Realization in *90+* for Piano (1994)

From the perspective of the listener, polymeter is most readily perceptible when other musical elements such as register, articulations, dynamics and intervals are used to reinforce the independence of superimposed rhythmic layers.[50] In this section, we will explore the influence of specific musical parameters used by Carter on the manipulation of *polymetric potential*, i.e., the interaction of structural presence and relative salience of concurrent metrical frameworks in a given passage. The full *realization* of the polymetric potential, that is, the likelihood that a given texture may give rise to simultaneous metrical interpretations or projections that are non-synchronous, will depend on a variety of factors, and more specifically on the weight attributed to particular musical parameters in the stratification of durational events.[51] Such weighting may be construed as occurring at three different, but interactive, stages in the unfolding of a musical object: (1) in the primary or *compositional* stage, musical elements are correlated by the composer and instructions are left to the performer in the form of a musical score; (2) in the secondary or *interpretative* stage, the instructions are decoded and subjected to expressive variation by the performer; and (3) in the tertiary or *subjective* stage, the performance is experienced by the listener whose

[50] Theoretically, even a single series of events may generate two separate streams if some of its elements are perceived as operating in different perceptual dimensions such as timbre or melody.

[51] Bregman, *Auditory Scene Analysis*, p. 665. While Bregman is mainly concerned with primitive processes, he points to the role of "schema-based segregation" in the hierarchical interpretation of musical structures through auditory streaming.

perception is influenced by mental schemas developed through exposure to various types of context-specific interpretative data.[52]

The following discussion of three examples from Elliott Carter's *90+* for piano will serve to illustrate how the various analytical tools presented up to now can be used to determine polymetric potential from a compositional point of view. Furthermore, these analyses will demonstrate that in Carter's use of structural pulse streams, the interaction of polyrhythms at the foreground and middleground levels plays a significant role in the realization of polymetric potential. Finally, while this section will focus on elements that belong to the first stage through score analysis, it will also suggest some ways in which performance practice and listening habits may come to influence the realization of polymetric potential.

3.4.1 Integrated Polymeter

Example 3.18 provides a representative sample of the *scherzando* section C^2 (mm. 49–68). In this passage, each 90+ pulsation gives rise to a metrical grouping or pulse-measure that is non-synchronous with the notated meter and is defined by eighth notes despite the triplet-eighth dominated texture. On the annotated score, surface events have been traced back to the corresponding elementary pulse streams: the upper and lower time-lines represent the triplet-eighth and the eighth pulse streams, respectively, and background pulsations have been circled and numbered. In addition, the resulting durational patterns have been transcribed above and below the grand staff with their resultant speeds and possible

[52] The three stages proposed here are related to, but not identical with, Nattiez's *tripartition sémiologique*; see Jean-Jacques Nattiez, *Musicologie générale et sémiologie* (Christian Bourgeois Éditeur, 1987), pp. 32-8.

metrical interpretations. The rhythmic pattern associated with the lower time-line (EPU = eighth) suggests a mensurally determinate grouping of eight eighth beats that is strongly supported by the recurrence of the 90+ pulse, with a possible intermediate grouping of 3+2+3.[53] At the same time, the upper time-line (EPU = triplet eighth) is interpreted as giving rise to two possible metrical projections that would compete with the 90+ pulse-measures. The partition of rhythmic events as belonging to this non-synchronous strand is supported by harmony and pitch recurrence. For example, the pitch sequence {C6, F#3, G6, A#2, D5, C#1}, which is found between 90+ pulsations 41 and 42, corresponds to hexachord 6-Z17 – a key sonority in this piece – as does the pitch sequence {C6, F#3, F4, Bb1, B3, D7, C6}, similarly found between pulsations 43 and 44 (not shown). Also, there are three instances of C6 aligning with the downbeat of the proposed measures, all of which are accented.

[53] In accordance with our exposition in chapter 2, the expression "mensurally determinate" refers to a metrical projection that is supported by the rhythmic organization of the musical surface. The use of the concept of *determinacy* in this context is also borrowed from Hasty.

Example 3.18: Integrated polymetric structure, *90+*, pulse-measures 41-43 (mm. 54-56)

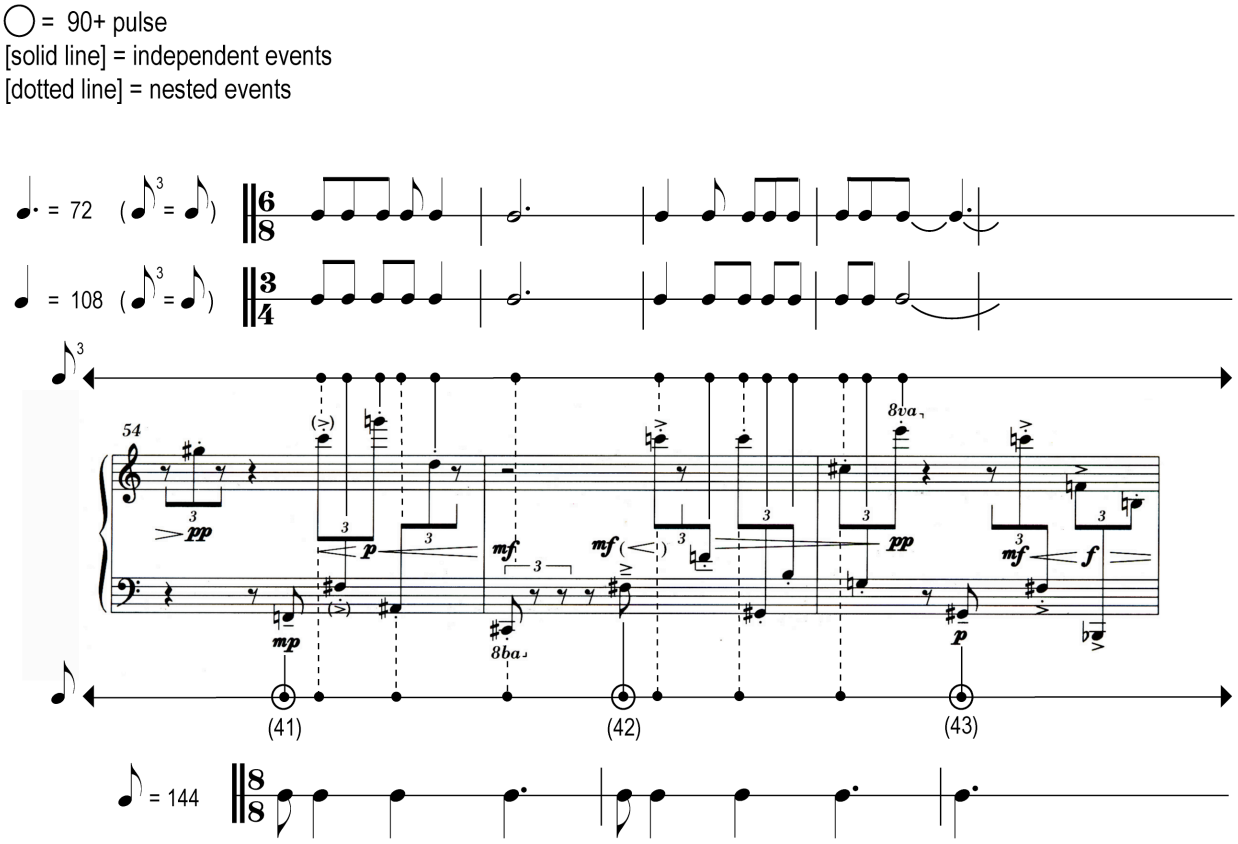


Table 3.12 presents a summary of the density and nestedness data for this passage. As we have seen in section 3.3, density and nestedness analyses are two tools that can be used to tackle the polymetric potential of a given passage. In the short excerpt under examination (**ex. 3.18**), the eighth pulse level shows that the three 90+ pulsations are independent and that six of the triplet-eighth pulsations are nested in the competing elementary pulse stream. On the other hand, the triplet-eighth pulse stream exhibits the highest number of independent pulsations as well as articulated time-points, which results in the highest polarity. However, given that the six nested pulsations corresponding to the notated quarter beat yield a 35% integration rate, the implicit conflict between structurally determining eighths and triplet-eighths is minimized.

Table 3.12: Density and nestedness analysis data, pulse-measures 41-43 (mm. 54.2.2–56.2.2)[54]

Pulse Level	Pulsation Total (P)	Articulated Time-Points (A)	Pulse Density (A/P)	Independent Events (I)	Nested Events (N)	Polarity (I/T_S)	Integration (N/T_S)
♪	17	9	0.53	3	6	18%	35%
♪³	24	14	0.58	8	6	47%	
♩	8	6	0.75	0	6	N/A	

Legend: T_S = Total number of synchronous events (= 17)
A = Number of articulated time-points at a specific pulse level
P = Number of implied pulsations (or time-points) at a specific pulse level
I = Number of independent events
N = Number of nested events

This interpretation is supported by the density analysis (**table 3.12**), which reveals that while the triplet-eighth elementary pulse stream exhibits a higher number of articulated time-points and the highest polarity, its time-point density index is actually closely comparable to that of the eighth elementary pulse stream and both pulse streams are relatively less active than the quarter pulse. This interaction suggests that while the structural predominance of the triplet-eighth has not been hampered by an excessively low degree of activity, the specific placement of articulated time-points favours the quarter pulse. Thus, the excerpt is much more likely to be perceived as a single stream of quarter beats with alternating eighth and triplet-eighth subdivisions, that is, an *integrated* polymetric structure. This property is reinforced compositionally by the use of a series of single pitches arranged dynamically in short linear phrases with irregular accentuation. Moreover, the presence of some events belonging to the

[54] The format used for bar-measure numbers includes beats and subdivisions for greater clarity. For example, "m. 54.2.2" corresponds to bar-measure 54, second beat, and second eighth note.

triplet-eighth pulse stream in the low register associated with the 90+ pulse (e.g., C#1 and G#2 in m. 55) results in stream segregation ambiguity at the expense of 90+ pulse salience.[55] As a result, the actual perceived metrical grouping will be largely dependent on the accentuation and timing patterns produced by the performer, and the eighth pulse articulations may even come to be perceived as expressive variations of the triplet-eighths, that is, as slightly elongated. This might favor a listener's interpretation of the meter as a 12-cycle with the downbeat corresponding to every third notated quarter beat, especially given the common practice of making the last part of a rhythmic group longer.[56]

3.4.2 Polarized Polymeter

The opening section of *90+* is representative of a *polarized* polymetric structure (**example 3.19**). This section presents an uninterrupted strand of 14 chords, which, with the exception of the last two occurrences, are expressed literally as chords of three or four pitches (boxed).[57] As we have seen in chapter 1 (subsection 1.2.1.1), the Chord pulse is operating against the first 11 pulsations of the 90+ pulse, or pulse-measures 1-11 (circled), and both pulses are initially accompanied by seven instances of a Rogue pulse (marked with asterisks). The polyrhythmic cycle at work here is considerably more complex than those shown in **table**

[55] This explains Schiff's assertion that in this section "the staccato pulses… are camouflaged by the sparse texture"; see Schiff, *The Music*, p. 220.

[56] This has been shown by empirical studies on timing variations applied by musicians; for one representative example, see Bruno H. Repp, "Pattern Typicality and Dimensional Interactions in Pianists' Imitation of Expressive Timing and Dynamics," in *Music Perception* 18/2 (2000): 173-211.

[57] In bar-measure 13, the Chord pulsation is a single accented legato note and in bar-measure 14 the single note event coinciding with the Chord pulse is held and becomes a four-note simultaneity on the second quintuplet-sixteenth.

3.1. Considering only the two most enduring structural pulse streams, the 90+ and Chord pulses, the corresponding abstract middleground structure represent a hypothetical polyrhythmic cycle with a timespan of 68 bar-measures between each coincidence point. The spacing of the two initiating pulsations by a sixteenth creates an even more complex out-of-phase polyrhythm that is ultimately left incomplete. Furthermore, the presence of the third structural pulse stream and the resulting 5:4:3 ratio between elementary pulse streams necessitates an actual subpulse stream with a pulsation total of 240spu, or 60spu per quarter beat.[58] While the resultant subpulse density of 240spu per bar-measure is difficult to imagine as being perceptible as such, the relation of the background pulse stream to the bar-measure's 48 subpulse units and the pattern of motion of pulsations toward or away from each other, that is, their convergence or divergence pattern, seems more tangible, especially as manifested by near-coincidence points (marked with arrows).

[58] The number of subpulse units (spu) per bar-measure is equal to the product of the number of elementary pulse units (epu) per quarter beat and the number of beats per bar-measures, i.e., 3 triplet eighths × 4 sixteenths × 5 quintuplet sixteenths = 60spu per quarter beat × 4 quarters per bar-measure = 240spu per bar-measure. The number of subpulse units (spu) equivalent to each elementary pulse unit (epu) will be equal to the product of the number of elementary pulse units (epu) per quarter beat for the two competing pulse streams, e.g., the value of a triplet eighth = 4epu × 5epu = 20spu.

Example 3.19: Structural polyrhythm, *90+*, pulse-measures 1-11 (mm. 1-14)

◯ = 90+ pulse

▯ = Chord pulse

* = Rogue pulse

↙ = near-coincidence

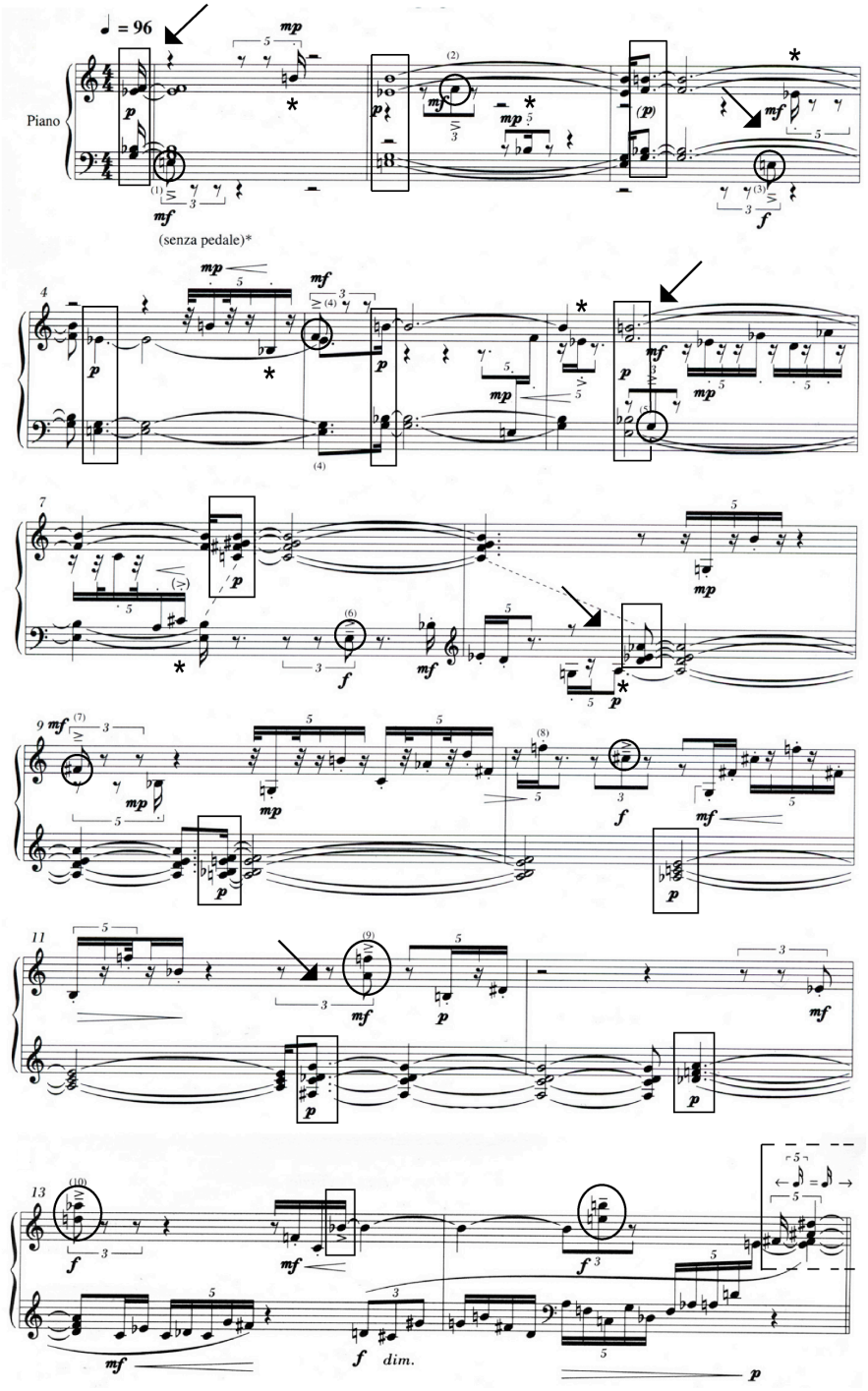
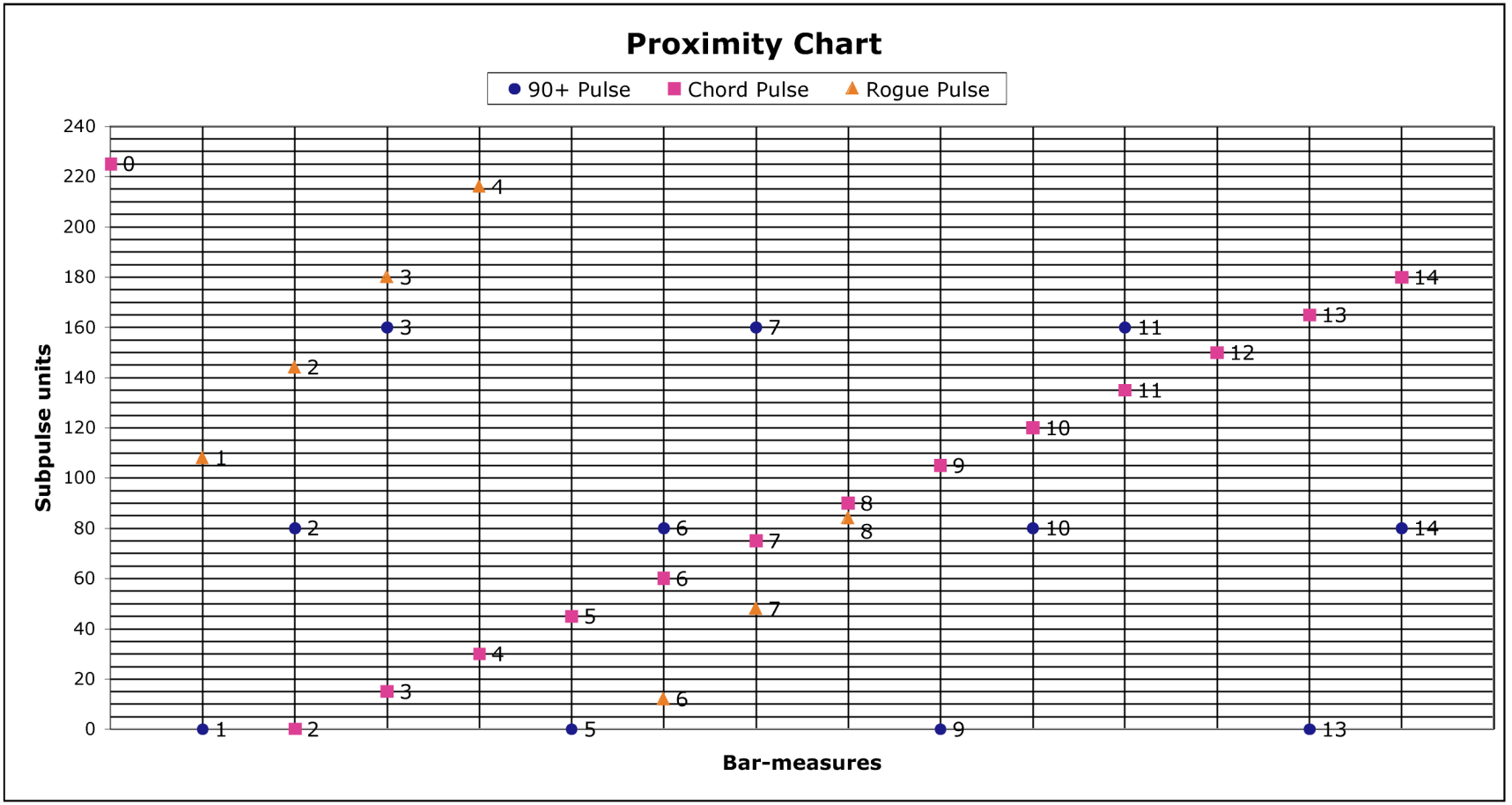


Figure 3.3: Proximity chart, *90+*, all structural pulse streams, bar-measures 0-14 (vertical axis = one bar-measure; horizontal axis = time-line of opening section)



This phenomenon can be readily observed in a *proximity chart* (**figure 3.3**). In this representation, the vertical axis corresponds to the timespan of one bar-measure calculated in subpulse units and the horizontal axis symbolizes the unfolding of the piece in time, with each background pulsation labeled with the bar-measure number where it is located. It can be observed that throughout this passage, the 90+ and Chord pulses participate in a larger divergence phase of +5spu, with near-coincidences gradually more spaced out, from 15spu (mm. 0-1) to 20spu (m. 6) and 25spu (m. 11). In addition, both these pulse streams experience one near-coincidence with the Rogue pulse, first at m. 3 (20spu) and then at m. 8 (6spu), the second of which corresponds to the dissolution of the last Rogue pulsation into the eighth Chord pulsation.

Table 3.13: Density and nestedness analysis data, pulse-measures 1-6 (mm. 0-8)

Pulse Level	Pulsation Total (P)	Articulated Time-Points (A)	Pulse Density (A/P)	Independent Events (I)	Nested Events (N)	Polarity (I/T_S)	Integration (N/T_S)
𝅘𝅥𝅯	129	14	0.11	7	7	20%	20%
𝅘𝅥𝅮 3	96	11	0.11	4	7	12%	
𝅘𝅥𝅯 5	161	19*	0.12	16	7	47%	
𝅗𝅥	32	7	0.22	0	7	N/A	

*Excludes two syncopated events.

Legend:
T_S = Total number of synchronous events (= 34)
A = Number of articulated time-points at a specific pulse level
P = Number of implied pulsations (or time-points) at a specific pulse level
I = Number of independent events
N = Number of nested events

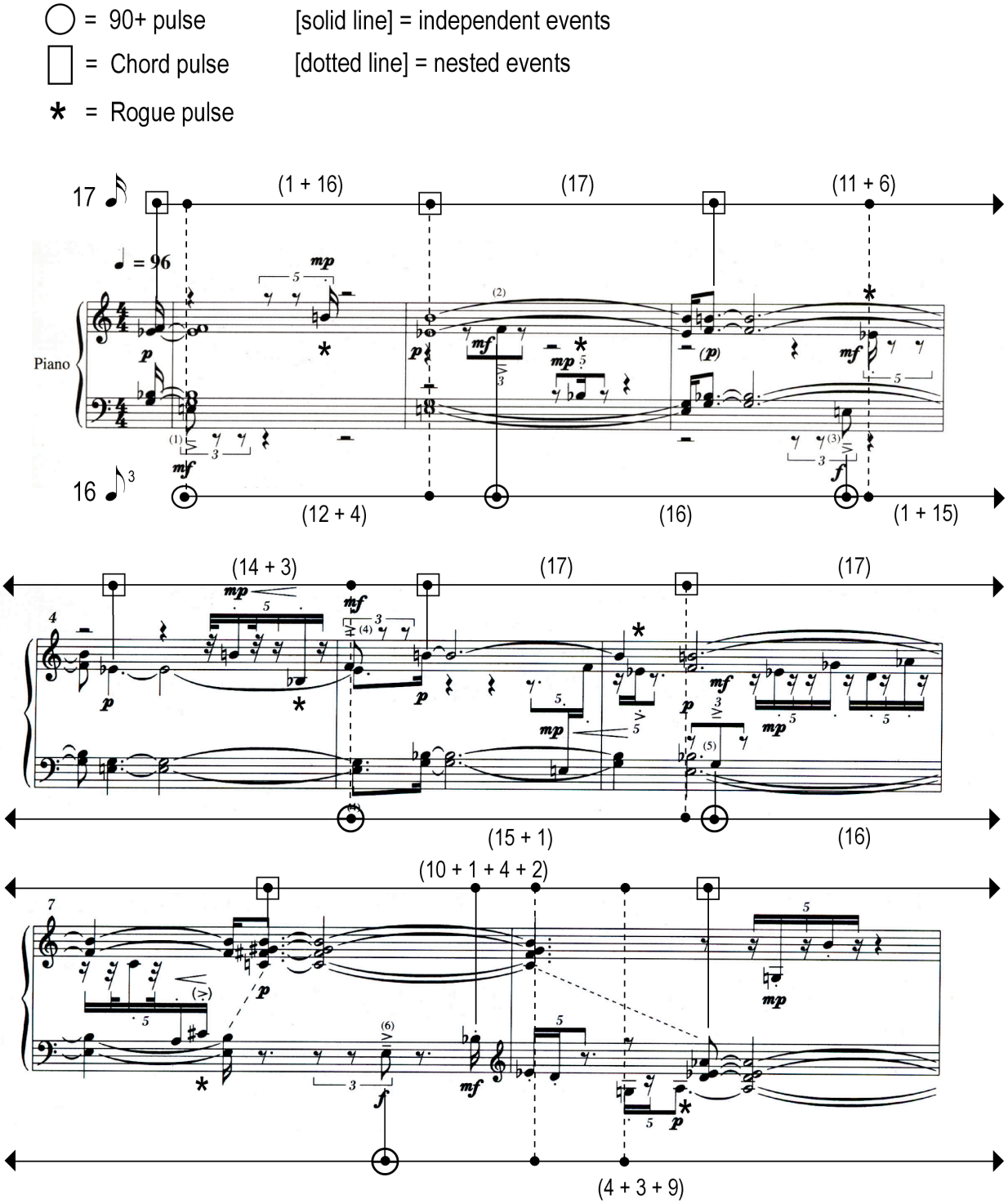
The density indexes for this passage exhibit sparse but carefully distributed pulse stream activity (see **table 3.13**). While the Rogue pulse, which disappears after only seven pulse articulations, fails to achieve structural significance, it contains the highest number of articulated time-points as well as independent events. As a result, its influence on the polymetric potential of this texture requires careful consideration. On one hand, the duration of its EPU suggests that it may be perceived as expressive variation of the elementary pulse stream associated with the Chord pulse. On the other, its rhythmic and textural make-up, i.e., short staccato single notes, indicates a close tie with the 90+ pulse with which it also engages in registral exchange. Rather than settling this conflict in favour of one of the two most stable structural pulse streams, the nestedness analysis confirms the prevalence of the Rogue pulse on account of its polarity, at least at the foreground level. This interpretation is reinforced by the clear segregation of the texture into legato chords and short single notes. Thus, of the two layers, the second, i.e., the short single notes associated with the 90+ and Rogue pulses (and their elementary pulse streams), would appear to be the most influential since it yields a combined polarity of 67%.

Given the relatively low density of the quarter beat and the stream segregation ambiguities, a *temporal-processive analysis*, that is, an analysis that takes into account the effect of temporally ordered events on an idealized listener, will prove to be helpful in evaluating the tendency for metrical dominance (or ambiguity), and by extension, polymetric potential. **Example 3.20** presents the details of the nestedness analysis for the first eight bars of the piece with the time-lines of the 90+ and Chord pulses as well as the rhythmic groupings they give rise to. The overall low density and irregularity of surface rhythm suggest that upon hearing this passage, the listener's metrical expectations may be prompted

to shift from one slow structural pulse stream to the other, a behaviour that is consistent with the experience of a polarized polymeter. Initially, the number of attacks, slightly faster speed, prescribed legato performance of the chords and the conventional association of harmonic rhythm with meter lend strong support to the perception of the chord pulse as dominant, a phenomenon which may prevent the emergence of a viable polymeter. Furthermore, the occurrence of the first 90+ pulsation in m. 1 defines the sixteenth as durationally determinate, and as such, it participates in the articulation of the Chord pulse.[59] Thus, as the quintuplet-sixteenth B is sounded, it is likely to be perceived as an expressive variation of the sixteenth pulse stream, that is, as being a little "late" (although the difference might not be noticeable for most listeners). While the next chord on the downbeat of m. 2 is nested in the 90+ pulse stream in a 3:1 subdivision, it also confirms the relevance of the 17 sixteenths timespan associated with the chord pulse, a measure that finds relatively little opposition from the non-synchronized F and Bb, especially given the emergence of a third chord pulsation in m. 3.

[59] The expression "durationally determinate" emphasizes the fact that a given duration has been realized or completed in the musical surface; it is borrowed from Hasty's work as represented by *Meter as Rhythm*.

Example 3.20: Polarized polymetric structure with nestedness details, *90+*, mm. 1-8



In the recorded performance of the opening by Charles Rosen, the predominance of the Chord pulse at the beginning of the piece comes across very clearly. This may be partly the result of the slower tempo and generally thicker texture exhibited by the recording.[60] Nonetheless, the accentuation, dynamic level, and generally lower register associated with the 90+ pulse are factors, which, in addition to the previously noted textural association of the Rogue and 90+ pulses, may prevent the latter from being completely superseded by the Chord pulse, as would be the case in an integrated polymetric structure. Moreover, in such a sparse texture, the particular durational patterns articulated by the events associated with the Rogue pulse may also become decisive in directing the focus of the listener. Thus, the surface activity in mm. 4-6 appears to increasingly favor the 90+ pulse, while the durational pattern of the following stream of quintuplet-sixteenth seems to lead to the Chord pulsation of m. 7. In listening to Ed Spanjaard's recording of the same passage, the changing relationship of these two middleground pulses is noticeably more graspable.[61]

---

[60] Elliott Carter, *The Complete Music for Piano: 90+; Piano Sonata; Night Fantasies; Elliott Carter in Conversation with Charles Rosen*, Charles Rosen, piano (Digital disc. Bridge Records, 9090, 1997). Whether the general quality of sound as 'thicker' is the result of the quality of the recording or of a heavier use of the pedal is not clear.

[61] Elliott Carter, *Homages & Dedications*, Nieuw Ensemble, cond. Ed Spanjaard (Digital disc. Montaigne/Naïve, MO 782089, 2003). Spanjaard's performance is further characterized by a generally faster tempo (a total duration of 4 minutes and 58 seconds as opposed to the 6 minutes and 12 seconds of Rosen's performance), less contrast in dynamic levels, and more precise beat subdivisions.

Figure 3.4: Contour graph of the structural polyrhythm as performed by Charles Rosen, *90+*, mm. 0-14

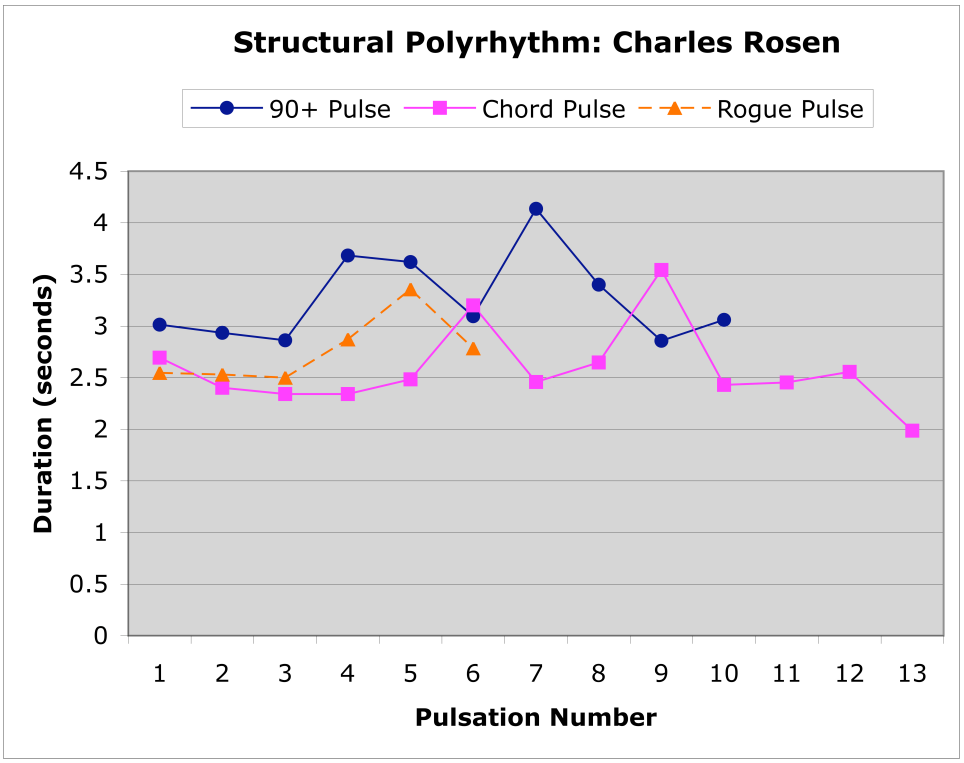
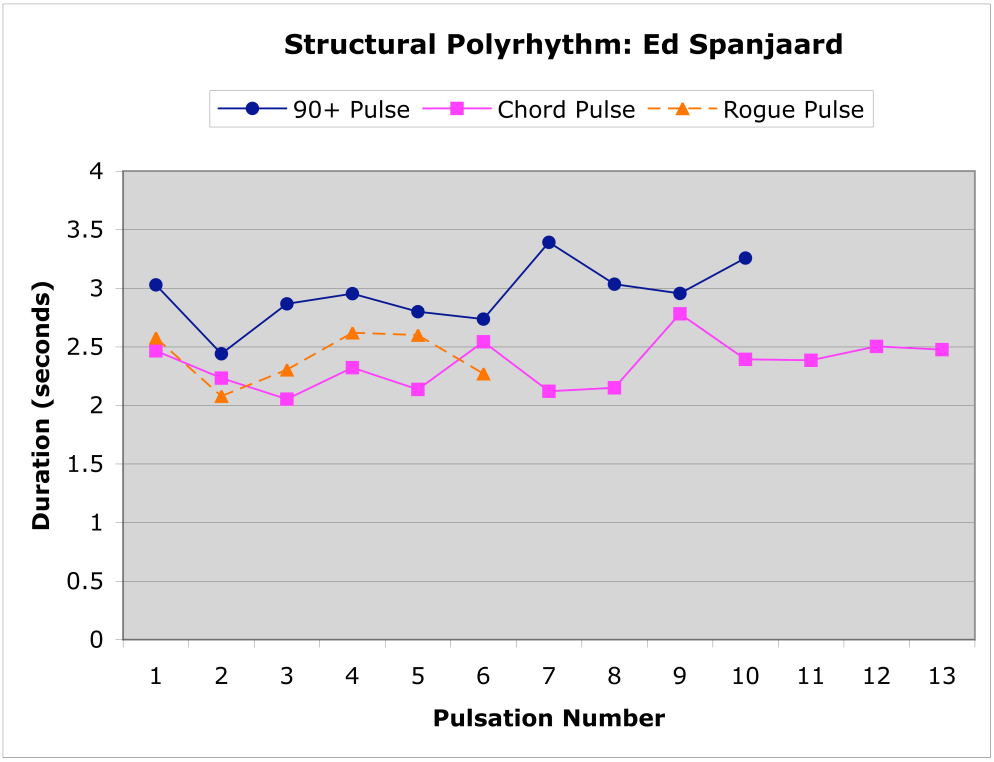


Figure 3.5: Contour graph of the structural polyrhythm as performed by Ed Spanjaard, *90+*, mm. 0-14



An examination of the contour graphs of the structural polyrhythm as performed by Rosen and Spanjaard sheds additional light on these perceived differences (see f**igures 3.4** and **3.5**). In a *contour graph*, series of isochronous pulsations are plotted according to their actual duration (vertical axis) and order of appearance in time (horizontal axis).[62] Contour graphs are useful for appraising general trends in the regularity and interaction of pulse streams; however, the inspection of specific interactions in time will necessitate their coordination with the musical score. As shown by **figure 3.4**, Rosen's performance of the structural polyrhythm in the opening of *90+* is characterized by more regularity in the timing of the Chord pulse overall. Thus, while the three pulses are very regular at the beginning, the fourth and fifth 90+ pulsations are much longer (mm. 5-7; see **ex. 3.20**, lower time-line). While, this slowing down of the 90+ pulse correlates with the increase of activity within the Rogue pulse (i.e., the increased number of quintuplet sixteenths in the musical surface), the fact that the periodicity of the Chord pulse is not affected until its sixth pulsation in bar-measure 6 (**ex. 3.20**, upper time-line) suggests that the latter might have served as the referential background pulse unit. In other words, it suggests that at least in this passage (i.e., mm. 5-6 in **ex. 3.20**) Rosen might have conceptualized the Chord pulse as the ground for the other pulses; this interpretation is supported by the regularity exhibited by Chord pulsations nos. 7-8 and 10-12 (mm. 7-12; refer back to **ex. 3.19** for the complete score).[63] Although Spanjaard's performance (**fig. 3.5**) shows similar fluctuations (e.g., 90+ and Chord pulsations nos. 6 and 9), it is more regular overall. In particular, 90+ pulsations nos. 3-6 (mm. 3-7; see

---

[62] I am indebted to Paul Riker for his technical assistance in extracting the corresponding time-points from the recordings.

[63] An analysis of covariance would provide more definite support for this interpretation. This type of analysis is included in the literature survey provided in chapter 4.

ex. 3.20, lower time-line) appear to preserve their independence and this might account for the greater perceptibility of the influence of the quintuplet sixteenths in directing the listener's attention to background pulsations from different pulse streams.[64]

Let us now return to our analysis of the score. Rather than being fully determining, the compositional choices made by Carter can be interpreted as reinforcing a tendency that is already at work through the interaction of polyrhythmic elements at different structural levels. While the textural distribution enhances the dominance of the Chord pulse, the participation of the 90+ pulse in a middleground polyrhythmic compound ensures its structural significance and also hinders its integration within the dominant pulse stream. Another factor that reinforces the mensural determinacy of the 90+ pulse is the initial alternation of the two associated pitch classes E and F, which yields a larger duple measure. An examination of the cyclic spectra associated with the middleground polyrhythm will clarify this point (see **figures 3.6** and **3.7**). In these representations, each pulse stream in the middleground structural polyrhythm can be understood as a *cycle*, which can then be represented in the form of a *cyclic spectrum,* a visual representation of the motion of pulsations in time using their notation as a basis for comparison. This representation emphasizes the cyclical behaviour of the two structural pulse streams in relation to the bar-measure and offers additional insights on the relation between polyrhythm and polymeter. As with the proximity chart, the vertical axis represents the duration of the bar-measure in subpulse units; the numbers around the circle represent the succession of bar-measures and should be read clockwise. Thus, the problem intrinsic to two-dimensional representations,

[64] The duple grouping exhibited by Chord pulsations nos. 3-6 is also suggestive of the application of our knowledge of expressive timing patterns to the analysis of performed rhythmic organization.

which emphasize time-point location within bar-measures at the expense of the unfolding of pulse-measures, is resolved by the use of a circular time-line.

Figure 3.6: Cyclic spectrum for the 90+ pulse

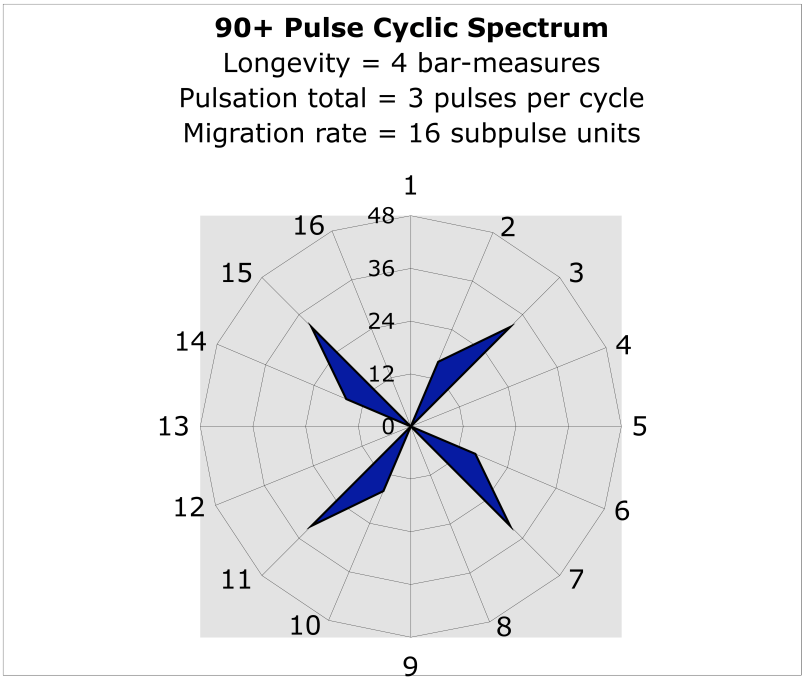
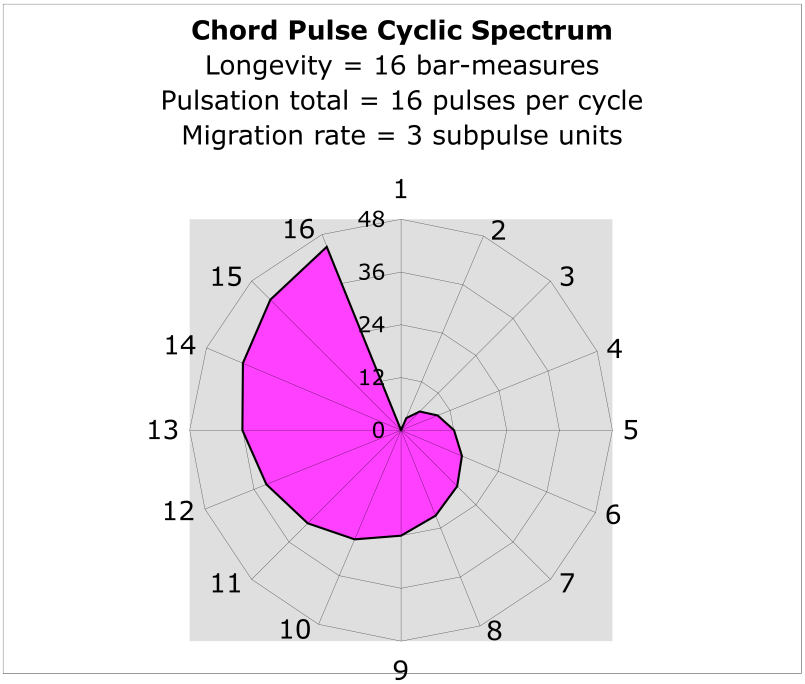


Figure 3.7: Cyclic spectrum for the Chord pulse



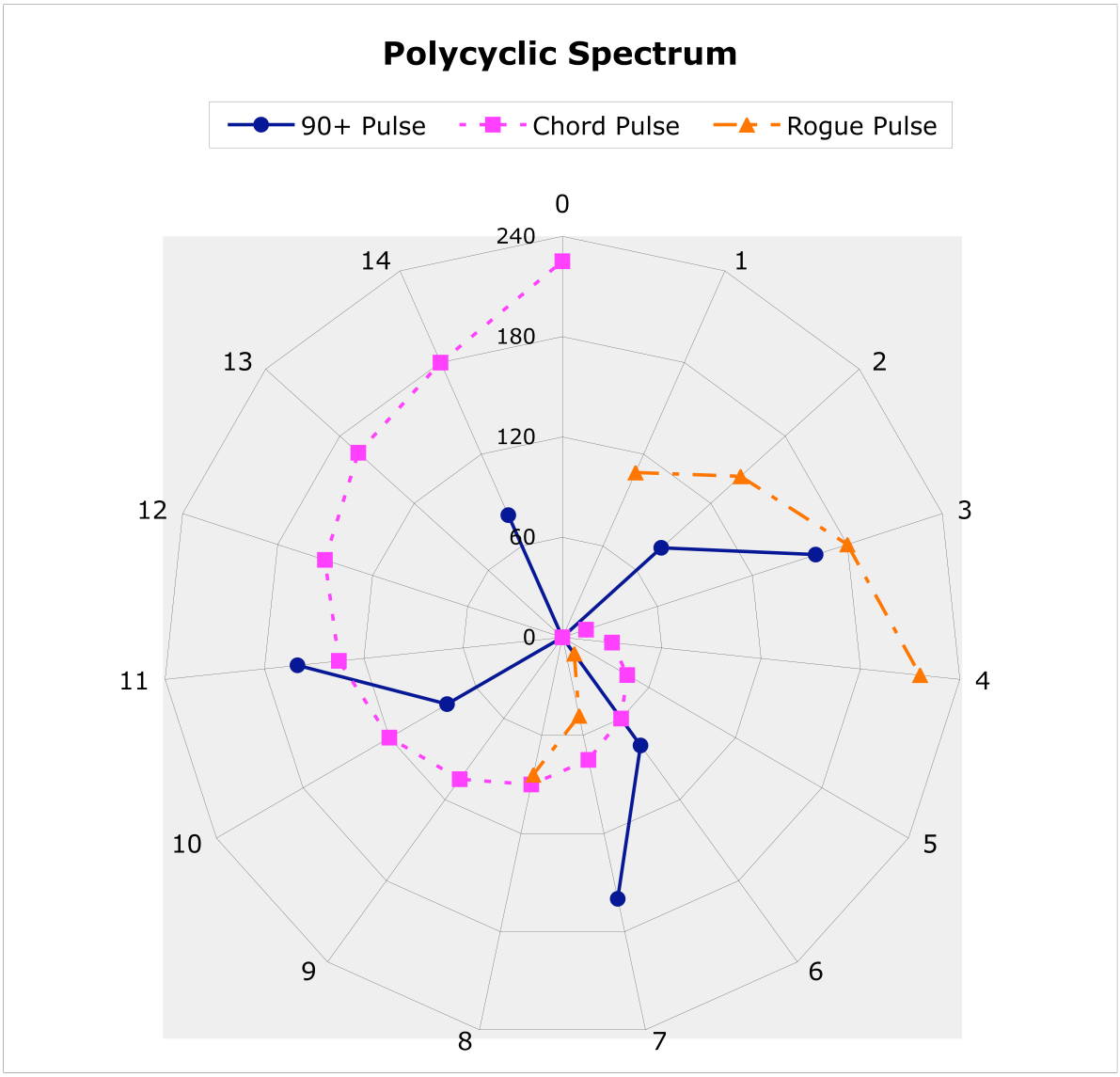
For comparison, both cyclic spectra are represented over a timespan of 16 bar-measures. The cyclic duration, or its *longevity*, corresponds to the number of bar-measures required for the cycle to return to its initial position within the bar-measure, while its *migration rate* is expressive of the progressive shift of the pulse in relationship with the subpulse stream associated with the bar-measure.[65] As can be observed from the cyclic spectra, while the duration of 16 bar-measures corresponds to a single instance of the Chord pulse cycle (**fig. 3.7**), it allows for four instances of the 90+ pulse cycle (**fig. 3.6**). Thus, the two pulse streams occupy the same notational space but they move quite differently in this space. The sharp angles of the 90+ pulse cyclic spectrum (**fig. 3.6**) are produced by its quick migration across the space of the four bar-measures it necessitates in order to complete its cycle, while the smooth shell-like design of the Chord pulse cyclic spectrum (**fig. 3.7**) is the result of the slow progression of its pulsations across its 16 bar-measures duration of its cycle. This particular interaction can also be expressed as a pulsation total ratio of 16:12, or more simply, 4:3.

Figure 3.8 presents the interaction of all structural pulse streams and provides a more definite illustration of the contribution of the proximity cycle on the discrimination between middleground and foreground levels. The expression "polycyclic spectrum" emphasizes the sense of directed motion displayed by structural polyrhythms. In considering the interaction of the 90+ and Chord pulse streams, if each near-coincidence point (see mm. 0–1, 6 and 11) is understood as a particularly accented event, it seems reasonable to posit that the two middleground pulsations can be perceived as one event, leading to the conceptualization of this passage as a slightly out-of-phase 4:3 polyrhythm. Thus, the SPU value attached to each

[65] The concept of the migration rate is also expressive of the visual experience of a performer or analyst upon reading the score.

convergence point represents a timespan imbued with tension that holds potential for expectation building and onto which the listener's attention can be directed. In this context, the Rogue pulse has only a very limited influence, which is more characteristic of the foreground level.

Figure 3.8: Polycyclic spectrum, *90+*, mm. 0-14

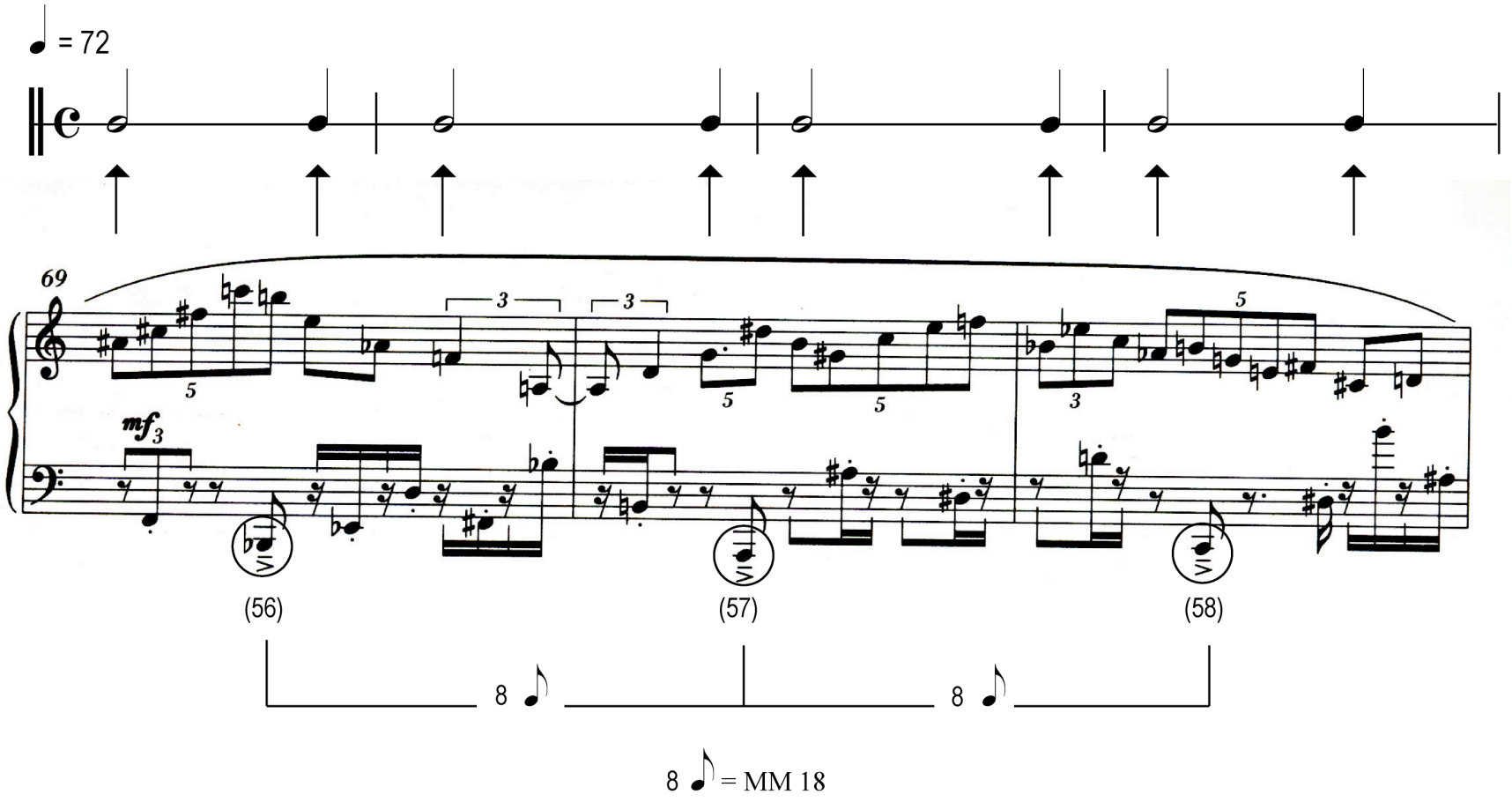


The two examples discussed so far share a sparse texture. In the first case, the resulting polymeter is one where levels have collapsed on one another and metaphorically speaking, its structure offers reduced metrical "depth." Conversely, in the opening section, there is a sense of space and the background pulsations can interact more perceptibly with the middleground and foreground. As will be seen, the third and final example presents us with a generally denser texture that gives rise to a different type of polymetric experience.

3.4.3 Balanced Polymeter

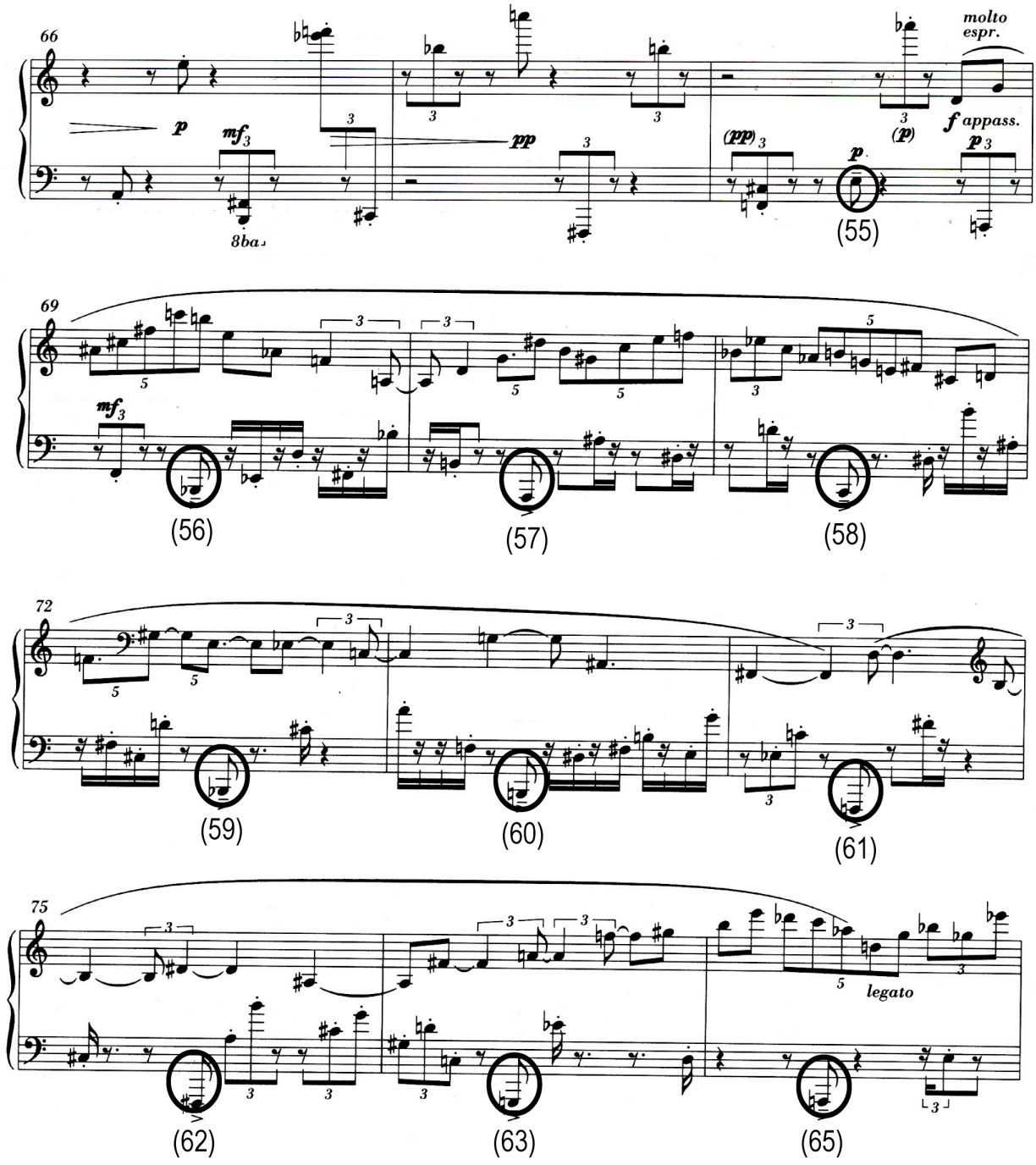
Example 3.21 offers a short excerpt from a passage (mm. 68-82) that exhibits a *balanced* polymetric structure. Here, the texture is organized into two main layers, each of which corresponding roughly to the right and left hand of the pianist, although the two parts cross in register a few times (see **example 3.22**). The right hand consists of a mostly stepwise but wide-ranging melody whose irregular subdivisions (ranging in values from triplet-crotchets to septuplet-sixteenths) produce the effect of a written-out rubato. The left hand provides an accompaniment with regularly accented tones in the lower register alternating with mostly irregular short staccato notes in the middle-low register higher. As shown by the reduction above the treble staff, the melody's surface irregular rhythms could be interpreted as strengthening the perception of the notated beat at the tempo of quarter = 72; the resulting durational pattern strongly suggests a triple metre. Concurrently, the left-hand accompaniment presents a series of non-synchronous measures that are structurally defined by the enduring 90+ pulse. **Examples 3.23** and **3.24** present reductions and metrical interpretations of the two rhythmic strands; each reduction is based on the most prevalent pulse level.

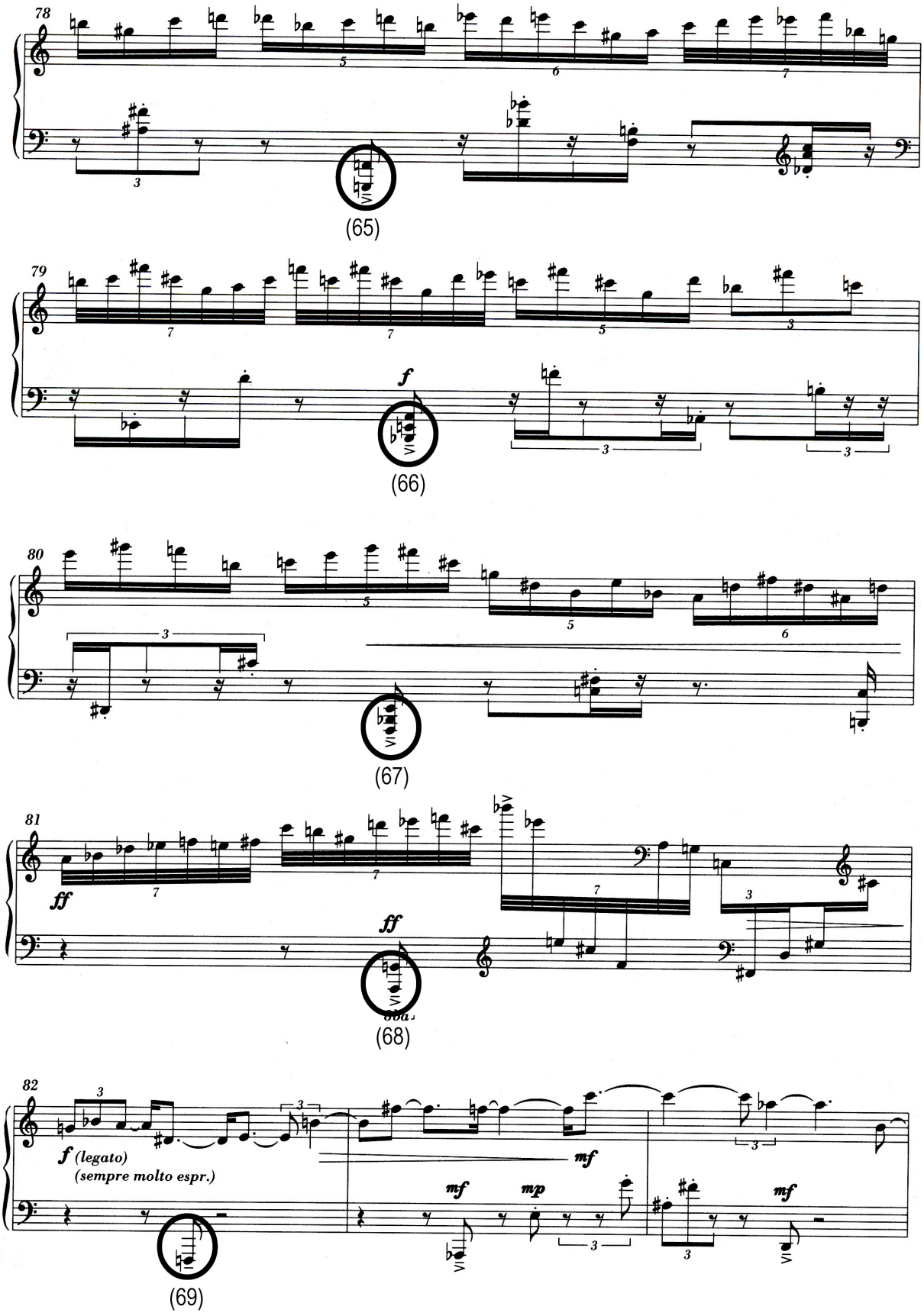
Example 3.21: Balanced polymetric structure, *90+*, pulse-measures 56-58 (mm. 69-71)



Example 3.22: Background pulsations, *90+*, pulse-measures 55-69 (mm. 68-82)

◯ = 90+ Pulse (MM 18)



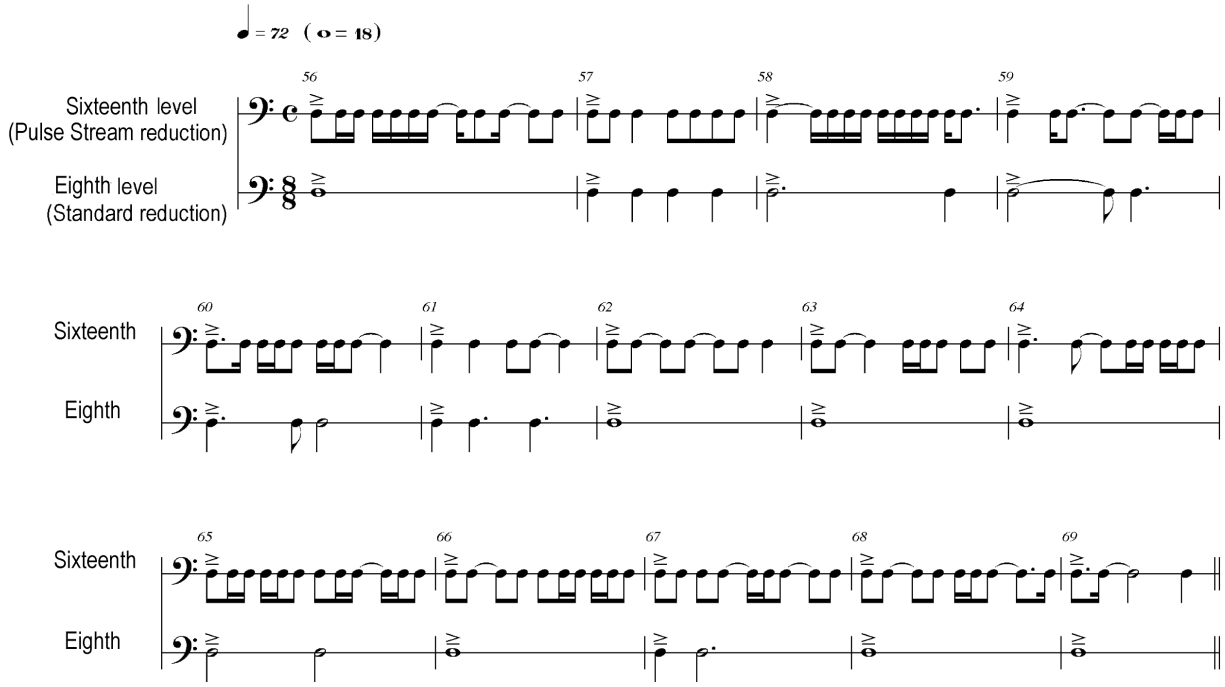

78
(65)
79
f
(66)
80
(67)
81
ff
ff
(68)
82
f (legato)
(sempre molto espr.)
mf
mf
mp
mf
(69)

As we have seen, reductive tools can be used to tackle the metrical potential of concurrent (or intertwined) layers whose events are rooted in competing pulse streams by focusing on successive isochronous timespans that are related by simple proportions. The quarter-level reduction of the left-hand accompaniment (**ex. 3.23**, bottom) shows that while the 90+ pulse contributes to articulate a series of downbeats, most of the resulting pulse-measures exhibit an internal make-up with a low degree of mensural determinacy and just as eighths are the dominant pulse subdivision, these are often syncopated. Consequently, although some pulse-measures suggest a quadruple meter, e.g., pulse-measures 57 and 65, most of these are metrically under-defined, especially since many reduce out to the value of an undivided whole note. Similarly, even though the beginning of the quarter level reduction of the right-hand melody (**ex. 3.24**, bottom) attests to relevance of a triple measure, the subsequent syncopated passage (mm. 72-74) and lack of definition in durational patterning (especially in mm. 78-81) contribute to create a certain amount of metrical ambiguity.

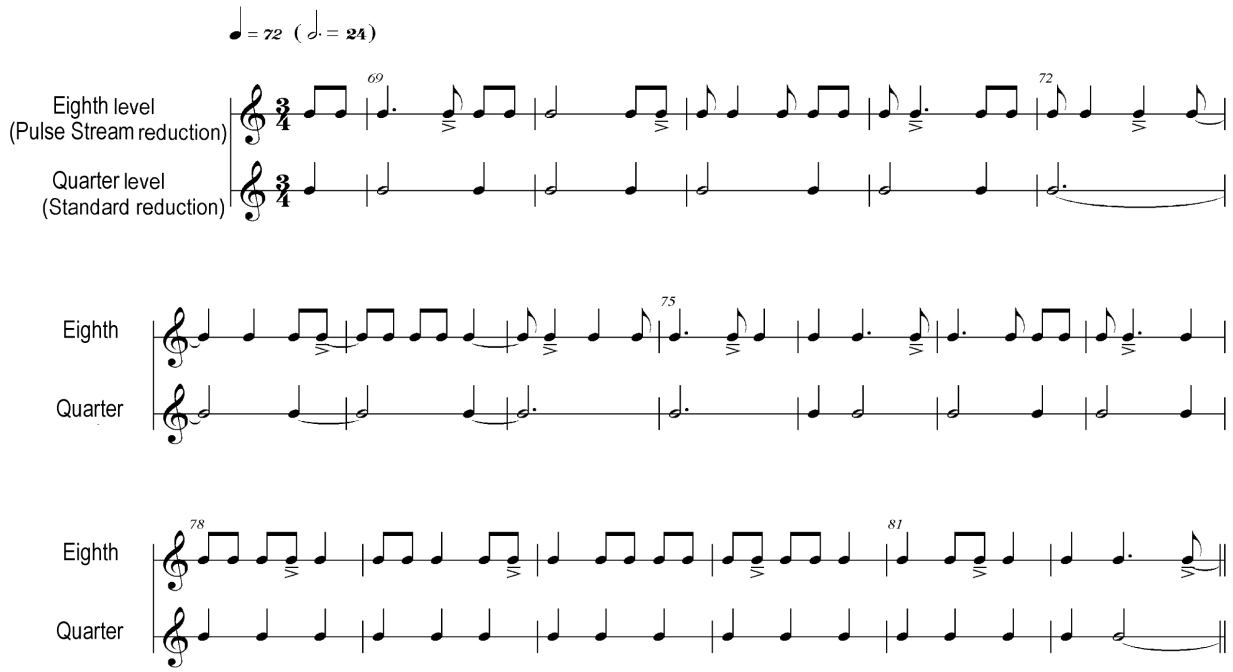
The melody and accompaniment structure of this texture would seem to suggest an interpretation of the notated bar as dominant and of the pulse-measure as syncopated. In spite of this, the pulse stream reductions (**ex. 3.23** and **ex. 3.24**, top) show that in this passage, nestedness contributes to enhance the mensural determinacy of both layers rather than leading to their integration (as in the *scherzando*) or to the dominance of one over the other (as in the opening section). As we have seen before, in a pulse stream reduction the particular make-up of measures can point to passages that either show evidence of mensural determinacy or lack thereof, *as supported by the degree of synchronicity between layers*. Thus, we have seen that several measures in the standard reductions showed mensural indeterminacy as a result of either a lack of subdivisions or the use of syncopation (ties).

Here, both pulse stream reductions (left-hand sixteenth level and right-hand eighth level) represent viable metrical layers with a higher level that is non-synchronous. In other words, while both examples express a speed of quarter = 72, the higher pulse levels represented by the whole note (**ex. 3.23)** and the dotted-half note (**ex. 3.24**) result in a 4:3 speed ratio (24:18), the same ratio that characterized the interaction of middleground pulse streams in the opening (**ex. 3.19**), and an unexpected structural connection that calls for further investigation.

Example 3.23: Eighth- and sixteenth-level reductions of left-hand accompaniment, *90+*, pulse-measures 56-69 (mm. 69-83)



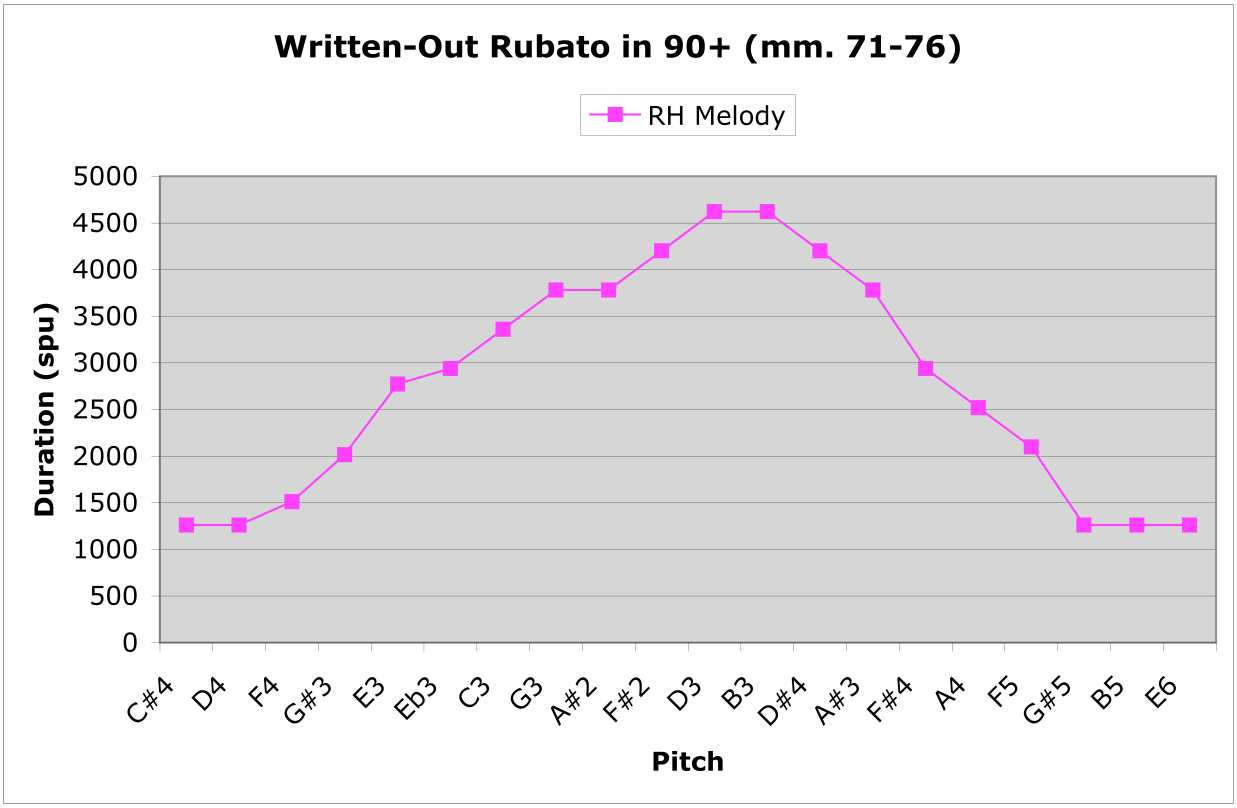
Example 3.24: Quarter- and eighth-level reductions of right-hand melody, *90+*, mm. 68-82



The interpretation of this texture as a balanced polymetric structure is further supported by the clear stream segregation it exhibits, except maybe for the few passages where the parts cross in register. However, depending on a particular listener's habits, it could be experienced as a polarized polymetric structure dominated by the right-hand melody, especially if the performer emphasizes the notated beat to the detriment of the strongly accented left-hand pulsations prescribed by the score. In the absence of clearly demarcated beats, a right-hand oriented listening mode might also induce mensural indeterminacy, especially given the extensive use of written-out rubato (and ties) in some passages. **Figure 3.9** provides a contour graph of a representative passage in the right-hand melody (mm. 71-76; see **ex. 3.22**); here, the durational values require the use of subpulse units (spu). In the absence of an underlying referential beat unit, the passage will be likely experienced as a gradual deceleration followed by acceleration. On the other hand, the results

of a beat-finding experiment I conducted with 32 musically trained participants in January 2007 show that synchronization was more successful with the left-hand quarter beat; these results suggest a structural bias in favor of left-hand dominance, at least for a mechanical performance. This experiment will be the focus of chapter 4.

Figure 3.9: Contour graph of the right-hand melody, *90+*, mm. 71-76



CHAPTER 4

The Realization of Polymetric Potential from a Psychological Perspective

The work presented in this chapter is situated at the boundary between music theory and music psychology. In many ways, it follows investigation practices of music psychology and the experiment it presents was partially modeled on recent experiments in beat perception and synchronization to music (Drake et al., 2000; Snyder & Krumhansl, 2001; Toiviainen & Snyder, 2003). However, it contrasts with such studies in two important ways: (1) it assumes a broader conception of meter than that represented by the current working definition in music psychology, and (2) its experimental design includes controlled and non-controlled variables (semi-factorial design).

With regard to the first aspect, it is my belief that with respect to music, meter needs to be defined more broadly than the way it is currently defined by music psychologists, if only to allow for the investigation of polymeter from a psychological perspective. *Definitions are intimately linked with purpose and method.* In contemporary music-psychological theory and experimental practice, meter is understood as a dynamic process by which one attunes to recurring periodicities (e.g., Jones, 1976; also see Large, 1994).[1] Through various experiments, it has been shown that when asked to reproduce or synchronize with rhythmic patterns, participants manage rhythmic information using a

[1] Edward W. Large's model of meter as a dynamical process is derived from principles associated with oscillators and is especially interesting because it allows the interpretation of the sum of a musical surface based on multiple (and simultaneous) resonance periods that relate to each other by ratios other than 2:1 and 3:1. For related research, see Eric D. Scheirer, "Tempo and Beat Analysis of Acoustic Musical Signals," *Journal of the Acoustical Society of America* 103/1 (1998): 588-601.

figure-ground paradigm in which foreground elements are hierarchically integrated into an underlying regularity; inability to infer such a figure-ground mental schema is also believed to have detrimental effects on participants' performance. This view of meter has serious implications for the study of polymeter. If meter necessarily involves entrainment and entrainment operates on figure-ground schemas, then polymeter, which is usually associated with the presence of concurrent metrical frameworks, is psychologically impossible.[2] Since twentieth-century composers have used this technique extensively, and given that they often do it with the purpose of creating a distinctive effect, it would be beneficial to come up with a definition and framework that would not only allow for empirical studies of polymetric experience, but also relate polymetric structures and their perception to our understanding of meter in a fruitful way, that is, without describing the resulting beat structure as merely "ambiguous" or "a-metric." While at times it may be the intention of the composer to create ambiguity or hamper the ability of the listener to latch on the kind of rhythmic regularity provided by a simple metric construction, the range of intentions and effects sought is much richer. As we have seen, the perception of contrasting temporal modes is an important expressive aspect of Carter's musical language and is not meant to be only conceptual.

With regard to the second aspect, given the lack of precedents for the study of polymetric (as opposed to polyrhythmic) textures, it appeared necessary to adopt an exploratory stance in designing the experiment. In the attempt to identify important aspects

---

[2] This is basically the argument presented by London in *Hearing in Time*, pp. 49-50. London also asserts that the use of the term "meter" in reference to phenomena outside the boundaries recognized in empirical studies are "metaphorical." On example of this would be the concept of "hypermeter," which can result in beats with durations longer than 2 seconds. In my opinion, this assertion is equivalent to an epistemological hold-up because it negates other uses of the term from historical, cultural, and disciplinary perspectives.

of polymetric experience, the net was cast over a broader set of means, measures, and parameters than is practical for statistical analysis. The main motivation in gathering such varied types of data was also to bypass limited definitions of meter that may be embedded in previously employed experimental designs. While some data can be analyzed using standard statistical methods, other types, such as participants' individual tapping responses, verbal commentary, and spontaneous body movements, are more appropriately considered using a descriptive approach. This chapter will focus mainly on the statistical analysis of tapping responses. Other types of data will be used to guide future experiments.

It is not the goal of this chapter to prove that polymeter "exists." From a music-theoretical perspective, there is no doubt that it does. On the other hand, to define polymeter from a psychological perspective is well beyond a single chapter in a dissertation and would necessitate an extensive program of experimental studies. Thus, *the larger goal of the study presented here was to explore how polymeter is experienced as well as to experiment with ways of measuring polymetric realization from the perspective of the listener*. The first section of this chapter presents a brief survey of previous studies relevant to the topic at hand; this will serve to introduce some of the issues relating to the study of polymeter from a psychological perspective. Then, the experimental methodology will be described followed by a presentation and discussion of the results. I will conclude with a brief summary of the findings and suggestions for future research. The implications of these results for music analysis and Carter research will be discussed in the conclusion of this dissertation.

4.1 Methodological Issues

In music psychology, rhythm and meter are studied by using both experimental and modeling (or computational) approaches. These investigations are not limited to perception and cognition but also include aspects of categorization, memory, performance, expressive timing, and motor processes. In the experimental domain, pulse or beat finding paradigms have been especially favored. Tapping and other forms of synchronization to music that result from attunement or *entrainment* to the recurrence of regularly spaced markers in the music (i.e., periodicities) is a widely observed spontaneous phenomenon and does not require musical training. Thus, in these experiments, participants are usually asked to communicate their understanding of the underlying pulse or beat by tapping with one or both hands or fingers on a computer key, a drumming pad, or keyboard. While beat tapping provides a useful way into the listener's perception of meter as manifested by the beat, it is obviously not meant to represent the totality of metric experience. Music psychologists, as music theorists, recognize that in music, there are often more than two pulse levels and that meter might be described as a *nested hierarchy* of pulse levels relating to each other by simple ratios such as 2:1 or 3:1.[3] As with any discipline, the object of study often finds itself defined by the methods used to investigate it. Even if *focal attention* tasks are used, that is, tasks by which participants are asked to tap at the next higher and/or lower pulse level(s) in the hierarchy (e.g., Drake et al., 2000; this study expands a framework proposed by Jones et al.

[3] In "Musical Meter in Attention to Multipart Rhythm," Keller and Burnham use 4:1 as an example of bar to beat simple ratio; see *Music Perception* 22/4 (2005): 630. Many music theorists would probably put 4:1 in a different category based on duple subdivision of beats and accentual structure.

1995), "metric experience" as such is only witnessed by comparison and deduction.[4] Thus, in psychological studies, metrical frameworks are understood as being generated by the entrainment process and meter has become tied to the *tactus*, i.e., a comfortable pulse level whose rate is closely related to bodily rhythms such as heart rate or walking speed.[5] The complexity of rhythmic patterns is directly related to goodness of fit, patterns that can accommodate more than one framework are deemed ambiguous, and those with no perceptible periodicity are regarded as non-metrical. It is my belief that a more fruitful approach may be to understand pulse (or beat) induction as one particular manifestation of meter.[6]

In the case of polymetric music, beat tapping is problematic. If participants are required to provide their response by tapping with one hand, the two competing pulse levels will often be fused into a composite rhythmic pattern, and then interpreted as "a-metric" by the analyst(s). If participants use both hands, the resulting pattern of competing beats might be too complex to perform spontaneously, especially if none of the pulse levels coincide, as is the case in out-of-phase meters (e.g., Beauvillain and Fraisse, 1984). A listener might be aware of two contrasting metric frameworks and even be able to identify the resulting structure but still be unable to perform it. Although the body of research on polyrhythm is

---

[4] "Focal attention" is related to *selective attention*, i.e., in which participants are prompted to attend to a specific strand in a given texture (Jones et al. 1995). It contrasts with *integrative* (or global) attention, which requires participants to listen to the stimulus as a whole.

[5] Clarke defines the *tactus* as the "most salient level of pulsation"; see Erik Clarke, "Rhythm and Timing," p. 486.

[6] The use of the concept of meter "induction" rather than "perception" stems from the debate about meter being "in the music" or "in the mind"; the latter perspective, which is the one favored in music psychology today, is tied to the conceptualization of meter as "an abstraction from stimulus properties"; Clarke, "Rhythm and Timing," p. 482.

considerable, there has been no study to date specifically designed to investigate polymeter.[7] In addition to studies on polyrhythmic perception and production, other relevant research includes work on auditory streaming, dynamic attending, expressive timing, and multipart rhythm. I will now offer a brief survey, with an emphasis on the concepts and methods most closely related to the present study.

4.1.1 Previous Research

The seminal work on polyrhythm perception was a series of studies conducted in the early 1980s by Handel and his associates (Handel and Oshinsky, 1981; Handel and Lawson, 1893; Handel, 1984). The goal of this research was to investigate how complex rhythmic patterns are interpreted through a spontaneous tapping task. Participants listened to two or three conflicting *pulse trains*, i.e., superposed sequences of isochronous pulsations such as 2:3:5, and were asked to "tap along with the perceived rhythm in any way they wished" (Handel and Lawson, 1983, p. 104). The authors reported that participants' tapping responses exhibited different strategies including *cross-rhythms* and *meter rhythms*. "Cross-rhythms" included events from at least two pulse trains and were therefore irregularly spaced, while the elements in meter rhythms were equally spaced; "meter rhythms" were isochronous sequences based either on one of the pulse train or the coincidence point at the beginning of each cycle. The authors found that rhythmic interpretation varied based on the relative amplitude (loudness), frequency (pitch), and duration of the rhythmic events. Results also

[7] In the first decade or so of studies with polyrhythms, "polyrhythm" and "polymeter" were sometimes used interchangeably (e.g., Grieshaber and Carlsen, 1996); the term "polymeter" is generally absent from more recent studies and temporal patterns that may fit more than one meter are usually characterized as "ambiguous." A possible exception is Keller and Burnham (2005), although they do not characterize the textures they use as polymetric.

suggested that the choice of a given interpretation was closely related to the rate of presentation as well as the level of complexity of the composite pattern. Cross-rhythms tended to occur at the slowest rate, while participants showed a preference for pulse trains of lower cardinality at faster rates and those of higher cardinality at relatively slower rates. More complex patterns such as 3:5 were less likely to be used as cross-rhythms than simpler ones like 2:5. There were also some more specific correlations: for example, participants tended to tap to the low-pitched pulse train when the 2:3 polyrhythm was presented at the fastest rate.

Some researchers followed the path of Handel and his colleagues in exploring various aspects of what the authors had referred to as "contextual interpretation," that is, focusing on the various parameters that would influence the (metric) interpretation of a given polyrhythm; these studies have made important contributions to research on meter as well as auditory streaming (e.g., Pitt and Monahan, 1987; Moelants and van Noorden, 2005). Other studies have tackled motor programming as relating to polyrhythmic performance (Shaffer, 1981; Jagacinski et al., 1988; Grieshaber and Carlsen, 1996) as well as the interaction of perception and production (Beauvillain and Fraisse, 1984; Klapp et al., 1985; Summers et al., 1993; Peper et al., 2000; Pressing et al., 1996).[8] Yet another strand of research has focused on attentional issues related to competing rhythmic patterns (Jagacinski et al., 1988; Jones et

[8] L. H. Shaffer's study of timing mechanisms in expert pianists' performance of music from contrasting styles predates Handel's work on polyrhythm. It should be noted that it remains the only study to date to have concluded that musicians could in fact achieve independent timing in bimanual coordination tasks.

al., 1995; Keller and Burnham, 2005).[9] Finally, one small but valuable area of music-psychological research has involved the production of non-Western rhythms by expert performers trained in these traditions and the cognitive models that support it (Magill and Pressing, 1997; Pressing, 2002).[10]

In general, these studies have concurred with the findings of Handel and his colleagues, although they have adopted also a more conservative stance on what constitutes an underlying beat structure. Thus, they treat cross-rhythms as composite patterns that integrate the musical surface and are not properly metric (because not based on isochronous pulse trains). Polyrhythmic production is theorized as involving an internal clock pulse as well as motor delay and subdivision variables; covariance analysis can be used to identify figure-ground interaction, with the ground usually exhibiting less variation than the figure (e.g., Pressing et al., 1996). Alternative models have been proposed, but experimental findings generally support various forms of *integrated processing*, which are usually based on a hierarchical view of meter that excludes non-isochronous beats, rather than *parallel processing*, which involves independent clocks (Klapp et al., 1985; Jagacinski et al. 1988; Jones et al., 1995; Keller and Burnham, 2005).[11] Even when participants were only required

---

[9] The latter study is of special interest because it is not limited to isochronous pulse streams but develops an approach whereby several rhythmic lines of varying degrees of alignment or metric compatibility are combined.

[10] I am limiting my survey to those studies that involve a music-psychology component. It should be mentioned, however, that in a recent study of African music, Kofi Agawu (*Representing African Music*) has taken a strong stance against the use of the concept of polymeter in this music based on what he describes as a "neo-colonialist" attitude on the part of the ethnomusicologists who study it.

[11] Although parallel processing is similar to my "balanced" polymetric model, and both my "integrated" and "polarized" polymetric structures could be associated with integrated processing, the conceptual framework behind my polymetric models is different from the one

to listen to polyrhythmic patterns and perform a detection task, integrative attending was found to yield better results than selective attending (Jones et al., 1995). Furthermore, when asked to memorize and reproduce embedded rhythmic patterns (recognition memory paradigm), participants performed better when concurrent rhythmic materials provide a good metrical match (Keller and Burnham, 2005). On the other hand, Magill and Pressing (1997) have proposed a form of cognitive clock that is not based on isochronous pulse streams; through structural equation modeling and covariance analysis, they were able to demonstrate that the model provided a good fit for most of the patterns performed by an expert master drummer from West Africa. Finally, studies on polyrhythm perception and production have found no essential difference in the timing mechanisms used by participants with no prior experience and those with more extensive experience, although experienced musicians, and especially percussion players, were more accurate and flexible in the production of more complex patterns, e.g., 3:4 as compared with 2:3 (Pressing et al., 1996).

It is not clear if and how these findings can be transposed onto the study of polymeter. In the simplest case, e.g., concurrent and in-phase duple and triple meters, the resulting beat structure would be equivalent to a polyrhythmic pattern (2:3) and its reproduction would most probably follow the principles outlined by the studies on the production of polyrhythms. But attention can take many forms and musicians are especially skilled at using various forms of selective attention, for example, tracking changes in an ensemble while performing their own part. There are still relatively few empirical studies that

---

supporting these psychological models. One of the main differences is that the categories I propose are not dependent on modes of production (i.e., motor processing) and/or attention independently of the musical structure, but presuppose an interactive process between composition, performance, and perception.

use actual music as source materials and even fewer that use complex textures such as those resulting from the use of polymeter as a structural device.[12]

4.1.2 Experimental Design[13]

The specific goal of this dissertation is to propose a theoretical framework for the analysis of musical textures built on polymetric structures that takes into account how polymeter is experienced. Given the state of research, the experiment was designed to address specific questions of (poly)meter induction as well as to explore various aspects of polymetric experience. There were three main questions that guided the design of this experiment:

1. What is the *relative degree of mensural determinacy* of each rhythmic layer and how does it change when layers are combined?
2. What are some of the *musical parameters* that influence participants' (poly)metric interpretation?
3. To what extent are participants able to follow *competing metrical frameworks* and how is this ability supported or hampered by various factors, including musical training and experience?

[12] The three studies used as a model for part of this experiment used contrasting musical materials ranging from J. S. Bach to ragtime and Bartók. However refined in construction these pieces are, none of them even comes close to the type of rhythmic complexity encountered in the work used as source for the sound materials in this experiment.

[13] This work was funded by a grant from the Graduate Research Grants Program from The Graduate Center of The City University of New York (2005-2006).

This section describes the general design and procedures employed in this study. Other details on specific analytical procedures will be given in the results section, when necessary.[14]

4.1.2.1 Participants

Volunteers were sought through an email advertisement on the music department list-serv and through a verbal announcement at a local semi-professional choir practice; thus, the sample for this experiment was convenient rather than random. Thirty-two participants (identified as Participants 1 to 32) were recruited based on musical training and experience, including the principal investigator. Because the first experiment (Participant 1) was conducted without headphones and some of the stimuli had to be modified after, this data set was discarded; in addition, the principal investigator's data (Participant 31) was used only for testing in the preliminary stages of data preparation and is not included in any of the analyses reported here. Thus, there were 30 participants whose tapping responses were considered (Male = 17, Female 13). The basic requirement was that participants should be actively engaged in musical activities and have received musical training equivalent to at least one year of college undergraduate studies in music. All participants except one were right-handed and no participant reported hearing problems.

Information on participants' musical training and experience was gathered, and this information was used for the exploratory analyses with *between-subjects factors*, that is,

---

[14] I will define basic statistical terms briefly as they are introduced, either in the main text or in the footnotes; these definitions are also reproduced in the glossary. Readers who are not familiar with statistical methods as used in experimental psychology and wish to have more detailed information can refer to Gene V. Glass and Kenneth D. Hopkins, *Statistical Methods in Education and Psychology*, 3rd ed. (Needham Heights, MA: Allyn & Bacon, 1996).

factors related to differences between individual participants that may have an effect on their performance of the experimental tasks. These analyses related directly to the third goal, which was to identify factors that might influence participants' ability to follow competing metric frameworks; they also provided additional information about differences in participants' performance as relating to mensural determinacy and musical parameters. Prior knowledge of polyrhythm and/or polymeter was not a requirement; however, 19 participants had experience with performing polyrhythm. Participants' musical training varied from no formal degree in music to Ph.D. degree in music (No degree = 6, Bachelor's or equivalent = 5, Master's = 18, Ph.D. = 1). Seven participants were amateur musicians and the remaining participants were all music students and/or professional musicians. Participants also identified themselves as primarily engaged in one of five areas of specialization: Composition (n = 2), Ethnomusicology (n = 1), Musicology (n = 5), Performance (n = 13), and Theory (n = 9). Where applicable, categorical data used for statistical analysis was transformed to promote more balanced groups. The four groups of musical degree were combined into two groups (Degree.2grp), i.e., No degree or Bachelor's (n = 11) and Master's or Ph.D. (n = 19). Area of specialization was also transformed in two ways (Specialization.4grp): (1) Musicology and Ethnomusicology were combined into one group (n = 6) as well as Theory and Composition (n = 11); (2) Amateur was integrated as the fourth group (n = 6). Participants' main instrument was also use as a between-subjects factor with three categories, based on performance manner and register: (1) piano or percussion (n = 9, with only one percussion player); (2) high melodic instrument, including flute, alto and soprano voice, and clarinet (n = 14); and (3) low melodic instrument, such as double bass, cello, and bass or tenor voice. Skill level for the main instrument ranges from beginner or

intermediate (n = 5), to advanced (n = 15) and expert (n = 10). Finally, piano experience was also used as a factor and was transformed (Piano.3grp) by combining "None" with Beginning (n = 10), and Advanced with Expert (n = 9); the other group was Intermediate (n = 11).

Some participants' data could not be transformed and was thus not used for statistical analysis, i.e., familiarity with source materials (only three participants knew about or had heard 90+ before, but none reported having performed or studied the piece extensively) and the composer's style (only two participants identified themselves as very familiar with his music), skill level at the main instrument, and familiarity with polyrhythm/polymeter (most participants were either somewhat familiar or very familiar with the concepts, and four participants reported no familiarity; the concepts were explained to these participants prior to the beginning of the experimental tasks).

4.1.2.2 Source Materials

The audio stimuli used in this study were derived from one passage from Elliott Carter's *90+* for solo piano (1994) in which competing metric structures had been identified by score analysis. The excerpt extends from m. 68, fourth eighth note, to m. 88, fifth sixteenth note, inclusive (the reader can refer back to **example 3.22** for an annotated version of the score up to m. 82). In the original score, the passage is notated with a time signature of 4/4 and a tempo of 72 quarter-note beats per minute (bpm); its mechanical performance lasts about one minute. The texture is organized into two main layers corresponding roughly to the right and left hands of the pianist, although the two parts cross in register a few times. The right hand consists of a mostly stepwise but wide-ranging melody that presents irregular subdivisions of quarter and half notes ranging in value from triplet quarter notes to septuplet

sixteenth notes (as shown previously in **example 3.21**), although it also includes a fair amount of syncopation. The design of this layer suggests at least two nested pulse levels with measures of three quarter notes at 72 bpm resulting in a series of dotted-half notes at 24 bpm (corresponding to chronometric durations of 0.832 and 2.5 seconds, respectively). The left hand consists of an accompaniment with regularly accented lower tones, whose inter-onset intervals correspond to whole notes at 18 bpm (3.333 seconds). These slow pulsations alternate with short staccato notes (notated as triplet eighth and sixteenth notes) that are higher in register and mostly irregular. The resulting compound layer suggests measures of four quarter notes at 72 bpm, although such a measure is articulated only once at the beginning of the passage (m. 70, fourth eighth note to m. 71, third eighth note). Because of the eighth note displacement between the quarter beats of the two main layers (right hand and left hand), there are no coinciding beats or measures. However, the two layers could be partially integrated at the eighth and sixteenth pulse levels (the reader may wish to turn back to the polarized reductions in **examples 3.23** and **3.24**).

4.1.2.3 Stimuli

Each layer was transcribed using the notation software Finale 2006 and converted into MIDI files, which were then manipulated using the sound editing software Cubase SX and samples from the Garritan Personal Orchestra VST.[15] Finally, all files were converted in WAV format and normalized to peak amplitude. There were a total of 18 sound files, all of which included the entire excerpt, although the exact duration varied depending on whether only one or both layers were present: the left hand alone and the combined layers lasted

[15] Paul Riker provided technical assistance for programming and sound engineering.

64.167 seconds; the right hand alone lasted 62.917 seconds). Sound files were performed stereophonically. For segregated layer conditions, each layer was piped through both channels equally; for combined layers conditions, stereo separation was created by panning one track 33% to the right and one track 33% to the left.[16]

The temporal structure of the notated score (duration and speed) was preserved but the excerpt was manipulated in various ways to allow for a comparative analysis of the influence of musical parameters on perceived temporal structure. The stimuli were organized into four blocks of trials. The purpose of this design is described briefly in this subsection, but the purpose of each manipulation will become clearer in the "Results" section (section 4.2). Generally, five musical parameters were manipulated, resulting in 18 stimuli or conditions (see **table 4.1** for a list of the stimuli and associated factors):

1. Texture: ten stimuli included segregated layers (left hand or right layers, equally distributed) and the remaining eight included the combined layers;
2. Timbre and frequency: eight stimuli used the reverberated teak stick sample (indeterminate pitch) and the remaining ten stimuli used the reverberated Steinway piano sample (determinate pitch);
3. Attack velocity: left hand notated accents were neutralized for seven stimuli, and accents were added to the right hand attacks that coincided with a notated quarter beat for four stimuli; all accents were produced by increasing attack velocity (on standard MIDI scale: flat velocity = 70, accented notes = 85-110);
4. Duration: half of the ten stimuli using the piano sample had their performance manner

[16] This was done to avoid incidental noise produced by the computer performance. Since all participants were exposed to the same computer-generated stimuli, there is no way to test whether panning had an effect on participants' behavior. Future studies could compare participants' reponses to computer-generated stimuli with those to live performances.

converted from staccato to legato and/or legato to staccato (staccato = 45% of quarter note) and tenuto accents were performed at 125% of notated value when included;

5. Register: within the ten stimuli with determinate pitch, the left hand's register was collapsed in two of them so that it would not exceed an octave (which affected the contour) and one stimulus with combined layers also had the two layers reversed (which pertains more directly to register).

The general organization into blocks was as follows: Blocks 1 and 2 used the segregated layers conditions (left hand or right hand layer), while Blocks 3 and 4 used the combined layers conditions; Blocks 1 and 3 included the indeterminate pitch conditions (teak timbre), and Blocks 2 and 4 presented the determinate pitch conditions (piano timbre); finally, Blocks 1 and 3 were fully factorial, while Blocks 2 and 4 were semi-factorial.

The input variables described above were used as *within-subjects factors* in the main analysis, that is, factors that apply to all the participants in a group.[17] Because of the particular design used here, i.e., all participants were subjected to all the conditions (*repeated measures design*), the within-subjects factors correspond to the controlled parameters in the stimuli; these factors will be referred to as "Version," "Layer," "Timbre," "Accent," "Articulation," and "Streaming." The Streaming factor relates to the contour change in the left hand layer only (segregated version). In the combined layers conditions with piano timbre, register reversal could not be used directly as a factor in the statistical analysis because of limitations in the experimental design; however, its influence could be assessed through comparison.

[17] In statistical analysis, *input* (or *independent*) *variables* refer to the factors and conditions that may influence participants' behavior; this is in contrast with *output* (or *dependent*) *variables*, which refer to the collected data in which changes may be observed.

This design allowed for within-block as well as across-blocks analysis. Generally, within-block analyses tested the influence of various factors on (poly)metric interpretation and participants' ability to follow competing metric frameworks: the within-block analyses for Blocks 1 and 2 aimed at identifying the influence of the factors Layer, Accent, Articulation, and Streaming (contour), and those of Blocks 3 and 4 tested the influence of layer-specific accentuation. In addition, the within-block analyses of Blocks 1 and 2 explored the relative degree of mensural determinacy of each rhythmic layer, while those for Blocks 3 and 4 explored the realization of polymetric potential and the possible influence of a combination of musical parameters on participants' ability to follow competing metrical frameworks. Across-blocks analyses allowed for the testing of the Timbre (Blocks 1&2 and 3&4) and Version (Blocks 1&3 and 2&4) factors; in **table 4.1**, across-blocks factors are italicized. Finally, in comparing selected conditions from Blocks 1 and 3 as well as Blocks 2 and 4, the question of how mensural determinacy changes when layers are combined could be addressed. Details on specific statistical methods are included in subsection 4.1.2.6.

Table 4.1: Stimuli description with factors and conditions (S = stimulus; LH = left hand and RH = right hand)

Block 1 *Version: segregated layers* *Timbre: teak (indeterminate pitch)* Layer: LH/RH Accent: off/on	**S1:** LH, accents off
	S2: LH, accents on
	S4: RH, accents off
	S5: RH, accents on
Block 2 *Version: segregated layers* *Timbre: piano (determinate pitch)* Layer: LH/RH Accent: off/on Articulation: staccato/legato *Contour: as written/collapsed	**S3a:** LH, accents off, staccato, as written
	S3b: LH, accents on, staccato, collapsed
	S3c: LH, accents off, legato, as written
	S6a: RH**,** accents off, legato
	S6b: RH, accents on, legato
	S6c: RH, accents off, staccato
Block 3 *Version: combined layers* *Timbre: teak (indeterminate pitch)* Accent: LH accents off/on, RH accents off/on	**S7:** LH accents off RH accents off
	S8: LH accents on RH accents on
	S10a: LH accents on RH accents off
	S10b: LH accents off RH accents on
Block 4 *Version: combined layers* *Timbre: piano (determinate pitch)* Accent: off/on Articulation: staccato/legato *Contour: as written/collapsed Register: low/high	**S9:** LH, accents on, staccato, low, as written RH, accents off, legato, high
	S11a: LH, accents off, legato, low, collapsed RH, accents off, legato, high
	S11b: LH, accents off, legato, low, as written RH, accents off, staccato, high
	S12: LH, accents on, legato, high, as written RH, accents off, staccato, low

Italics denote factors that were used in across-blocks analyses.
*. Applies to the left hand layer conditions only.

4.1.2.4 Apparatus and Experimental Procedure

The experiment took place in a private room in the music department. Participants were sitting at a desk and the principal investigator (this author) was sitting about 5 feet away on the left side of the desk. All sound files were placed on an interface controlled by a program created with Max/MSP and run with a MacBook 1.83 GHz Intel Core Duo.[18] The sound files were played through standard quality headphones (SENTRY HO268) and participants were instructed to adjust the volume to a comfortable level. The participants controlled all manual operations and tapping was performed using the right arrow key on the computer keyboard.

Prior to the experiment, participants were asked to fill-out a questionnaire about their general musical experience and training, experience with twentieth-century music, and familiarity with polyrhythm and/or polymeter. Participants who had performed music that made use of these techniques were also asked to provide some details about learning and performance strategies. Before beginning the experiment, participants signed a consent form and the principal investigator read an introductory statement about the experiment procedure. This was followed by a short practice session with a musical stimulus not included in the experiment. During the experiment, all instructions were given verbally and participants could ask for clarifications at any time as well as provide verbal comments in between listening and tapping tasks. For the listening task, participants were instructed to move with the music (or not) as they would normally do while listening to music. For the tapping task, participants were instructed to use the tapping to express their interpretation of the "underlying pulse or beat," which might involve not tapping at all if they felt there was no

[18] This program was designed by Paul Riker.

underlying pulse, or altering the speed or duration associated with the beat if they felt that it was altered.

In the first three blocks, participants listened to each stimulus twice, first without tapping and then tapping along.[19] Although participants were instructed to first listen without tapping, many used this first listening as a "practice" session, trying out a given beat or pulse level by tapping on the desk or conducting with one or more limbs (hands or feet). In the last section, participants listened to each sound file three times. After the first listening, they were asked to describe the texture. The main purpose of this task was to establish a common vocabulary that could then be used by the principal investigator to subsequently direct the participants' attention. For the second listening, participants were instructed to focus on one specific layer and tap along "its main beat or underlying pulse." When a participant had identified more than two layers, the principal investigator asked the participant to focus on a given combination of layers. For the third listening, participants were asked to focus on the contrasting layer and tap along "its main beat or underlying pulse."

It was assumed that order of presentation of left hand and right hand stimuli could also have an effect; order of presentation was counterbalanced in various ways between participants throughout the experiment. In Blocks 1, 2, and 4, the order was counterbalanced based on Layer (left hand first or right hand first). In Blocks 1 and 3, unaccented (flat velocity) stimuli were also always presented first. In addition, in all the blocks, the order was

[19] Although participants were given the option to repeat the tapping task if they were not satisfied with their interpretation of the underlying pulse, most of them did not elect to do so. A record was kept of repeat performances but only the first tapping response was used in the analyses presented here; repeat performances may be used at a later point for further analysis. There were also a small number of "false starts" due to technical issues; the computer saved these as empty files, but they were not used for analysis.

systematically varied to maximize presentation order across all participants and avoid a confounding effect of order in the statistical analyses. The experiment lasted roughly between 1 and 2 hours depending on the number of verbal comments made by each participant, and participants were instructed to take a short break between the second and third blocks.

4.1.2.5 Data Collection and Measures

Three types of data were gathered: (1) participants' verbal comments about their listening experience and/or tapping strategies were recorded using a SONY ICD-P320 digital voice recorder; (2) observations on the body movements of the subjects during listening and tapping were written down by the principal investigator during the experiment; and (3) tapping responses were recorded as series of time points (measured in milliseconds) and saved as text files. The results presented in this chapter are largely based on tapping responses, which were used to generate two basic types of performance measures: tapping consistency and synchronization measures. Tapping consistency was measured with respect to variation in tapping period and synchronization measures focused on tapping phase; tapping periods for each participant were also represented graphically and used in the preliminary stages of data analysis. Both of these types of measures allowed for the exploration of the three goals pertaining to relative degree of mensural determinacy, musical parameters, and ability of follow competing metrical frameworks.

The first step of data processing consisted in importing the time-points text files into Excel and generating *inter-tap interval* (ITI) series for each tapping response, which were used to create tapping contour graphs. A *tapping contour graph* is a visual representation of

an ITI series, with the series of taps on the horizontal axis and the duration of each tap on the vertical axis (examples of tapping contour graphs are discussed in subsection 4.2.1 and section 4.4). These graphs were used in preliminary analyses for visual appraisal of changes in tapping behavior within and across stimuli and to identify segments exhibiting regular tapping periods corresponding to the perceived "pulse" or "beat." The mean ITI corresponds to the average tempo expressed by a given tapping response; it is representative of the period associated with a given pulse level. ITI series were also used as a basis for the first basic performance measure aimed at assessing tapping consistency. Consistency measures in pulse-finding experiments are typically based on the *standard deviation* (SD) of an ITI series, i.e., the root-mean-square (RMS) deviation of the ITI values from their mean. The measure used here was the *coefficient of variability* (CV), which is calculated by dividing the standard deviation by the mean inter-tap interval (CV = SD/mean ITI) so that it will allow for comparison across tempi and participants.[20]

The second stage of data preparation involved the conversion of the Excel data sets to text files and their processing by a computer program designed to identify those taps that were synchronized with the various pulse streams, as determined through score analysis.[21]

---

[20] The form used here is borrowed from Joel Snyder and Carol L. Krumhansl, "Tapping to Ragtime: Cues to Pulse Finding," in *Music Perception* 18/4 (Summer 2000): 455-489. Given that the goal of the experiment was not to measure participants relative success in synchronization to "the" beat, but to explore tapping performance in a polymetric texture, my use of CV differs from these authors' in one important respect: it includes all taps, whether or not they are actually synchronized with an actual beat. Therefore, the results based on CV measures presented in this chapter are not directly comparable with those in the literature. The advantage of using this form is that it is more sensitive to tempo/period fluctuations. The disadvantage is that the wider variance may hamper statistical significance.

[21] I am indebted to Panayotis Mavromatis who created the basic scripts used for processing the data.

The tolerance window for synchronized taps was set to ± 100ms.[22] The main pulse streams used at this stage included two different phases of quarter-note beats at the lower level (corresponding to the left hand and right hand lower pulse level) and two higher level pulses, i.e., the dotted-half note (right hand) and whole note (left hand) beats. The resulting text files were once again imported into Excel and formatted for the basic synchronization measures; this involved matching the relevant pulse streams to each stimulus and standardizing the cut-off points across stimuli.

While closely related, the two basic synchronization measures used assess participants' accuracy from complementary perspectives; these measures proved to be especially relevant for the consideration of relative degree of mensural determinacy. *Tapping synchronization* (Tsyn) corresponds to the ratio of synchronized taps to total number of taps at a specific pulse level. *Beat synchronization* (Bsyn) measures the number of synchronized taps against the total number of available beats in a give pulse stream. Thus, while tapping synchronization is a measure of participants' involvement with a given pulse level, beat synchronization measures the overall rate of successful synchronization to this pulse level; both measures assess the influence of a given pulse level on participants' tapping responses and as such they are expressive of the mensural determinacy of a rhythmic strand, that is, the relative strength of a given beat structure as projected by a unified sequence of rhythmic events.[23] The combined usage of the two measures allows for a more nuanced interpretation

---

[22] This is based on the criteria established by Carolyn Drake, Amandine Penel, and Emmanuel Bigand in "Tapping in Time With Mechanically and Expressively Performed Music," *Music Perception* 18/1 (2000): 1-23, as well as Joel S. Snyder and Carol L. Krumhansl in "Tapping to Ragtime: Cues to Pulse Finding," *Music Perception* 18/4 (2001): 455-489.

[23] The concept of "mensural determinacy" is adapted from Hasty, *Meter as Rhythm*.

of tapping behavior. Both tapping and beat synchronization measures were calculated for several pulse levels, from the common eighth note pulse to the competing and out-of-phase dotted-half and whole notes (identified in the analysis). However, the analyses reported in this chapter are based on the quarter-note pulse level (or beat) only, although both downbeat and upbeat phases will be considered.

4.1.2.6 Data Analysis[24]

The results presented in this chapter focus on *inferential* statistical methods.[25] Although expected results will sometimes be discussed, all of the analyses were conducted based on the *null hypothesis*, i.e., using the basic assumption that the independent variables or factors did not have an effect on the dependent variables.[26] The significance level for all these analyses was based on a 95% confidence interval (i.e., $p < .05$). In general, trends will be reported only for p values between .05 and .059, and significance levels will be reported in full whenever possible (e.g., $p = .02$).

Three types of statistical analyses were used: (1) Pearson's correlation, (2) paired-samples t tests, and (3) univariate analyses of variance (ANOVA). Since the number of hours

---

[24] In the course of this process, I received the help of several persons. I am especially indebted to Tina Zottoli for her guidance through the steps of the statistical analysis.

[25] This is in contrast to *descriptive* statistics. The main difference between these two approaches is that inferential statistics aim to establish a causal relationship in order to make predictions about the general population based on the response patterns observed in a random population sample, while descriptive statistics provide quantitative information about the population sample.

[26] The "goal" of inferential statistics is thus to determine whether the null hypothesis can be rejected based on a given confidence window or *significance level* (*p* value)

spent weekly with twentieth-century music was the only between-subjects factor that was a continuous variable, a Pearson's correlation was used only for assessing the relationship between this factor and overall tapping consistency (CV).[27] Paired-samples t tests were used as necessary to determine significant differences between sets of conditions. ANOVAs were used for factorial and semi-factorial designs within and across blocks to assess significant main effects and interactions of within- (main analysis) and between-subjects factors (exploratory analyses). Greenhouse-Geisser adjustments for sphericity were applied whenever necessary; pairwise comparisons and post-hoc tests were performed using a Bonferroni adjustment for multiple comparisons. The choice of adjustments was guided by the repeated measures design, with a preference for a more conservative adjustment (Bonferroni) because of the relatively small sample size. Finally, given that the sample was convenient rather than random, generalizations are applicable only for the population represented by the sample, i.e., music students, amateur and professional musicians with at least one year of formal musical training who are actively engaged in musical activities (including performance, analysis, and composition).[28]

---

[27] A *continuous variable* is one that can be measured according to a scale, in contrast to a *discrete variable*, which has a finite value; the latter type also includes categorical data such as type of instrument played and whether or not participants had prior-experience with the performance of polyrhythm (based on "yes" or "no" answers).

[28] This is the case in most music-psychological studies, even when they use both musicians and non-musicians. Thus, it should be kept in mind when evaluating the results obtained in experimental studies.

4.2 Results

The first subsection presents the results from the correlation of the overall coefficient of variability and hours spent weekly with twentieth-century music (inferential) for each participant, followed by a survey of selected tapping contour graphs illustrating different tapping strategies and the wide range of individual responses (descriptive). The four remaining subsections will report the results of the statistical analysis of Blocks 1 through 4. Each section will begin with a preliminary analysis, followed by the main analysis with within-subjects factors. The main analysis will be followed by selected results from the exploratory analyses with between-subjects factors. Finally, analyses across blocks will be conducted where applicable (i.e., Blocks 1 & 2, 1 & 3, 2 & 4, and 3 & 4). In order to facilitate interpretation, brief discussions of the results have been integrated within each subsection and short summaries are provided whenever deemed necessary.[29]

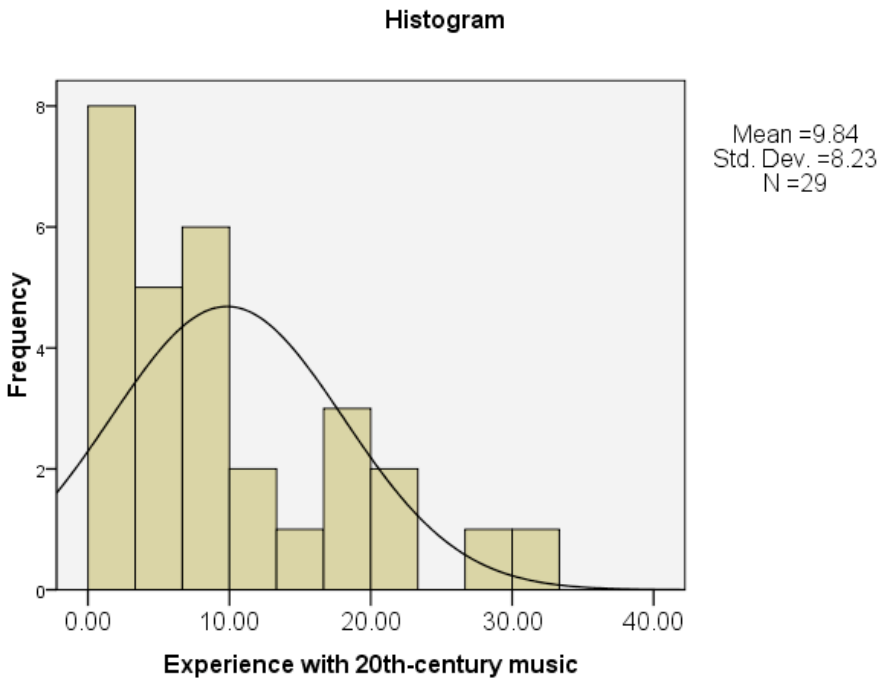
4.2.1 Overall Tapping Consistency and Tapping Contour Graphs

The tapping consistency measure showed a wide variation across participants, with overall coefficient of variability values (CV = SD/mean ITI) ranging from 0.088 (most stable) to 0.880 (mean = 0.379, SE = 0.034). In order to assess the effect of participants' musical experience on tapping consistency, a Pearson's correlation was performed using the overall CV value and the number of hours spent weekly with twentieth century music (which included performance, analysis as well as listening) for each participant. **Figure 4.1** provides a histogram (with normal distribution curve superposed) of participants' data and **figure 4.2**

[29] This is not the standard practice in psychological literature and this choice was made because of the number of results as well as exploratory nature of many of the components.

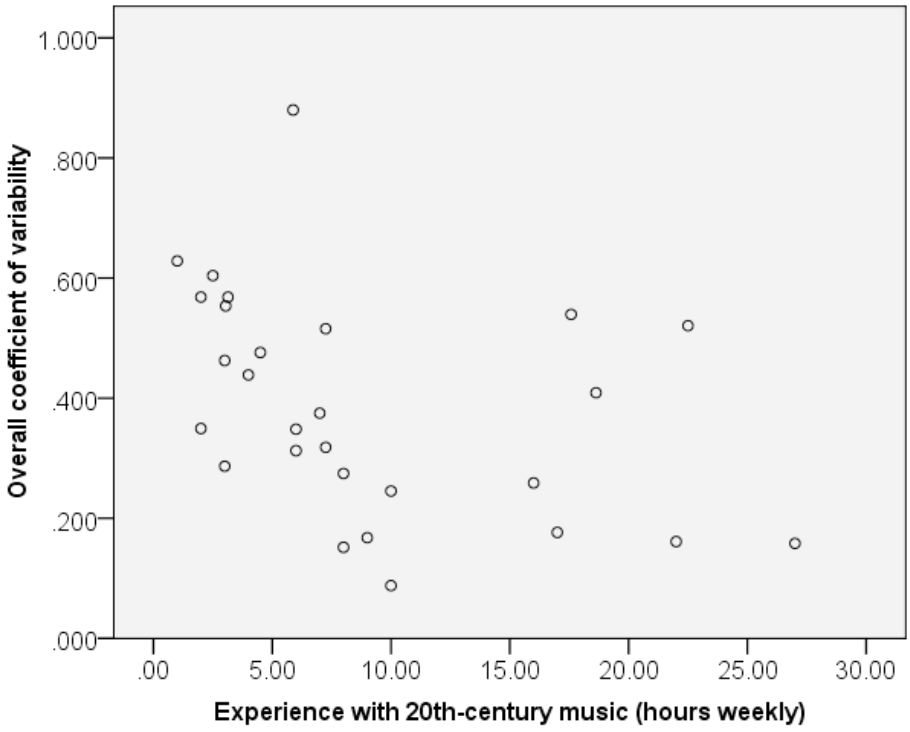
provides a plot of the correlation. The results show that there was a significant negative correlation between these two variables, r = -.406, p = .032 (number of valid responses = 28). Thus, participants who spent more time involved with the musical repertoire to which the source materials belonged tapped in a more stable way (lower CV value) than those participants who were less engaged with it. However, there appears to be a number of outliers as well as a ceiling effect situated around 10 hours per week, which suggests that there might be other between-subjects factors related to musical training and experience that have a significant effect on participants' tapping performance.[30] As we will see, more support in that direction was provided by the exploratory analyses within each block.

Figure 4.1 Participants' experience with twentieth-century music (hours weekly)



[30] It is also possible that some participants overestimated the number of hours, but this is a problem that is inherent to this type of measure.

Figure 4.2 Correlation of tapping consistency and musical experience



In the preliminary stages of data analysis, tapping contour graphs were surveyed to assess individual differences as well as general trends in tapping behavior. This survey confirmed that there was a wide range of tapping modes across participants for the same stimulus as well as between stimuli for individual participants. For example, **figure 4.3** provides a tapping contour graph comparing the tapping responses of Participant 30 to the four stimuli of Block 1 (segregated layers with teak timbre/indeterminate pitch). As seen in this graph, the two basic structures (left hand and right hand layers) gave rise to two distinctive tapping contours. Stimuli 1 and 2 (left hand layer) were interpreted as more periodic than stimuli 4 and 5 (right hand layer), the latter showing tempo fluctuations (i.e., acceleration and deceleration patterns) as well as more erratic tapping. For the same stimuli, tapping responses showed other modes of responses, from mostly regular to mostly irregular

across the two layer (left/right) and accent (off/on) conditions. These tapping contour graphs also provided insights into the possible effects of musical parameters on rhythmic interpretation. In **figure 4.3**, the period shift from a mean ITI of 418ms to 829ms from stimulus 1 to stimulus corresponding to a 2:1 ratio correlates with the addition of accents to the left hand layer. In other words, *hearing the accented whole-note pulse in the left hand layer caused this participant to tap slower*, shifting from an eighth-note pulse to a quarter-note pulse.[31] In contrast, the addition of accents in the right hand layer (stimulus 5) does not appear to have had an effect on this participant's beat interpretation.

[31] At the notated tempo of 72 quarter-beats per minute (bpm), the inter-onset interval associated with the quarter-note pulse is 833ms (60s/74bpm), while that of the eighth-note pulse is 417ms (60s/144bpm).

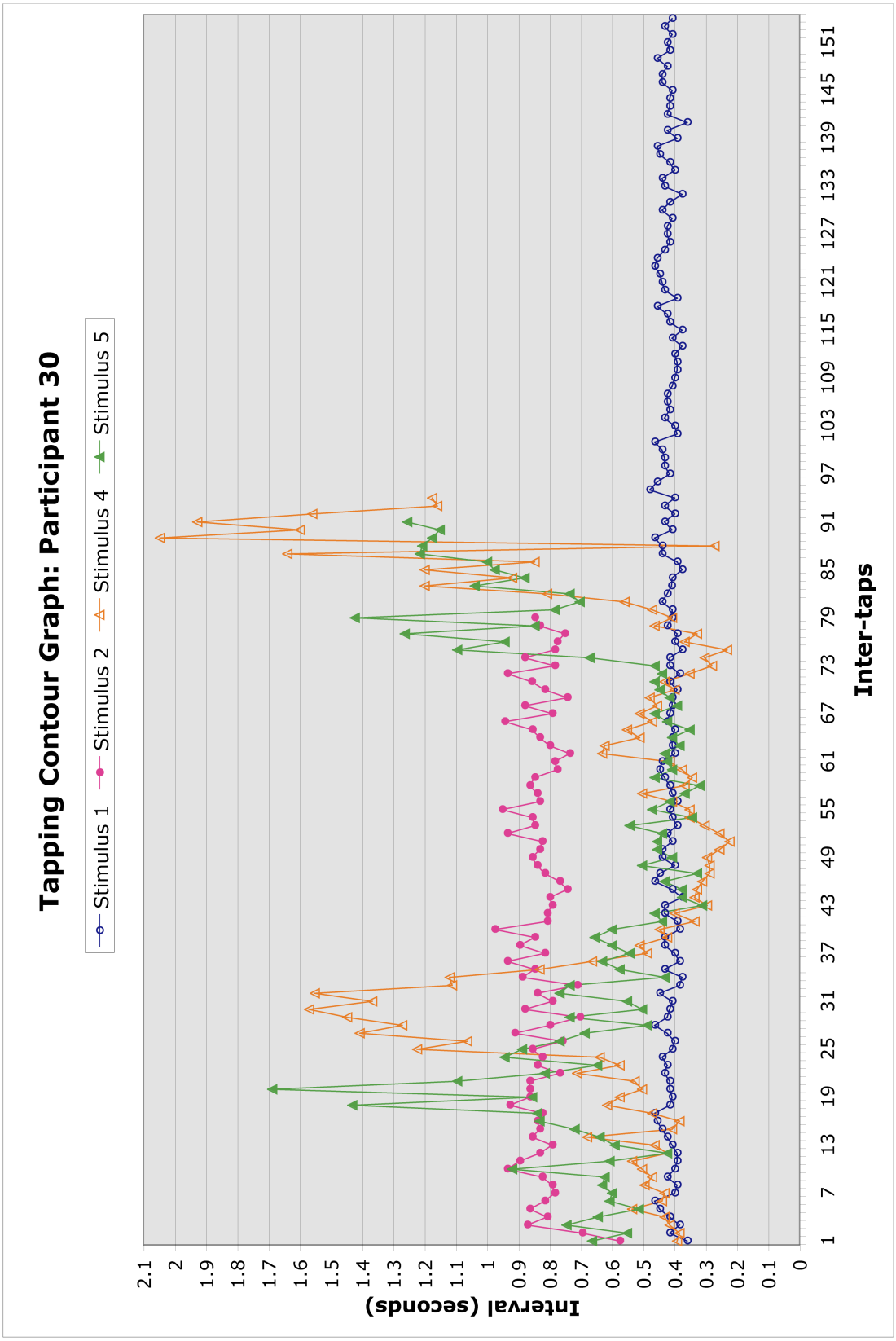
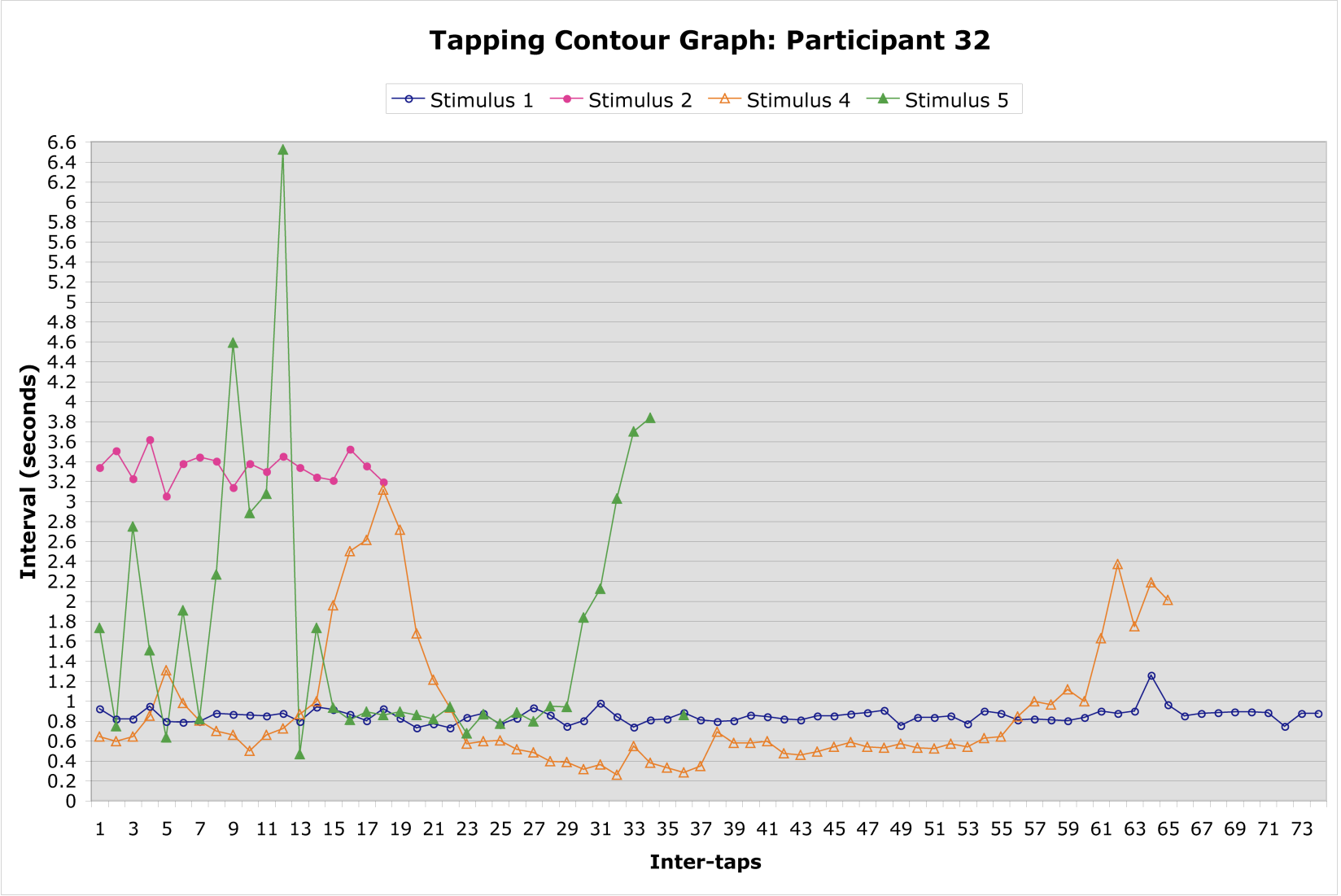


Figure 4.3: Tapping contour graph of Participant 30 for Block 1 stimuli (1 = unaccented left hand layer, 2 = accented left hand layer, 4 = unaccented right hand layer, and 5 = accented right hand layer)

Participant 32 (**figure 4.4**) performed a similar shift in tapping period when accents were added to the left hand layer (with a 4:1 ratio), successfully synchronizing with the left hand whole-note beats and exhibiting a mean ITI of 3.3s that was subject to very little variability (mean CV = 0.043).[32] Based on my observations of this participant's body movements during the experiment, this tapping period was assisted by intermediate beats subdividing this timespan in four equal parts, as marked by a conducting pattern with the hand used to tap the beat (right hand). This graph also shows contrasting responses to the right hand layers for the two accents conditions. Thus, when the right hand layer is unaccented (stimulus 4), the tapping contour exhibits smooth variations in tempo: for example, inter-taps 13 to 23 show a deceleration followed by an acceleration. When the right hand layer is accented (stimulus 5), the resulting tapping contour is more rugged at the beginning, that is, it is characterized by unequal inter-tap intervals, although some of these intervals exhibit a timespan ratio of about 2:1 (e.g., inter-taps 1-2 and 6-7). This contour suggests that *this participant interpreted the alternating series of half and quarter notes in the right-hand layer as the governing temporal pattern* (this is illustrated in **ex. 3.21**). In contrast, the tapping segment including inter-taps 15 to 29 is much more regular; here, the participant appears to have settled for a pulse close to the quarter-note beat (mean ITI = 855ms). Finally, the last segment (inter-taps 29-34) is characterized by a gradual deceleration from 936ms to 3.832s.

[32] As mentioned in the "Experimental Design" section (subsection 4.1.2.1), despite that fact that there are only 30 valid responses overall, participants are identified using numbers 1 to 32 based on the original number of volunteers.

Figure 4.4: Tapping contour graph of Participant 32 for Block 1 stimuli (1 = unaccented left hand layer, 2 = accented left hand layer, 4 = unaccented right hand layer, and 5 = accented right hand layer)



Thus, it appears that overall *these two participants' tapping behavior was strongly influenced by layer and accentuation but the specific effect of accentuation was dependent on layer*; each participant also appears to have responded differently to specific combinations of layer and accentuation. Both participants showed contrasting tapping contours for the two layers, with the left hand layer being generally more regular than the right hand layer. The addition of accents in the left hand layer prompted both participants to tap a slower beat and both of them appear to have latched on the quarter-note beat, but Participant 30 used it as the main beat while Participant 32 integrated it into a larger gestural pattern. In contrast, the effect of accents in the right hand layer was not consistent across these two participants, which is symptomatic of an interaction between accentuation and between-subjects factors, that, those factors linked to the specific musical training and experience of individual participants. While the first participant (Participant 30) showed no difference in tapping behavior when the right hand layer was accented, the second (Participant 32) appears to have interpreted the accented events as signals for the underlying temporal structure.

This discussion suggested ways in which musical parameters can influence participants' tapping responses. It also pointed to differences between participants. The goal of the statistical analyses presented in the remaining subsections was to identify specific within- and between-subjects factors as well as to determine whether the effect of these factors was statistically significant.

4.2.2 Block 1

Block 1 presented each layer of the polyrhythmic compound separately, first with and then without accents; the stimuli were prepared using a teak stick sound sample (Teak timbre condition) so that participants would be exposed to the temporal structure of each layer successively without the interference of pitch. This particular design was aimed at testing the relative degree of mensural determinacy in each rhythmic strand independently of overall pitch structure. It also allowed for a statistical analysis of the influence of accentuation as well as possible interaction between temporal structure (as represented by layer) and accentuation. The analysis of between-subjects factors (exploratory analyses, subsection 4.2.2.3), which sought to identify aspects of participants' musical training and experience that could either have an effect on their interpretation of the beat structure (as represented by phase synchronization) or tapping consistency.

4.2.2.1 Preliminary Analysis: Order of Presentation and Phase Discrimination

As discussed in the "Experimental Design" section, the order of presentation of stimuli in Block 1 (Order.Block1) was counterbalanced between layers based on the assumption that order would have an effect. A preliminary univariate repeated measures analysis of variance (ANOVA) was performed with Layer (left/right) and Accent (off/on) as within-subjects factors, and Order.Block1 as between-subjects factor (LH-RH = 15 and RH-LH = 14, for a total of 29 valid responses). Order of presentation had a significant main effect on participants' synchronization to both the downbeat phase, TsynQD, $F(1, 27) = 11.004$, $p < 0.01$, and upbeat phase, TsynQU, $F(1, 27) = 10.959$, $p < 0.01$, of the quarter-note

beat.[33] However, it did not interact significantly with either Layer or Accent ($p \geq 0.05$), and thus, it was not included in subsequent analyses.

Table 4.2: Estimated marginal means and mean differences for downbeat and upbeat synchronization (TsynQD and TsynQU) based on order of presentation in Block 1 (Order.Block1)

Measure	Order.Block1 (I) (J)	Mean	Standard Error	Mean Difference (I – J)	Standard Error	Significance (*p* value)
TsynQD	LH-RH	.377	.025	.122*	.037	.003
	RH-LH	.256	.026			
TsynQU	LH-RH	.153	.015	-.072*	.022	.003
	RH-LH	.225	.016			

*. The mean difference is significant at the .05 level.

The significant main effect of order of presentation (Order.Block1) on participants' synchronization rate is suggestive of the influence of context on listeners' interpretation of beat structure. The participants who heard and tapped to the left hand conditions first (LH-RH) synchronized with the downbeat of both layers at a significantly higher rate (37.7%) than those who performed these tasks in the reverse order (25.6%), resulting in an estimated mean difference of 12.2% (see **table 4.2**).[34] Furthermore, hearing and tapping to the right hand conditions first (RH-LH) was correlated with a significantly higher rate of synchronization to the upbeat, although the estimated mean difference is smaller (7.2%). The significant difference in downbeat tapping synchronization rates (TsynQD) suggests that

[33] Although the left hand downbeat is equivalent to the right hand upbeat, the labels "D" (downbeat) and "U" (upbeat) attached to the tapping and beat synchronization measures (Tsyn and Bsyn) for segregated layers indicate that the alignment of participants' taps with the quarter-note beat (Q) has been determined within each layer.

[34] Pairwise comparisons are based on estimated marginal means. For ease of reading, values in all results tables are given without a zero to the left of the decimal and converted into percentage in the text when applicable (1 = 100%).

performing the listening and tapping tasks in the left-hand-first order provided the participants with a temporal template with which to interpret the fluctuating surface of the right hand layer more successfully. Given that the surface irregularity that also characterizes the left hand layer is mitigated by a regular whole note beat, this result also suggests that the background pulse of the left hand layer, which was heard first without and then with accents, acted as a temporal anchor for the more foreground quarter-note beat; however, because there was no control for a carry-over effect in this experiment, it is not possible to test for this hypothesis.

The significantly higher synchronization rate for the upbeat phase of the quarter-note beat that was observed in the right-hand-first order (RH-LH) is more difficult to interpret. The synchronization measures used in this analysis do not allow us to determine the exact period of participants' tapping, but only its phase in relation to the quarter-note beat. Thus, even a higher downbeat synchronization rate does not guarantee a tapping period equivalent to a quarter-note beat, although it does suggest a period "consonant" with it, which is to say, one that is a multiple of the quarter-note beat and is in phase with it. One possible explanation for the higher number of taps successfully synchronized with the upbeat (TsynQU) is that *participants who heard the right hand layers first may have been induced into tapping at a faster rate*, e.g., an eighth note beat, although other tapping periods could produce a similar result, e.g., a dotted-quarter-note beat that would coincide with every other downbeat and upbeat. The following tests will shed more light on this issue.

Table 4.3: Paired samples t-tests results for tapping synchronization to the downbeat and upbeat phases (TsynQD and TsynQU) in Block 1 (S = stimulus)

Paired Samples		Mean Difference	Standard Deviation	Standard Error Mean	Significance (2-tailed)
Pair 1	S1.TsynQD - S1.TsynQU	.129*	.248	.045	$t(29) = 2.861$, $p = .008$
Pair 2	S2.TsynQD - S2.TsynQU	.219*	.345	.063	$t(29) = 3.473$, $p = .002$
Pair 3	S4.TsynQD - S4.TsynQU	-.006	.138	.025	$t(29) = 0.238$, $p = .813$
Pair 4	S5.TsynQD - S5.TsynQU	-.169*	.173	.032	$t(28) = -5.245$, $p = .000$

*. The mean difference is significant at the .05 level.

In order to determine whether downbeat and upbeat synchronization differed significantly within each stimulus, paired-samples t tests were performed for the two synchronization measures (Tsyn and Bsyn); here, I will only report the results for the tapping synchronization measure (Tsyn), which is more directly relevant to the present discussion. As can be observed from **table 4.3**, the two left hand stimuli (S1 and S2) exhibit a significantly higher tapping synchronization to the downbeat phase of the quarter beat, while the two right hand stimuli yield a lower synchronization rate for the downbeat phase, only one of which is significantly lower. In other words, *while tapping to the left hand layer, participants showed a significant preference for the downbeat phase of the quarter note* (a mean difference of about 12.9% and 21.9% for the two stimuli; the higher of the two corresponds to the accented condition). *On the other hand, participants' tapping to the downbeat and upbeat phases for the right hand layer was either not significantly different (S4) or significantly favored the upbeat phase (S5).* Based on these results, the explanation suggested above for higher upbeat synchronization in the right-hand-first order (RH-LH) as resulting from a tapping period

equivalent to an eighth note beat is possible. However, the significant preference for the upbeat phase in the accented right hand condition (S5) also indicates that participants might have been induced into syncopated tapping, i.e., out-of-phase synchronization. Furthermore, the lack of a significant difference between downbeat and upbeat phase synchronization could also be explained by a shifting phase. Thus, although suggestive, the evidence is not conclusive and more analysis (and different measures) would be required to settle this question.

4.2.2.2 Main Analysis: The Effect of Layer and Accent

I will now present the results from the main analysis of within-subjects factors in Block 1, followed by selected results from the exploratory analyses with mixed factors (within- and between-subjects). Because of the relatively small sample size, this approach was chosen to insure greater reliability of the statistical results for within-subjects factors.

Table 4.4: Descriptive statistics for Block 1 (Legend for stimuli descriptors: S = stimulus; LH = left hand layer; RH = right hand layer; Tk = teak; "off"/"on" applies to accents)

Measure	Stimulus	Mean	Standard Deviation
CV	S1.LH.Tk.off	.349	.323
	S2.LH.Tk.on	.256	.181
	S4.RH.Tk.off	.478	.277
	S5.RH.Tk.on	.564	.296
TsynQD	S1.LH.Tk.off	.323	.174
	S2.LH.Tk.on	.386	.223
	S4.RH.Tk.off	.240	.069
	S5.RH.Tk.on	.325	.117
BsynQD	S1.LH.Tk.off	.258	.208
	S2.LH.Tk.on	.317	.248
	S4.RH.Tk.off	.176	.102
	S5.RH.Tk.on	.216	.110
TsynQU	S1.LH.Tk.off	.198	.112
	S2.LH.Tk.on	.153	.136
	S4.RH.Tk.off	.244	.095
	S5.RH.Tk.on	.157	.078
BsynQU	S1.LH.Tk.off	.152	.142
	S2.LH.Tk.on	.106	.117
	S4.RH.Tk.off	.171	.104
	S5.RH.Tk.on	.116	.087

Table 4.4 presents the means and standard deviations of each individual condition across participants for the five performance measure. As can be observed from this table, higher coefficient of variability (CV) levels are associated with the right hand conditions (less consistent tapping), while lower CV corresponds to the left hand conditions (more consistent tapping), although the association of the highest standard deviation with the first

left hand layer condition (S1) suggests a greater variation in participants' performance for this condition. Synchronization measures also reveal that *participants' tapping was most representative of the quarter-note downbe*at (TsynQD and BsynQD) *when tapping to the accented left hand condition* (S2): successful synchronization to the quarter downbeat in S2 represents about 38.55% of all participants' taps, corresponding to 31.66% of all available quarter downbeats.[35] *The least representative was the unaccented right hand condition* (S4), in which only 24.03% of all participants' taps were synchronized to the quarter downbeat and only 17.62% of all available quarter downbeats were synchronized to; this condition also exhibited the highest levels of quarter upbeat synchronization (24.41% and 17.1% for TsynQU and BsynQU, respectively). This is consistent with the finding that there was no significant difference between downbeat and upbeat phases in this condition and that participants might have been tapping at a faster rate (e.g., a eighth note beat) or mitigating pulse level (e.g., a dotted-quarter note), or were shifting in phase. These descriptive statistics also show that downbeat synchronization was generally higher than upbeat synchronization, although standard deviation levels were also generally greater in the left hand conditions, with the largest variance corresponding to the condition with accents on (S2 = 22.3%). The high standard deviation suggests that *the presence of accents in the left hand layer might have been interpreted differently depending on participants' musical training and experience*. This proposition will be tested by the exploratory analyses with between-subjects factors (subsection 4.2.2.3).

[35] Because of the way they are calculated (using all taps, including interrupted and/or intermittent tapping), tapping synchronization (Tsyn) rates are generally higher than beat synchronization (Bsyn) rates. Thus, the two measures will usually be examined together by phase (QD or QU).

Table 4.5: Effect of Layer (LH = left hand and RH = right hand) in Block 1

Measure	**Mean Difference** (LH - RH)	**Standard Error**	**Significance**[a] (*p* value)
CV	-.219*	.044	.000
TsynQD	.072*	.028	.015
BsynQD	.091*	.025	.001
TsynQU	-.025	.016	.126
BsynQU	-.015	.019	.443

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

Table 4.6: Effect of Accent ("off"/"on") in Block 1

Measure	**Mean Difference** (Off - On)	**Standard Error**	**Significance**[a] (*p* value)
CV	.003	.035	.922
TsynQD	-.073*	.023	.003
BsynQD	-.049*	.021	.027
TsynQU	.066*	.018	.001
BsynQU	.051*	.021	.024

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

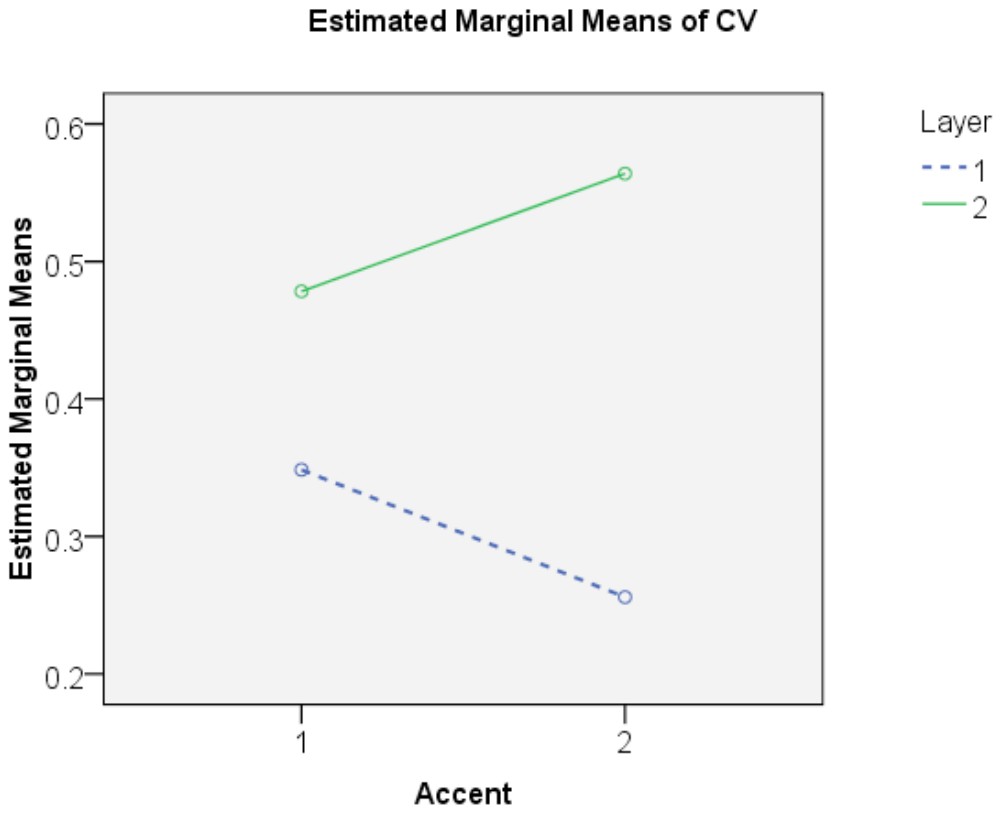
a. Adjustment for multiple comparisons: Bonferroni.

In order to determine the statistical significance of the effect observed above layer, a series of univariate repeated measures analyses of variance (ANOVA) was conducted with Layer (LH/RH) and Accent (off/on) as within-subjects factors. It was found that Layer had a significant main effect on tapping consistency (CV), $F(1, 28) = 25.130$, $p < .001$, downbeat tapping synchronization (TsynQD), $F(1, 28) = 6.719$, $p < .05$, and overall downbeat

synchronization (BsynQD), $F(1, 28) = 13.279$, $p < .01$, but not on upbeat synchronization measures (TsynQU and BsynQU). In other words, *participants' tapping behavior was influenced by the specific temporal structure exhibited by each layer only with respect to tapping consistency and synchronization to the downbeat phase*. The negative mean difference for CV in **table 4.5** shows that *participants' tapping behavior was significantly less subject to change (e.g., tempo fluctuations, interruption, etc.) in the left hand layer conditions*, which suggests that the temporal structure of this layer was perceived as more regular. This interpretation finds additional support in that synchronization to the downbeat phase of the quarter note (TsynQD and BsynQD) was also significantly higher for the left hand conditions. As observed previously, synchronization rates for the upbeat phase were higher for the right hand conditions, but differences in estimated marginal means were not statistically significant.

In contrast, accentuation did not have a significant main effect on tapping consistency (CV), but it was significant for all synchronization measures: TsynQD, $F(1, 28) = 10.631$, $p < .01$; BsynQD, $F(1, 28) = 5.485$, $p < .05$; TsynQU, $F(1, 28) = 13.439$, $p < .01$; BsynQU, $F(1, 28) = 5.679$, $p < .05$. Mean differences also show that the effect of accentuation was different depending on phase (see **table 4.6**). Thus, synchronization rates for the downbeat phase were significantly higher when accents were "on" and synchronization rates for the upbeat were significantly higher when accents were "off." These results confirm that *the presence of accents, which in these layers coincided with the downbeat phases of their respective layer, contributes to mensural determinacy as manifested by participants' tapping performance, regardless of layer.*

Figure 4.5: Layer × Accent interaction for tapping consistency (CV) in Block 1 (Layer 1 = left hand and Layer 2 = right hand; Accent 1 = "off" and Accent 2 = "on")



Measure	Layer	Accent	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
CV	LH	Off	.349	.060	.226	.471
		On	.256	.034	.187	.325
	RH	Off	.478	.051	.373	.584
		On	.564	.055	.451	.677

Finally, although there was no significant interaction of Layer × Accent, there was a trend toward significance for CV, i.e., $F(1, 28) = 3.969$, $p = .056$. The trend towards a significant interaction between the factors of Layer and Accent for the coefficient of variability measures is puzzling and should be examined more closely. **Figure 4.5** is a

graphic representation of this interaction, with pairwise comparisons results included below. As can be observed from the graph, while the participants' tapping to the left hand layer shows less variability when the accents are "on," this relationship is reversed in the right hand layer conditions. This finding suggests that the effect of accent on tapping consistency might have been different based on specific layer. In other words, *when tapping to the unaccented right hand layer, participants' had a generally more stable tapping period, but when tapping to the unaccented left hand layer, participants' showed a less stable tapping period.* Nevertheless, the wide variance across participants prevented this result to achieve statistical significance.

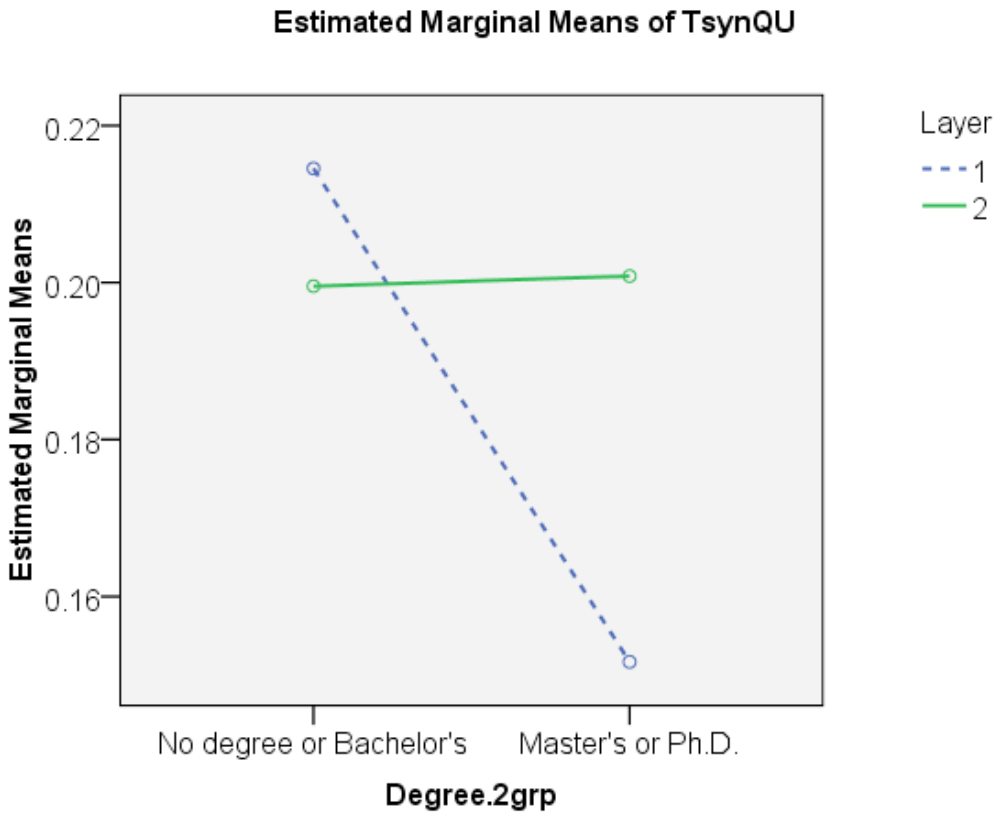
To summarize: *Although there was a significant main effect found for order in Block 1, there was no significant interaction with either layer or accentuation*; thus, these factors could be analyzed independently from order of presentation. With regard to the measuring of the relative degree of mensural determinacy projected by each layer, *the results from the statistical analysis support an interpretation of the temporal structure of the left hand layer as more mensurally determinate than the right hand layer*, *as represented by a more consistent tapping period and greater number of synchronized taps to the downbeat phase of the left hand layer*. This performance was enhanced by the presence of regularly occurring accents coinciding with the whole-note background pulse, which was previously analyzed as the downbeat of a measure of four quarter-note beats. On the other hand, *synchronization to the upbeat phase of the quarter note is less clearly defined.* The observed higher occurrence of upbeat synchronization in the right hand layer did not reach significant levels (i.e., the null hypothesis could not be rejected), but accents did have a significant main effect, with the unaccented conditions corresponding to higher upbeat synchronization. Thus, it would appear

that in this passage, *accentuation was a significant factor in participants' interpretation of the beat structure in a manner consistent with the analyzed structure*, although the observed trend toward a less consistent tapping to the accented right hand layer suggests that these performed accents, which were not written in the score but added by the analyst, might have conflicted with the participants' interpretation of the metrical structure. Furthermore, the lack of a significant difference between downbeat and upbeat synchronization in the unaccented right hand layer suggests that participants might have responded to the lower level of mensural determinacy of this layer by tapping at a faster rate.

4.2.2.3 Exploratory Analyses: Selected Between-Subjects Factors

It was suggested above that musical training and experience might have had an effect on participants' tapping behavior as greater variability was observed in the left hand condition with accents on. The influence of training and experience as well as possible interactions of within- and between-subjects factors were explored by conducting a series of univariate repeated measures analyses of variance (ANOVA) with Layer and Accent as within-subjects factors and each between-subjects factor individually, unless a specific interaction was suspected. I will now report the results that were judged most significant and interesting for future studies. It should be noted that because of the increased probability of error when conducting repeated analyses with the same variables, these results should not be taken as conclusive. However, they are suggestive of the many factors that need to be taken into consideration when studying listeners' rhythmic percept through performance tasks such as beat-finding experiments using musical stimuli.

Figure 4.6: Layer × Degree.2grp interaction for upbeat tapping synchronization (TsynQU) in Block 1 (Layer 1 = left hand and Layer 2 = right hand)



Measure	Musical Degree (2 groups)	Layer	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
TsynQU	No degree or Bachelor's	LH	.215	.029	.156	.273
		RH	.200	.017	.164	.235
	Master's or Ph.D.	LH	.152	.022	.106	.197
		RH	.201	.014	.173	.229

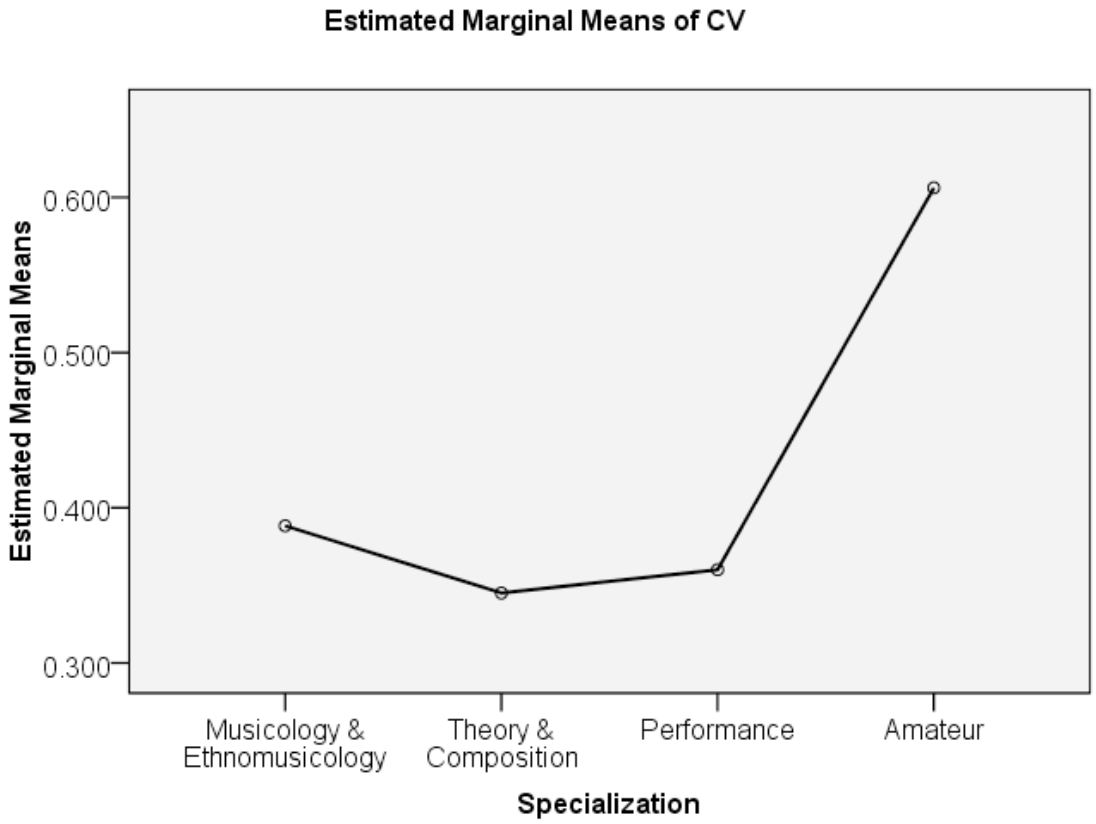
The analysis with musical degree (Degree.2grp) as between-subjects factor (No degree or Bachelor's = 11, Master's or Ph.D. = 18) showed a significant interaction of Layer × Degree.2grp for TsynQU, $F(1, 27) = 4.385$, $p = .046$, although there was no significant main effect of musical degree or interaction with Accent. **Figure 4.6** presents a graphic

representation of this interaction, with pairwise comparisons results below. As shown this graph, the effect of musical degree on upbeat synchronization appears to be layer-specific. Thus, while musical degree didn't appear to have an effect on participants' tapping on the upbeat phase of the right hand layer conditions, *participants with a master's or Ph.D. in music synchronized significantly less often with the upbeat phase of the left hand layer*. One possible explanation for these results is that more highly trained musicians tapped at a slower rate than less highly trained musicians; another explanation is that they were less frequently induced into syncopated tapping.[36]

The analysis with area of specialization as between-subjects factor (Musicology & Ethnomusicology = 6, Theory & Composition = 10, Performance = 7, Amateur = 6) yielded a significant main effect of area of specialization on tapping consistency (CV), $F(3, 25) = 3.169$, $p = .042$, but there was no significant interaction with either layer or accentuation. A significant difference was found between the groups Theory & Composition and Amateur (mean difference = 0.261, SE = 0.091). Thus, *participants specialized in theory or composition exhibited much greater regularity in tapping period than participants who are amateur musicians*. Inspection of the graph and table provided in **figure 4.7** reveals that the estimated marginal means of the first three groups were relatively close to each other, although both the Musicology & Ethnomusicology and Performance groups exhibited higher standard errors, which prevented these two groups from achieving statistically significant levels of mean difference with the Amateur group.

[36] This finding has precedents in the literature. Thus, Drake et al. (2000) report that musicians synchronize at slower rates than non-musicians; they also showed a greater range of pulse levels to which they could synchronize successfully.

Figure 4.7: Main effect of area of specialization (Specialization.4grp) on tapping consistency (CV) in Block 1

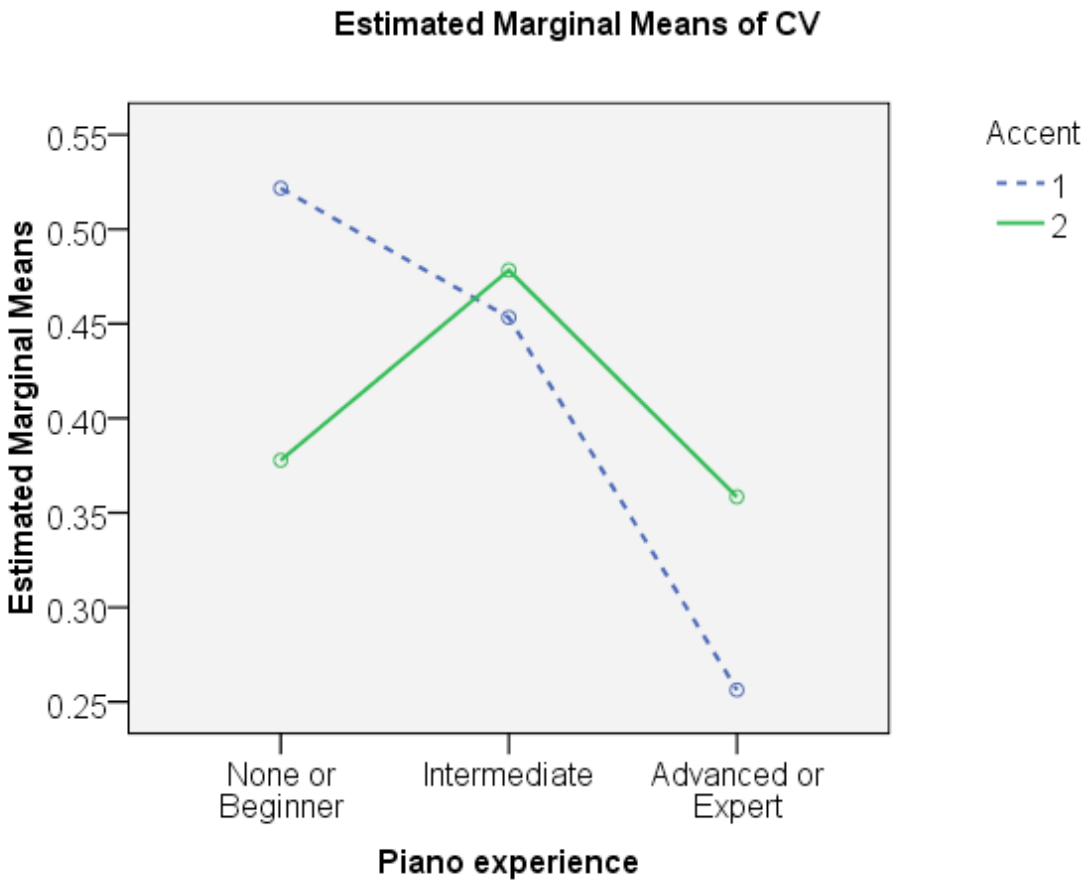


Measure	Specialization (4 groups)	Mean	Standard Error	95% Confidence Interval	
				Lower Bound	Upper Bound
CV	Musicology & Ethnomusicology	.388	.072	.241	.536
	Theory & Composition	.345	.056	.231	.459
	Performance	.360	.066	.223	.497
	Amateur	.606	.072	.458	.754

Although there was no main effect or interaction based on participants' main instrument (Piano/Percussion, High melodic, Low melodic), the analysis with Piano.3grp as between-subjects factor (None or Beginner = 9, Intermediate = 11, Advanced or Expert = 9) revealed a significant interaction of piano experience with accentuation for tapping

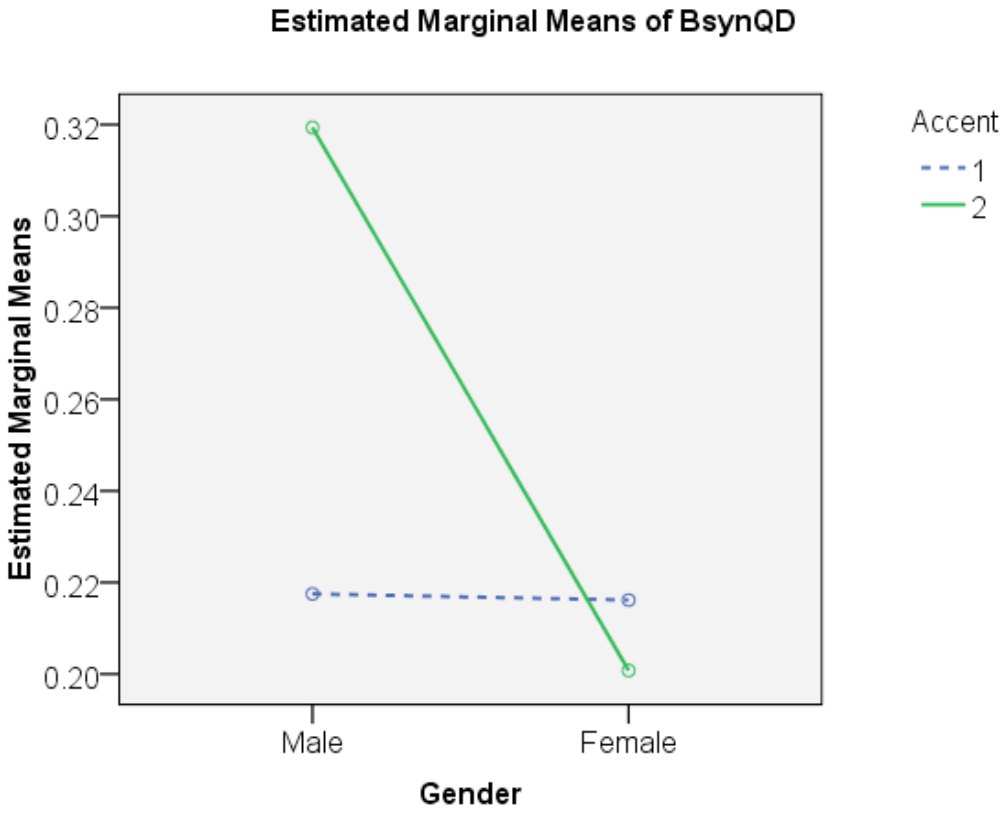
consistency (CV), $F(2, 26) = 5.150$, $p = .013$; there was no main effect of Piano.3grp on any of the dependent variables. Thus, *participants who rated their skill level at the piano as intermediate or advanced/expert showed similar levels of tapping consistency, but participants with little or no piano experience tapped more regularly when the layers were accented* (see **figure 4.8**). Although these results suggest that accents might help participants with less musical experience, it is not clear what particular aspects of piano instruction are responsible for this effect.

Figure 4.8: Accent × Piano.3grp interaction for tapping consistency (CV) in Block 1 (Accent 1 = "off" and Accent 2 = "on")



Finally, the analysis with Gender as between-subjects factor (Male = 16, Female = 13) revealed an unexpected main effect of Gender on TsynQU, $F(1, 27) = 6.073$, $p = .02$, as well as a significant interaction Accent × Gender for BsynQD, $F(1, 27) = 10.188$, $p = .004$. Thus, women's tapping performance was characterized overall by a significantly higher number of taps synchronized to the upbeat (mean = 22%, standard deviation = 1.7%) than that of men (mean = 16.2%, standard deviation = 1.6%), resulting in a mean difference of 5.7% and standard error of 2.3%. Thus, *the interaction of gender with accentuation suggests significant differences in phase synchronization based on accentuation*. **Figure 4.9** is a graphic representation of this significant interaction, with pairwise comparisons results shown below. As can be observed from the graph, when accents are off, downbeat synchronization rates for men and women are about the same (21.6% and 21.8%, respectively) but when accents are on, men's downbeat synchronization appears to be significantly higher (31.9% for men as compared with 20.1% for women). Gender differences have not been explored in the literature on related topics and there is no theoretical basis for this finding. Given that this is a relatively small sample size, *the possibility of a sampling error due to the characteristics of the female participants in this study cannot be discounted*. However, because these effects could not be tested directly, future research should examine the influence of gender.

Figure 4.9: Accent × Gender interaction for overall downbeat synchronization (BsynQD) in Block 1 (Accent 1 = "off" and Accent 2 = "on")



Measure	Gender	Accent	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
BsynQD	Male	Off	.218	.036	.143	.292
		On	.319	.039	.239	.399
	Female	Off	.216	.040	.133	.299
		On	.201	.043	.112	.290

To summarize: This exploration of between-subjects factors suggests various ways in which musical training and experience might have an effect on participants' tapping behavior and metric interpretation as represented by phase synchronization. First, the ambiguity of the results for upbeat tapping synchronization was further qualified by the observation of a mean difference specifically located in the left hand layer. *Participants with a master's or Ph.D. degree in music tapped significantly less often with the upbeat phase of the quarter-note beat, suggesting that these participants adopted a different tapping strategy that may have involved a slower tapping rate.* The previously observed significant difference between downbeat and upbeat phases in the left hand layer supports this hypothesis; there is also a precedent in the literature for this finding (e.g., Drake et al. 2000). Tapping consistency also appears to have been influenced by musical experience. Thus, *participants who specialized in theory and composition exhibited a tapping behavior that was significantly more stable than that of the amateur musicians.* On the other hand, *the presence of accents benefited the tapping behavior of participants with little or no piano experience*, but not that of participants with more piano experience. Finally, participants' gender seems to have also played a role. *Overall, women synchronized more often with the upbeat phase of the quarter-note beat and men tapped a higher number of downbeats in the accented conditions,* although the gender effect may well be an artifact of this sample and disappear in a more formal investigation.

4.2.3 Block 2

In contrast to the first block, the stimuli in Block 2 made use of determinate pitch. Thus in this block, the relative degree of mensural determinacy for each layer could be investigated with regard to overall pitch structure. As with Block 1, accents were manipulated by turning them "off" or "on." In addition, the use of determinate pitch and piano timbre allowed for the manipulation of register and contour as well as articulation (staccato and legato playing). Because of the increased number of parameters and the necessary limitations of a single experimental session, the conditions were not fully balanced within the block. This semi-factorial design necessitated the use of selected stimuli in analyses, but it still allowed for the exploration of the influence of several musical parameters on participants' tapping behavior (consistency of tapping period and phase synchronization), including articulation, which relates to the duration of rhythmic events independently of inter-onset interval.

4.2.3.1 Preliminary Analysis: Order of Presentation and Phase Discrimination

The order of presentation for Block 2 was counterbalanced in two ways, between layer and within layer conditions. As in the first block, half of the participants listened and tapped to the left hand stimuli first (S3a, S3b, and S3c), while the other half engaged the right hand stimuli first (S6a, S6b, and S6c). In addition, the order of presentation of stimuli was also counterbalanced within each layer ($3 \times 2 = 6$ possible orderings). It is assumed that this counterbalancing has accounted for possible interactions of presentation order with within-

subjects factors, but because of the relatively small sample size and the 12 possible order combinations overall (2 layers × 6 orderings), this assumption cannot be tested statistically.[37]

Table 4.7: Estimated marginal means and mean differences for upbeat synchronization (TsynQU and BsynQU) based on order of presentation in Block 2 (Order.Block2; LH-RH = left hand layers first and RH-LH = right hand layers first)

Measure	Order.Block2 (I) (J)	Mean	Standard Error	Mean Difference (I – J)	Standard Error	Significance (*p* value)
TsynQU	LH-RH	.223	.012	.035*	.017	.045
	RH-LH	.187	.011			
BsynQU	LH-RH	.189	.016	.083*	.022	.001
	RH-LH	.106	.015			

*. The mean difference is significant at the .05 level.

To test the effect of order between layers, a preliminary univariate repeated measures analysis of variance (ANOVA) was performed with Order.Block2 as between-subjects factor (LH-RH = 13 and RH-LH = 15, for a total of 28 valid responses) and the within-subjects factors Layer (LH/RH) and Articulation (staccato/legato), controlling for Accent (off), i.e., excluding stimuli that featured accents (S3b and S6b; see **table 4.1**). There was a main effect of order on both types of synchronization measures (Tsyn and Bsyn) but no significant interaction with either of the two within-subjects factors. Significant mean differences were found in the upbeat phase of the quarter-note beat (QU), i.e., TsynQU, $F(1, 26) = 4.415$, $p < 0.05$, and BsynQU, $F(1, 26) = 13.725$, $p < 0.01$, and both of these differences showed that *the left-hand-first order yielded a significantly higher upbeat synchronization*. As shown by **table 4.7**, synchronization rates for the upbeat phase were higher by 3.5 to 8.3% when participants were exposed to the left hand conditions before those of the right hand.

[37] The same applies to possible interactions with between-subjects factors in the exploratory analyses (subsection 4.2.3.3).

However, this finding contrasts with the results obtained in Block 1, where the left-hand-first order resulted in significantly higher *downbeat* synchronization (see subsection 4.2.2.1, **table 4.2**). A possible explanation for this discrepancy is that the presence of determinate pitch might have interacted with order of presentation; given that the analysis included only unaccented layers, another possibility is that the previously observed effect of left-hand-first order on phase synchronization was not independent of accentuation. The influence of pitch structure will be explored in the across-blocks analysis (subsection 4.2.3.4), although the specific interaction with order of presentation cannot be tested within this design.

As in Block 1, paired-samples t tests were performed to assess the significance of mean differences between downbeat and upbeat synchronization (phase discrimination). **Table 4.8** presents the results for the beat synchronization measures (Bsyn), including those for the accented stimuli (S3b and S6b).[38] As can be observed, phase synchronization measures differ significantly only for the left hand conditions, where participants consistently tapped more downbeats than upbeats (mean differences of 10.7% to 15.7%, with the largest difference belonging to the accented condition, S3b). In contrast, within the right hand layer, there was no significant difference between participants' downbeat and upbeat phase synchronization. As in Block 1, this result suggests a faster tapping rate, partially syncopated period, or shifting phase. However, the implications of this synchronization pattern for the interpretation of the main effect of order in Block 2 are not clear, especially given that the left hand layers exhibited significantly more downbeat tapping. What might have induced participants in the left-hand-first condition to synchronize more often with the upbeat phase

[38] The results for taping synchronization (Tsyn) showed the same pattern of results and are not presented here.

overall? One possible explanation is that since the right hand layer begins with events moving at a generally faster rate, more participants who were exposed to the slower moving layer first (LH-RH) might have opted for a faster tapping rate when tapping to the denser layer than participants who began tapping to the right hand layers (RH-LH). Given that only four of the six stimuli were used in this preliminary analysis and that participants' tapping behavior was more variable in the left hand layers, it is also possible that the main effect of order is related to the tapping behavior of those specific participants in the left-hand-first order and is not representative of the population as a whole (a *Type I error*, i.e., rejecting the null hypothesis when it is true).

Table 4.8: Paired samples t-tests results for downbeat and upbeat beat synchronization (BsynQD and BsynQU) in Block 2 (S = stimulus; see **table 4.1** for individual details)

Paired Samples		**Mean Difference**	**Standard Deviation**	**Standard Error Mean**	**Significance** (2-tailed)
Pair 1	S3a.BsynQD - S3a.BsynQU	.115*	.208	.038	$t(29) = 3.032$, $p = .005$
Pair 2	S3b.BsynQD - S3b.BsynQU	.157*	.244	.045	$t(28) = 3.459$, $p = .002$
Pair 3	S3c.BsynQD - S3c.BsynQU	.107*	.186	.034	$t(29) = 3.148$, $p = .004$
Pair 4	S6a.BsynQD - S6a.BsynQU	-.009	.056	.011	$t(27) = -0.848$, $p = .404$
Pair 5	S6b.BsynQD – S6b.BsynQU	.007	.062	.011	$t(28) = 0.631$, $p = .533$
Pair 6	S6c.BsynQD – S6c.BsynQU	-.011	.091	.017	$t(29) = -0.684$, $p = .499$

*. The mean difference is significant at the .05 level.

4.2.3.2 Main Analysis: The Effect of Layer and Articulation

In Block 2, the semi-factorial design necessitated three sets of analysis with within-subjects factors. The first set followed the same basic design as the preliminary analysis above, i.e., it used only unaccented conditions and investigated the effects of layer (in tandem with pitch structure) and articulation; the other two sets (described in the text below) focused on a single layer and within-subjects factor, each of them including the data from only two conditions.

When accents were "off," layer had a significant main effect on all dependent variables, except for overall downbeat synchronization (BsynQD), where it only approached significance, $F(1, 27) = 3.936$, $p = .058$. Articulation and Layer × Articulation were not significant. **Table 4.9** presents a summary of the results for the effect of the Layer factor on all dependent variables. Tapping consistency measures (CV) were significantly higher for the right hand than for the left hand conditions, $F(1, 27) = 17.904$, $p < .001$, a result that is in line with the results obtained in Block 1. Thus, *participants' tapping to the left hand layer showed greater consistency that that of the right hand layer*. *Participants also exhibited a greater percentage of downbeat tapping in the left hand conditions*, TsynQD, $F(1, 27) = 8.234$, $p < .01$. Thus, when synchronizing to the left hand layer, 32.7% of participants' taps were on the downbeat as compared with 23.3% for the right hand layer. The discrepancy between tapping and beat synchronization measures for the downbeat phase suggests that although a significantly higher portion of taps in the left hand layers were synchronized with the downbeat phase, these taps represented a relatively lower number of actual beats in relation to the total number of available beats. Possible tapping strategies consistent with these results include tapping at a slower rate (e.g., a half note beat) or intermittent tapping. It

should also be noted that mean differences for the upbeat synchronization phase were generally greater and subject to less variance. Finally, the mean differences for upbeat synchronization were also statistically significant, i.e., TsynQU, $F(1, 27) = 34.725$, $p < .001$, and BsynQU, $F(1, 27) = 12.416$, $p < .01$. Thus, *upbeat phase synchronization was significantly higher when participants tapped to the right hand conditions* (mean differences = 10.5% and 7.3%).

Table 4.9: Effect of Layer on all dependent variables in Block 2 (LH =left hand layer and RH = right hand layer)

Measure	**Layer**	**Mean**	**Mean Difference** (LH - RH)	**Standard Error**	**Significance**[a] (*p* value)
CV	LH	.288	-.173*	.041	.000
	RH	.461			
TsynQD	LH	.327	.094*	.033	.008
	RH	.233			
BsynQD	LH	.214	.047	.024	.058
	RH	.167			
TsynQU	LH	.151	-.105*	.018	.000
	RH	.256			
BsynQU	LH	.108	-.073*	.021	.002
	RH	.181			

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

The second analysis of within-subjects factors involved assessing the effect of streaming in the left hand layer. As mentioned in the "Experimental Design" section (subsection 4.1.2.3), the left hand layer in this passage could be understood as two interleaved layers based on register and rhythmic behavior (see **example 3.22**). One possible

effect of streaming on participants tapping is lower rates of successful synchronization (Jones et al., 1995). For this particular layer, it was hypothesized that streaming would increase the effect of rhythmic irregularity. Thus, in an attempt to counteract this effect, two parameters (accents and contour) were manipulated and compounded into the Streaming (on/off) factor. In the preparation of the stimuli for this experiment, it was assumed that in this layer, accents contributed to rhythmic regularity and thus, in the Streaming "off" condition (S3b), accents were on and registers were collapsed so that all frequencies (pitches) fitted in an octave span. The association of downbeat accents and regularity is a widely spread practice in musical composition, performance, and analysis. There are also many empirical studies that support this assumption. In Block 1, the effect of accentuation was found to be significant for both downbeat and upbeat phases, and the specific effect of accents in the left hand layer was to increase the number of downbeat taps and decrease the number of upbeat taps. Given that all accents were on the downbeat phase, these results offer additional support for the assumption of downbeat accents and metric regularity. Thus, in the streaming "on" condition (S3a), accents were suppressed and registers were presented as written in the score. A series of paired-samples t tests was performed (one for each dependent measure); although the streaming "off" condition showed slightly higher downbeat and upbeat synchronization rates, the mean differences were not statistically significant. Thus, the null hypothesis could not be rejected and the results are inconclusive in terms of an effect of streaming on successful synchronization in the segregated left hand layer (see **table 4.10**).

Table 4.10: Paired-samples t tests results based on streaming in the left hand layer in Block 2 (S = stimulus; 3a = streaming "on" and 3b = streaming "off")

Paired Samples		**Mean Difference** (On - Off)	**Standard Deviation**	**Standard Error Mean**	**Significance** (2-tailed)
Pair 1	CV.S3a - CV.S3b	.040	.223	.041	$t(28) = 0.963$, $p = .344$
Pair 2	TsynQD.S3a - TsynQD.S3b	-.015	.120	.022	$t(28) = -0.679$, $p = .503$
Pair 3	BsynQD.S3a - BsynQD.S3b	-.045	.141	.026	$t(28) = -1.718$, $p = .097$
Pair 4	TsynQU.S3a - TsynQU.S3b	-.000	.103	.019	$t(28) = -0.018$, $p = .986$
Pair 5	BsynQU.S3a - BsynQU.S3b	-.008	.141	.026	$t(28) = -0.315$, $p = .755$

Table 4.11: Paired samples t-tests results based on accentuation in the right hand layer in Block 2 (S = stimulus; 6a = accent "off" and 6b = accents "on")

Paired Samples		**Mean Difference** (Off - On)	**Standard Deviation**	**Standard Error Mean**	**Significance** (2-tailed)
Pair 1	CV.S6a - CV.S6b	.040	.181	.034	$t(27) = 1.169$, $p = .253$
Pair 2	TsynQD.S6a - TsynQD.S6b	.016	.101	.019	$t(27) = 0.827$, $p = .415$
Pair 3	BsynQD.S6a - BsynQD.S6b	.007	.086	.016	$t(27) = 0.437$, $p = .665$
Pair 4	TsynQU.S6a - TsynQU.S6b	-.037	.120	.023	$t(27) = -1.635$, $p = .114$
Pair 5	BsynQU.S6a - BsynQU.S6b	-.011	.066	.013	$t(27) = -0.910$, $p = .371$

The last step in the main analysis of within-subjects factors involved assessing the effect of accentuation in the right hand layer. As mentioned previously, these accents were

added by the analyst to enhance the perceptibility of one possible interpretation of the metric structure in that layer.[39] However, the possibility of a discrepancy between analyzed and perceived structure was already presented as one possible explanation for the trend toward significance for the Layer × Accent interaction for tapping consistency in Block 1 that showed increased variability in the accented right hand accents conditions (see subsection 4.2.2.2). In order to test the null hypothesis for the Accent factor, another series of paired-samples t tests was performed, controlling for Layer (right hand) and Articulation (legato). **Table 4.11** presents the results. Overall, tapping variability and downbeat synchronization were slightly higher in the unaccented condition (and upbeat synchronization was slightly higher in the accented condition), but these differences were not significant and the null hypothesis could not be rejected.

To summarize: The design of Blocks 2 allowed for the assessment of the relative mensural determinacy of each layers with determinate pitch as well as for the exploration of the influence of three musical parameters on participants' tapping behavior within certain limits, i.e., articulation and layer in the unaccented conditions, streaming in the left hand layer with staccato articulation, and accentuation in the right hand layer with legato articulation. The first analysis controlling for accents found that *layer was the only factor that had a significant main effect, and this effect was significant for most of the dependent variables, including tapping consistency and synchronization to both downbeat and upbeat phases*, although the mean difference in overall downbeat synchronization (BsynQD) only approached significance. The resulting pattern for the effect of layer was consistent with the

[39] As discussed in chapter 3, subsection 3.4.3, this interpretation is based on an underlying temporal pattern characterized by a series of short-long durations subdivided into isochronous values giving rise to tempo fluctuations at the surface.

results from Block 1, i.e., there was a *higher downbeat synchronization in the left hand conditions* and *higher upbeat synchronization in the right hand conditions*, with the added feature that mean differences in upbeat synchronization were now significant. These results suggest that *the phase synchronization aspect of participants' tapping performance was not only consistent across these two blocks but that the observed tendencies in Block 1 were strengthened.* One possible explanation is that participants' benefited from repeating the tapping tasks (a learning effect); another possible explanation is that the presence of determinate pitch enhanced the effect of layer on adopted tapping strategy. The second possibility will be explored further in the comparative analysis of these two blocks that will follow the presentation of selected results from the exploratory analyses. The null hypothesis could not be rejected for articulation as well as for streaming in the left hand layer and accentuation in the right hand layer; thus, the results are inconclusive as to the possible effect of these musical parameters on participants' interpretation of the temporal structure of the two segregated layers of this texture when presented in conjunction with pitch structure (i.e., piano timbre or determinate pitch).

4.2.3.3 Exploratory Analyses: Selected Between-Subjects Factors

In the main analysis of Block 2, we have seen that only the Layer factor achieved statistical significance. However, since the stimuli in Block 2 use determinate pitch (piano timbre), factors such as main instrument and piano experience might play a role in participants' tapping behavior, in addition to general musical training and experience (e.g., musical degree, specialization, etc.). Exploratory analyses with between-subjects factors revealed that in Block 2, musical degree (Degree.2grp), area of specialization

(Specialization.4grp), main instrument (Instrument), and participants' gender (Gender) were all significant factors for at least one dependent variable. Again, because of the possibility of a Type I error (or sampling error), these results should be understood as suggestive rather than conclusive.

The first series of exploratory analyses with between-subjects factors was conducted with Layer and Articulation as within-subjects factors, controlling for Accent (off). When accents were off, musical degree (No degree or Bachelor's = 10, Master's or Ph.D. = 18) yielded a significant main effect on tapping consistency (CV), $F(1, 26) = 6.094$, $p < .05$, and overall downbeat (BsynQD) as well as upbeat (BsynQU) synchronization, i.e., $F(1, 26) = 4.354$, $p < .05$, and $F(1, 26) = 9.943$, $p < .01$, respectively. Mean differences for tapping synchronization were not significant and musical degree (Degree.2grp) did not interact significantly with within-subjects factors. A summary of the results is presented in **table 4.12**. As could be expected, *participants with a more advanced musical degree were significantly more consistent in their tapping and achieved higher synchronization rates to both the downbeat and upbeat phases of the quarter-note beat*. Thus, while participants with a master's or Ph.D. in music tapped an average of 22.2% and 17.2% of all available pulsations in the downbeat and upbeat phases, respectively, those with either no degree or a bachelor's tapped only 13.5% and 9.5% for the same measures. In addition, there was also consistently less variance among participants with a higher degree. The discrepancy between tapping and beat synchronization, which in the case of downbeat tapping synchronization (TsynQD) yielded a higher, although not significantly so, tapping synchronization rate for participants with less musical training (1 = 29.4% and 2 = 27.2%) could be explained by participants tapping at a slower rate than the quarter-note beat or intermittently. Given that

participants were instructed to tap "the underlying pulse or beat," which involved not tapping at all if they thought there was no underlying beat structure, intermittent tapping could mean that there were segments where participants experienced some segments as mensurally indeterminate.

Table 4.12: Estimated marginal means and mean differences for all dependent variables based on musical degree (Degree.2grp) in Block 2 (1 = No degree or Bachelor's and 2 = Master's or Ph.D.)

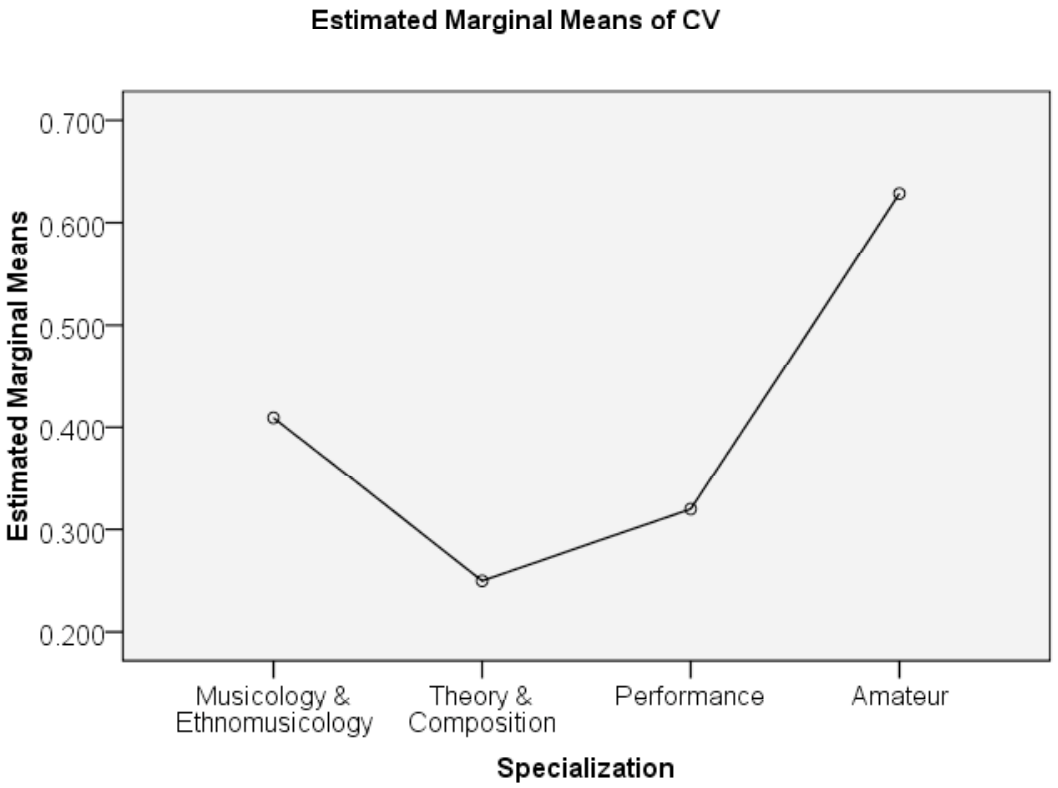
Measure	**Degree** (2 groups)	**Mean**	**Standard Error**	**Mean Difference** (1 - 2)	**Standard Error**	**Significance**[a] (*p* value)
CV	1	.492	.059	.183*	.074	.020
	2	.309	.044			
TsynQD	1	.294	.027	.021	.034	.534
	2	.272	.020			
BsynQD	1	.135	.034	-.087*	.042	.047
	2	.222	.025			
TsynQU	1	.187	.015	-.026	.018	.172
	2	.213	.011			
BsynQU	1	.095	.020	-.077*	.025	.004
	2	.172	.015			

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

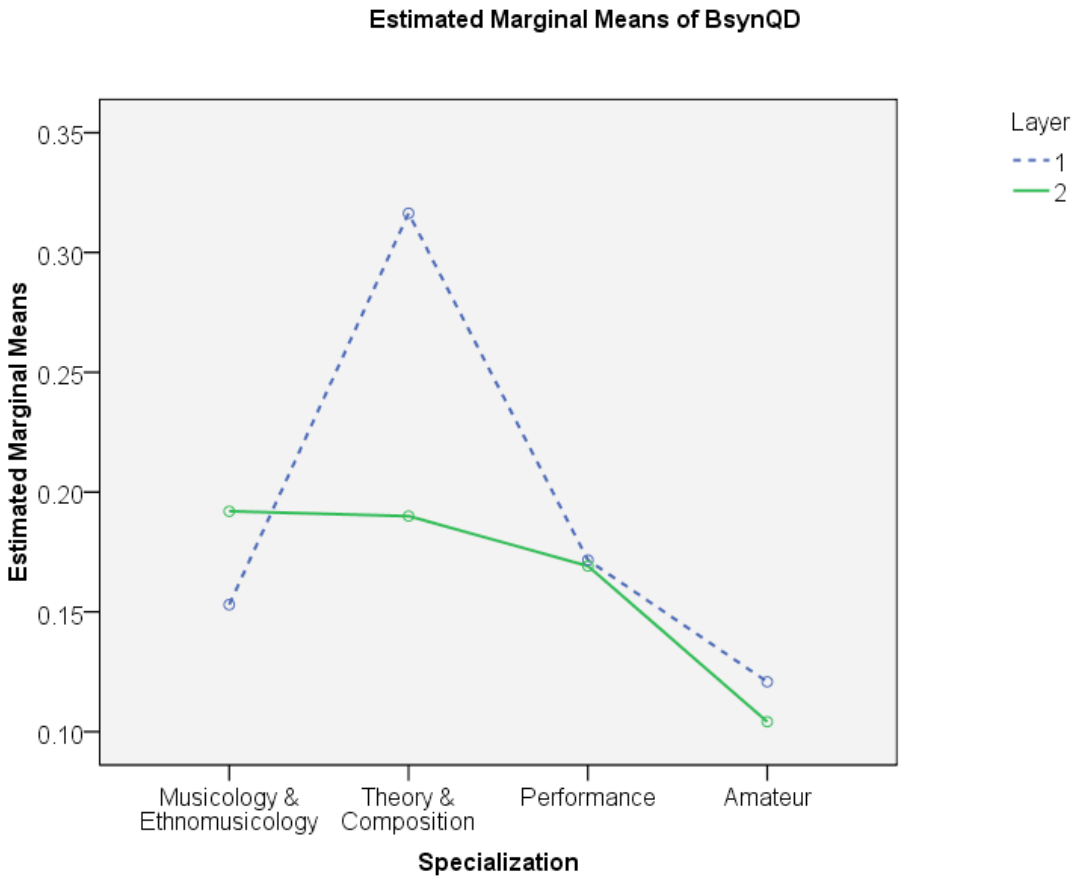
Figure 4.10: Main effect of area of specialization (Specialization.4grp) on tapping consistency (CV) in Block 2



Area of specialization (Musicology & Ethnomusicology = 5, Theory & Composition = 11, Performance = 6, Amateur = 6) also appears to have played a significant role on tapping consistency and there was an interesting significant interaction with the Layer factor involving participants who specialized in theory and composition. **Figure 4.10** presents a graphic representation of the estimated means for tapping consistency (CV) for each group. Post-hoc tests show that apart from the significant difference between Theory & Composition and Amateur, which was already observed in Block 1, the difference between Performance and Amateur also achieved significance, $F(3, 24) = 8.427$, $p = .001$. In addition, there was a significant interaction of Layer × Specialization.4grp for overall downbeat synchronization

(BsynQD), $F(3, 24) = 3.333$, $p = .036$. A graph of this interaction is presented in **figure 4.11** along with the data for BsynQD. Inspection of the graph suggests that while most participants' synchronization with the downbeat remained largely unchanged across layers, *theorists and composers synchronized significantly more often with downbeat of the unaccented left hand layer than other participants*. This result is consistent with the fact that musicians specializing in theory and composition are much more likely to engage the type of materials used in this experiment, and suggests that this experience allows them to be more sensitive to the rhythmic organizational cues given by the composer even when accents are "off" (as in this particular series of tests), which for the left hand conditions included here consisted in chronometrically regular whole-note pulsations in the lower register.

Figure 4.11: Layer × Specialization.4grp interaction for overall downbeat synchronization (BsynQD) in Block 2 (Layer 1 = LH [left hand] and Layer 2 = RH [right hand])

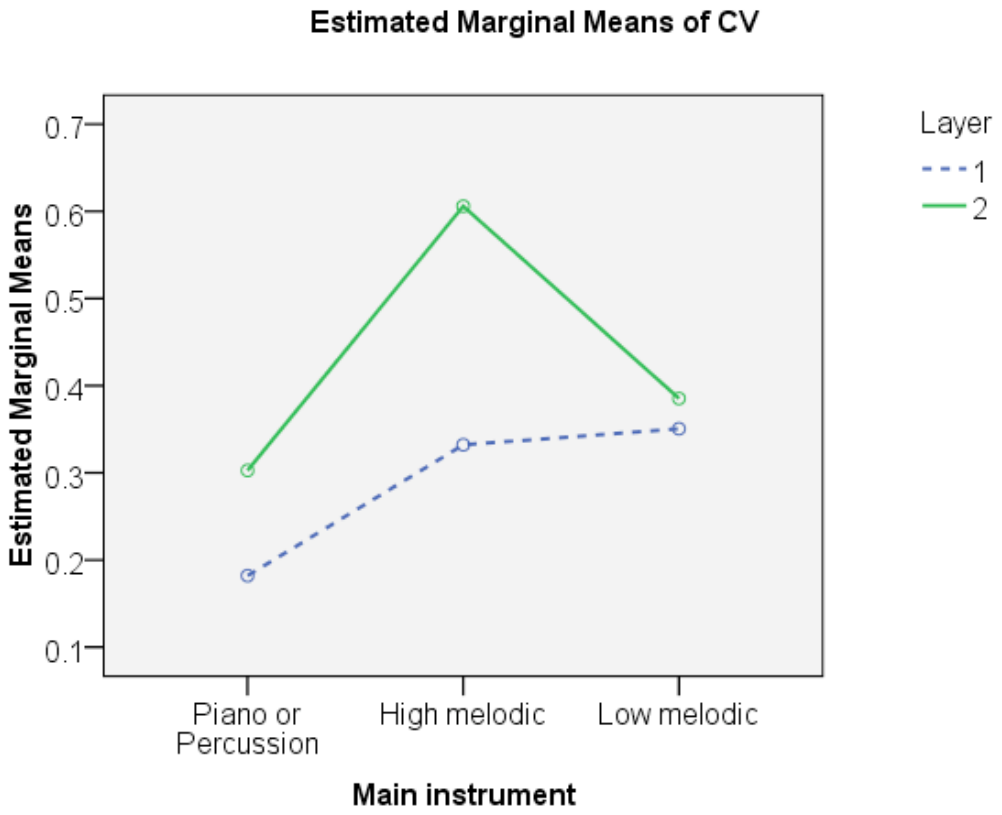


Measure	Specialization	Layer	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
BsynQD	Musicology & Ethnomusicology	LH	.153	.064	.022	.284
		RH	.192	.038	.113	.271
	Theory & Composition	LH	.316	.043	.228	.405
		RH	.190	.026	.136	.244
	Performance	LH	.172	.058	.052	.292
		RH	.169	.035	.097	.242
	Amateur	LH	.121	.058	.001	.241
		RH	.104	.035	.032	.177

As expected, there was a significant effect involving participants' instrumental training (Piano or Percussion = 9, High melodic = 13, Low melodic = 9). Main instrument had a significant main effect on tapping consistency (CV), with participants for whom piano or percussion was the main instrument showing greater consistency (mean difference = 0.227, SE = 0.080) than those who played a melodic instrument with a high register (e.g., soprano voice, flute, clarinet, etc.), $F(2, 25) = 3.988$, $p = .031$. These two groups also exhibited a trend toward significant difference for the overall upbeat synchronization measures (BsynQU), with pianists and percussionists tapping a greater number of upbeats (mean difference = 7.28%), $F(2, 25) = 3.224$, $p = .057$. These results suggest that *when accents were "off," piano and percussion players were more confident in their tapping behavior and/or less subject to shifts and interruptions*; the trend toward significance for overall upbeat synchronization also suggests that one possible factor in this enhanced tapping consistency might have been related to a subdivision strategy, although the current measures (i.e., phase measures) do not allow us to make that determination in any conclusive way. Main instrument also approached significance in its interaction with layer when considering the tapping consistency measure (CV), $F(2, 25) = 3.396$, $p = .05$. **Figure 4.12** presents a graphic representation of this interaction with estimated marginal means below. In the main analysis of within-subjects factors, we saw that layer has a significant main effect on tapping consistency, i.e., participants' tapping behavior was more stable when tapping to the left hand layers. The results presented here suggest that the effect of layer on tapping consistency (CV) is not independent of main instrument, with participants who play a high melodic instrument tapping less regularly to the right hand layer. Given this layer exhibits an overall higher register, this result suggests that *participants who play a high melodic instrument*

might have been more sensitive to the surface tempo fluctuations, and more specifically, that the surface fluctuations were interpreted in a metrical way leading these participants to adjust their tapping to a changing metrical interpretation.

Figure 4.12: Layer × Instrument.3grp interaction for tapping consistency (CV) in Block 2 (Layer 1 = LH [left hand] and Layer 2 = RH [right hand])

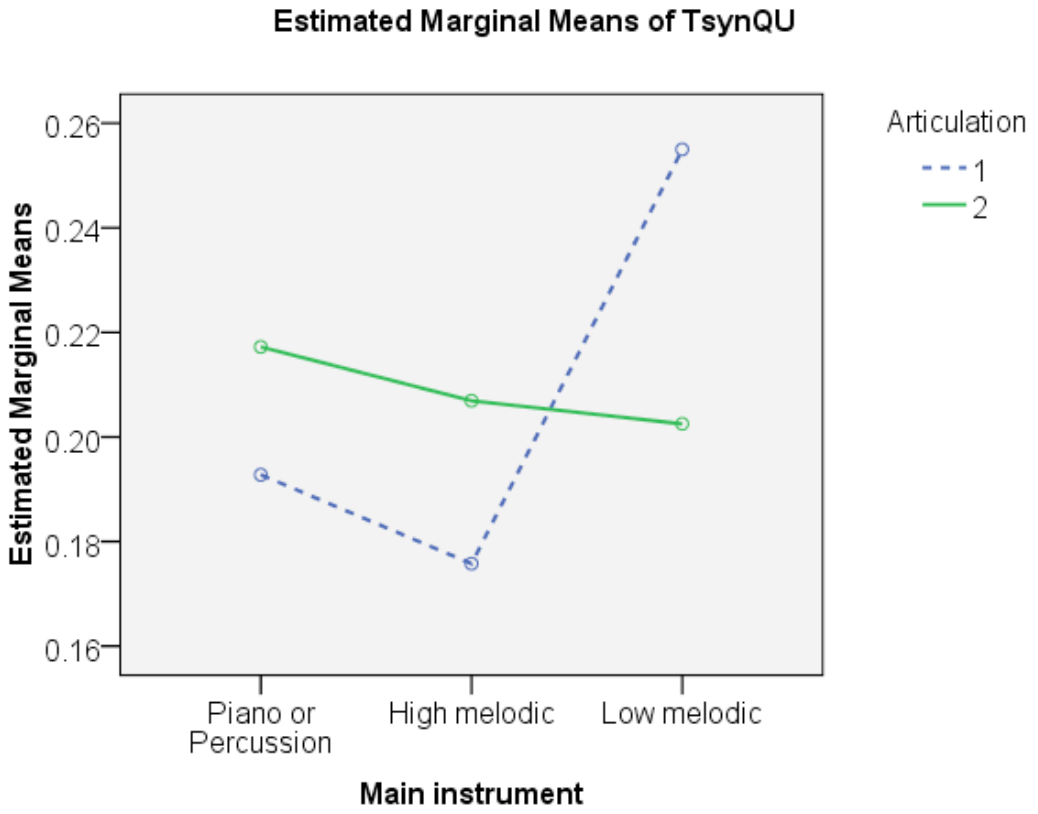


Measure	Main instrument	Layer	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
CV	Piano or Percussion	LH	.182	.063	.053	.311
		RH	.303	.077	.145	.461
	High melodic	LH	.332	.052	.225	.440
		RH	.606	.064	.474	.738
	Low melodic	LH	.351	.077	.192	.509
		RH	.385	.094	.192	.579

Furthermore, although articulation was not a significant factor in the analysis of within-subjects factors, there was also a trend toward significance in the interaction of Articulation × Instrument for upbeat tapping synchronization (TsynQU), $F(2, 25) = 3.177$, $p = .059$. On the graph (see **figure 4.13**), upbeat tapping synchronization patterns are different based on participants' main instrument. Thus, while upbeat tapping synchronization does not differ across instrument groups in the legato conditions, *participants who play a low melodic instrument (e.g., bass or tenor voice, cello or double bass, etc.) tapped a significantly higher number of upbeats in the staccato conditions than those of the two other instrument groups*.[40] Given that this test averages downbeat and upbeat synchronization rates across layers, it is not possible to determine whether these participants were more likely tapping at a faster rate (e.g., eighth-note beats), as the lack of a significant difference between these two phases for the right hand layers suggests, or tapping in a syncopated manner; the latter would be consistent with the preliminary finding that there was a significantly higher number of taps in the downbeat phase of the left hand layers (subsection 4.2.3.1). Either way, this result points to a particular way of conceptualizing the beat that is linked with specialized training and experience, i.e., playing a low melodic instrument.

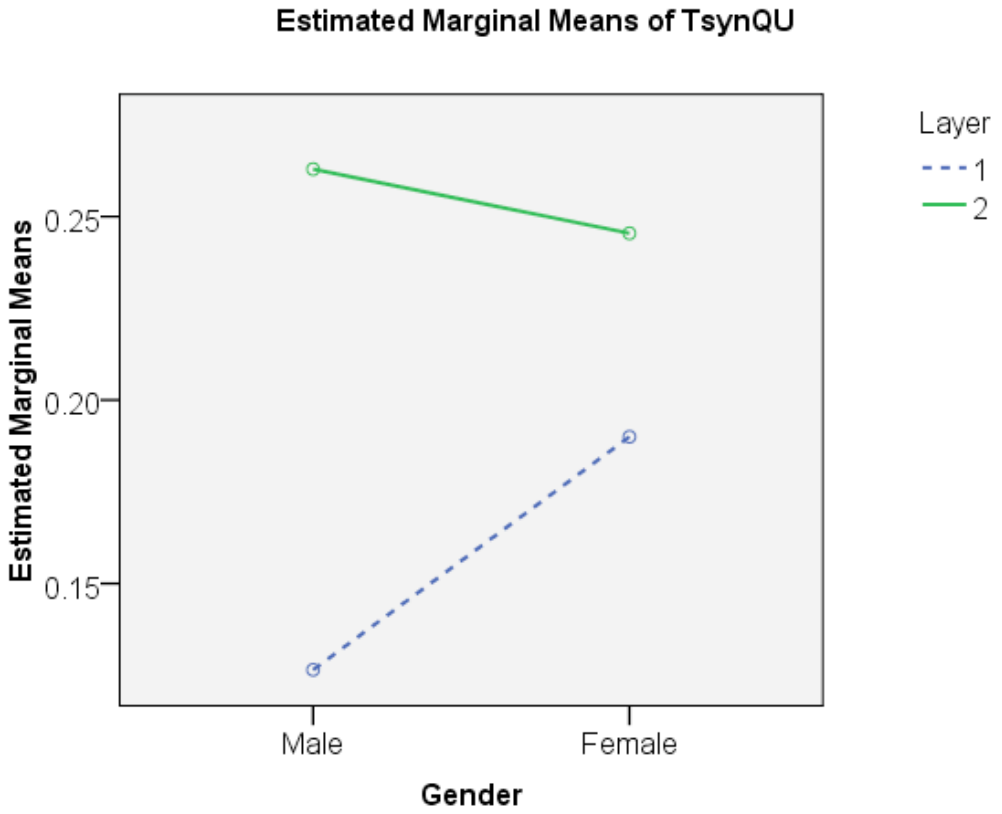
[40] It should be kept in mind that this finding is valid only for the unaccented layers, as the specific design of this block required the Accent factor to be controlled, i.e., including only the conditions with accents "off".

Figure 4.13: Articulation × Instrument.3grp interaction for upbeat tapping synchronization (TsynQU) in Block 2 (Articulation 1 = staccato and Articulation 2 = legato)



Measure	Main Instrument	Articulation	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
TsynQU	Piano or Percussion	Staccato	.193	.017	.157	.228
		Legato	.217	.021	.173	.261
	High melodic	Staccato	.176	.014	.146	.205
		Legato	.207	.018	.170	.244
	Low melodic	Staccato	.255	.021	.212	.298
		Legato	.203	.026	.148	.257

Figure 4.14: Layer × Gender interaction for upbeat tapping synchronization (TsynQU) in Block 2 (Layer 1 = LH [left hand] and Layer 2 = RH [right hand])

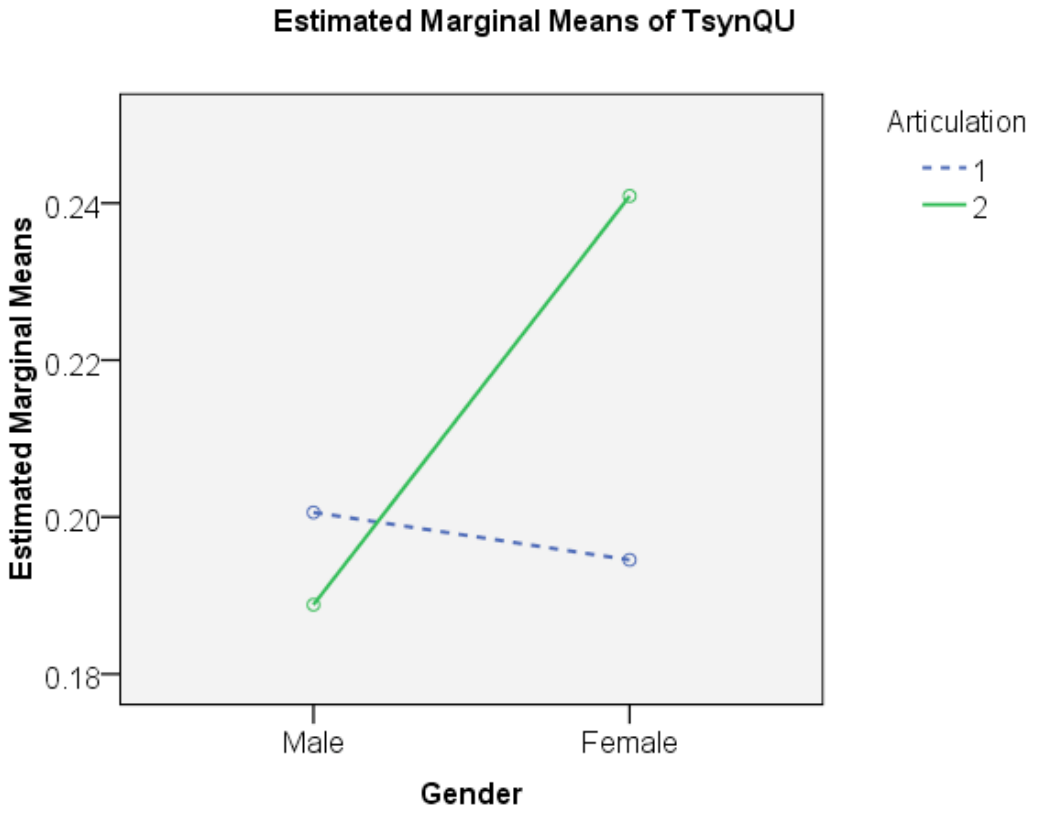


Measure	Gender	Layer	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
TsynQU	Male	LH	.126	.020	.086	.167
		RH	.263	.009	.244	.282
	Female	LH	.190	.024	.140	.240
		RH	.245	.012	.222	.269

Finally, participants' gender also turned out to be a potentially significant factor. First, women's tapping behavior appeared to be significantly more variable than men's, $F(1, 26) = 4.429$, $p = .045$. Second, gender interacted significantly with both layer and articulation for upbeat tapping synchronization (TsynQU). Thus, when accents were "off," women synchronized more often with the upbeat phase of the quarter-note beat in the left hand

conditions, $F(1, 26) = 5.858$, $p = .023$; synchronization rates for men and women did not differ significantly in the right hand conditions, where synchronization to the upbeat phase was generally more frequent (see **figure 4.14**). Similarly, women synchronized with the upbeat phase significantly more often than men in the legato conditions, but men and women showed similar synchronization rates in the staccato conditions, but (see **figure 4.15**), $F(1, 26) = 4.482$, $p = .044$). The same cautionary note regarding the possibility of a sampling error applies to these findings; nevertheless, the recurrence of an effect of gender on upbeat synchronization and the presence of potentially significant interactions with within-subjects factors linked to the musical structure call for an investigation of this factor.

Figure 4.15: Articulation × Gender interaction for upbeat tapping synchronization (TsynQU) in Block 2 (Articulation 1 = staccato and Articulation 2 = legato)

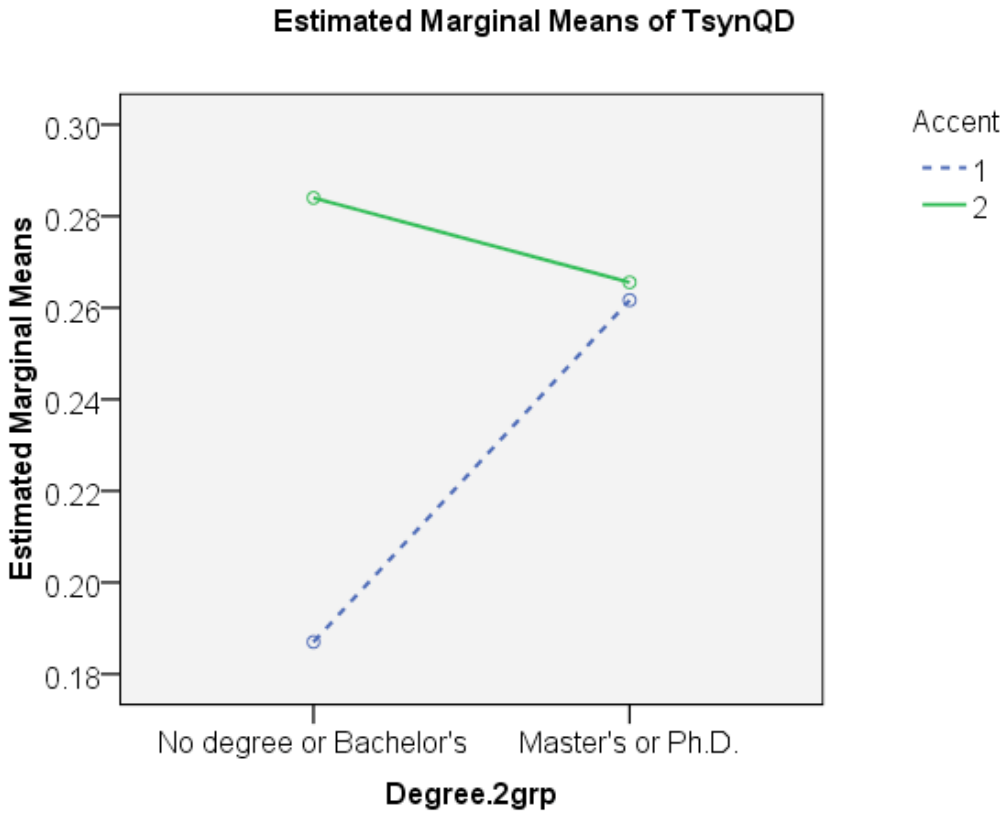


Measure	Gender	Articulation	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
TsynQU	Male	Staccato	.201	.014	.171	.230
		Legato	.189	.014	.160	.218
	Female	Staccato	.195	.018	.158	.231
		Legato	.241	.017	.205	.277

Up to this point, all between-factors subjects were analyzed with Layer and Articulation as within-subjects factors, controlling for Accent (off). Exploratory analyses were also conducted for with the second and third group of conditions, which involved streaming (left hand layer only) and accentuation (right hand layer only). Overall, between-subjects factors (e.g., Specialization, Instrument, and Gender) also appeared to have a main

effect on several of the measures, but there was no significant interaction with either the Streaming or Accent factor, except for a significant interaction of Accent × Degree.2grp in the right hand layer when controlling for Articulation (legato). When comparing the right hand conditions on the basis of accents, musical degree still had a significant main effect on beat synchronization with both the downbeat and upbeat phases, i.e., BsynQD, $F(1, 26) = 13.546$, $p = .001$, and BsynQU, $F(1, 26) = 12.112$, $p = .002$. In both cases, *participants with a higher musical degree synchronized with significantly more beats overall*; this result was consistent with the results from the previous analysis across layers with accents "off." Musical degree also interacted with accentuation for downbeat tapping synchronization. **Figure 4.16** provides a graphic representation of this interaction, with estimated marginal means for each group below. As shown by the graph, although the accented condition appears to have induced participants with less musical training into tapping significantly more often to the downbeat phase, (TsynQD), $F(1, 26) = 4.335$, $p = .047$, accents did not bring about a significant change in downbeat synchronization for participants who hold a higher degree. This result confirms the findings of the exploratory analyses for Block 1.

Figure 4.16: Accent × Degree.2grp interaction for downbeat tapping synchronization (TsynQD) in Block 2 (Accent 1 = "off" and Accent 2 = "on")

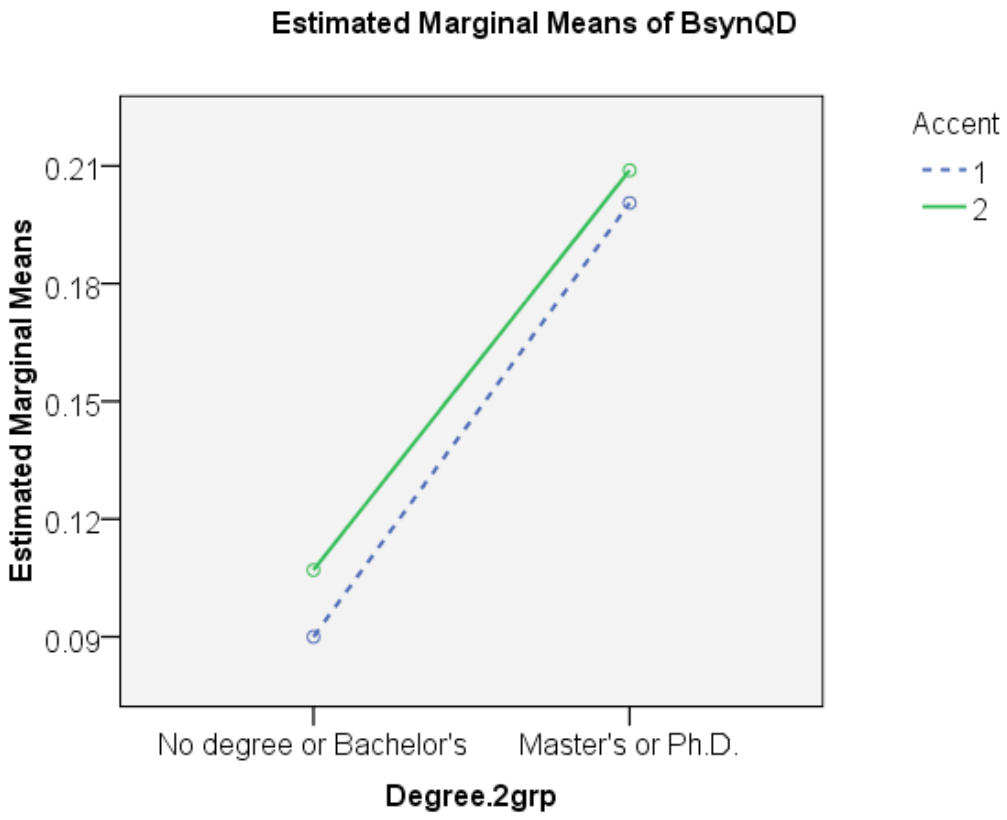


Measure	Degree (2 groups)	Accent	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
TsynQD	No degree or Bachelor's	Off	.187	.016	.155	.219
		On	.284	.029	.224	.344
	Master's or Ph.D.	Off	.262	.012	.238	.286
		On	.266	.022	.221	.310

The distinction between tapping and beat synchronization measures also comes to life in this specific interaction. The results for the corresponding interaction for overall downbeat synchronization measures (BsynQD) offer a meaningful contrast (see **figure 4.17**). Here, although participants with a master's or Ph.D. degree in music still synchronize with the downbeat phase for a significantly higher percentage of the total number of available beats

than participants with no musical degree or only a bachelor's degree in music, there is no difference in the overall downbeat synchronization based on accentuation, $F(1, 26) = 0.106$, $p = .748$. Thus, the effect of accents in interaction with musical degree is limited to synchronization measures based on the total number of taps (TsynQD). This contrast in synchronization measure suggests that *participants' with less formal training tapped only intermittently, resulting in a higher rate of downbeat tapping but a lower rate of overall downbeat synchronization*.

Figure 4.17: Accent × Degree.2grp interaction for overall tapping synchronization (BsynQD) in Block 2 (Accent 1 = "off" and Accent 2 = "on")



To summarize: The exploratory analyses conducted in Block 2 confirm that between-subjects factors beyond the general differentiation of non-musicians versus musicians should be considered when conducting beat-finding experiments with complex musical stimuli. Here, while only the temporal structure (in tandem with pitch structure) appeared to have a significant effect on participants' tapping consistency and synchronization, the inclusion of between-subjects factors revealed interactions with articulation and accentuation. In general, it was observed that *participants with a higher musical degree achieved higher rates of synchronization to both upbeat and downbeat*. In addition, interesting interactions between specialized training as represented by main instrument and area of specialization were also discovered. *Participants who played a high melodic instrument tapped in a less stable manner (higher variability) when synchronizing with the right hand (higher register) layers and participants who play a low melodic instrument synchronized significantly more often with the upbeat phase when the stimuli were performed with staccato articulation.* Theorists and composers also synchronized significantly more often with the downbeat of the left hand layer than participants who specialized in other areas. Finally, gender was also identified as a potentially significant between-subjects factor. The specific relationships outlined were similar to those found in Block 1: for example, women showed a tendency toward higher rates of tapping synchronization to the upbeat phase, while men did not.

4.2.3.4 Across-Blocks Analysis: Blocks 1&2 (Timbre)

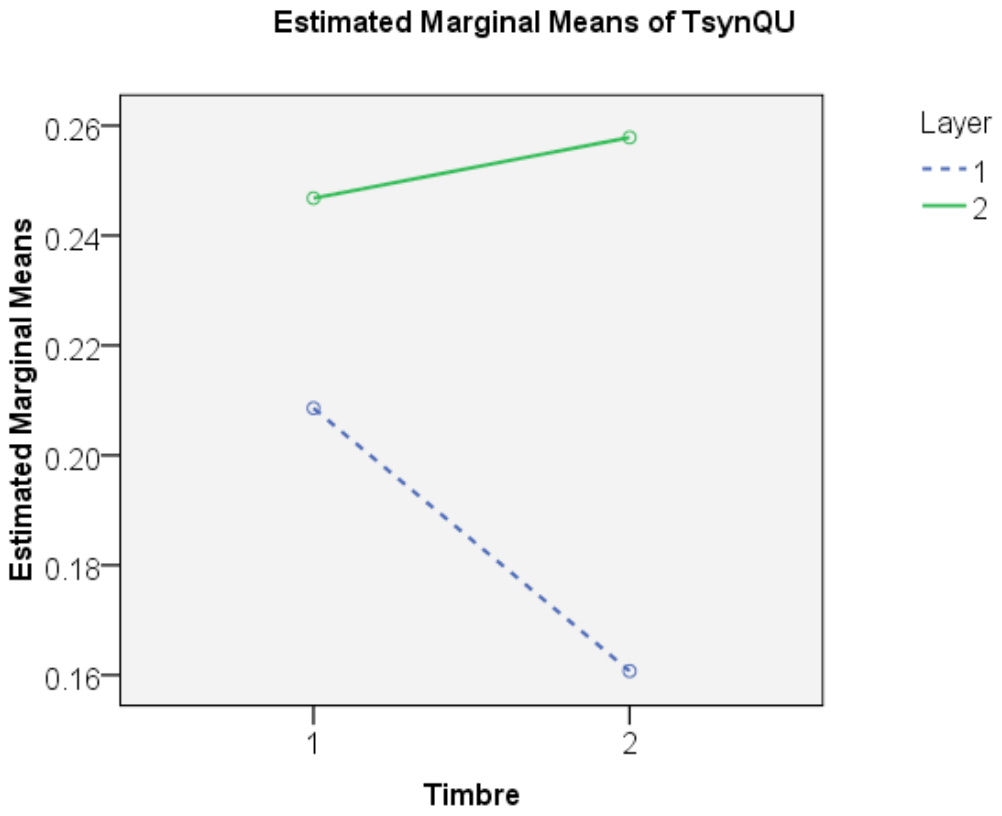
In the main analysis, we found that there was a significantly higher downbeat synchronization in the left hand layers as well as a significantly higher upbeat

synchronization in the right hand layers, which was consistent with Block 1. It was also suggested that the more significant upbeat synchronization mean difference of Block 2 might have been related to the use of determinate pitch. In order to determine whether there was an effect of determinate pitch in general (as represented by the timbre) on tapping consistency and synchronization, a series of univariate analyses of variance was conducted across Blocks 1 and 2. Because of the semi-factorial design of Block 2, two groups of tests were performed with Layer (LH/RH) and Timbre (Teak/Piano) as within-subjects factors, controlling for Accent (off) and Articulation (the first group used only the staccato conditions and the second group used only the legato conditions); accented conditions were not used. Because the purpose of these tests was to determine the effect of timbre, only the results for main effect and interactions involving the Timbre factor will be reported.

The two series of tests yielded very similar results, which is consistent with the finding that there was no significant main effect or interaction involving articulation in the within-subjects factorial analysis. Overall, timbre did not have a main effect but it interacted significantly with layer for at least one dependent variable. In the second group of tests with the unaccented layers with legato articulations (S1, S4, S3c, and S6a; see **table 4.1** for details), timbre interacted significantly with layer for the upbeat tapping synchronization (TsynQU), $F(1, 27) = 4.838$, $p = .037$. **Figure 4.18** provides a graphic representation of this interaction, including pairwise comparisons based on estimated marginal means. As shown by this graph, when participants tapped to the piano timbre conditions (Timbre 2), they synchronized significantly less with the upbeat than when they tapped to the teak condition (Timbre 2), but only when synchronizing to the left hand layer (Layer 1). When synchronizing with the right hand layer (Layer 2), there was an opposite trend, although the

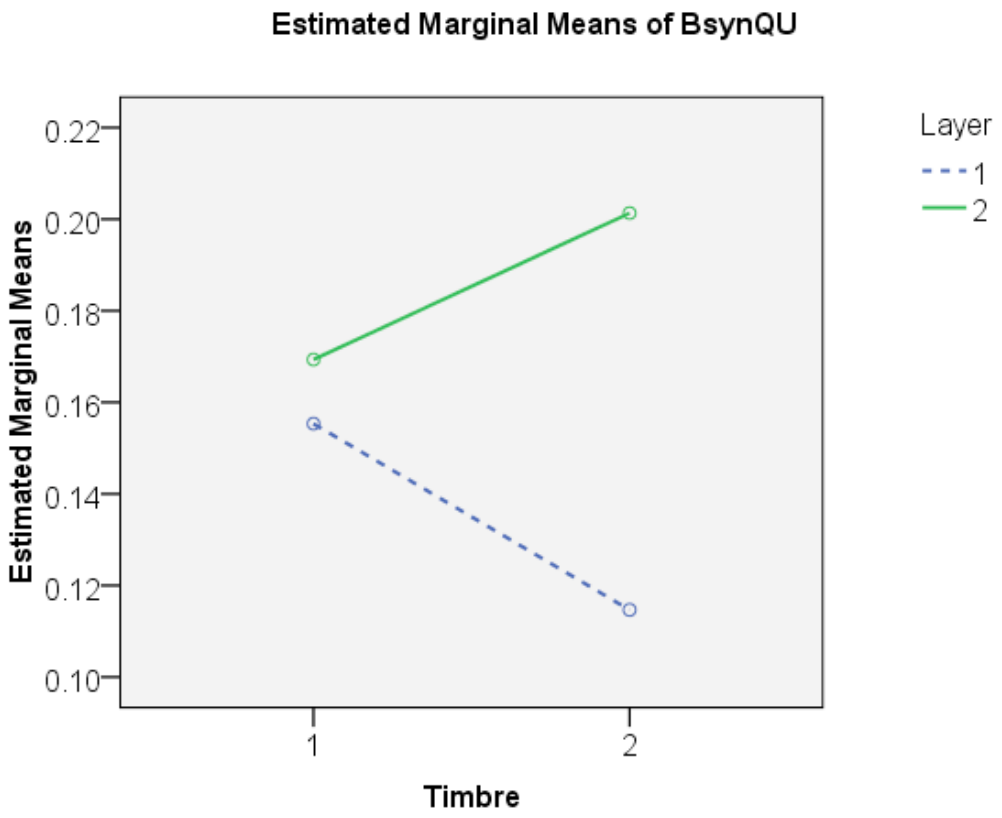
synchronization rates were much more similar across timbre conditions. The same interaction was observed in the first group of tests with accents off and staccato conditions for both tapping and overall upbeat synchronization (TsynQU and BsynQU), but these interactions did not reach statistical significance, i.e., $F(1, 29) = 4.125$, $p = .052$, and $F(1, 29) = 4.033$, $p = .054$. The pattern of responses was basically the same for tapping synchronization rates, but the results were more symmetrical for overall upbeat synchronization rates (see **figure 4.19**), with a more pronounced trend toward higher upbeat synchronization rates when participants tapped to the right hand layer in the piano timbre condition. The difference between participants tapping and overall upbeat synchronization rates (TsynQU and BsynQU), which results from measuring the number of synchronized taps over the total number of taps (Tsyn) and the total number of available beats (Bsyn), suggests that participants might have been tapping relatively less (intermittently) or slower when tapping to the right hand layer in the teak timbre condition.

Figure 4.18: Layer × Timbre interaction for upbeat tapping synchronization (TsynQU) across Blocks 1 and 2 (Layer 1 = left hand and Layer 2 = right hand; Timbre 1 = teak and Timbre 2 = Piano)



Measure	Layer	Timbre	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
TsynQU	LH	Teak	.209	.021	.166	.251
		Piano	.161	.020	.120	.201
	RH	Teak	.247	.018	.209	.285
		Piano	.258	.011	.236	.280

Figure 4.19: Layer × Timbre interaction for overall upbeat synchronization (BsynQU) across Blocks 1 and 2 (Layer 1 = left hand and Layer 2 = right hand; Timbre 1 = teak and Timbre 2 = Piano)



Measure	Layer	Timbre	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
BsynQU	LH	Teak	.155	.026	.103	.208
		Piano	.115	.023	.068	.161
	RH	Teak	.169	.019	.131	.208
		Piano	.201	.025	.150	.253

To summarize: Although the presence of determinate pitch (piano timbre) did not have a main effect on tapping consistency and synchronization, it influenced the way in which participants synchronized to at least one layer, that is, the left hand layer. More

specifically *when determinate pitch was used in the unaccented left hand conditions, participants synchronized less often with the upbeat phase when they tapped to the left hand layer*, regardless of articulation. The pattern of responses was similar for upbeat tapping synchronization across articulations conditions (staccato and legato), but the effect was more pronounced with the legato articulation. This finding suggests that *the longer duration of sound events might have had some influence on the tapping period for the unaccented left hand layer*, either because it helped participants to predict where the beats would fall or induced them into tapping more slowly.

4.2.4 Block 3

Block 3 presents the combined left hand and right hand layers in the Teak timbre condition (indeterminate pitch); the only other musical parameter that was altered was accentuation. Because accents could be manipulated in each hand independently of the other, this block was a 2 × 2 factorial design, with Laccent (left hand accents "off" or "on") and Raccent (right hand accents "off" or "on") as the two within-subjects factors. For combined layers, synchronization phase is designated by the labels "L" (left hand) and "R" (right hand) for the layer associated with the downbeat alignment for the quarter-note beat (Q); for example, TsynQL refers to tapping synchronization with the left hand downbeat and BsynQR refers to beat synchronization with the right hand downbeat.

This particular design allowed for the testing of the effect of layer-specific accentuation on participants' metric interpretation when layers were combined. Given that pitch was indeterminate, the potential for streaming was also very low. Thus, these conditions promoted global attention and offered a raw version of the overall temporal

structure. The use of four different combinations of accents also allowed for the testing of the flexibility of participants' metric interpretation. Finally, the analysis across Blocks 1 and 3 made it possible to measure how each layer's relative degree of mensural determinacy changed when layers where combined.

4.2.4.1 Preliminary Analysis: Order of Presentation and Phase Discrimination

All participants first listened and tapped to the combined layers stimulus with both left hand and right hand accents "off" (S7.LoffRoff). Given the possibility of a significant main effect linked to order of presentation, the order of the remaining stimuli (S8.LonRon, S10a.LonRoff, and S10b.LoffRon) was counterbalanced so that all possible orders were used for an equal number of times, for a total of six possible orderings (1 through 6), with five cases each (see **table 4.13**). It was assumed that this procedure would account for the confounding effect of an interaction of Order.Block3 with within-subjects factors. Because of the large number of cells (Laccent × Raccent × Order.Block3 = 2 × 2 × 6 = 24) and small number of cases for each group in this analysis, there is an increased chance of statistical error, especially since the small number of cases per order group also increases the influence of individual participants' performance (and potential outliers) on group means and standard variation. Thus, it was not practical to test the original assumption as to a significant main effect of order (Order.Block3).

Table 4.13: Order of presentation in Block 3 (Legend for stimuli descriptors: S = stimulus; L = left hand layer and R = right hand layer; "off"/"on" applies to accents)[41]

Order	Stimuli	Accent Conditions	Number of Cases
1	S7, S8, S10a, S10b	LRoff, LRon, LonRoff, LoffRon	5
2	S7, S8, S10b, S10a	LRoff, LRon, LoffRon, LonRoff	5
3	S7, S10b, S8, S10a	LRoff, LoffRon, LRon, LonRoff	5
4	S7, S10a, S8, S10b	LRoff, LonRoff, LRon, LoffRon	4
5	S7, S10a, S10b, S8	LRoff, LonRoff, LoffRon, LRon	5
6	S7, S10b, S10a, S8	LRoff, LoffRon, LonRoff, LRon	5

When layers are combined, each layer provides a certain level of interference based on streaming and accentuation (see **table 4.14**). Streaming was predicted to influence phase synchronization patterns within stimuli, i.e., whether participants would synchronize more often to one phase over another (moderate streaming) or would not show a preference of one phase over the other (low streaming). Thus, low streaming here corresponds to a tendency toward stream fusion and moderate streaming corresponds to a tendency toward stream segregation. Given that the pitch content has been suppressed (Teak timbre condition), there is no high streaming category. When accents are "off" in both layers, streaming is possible only on the basis of the panning offered by the stereophonically performed stimuli (low streaming); a similar situation arises when combining two accented layers. The presence of accents was also thought to influence which layer participants would be most likely to synchronize with. Given that all accents fall on the downbeat of their respective layer, it was predicted that participants would synchronize more often with the accented layer than with the unaccented one. Finally, from the perspective of a given layer, the interaction of accentuation and streaming was theorized to result in a particular level of interference. From

[41] Since there were only 29 valid responses saved by the computer, the final number of responses for order group 4 was four.

these ratings, predictions could be made as to the probable result of combining two specific accent conditions. For example, given a condition where the main accents are "on" and the competing accents are "off," which corresponds to a condition with moderate streaming possibility in favor of the main layer, the influence of the competing layer is minimized (low interference level). On the other hand, if both accents are "off" or "on" (low streaming), the influence of the competing layer will be greater (moderate interference level). Finally, the highest level of interference would result from no accentuation in the main layer and accents "on" in the competing layer, where stream segregation would be working against the main layer. Because interference level is determined from the perspective of a given layer, these predictions will be tested in the analyses across Blocks 1 and 3 in subsection 4.2.4.4.

Table 4.14: Interference levels based on accentuation and streaming

Main Accents	**Off**	**On**	**On**	**Off**
Competing Accents	Off	Off	On	On
Streaming	Low	Moderate	Low	Moderate
Interference Level	Moderate	Low	Moderate	High

A series of paired-samples t tests was performed to evaluate the differences in left hand and right hand downbeat synchronization (the combined layers equivalent to downbeat and upbeat phase synchronization for the segregated layers). **Table 4.15** presents the results of this analysis for the tapping synchronization measures (TsynQL and TsynQR).[42] As shown by this table, synchronization with the left hand and right hand downbeats differed significantly only for the two stimuli with right hand accents "on" (S8.LRon and S10b.LoffRon), where participants synchronized significantly more often with the right hand

[42] The results for beat synchronization measures (BsynQL and BsynQR) were very similar so only tapping synchronization results are presented as representative of both types of synchronization.

downbeat, with mean differences of 28.7% and 18.4%, respectively. In addition, there was a trend for participants to synchronize more often with the left hand downbeat when only the left hand accents were "on" (S10a). Finally, when both layers were presented without accents (S7), there was no significant difference between left hand and right hand downbeat synchronization rates and the comparatively low standard deviation (9.3%) suggests that there was a strong agreement in this respect between participants. As proposed previously, possible explanations for a lack of significant difference between two phases include a faster "consonant" rate, such as that represented by an in-phase eighth note beat, and phase shifting, which can include intermittent as well as syncopated tapping. Unlike in the first two blocks where layers were presented separately, the emergence of tapping patterns representative of an eighth note beat in combined layers suggests an integrated perception of the musical surface. Thus, these results confirm the assumption that the simultaneous presentation of the two layers without accents (and without determinate pitch) favors stream fusion (low streaming). Furthermore, the mean differences when left hand and right hand accents were "on" provide support for the assumption that the presence of accents in one of the layers would contribute to stream segregation in favor of the accented stream. However, given that statistical significance was only reached when the right hand accents were "on" and given that the synchronization to the right hand downbeat was also significantly higher when both layers' accents were "on," it would appear that right hand accents had more influence on participants' interpretation of the underlying beat. This issue will be further clarified by the main analysis in the next subsection.

Table 4.15: Paired-samples t tests results for tapping synchronization to left hand and right hand downbeat (TsynQL and TsynQR) in Block 3

Paired Samples		Mean Difference (L - R)	Standard Deviation	Standard Error Mean	Significance (2-tailed)
Pair 1	TsynQL.S7.LRoff - TsynQR.S7.LRoff	.019	.093	.017	$t(29) = 1.120$, $p = .272$
Pair 2	TsynQL.S8.LRon - TsynQR.S8.LRon	-.287*	.247	.045	$t(29) = -6.364$, $p < .001$
Pair 3	TsynQL.S10a.LonRoff - TsynQR.S10a.LonRoff	.095	.252	.047	$t(28) = 2.037$, $p = .051$
Pair 4	TsynQL.S10b.LoffRon - TsynQR.S10b.LoffRon	-.184*	.316	.059	$t(28) = -3.142$, $p = .004$

*. The mean difference is significant at the .05 level.

4.2.4.2 Main Analysis: Layer-Specific Accentuation (Laccent and Raccent)

In order to determine the effect of layer-specific accentuation, a series of univariate analyses of variance with Laccent and Raccent as within-subjects factors was conducted. Overall, these analyses showed a significant main effect only for Raccent, but the interaction of Laccent × Raccent was significant. Although there was no significant mean difference for tapping consistency (CV), each tapping and beat synchronization measure exhibited at least one statistically significant mean difference. First, I will review the data for the two within-subjects factors (Laccent and Raccent) and their main effect. Then, in order to make the interpretation of the results more meaningful, I will report the results for significant interactions by dependent measure.

Table 4.16: Estimated means and mean differences based on left hand accents (Laccent)

Measure	Laccent	Mean	Standard Error	Mean Difference (Off - On)	Standard Error	Significance[a] (*p* value)
CV	Off	.371	.049	.009	.033	.792
	On	.362	.054			
TsynQL	Off	.227	.017	.003	.021	.901
	On	.224	.016			
BsynQL	Off	.231	.026	.027	.023	.255
	On	.204	.025			
TsynQR	Off	.308	.016	-.007	.026	.801
	On	.315	.020			
BsynQR	Off	.298	.023	.011	.029	.719
	On	.287	.029			

a. Adjustment for multiple comparisons: Bonferroni.

Tables 4.16 and **4.17** present the estimated means and mean differences based on layer accentuation (Laccent and Raccent). As opposed to tapping consistency results for segregated layers (Blocks 1 and 2), the means for the coefficient of variability measure (CV) are very close across accent conditions, i.e., within 0.362-0.371 for both left hand and right hand accents conditions). This suggests that when layers are combined, participants' tapping behavior is more consistent across stimuli. The data based on left hand accentuation (Laccent) also show mostly similar synchronization rates across measures (e.g., comparing TsynQL and BsynQL) and no significant mean difference within each measure (left hand accents "off"/"on"). Synchronization rates fall within less than 12% of each other, i.e., from 20.4% for beat synchronization to the left hand downbeat (BsynQL) with left hand accents "on," to 31.5% for tapping synchronization to the right hand downbeat (TsynQR) with left hand accents "on" (see **table 4.16**). Thus, it would appear that not only was there no

significant difference in left hand downbeat synchronization rates based on left hand accentuation (e.g., comparing TsynQL with accents "off" and "on"), but under some circumstances left hand accents actually contributed to increase participants' synchronization with the right hand downbeat (e.g., comparing TsynQR with TsynQL). However, as it is not possible to distinguish between tapping periods, this result could also correspond to an integrated or partly syncopated form of tapping (e.g., an eighth note or dotted-quarter note beat) in which participant's would have synchronized more often with the right hand quarter downbeat.

When comparing the means of synchronization values from the perspective of the right hand layer (**table 4.17**), the range of variation in synchronization rates based on downbeat alignment and accentuation is much greater, i.e., from 18.5% for tapping synchronization to the left hand downbeat (TsynQL) with right hand accents "on," to 41.7% for tapping synchronization to the right hand downbeat (TsynQR) with right hand accents "on." Three of the four mean differences for synchronization measures based on right hand accentuation (Raccent) are also statistically significant. Raccent was significant for the tapping synchronization to the left hand downbeat (TsynQL), $F(1, 28) = 6.442$, $p < .05$, but not for the beat synchronization to the corresponding phase (BsynQL). In contrast, Raccent was significant for both tapping and beat synchronization to the right hand downbeat, i.e., TsynQR, $F(1, 28) = 47.977$, $p < .001$, and BsynQR, $F(1, 28) = 54.934$, $p < .001$. These significant mean differences indicate that participants synchronized considerably more often with the right hand downbeat when right hand accents were "on," i.e., 21.1% and 21.5% for tapping and beat synchronization measures, respectively. The mean difference in tapping synchronization to the left hand downbeat (TsynQL) also suggests that when right hand

accents are "off," the left hand events interfere significantly with the right hand events, which induced participants' to synchronize more often with the left hand downbeat (mean difference of 7.8%).

Table 4.17: Estimated means and mean differences based on right hand accents (Raccent)

Measure	Raccent	Mean	Standard Error	Mean Difference (Off - On)	Standard Error	Significance [a] (*p* value)
CV	Off	.368	.056	.002	.053	.967
	On	.366	.055			
TsynQL	Off	.265	.015	.078*	.031	.017
	On	.187	.025			
BsynQL	Off	.230	.019	.025	.041	.547
	On	.205	.039			
TsynQR	Off	.206	.012	-.211*	.030	.000
	On	.417	.025			
BsynQR	Off	.185	.020	-.215*	.029	.000
	On	.400	.031			

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

Thus, it would appear that *overall, greater synchronization was achieved with the right hand downbeat regardless of accentuation*, although the percentage difference for synchronization rates for the right hand accents "on" as compared with left hand accents "on" was about 20%. In addition, when considering the results from the analysis of the first two blocks, the influence of the right hand accents on metrical interpretation as represented by participants' tapping performance in Block 3 is especially suggestive of the effect of combining layers on the mensural determinacy of each layer. *While the left hand layer*

showed greater mensural determinacy than the right layer when layers were presented separately, i.e., participants synchronized more successfully with the left hand layer than the right hand layer, the superposition of layers not only improved the right hand layer's mensural determinacy significantly, but under some specific accentual conditions, participants synchronized more often with the right hand layer. This is an important point with particular implications for the music-theoretical approach proposed in this dissertation, and I will come back to it later.

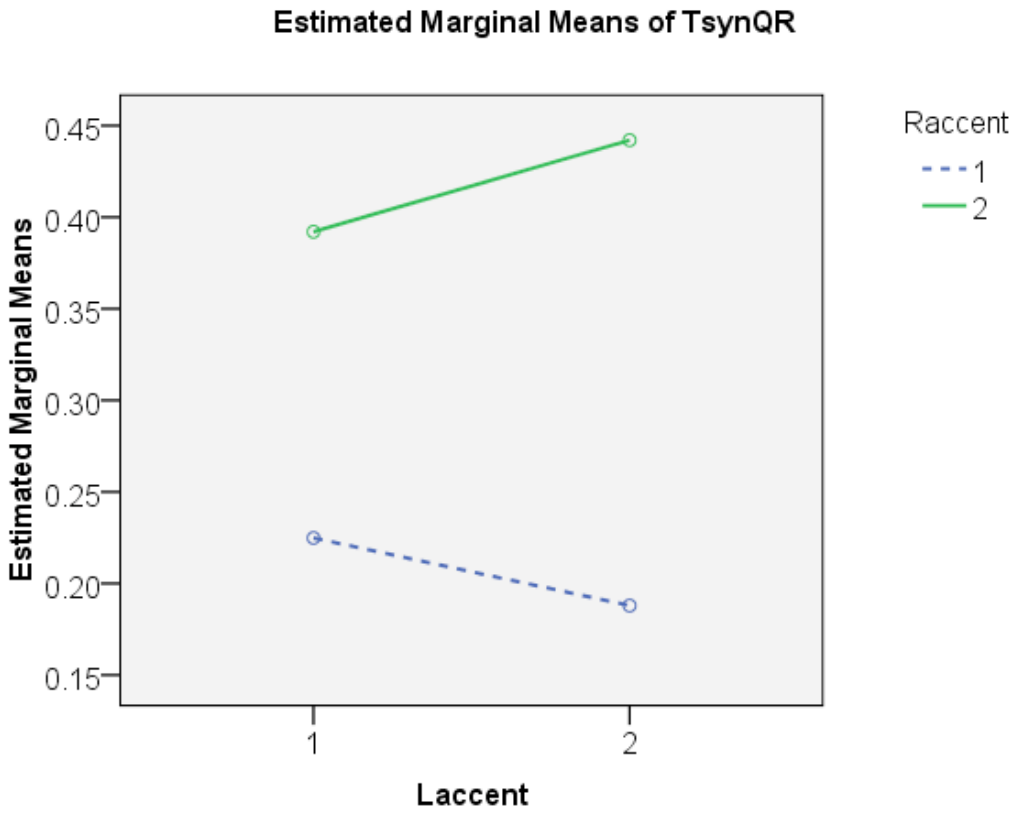
The examination of the interaction between left hand and right hand accents (Laccent × Raccent) will contribute to elucidating the relative influence of accentuation in relation to the temporal structure exhibited by each layer. The interaction Laccent × Raccent was significant for both tapping and beat synchronization with the right hand downbeat (TsynQR and BsynQR), i.e., TsynQR, $F(1, 28) = 5.659$, $p = 0.024$, and BsynQR, $F(1, 28) = 4.858$, $p = .036$. **Figure 4.20** is a graphic representation of the interaction for tapping synchronization with the right hand downbeat (TsynQR); pairwise comparisons for both tapping and beat synchronization are shown below.[43] Although the Laccent factor was found to have no significant main effect on either phase, which is to say that it yielded no significant mean difference across Raccent conditions, there was a significant interaction between these two factors for synchronization rates with the right hand downbeat. Inspection of the graphic representation and pairwise comparisons results show that although the effect of right hand accentuation (Raccent) is more substantial, there is still an effect of left hand accentuation (Laccent) based on the specific accent condition of the right hand layer. Thus, when right

[43] Since both synchronization measures yield a similar pattern, the graphic representation of the interaction for tapping synchronization (TsynQR) is presented as representative of both types.

hand accents are "off" (Raccent 1), the addition of left hand accents (Laccent 2) has a tendency to decrease tapping synchronization with the right hand downbeat. In contrast, when right hand accents are "on" (Raccent 2), the addition of left hand accents increases participants' synchronization with the right hand downbeat. In other words, *participants synchronize more often with the right hand downbeat when both layers' accents are "off" than when only the left hand accents are "on" and they synchronize more often to the right hand downbeat when both layers' accents are "on" than when only the right hand accents are "on."*

To summarize: The analysis results presented in this section demonstrate that *when pitch is indeterminate and the layers are combined, participants' synchronization to competing layers can be successfully manipulated by accentuation*. In this particular polyrhythmic compound, the right hand layer's accentual structure proved to have a more substantial effect on participants' tapping behavior than the left hand layer's accentual structure, although there was an interaction between the two and thus the effect of left hand accentuation was not negligible. Paired-samples t tests showed that mean differences between right hand and left hand downbeat synchronization were statistically significant only when accents were "on." In other words, *participants synchronized more often with the right hand downbeat only when the right hand layer was accented*. The specific patterns of fluctuation in phase that were observed in this analysis provide valuable information on the relative metrical strength of competing layers under certain accentual conditions. The observed shift in mensural determinacy when comparing segregated and combined layers is especially suggestive. We will now turn to salient results from the exploratory analyses with between-subjects factors; a more formal comparison of Blocks 1 and 3 will follow.

Figure 4.20: Laccent × Raccent interaction for tapping and beat synchronization to the right hand downbeat (TsynQR and BsynQR) in Block 3 (L = left hand and R = right hand; Raccent 1 = "off" and Raccent 2 = "on")



Measure	Laccent	Raccent	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
TsynQR	Off	Off	.225	.010	.205	.245
		On	.392	.032	.328	.457
	On	Off	.188	.023	.141	.235
		On	.442	.031	.379	.505
BsynQR	Off	Off	.217	.023	.171	.263
		On	.379	.037	.303	.454
	On	Off	.153	.024	.104	.202
		On	.421	.041	.337	.505

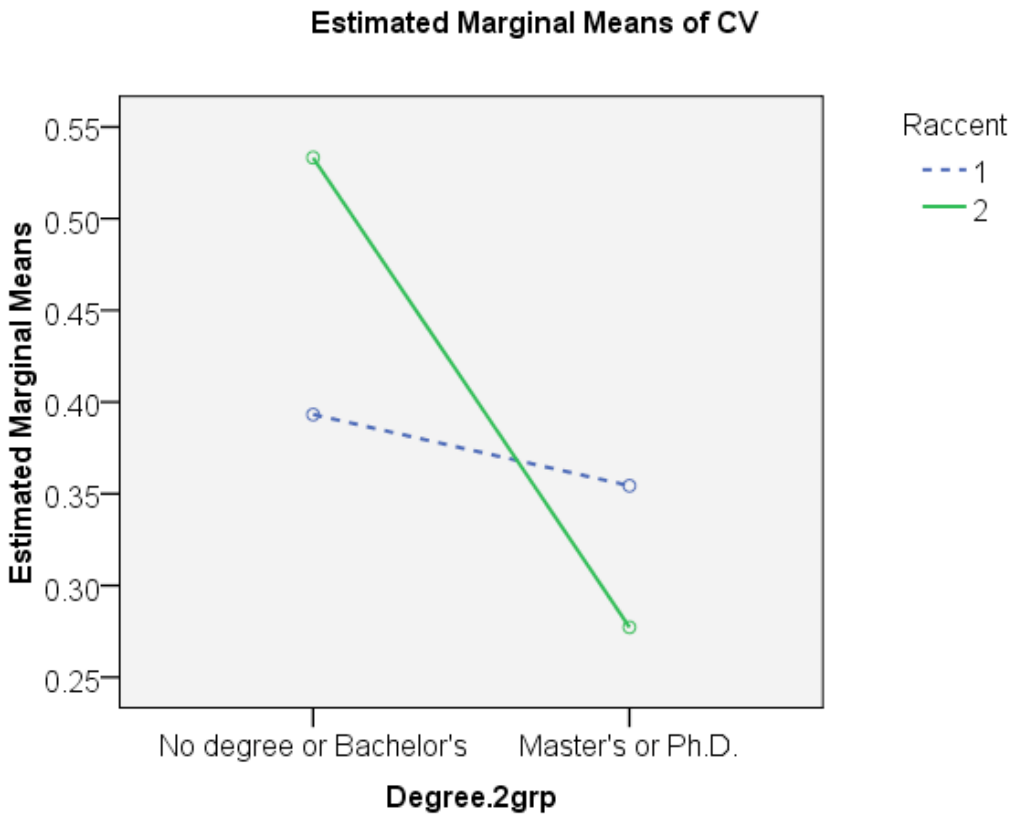
4.2.4.3 Exploratory Analyses: Selected Between-Subjects Factors

Exploratory analysis results will be reported only for four between-subjects factors, i.e., musical degree, experience with polyrhythmic performance, area of specialization, and gender. Given the complexity of the temporal structure of the combined layers, it was predicted that participants with more training and experience as represented by a more advanced musical degree (Master's or Ph.D. group) and prior experience with polyrhythmic performance, would achieve higher tapping consistency (low CV) and higher synchronization rates. Given that this task involved global rather than focal attention, there was no specific prediction made as to the specific synchronization phase. Finally, area of specialization and gender were investigated in order to test whether the particular pattern of responses obtained in Blocks 1 and 2 would still hold when layers were combined.

Contrary to predictions, there was no significant main effect or interaction of musical degree (No degree or Bachelor's = 10, Master's or Ph.D. = 19) with the within-subjects factors Laccent and Raccent. However, there was a trend toward a significant main effect on tapping synchronization with the left hand phase. Thus, participants who had a master's or Ph.D. degree in music synchronized more often with the left hand phase (mean difference = 5.3%, standard error = 2.6%) than thus who had either no musical degree or only a bachelor's, $F(1, 27) = 4.092$, $p = .053$. There was also a trend toward a significant interaction of musical degree with right hand accents (Raccent) for tapping consistency (CV). This interaction suggests that *the effect of right hand accentuation on participants' tapping behavior varied depending on participants' musical training*. As shown by **figure 4.21**, while participants showed a similar tapping consistency rate across degree groups when the right hand accents were "off" (Raccent 1), the addition of accents in the right hand layer

(Raccent 2) correlated with a lower coefficient of variability (more stable tapping) for participants with a higher musical degree but greater variability for those with less training, $F(1, 27) = 4.155$, $p = .051$. These results suggest that when layers were combined synchronization with the left hand phase was more difficult than with the right hand phase and required more advanced musical skills. This hypothesis would need to be tested again using a more balanced sample.

Figure 4.21: Raccent × Degree.2grp interaction for tapping consistency (CV) in Block 3 (Raccent 1 = "off" and Raccent 2 = "on")

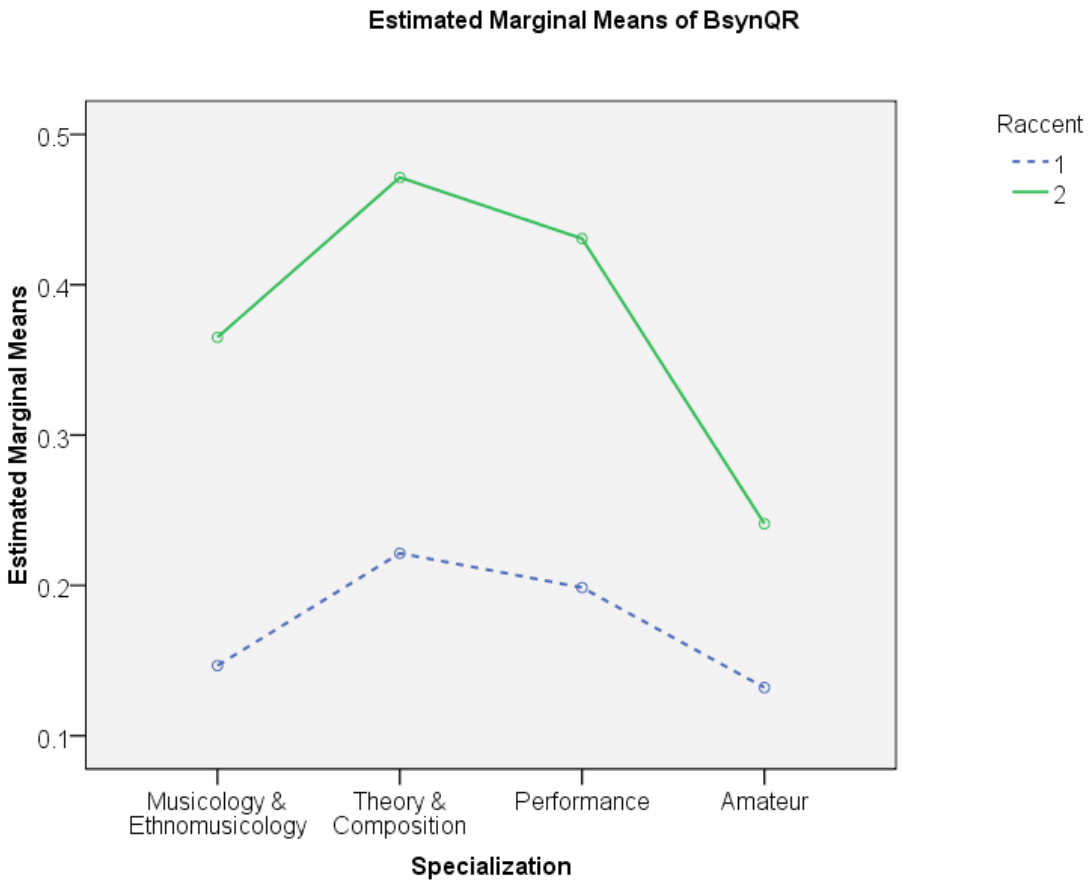


On the other hand, area of specialization (Musicology & Ethnomusicology = 6, Theory & Composition = 11, Performance = 7, Amateur = 5) was found to have a significant

main effect on tapping consistency (CV), F(3, 25) = 4.823, p = .009. Participants who were amateur musicians showed a higher tapping variability than those who specialized in theory and composition (mean difference = 0.430, SE = 0.119, p = .008) as well as those who specialized in performance (mean difference = 0.397, SE = 0.130, p = .031). The pattern of tapping behavior based on area of specialization was very similar to the one for Block 1, with the difference that the mean difference between performers and amateur musicians was now significant. In addition, there was a trend toward a main effect of specialization on overall beat synchronization with the right hand phase (BsynQR), F(3, 25) = 2.988, p = .05. **Figure 4.22** is a graphic representation of the effect of specialization plotted against right hand accents. As shown by this graph, theorists and composers achieved a higher rate of synchronization than amateurs, with a mean difference of 16% (standard error = 4.1%, p = .057). The similar shape of the two right hand accents (Raccent) conditions confirms that area of specialization has a similar effect on synchronization with the right hand phase (no statistical significance), regardless of accentuation, but that participants generally tapped more right hand downbeats when the right hand accents were "on" (Raccent 2).

The effect of prior experience with polyrhythm was also tested (No = 10, Yes = 19), but no main effect or interaction with either of the two accentuation factors was found. Given that the texture used consisted of rhythmic strands rather than superimposed pulse trains such as 2:3 or 5:4, it may be that the particular skills developed in the performance of polyrhythmic passages did not apply to the interpretation and performance of an underlying beat structure in this more loosely constructed (and less regular) polyrhythmic compound. It is also possible that in order to see an effect, the specific experience of participants would need to be better assessed and groups more clearly defined.

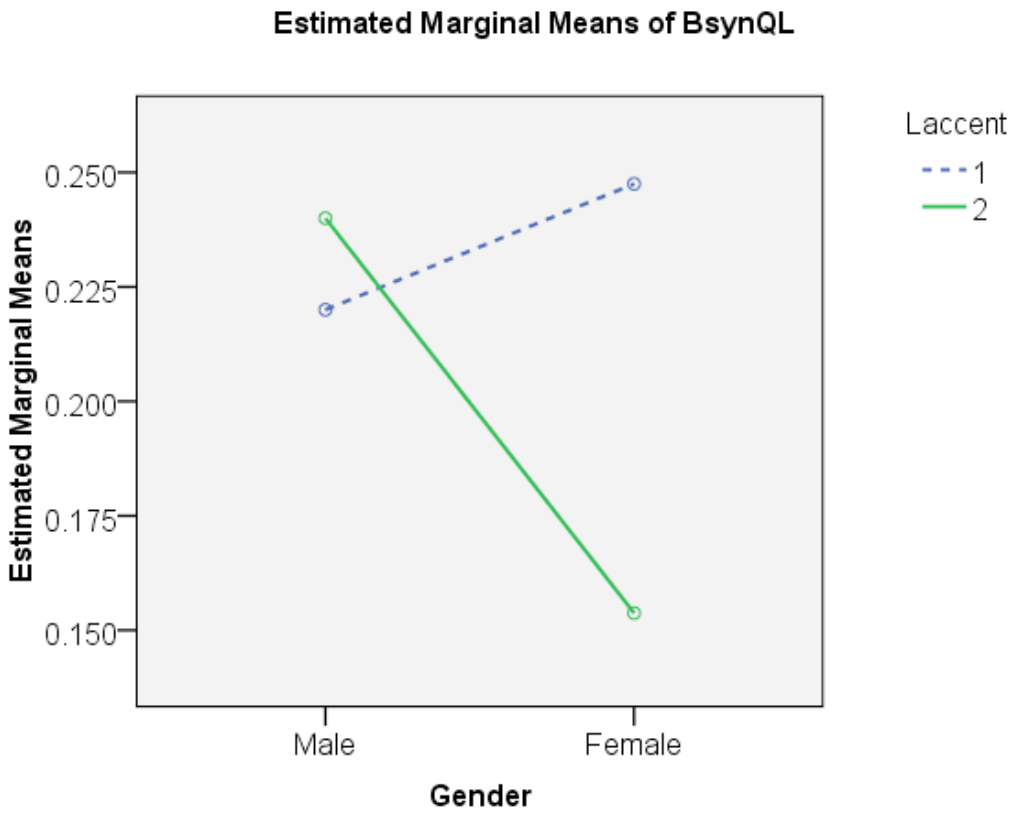
Figure 4.22: Main effect of area of specialization (Specialization.4grp) on overall beat synchronization to the right hand phase (BsynQR) in Block 3 (Raccent 1 = "off" and Raccent 2 = "on")



Finally, it was found that participants' gender (Male = 17, Female = 12) had a significant main effect on tapping synchronization and interacted significantly with left hand accentuation for overall beat synchronization rates. Women synchronized less often than men with the right hand phase of the quarter-note beat (TsynQR), $F(1, 27) = 4.794$, $p = .037$. The significant interaction Gender × Laccent is represented graphically in **figure 4.23**. When the left hand accents were "off" (Laccent 1), participants appear to have tapped a similar number of left hand downbeats. However, when the left hand accents were "on," women tapped a

significantly lower number of left hand downbeats, suggesting that the women in this experiment were less sensitive to the influence of the left hand layer in the combined layers conditions.

Figure 4.23: Laccent × Gender interaction for overall synchronization with the left hand phase (BsynQL) in Block 3 (Laccent 1 = "off" and Laccent 2 = "on")



Measure	Gender	Laccent	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
BsynQL	Male	Off	.220	.035	.149	.291
		On	.240	.031	.176	.304
	Female	Off	.247	.041	.163	.332
		On	.154	.037	.077	.230

To summarize: It appears that *in Block 3, there was a higher level of agreement between participants, although there were trends toward a significant main effect of musical degree on tapping synchronization with the left hand phase as well as an interaction with right hand accents involving tapping consistency.* In addition, the null hypothesis could not be rejected for prior experience with polyrhythmic performance, which suggests that these specialized skills were not applicable to the specific task and polyrhythmic texture in this block. On the other hand, area of specialization and gender both had a significant main effect and interacted with accentuation, confirming the relevance of musical training on synchronization to musical stimuli and providing additional incentive for a more formal investigation of a possible effect of gender on synchronization to music.

4.2.4.4 Across-Blocks Analysis: Blocks 1&3 (Version)

Let us now re-visit the question of mensural determinacy as expressed by participants' tapping performance in a more formal way. In the preliminary analyses to Blocks 1 and 3, paired-samples t tests were conducted to determine whether there was a statistically significant mean difference between synchronization phases for each stimulus. Although the performance measures used in this study don't allow us to identify the specific tapping period, a significant difference in phase can be understood to indicate phase dominance and thus higher mensural determinacy for a given layer under a particular set of conditions. On the other hand, the lack of statistical significance between downbeat and upbeat phase can be correlated with a certain amount of ambiguity since it results from either integrated, intermittent, and/or phase shifting tapping. Integrated tapping applies mostly to combined layers and would involve tapping at a faster or slower rate that allows for

synchronization with alternating downbeats, such as would result from an eighth note or dotted-quarter-note period. It is important to note that ambiguity here is determined in function of specific patterns of synchronization to the quarter-note beat; in this sense, it is the product of participants' tapping responses and the limitations imposed by the specific synchronization measures used. It is possible that some of these responses exhibit a regular tapping pattern that is not in phase with the analyzed beat structure; it is also possible that some of these responses are random. In all cases, an ambiguous response is not less "correct" than one that is deemed representative of left hand or right hand dominance. It is beyond the scope of this dissertation to engage with individual responses.

Table 4.18: Phase dominance and metrical profiles in Blocks 1 and 3

Metrical Profile	Stimulus	Version & Accent Conditions
Left hand phase dominance (QL)	S1, S2	Left hand only & Accents off/on
Right hand phase dominance (QR)	S5	Right hand only & Accents on
	S8, S10b	Combined layers & Right hand accents on
Ambiguous with left hand phase tendency (A/QL)	S10a	Combined layers & Left hand accents on
Ambiguous phase (A)	S4	Right hand only & Accents off
	S7	Combined layers & Accents off

Table 4.18 presents a summary of the results for the paired-samples t tests for Blocks 1 and 3. Metrical profile is shown as relating to quarter beat (Q) phase dominance as determined from a statistically significant mean differences between downbeat and upbeat (Block 1) or left hand and right hand downbeat (Block 3) phase synchronization; corresponding stimuli are described according to Version (segregate or combined layers) and

Accent (left hand and/or right hand accents "off"/"on") conditions. As shown by this table, Left hand phase dominance (QL) was observed in both of the segregated the left hand conditions (without and with accents, S1 and S2), while right hand phase dominance was observed only in the stimuli where the right hand was presented by itself with accents "on" (S5). When accents were "off" (S4), there was no significant difference in phase, thus the determination of this stimulus as having an "ambiguous phase" (a weak metrical profile). As mentioned above, these results suggest that the left hand layer offers a stronger metrical profile, at least as demonstrated by the participants' tapping behavior in this experiment and based on the analyzed beat structure. Interestingly, when layers are combined there seems to be a shift in the relative dominance of the two layers. Thus, the right hand downbeat emerged as the preferred phase in two of the four stimuli of Block 3 (S8 and S10b), those two stimuli sharing the right hand accents "on" condition (regardless of left hand accentuation). Although there was no statistically significant mean difference between the left hand and right hand phases in the two remaining stimuli (S7 and S10a), when the left hand accents were "on" and the competing right hand accents were "off" (S10a), there was a trend toward significance in favor of the left hand downbeat phase.[44] Given that the combined layers are a simple superposition of one of each of the segregated layers' conditions, another way to represent the results of the paired-samples t tests summarized above is to use a simple addition operation (see **table 4.19**)

[44] It is interesting to note that if one disregards pitch content, this condition (S10a) is the one that most closely resemble the source materials.

Table 4.19: Superposition operation, interference levels, and resulting metrical profiles (Legend for stimuli descriptors: S = stimulus, L = left hand, R = right hand, off/on applied to accents "off"/"on"; Legend for metrical profiles descriptors: QL = left hand phase dominance, QR = right hand phase dominance, A = ambiguous phase)

Superposition Operation	Interference Levels[a]		Resulting Metrical Profiles
	L	R	
S1.Loff + S4.Roff = S7.LRoff	Moderate	Moderate	QL + A = A
S2.Lon + S5.Ron = S8.LRon	Moderate	Moderate	QL + QR = QR
S2.Lon + S4.Roff = S10a.LonRoff	Low	High	QL + QA = A/QL
S1.Loff + S5.Ron = S10b.LoffRon	High	Low	QL + QR = QR

a. Interference levels are given from the perspective of each layer.

The resulting metrical profiles, which are based on statistically significant mean differences in phase synchronization, are somewhat different from what one might have expected if the superposition of two competing layers was in fact a simple addition. For example, given two strong metrical profiles such as those offered by the left hand and right hand layers with accents "on" (S2.Lon and S5.Ron), one might have expected some form of integrated response, which here would correspond to an ambiguous profile (A). The actual result, which corresponds to right hand phase dominance, suggests that in superposing these two specific layers, there is some process at work that gives an advantage to the right hand layer at the quarter-note beat level. To determine whether there was a significant effect of Version (segregated versus combined layers), a series of paired-samples t tests was conducted comparing the means of segregated and combined layers under various accent conditions. Five separate tests were performed, one for tapping consistency (CV) and two for each type of synchronization measures (Tsyn and Bsyn) and phase (downbeat and upbeat); each of the five tests includes eight paired samples that account for each of the interference levels introduced above. The results of this analysis, which I will now present, confirm the

observations made above and provide additional information, including upbeat synchronization patterns in relation to the effect of Version.

The results from the paired-samples t tests based on tapping consistency measures (not shown) reveal that there were only two cases for which there was a significant mean difference. In both of these cases, participants' tapping to the right hand layer was significantly more stable when layers were combined. Thus, participants' tapping showed more variation in their tapping response to the right hand with accents "on" (S5.Ron) than to two of the combined layers conditions, i.e., left hand layers with accents "off" and right hand accents "on" (S10b.LoffRon), and left hand and right hand layers with accents "on" (S8.LRon); the results for these two paired-samples t tests were $t(27) = 3.475$, $p = .002$, and $t(28) = 2.881$, $p = .008$, respectively. Incidentally, these two more stable tapping responses correspond to the two combined layers conditions in which participants' tapping exhibited right hand phase dominance.

Table 4.20 presents a summary of the results from the paired-samples t tests for downbeat tapping synchronization (TsynQD), including metrical profiles for segregated and combined layers versions as well as expected interference levels. In this test, the downbeat phase corresponds with the segregated layer's downbeat, e.g., since the first pair includes the segregated left hand layer without accents (S1.Loff), the quarter-note downbeat accounted for in the combined layers condition is the one that corresponds to the left hand downbeat. As shown in the table, there were five paired-samples with significant mean differences (marked with asterisks; pairs that exhibit significant mean differences are also color coded based on the layer being compared, i.e., green = left hand and yellow = right hand). Overall, participants' synchronized significantly more often to the left hand downbeat when it was by

itself than when it was combined with the right hand layer (positive mean difference), regardless of accentuation (S1.Loff and S2.Lon) and interference level, with mean differences between 8.4% and 21.6%. In contrast, participants' synchronized significantly more often with the right hand downbeat only when it was combined with the left hand layer (negative mean difference) and both the right hand and left hand accents were "on" (S8.LRon), with a mean difference of 12.3%. There was also a trend toward significantly higher synchronization with the unaccented right hand downbeat (S4.Roff) in the segregated layers version than in the combined layers version with the accented left hand layer (S10a.LonRoff), where right hand downbeat phase tapping might have been hampered by the high-level interference of the accented left hand layer.

The results of the paired-samples t tests for overall downbeat synchronization (BsynQD) are presented in **table 4.21**. When measuring participants' synchronized taps against the total number of available beats rather than the total number of taps, several of the statistically significant differences reported above were no longer significant, which suggests that some participants might have been tapping sporadically; all of these involved cases where the left hand in the segregated version yielded a higher downbeat tapping synchronization rate. On the other hand, there was an additional significant mean difference in overall beat synchronization with the right hand downbeat phase when the accented right hand layer (S5.Ron) when it was combined with the unaccented left hand layer (S1.Loff), suggesting that the total number of participants' taps in this condition (S10b.LoffRon) might have included a sizeable portion of "dissonant" taps, i.e., taps that were not synchronized with either phase of the quarter-note beat.

Tables 4.22 and **4.23** present the results of the paired-samples t tests for the upbeat synchronization measures. The results for tapping and beat synchronization (TsynQU and BsynQU) are very similar. Overall, participants synchronized more often to the upbeat phase of both layers when the two layers were combined (negative mean differences). However, significant mean differences in upbeat tapping synchronization (TsynQU) occurred only in paired-samples with the segregated left hand layers (S1.Loff and S2.Lon), and at moderate and high interference levels. When considering overall upbeat synchronization (BsynQU), there was an additional significant mean difference involving the accented right hand layer (S5.Ron) and the combined layers version with the unaccented left hand layer, which provided only low-level interference. As anticipated, participants synchronized more often with the upbeat phase of the unaccented left hand (i.e., the right hand downbeat) when it was combined with the accented right hand (S10b.LoffRon) than when it was presented in its segregated version (S1.Loff). Similarly, participants synchronized more often to the upbeat phase of the accented left hand (S1.Lon) when it was combined with the accented right hand (S8.LRon). In both of these cases, the resulting metrical profile for the combined layers version was a right hand phase dominance (QR), which corresponded to a phase shift in relation to the left segregated layers version (from QL to QR). In addition, while the superposition of the unaccented left hand layer (S1.Loff) to the accented right hand layer (S5.Ron) produces only low interference, participants tapped significantly more upbeats to the right hand layer (i.e., more left hand downbeats) in the combined layers version (S10b.LoffRon). Since the resulting metrical profile for the combined layers is one of right hand phase dominance (QR), this suggests that the left hand downbeats became subservient to the right hand layer (i.e., hierarchically integrated).

Table 4.20: Paired-samples t tests results for downbeat tapping synchronization (TsynQD) across Blocks 1 and 3

Paired Samples				Interference Level	Mean Difference (B1 - B3)	Standard Deviation	Standard Error Mean	Significance (2-tailed)
Block 1		Block 3						
Condition	Profile	Condition	Profile					
S1.Loff	QL	S7.LRoff	A	Moderate	.084*	.182	.033	$t(29) = 2.531$, $p = .017$
S4.Roff	A	S7.LRoff	A	Moderate	.013	.072	.013	$t(29) = 1.012$, $p = .320$
S2.Lon	QL	S10a.LonRoff	A/QL	Low	.098*	.235	.044	$t(28) = 2.238$, $p = .033$
S5.Ron	QR	S10b.LoffRon	QR	Low	-.080	.221	.042	$t(27) = -1.913$, $p = .066$
S2.Lon	QL	S8.LRon	QR	Moderate	.216*	.258	.047	$t(29) = 4.595$, $p < .001$
S5.Ron	QR	S8.LRon	QR	Moderate	-.123*	.166	.031	$t(28) = -4.011$, $p < .001$
S1.Loff	QL	S10b.LoffRon	QR	High	.130*	.163	.030	$t(28) = 4.276$, $p < .001$
S4.Roff	A	S10a.LonRoff	A/QL	High	.054	.145	.027	$t(28) = 2.009$, $p = .054$

*. The mean difference is significant at the .05 level.

Table 4.21: Paired-samples t tests results for overall downbeat synchronization (BsynQD) across Blocks 1 and 3

Paired Samples				Interference Level	Mean Difference (B1 - B3)	Standard Deviation	Standard Error Mean	Significance (2-tailed)
Block 1		Block 3						
Condition	Profile	Condition	Profile					
S1.Loff	QL	S7.LRoff	A	Moderate	.040	.195	.036	$t(29) = 1.133$, $p = .267$
S4.Roff	A	S7.LRoff	A	Moderate	-.039	.145	.027	$t(29) = -1.470$, $p = .152$
S2.Lon	QL	S10a.LonRoff	A/QL	Low	.090	.243	.045	$t(28) = 1.997$, $p = .056$
S5.Ron	QR	S10b.LoffRon	QR	Low	-.159*	.242	.046	$t(27) = -3.471$, $p = .002$
S2.Lon	QL	S8.LRon	QR	Moderate	.139*	.286	.052	$t(29) = 2.652$, $p = .013$
S5.Ron	QR	S8.LRon	QR	Moderate	-.186*	.153	.029	$t(28) = -6.533$, $p < .001$
S1.Loff	QL	S10b.LoffRon	QR	High	.041	.194	.036	$t(28) = 1.130$, $p = .268$
S4.Roff	A	S10a.LonRoff	A/QL	High	.024	.142	.026	$t(28) = 0.927$, $p = .362$

*. The mean difference is significant at the .05 level.

Table 4.22: Paired-samples t tests results for upbeat tapping synchronization (TsynQU) across Blocks 1 and 3

Paired Samples				Interference Level	Mean Difference (B1 - B3)	Standard Deviation	Standard Error Mean	Significance (2-tailed)
Block 1		Block 3						
Condition	Profile	Condition	Profile					
S1.Loff	QL	S7.LRoff	A	Moderate	-.026	.121	.022	$t(29) = -1.195$, $p = .242$
S4.Roff	A	S7.LRoff	A	Moderate	.000	.108	.020	$t(29) = 0.017$, $p = .987$
S2.Lon	QL	S10a.LonRoff	A/QL	Low	-.027	.157	.029	$t(28) = -0.909$, $p = .371$
S5.Ron	QR	S10b.LoffRon	QR	Low	-.045	.197	.037	$t(27) = -1.220$, $p = .233$
S2.Lon	QL	S8.LRon	QR	Moderate	-.289*	.240	.044	$t(29) = -6.590$, $p < .001$
S5.Ron	QR	S8.LRon	QR	Moderate	.006	.124	.023	$t(28) = 0.239$, $p = .813$
S1.Loff	QL	S10b.LoffRon	QR	High	-.192*	.228	.042	$t(28) = -4.545$, $p < .001$
S4.Roff	A	S10a.LonRoff	A/QL	High	-.040	.194	.036	$t(28) = -1.102$, $p = .280$

*. The mean difference is significant at the .05 level.

Table 4.23: Paired-samples t tests results for overall upbeat synchronization (BsynQU) across Blocks 1 and 3

Paired Samples				Interference Level	Mean Difference (B1 - B3)	Standard Deviation	Standard Error Mean	Significance (2-tailed)
Block 1		Block 3						
Condition	Profile	Condition	Profile					
S1.Loff	QL	S7.LRoff	A	Moderate	-.057	.167	.030	$t(29) = -1.875$, $p = .071$
S4.Roff	A	S7.LRoff	A	Moderate	-.055	.148	.027	$t(29) = -2.031$, $p = .052$
S2.Lon	QL	S10a.LonRoff	A/QL	Low	-.037	.154	.029	$t(28) = -1.279$, $p = .212$
S5.Ron	QR	S10b.LoffRon	QR	Low	-.104*	.235	.044	$t(27) = -2.351$, $p = .026$
S2.Lon	QL	S8.LRon	QR	Moderate	-.306*	.256	.047	$t(29) = -6.541$, $p < .001$
S5.Ron	QR	S8.LRon	QR	Moderate	-.032	.190	.035	$t(28) = -0.897$, $p = .377$
S1.Loff	QL	S10b.LoffRon	QR	High	-.219*	.251	.047	$t(28) = -4.697$, $p < .001$
S4.Roff	A	S10a.LonRoff	A/QL	High	-.059	.166	.031	$t(28) = -1.896$, $p = .068$

*. The mean difference is significant at the .05 level.

To summarize: It appears that the combined layers version provides an advantage to the right hand layer in such a way that when the two layers are combined and are both presented without accents, the right hand layer interferes with the left hand layer, resulting in ambiguous phase dominance (A), which denotes either integrated, intermittent, and/or phase shifting tapping. Furthermore, when the two layers are combined and the right hand accents are "on" (with or without left hand accents "on"), the presence of the right hand layer produces a significant shift in phase dominance, from left hand phase to right hand dominance (QL → QR). Finally, when the left hand is accented accents and the right hand is unaccented, the interference of the right hand is somewhat alleviated and there is a tendency toward left hand phase dominance (A/QL). Finally, with regard to the music-theoretical part of this dissertation, these results are particularly meaningful as they corroborate the assumptions driving the analytical methodology. More specifically, *these results confirm the relevance of the concept of nestedness for the perceptual processes at work in synchronizing to polymetric textures and provide empirical evidence of the relevance of reductive methods that account for competing layers*, such as the polarized reductions used in chapter 3. In other words, *reductive methods that take the interaction of competing layers into consideration are more likely to produce analyses that are consistent with human experience*.

4.2.5 Block 4

Block 4 presents the stimuli in the combined layers version with piano timbre (determinate pitch). It is the most loosely constructed block but also the one that resembles most closely the source materials. Within each stimulus, the manipulation of musical parameters was modeled on Block 2, i.e., there were variations in accentuation (off/on),

articulation (staccato/legato), and contour (as written/collapsed; left hand layer only). **Table 4.24** presents a summary description of the four stimuli (or conditions). In addition, the combined layers version allowed for a manipulation of the relative register of the two layers, i.e., a reversal between relatively lower (left hand layer) and relatively higher (right hand layer) register (S9 and S12).[45] One notable difference was the consistent presentation of the right hand layer without accents; this was done to mimic more closely the source materials.

Table 4.24: Summary description of Block 4 stimuli

Condition	Stimulus	Description
1	S9*	LH: accents on, staccato, low, as written RH: accents off, legato, high
2	S11a	LH: accents off, legato, low, collapsed RH: accents off, legato, high
3	S11b	LH: accents off, legato, low, as written RH: accents off, staccato, high
4	S12	LH: accents on, legato, high, as written RH: accents off, staccato, low

*. This stimulus is a reproduction of the source materials.

Thus, in contrast to Block 3 where there was no pitch content (teak timbre/indeterminate pitch), the four stimuli allowed for moderate to high streaming for at least parts of the entire duration of the musical excerpt.[46] Given the number of manipulated elements and the necessary limitations of one experimental session, within-layers and

[45] In addition to the register reversal, the articulation was also reversed to facilitate streaming and the combined effect of register and articulation, i.e., the sparser left hand layer in the high register was performed with legato articulation to make it more audible and the denser right hand in the low register was performed with staccato articulation to alleviate the interference with the left hand layer.

[46] As we have seen in the "Experimental Design" section, the particular musical design of the excerpt results in passages where the two layers are not easily separated from each other, as when the two layers cross in registers or one layer is relatively much more active then the other.

between-layers conditions were not balanced (non-factorial design). However, the experimental procedure, which required participants to tap to one layer and then the other allowed for a 2 × 4 factorial design with Attention (L/R) and Condition (each of the four stimuli taken as a whole) as within-subjects factors. This design allowed for the investigation of the relative salience of each layer as well as exploration of possible combinations of musical parameters that could have a significant effect on participants' ability to follow competing metrical frameworks. Across-blocks analysis also allowed for the assessment of the influence of pitch structure (Blocks 3 and 4) on participants' polymetric interpretation, each stimulus being subjected to two interpretations based on attended layer (left hand/right hand). Finally, the issue of changes in relative degree of mensural determinacy (and resulting metrical profiles) could be further explored in the context of determinate pitch (analysis across Blocks 2 and 4).

4.2.5.1 Preliminary Analysis: Order of Presentation and Phase Discrimination

Although the order of presentation was varied between participants, the large number of possible orderings (a total of 24 possible ordering when considering only the four conditions) made it impossible to fully counterbalance it across participants. Each of the 24 possible orderings was used at least once and the order for attended layer (tapping to the left hand or right hand layer first) was varied for each participant across stimuli, but there is no possible way to determine if order had an effect and if this variation in ordering prevented order from interacting with within-subjects factors.

In order to investigate the differences in phase synchronization within conditions, a preliminary analysis consisting of two groups of paired-samples t tests was conducted, each

group assessing the mean difference between tapping and overall downbeat and upbeat phase synchronization (Tsyn and Bsyn) for one of the attended layers (L = left hand and R = right hand). **Tables 4.25** and **4.26** present the results for each group.

Table 4.25: Paired-samples t tests results for phase differentiation in tapping and overall beat synchronization (Tsyn and Bsyn) by condition (S) for attended left hand layer (LH) in Block 4 (QD = quarter-note downbeat and QU = quarter-note upbeat)

Paired Samples		**Mean Difference** (QD - QU)	**Standard Deviation**	**Standard Error Mean**	**Significance** (2-tailed)
Pair 1	TsynQD.S9.LH - TsynQU.S9.LH	.196*	.275	.052	$t(27) = 3.763$, $p = .001$
Pair 2	BsynQD.S9.LH - BsynQU.S9.LH	.163*	.262	.050	$t(27) = 3.283$, $p = .003$
Pair 3	TsynQD.S11a.LH - TsynQU.S11a.LH	.105*	.177	.033	$t(27) = 3.155$, $p = .004$
Pair 4	BsynQD.S11a.LH - BsynQU.S11a.LH	.094*	.143	.027	$t(27) = 3.467$, $p = .002$
Pair 5	TsynQD.S11b.LH - TsynQU.S11b.LH	.154*	.260	.048	$t(28) = 3.187$, $p = .004$
Pair 6	BsynQD.S11b.LH - BsynQU.S11b.LH	.117*	.218	.040	$t(28) = 2.881$, $p = .008$
Pair 7	TsynQD.S12.LH - TsynQU.S12.LH	-.054	.227	.042	$t(28) = -1.291$, $p = .207$
Pair 8	BsynQD.S12.LH - BsynQU.S12.LH	-.065*	.168	.031	$t(28) = -2.083$, $p = .047$

*. The mean difference is significant at the .05 level.

Table 4.26: Paired-samples t tests results for phase differentiation in tapping and overall beat synchronization (Tsyn and Bsyn) by condition (S) for right hand layer attention (R) in Block 4 (QD = quarter-note downbeat and QU = quarter-note upbeat)

Paired Samples		**Mean Difference** (QD - QU)	**Standard Deviation**	**Standard Error Mean**	**Significance** (2-tailed)
Pair 1	TsynQD.S9.RH - TsynQU.S9.RH	-.008	.189	.036	$t(26) = -0.224$, $p = .824$
Pair 2	BsynQD.S9.RH - BsynQU.S9.RH	-.033	.132	.025	$t(26) = -1.286$, $p = .210$
Pair 3	TsynQD.S11a.RH - TsynQU.S11a.RH	-.002	.116	.022	$t(28) = -0.096$, $p = .924$
Pair 4	BsynQD.S11a.RH - BsynQU.S11a.RH	-.009	.103	.019	$t(28) = -0.469$, $p = .643$
Pair 5	TsynQD.S11b.RH - TsynQU.S11b.RH	.009	.153	.028	$t(29) = 0.323$, $p = .749$
Pair 6	BsynQD.S11b.RH - BsynQU.S11b.RH	-.010	.102	.019	$t(29) = -0.553$, $p = .585$
Pair 7	TsynQD.S12.RH - TsynQU.S12.RH	-.012	.162	.030	$t(29) = -0.406$, $p = .688$
Pair 8	BsynQD.S12.RH - BsynQU.S12.RH	.021	.105	.019	$t(29) = 1.083$, $p = .288$

When attending and tapping to the left hand layer (**table 4.25**), participants synchronized significantly more to the downbeat phase of the quarter-note beat than to the upbeat phase (with mean differences between 9.4% and 19.6%), except for tapping synchronization to S12, i.e., the condition in which there was a register reversal. When tapping to S12, participants tapped significantly more upbeats than downbeats (Bsyn), suggesting that the left hand beats were subservient to the right hand phase.[47] The standard

[47] The lack of significance for the tapping synchronization measure (Tsyn) might be due to the comparatively lower mean difference and higher variance in participants' performance; the lower mean difference also suggests that participants performed more dissonant (out-of-phase) taps, but that when considering the number of synchronized taps over the total available number of beats, participants tapped significantly more upbeats than downbeats.

deviations were also relatively high, i.e., 14.3% to 27.5%, pointing to wide differences between participants' responses to each condition when attending the left hand layer. Thus, left hand phase dominance (QL) was the metrical profile for at least three of the stimuli (S9, S11a, and S11b). For the fourth stimulus (S12), the metrical profile was either ambiguous (A) or integrated within the right hand phase (QR). On the other hand, when attending and tapping to the right hand layer (**table 4.26**), participants performed a similar number of taps in the downbeat and upbeat phases (no significant mean difference) for all conditions, suggesting that either participants integrated the two layers together by tapping at a faster rate (e.g., an eighth note beat) or shifted in phase, which can include intermittent tapping as well as integrated tapping (e.g. with a dotted-quarter-note beat). Thus, the resulting metrical profile for all four conditions when attending to the right hand layer was ambiguous (A) and possibly integrated. These results contrast with those from Block 3, where it was found that when layers were combined, participants synchronized more often with the right hand downbeat. They also suggest that when attending and tapping to the left hand layer, participants were relatively more successful in streaming out the competing right hand layer, as represented by the low occurrence of an ambiguous metrical profile. One possible explanation is that the addition of pitch was highly beneficial to the left hand layer; the question of timbre (determinate pitch) effect will be addressed later in the across-blocks analysis section.

Finally, **table 4.27** provides the estimated marginal means of the downbeat and upbeat synchronization measures for each condition across attended layers. These descriptive statistics suggest possible differences based on specific condition. Thus, downbeat tapping

synchronization rates appear to be relatively higher for stimuli S9 and S11b (29.1%), although these two conditions also show the highest standard errors 1.8% and 1.6%, respectively). Both of these conditions exhibit the left hand layer in the lower register with streamed contour (as written). As expected from the phase comparisons, stimulus S12 shows the lowest downbeat tapping synchronization rate (23.3%, SE = 0.9%) and the highest upbeat tapping synchronization rate (26.3%, SE = 1.2%).

Table 4.27: Estimated marginal means of each condition for all synchronization measures across attended layers in Block 4

Measure	Condition	Mean	Standard Error
TsynQD	1: S9	.291	.018
	2: S11a	.258	.012
	3: S11b	.291	.016
	4: S12	.233	.009
BsynQD	1: S9	.237	.026
	2: S11a	.227	.024
	3: S11b	.244	.024
	4: S12	.190	.017
TsynQU	1: S9	.197	.011
	2: S11a	.210	.010
	3: S11b	.213	.011
	4: S12	.263	.012
BsynQU	1: S9	.170	.016
	2: S11a	.186	.020
	3: S11b	.191	.020
	4: S12	.211	.021

4.2.5.2 Main Analysis: The Effect of Attention and Condition

In comparing downbeat and upbeat synchronization within conditions and in each attention group (LH/RH), the preliminary analysis revealed that phase synchronization varied based on attended layer and condition, i.e., participants synchronized significantly more to the downbeat when attending and tapping to the left hand layer, but this distinction did not hold for tapping synchronization to the stimulus with reversed registers (S12). Descriptive statistics also suggested that participants' tapping behavior varied depending on the specific stimulus, with register playing an important role as downbeat locator and streaming cue.

In order to assess the significance of an effect of attended layer and condition on participants' tapping behavior, a univariate analysis of variance was performed with attended layer (Attention) and stimuli (Condition) as within-subjects factors. Overall, both attended layer and conditions had a significant main effect on at least two of the five dependent variables, but attended layer also interacted significantly with condition for three of the four synchronization measures. I will first present the results for the main effect of attended layer; then, I will present the results for the condition factor together with significant interactions, which alter the specific effect of each factor.

Table 4.28: Estimated means and mean differences based on attended layer (Attention) in Block 4 (LH = left hand and RH = right hand)

Measure	Attended Layer	Mean	Standard Error	Mean Difference (LH - RH)	Standard Error	Significance[a]
CV	LH	.306	.036	-.157*	.045	F(1, 26) = 11.977, *p* = .002
	RH	.462	.054			
TsynQD	LH	.300	.017	.064*	.021	F(1, 26) = 8.980, *p* = .006
	RH	.236	.008			
BsynQD	LH	.236	.026	.024	.021	F(1, 26) = 1.243, *p* = .275
	RH	.213	.017			
TsynQU	LH	.200	.010	-.041*	.014	F(1, 26) = 8.546, *p* = .007
	RH	.241	.007			
BsynQU	LH	.158	.017	-.064*	.017	F(1, 26) = 13.573, *p* = .001
	RH	.221	.021			

b. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

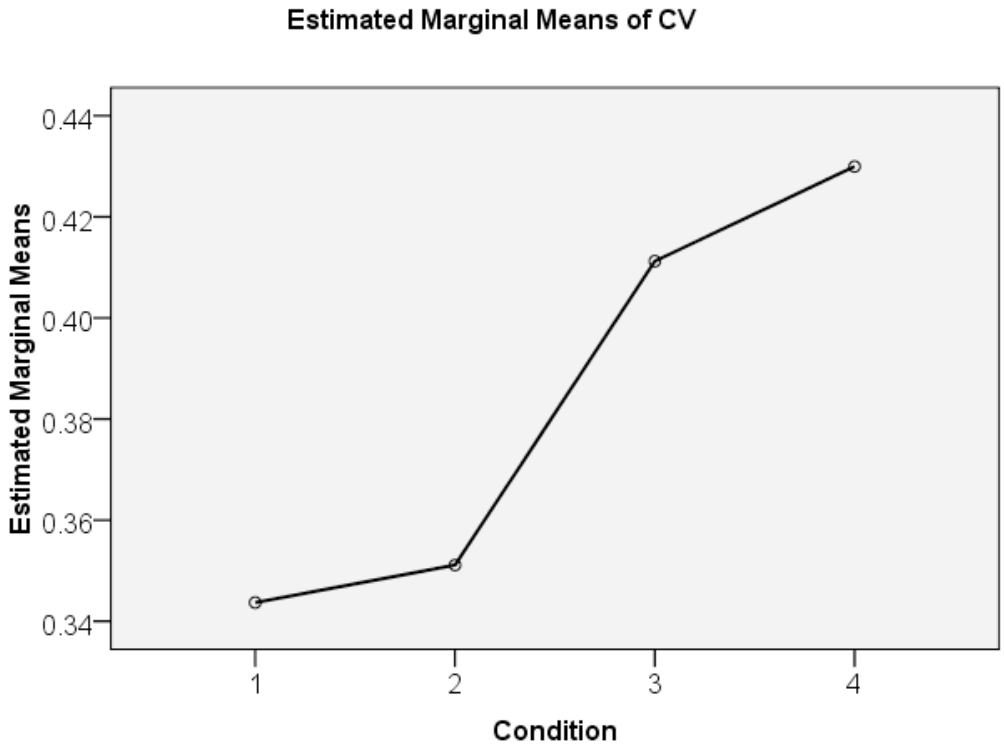
Table 4.28 presents a summary of the results for each attended layer and each dependent measure, including estimated means and mean differences. As shown by this table, attended layer had a significant main effect on all the dependent measures, except for overall downbeat synchronization (BsynQD). In general, participants' tapping performance was less stable when attending the right hand layer (higher CV). As observed in the preliminary analysis, participants also synchronized more often with the quarter-note downbeat when they were attending to the left hand layer, i.e., the tapping synchronization rate (TsynQD) was 30% (SE = 1.7%), as compared with 23.6% (SE = 0.8%) for the right hand layer, and this result was statistically significant. However, when considering overall downbeat synchronization there was no significant difference based on attended layer. The

discrepancy can be explained by the fact that although left hand downbeats represented a higher portion of participants' total taps than the right hand downbeats, when measuring the taps against the total number of available beats, the difference was relatively smaller (only 2.4% as compared with 6.4%; the standard error was the same, i.e., 2.1%). This suggests that participants synchronized with the left hand downbeat for relatively shorter segments of the total duration of the musical excerpt than they did for the right hand; the variance for overall downbeat synchronization (BsynQD) was also greater when participants attended the left hand layer, i.e., 2.6% as compared with 1.7% for the right hand layer. Finally, both tapping and overall upbeat synchronization rates (TsynQU and BsynQU) showed significant mean differences. Participants synchronized significantly more often with the upbeat phase and tapped a significantly higher number of upbeats when attending to the right hand layer than when attending to the left hand layer (mean difference = 4.1% and 6.4%, SE = 1.4% and 1.7%), which is consistent with the lack of a significant mean difference between phases observed in the preliminary analysis. It should be noted that here, the larger mean difference was for the overall upbeat synchronization measure (BsynQU), resulting once more from a relatively higher number of out-of-phase taps when attending to the left hand layer.

A graph of the main effect of condition on tapping consistency (CV) is shown in **figure 4.24**, with estimated marginal means and standard errors. Interestingly, participants exhibited the highest coefficient of variability, or least stable tapping performance, when tapping to the stimulus with register reversed (S12), and the lowest variability (most stable tapping behavior) when tapping to the stimulus that was a reproduction of the source materials (S9). Although condition did not have a statistically significant main effect on tapping consistency, $F(3, 78) = 1.943$, $p = .164$, the mean difference between those two

conditions was significant, i.e., 0.086 (SE = .026, p = .017), and thus the effect of condition on tapping consistency was not negligible. This result also suggests that register reversal had a detrimental effect on tapping consistency as well as downbeat synchronization. However, the relatively high variability also exhibited by participants' tapping to condition 3 (S11b), where there is no register reversal, is more difficult to interpret. When looking at the specific variables of each condition (**table 4.24**), it can be observed that S11b and S12 share the same articulation in the right hand layer (staccato), which contrasts with that used in S9 (most stable tapping) and S11a. In addition, in S11b the left hand layer is unaccented, which also contrasts with S9. Given that there was no main effect of condition on tapping consistency, these observations are far from being conclusive; however, they suggests possible elements for further investigations.

Figure 4.24: Main effect of condition on tapping consistency (CV) in Block 4, with significant mean difference between conditions 1 and 4 (1 = S9, 2 = S11a, 3 = S11b, and 4 = S12)



Measure	Condition	Mean	Standard Error
CV	1: S9	.344	.043
	2: S11a	.351	.035
	3: S11b	.411	.061
	4: S12	.430	.050

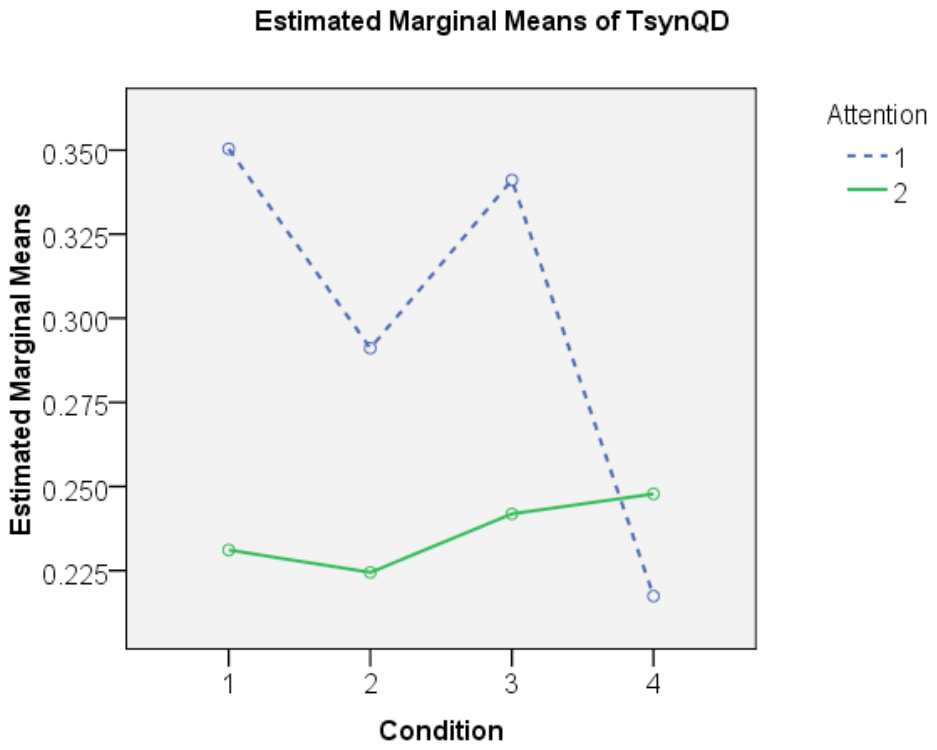
On the other hand, condition had a significant effect on tapping synchronization with the downbeat phase of the quarter-note beat (TsynQD), $F(3, 78) = 4.453$, $p = .012$. Given that there was also a significant interaction of attended layer and condition for downbeat synchronization measures (TsynQD and BsynQD), the former will need to be examined in tandem. The main effect of condition on downbeat tapping was localized between conditions 1 (S9) and 4 (S12), confirming our earlier observation that participants synchronized significantly less often with the downbeat in condition 4, with a mean difference of 5.9% (SE

= 2%, p = .036).[48] **Figure 4.25** provides a graph of the interaction of attended layer and condition for downbeat tapping synchronization (TsynQD), with pairwise comparisons results below.[49] As shown by the graph, when participants attended and tapped to the right hand layer, the synchronization rates were relatively similar across conditions and generally lower than those for the left hand layer. On the other hand, when participants tapped to the left hand layer, the synchronization rates were more variable and exhibited a steep decrease in downbeat synchronization for condition 4 (S12), resulting in a reversal in synchronization performance for right hand layer tapping. In other words, although participants synchronized more successfully with the downbeat when attending to the left hand layer in general (main effect of attended layer), this was not the case for condition 4 (S12) due to the significant interaction of attended layer with condition. Finally, the relatively lower downbeat synchronization rate for condition 2 (S11a) also suggests that the absence of accents and collapsed register in the left hand also had a detrimental effect.

[48] The corresponding beat synchronization measure (BsynQD) showed only approaching significance, $F(3, 78) = 2.843$, p = .056, but there was no significant effect on overall upbeat synchronization (BsynQU).

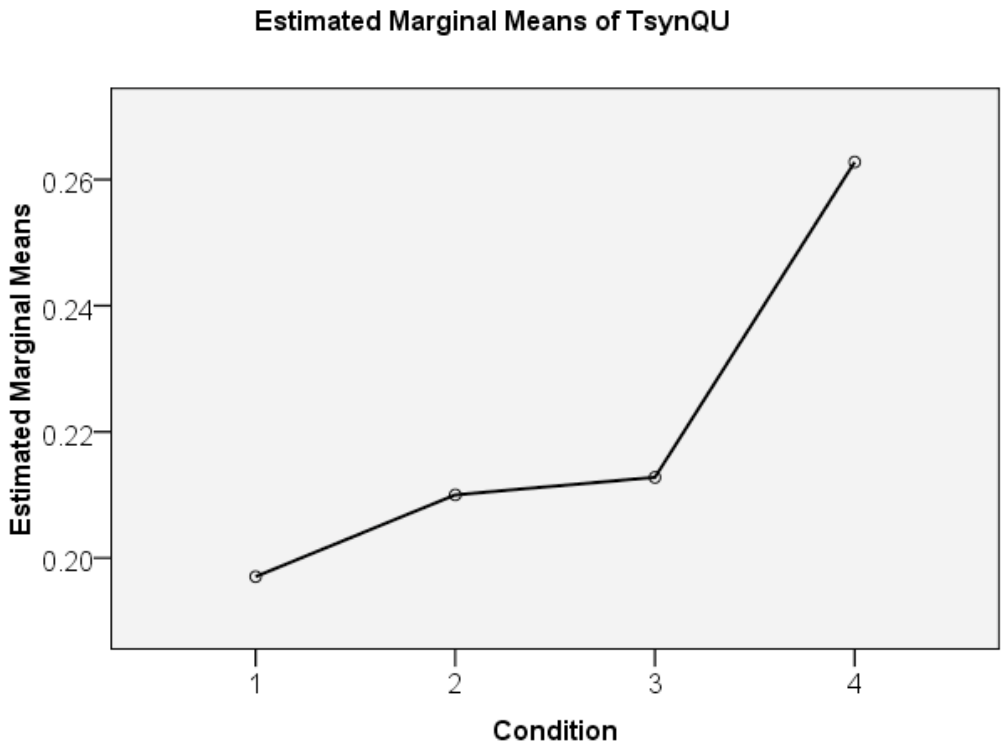
[49] As mentioned above, there was also a significant interaction for overall downbeat synchronization (BsynQD). Since these results were very similar to those for downbeat tapping synchronization (TsynQD), they will not be presented here.

Figure 4.25: Attention × Condition interaction for downbeat tapping synchronization (TsynQD) in Block 4 (Attention: 1 = LH and 1 = RH; Condition: 1 = S9, 2 = S11a, 3 = S11b, and 4 = S12)



Measure	Attended Layer	Condition	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
TsynQD	LH	1: S9	.350	.036	.276	.425
		2: S11a	.291	.023	.244	.339
		3: S11b	.341	.033	.273	.409
		4: S12	.217	.020	.176	.258
	RH	1: S9	.231	.020	.190	.272
		2: S11a	.224	.014	.196	.253
		3: S11b	.242	.013	.214	.269
		4: S12	.248	.017	.213	.282

Figure 4.26: Significant main effect of condition on upbeat tapping synchronization (TsynQU) in Block 4 (1 = S9, 2 = S11a, 3 = S11b, and 4 = S12)



Measure	Condition	Mean	Standard Error
TsynQU	1: S9	.197	.011
	2: S11a	.210	.010
	3: S11b	.213	.011
	4: S12	.263	.012

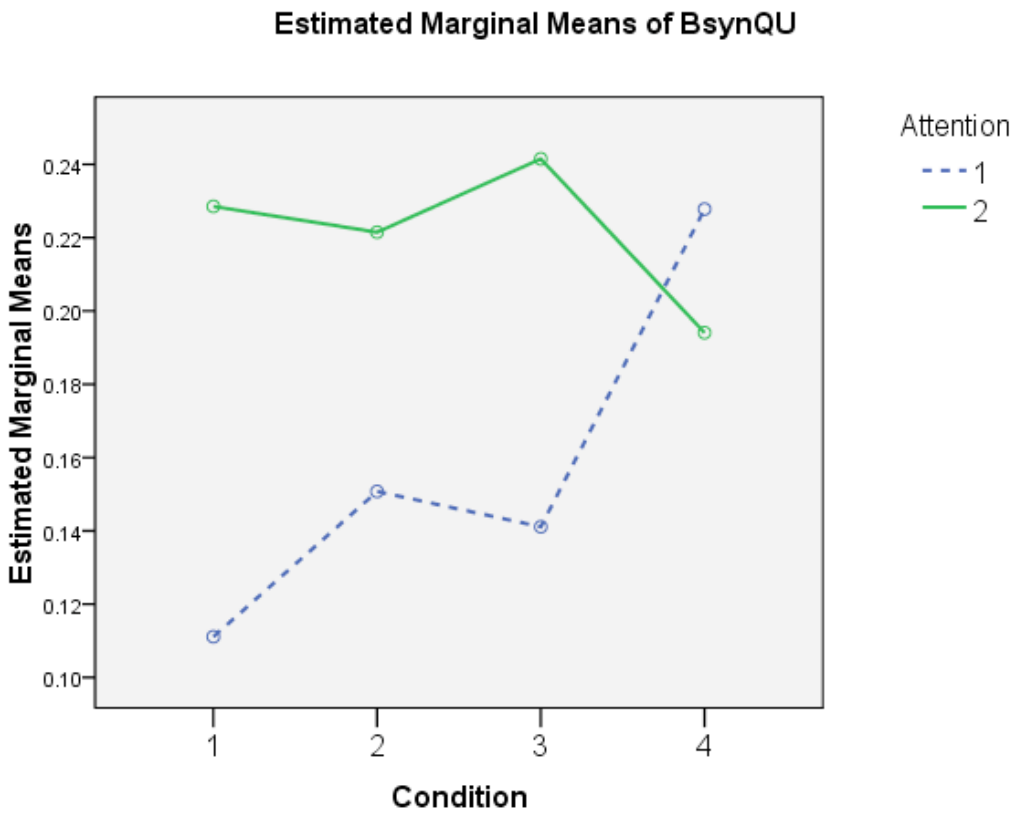
Condition also had a significant main effect on upbeat tapping synchronization (TsynQU), $F(3, 78) = 5.916$, $p = .002$. **Figure 4.26** provides a graphic representation of this main effect, with estimated means below. Participants synchronized more often with the upbeat when tapping to condition 4 (S12) than when tapping to conditions 1 (S9) and 2 (S11a), with mean differences of 6.6% (SE = 1.9%, $p = .014$) and 5.3% (SE = 1.8%, $p =$

.042).[50] Given that condition 4 (S12) also showed the highest variation in tapping (CV) and that there was no significant difference found between phase tapping for this condition, these results suggest that participants were not successful in attending to a specific layer and were induced into frequent phase shifting as well as integration of beats from the competing layer.

Finally, attended layer and condition also significantly interacted for overall upbeat synchronization (BsynQU). **Figure 4.27** presents a graph of this interaction with the pairwise comparison results. As expected, the interaction of attended layer and condition for upbeat phase synchronization closely mirrors that for downbeat phase synchronization, with right hand layer tapping acting as an axis of symmetry (compare with **figure 4.25**).

[50] There was also a trend in the same direction between conditions 4 (S12) and 3 (S11b), p = .057.

Figure 4.27: Attention × Condition interaction for overall upbeat synchronization (BsynQU) in Block 4 (Attention: 1 = LH and 1 = RH; Condition: 1 = S9, 2 = S11a, 3 = S11b, and 4 = S12)



Measure	Attended Layer	Condition	Mean	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
BsynQU	LH	1: S9	.111	.017	.076	.146
		2: S11a	.151	.022	.106	.195
		3: S11b	.141	.027	.085	.197
		4: S12	.228	.031	.165	.291
	RH	1: S9	.229	.025	.176	.281
		2: S11a	.221	.026	.168	.275
		3: S11b	.241	.026	.188	.295
		4: S12	.194	.018	.156	.232

To summarize: The statistical analysis of the main effects and interaction of attended layer and condition (as represented by the particular make-up of each stimulus) confirmed our preliminary observations. *When attending and tapping to the left hand layer, participants appeared to have been successful in streaming out the right hand layer, except for the condition in which there was a reversal in register between the two layers. In contrast, participants' tapping behavior was less regular when attending and tapping to the right hand layer, suggesting that they either had difficulty in synchronizing with the beat or interpreted this layer as fluctuating in phase and/or period.* The latter possibility is supported by the lack of significance in mean differences between downbeat and upbeat phase synchronization for this layer, regardless of condition. Finally, the significant interaction between attended layer and condition for both downbeat and upbeat phase synchronization suggests that when the right hand layer was in the lower register, it interfered with the left hand layer and induced a shift in metrical interpretation for the left hand layer, from left hand phase dominance (QL) to an ambiguous (or integrated) phase (A).[51] *These results are particularly suggestive of the influence of specific musical parameters on participants' ability to follow competing metrical frameworks.* They also indicate specific elements that can be used in the analytical determination of polymetric realization.

4.2.5.3 Exploratory Analyses: Selected Between-Subjects Factors

In an attempt to identify some of the factors that might have contributed to the wide variation observed in participants' tapping performance in this block, exploratory analyses

[51] This effect may also be indirectly related to homogeneity in timbre (i.e., both layers are performed using a piano sample); it is possible that the effect of register would not be significant if the two layers were performed using contrasting timbres (e.g. cello and clarinet samples).

were conducted with the between-subjects factors musical degree, area of specialization, main instrument, piano experience, experience with polyrhythmic performance, and gender. It was hypothesized that musical training in general (musical degree and specialization) would have an effect on participants' ability to engage in focal attention and that the presence of determinate pitch (piano timbre) might lead to significant interactions of within-subjects factors (Attention and/or condition) with participants' main instrument and piano experience. In other words, participants' ability to follow competing musical frameworks embedded in particular pitch structures could also vary based on instrumental training and experience. As with Block 3, it was also thought that prior experience with polyrhythmic performance could have an effect on participants' tapping performance. Finally, the effect of gender was investigated from a purely exploratory perspective.

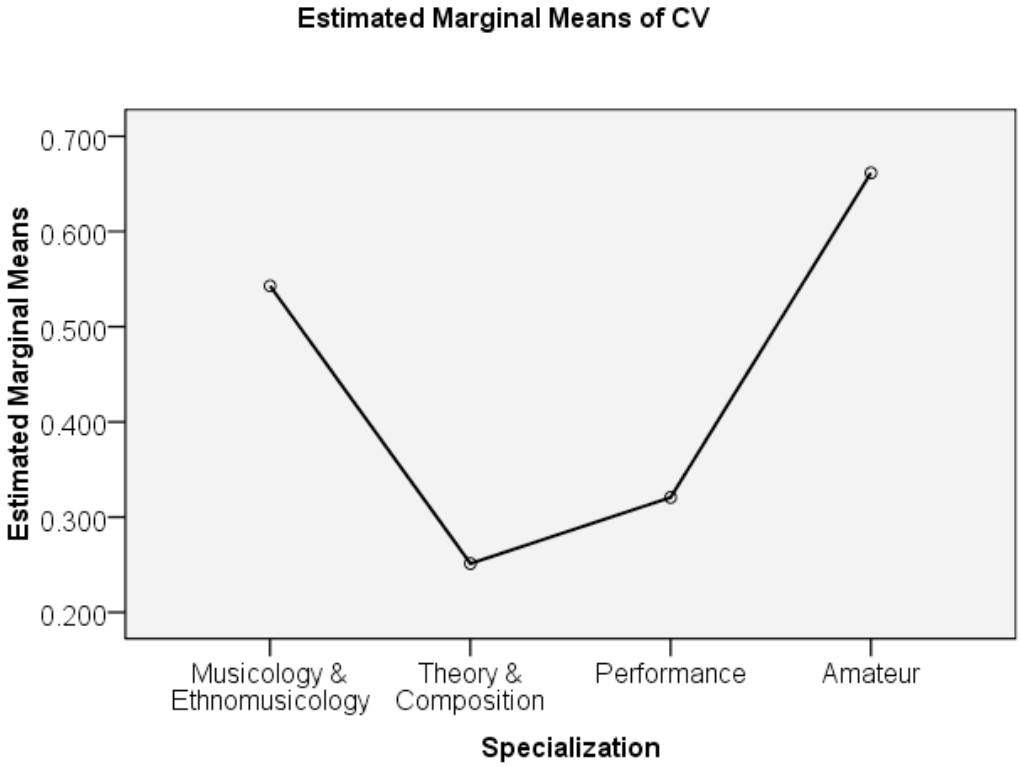
Overall, only musical degree (Degree.2grp) and area of specialization (Specialization.4grp) had a main effect on some of the dependent measures. Main instrument, piano and polyrhythmic performance experience, and gender had no significant effect; in addition, there was no significant interaction of any of the between-subjects factors with either attended layer or conditions. These results might be related to the relatively small sample size (27 valid responses) or methodology (exploratory analyses), and thus are inconclusive.

Musical degree (No degree or Bachelor's = 8, Master's or Ph.D. = 19) had a main effect only on tapping consistency (CV), $F(1, 25) = 4.804$, $p = .038$. Participants without a musical degree or with only a bachelor's showed more variability in their tapping performance (mean = 0.510, SE = 0.045) than those with a higher musical degree (mean = 0.331, SE = 0.069), although there was more variation within the second group, suggesting

that other factors were at play. Area of specialization (Musicology & Ethnomusicology = 5, Theory & Composition = 11, Performance = 7, Amateur = 4) also had a significant main effect on tapping consistency (CV), i.e., $F(3, 23) = 10.710$, $p < .001$, as well as on overall downbeat synchronization (BsynQD), i.e., $F(3, 23) = 3.196$, $p = .045$. **Figures 4.28** and **4.29** provide graphic representations of the main effect of specialization on these performance measures, with selected results from the post-hoc tests. Although the pattern of the effect of specialization on tapping consistency (CV) in Block 4 resembles that of the previous blocks, i.e., highest variability for the Amateur group and lowest for the Theory & Composition group, there is a significant change here involving the Musicology & Ethnomusicology group (**figure 4.28**). In Block 4, *participants who were musicologists or ethnomusicologists showed a greater variability in tapping period,* which led to a significant mean difference with theorists and composers. Theorists and composers also tapped more downbeats overall than other groups, with the greatest mean difference being with amateur musicians (mean difference = 14.6%, SE = 5.3%; see **figure 4.29**).[52]

[52] The discrepancy between significance levels for the main effect of specialization ($p = .045$) and means comparison of the Theory & Composition and Amateur groups ($p = .065$) is related to the use of Bonferroni adjustment; it should also be noted that there were only four valid responses from the Amateur group for this block, and thus the loss of significance might be due to the small sample size.

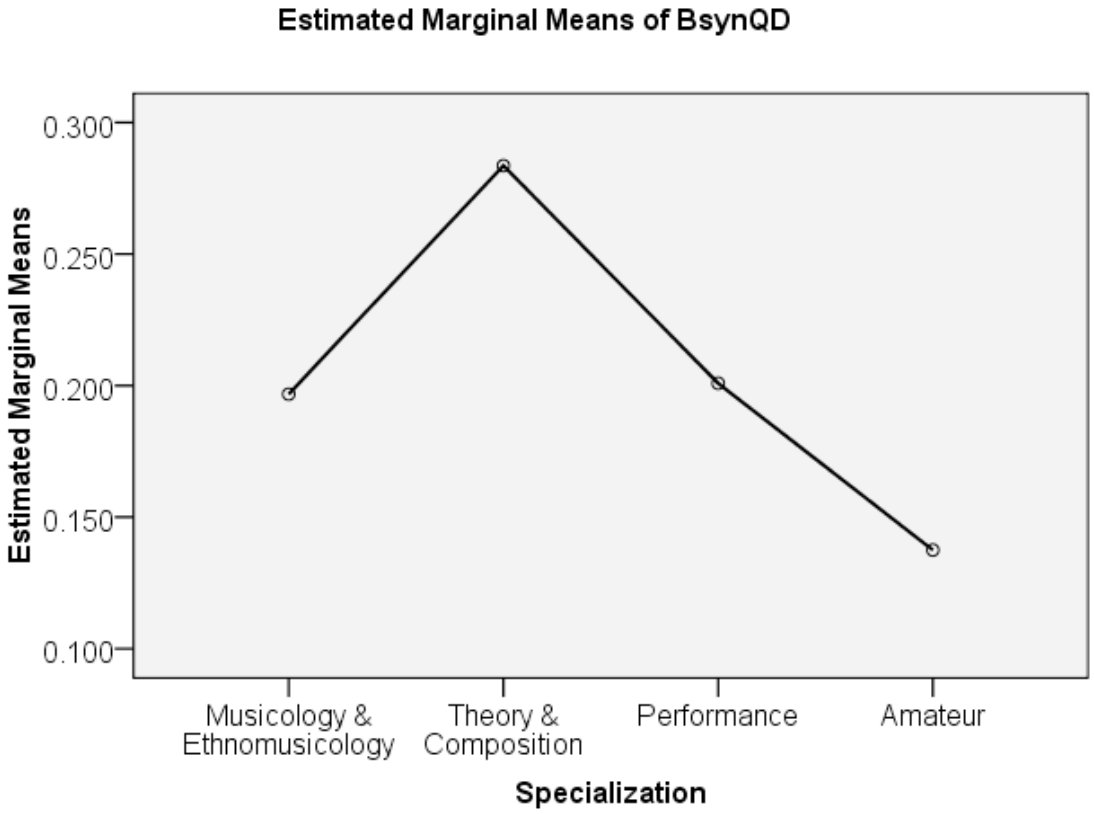
Figure 4.28: Significant main effect of area of specialization (Specialization.4grp) on tapping consistency (CV) in Block 4



Measure	Specialization (I)	Specialization (J)	Mean Difference (I-J)	Standard Error	Significance[a]
CV	Musicology & Ethnomusicology	Theory & Composition	.292*	.077	$p = .006$
	Amateur	Musicology & Ethnomusicology	.119	.096	$p = 1.000$
		Theory & Composition	.410*	.083	$p < .001$
		Performance	.341*	.090	$p = .005$

a. Adjustment for multiple comparisons: Bonferroni.
*. The mean difference is significant at the .05 level.

Figure 4.29: Significant main effect of area of specialization (Specialization.4grp) on overall downbeat synchronization (BsynQD) in Block 4



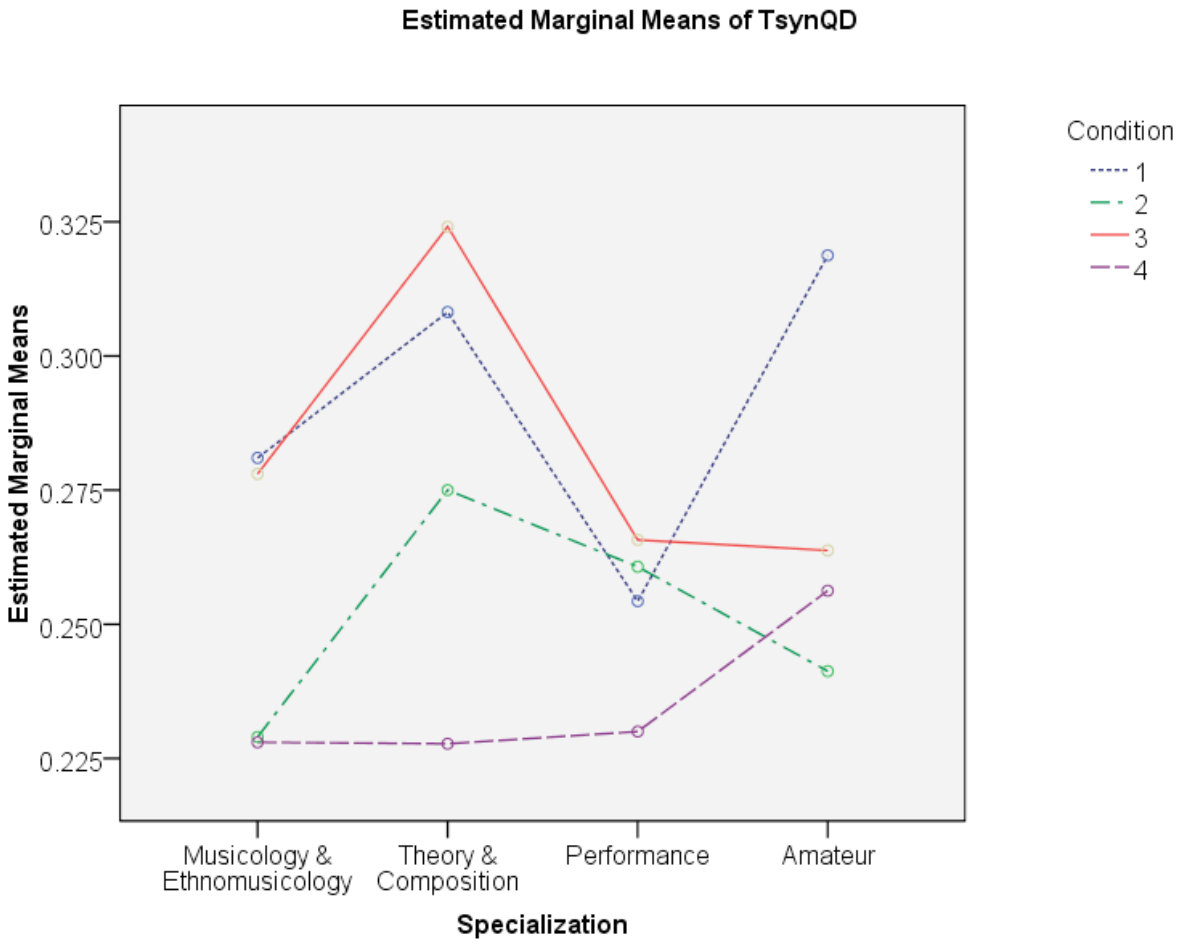
Measure	Specialization (I)	Specialization (J)	Mean Difference (I-J)	Standard Error	Significance[a]
BsynQD	Theory & Composition	Musicology & Ethnomusicology	.087	.049	$p = .523$
		Performance	.083	.04360	$p = .422$
		Amateur	.146	.05265	$p = .065$

a. Adjustment for multiple comparisons: Bonferroni.
*. The mean difference is significant at the .05 level.

Finally, although there was no significant interaction of within- and between-subjects factors, I have included a graphic representation of the interaction of condition and specialization for downbeat tapping synchronization that shows interesting trends (**figure**

4.30). Some of the patterns represented here are similar to those observed in the main analysis, suggesting that specialization did not have a significant effect; for example, condition 4 (S12) yielded the lowest downbeat synchronization phase, and this appeared to be the case across groups. It is especially interesting to note that this low performance applies to theorists and composers as well, a group that typically achieved higher rates of synchronization than other groups. Finally, the spike in performance for amateur musicians in condition 1 (S9) is also interesting given that this stimulus is a reproduction of the score. However, given the small number of valid responses in this group (i.e., only four), this result would require further investigation. One possible explanation is that these participants tapped a relatively smaller number of taps in total, but that a relatively larger portion of them were synchronized with the quarter-note downbeat; this seems to be supported by the results for the overall downbeat synchronization measures (BsynQD). A lower number of taps could be explained by participants' interpretation of segments of the musical surface as non-metric as well as a failure to attend to a given stream.

Figure 4.30: Interaction between area of specialization (Specialization.4grp) and condition for downbeat tapping synchronization (TsynQD) in Block 4 (Condition: 1 = S9, 2 = S11a, 3 = S11b, and 4 = S12)



4.2.5.4 Across-Blocks Analyses: Blocks 2&4 (Version) and 3&4 (Timbre)

In the three previous blocks, analyses across blocks allowed for the investigation of additional within-subjects factors. The across-blocks univariate analysis of variance with Blocks 1 and 2 explored the main effect of timbre (indeterminate versus determinate pitch) and its interaction with layer, controlling for accents (i.e., unaccented conditions only). Similarly, the effect of version (segregated versus combined layers) was investigated by

comparing stimuli from Blocks 1 and 3 in a series of paired-samples t tests. Block 4 was designed to allow for across-blocks analysis in two ways. In order to determine the effect of version (segregated/combined layers), stimulus 11b was built as a simple superposition of Block 2 stimuli 3c (unaccented left hand layer with legato articulation and contour as written) and 6c (unaccented right hand layer with staccato articulation). In addition, conditions can be paired across Block 3 and 4 based on two accentuation models: (1) accents "off" in both layers (S7.LRoff), which corresponded to two of Block 4 conditions, i.e., S11a and S11b; (2) accented left hand with unaccented right hand (S10a.LonRoff), which corresponded to the two remaining Block 4 conditions, S9 and S12. Readers will find it useful to refer back to **table 4.24** at the beginning of this section for a description of the stimuli. Both of these analyses were performed using paired-samples t tests. For each analysis, I will review the tapping consistency analysis first, followed by synchronization measures.

Table 4.29: Paired-samples t tests results for the effect of version on tapping consistency (CV) across Blocks 2 and 4 (S3c = accents "off," legato articulation, and contour as written; S6c = accents "off" and staccato articulation; S11b combines S3c and S6c)

Paired Samples		**Mean Difference**	**Standard Deviation**	**Standard Error Mean**	**Significance** (2-tailed)
Pair 1	CV.S3c.LH - CV.S6c.RH	-.204*	.260	.047	$t(29) = -4.303$, $p < .000$
Pair 2	CV.S11b.LH - CV.S11b.RH	-.198*	.485	.090	$t(28) = -2.201$, $p = .036$
Pair 3	CV.S3c.LH - CV.S11b.LH	-.038	.181	.034	$t(28) = -1.135$, $p = .266$
Pair 4	CV.S6c.RH - CV.S11b.RH	-.042	.503	.092	$t(29) = -0.456$, $p = .652$

*. The mean difference is significant at the .05 level.

In the comparison of segregated and combined layers versions across Blocks 1 and 3 (both with teak timbre or indeterminate pitch), it was found that there were only two cases for which there was a significant effect of version, both of which involved a more stable tapping performance for the accented right hand layer when it is combined with the left hand (accented or unaccented). Similarly, the combined layers version gave an advantage to right hand phase synchronization, especially when the right hand was accented. An unaccented right hand resulted in phase ambiguity (integrated, intermittent, or partly syncopated tapping), which was somewhat lessened when the left hand layer's accents were "on." In the comparisons across Blocks 2 and 4, we will consider only conditions where the right hand accents are "off." Thus, it will be interesting to see how the determinate pitch (piano) conditions compare with the indeterminate pitch (teak timbre) conditions. In order to assess the effect of version on tapping consistency (CV) segregated layers in Block 2 and attended layers in Block 4 were first compared to each other, and then to their corresponding versions across blocks. **Table 4.29** presents a summary of the results of this analysis. As can be observed from the data for the first two pairs (review or earlier results), participants tapped significantly more regularly (lower CV) to the segregated left hand layer (S3c) than to the segregated right hand layer (S6c). When attending and tapping to each layer within the combined layers version (S11b), participants showed the same bias, i.e., their tapping was less stable (higher CV) when attending to the right hand layer than the left hand layer. These results suggest that version did not have an effect on tapping consistency for these two layers. The results for the across-blocks comparison confirms this assessment, showing no significant difference in tapping consistency between each segregated layer and its combined version equivalent. Thus, it would appear that the lack of effect of version when the right

hand accents are "off" (as seen in the analysis across Blocks 1 and 3) was not changed by the presence of determinate pitch, but the effect of timbre will be re-visited below in the analysis across Blocks 3 and 4.

The comparison of synchronization measures across Blocks 2 and 4 showed similar results to those for the tapping consistency measure. **Table 4.30** presents a summary of the results with the metrical profile for each condition. As shown by this table, there was no significant difference between any of the synchronization measures, including downbeat and upbeat phase synchronization.[53] In other words, participants showed no significant increase or decrease in the number of synchronized taps to either the downbeat or upbeat phase of a given layer regardless of whether this layer was segregated or presented in combination with another. Thus, it would appear that version had no significant effect, at least for these two specific conditions (i.e., S3c and S6c), which featured unaccented layers with legato articulation in the left hand (lower register) and staccato articulation in the right hand (higher register). These results are also consistent with the unchanging metrical profiles associated with each condition (based on phase synchronization comparisons within conditions performed as preliminary analysis for each block). The experimental design does not allow us to test possible interaction of version with either articulation or register. These results are also consistent with the finding across Blocks 1 and 3 that the left hand layer showed an advantage in combined layers only when it was accented.

---

[53] Although downbeat and upbeat phases are symmetrically related to each other (e.g., left hand upbeat = right hand downbeat), downbeat and upbeat phases were compared directly across blocks based on attended layer. For example, the upbeat synchronization measure for the left hand layer S3c was paired with the upbeat synchronization measure for the attended left hand layer within the combined layers condition S11b.

Table 4.30: Paired-samples t tests results for the effect of version on tapping and beat synchronization (Tsyn and Bsyn) across Blocks 2 and 4 (S3c = accents "off," legato, and contour as written; S6c = accents "off" and staccato; S11b combines S3c and S6c)

Measure	Paired Samples				Mean Difference (B2 - B4)	Standard Deviation	Standard Error Mean	Significance (2-tailed)
	Block 2		Block 4					
	Condition	Profile	Condition	Profile				
TsynQD	S3c.LH	QL	S11b.LH	QL	-.012	.182	.034	t(28) = -0.357, p = .724
BsynQD	S3c.LH	QL	S11b.LH	QL	-.043	.143	.026	t(28) = -1.615, p = .118
TsynQU	S3c.LH	QL	S11b.LH	QL	-.032	.131	.024	t(28) = -1.303, p = .203
BsynQU	S3c.LH	QL	S11b.LH	QL	-.034	.151	.028	t(28) = -1.203, p = .239
TsynQD	S6c.RH	A	S11b.RH	A	-.015	.082	.015	t(29) = -0.981, p = .335
BsynQD	S6c.RH	A	S11b.RH	A	-.024	.086	.016	t(29) = -1.501, p = .144
TsynQU	S6c.RH	A	S11b.RH	A	-.006	.103	.019	t(29) = -0.338, p = .738
BsynQU	S6c.RH	A	S11b.RH	A	-.023	.118	.022	t(29) = -1.050, p = .303

Table 4.31: Paired-samples t tests results for the effect of timbre on tapping consistency (CV) across Blocks 3 and 4

Paired Samples		Mean Difference	Standard Deviation	Standard Error Mean	Significance (2-tailed)
Pair 1	S7 - S11a.LH	.101*	.166	.031	$t(27) = 3.228$, $p = .003$
Pair 2	S7 - S11a.RH	-.029	.327	.061	$t(28) = -0.483$, $p = .633$
Pair 3	S7 - S11b.LH	.078	.216	.040	$t(28) = 1.950$, $p = .061$
Pair 4	S7 - S11b.RH	-.132	.539	.098	$t(29) = -1.345$, $p = .189$
Pair 5	S10a - S9.LH	.087	.319	.061	$t(26) = 1.422$, $p = .167$
Pair 6	S10a - S9.RH	-.082	.430	.083	$t(26) = -0.988$, $p = .332$
Pair 7	S10a - S12.LH	-.021	.424	.080	$t(27) = -0.258$, $p = .798$
Pair 8	S10a - S12.RH	-.143*	.343	.064	$t(28) = -2.251$, $p = .032$

*. The mean difference is significant at the .05 level.

As we have seen in the analysis across Blocks 1 and 2 (controlling for accents), although timbre did not have a main effect on tapping consistency and synchronization, it was found to interact significantly with layer for upbeat synchronization. More specifically, when participants tapped to the unaccented left hand layer with piano timbre (determinate pitch, they synchronized more often with the upbeat phase than in the corresponding teak (indeterminate pitch) condition. **Tables 4.31** to **4.33** present the results of the analyses performed to assess the effect of timbre for combined layers across Blocks 3 and 4; significant mean differences are highlighted in yellow for ease of reading. Because the task in Block 3 involved global rather than focal attention, each phase synchronization measure in Block 2 (left hand or right hand downbeat) was paired with the corresponding attended layer's metrical function in Block 4 (left hand or right hand downbeat/upbeat); for tapping consistency comparisons, each CV value for Block 3 was used in two comparisons, one for each of the attended layers. As shown by **table 4.31**, participants' tapping performance was significantly different in only two of the eight pairs. Participants' performance was less stable (higher CV value) when they tapped to the unaccented layers with indeterminate pitch (S7) than when they attended and tapped to the unaccented legato left hand in the low register with collapsed contour combined with the unaccented legato right hand in the higher register (S11a); there was no significant difference in participants' tapping consistency when they attended to the right hand layer in the same condition. In contrast, participants' tapping performance was more stable when they tapped to the indeterminate pitch condition with the accented left hand (S10a) than when they attended and tapped to the right hand layer within the determinate pitch condition that combined an unaccented staccato right hand in the lower

register with the accented legato left hand in the higher register (S12). The latter result is not surprising given the detrimental effect already observed when the two layers are swapped in register. The first result, however, provides additional support for a *beneficial effect of determinate pitch on participants' metric interpretation of the unaccented left hand layer*.

Let us now turn to the analyses based on synchronization measures (Tsyn and Bsyn). The two stimuli from Block 3 that are used here for comparison with Block 4 stimuli, i.e., S7.LRoff and S10a.LonRoff, had shown somewhat ambiguous metrical profiles in the preliminary analysis comparing phase synchronization, with the combined layers condition with accented left hand showing a trend toward left hand phase dominance (A/QL). Metrical profiles and results from the paired-samples t tests for downbeat synchronization measures (TsynQD and BsynQD) are shown in **table 4.32**. A comparison of metrical profiles from Blocks 3 and 4 conditions offers a preview of the expected results. For example, the paired samples S7-S11a.RH and S7-S11b.RH show no change in metrical profile (A = A), suggesting that *when both layers were unaccented, participants' synchronization performance when attending to the right hand layer was relatively consistent across timbre as well as articulation conditions* (legato/staccato contrast in right hand). A look at the significance results in the far-right column confirms this assessment (significant mean differences are also marked with an asterisk in the sixth column). Similarly, the significantly higher synchronization rate for downbeat tapping (TsynQD) in condition S11b.LH (mean difference between S11b.LH and S7 = 9.3%) is matched by a shift from ambiguous to left hand phase dominance (A → QL), suggesting that for these particular conditions, *when the layers were combined the switch to determinate pitch favored the mensural determinacy of*

the left hand layer.[54] Finally, these results also contrast with participants' tapping behavior in conditions S12.LH and S12.RH (paired with S10a). In these two conditions, there is a shift from an ambiguous metrical profile with left hand phase tendency (A/QL) to either ambiguous with right hand phase tendency (A/QR) or simply ambiguous (A) metrical profiles. These statistical results also show that in both cases, participants tapped a significantly smaller number of downbeats, suggesting that *when registers were reversed, determinate pitch had a detrimental effect on downbeat synchronization with the left hand layer*.[55]

There remain three cases in which a shift in metrical profile was not explained by significant mean differences in participants' downbeat synchronization, i.e., S7-S11a:LH and the two S10a-S9 pairs (LH/RH). Thus, the profile for the first pair shows a metrical shift from ambiguous phase to left hand phase dominance (A → QL) that is not supported by a significant mean difference in downbeat synchronization. Similarly, when attending to the left hand layers of S9, participants' tapping response appeared to shift from ambiguous phase with a tendency toward left hand dominance (the profile for S10a) to a clear left-hand phase dominance (A/QL → QL) even though participants' synchronization rate was not significantly different; the same issue arise with the contrasting layer, where attending to the right hand layer seemed to correlate with an increase in phase ambiguity (A/QL → A). Given

[54] As this aspect of the synchronization measures has already been discussed several times, discrepancies in significance between tapping and beat synchronization measures will not be discussed here. I will consider a result as significant as long as one of the two synchronization measures (Tsyn or Bsyn) shows a statistically significant mean difference.

[55] It should be kept in mind though that pitch was not controlled beyond the most general consideration of indeterminate or determinate pitch and that a specific interpretation of pitch effects can only be speculative.

that phase dominance is determined by comparing downbeat and upbeat taps, one possible explanation is that these shifts are related to a change in upbeat, rather than downbeat, synchronization. **Table 4.33** presents the results of the paired-samples t tests for the upbeat synchronization measures (TsynQU and BsynQU), along with metrical profiles. Here, we see that the first pair (S7-S11a.LH) is indeed characterized by a significantly higher rate of upbeat synchronization. In other words, *when attending to the left hand layer with determinate pitch, participants tapped significantly fewer upbeats than when tapping to the left hand layer with indeterminate pitch*. The S10a-S9 pairs also show a higher rate of synchronization to the upbeat for the right-hand-layer attending mode, giving more support to the phase shift in the direction of the right hand phase (i.e., A/QL → A/QR). On the other hand, both attended layers for the S10a-S9 pair still show no significant mean difference. Given that the corresponding metrical profiles involved statistical trends, one possible explanation is that these results are related to changing levels of variance among participants, resulting in a change in statistical significance. It is also interesting to note that this pair involves the condition that reproduces the notated score. In other words, it appears that *the condition that most resembles Carter's layout is the one for which there is less agreement among participants*.

Table 4.32: Paired-samples t tests results for the effect of timbre on downbeat synchronization (TsynQD and BsynQD) across Blocks 3 and 4 (selected conditions)

Measure	Paired Samples				Mean Difference (B3 - B4)	Standard Deviation	Standard Error Mean	Significance (2-tailed)
	Block 3		Block 4					
	Condition	Profile	Condition	Profile				
TsynQD	S7	A	S11a.LH	QL	-.039	.135	.025	$t(27) = -1.544$, $p = .134$
BsynQD	S7	A	S11a.LH	QL	-.010	.163	.031	$t(27) = -0.325$, $p = .748$
TsynQD	S7	A	S11a.RH	A	-.003	.083	.015	$t(28) = -0.179$, $p = .859$
BsynQD	S7	A	S11a.RH	A	.010	.125	.023	$t(28) = 0.447$, $p = .658$
TsynQD	S7	A	S11b.LH	QL	-.093*	.171	.032	$t(28) = -2.951$, $p = .006$
BsynQD	S7	A	S11b.LH	QL	-.026	.205	.038	$t(28) = -0.671$, $p = .507$
TsynQD	S7	A	S11b.RH	A	-.021	.095	.017	$t(29) = -1.186$, $p = .245$
BsynQD	S7	A	S11b.RH	A	-.001	.119	.022	$t(29) = -0.061$, $p = .952$
TsynQD	S10a	A/QL	S9.LH	QL	-.064	.188	.036	$t(26) = -1.784$, $p = .086$
BsynQD	S10a	A/QL	S9.LH	QL	-.044	.207	.040	$t(26) = -1.095$, $p = .284$
TsynQD	S10a	A/QL	S9.RH	A	-.047	.146	.028	$t(26) = -1.657$, $p = .109$
BsynQD	S10a	A/QL	S9.RH	A	-.045	.126	.024	$t(26) = -1.847$, $p = .076$
TsynQD	S10a	A/QL	S12.LH	A/QR	.064	.220	.042	$t(27) = 1.530$, $p = .138$
BsynQD	S10a	A/QL	S12.LH	A/QR	.078*	.172	.033	$t(27) = 2.378$, $p = .025$
TsynQD	S10a	A/QL	S12.RH	A	-.054	.170	.032	$t(28) = -1.717$, $p = .097$
BsynQD	S10a	A/QL	S12.RH	A	-.063*	.134	.025	$t(28) = -2.530$, $p = .017$

*. The mean difference is significant at the .05 level.

Table 4.33: Paired-samples t tests results for the effect of timbre on upbeat synchronization (TsynQU and BsynQU) across Blocks 3 and 4 (selected conditions)

Measure	Paired Samples				Mean Difference (B3 - B4)	Standard Deviation	Standard Error Mean	Significance (2-tailed)
	Block 3		Block 4					
	Condition	Profile	Condition	Profile				
TsynQU	S7	A	S11a.LH	QL	.039	.107	.020	$t(27) = 1.935$, $p = .064$
BsynQU	S7	A	S11a.LH	QL	.066*	.145	.027	$t(27) = 2.418$, $p = .023$
TsynQU	S7	A	S11a.RH	A	.019	.076	.014	$t(28) = 1.368$, $p = .182$
BsynQU	S7	A	S11a.RH	A	.018	.127	.024	$t(28) = 0.743$, $p = .464$
TsynQU	S7	A	S11b.LH	QL	.039	.134	.025	$t(28) = 1.574$, $p = .127$
BsynQU	S7	A	S11b.LH	QL	-.012	.124	.023	$t(29) = -0.517$, $p = .609$
TsynQU	S7	A	S11b.RH	A	.007	.100	.018	$t(29) = 0.401$, $p = .691$
BsynQU	S7	A	S11b.RH	A	.000	.121	.022	$t(29) = 0.015$, $p = .988$
TsynQU	S10a	A/QL	S9.LH	QL	.030	.152	.029	$t(26) = 1.010$, $p = .322$
BsynQU	S10a	A/QL	S9.LH	QL	.040	.142	.027	$t(26) = 1.464$, $p = .155$
TsynQU	S10a	A/QL	S9.RH	A	.047	.140	.027	$t(26) = 1.738$, $p = .094$
BsynQU	S10a	A/QL	S9.RH	A	.006	.150	.029	$t(26) = 0.192$, $p = .849$
TsynQU	S10a	A/QL	S12.LH	A/QR	-.093*	.199	.038	$t(27) = -2.475$, $p = .020$
BsynQU	S10a	A/QL	S12.LH	A/QR	-.071*	.176	.033	$t(27) = -2.139$, $p = .042$
TsynQU	S10a	A/QL	S12.RH	A	.029	.182	.034	$t(28) = 0.846$, $p = .405$
BsynQU	S10a	A/QL	S12.RH	A	.036	.160	.030	$t(28) = 1.218$, $p = .233$

*. The mean difference is significant at the .05 level.

To summarize: *When attending to the left hand layer participants' synchronization performance seemed to have benefited from the presence of pitch structure*, as shown by the higher downbeat and lower upbeat phase synchronization' participants' tapping period was also more stable for the unaccented left hand layer when there was determinate pitch. *On the other hand, when attending the right hand layer participants' synchronization performance was relatively consistent across timbre as well as articulation conditions.* Furthermore, *register was especially influential and the specific manipulation of reversed register seemed to have had a detrimental effect on synchronization with the left hand downbeat.* Contour also seemed to have had an effect, as attested by the significantly lower upbeat synchronization rate attending to the unaccented legato left hand with collapsed contour in the low register with the unaccented legato right hand in the higher register (S7-S11a.LH is marked with a shift in metrical profile from A to Q/L). In other words, *participants' might have tapped slower (or in a less syncopated manner) when the left hand layer induced less streaming*. Finally, *participants' ability to synchronize with the attended left hand layer appeared to have benefited from a contrast in articulation and contour* when the unaccented legato left hand with streamed contour was combined with the unaccented staccato right hand in the higher register (as suggested by a metrical profile shift from ambiguous to left hand dominant in the paired sample S7-S11a.LH).

Overall, these results suggest that *there was in fact a significant effect of pitch on participants' ability to follow competing metrical frameworks as well as on the relative degree of mensural determinacy of competing layers, but this effect interacted with layer as well as other non-controlled factors of pitch organization.*

4.4 Final Discussion: Assessing Polymetric Realization

This section presents a brief discussion of the findings of the experiment reported in this chapter as well as possible avenues for future research. The final chapter will provide a complementary discussion of the implications of these findings as related to the conceptual framework introduced in the previous chapters and the overall goal of the dissertation.

The approach presented in this chapter not only addresses important issues relating to the technique of polymeter but it also offers a valuable complement to score analysis, within a conceptual framework that provides constructive differentiation in the realm of polymetric experience. Nonetheless, it is important to emphasize that the analytical interpretation used as a basis for the synchronization measures is not the only possible interpretation of this polymetric structure. In fact, participants' tapping responses (as illustrated by the tapping contour graphs) are another form of analysis and in many ways the experiment was conducted in the spirit of participants providing their own analysis of the texture. The musical excerpt, which corresponds to stimulus 9, was previously analyzed as a balanced polymetric structure (chapter 3, subsection 3.4.3). Based on the comparative analysis of phase synchronization in relation to the quarter beat of the left and right hand layers (i.e., paired-samples t tests), the passage appears to have been experienced most often as a polarized polymetric structure dominated by the lower stream. However, there was a significant degree of ambivalence in participants' tapping responses and a considerable number of taps either fell outside of the set limit of 100ms or belong to other pulse patterns.

Figure 4.31: Tapping contour graph of Participant 8 for Stimulus 9 in Block 4
LH = accented left hand, staccato, lower register; RH = unaccented right hand, legato, high register)

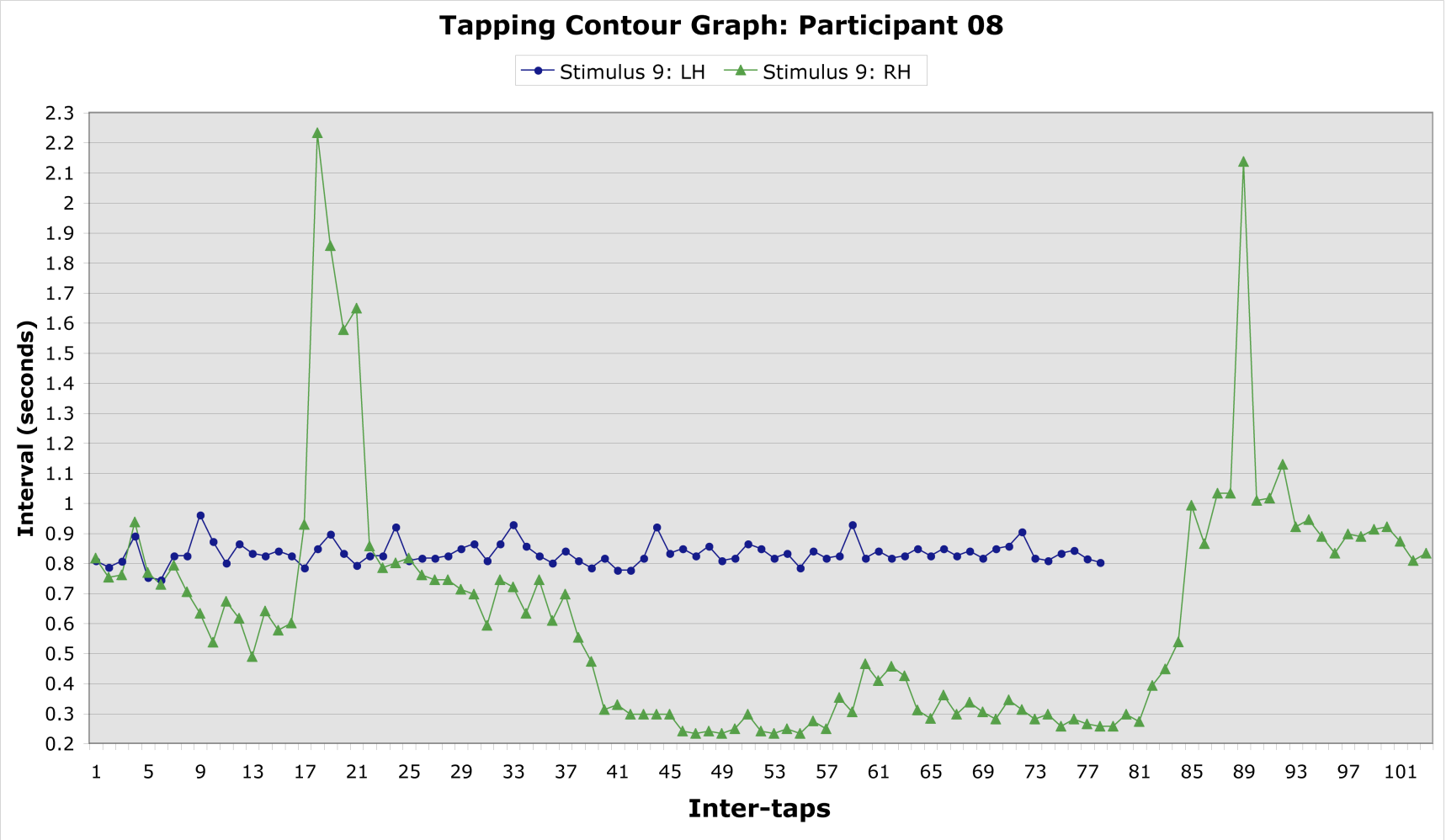


Figure 4.31 presents an exceptionally clear representation of this polymetric interpretation. The tapping contour graph of Participant 8 shows left hand dominance (QL) when attending the left hand layer, with an overall downbeat synchronization rate of 86% (mean ITI = 832ms, CV = 0.046). However, when attending to the right hand layer, this participant's tapping response showed ambiguous phase dominance (A), achieving a mean of downbeat/upbeat synchronization rate of only 31% (mean ITI = 628ms, CV = 0.621). In other words, this participant's tapping shows that the two layers were perceived as expressing contrasting beat structures, with the left hand layer being more strongly determined by the quarter-note beat. This example also demonstrates the limitations of an analytical method based on pre-determined periods. Although the right hand is more variable in its inter-tap intervals (including a hiatus of 2.136 seconds at inter-tap 89), it is not mensurally indeterminate. For example, in this participant's interpretation of the beat structure, the segment including inter-taps 40-81 exhibits a mean ITI of 297ms (SD = 58ms).[56] Thus, while the right hand layer showed ambivalence with regard to the quarter-note beat, it also induced alternative periodicities (or measures) that were at least partially dissonant (i.e., non-synchronized and/or non-isochronous) with the beat structure of the left hand layer.[57] In other words, *not only were participants able to follow competing streams, but when attending to competing layers successively they were also able to experience the same texture according to contrasting metrical frameworks.*

[56] This segment also includes a temporary period of 438ms (SD = 26ms; see **fig. 4.31**, inter-taps 60-63), which is closely related to the quarter-note beat (i.e., 2:1 ratio).

[57] At the notated tempo of quarter = MM 72, the alternative period of 297ms corresponds roughly to a triplet eighth note (60s / 72bpm × 3 = 278ms).

In keeping with previous studies on polyrhythm, it would appear that polymetric textures tend to produce downbeat ambivalence. In a situation where listeners are required to perform pulse-finding tasks, the emergence of a mitigating pulse might be favored (as shown by the shifts in metrical profile from ambivalence to right or left hand dominance), but this experiment shows that polymetric interpretations are also possible, at least if one considers successive interpretations using focal attention. However, a more direct way of testing polymetric perception is still lacking. On the other hand, verbal comments and observations participants' spontaneous synchronization through body movements support the assessment that in this excerpt, there is a perceptible temporal conflict between at least two streams. Based on the verbal comments gathered during the experiment, two specific musical factors appear to have had a determining effect on the cognition of the contrasting metric layers:

1. The presence of clearly segregated and regular slow pulsations with an inter-onset interval of about 3.3seconds or 18 beats per minute in the lower stream, which seemed to be salient enough to counteract the irregularity produced by quasi-isochronous durations at the tactus-level of 72 beats per minute.
2. The attraction of the rapid figuration in the upper stream, which induced a sense of fluctuating pulse and created a strong pull away from the lower stream regularity and its associated beat boundaries.

These two factors point to an interaction between two different levels of attention in the auditory streaming of metrical structures: a low-level, tempo driven, perception of the beat (Scheirer, 1998), and a higher-level awareness of periodicity that calls on stream analysis and memory. These results also support the assumption that it may be beneficial to investigate

pulse levels other than those currently associated with psychological meter (i.e., directly above of below the tactus level).

Overall, the results of this experiment also suggest that the production of a tapping pattern expressive of polymeter does require a particular type of attention and specialized skills, as demonstrated by statistically significant differences in tapping consistency and synchronization measures based on between-subjects factors such as musical degree and area of specialization. It also appeared that while temporal structure is the most consistent factor influencing participants' tapping performance (as shown by the influence of the within-subjects factor of "Layer"), other musical parameters (such as accentuation and register) may be determining and/or interact with each other and with between-subjects factors in various ways. More analysis is required to determine the degree of integration of the two streams, the exact nature of the alternative periodicities, and the specific relationship of taps and periodicities to the musical surface. Furthermore, questions of variability in tapping rate and beat duration also need to be considered, more specifically, what timing boundaries can be used in matching taps with beats and identifying periodicities, at what point does tapping become a reactive task, and how is prediction rate influenced by tempo; this is especially important to allow for comparison with other musical styles. Research on synchronization to expressive musical performances has shown that it is much more difficult for participants to synchronize to performance that include the kind of timing variations applied by human performers (Drake et al., 2000). This is partly due to the fact that such timing variations often go beyond the usual window of tolerance (100ms or 12% of beat value) used as reference for successful synchronization. In many ways, Carter's notated score reproduces such timing variations, which contributed to the difficulty of the task. Finally, it would appear that an in-

depth analysis of the tapping patterns of a selected group of participants in tandem with the musical score would allow the analyst to identify specific elements at work, especially in terms of pitch structure, which was not considered in this experiment other than in the most general manner (i.e., indeterminate versus determinate pitch).

Other types of collected data (i.e., observations of body movements and verbal comments), suggested various avenues of research for future studies. Since many of the participants spontaneously engaged in complex bodily patterns involving more than one motor system, future studies could include a more formal analysis of these patterns. This approach could prove particularly fruitful for the study of polymeter, where tapping superimposed beat structures seems to require integrative processing and might not be fully representative of "experienced" temporal structures. Interpretation of the results also suggested an interaction between tapping strategy and variability; future research could address the influence of counting and grouping strategies on tapping consistency. Another potentially rich field of investigation suggested by participants' verbal comments (not reported) is the study of spontaneous imagery. Overall, it would seem that multi-modal approaches might help us to move closer to a meaningful description of polymeter from a psychological perspective. Furthermore, given the wide variations in participants' musical training and the far-reaching implications of daily involvement with music, these studies should not limit their investigation to a statistical evaluation of the responses but also pay attention to individual responses.

CHAPTER 5

Concluding Remarks: Toward a General Theory of Polymeter

Throughout this work, I have made the point that definitions are shaped by the questions we ask and the methods we choose to provide answers to these questions. We saw that this was especially true in empirical research. However, no discipline escapes this fact. In Western music theory, meter is most often described as a hierarchical accentual structure that often presupposes isochronous beats and implies recursiveness. However, this view of meter is certainly not universal, and even from a psychological perspective (as suggested by work on memory) it would appear that some differences between successive layers of temporal organization are better understood as differences in kind, rather than degree. Meter as entrainment can be understood as one type of meter, situated at the specific level of the tactus (as manifested by spontaneous pulse); it is the meter of the psychological present. However, perception is not tied to action (as when listeners synchronize to music), and other modes of psychological experience need to be investigated, from subliminal awareness to schema-based cognition.

An important part of Carter's modernist aesthetics is to "push" the limits of temporal experience, to create a musical object that appeals to simultaneous and contrasting experiences of time. While it presents several challenges, the study of polymeter from a psychological perspective would seem to hold much promise. It can also help re-shape our understanding of meter so that it will serve the needs of musicians and scholars involved with contemporary musical practices better. Within this perspective, a more general working definition of *meter as a way of measuring and organizing the recurrence of events in time*

appears to be flexible enough to accommodate contrasting conceptualizations of musical past and present as well as to embrace possible future developments. Finally, psychology and empirical methods need not dictate the psychological relevance of polymeter. I have made the point that it may be beneficial for music theory and music psychology to influence one another, but advances in one should not be limited by the state of research in the other. A more fruitful collaboration can be envisioned in which the fruits of psychological research and music analysis are used to develop conceptual frameworks that are informed by both perspectives.

The stated purpose of this dissertation is to present a conceptual framework toward a general theory of polymeter applied to a specific subset of works representative of one musical practice in which the use of contrasting metrical frameworks is an important part of the discourse, from composition to expression and reception. In chapter 1, Carter's solo and chamber instrumental works after 1980 were identified as a repertoire where the technique of metric layering as represented by structural polyrhythms had proven to be especially relevant to both compositional process and aesthetic goals; this assessment was supported by scholarly literature, musical examples, comments made by the composer as well as my survey of his sketches for these works.

After this introduction, the first part of the dissertation (chapters 2 and 3) presented the conceptual framework applied to (and inspired by) Carter's music. The focus of the analytical discussion was on three representative works, in which structural polyrhythms were found to be especially relevant to the musical experience: *Changes* for guitar (1983), and *Esprit rude/Esprit doux I* for flute and clarinet (1985), and *90+* for piano (1994). A contrast was drawn between textural streaming and metric layering. Polymetric structures

were defined as contrasting rhythmic strands built on competing pulse streams and giving rise to non-isochronous beat structures and non-synchronous metrical projections. Three models of polymetric structures were introduced. In *balanced* polymetric structures, the rhythmic strands are nearly equal in metrical strength; *polarized* structures either involve shifting dominance or one pulse stream that is dominant while others provide periodic syncopation; finally when streams are fused, the polymetric structure is *integrated*.

By using a combination of analytical tools, concepts, and visual representations, it was demonstrated that Carter's multileveled polyrhythmic structures are set up to temporal expectations in a variety of ways. The composer's particular use of polymetric structures involves temporal processes operating at different structural levels (*foreground*, *middleground*, *background*) and relating to each other in such a way that their interaction contributes to the realization of polymetric potential. On one hand, while Carter's foreground level is characterized by a limited projective field, isochronous series direct the listener's attention to important moments in the musical surface that often correspond to background pulsations or near-coincidences. On the other hand, background irregularities can be interpreted rhetorically and contribute to a musical narrative articulated at the surface by character-patterns and subjected to various tempo manipulations. Finally, although it is sometimes subject to more local polyrhythms that enhance the effect of background structural polyrhythms, the middleground level is more elusive. It was proposed that this aspect of temporal organization might be a necessary characteristic of polymetric structures. Furthermore, it was shown that the investigation of the density of elementary pulse streams and their degree of nestedness as represented by the middleground pulse (polarity versus

integration) provides a practical tool for determining the specific type of polymetric structure at play.

The experiment described in chapter 4 was guided by three specific goals geared toward an evaluation of the realization of polymetric potential in one specific polymetric structure from the focus repertoire, i.e., the *molto espressivo* passage in Carter's 90+ for piano. After listening to manipulation of this musical excerpt, participants were asked to perform an interpretation of the underlying beat structure by tapping along (pulse-finding paradigm). The first goal involved two related steps: (1) determining the relative degree of mensural determinacy of two layers independently from each other, and (2) determining if the relative degree of mensural determinacy of these two layers changed when they were combined. These two steps allowed us to test assumptions about the influence of nestedness on the realization of polymetric potential. This investigation was based on statistical analysis of the tapping consistency and phase synchronization measures comparing participants' responses to segregated and combined layers. In the music-theoretical chapters (chapter 2 and 3), it was posited that nestedness, i.e., the property of rhythmic events that belong to competing pulse streams, would be determining for the realization of polymetric potential. Given that the two layers exhibited a common period, i.e., the quarter-note beat, this passage provided an interesting case study for competing metrical frameworks. It was found that in the passage used as source material, the left hand layer showed a relatively higher degree of mensural determinacy when it was segregated. On the other hand, the combining of the two parts appeared to be beneficial to the right hand layer, and specific interactions were investigated from the perspective of streaming, interference level, and resulting metrical profiles. The findings provided support for the relevance of the tools introduced in the music-

analytical part of the dissertation, especially nestedness analysis and pulse stream reductions, which investigate the influence of rhythmic events that can be conceptualized to operate in concurrent (and otherwise contrasting) pulse streams.

The second and third goals were motivated by the need to explore more formally the factors that influence listeners' interpretations of competing metrical frameworks and their ability to follow them. Various aspects of the sample texture were manipulated in order to establish whether a causal relationship could be posited between selected musical factors (i.e., accentuation, articulation, overall pitch structure, contour, and register) and participants' tapping performance. It was shown that these factors did have an effect and/or interacted with each other in various ways. Most salient among the results was the finding that accentuation has a main effect on synchronization patterns, but that the effect was not the same for the two layers. This finding points to an interaction between accentual structure and durational structure. The effect of register reversal was also particularly pronounced. When participants were asked to attend to one layer within the polyrhythmic compound, the placement of the left hand layer in the higher register had a detrimental effect on participants' performance. Overall pitch structure had a beneficial effect on participants' interpretation of the left hand layer's beat structure. On the other hand, the effect of articulation and contour (left hand only) was relatively weak.

Data was also collected about the musical training and experience of participants. Results point to a correlation between the amount of time participants spend weekly with twentieth-century music and the consistency of tapping period. Both tapping consistency and synchronization were also influenced by musical degree and area of specialization, with higher degree and theory/composition specialization having a beneficial effect on

participants' performance. Exploratory analyses identified a variety of effects and interactions for factors such as main instrument, piano experience, and even gender (surprisingly), but because of the small sample size, more research is needed.

This study contributes to the study of polymeter both from music-theoretical and music psychological perspectives. In order to move closer to a general theory of polymeter, however, other important aspects of musical structures and practice will need to be addressed. Foremost is the need to move toward analytical practices that can investigate the interaction of pitch and rhythm. The contribution of performers to the realization of polymetric potential also deserves more attention as well as the particularity of their experience. In the performance practice of Carter's works, there appears to be a wide range of awareness of the structures underlying the musical surface, even among expert performers.[1] Finally, a general theory of polymeter necessarily requires an investigation of contrasting repertoires. It is hoped that this study and the tools it provide will inspire others to pursue this project.

[1] This is based on a series of interviews conducted in New York City; these are not reported here.

BIBLIOGRAPHY

I. General Theoretical and Analytical Approaches

Agawu, V. Kofi. 2003. *Representing African Music: Postcolonial Notes, Queries, Positions*. New York: Routledge.

Anku, Willi. 2000. "Circles and Time: a Theory of Structural Organization of Rhythm in African Music." *Music Theory Online* 6 (1), http://www.societymusictheory.org/mto/issues/mto.00.6.1/mto.00.6.1.anku_frames.html (accessed August 6, 2008).

Arom, Simha. 1991. *African Polyphony and Polyrhythm: Musical Structure and Methodology*. Trans. Martin Thom, Barbara Tuckett, and Raymond Boyd. Cambridge, UK: Cambridge University Press.

Berry, Wallace. 1976. *Structural Functions in Music*. Englewood Cliffs, N.J.: Prentice Hall.

Byrne, Andrew McCartin. 1999. "Force: A Study of Polyrhythm." Ph.D. dissertation, Columbia University.

Clayton, Martin. 1997. "Le mètre et le *tãl* dans la musique de l'Inde du Nord." *Cahiers de Musiques Traditionelles* 10: 169-189.

Clendinning, Jane Piper. 1993. "The Pattern-Meccanico Compositions of György Ligeti." *Perspectives of New Music* 31 (1): 192-234.

Erickson, Robert. 1963. "Time-Relations." *Journal of Music Theory* 7 (2): 174-192.

Folio, Cynthia. 1995. "An Analysis of Polyrhythm in Selected Improvised Jazz Solos," in *Concert Music, Rock, and Jazz Since 1945: Essays and Analytical Studies*. Eds. Elizabeth West Marvin and Richard Hermann. Rochester, NY: University of Rochester. 103-134.

Günther, Ursula. 1995. "Some Polymetric Songs in the Manuscript Torino J.II.9," in *The Cypriot-French repertory of the manuscript Torino J.II.9.* Eds. Ludwig Finscher and Ursula Günther. International musicological conference, 20-25 March 1992. Musicological studies and documents Series, No.45. Neuhausen-Stuttgart: Hänssler. 463-489.

_____. 1990. "Polymetric Rondeaux from Machaut to Dufay: Some Style-Analytical Observations," in *Studies in Musical Sources and Styles: Essays in Honor of Jan LaRue.* Madison: A-R. 75-108.

Hanninen, Dora A. 1996. "A General Theory for Context-Sensitive Music Analysis: Applications to Four Works for Piano by Contemporary American Composers." Ph.D. dissertation, University of Rochester.

Hasty, Christopher F. 1997. *Meter as Rhythm*. New York: Oxford University Press.

_____. 1981. "Segmentation and Process in Post-Tonal Music." *Music Theory Spectrum* 3 (1): 54-73.

Hatten, Robert. 1994. *Musical Meaning in Beethoven: Markedness, Correlation, and Interpretation*. Bloomington, IN: Indiana University Press.

Hudson, Richard. 1994. *Stolen Time: the History of Tempo Rubato*. Oxford: Clarendon Press.

Kramer, Jonathan D. 1988. *The Time of Music: New Meanings, New Temporalities, New Listening Strategies*. New York: Schirmer Books.

Krebs, Harald. 1999. *Fantasy Pieces: Metrical Dissonance in the Music of Robert Schumann*. New York: Oxford University Press.

_____. 1987. "Some Extensions of the Concept of Metrical Consonance and Dissonance." *Journal of Music Theory* 31 (1): 99-120.

Lerdahl, Fred and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Locke, David. 1982. "Principles of Offbeat Timing and Cross-Rhythm on Southern Ewe Dance Drumming." *Ethomusicology* 26 (2): 217-246.

London, Justin. "Rhythm," in *Grove Music Online*. Ed. L. Macy (Accessed November 19, 2006), <http://www.grovemusic.com>.

_____. 2002. "Rhythm in Twentieth-century Theory," in *The Cambridge History of Western Music Theory*. Ed. Thomas Christensen. Cambridge, UK: Cambridge University Press. 695-725.

Marvin, Elizabeth West. 1991. "The Perception of Rhythm in Non-Tonal Music: Rhythmic Contours in the Music of Edgard Varèse." *Music Theory Spectrum* 13 (1): 61-78.

Mead, Andrew. 2007. "On Tempo Relations." *Perspectives of New Music* 45 (1): 64-108.

Nattiez, Jean-Jacques. 1987. *Musicologie générale et sémiologie*. Paris: Christian Bourgeois Éditeur.

Robison, Brian Carl. 1999. "*Carmen Arcadiae Mechanicae Perpetuum*: Towards a Methodology for Analyzing Harrison Birtwistle's Music Since 1977." D.M.A. dissertation, Cornell University.

Roeder, John. 1994. "Interacting Pulse Streams in Schoenberg's Atonal Polyphony." *Music Theory Spectrum* 16 (2): 231-249.

Rothstein, William. 2008 (in press). "National Metrical Types in Music of the Eighteenth and Early Nineteenth Centuries," in *Communication in Eighteenth-Century Music*. Eds. Kofi Agawu and Danuta Mirka. Cambridge, UK: Cambridge University Press. 112–59.

Smither, Howard Elbert. 1960. "Theories of Rhythm in the Nineteenth and Twentieth Centuries with a Contribution to the Theory of Rhythm for the Study of Twentieth Century Music." Ph.D. dissertation, Cornell University.

Stone, Kurt. 1963. "Problems and Methods of Notation." *Perspectives of New Music* 1 (2): 9-31.

Thomas, Margaret E. 1996. "Conlon Nancarrow's "Temporal Dissonance": Rhythmic and Textural Stratification in the Studies for Player Piano." 2 vols. Ph.D. dissertation, Yale University.

Trucks, Amanda L. 1992. "The Metric Complex in Johannes Brahms's Klavierstücke, Op. 76." Ph.D. dissertation, University of Rochester, Eastman School of Music.

Yeston, Maury. 1976. *The Stratification of Musical Rhythm*. New Haven, CT: Yale University Press.

II. Psychological Studies

Beauvillain, C. and P. Fraisse. 1984. "On the Temporal Control of Polyrhythmic Performance." *Music Perception* 1 (4): 485-499.

Bregman, Albert S. 1990. *Auditory Scene Analysis: the Perceptual Organization of Sound*. Cambridge, MA: MIT Press.

Brower, Candace. 1993. "Memory and the Perception of Rhythm," *Music Theory Spectrum* 15 (1): 19-35.

Clarke, Carolyn and Richard Parncutt. "Psychology of Music," in *Grove Music Online*. Ed. L. Macy (Accessed November 19, 2006), <http://www.grovemusic.com>.

Clarke, Eric F. 1999. "Rhythm and Timing in Music," in *The Psychology of Music*. Ed. Diana Deutsch. San Diego: Academic Press. 473-500.

Clarke, Eric F. and Nicholas Cook, eds. 2004. *Empirical Musicology: Aims, Methods, Prospects*. Oxford, UK: Oxford University Press.

Clayton, Martin. 2007. "Time, Gesture and Attention in a *Khyãl* Performance." *Asian Music* 38 (2): 71-96.

Drake, Carolyn, Amandine Penel, and Emmanuel Bigand. 2000. "Tapping in Time With Mechanically and Expressively Performed Music." *Music Perception* 18 (1): 1-23.

Glass, Gene V. and Kenneth D. Hopkins. 1996. *Statistical Methods in Education and Psychology*. 3rd ed. Needham Heights, MA: Allyn & Bacon.

Grieshaber, Kate and James C. Carlsen. 1996. "Musicians' Polymetric Performances: Effects of Hand Dominance, Meter Speed, Pattern Complexity, And Cueing Context." *Psychomusicology* 15: 46-68.

Handel, Stephen. 1989. *Listening: an Introduction to the Perception of Auditory Events*. Cambridge, MA: MIT Press, 1989.

Handel, Stephen. 1984. "Using Polyrhythms to Study Rhythm." *Music Perception* 1 (4): 465-484.

Handel, Stephen and Gregory R. Lawson. 1983. "The Contextual Nature of Rhythmic Interpretation." *Perception and Psychophysics* 34 (2): 103-120.

Handel, Stephen and James S. Oshinsky. 1981. "The Meter of Syncopated Auditory Polyrhythms." *Perception & Psychophysics* 30 (1): 1-9.

Jagacinski, Richard J., Elizabeth Marshburn, Stuart T. Klapp, and Mari Riess Jones. 1988. "Tests of Parallel Versus Integrated Structure in Polyrhythmic Tapping." *Journal of Motor Behavior* 20 (4): 416-442.

Jones, Mari Riess, Richard J. Jagacinski, William Yee, Richard L Floyd, and Stuart T. Klapp. 1995. "Tests of Attentional Flexibility in Listening to Polyrhythmic Patterns." *Journal of Experimental Psychology: Human Perception and Performance* 21 (2): 293-307.

Keller, Peter E. and Denis K. Burnham. 2005. "Musical Meter in Attention to Multipart Rhythm." *Music Perception* 22 (4): 629-661.

Klapp, Stuart T., Martin D. Hill, John G. Tyler, Zeke E. Martin, Richard J. Jagacinski, and Mari Riess Jones. 1985. "On Marching to Two Different Drummers: Perceptual Aspects of the Difficulties." *Journal of Experimental Psychology: Human Perception and Performance* 11 (6): 814-827.

Large, Edward. 1994. "Models of Metrical Structure in Music," in *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society: August 13-16, 1994, Cognitive Science Program, Georgia Institute of Technology*. Hillsdale, N.J.: Lawrence Erlbaum Associates. 537-542.

London, Justin. 2004. *Hearing in Time: Psychological Aspects of Musical Meter*. Oxford, UK: Oxford University Press.

Magill, Jonathan M. and Jeffrey L. Pressing. 1997. "Asymmetric Cognitive Clock Structures in West African Rhythms." *Music Perception* 15 (2): 189-222.

Moelants, Dirk and Leon van Noorden. 2005. "The Influence of Pitch Interval on the Perception of Polyrhythms." *Music Perception* 22 (3): 425-440.

Peper, C.E., Peter J. Beek, and Andreas Daffertshofer. 2000. "Considerations Regarding a Comprehensive Model of (Poly)rhythmic Movement," in *Rhythm Perception and Production*. Studies on New Music Research, 3. Ed. Peter Desain and Luke Windsor.

Lisse: Swets & Zeilinger Publishers. 35-49.

Pitt, Mark A. and Caroline B. Monahan. 1987. “The Perceived Similarity of Auditory Polyrhythms.” *Perception & Psychophysics* 41 (6): 534-546.

Povel, Dirk-Jan. 1984. “A Theoretical Framework for Rhythm Perception.” *Psychological Research* 45: 315-337.

Pressing, Jeff. 2002. “Black Atlantic Rhythm: Its Computational and Transcultural Foundations.” *Music Perception* 29 (3): 285-310.

Pressing, Jeff, Jeff Summers, and Jon Magill. 1996. “Cognitive Multiplicity in Polyrhythmic Pattern Performance.” *Journal of Experimental Psychology: Human Perception and Performance* 22 (5): 1127-1148.

Repp, Bruno H. 2000. “Pattern Typicality and Dimensional Interactions in Pianists’ Imitation of Expressive Timing and Dynamics.” *Music Perception* 18 (2): 173-211.

Shaffer, L.H. 1981. “Performances of Chopin, Bach, and Bartok: Studies in Motor Programming.” *Cognitive Psychology* 13: 326-376.

Scheirer, Eric D. 1998. “Tempo and Beat Analysis of Acoustic Musical Signals.” *Journal of the Acoustical Society of America* 103 (1): 588-601.

Sloboda, John A. and Andreas C. Lehman. 2001. “Tracking Correlates of Changes in Perceived Intensity of Emotion During Different Interpretations of a Chopin Prelude.” *Music Perception* 19 (1): 87-120.

Snyder, Joel S. and Carol L. Krumhansl. 2001. “Tapping to Ragtime: Cues to Pulse Finding.” *Music Perception* 18 (4): 455-489.

Summers, Jeffery J., Jennifer A. Todd, and Young H.Kim. 1993. “The Influence of Perceptual and Motor Factors on Bimanual Coordination in a Polyrhythmic Tapping

Task." *Psychological Research* 55 (2): 107-115.

Toiviainen, Petri and Joel S. Snyder. 2003. "Tapping to Bach: Resonance-Based Modeling of Pulse." *Music Perception* 21 (1): 43-80.

III. Carter: Scores, Writings, Interviews, and Research Guides

Bernard, Jonathan W. 1990. "An Interview with Elliott Carter." *Perspectives of New Music* 28 (2): 180-214.

Boretz, Benjamin. 1970. "Conversation with Elliott Carter." *Perspectives of New Music* 8 (2): 1-22.

Carter, Elliott. 2002. *Harmony Book*. Ed. Nicholas Hopkins and John F. Link. New York: Carl Fisher.

_____. 1997. *Elliott Carter: Collected Essays and Lectures, 1937-1995*. Ed. Jonathan W. Bernard. Rochester, NY: University of Rochester Press.

_____. 1994. *90+ for Piano*. New York: Hendon Music, Boosey & Hawkes.

_____. 1991. *Elliott Carter: in Conversation with Enzo Restagno for Settembre Musica 1989*. Trans. Katherine Silberblatt Wolfthal. Brooklyn: Institute for Studies in American Music, Conservatory of Music, Brooklyn College of the City University of New York.

_____. 1985. *Esprit rude/Esprit doux I for Flute and Clarinet*. New York: Hendon Music, Boosey & Hawkes.

_____. 1983. *Changes for Guitar*. New York: Hendon Music, Boosey & Hawkes.

_____. 1977. *The Writings of Elliott Carter.* Ed. Else Stone and Kurt Stone. Bloomington and London: Indiana University Press.

Doering, William T. 1993. *Elliott Carter: A Bio-Bibliography.* Bio-Bibliographies in Music,

No. 51. Westpark, CT: Greenwood.

Edwards, Allen. 1971. *Flawed Words and Stubborn Sounds: A Conversation with Elliott Carter*. New York: W. W. Norton.

Link, John F. 2000. *Elliott Carter: A Guide to Research*. New York: Garland Publishers.

Pozzi, Raffaele. 1988. "Elliot Carter Talking." Tr. Louise Forster. *Tempo* 167: 14–17.

IV. Analytical Studies of Carter's Music

Bernard, Jonathan W. 1998. "The Evolution of Elliott Carter's Rhythmic Practice." *Perspectives of New Music* 26 (2): 164-203.

_____. 1998. "Elliott Carter: Concerto for orchestra (1965-69)," in *Settling New Scores: Music Manuscripts from the Paul Sacher Foundation.* Switzerland: Schott Mainz, Germany, and Paul Sacher Stiftung. 121-126.

_____. 1995. "Elliott Carter and the Meaning of Modern Time." *The Musical Quarterly* 79 (4): 644-82.

_____. 1993. "Problems of Pitch Structure in Elliott Carter's First and Second String Quartets." *Journal of Music Theory* 37 (2): 231-266.

_____. 1983. "Spatial Sets in Recent Music of Elliott Carter." *Music Analysis* 2 (1): 5-34.

Boland, Marguerite. 2006. "'Linking' and 'Morphing': Harmonic Flow in Elliott Carter's '*Con leggerezza pensosa*'." *Tempo* 60 (237): 33-43.

Bye, Antony. 1994. "Carter's 'Classic' Modernism." *Tempo* 189: 2-5.

Capuzzo, Guy. 1999. "Variety Within Unity: Expressive Ends and their Technical Means in the Music of Elliot Carter, 1983-1994." Ph. D. dissertation, University of Rochester, Eastman School of Music.

_____. 1998. "Letters to the Edition: *From Guy Capuzzo*." *Tempo* 204: 53.

Derby, Richard. 1981-82. "Carter's *Duo for Violin and Piano*." *Perspectives of New Music* 20: 149-68.

Garrison, Leonard L. 1994. “Elliott Carter’s *Scrivo in vento*. I: Historical and Analytical Notes.” *The Flutist Quarterly: The Official Magazine of the National Flute Association* 19 (4): 86-92.

_____. 1994. “Elliott Carter’s *Scrivo in vento*. II: Performance Notes.” *The Flutist Quarterly: The Official Magazine of the National Flute Association* 20 (1): 75-80.

Gass, G. 1981. “Elliott Carter’s Second String Quartet: Aspects of Time and Rhythm.” *Indiana Theory Review* 4 (3): 12-23.

Godfrey, Daniel. 1987. “A Unique Vision of Musical Time: Carter’s String Quartet No. 3.” *Sonus* 8 (1): 40-59.

Goldman, Richard Franko. 1957. “The Music of Elliott Carter.” *Musical Quarterly* 43 (2): 151-170.

Harvey, David I. H. 1989. *The Later Music of Elliott Carter: A Study in Music Theory and Analysis*. New York: Garland Publishers.

Hollinger, Heinz. 1992. “Elliott Carter: écrire par couches superposées.” Trans. Daniel Haefliger. *Dissonance* 31: 10-13.

Jemian, Rebecca. 2001. “Rhythmic Organization in Works by Elliott Carter, George Crumb, and John Adams: Rhythmic Frameworks, Timepoints, Periodicity, and Alignment.” Ph. D. dissertation, Indiana University.

Kozinn, Allan. 1984. “Elliott Carter’s *Changes*.” *The Guitar Review* 57: 1-4.

Link, John F. 1994. “Long-Range Polyrhythms in Elliot Carter’s Recent Works.” Ph. D. dissertation, City University of New York.

_____. 1994. “The Composition of Elliott Carter’s ‘Night Fantasies’.” *Sonus* 14 (2): 67-89.

Lochhead, Judith Irene. 1994. “On the ‘Framing’ Music of Elliott Carter’s First String Quartet,” in *Musical Transformation and Musical Intuition: Essays in Honor of David Lewin.* Eds. Raphael Atlas and Michael Cherlin. Roxbury, MA: Ovenbird Press. 179-198.

_____. 1986. “Temporal Structure in Recent Music.” *Journal of Musicological Research* 6 (1-2): 49-93.

Mead, Andrew. 1995. “Twelve-Tone Composition and the Music of Elliott Carter,” in *Concert Music, Rock, and Jazz since 1945: Essays and Analytical Studies*. Rochester, NY: University of Rochester Press. 67-102.

_____. 1994. “The Role of Octave Equivalence in Elliott Carter’s Recent Music: A Birthday Celebration.” *Sonus* 14 (2): 13-37.

_____. 1984. “Pitch Structure in Elliott Carter’s String Quartet No. 3.” *Perspectives of New Music* 22 (1-2): 31-60.

Noubel, Max. 2000. *Elliott Carter, ou Le temps fertile*. Genève, Switzerland: Contrechamps.

Poudrier, Ève. 2009 (in press). “Local Polymetric Structures in Elliott Carter’s *90+* for Piano (1994),” in *Modernist Legacy: Essays on New Music*. Ed. Björn Heile. Aldershot, UK: Ashgate Publishing Ltd. [292-332.]

Ravenscroft, Brenda. 2003. “Setting the Pace: the Role of Speeds in Elliott Carter’s *A Mirror on Which to Dwell*.” *Music Analysis* 22 (3): 253-282.

_____. 1992. “Texture in Elliott Carter’s *A Mirror on Which to Dwell*.” Ph.D. dissertation, University of British Columbia.

Roeder, John. 2006. ““Autonomy and Dialog in Elliott Carter’s *Enchanted Preludes*,” in

Analytical Studies in World Music. Ed. Michael Tenzer. New York: Oxford University Press. 377-414.

Rosen, Charles. 1984. *The Musical Languages of Elliott Carter*. Washington, D.C.: Music Division, Research Services, Library of Congress.

Sallmen, Mark. 2007. "Listening to the Music Itself: Breaking Through the Shell of Elliott Carter's *In genesis*." *Music Theory Online* 13 (3), http://mto.societymusictheory.org/issues/mto.07.13.3/mto.07.13.3.sallmen.html (accessed August 7, 2008).

_____. 1998. "A Flexible Approach to Ordering and Grouping in Atonal Music in General: Text-Music Relations in Elliott Carter's *In Sleep, In Thunder* in Particular." Ph.D. dissertation, University of Rochester.

Schiff, David. "Elliott Carter," in *Grove Music Online*. Ed. L. Macy (Accessed 30 May 2007), <http://www.grovemusic.com>.

_____. 1998. "A Paper Mountain: Elliott Carter's Sketches," in *Settling New Scores: Music Manuscripts from the Paul Sacher Foundation.* Switzerland: Schott Mainz, Germany, and Paul Sacher Stiftung. 115-118.

_____. 1998. *The Music of Elliott Carter*. Rev. 2nd ed. Ithaca: Cornell University Press.

_____. 1988. "Elliott Carter's Harvest Home." *Tempo* 167: 2-13.

_____. 1983. *The Music of Elliott Carter*. 1st ed. London and NewYork: Eulenburg Books and DaCapo Press.

Theisen, Alan Jude. 2006. "Emancipated Discourse, Temporal Unfolding, and The Articulation of Harmonic Regions in Selected Chamber Compositions by Elliott

Carter." M.M. Thesis, University of Southern Mississippi.

Tingley, George Peter. 1982. "Metric Modulation and Elliott Carter's First String Quartet." *Indiana Theory Review* 4 (3): 3-11.

Truniger, Matthias. 2005. "Simultaneous Streams: Philosophies of Time and the Music of Elliott Carter." D.M.A. dissertation, New England Conservatory of Music.

_____. 1998. "Elliott Carter: *Esprit Rude/Esprit Doux*." *Sonus* 19/1 (1998): 26-52.

Vermaelen, Denis. 2004. "Elliott Carter's Sketches: Spiritual Exercises and Craftmanship," in *A Handbook to 20th-Century Musical Sketches*. Eds. Patricia Hall and Friedemann Sallis. New York: Cambridge University Press. 161-175.

Weston, Craig Alan. 1992. "Inversion, Subversion, and Metaphor: Music and Text in Elliott Carter's *A Mirror on Which to Dwell*." D.M.A. dissertation, University of Washington.

Whittall, Arnold. 1997. "Modernist Aesthetics, Modernist Music: Some Analytic Perspectives," in *Music Theory in Concept and Practice*. Eds. David W. Beach, James M. Baker, and Jonathan W. Bernard. Rochester, NY: University of Rochester Press. 157-180.

V. Audio and Video Recordings

Carter, Elliott. 2008. *String Quartets Nos. 1 and 5*. Pacifica String Quartet. Naxos, American Classics 8.559362.

_____. 2005. *Ravel: Gaspard de la nuit; Carter: Night Fantasies; Two Diversions; 90+*. Pierre-Laurent Aimard, piano. Digital disc. Warner Classics, 2564 62160-2.

_____. 2005. *The Music of Elliott Carter, Volume 6: Violin Concerto, Four Lauds, Holiday Overture*. Rolf Schulte, violin; Odense Symphony Orchestra; Justin Brown,

conductor; Don Palma, conductor. Digital disc. Bridge Records, BRIDGE 9177.

_____. 2004. *A Labyrinth of Time*. Directed by Frank Scheffer. Digital video disc. Juxtapositions, Idéale Audience DVD9DS17.

_____. 2003. *Complete Piano Works*. Winston Choi, piano. Digital disc. L'empreinte digitale, ED 13164.

_____. 2003. *Homages & Dedications: Luimen; Scrivo in vento; Con leggerezza pensosa, Changes, Esprit rude/Esprit doux II, Bariolage; Inner Song; Immer Neu; Gra; Enchanted Preludes; 90+; Canon for 4.* Nieuw Ensemble, Ed Spanjaard, conductor. Digital disc. Montaigne/Naïve, MO 782089.

_____. 2003. *The Music of Elliott Carter, Volume Five, Nine Compositions (1994-2002): Steep Steps; Two Diversions; Retrouvailles; Oboe Quartet; Figment No. 1; Figment No. 2 (Remembering Mr. Ives); Au Quai; Of Challenge and of Love.* V. Blackwell, Bass clarinet; C. Rosen and J. Greenberg, piano; F. Sherry, cello; M. Gallagher, viola; P. Kolkay, bassoon; T. Arnold, soprano; C. Neidich and A. Oshima, clarinet; Speculum Musicae. Digital disc. Bridge Records, BRIDGE 9128.

_____. 2003. *Quintets and Voices: Fragment II; Quintet for Piano & Strings; Syringa; Tempo e Tempi; Quintet for Piano & Winds; Retrouvailles*. The Arditti String Quartet; Ursula Oppens, piano; Lucy Shelton, soprano; Andre Solomon-Glover, baritone; Ensemble Sospeso, Stefan Ashbury, conductor. With Steve Taylor, Charles Neidich, William Purvis, and Frank Morelli. Films by Frank Scheffer and Tim Chu; interview by Joshua Cody; liner notes by Paul Griffiths. Digital video disc. Mode Records, Mode 128.

_____. 2001. *Clarinet Concerto, Triple Duo, Enchanted Preludes, Con leggerezza pensosa,*

Gra, Riconoscenza per Goffredo Petrassi, 90+, Esprit rude/Esprit doux I and II. Nouvel Ensemble Moderne, Lorraine Vaillancourt, conductor. Digital disc. ATMA Classique, ACD2 2280.

_____. 2001. *The Music of Elliott Carter, Volume 4: Shard; Luimen; Tempo e Tempi; Eighth Pieces for Four Timpani.* David Starobin, guitar; Susan Narucki, soprano; Daniel Druckman, timpani; Speculum Musicae, William Purvis, conductor. Digital disc. Bridge Records, BRIDGE 9111.

_____. 2000. *Chamber Music: 5th String Quartet, 90+, Sonata for Cello and Piano, Figment for Cello Alone, Duo for Violin and Piano, Fragment for String Quartet*. Ursula Oppens, piano; Arditti String Quartet. Digital disc. Montaigne/Naïve, MO 782122.

_____. 2000. *Chamber Music: Quintet for Piano and Winds; Scrivo in vento; Con leggerezza pensosa; Inner Song; Esprit rude/Esprit doux* [I]*; Gra; Enchanted Preludes; Duo*. M. Bourgue, oboe; A. Cazalet, French horn; C. Desert, piano; P. Gallois, flute; M. Lethiec, clarinet; A. Noras, cello; G. Poulet, violin; A. Wallez, bassoon. Digital disc. ARION, ARN 68495.

_____. 1998. *Chamber Music for Winds: Woodwind Quintet; Esprit rude/Esprit doux I; Eight Etudes and a Fantasy; Enchanted Preludes; Sonata for Flute, Oboe, Cello, and Harpsichord.* Ensemble Contrasts. Digital disc. Co-production: Classic Produktion Osnabrück & Westdeutscher Rundfunk Köln, cpo 999 453-2.

_____. 1998. *Night Fantasies, [90+]*. Florence Millet, piano. Digital disc. Pianovox, PIA 501-2.

_____. 1997. *The Music of Elliott Carter, Volume 3, The Complete Music for Piano: 90+; Piano Sonata; Night Fantasies; Elliott Carter in Conversation with Charles Rosen*.

Charles Rosen, piano. Digital disc. Bridge Records, BRIDGE 9090.

_____. 1994. *The Music of Elliott Carter, Volume 2, Eight Compositions (1948-1993): Gra, Enchanted Preludes, Duo, Changes, Scrivo in vento, Con legerezza pensosa, Cello Sonata, Riconoscenza per Goffredo Petrassi*. The Group for Contemporary Music. Digital disc. Bridge Records, BCD 9044.

_____. 1992. *Sonata for Flute, Oboe, Cello, and Harpsichord; Sonata for Cello and Piano; Double Concerto for Harpsichord and Piano with Two Chamber Orchestras*. Harvey Sollberger, flute; Charles Kuskin, oboe; Fred Sherry and Joel Kosnick, cello; Paul Jacobs, harpsichord and piano; Gilbert Kalish, piano; The Contemporary Chamber Ensemble, Arthur Weisberg, conductor. Digital disc. Elektra/Nonesuch, 79183-2.