UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

PROLOGUE TO SOUNDEN HORN:

A PHONETIC COMPOSITION FOR CHAMBER ORCHESTRA

A DOCUMENT

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF MUSICAL ARTS

By

KENNETH EDWARD MCSPERITT, III
Norman, Oklahoma
2017

PROLOGUE TO SOUNDEN HORN: A PHONETIC COMPOSITION FOR CHAMBER ORCHESTRA

A DOCUMENT APPROVED FOR THE
SCHOOL OF MUSIC

BY

_____
Dr. Konstantinos Karathanasis, Chair

_____
Dr. Marvin Lamb

_____
Dr. Sarah Ellis

_____
Dr. Paula Conlon

_____
Mrs. Karen Antell

© Copyright by KENNETH EDWARD MCSPERITT, III 2017
All Rights Reserved.

Acknowledgements

The composer would like to dedicate this document to the individuals who supported me in so many ways throughout the course of my graduate experience: To my remarkably understanding and patient wife, Carin, and my inspirational children, Bethany, Ian, Tiernan, and Isobel… thank you and I love you! To my parents, Ken and Jean McSperitt, I owe you my life and the basis for who I am. To my composition professors, Dr. Marvin Lamb and Dr. Kostas Karathanasis, you are truly wonderful mentors. Thank you for your support and for being my professional "sounding boards." To my other D.M.A. Committee members, Dr. Sarah Ellis, Dr. Paula Conlon, and Prof. Karen Antell, thank you for challenging me academically to perform to my highest potential. Your impressions will live with me, guiding my future standards. To Clarence Barlow, for being a kindred spirit, both in the world of musical applications of the phonetic sciences and in sharing your wit and word-smithy-ness with me. You are a joyous soul! And finally, to all those who touched me along the way, who fed my curiosity and challenged my scholarship, thank you for your dedication and encouragement.

Table of Contents

Acknowledgements iv

List of Figures vi

Abstract viii

Chapter 1: Introduction 1

Chapter 2: Approaches to Orchestration of Phonetic Techniques 5

Chapter 3: Spectral Mass Technique: 15

Chapter 4: Sound Color Technique: 28

Chapter 5: Phonetic-Event Technique: 41

Chapter 6: Reflections on *Prologue to Sounden Horn* 54

References 58

Appendix A: Annotated List of Notable Compositions 63

Appendix B: Complete Score to *Prologue to Sounden Horn* 67

List of Figures

Figure 2. 1 Establish a Rhythmic Context 9

Figure 2. 2 Instrument Assignments by Harmonic Zones 11

Figure 2. 3 Sample Phonetic Map from Mvt. III of *Sounden Horn* 13

Figure 2. 4 Percussion Key 14

Figure 3. 1 Sample Spectrogram of Excerpt from Mvt I 17

Figure 3. 2 A Sample Phonetic Map Used in Mvt I 18

Figure 3. 3 Percussion Key Used in Phonetic Maps 19

Figure 3. 4 Spectrograms of Mvt I, Phrase 3 20

Figure 3. 5 Notation of Bassoon Line Representing F1 in Phrase 3 25

Figure 3. 6 Sample of Arpeggiations Used to Represent Multiple Formants in One Instrument 26

Figure 4.1 Text and Phonetic Transcription for Mvt II Source Material 30

Figure 4.2 Example of Mvt II Rhythmic Context from Phrase 8 31

Figure 4.3 Sample Spectrograms of Oboe and Flute Instrument-Pitch and Vocal [ɜ:] timbre for Visual Comparison 32

Figure 4.4 Sample Instrument-Specific Timbre Assignment Chart 34

Figure 4.5 Sample of Phonetic Instrument 35

Figure 4.6 Notated Graphic Combining Rhythmic Construct and Instrument Pitch Assignments 36

Figure 4.7 Mvt II Phrase 13 Pitch and Rhythm Construct 38

Figure 5. 1 Sample Rhythmic Context from Mvt III Text Excerpt, Phrase 17 45

Figure 5. 2 Sample Spectrogram from Mvt III, Phrase 17 45

Figure 5. 3 Sample Phonetic-Event Map from Mvt. III, Phrase 17 46

Figure 5. 4 Sample Phonetic-Event Map from Mvt III, Phrase 17 47

Figure 5. 5 Sample Phonetic-Event Map from Mvt III, Phrase 17 48

Figure 5. 6 Phoneme Formant Pitch Relationship Chart 49

Figure 5. 7 Flow Chart for Mvt III 50

Abstract

Prologue to Sounden Horn is a phonetic composition for chamber orchestra. Phonetic Composition is the orchestration of speech sounds for standard acoustic instruments. Spectrograms of audio recordings are used as a basis for musical content. The primary elements derived from spectrograms are speech formants – their relative ratios, transitions, and amplitudes. Function is then established based on applications of phonetics and phonology to determine organizational strategies including durational parameters and orchestration tools. Sculpting this data into a musically aesthetic work, based on and resembling the original speech recording, is the ultimate goal of Phonetic Composition. This document details that process.

Chapter 1: Introduction

Phonetic Composition is the orchestration of speech patterns for standard instruments of the orchestra. Currently, the replication of speech patterns is of particular interest to the computer software, communications, and gaming industries (among others) due to the integration of advancing technologies in speech recognition software and text-to-speech synthesis. Viewing the orchestra as a type of synthesizer, and a musical score as a type of software program, the applications of text-to-speech synthesis have been modified in this document for their use as compositional and orchestrational techniques. This text-to-speech approach to composition, using the phonetic sciences and spectral analysis as a foundation, is referred to here as Phonetic Composition.

The research contained in this document is a continuation of research I began in 2012-13 while a master's student at Stephen F. Austin State University in Nacagdoches, Texas. As an extension, applicable results from that research – presented in my master's thesis – are included in this document. That thesis document is entitled *Perigee and Apogee, A Phonetic Composition for Chamber Orchestra* and is cited in the bibliography of this document. The data collected in this document specific to *Prologue to Sounden Horn* was acquired from sound recordings made prior to the initiation of research by external entities, and no human studies were conducted in its acquisition.

Previous notable speech-based compositions from historic literature include works by the composers Clarence Barlow, Jonathan Harvey, and Peter Ablinger. The methods and applications used by these pioneers of speech-based composition were not used in the processes of Phonetic Composition as prescribed in this document, though some similarities exist, discussed throughout this document.

Clarence Barlow's work in speech-based composition was coined *synthrumentation* by the composer, having composed *Im Januar am Nil* in 1981. Barlow's work is the most similar to Phonetic Composition, in that it orchestrates formant frequencies of speech sounds, transposed for performance by acoustic instruments[1]. It differs in many ways, including his orchestration of only vowel sounds, limiting the assignment of speech harmonics to scordatura-tuned string instruments, and avoiding aperiodicity. Barlow used computer software to analyze spectral patterns in speech recordings, which then calculated frequencies for output as MIDI note values. These MIDI note values were then assigned to each string instrument according to its scordatura tuning, allowing for harmonic tuning variances. This method allows for static representations of speech formants in a very detached, disjunct fashion.

Johnathan Harvey created *Speakings* in 2007. This piece uses computer software to map human speech formants (overtones or groups of overtones that characterize a sound[2]) onto amplified instruments during live performance, giving the illusion of "speaking instruments."[3] This is not an orchestrational approach to speech patterns and their formants in the same manner as Phonetic Composition. Harvey's applications depend on electronic and computer-aided effects to manipulate the instrument sounds via extensive signal processing.

Peter Ablinger created a series of compositions for mechanized piano and speech recordings in 2009, which he entitled *Voices and Piano*. Using computer software to parse speech frequencies from recordings of famous historic speeches, such

[1] Bob Gilmore, "Clarence Barlow Interview," Paris Transatlantic Magazine, http://www.paristransatlantic.com/magazine/interviews/barlow.html (accessed March 18, 2017).
[2] Peter Ladefoged. *Vowels and Consonants*, 5th ed. (Los Angeles: Blackwell Publishing, 2005), 200.
[3] Jonathan Harvey, "Jonathan Harvey," Jonathan Harvey, http://www.vivosvoco.com/index.html (accessed March 18, 2017).

as those orated by JFK and Martin Luther King, Jr., Ablinger interpolated this data to servomotors which control the hammers of a mechanized piano, which then reproduces the speech patterns on the 88 keys available. This results in pseudo-speech sounds in real time. No written music and no interpretive analysis is necessary for this application, as the sound recording feeds directly into the computer software which triggers the servos connected to the piano.[4] No human interaction is required in the execution of such a piece.

Multiple forms of analysis and synthesis have been employed in speech research labs since the late 1800's. Processes such as Sinewave Synthesis[5], imaging and modeling techniques[6], Automatic Speech Recognition[7], the Source-Filter Theory[8], the Computational Segmentation Method[9], Linear Predictor Coefficient (LPC) analysis using digital filters, and windowing using Fourier Analysis[10] represent a sampling of techniques. With varying degrees of success, certain elements of each have been found to produce intelligible speech, or relative approximations thereof.[11] Since different

---

[4] Aljoscha Hoffman, "Peter Ablinger," Aljoscha Hoffman, http://www.ablinger.mur.at/ (accessed March 18, 2017).
[5] Philip Rubin, Robert E. Remez, Jennifer Pardo, "Sinewave Synthesis," Haskins Laboratories, Yale University, http://www.haskins.yale.edu/research/sws.html (accessed March 18, 2017).
[6] Brad H. Story, "Using Imaging and Modeling Techniques to Understand the Relation Between Vocal Tract Shape to Acoustic Characteristics," Department of Speech and Sciences, University of Arizona, http://sal.shs.arizona.edu/~bstory/StorySMAC03.pdf (accessed June 11, 2013).
[7] Massachusetts Institute of Technology, "Automatic Speech Recognition," Department of Electrical Engineering and Computer Science, http://1337institute.com/c/62/automatic-speech-recognition (accessed March 18, 2017).
[8] Indian Institute of Technology, Bombay, "Source-Filter Theory of Speech Production," Indian Institute of Technology, Department of Humanities and Social Sciences, http://wwwiitb.ac.in/courses/HS435/Acousticslecture_files/frame.htm (accessed June 7, 2013).
[9] Jordi Janer, "Syllabling on Instrument Instrumentation: Case Study and Computational Segmentation Method," Music Technology Group, Pompeu Fabra University, Barcelona, http://uni-graz.at/~parncutt/cim07/CIM07%20Proceedings/CIM07)Janer-Penalba_Syllabling%20on%20instrument%20imitation.pdf (accessed June 11, 2013).
[10] Peter Ladefoged, *Elements of Acoustic Phonetic* (Chicago: The University of Chicago Press, 1996)
[11] Rubin, "Sinewave Synthesis," http://www.haskins.yale.edu/research/sws.html (accessed March 18, 2017).

disciplines have produced varied approaches to the speech synthesis problem, it seems logical to think that more than one approach to speech synthesis for orchestra could be devised. Incorporating elements of spectral analysis, the phonetic sciences, and phonology, this document prescribes three approaches to Phonetic Composition.

For the purposes of this study, Phonetic Composition is represented by three distinct methods, referred to here as Spectral Mass Technique, Sound Color Technique, and Phonetic-Event Technique. The prescribed methods of Phonetic Composition are all based on the same fundamental principles of phonetic science, but each is unique in its focus and implementation. Each has its own set of guidelines. These guidelines serve as parameters which differentiate each discipline.

The three approaches to Phonetic Composition, as mentioned above, are applied in this document to excerpts from an audio recording of composer Marvin Lamb orating program notes to his own composition entitled *Sounden Horn* prior to its premier performance at the University of Oklahoma in October of 2015.[12] The excerpts from this recording have been divided as a textual basis for the three movements of the composition represented in this document, *Prologue to Sounden Horn*. These three movements each represent one phonetic composition technique as named above. Each of these movements will be addressed in their corresponding chapters.

The applications of these approaches are represented as follows:

Mvt. I. Blow This Sounden Horn	Spectral Mass Technique
Mvt. II. Weird Haunting Bizarre	Sound Color Technique
Mvt. III. No One Would Perish	Phonetic-Event Technique

[12] Marvin Lamb, "Sounden Horn" (presented at the University of Oklahoma, October 2015).

Chapter 2: Approaches to Orchestration of Phonetic Techniques

In this chapter, the fundamental elements of Phonetic Composition are described in general terms for overall comparison of applied techniques. Here the commonalities between the three methods will be addressed. Subsequent chapters will address the elements and applications unique to each. The explanations found here will suffice to give a general overview of the fundamental processes used for the creation of *Prologue to Sounden Horn*.

The three approaches are partially defined by force (ensemble size). The term "force" refers to the number of voices used to represent a phoneme or sequence of phonemes in Phonetic Composition. The term "voice" (horizontal musical lines following a melodic shape and motion) may refer to a single instrument or a combination of instruments, acting in tandem to represent one formant zone. The advantage of using these terms is their flexibility when speaking generically about the application of similar methods to differing techniques of phonetic orchestration.

Based on similar existing speech research methods, such as spectral analysis, phonetics, and phonology, these techniques – Spectral Mass, Sound Color, and Phonetic-Event – share certain fundamental elements. They diverge, however, in distinct and defining ways. Spectral Mass Technique is characterized by a relatively large number of voices simultaneously coloring the sound spectrum, using the harmonic series as a fundamental construct, and intensifying around formant zones. Sound Color Technique uses as few as one or two voices simultaneously, focusing on the timbre of each individual instrument pitch to represent speech sounds. Phonetic-Event Technique is unique in that it creates a prescribed set of sonic moments, in the form of an

orchestration database, which are then sequenced as called for in the text. These sonic moments are not bound to harmonics of a fundamental frequency or individual instrument timbre. They are orchestrated in a manner that represents the spectral elements of speech sounds in interpretive ways. These approaches are described in more detail in subsequent chapters.

Several steps of the spectral, phonetic, and phonological analyses processes are common to all three techniques. These steps include selecting source material, constructing a rhythmic context, creating a spectrogram, assigning instruments to frequency zones, creating a phonetic map, and assigning percussive effects. After data from these steps is secured, the information is applied in different manifestations, based on the technique prescribed for a certain movement, section within a movement, or specific component within a section of a movement. Techniques can be layered effectively, using more than one technique simultaneously within the context of a larger scheme to enhance the effectiveness of the orchestration when deemed appropriate and desirable.

The source recording was submitted for use in this research by Dr. Lamb and was not made for the purposes of this research. The source recording was divided into short excerpts for use as sonic material for the content of *Prologue to Sounden Horn*. These excerpts were assigned to three separate movements of the piece. Formal analysis of the recordings stemmed from this initial selection process. The text for each movement will be provided in subsequent chapters.

In the construction of a rhythmic context, each line or phrase of text from the audio recording was analyzed for metric scheme, tempo, and rhythm. Metric schemes

were established based on the ratio of stressed syllables to weak syllables. Tempos were established based on the prescribed metric schemes of each phrase and the pace of those stresses on the source recording. At times a single meter and/or tempo could successfully represent a phrase. At other times multiple meters and tempos were employed to accurately represent the pacing of a phrase. Rhythms were determined within the metric scheme and tempos were established to represent the pacing of each word. This rhythmic basis was referenced in the application of the overarching rhythmic schemes prescribed and scaled for sonic renderings ranging from real time to up to thirty-two times the original duration of source material. An example is found in Figure 2.1 below. This step will be detailed per technique in corresponding chapters to follow.

In Speech Analyzer v.3.1.0, formant tracks are calculated using a bank of first order LPC tracking filters.[13] Spectrogram creation is based on a Fast Fourier transform (FFT) – an efficient algorithm which decomposes a sound into a spectrum of its frequency components. In Speech Analyzer, the FFT is calculated using 4x oversampling and linear interpolation. A frequency adjustable Hanning window is applied to the data before the FFT is calculated.[14] For the spectrograms used in this research, the Display Frequency was set to 8150 Hertz, the thresholds were set to a range of -3.0 dB to -14.0 dB, and the spectrum resolution was set to "Wide Band Filter" mode of 300 Hertz.

Each image was analyzed for formant zones, formant ratios, and formant transitions. Textual representations of the source material were added subsequently in

[13] Kamron Mustafa and Ian C. Bruce, "Robust Formant Tracking for Continuous Speech with Speaker Variability," *IEEE Transactions on Audio, Speech, and Language Processing 14*, no. 2 (March 2006).
[14] SIL International, Speech Analyzer v.3.1.0, copyright 1996-2012.
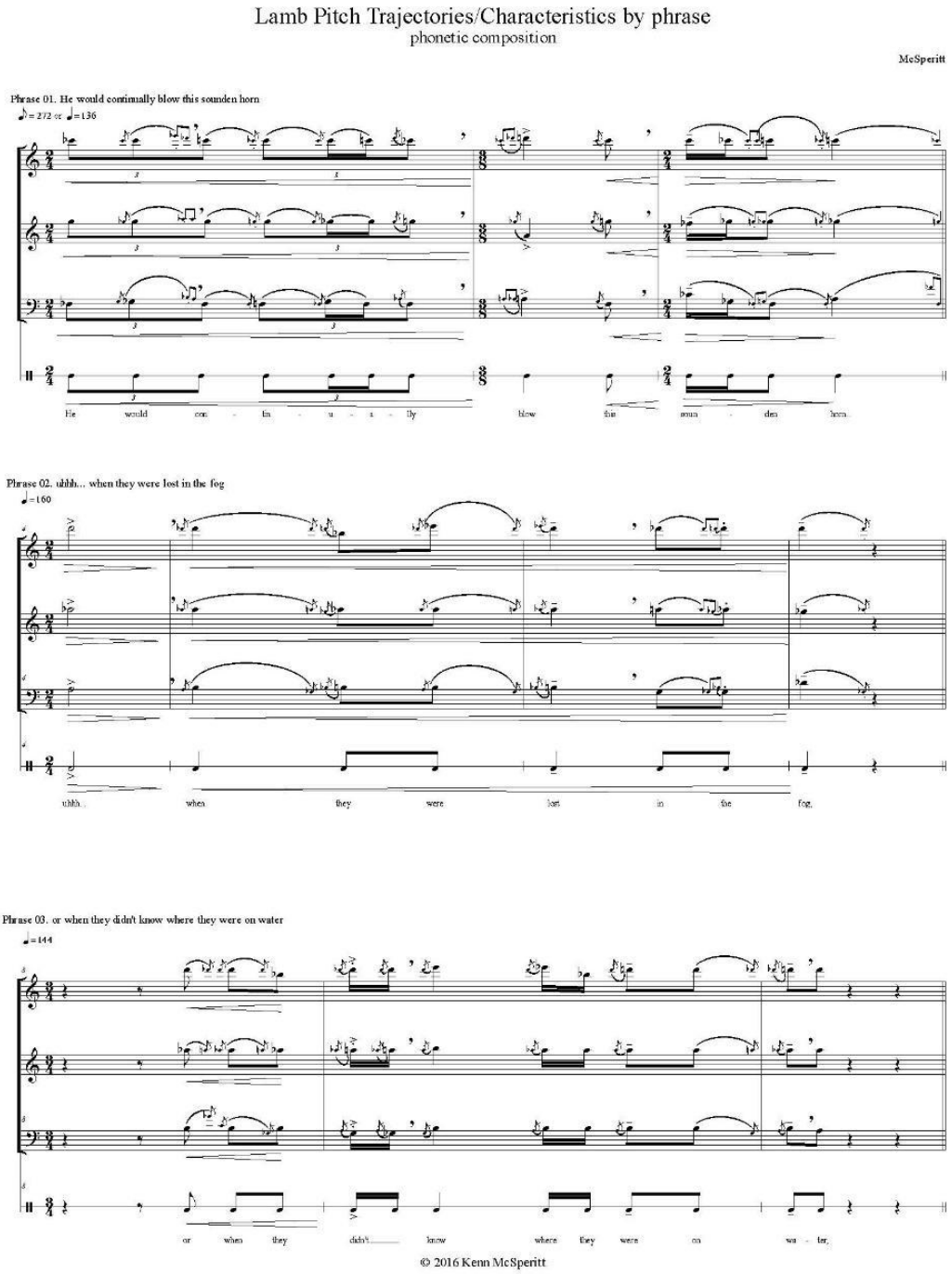
alpha-numeric symbols as well as in phonetic transcriptions to prepare for the creation of a phonetic map, discussed in more detail below.

The two major categories of speech sounds are vowels and consonants. The sub-categories of speech sounds are based on manner of execution. The vowels are divided into three groups, based on tongue position in the mouth at the time of articulation: high vowels, low vowels, and diphthongs/triphthongs. The consonants are subcategorized based on their manner of articulation: stops, fricatives, and approximants.[15] The major factor in distinguishing vowel sounds one from the other is the relationship (ratio) of the three primary formants (formants 1, 2, and 3). The fundamental formant (f0) does not affect vowel color. It is responsible only for the lowest-sounding frequency (or pitch). The overtone relationships above the fundamental are responsible for vowel distinctions. Consonants are the result of the physical proximity and manipulation of mechanisms in the vocal tract. Details of speech sound formation is not within the scopr of this document and can be researched in thorough texts such as Ladefoged.[16]

Characteristics of speech sounds are analyzed and dissected for their most identifying features. These features constitute what amounts to “building blocks” or “bricks” in the construction process. To use a metaphor, the bricks are not the house. In order to be perceived as a house, to serve the function of a house, the bricks must be modified for context, placed properly, and linked together in an organized, functional, and preferably aesthetic manner.

[15] Ladefoged, *Vowels and Consonants*.
[16] Ibid.

Figure 2. 1 Establish a Rhythmic Context



In the same way, phonemes are not language. Phonemes in isolation have no context. Phonemes in the abstract carry insufficient linguistic information. Phonemes

linked out of proper sequence are, at best, confusing. If the stresses within a phrase are illogical, meaning is difficult to grasp and the listener must recreate the phrase in their mind, slowing the communication process.[17] Phonemes which are not linked using characteristic transitions distract from their linguistic context.[18] To be perceived as language and serve as a viable form of communication, phonemes must be organized, functional, and preferably aesthetic.[19]

So, why is phonology so important to the process and end result of Phonetic Composition? First, phonology is the science of understanding, or making sense of, what is heard. The brain interprets hums, whistles, buzzes, hisses, and clicks (which in and of themselves carry no linguistic function), combines them into useful data, and links their combination to ideas or images in the mind.[20] In order to successfully create a composition that takes these same elements, albeit from a different sound source than a human vocal tract, and recreates their linguistic effect in the human mind, it is necessary to understand how the mind receives, perceives, and processes sound.[21]

Moving forward in the steps of Phonetic Composition common to all techniques, established above, is the assignment of instruments to frequency zones. In preparation for the creation of *Prologue to Sounden Horn*, the instrumentation was established as a chamber orchestra, to include two French horns and percussion. Chamber orchestra was selected as the preferable medium to accommodate the demands of the three orchestrational approaches in Phonetic Composition – Spectral Mass Technique, Sound

---

[17] Kathryn LaBouff, *Singing and Communicating in English: A Singer's Guide to English Diction* (Oxford: Oxford University Press, 2008), 18.
[18] Robert Mannell, "TTS: An Overview of Concatenative Approaches to Speech Synthesis," Macquarie University, http://clas.mq.edu.au/speech/synthesis/tts_concat/index.html (accessed March 27, 2017).
[19] Ibid.
[20] Robert E. Remez, "Sine-wave Speech," Department of Psychology, Barnard College, Columbia University, http://www.scholarpedia.org/article/Sine-wave_speech (accessed March 13, 2017)
[21] Ladefoged, *Vowels and Consonants*, page 109.

Color Technique, and Phonetic-Event Technique. The strings facilitated the extremes of register required to represent upper formants via harmonics, as well as a saturation of the spectrum carried by similarity in timbre and resonance for the purposes of Spectral Mass Technique. The winds provided the necessary timbral diversity to represent vowel colors used as the pitch-determinant in Sound Color Technique. The range and agility provided by the combination of these instrument families, supported by percussion, served the dynamic contrasts required of Phonetic-Event Technique. These applications will be discussed in Chapters 3, 4, and 5.

The instruments were broken down by tessituras and compared to frequency ranges as represented by formant zones in the spectrograms, discussed above. An example of this analysis and distribution of instruments by formant zones is found in Figure 2.2 below.

Figure 2. 2 Instrument Assignments by Harmonic Zones

Instrument	Harmonic Range	Pitch Area (in concert pitch)
Piccolo	24-32	C6-C8
Flute	6-26	C4-C7
Oboe	5-22	C4-C6
Bassoon	2-8	F2-A#4
Bb Clarinet	3-17	D3-B5
Horn	3-12	C3-C5
Violin	24-40	C6-C9
Viola	3-32	C3-B7
Cello	2-32	C2-A#7
Dbl Bass	2-6	E1-C4

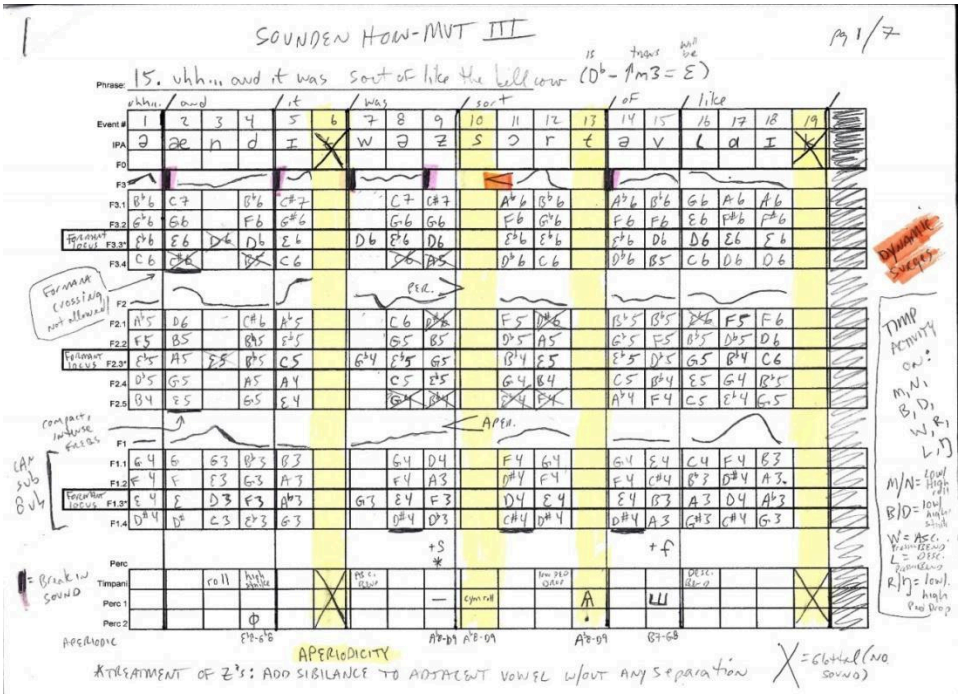
The next step in the process is to create a phonetic map. The phonetic map serves as a timeline from which to draw details about the most important elements of speech, combining data from spectral, phonetic, and phonological analyses into one

resource document. Here, the formant frequencies are converted to pitch, formant transitions are diagrammed, formant amplitude is notated, phrases are represented for flow and form, and percussive effects are assigned based on consonant considerations. This valuable information is coalesced into the phonetic map. The map serves as a sort of "tool box" or "color palette" from which to begin the compositional construction of speech sounds in a phonological sequence. A sample page from a phonetic map is demonstrated below in Figure 2.3.[22]

The final stage common to all Phonetic Composition techniques employed in this document is the assigning of percussive instruments, implements, and attack-types to represent the various consonants. The use of percussion is optional and used here only to enhance the consonants as represented in the music. Referencing back to earlier in this chapter, it was stated that all speech sounds carry a spectral signature, even the aperiodic sources (static noise). The consonants can all be represented by wind and string instruments to some degree. The use of percussion in the creation of *Prologue to Sounden Horn* was based on musical considerations to enhance certain moments within the piece. Percussion instruments were either omitted, used as support for wind and string instruments, or used in isolation to represent consonants strictly based on musical

Figure 2. 3 Sample Phonetic Map from Mvt. III of *Sounden Horn*

[22] Certain documents which served important roles in the compositional process of *Prologue to Sounden Horn* are included as figures in this paper and have been preserved in their original state for authenticity and as an insight to the logic under-girding of the process. Figure 2.3 is one such document.



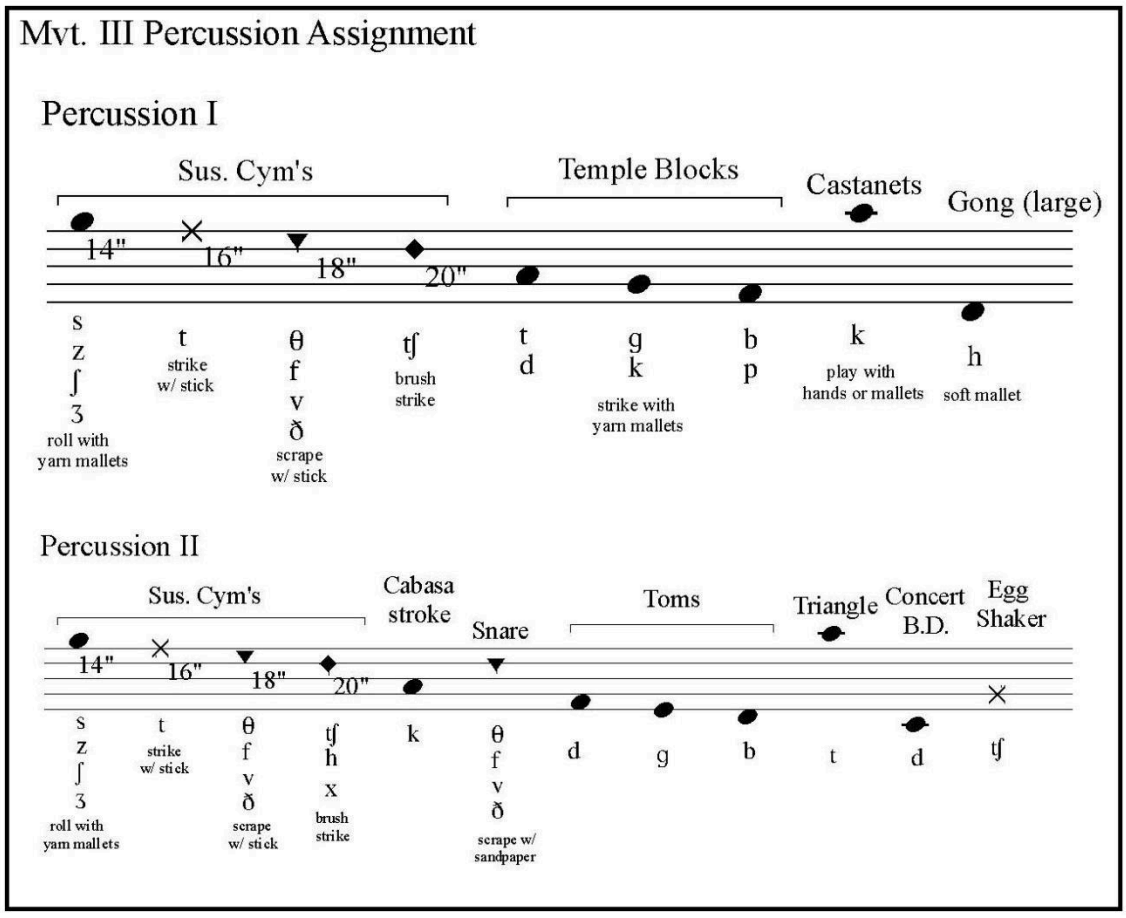
aesthetics. One of several percussion keys used for this composition can be seen below in Figure 2.4.

Other characteristics of *Prologue to Sounden Horn* which serve an overarching function should also be noted at this point. When string harmonics are present, they always represent frequency ratios between formants 4 through 7. These are the very highest frequency ranges that acoustic instruments can produce within audible range to humans. These formants assist in the identification of voice timbres.[23] These high-energy areas above the three primary formants are unique to each individual. Pitch considerations are based solely upon content found within the spectral signatures of the phonemes themselves. Any aural instance of apparent "tonality" is purely coincidental and is embedded within the formant relationships themselves. The subjective choices made by

[23] Georgia State University, "Timbre," Department of Physics, http://hyperphysics.phy-astr.gsu.edu/hbase/Sound/timbre.html (accessed March 18, 2017).

the composer are components such as durational ratios, transpositions of sectional material, instrumentation, number of voices present at a time, and which formants to represent at any given moment. These general considerations are largely true for all three techniques of Phonetic Composition.

Figure 2. 4 Percussion Key



Chapter 3: Spectral Mass Technique:
An Analysis of Mvt. I. Blow This Sounden Horn

The opening movement of *Prologue to Sounden Horn*, 'Blow This Sounden Horn,' is based on excerpts from verbal program notes to a performance of Marvin Lamb's *Sounden Horn*, as delivered by the composer. The Phonetic Composition technique primarily used in this movement is Spectral Mass Technique.

The distinctive elements of Spectral Mass Technique (SMT) are as follows:

- Large number of voices (sound saturation)
- Harmonic series as a primary character of the fundamental (F0), applied in conjunction with formant characteristics of corresponding speech sounds
- Abrupt changes in pitch without smooth, connective glissandi
- Harmonic-voicing component integral to character
- An allusion to the F0 without directly quoting it in the phonetic component

The fundamental processes for extracting data and constructing compositional tools for all three techniques employed in this document were covered in Chapter 2. Therefore these steps will not be discussed at length here, though piece-specific information will be presented as is appropriate.

The excerpt selected from the source recording contains the following text, which is represented below accompanied by its phonetic transcription.

1. He would continually blow this sounden horn…
   hi wud kən'tɪnjuəli blo ðɪs saundən hɔrn

2. uhhh… when they were lost in the fog,
   ə hwɛn ðe wɝ lɔst ɪn ðə fɑg

3. or when they didn't know where they were on water,

ɔr hwɛn ðe dɪdənt no hwɛr ðe wɝ ɔn wɔtɚ

4. or when they were caught in a snow storm,
   ɔr hwɛn ðe wɝ cɔt ɪn ə sno stɔrm

5. or when all manner of things would happen that would
   ɔr hwɛn ɔl mænɚ əv θɪŋz wud hæpən ðæt wud

   cause them to get caught in the rivers in the dark
   kɔz ðɛm tu gɛt kɔt ɪn ðə rɪvɚz ɪn ðə dɑrk

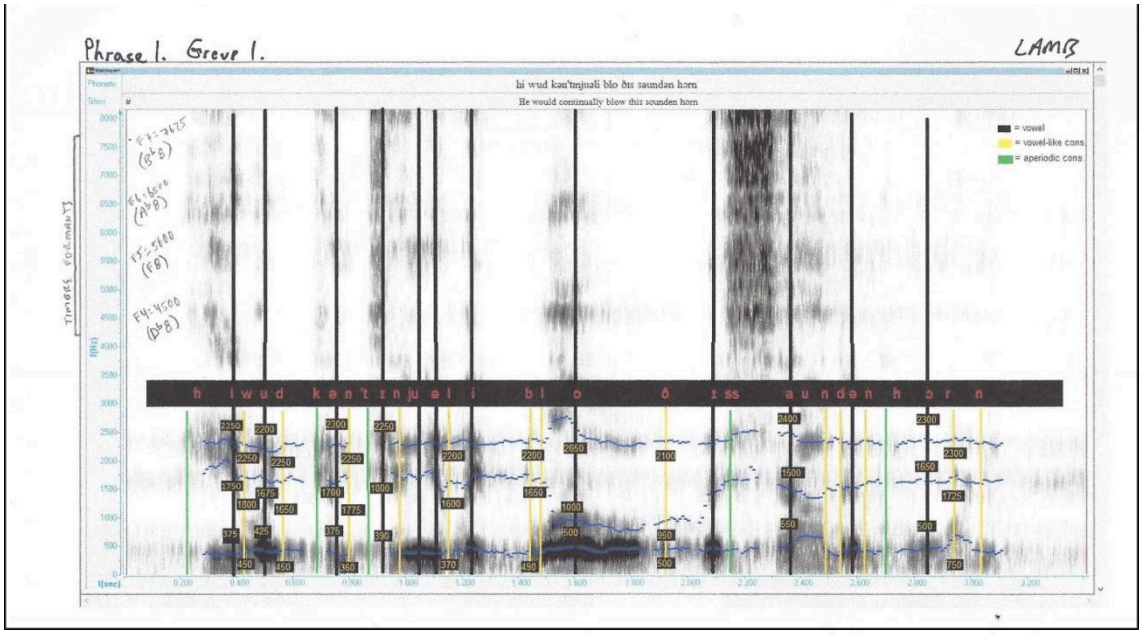
6. not knowing what the next second might bring.
   nɔt noɪŋ hwɑt ðə nɛkst 'sɛkənd mɑɪt brɪŋ

The text was transcribed using the International Phonetic Alphabet and *A Pronouncing Dictionary of American English* by John Kenyon and Thomas A. Knott.[24] The transcription was modified according to guidelines found in Kathryn LaBouff's *Singing and Communicating in English: A Singer's Guide to English Diction*.[25] These modifications were found to be significant in the mapping process, as they effectively coordinated the information provided by the spectrograms with the information gathered from Ladefoged's *Vowels and Consonants* (referenced earlier) about human speech characteristics.

The rhythm scheme was then derived from the source recording according to stresses and pace (See Figure 2.1). Spectrograms of each phrase within this excerpt were then created for the purpose of formant analysis. A sample spectrogram is found below in Figure 3.1.

[24] John S. Kenyon and Thomas A. Knott, *A Pronouncing Dictionary of American English* (Springfield, MA: Merriam-Webster, Inc., publishers, 1953).
[25] Kathryn LaBouff, *Singing and Communicating in English: A Singers Guide to English Diction* (Oxford: Oxford University Press, 2008).

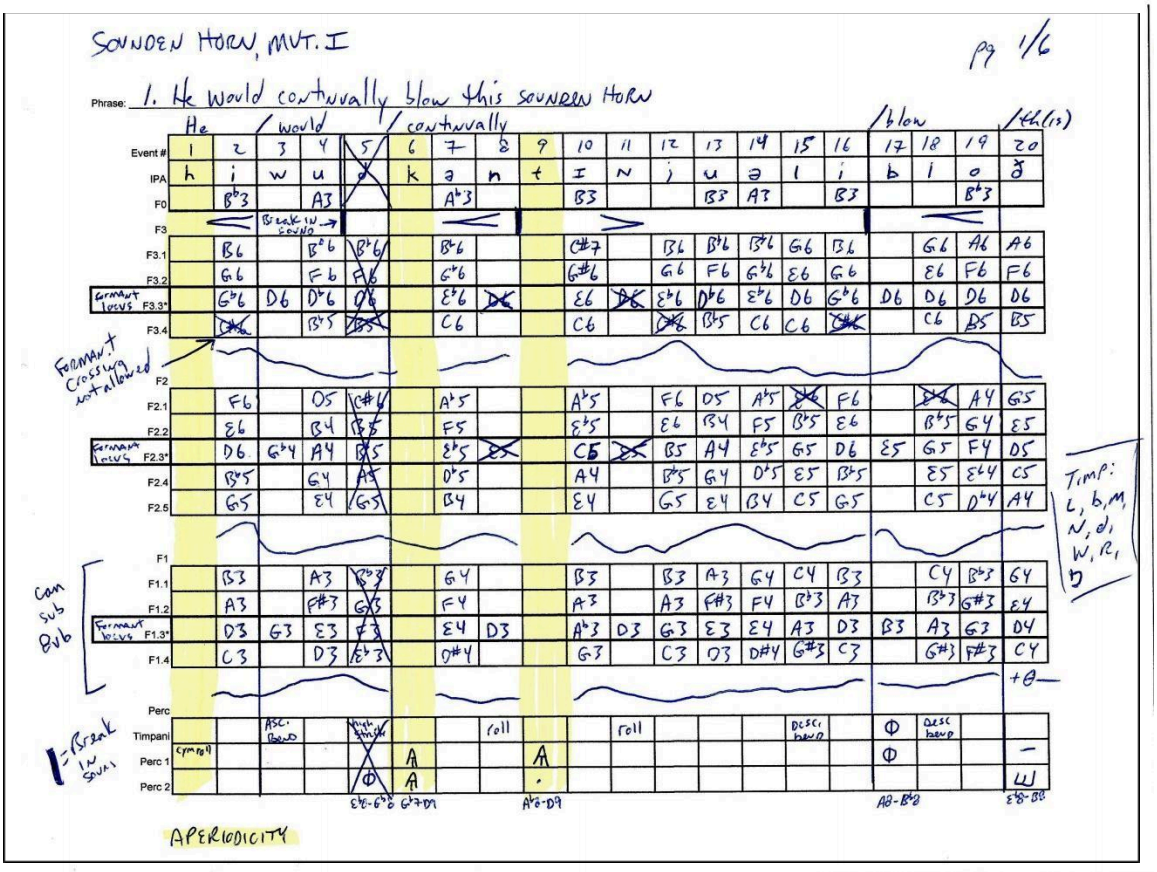
Figure 3. SEQ Figure_3. * ARABIC 1 Sample Spectrogram of Excerpt from Mvt I



The spectrogram was analyzed for formant properties such as frequency (indicated here by the black-box numbers), amplitude (represented in shades of grayscale to indicate differing degrees of intensity), and transitions (indicated by the wandering vertical blue lines). This information was then used in the assignment of pitch structures for each formant voice for the purposes of musical line and contour, pitch relationships between voices, and dynamic contrast with respect to phrasing.

The information extracted from the spectrograms was then used to create a phonetic map (described in Chapter 2). A sample of the phonetic map used as a sort of "infrastructure" for Mvt. I. is found below, in Figure 3.2.

Figure 3. SEQ Figure_3. * ARABIC 2 A Sample Phonetic Map Used in Mvt I

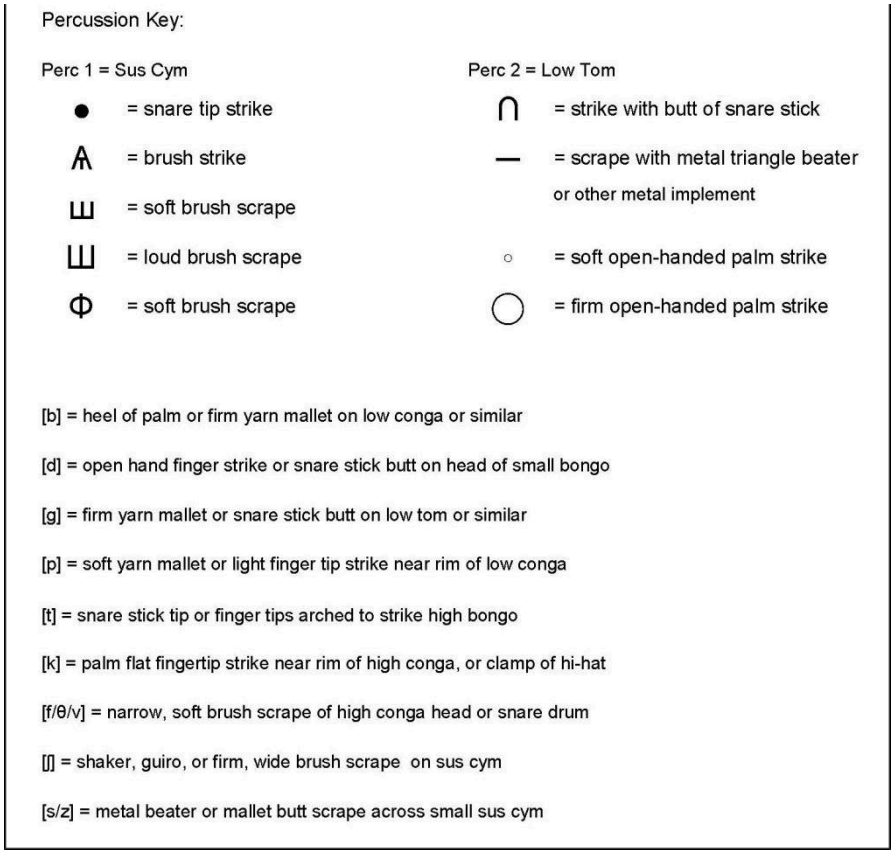


The percussion key for all phonetic maps used in this composition can be found below in Figure 3.3.

The phonetic map gives a fundamental basis for the overall infrastructure of the pitch construct found within each phrase. Phonetic mapping is a process original to research in Phonetic Composition.[26] It serves as a system of organization, from which compositional decisions regarding pitch structure, flow, and phrase shaping are based. It is identical between the diverse techniques employed in Phonetic Composition as detailed in Chapters 3 through 5 of this document. Given that the process holds across platforms, the discussion of this process will be detailed here and briefly referenced in subsequent Chapters 4 and 5.

[26] Kenneth E. McSperitt, "Perigee and Apogee: A Phonetic Composition for Chamber Orchestra" (master's thesis, Stephen F. Austin State University, 2013).

Figure 3. SEQ Figure_3. * ARABIC 3 Percussion Key Used in Phonetic Maps



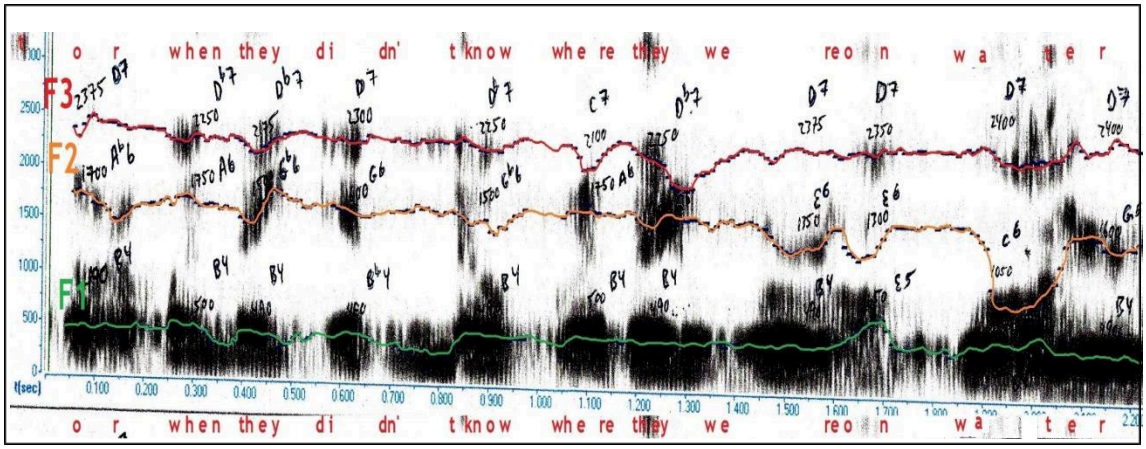
Each page of the phonetic map is labeled by movement and phrase(s) represented. Above the table, the alphanumeric representation of the text analyzed is written above the event number where it initiates. Beginning at the top of the formal table, the event number is listed from left to right beginning with event number one. The term "event" simply refers to a vertical location on the spectrogram where the most stable version of the corresponding phoneme resides. Next is the IPA symbol (the phonetic alphabet sanctioned by the International Phonetic Association[27]), indicating the pronunciation of the phoneme in question. Each column represents one phonetic event,

[27] Knott, xvi.

coinciding with a single phoneme. Below the IPA symbol is the fundamental frequency (F0) derived from the original source recording. Since we speak in microtones and not semitones, these pitches reflect the closest pitch to that sounded in the recording. The F0 serves as the basis of pitch and inflection for the fundamental, but does not directly affect the formant loci, as formant loci do not follow the F0 as it oscillates through pitches, but reflects the overtone series created by the shaping of the many components of the vocal tract.

Next, in the gap between the F0 row and the F3 formant block found beneath it, the amplitude scheme (or dynamic motion) is expressed in shapes such as crescendos and decrescendos. Breaks in sound are also designated between phonemes via thick, dark lines. These breaks indicate complete stops in sound, whether due to actual linguistic phrasing, or in conjunction with cessation of sound affiliated with silent consonants or closures of the vocal apparatus in the midst of a group of phonemes not at the ends of phrases. An example of this can be seen in Figure 3.4, at the conclusion of the word "would" and before the word "continuously." The final *d* in "would" is a stop

Figure 3. SEQ Figure_3. * ARABIC 4 Spectrograms of Mvt I, Phrase 3



(phonetically speaking), in that there is complete closure before the subsequent word "continuously." However, this is not the only factor in play, as the consonant *d* in "would" is also a final stop and does not actually sound in the source recording. This is a phenomenon common in human language, where final stop consonants such as *t*'s and *d*'s are implied rather than expressed. In other words, they are not sounded, though our brains instinctively recognize the word in context regardless.

The next three blocks (in descending order) represent the lower three formant zones – the locus and nearest surrounding harmonics. These formant loci are derived from the spectrograms, estimated for location based on the center of the most intense location of each formant. The loci are not stationary, as can be seen in Figure 3.4 represented by the blue horizontal, wandering lines. Formant loci move continuously, even in vowels, which are the most stable phonemes. The boxes in extreme positions are subject to cancellation, as witnessed by the X's within boxes previously filled with pitch information. Cancellation is most often related to instances of formant crossing, that is, when two adjacent formants overlap (i.e. phonetic event #2, where the lowest formant voice in F3, located in the row F3.4, is lower than the F2.1 pitch). The decision on which pitch to eliminate has to do with formant strength. In this case, formant F2 has higher energy than F3, so the lowest F3 pitch is eliminated. Formants are compared in the spectrogram for energy hierarchy where voice crossing is concerned. Elimination of crossing voices is an executive decision of the composer to allow for clarity of voice lines in the composition.

Between the formant blocks, formant transitions are indicated via sketched shapes, descending and ascending across strings of phonemes. These transitions are critical to intelligibility, as the direction, trajectory, and rate of change function in tandem between formants to create the perception of vowel and voiced consonants in speech sounds. These transitions carry the linguistic characteristics associated with intelligible speech more than the formant pitches alone.[28]

The lowest block contains percussion assignments for each appropriate phonetic event. Figure 3.3, above, details the symbols used for each consonant, both periodic (voiced, or regular oscillations of waveforms) and aperiodic (voiceless, or random pitch generation). The percussive effects associated with each consonant are selected subjectively based on the audible similarity between them. These assignments can also vary based on the needs of the music, with considerations concerning thickness of texture, intensity of sound, and timbral character serving as the determining factors.

Once these fundamental characteristics are established and documented, the phonetic event map serves as a reference for the composing stage. Choices can be made regarding what information to use and where, based on artistic license and aesthetic considerations. Not all formants need to be represented at all times, nor do all consonants. Consonants can be represented by periodic instruments alone, aperiodic instruments alone, or a combination of the two. One formant can carry the melodic motive, or two, or three. These parameters are subject only to the discretion of the composer. However, all material is derived from information found within the phonetic event map.

[28] Ladefoged, *Phonetic Data Analysis*, 105.

Characteristics in Spectral Mass Technique which are unique to this method include the sudden shifting of pitch areas without connective glissandi or stepwise motion, saturation of formant zones or the sound spectrum at large, and pitch selection based on the F0 found for each phoneme in the source recording. The following description of compositional choices explains this process, which is unique to Spectral Mass Technique. As an example, discussion of the construction of Phrase 3 below will serve as a model for the methods employed.

Phrase 3, "Or when they didn't know where they were on water," which can be found in the score (found in Appendix B) in Mvt. I., ppg. 12-16, mms. 60-77). This section transitions in character from that of "fog" in the previous phrase (mms. 39-59) to "water" in Phrase 3. Certain parameters were predetermined, as is the case with each phrase throughout the work. These parameters include:

- formants will be represented in the bassoon (F1 at 8vb/15vb), oboes (F2 at 8vb) and flute/piccolo (F3 at 8vb)
- wind trills and flourishes used to represent "water"
- non-formant-carrying winds along with thick string orchestration represent vowel-like (periodic) consonants
- Percussion representing all consonants
- Timpani active on phonemes *n, w,* and *r*
- One iteration at 24:1 augmentation ratio for an approximate duration of 54 seconds

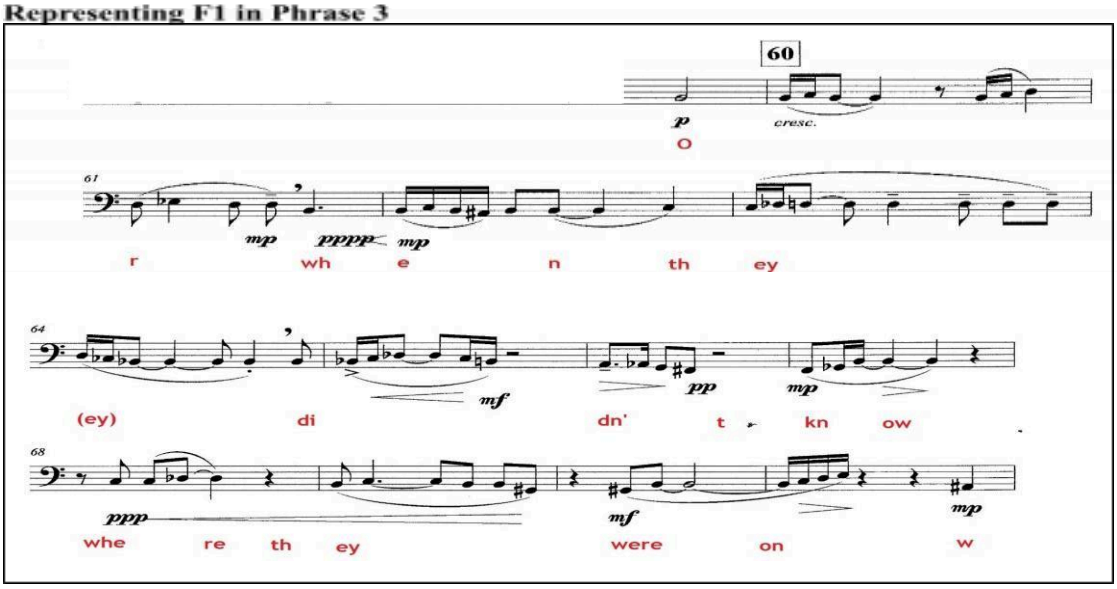
The orchestration considerations expressed above are specific to Phrase 3. Each phrase within the piece receives its own unique phrase characteristics. These choices are subjective and serve the aesthetic intentions of the composer.

In the construction of Phrase 3, and in keeping with the predetermined orchestration parameters, bassoon (representing F1) oscillates around B2. Starting at a *piano* dynamic level, a crescendo reflects the growing intensity of F1 from phonemes *o* at m. 60 (which is relatively static with regards to pitch) into *r* of "or" at m. 61, descending at the word "when." Figure 3.4 above shows the spectrogram of this phrase. This is accompanied in Figure 3.5 by a notated representation of the bassoon line representing F1 in this section.

Energized by rapidly ascending chromatics in m. 63 into the word "they" (especially noticeable in F2 represented by the oboes), the line dips abruptly, breaking sound continually prior to the word "didn't." Formant F1 begins its ascent from Bb2 in m. 65, descending prior to the advent of the second *d* in the word "didn't," increasing in energy across the ensemble. In m. 68, F1 continues its descent after the break in sound after the word "didn't." An increase in frequency is seen going from the *kn* of "known" into the phoneme *o* in m. 67, along with a surge in intensity, which subsides quickly. Rests lead into m. 68, where the phoneme *wh* of "where" occurs, followed by an ascending line. After the break between the words "where" and "they" at the end of m. 68, a subtle arcing line grows in intensity. A more stressed word "were" in m. 70 is intensified in energy and ascends in pitch, moving to E3, significantly higher than previous phonemes, as it merges with the word "on," followed by the trailing phoneme *n* to complete the motion. In m. 72, the syllable *wa* of "water" is energized, represented

by the increased rhythmic motion. A significant break between *wa* and *ter* of "water" is represented by the two beats of rest, followed by the intensified syllable *ter* in m. 74, which peaks at m. 75, subsequently trailing off toward m. 77.

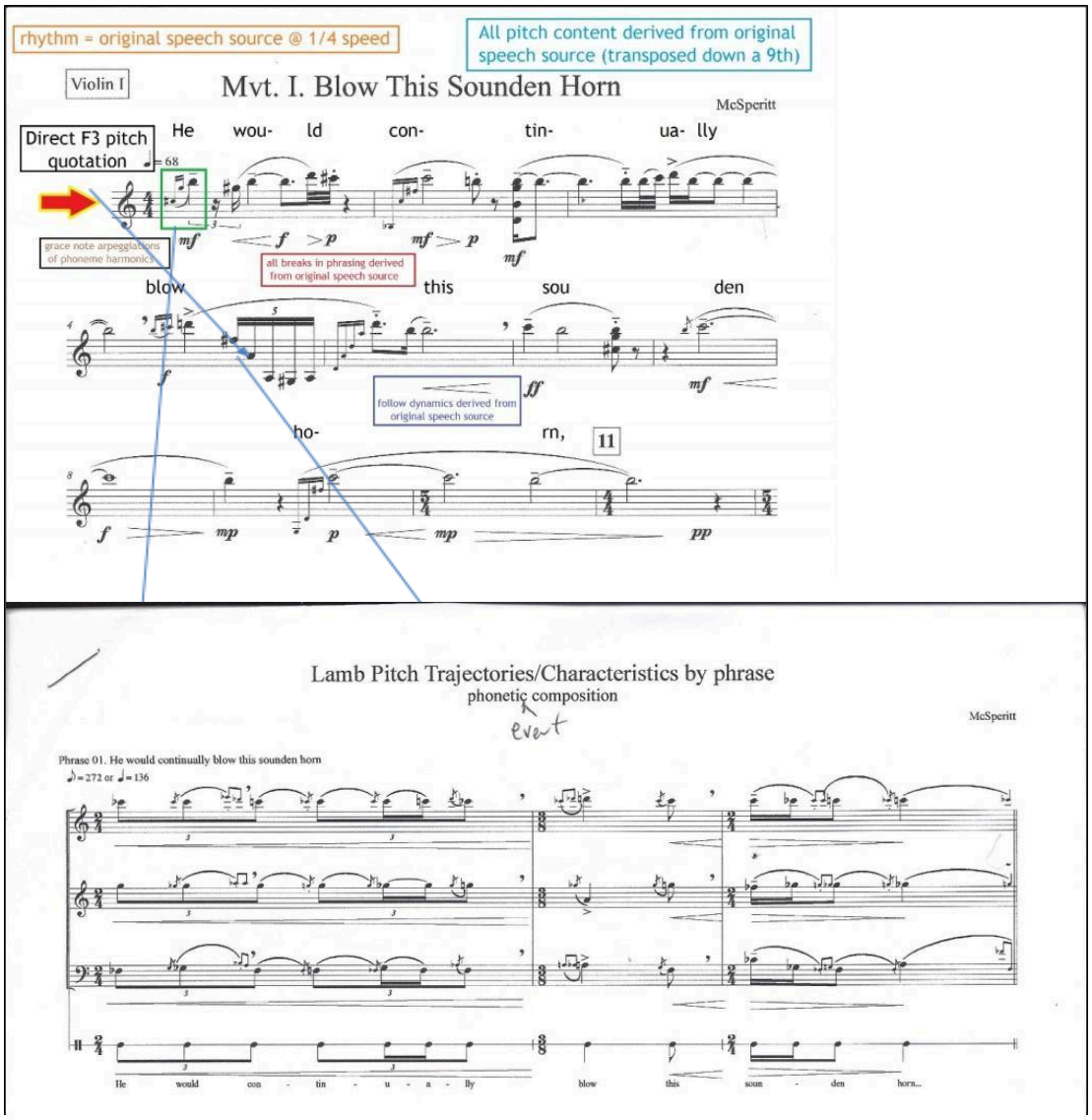
Figure 3. SEQ Figure_3. * ARABIC 5 Notation of Bassoon Line Representing F1 in Phrase 3



When using one solo instrument to represent more than one formant, arpeggiations are used to outline multiple pitch levels, as seen in figure 3.6 below.

The process for orchestrating F1 (as witnessed in the description above) is consistent for all formants, throughout not only Phrase 3, but the entire composition *Prologue to Sounden Horn*. The major differences between the techniques of Spectral Mass, Sound Color, and Phonetic-Event – mentioned earlier in this document – is the process for which pitches are assigned. All other considerations, such as phrasing, note durations, etc., hold true from one technique to the next.

Figure 3. Sample of Arpeggiations Used to Represent Multiple Formants in One Instrument



The form in Mvt. I is primarily through-composed and follows the text with only minor deviations to create musical coherence. In Sound Color Technique (SCT), the process of selecting pitch structures is based on the timbral signatures of individual instrument pitches as they compare to the timbres of phonemes, as determined by spectral analysis and audition, which will be discussed in Chapter 4. In Phonetic-Event

Technique (PET), the formants and their transitional relationships are the primary compositional considerations. Like SMT, the formant loci are used as the basis for frequency interval ratios; however, in PET, transpositions other than octaves are not out of character, transitions are much more freely interpreted, and repetition of phonemes, syllables, words, and phrases are used frequently to establish thematic repetition and consistent rhythmic gestures. This technique will be more thoroughly addressed in Chapter 5.

Upon completion, the piece was reviewed according to guidelines prescribed in Chapter 2. Exceptions to the guidelines were used when the results would have yielded unreasonable demands of the performers and/or would have not been detectible within the orchestration.

In conclusion, the elements that distinguish Spectral Mass Technique, such as large number of voices, saturation of the spectrum with harmonics of a fundamental pitch construct, lack of connective glissandi between pitches, and the use of harmonic voicing, were applied in *Prologue to Sounden Horn*, Mvt. I: "Blow This Sounden Horn." These elements, though based on the same scientific premise as the remainder of *Prologue to Sounden Horn*, define SMT as unique in its treatment of and approach to the common tools of Phonetic Composition.

Chapter 4: Sound Color Technique:
An Analysis of Mvt. II. Weird Haunting Bizzare

The second movement of *Prologue to Sounden Horn*, 'Weird Haunting Bizarre,' is based on the spectral characteristics of individual instrument pitches. The wind instruments carry the phonetic content, while the strings mimic that content. The strings do not carry the phonetic content in this technique due to the uniformity of timbral characteristics in the string family. Individual pitches do not carry the same changes in overtones and formant zones as wind instruments. Timbral variance from pitch to pitch within a single instrument is the most crucial factor in matching phoneme spectral characteristics.

The spectral characteristics of each wind instrument have been analyzed and compared to those of the periodic speech sounds of American Standard English.[29] The analysis process used for this purpose will be discussed in more detail later in this chapter. Once phonemes and instrument pitches of similar timbre were identified they were organized into tables for application in phonetic orchestration. The Phonetic Composition technique used in this movement is referred to as Sound Color Technique.

The distinctive elements of Sound Color Technique (SCT) are as follows:

- Solo and small ensemble execution of the phonetic component (few simultaneous voices carry the melodic content)
- Pointillistic melodic content based on prescribed pitches which emulate vowels and vowel-like consonant timbres
- Generally a thin sonic texture

[29] Kenyon, *A Pronouncing Dictionary of American English*.

The instrumentation for Mvt. II is the same as Mvt. I. The phonetic component, however, is represented directly only by the wind instruments. Use of the string instruments is strictly to reiterate pitch content prescribed for the winds in matching phoneme timbres, as discussed in the process below. Percussion serve a consistent purpose throughout the composition, that is, to support and color the consonant phonemes and provide basic rhythmic pulse where deemed necessary.

The phonetic component of SCT is fundamentally represented by an orchestration technique known as pointillisim. The basis for the pointillistic treatment of the melodic construct is pitch timbre as related to speech sounds. Timbral considerations were determined from extensive aural and visual analysis of speech sounds, instrument pitches, and spectrograms.

While pointillism in music composition based on tone color is not a new technique, the basis of SCT is unique. The sonic replication of a pre-selected text using pitch colors associated with speech sounds in a pointillistic fashion is represented by a very limited repertoire in the literature. To my knowledge, the only composer to have published a work using this technique is Clarence Barlow, who is the Corwin Endowed Chair of Composition at the University of California – Santa Barbara.[30] His piece, *Im Januar am Nil*, is addressed in Chapter 1.

Advances in the use of linguistics and speech sounds in contemporary works are largely in the electronic and computer-generated realm of art music. Applications of these scientific disciplines for traditional orchestral instruments are rare. A significant percentage of the body of literature which uses speech sounds as its basis within this

[30] Bob Gilmore, "Clarence Barlow Interview" (accessed March 18, 2017).

medium was researched and developed before the mid 1980's.[31] Technological tools and speech science have advanced significantly since that time. These breakthroughs allow for new insights into the possible applications of phonetics in orchestration. Thoughts on the contribution of this method to the evolution of music composition are discussed in Chapter 6.

Movement II is largely through-composed, following the prescribed text, as expressed below. Some repetitions of phrases were employed to add a sense of continuity and familiarity with motivic material, but these repetitions – when they occur – are not formatted to fit any classically formal framework. Repetitions of phrases were conjoined and determined based on durational considerations, not on formal substance.

The steps of the compositional process for constructing Mvt. II are similar to that of the other two techniques of Phonetic Composition. These steps were outlined in Chapter 2. The following summary of the SCT process gives an brief account of processes common to all three techniques employed in the creation of the overarching composition as well as a relatively exhaustive account of the processes unique to SCT.

The source material for Mvt. II, also derived from the original recording of Marvin Lamb's spoken program notes to a premiere performance of *Sounden Horn*, is shown below in Figure 4.1, accompanied by its phonetic transcription.

Figure 4. SEQ Figure_4. * ARABIC 1 Text and Phonetic Transcription for Mvt II Source Material

7. A'… and they just… uh… used their ears instead of their eyes,
   æ ænd ðe d͡ʒust ə juzd ðɛr ɪrz ɪnstɛd əv ðɛr ɑɪz

8. which I thought was a wonderful musical metaphor.
   hwɪtʃ aɪ θɔt wəz ə wʌndə-fəl mjuzɪkəł mɛtəfɔr

[31] Bob Gilmore, "Clarence Barlow Interview" (accessed March 18, 2017).

9. uhhh… And they… wuh… kept hearing this sound, it'll go…
   ə ænd ðe wə kɛpt hɪrɪŋ ðɪs sɑund ɪtɬ go

10. it's an undulating descending minor third
    ɪts æn 'ʌndjə˛letɪŋ dɪsɪndɪŋ mɑɪnəˠ θɝd

11. "uhh uhh uhh uhhh uhh uhh"
    ə ə ə ə ə ə

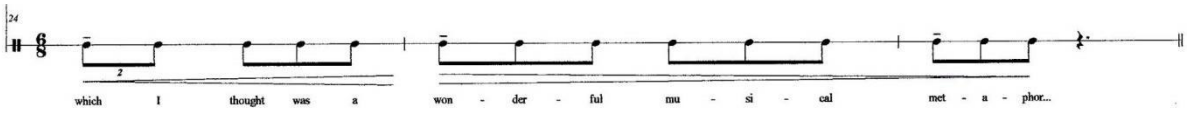
12. and it's this weird haunting bizarre sound
    ænd ɪts ðɪs wɪrd hɔntɪŋ bɪzɔr sɑund

13. that just goes forever when you hear it played…
    ðæt d͡ʒust goz ɔn fɔr'ɛvəˠ hwɛn ju hɪr ɪt pled

14. uhh… on a river…
    ə ɔn ə rɪvəˠ

An example of the rhythmic context is found in Figure 4.2 below.

Figure 4. SEQ Figure_4. * ARABIC 2 Example of Mvt II Rhythmic Context from Phrase 8

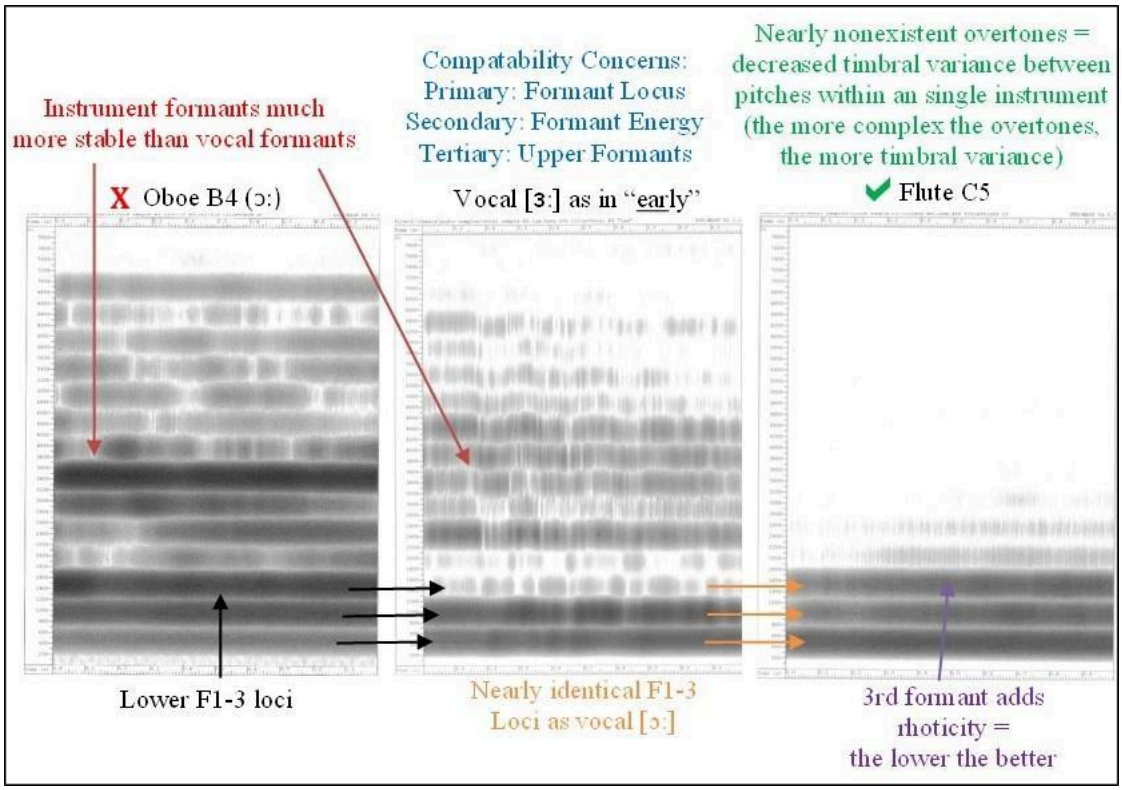


In determining pitch content to apply to the rhythmic scheme (Figure 4.2), SCT calls for instrument-specific pitch timbres, distinguished as having similar spectral characteristics to specific phonemes. This step replaces the spectrogram creation in the other two techniques. In the process of determining pitch-phoneme corollaries, two methods of research were conducted – spectral analysis and aural comparison.

In the spectral analysis stage, spectrograms of steady-state phonemes from the source recording were compared with steady-state individual instrument timbres, pitch

by pitch across the full range of each wind instrument. Sample spectrograms are found below in Figure 4.3. The spectrograms were compared for formant loci, formant strength, and matching upper formants. The hierarchy of timbral-matching considerations is in the order listed in the previous statement.

Figure 4.3 Sample Spectrograms of Oboe and Flute Instrument-Pitch and Vocal [ɜ:] timbre for Visual Comparison



The measurements taken regarding formant compatibility are done with computer-aided measuring tools, including those found in the Speech Analyzer 3.1 software mentioned previously. The formant frequencies being analyzed are so small and close together that simply viewing the formants side-by-side does not adequately compare multiple spectrograms.

In the spectrograms found in Figure 4.3 above, a comparison of formant loci 1, 2, and 3 shows a slight difference between the oboe pitch B4 and those of the vocal [3:] (sounding "her"). The formant loci of the oboe B4 are lower than those of the vocal, while those of the flute C5 are nearly identical. Again, the graphic representation above is only an approximate comparison and not regulated by measurements within the software, which are far more detailed and accurate.

Secondly, we compare formant energies, as indicated by the darker and lighter black and grey horizontal bars exhibited in the spectrograms. In comparing the three spectrograms, formants one and two appear to maintain very similar strength. The third formant, however, is weak in the vocal spectrogram, fairly strong in the oboe, and slightly weaker in the flute. This difference between oboe and flute, while visually not highly different, is audibly perceivable, as will be discussed below. Once a phoneme is measured for formant loci and intensity, the original source recordings of all steady-state vowels and instrument pitches are compared audibly.

This audible comparison works in much the same way that tuning an instrument by ear occurs – beat matching. That is, when two pitches are "out of tune" their waveforms interfere with one another, causing dissonance (asynchronicity or the quality of being asynchronous).[32] Similarly, when two waveforms of the same frequency but of differing timbres are suspended in close proximity, their overtones (formant zones) either reinforce one another or conflict with one another. This difference is audible. This technique exists and is practiced frequently in choral ensembles, where vowel-matching is critical to ensemble blending and intonation. When two waveforms with compatible

[32] Georgia State University, "Timbre," Department of Physics, http://www/hyperphysics.phys-astr.gsu.edu/hbase?Sound/timbre.html (accessed March 18, 2017).

formant zones merge, they enhance and amplify one another. The difference between compatible and incompatible timbres is readily apparent once one's ears are trained to distinguish the two.

Once compatible instrument-pitch and vocal timbres are determined, they are compiled into a document which can be used in the assignment of pitches to the rhythmic context described previously. An example of an instrument-specific timbre assignment chart can be found in Figure 4.4 below. The phonemes are listed in the column to the far left. The pitches of the full range of the flute are listed horizontally at the top of the chart. The 'x's indicate timbral compatibility.

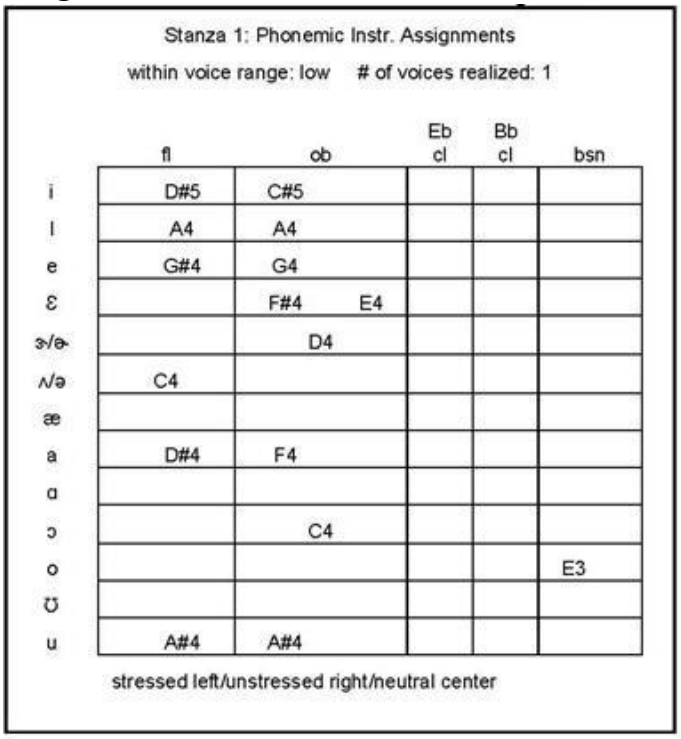
Figure 4.4 Sample Instrument-Specific Timbre Assignment Chart

FLUTE	C4	C#4	D4	D#4	E4	F4	F#4	G4	G#4	A4	A#4	B4	C5	C#5	D5	D#5	E5	F5	F#5	G5	G#5	A5	A#5	B5	C6
[i:] as in "peep"																x									
[I] as in "pit"										x					x		x			x					
[e] as in "paper"			x						x					x							x		x		
[ɛ] as in "pet"		x																							
[ɜ:] as in "bird"					x	x							x												
[ə] as in "about"	x																								
[æ] as in "pat"																									
[ɑ] as in "hot"																			x						x
[ɑ:] as in "aunt"				x																					
[ɔ:] as in "law"												x													
[ɔ] as in "roll"																		x						x	
[o] as in "wrote"																						x			
[ʌ] as in "putt"																									
[ʊ] as in "put"																									
[u:] as in "school"							x	x																	
[u] as in "rude"											x														
[ɻ] as in "rat"																									
[ɬ] as in "scold"																									
[l] as in "lift"																									
[n] as in "not"																									
[m] as in "map"																									

These pitch assignments can then be collected into charts used in the construction of each phrase, based on the phonemes in each phrase and the instrument pitches assigned to those phonemes, giving a ready-to-access palette of instrument-specific pitches to select from when creating the melodic component in each

section. A sample of a Phonetic Instrument Assignment Chart can be found below in Figure 4.5.

Figure 4.5 Sample of Phonetic Instrument Assignment Chart for Phrase 1



Stanza 1: Phonemic Instr. Assignments

within voice range: low # of voices realized: 1

	fl	ob		Eb cl	Bb cl	bsn
i	D#5	C#5				
ɪ	A4	A4				
e	G#4	G4				
ɛ		F#4	E4			
ɝ/ɚ		D4				
ʌ/ə	C4					
æ						
a	D#4	F4				
ɑ						
ɔ		C4				
o						E3
ʊ						
u	A#4	A#4				
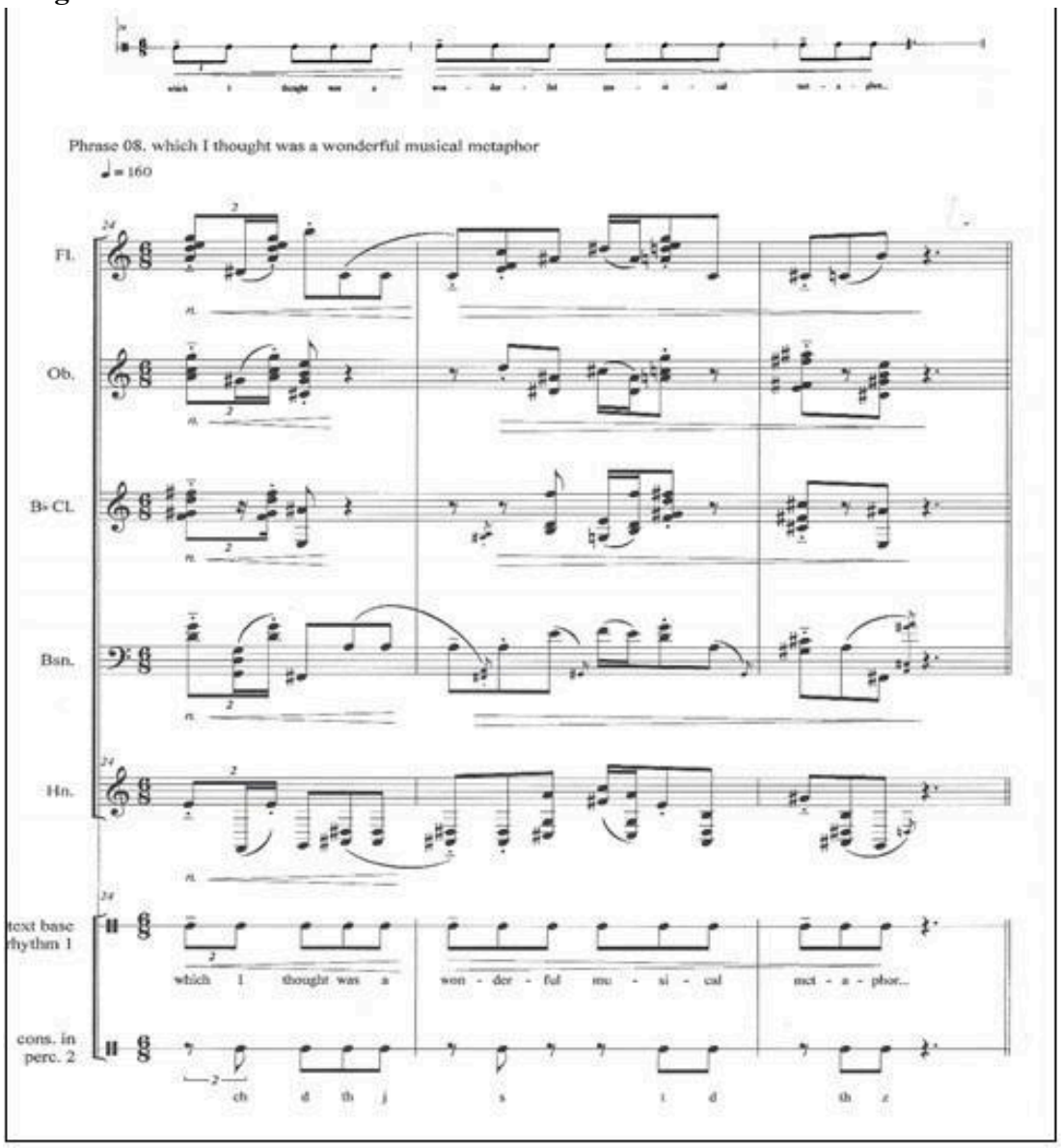
stressed left/unstressed right/neutral center

These pitches can also be combined with the rhythmic construct to produce the following notated graphic for use as an easily-resourced compositional tool. An example is found below in Figure 4.6.

Percussive effects are then assigned. This step was covered in Chapter 3 and remains similar in SCT. The charts produced as described in this chapter are then used to create the composition, both for rhythm and pitch. Spectrograms of each phrase are used to determine flow and phrasing (as described previously). The data collected from spectral and aural analysis was then used in discrete ways, selecting orchestration and durational considerations based on creative license of the composer, though adhering to

the parameters of the tools provided by the data extraction process. Therefore, all material is either directly or indirectly related to the source recording and manipulated only to the degree that is deemed necessary for aesthetic purposes.

Figure 4.6 Notated Graphic Combining Rhythmic Construct and Instrument Pitch Assignments



This approach serves the purpose of Phonetic Composition, in that the end result is intended to provide an aesthetic experience to the audience, regardless of their knowledge or awareness of the technical basis for the artwork created.
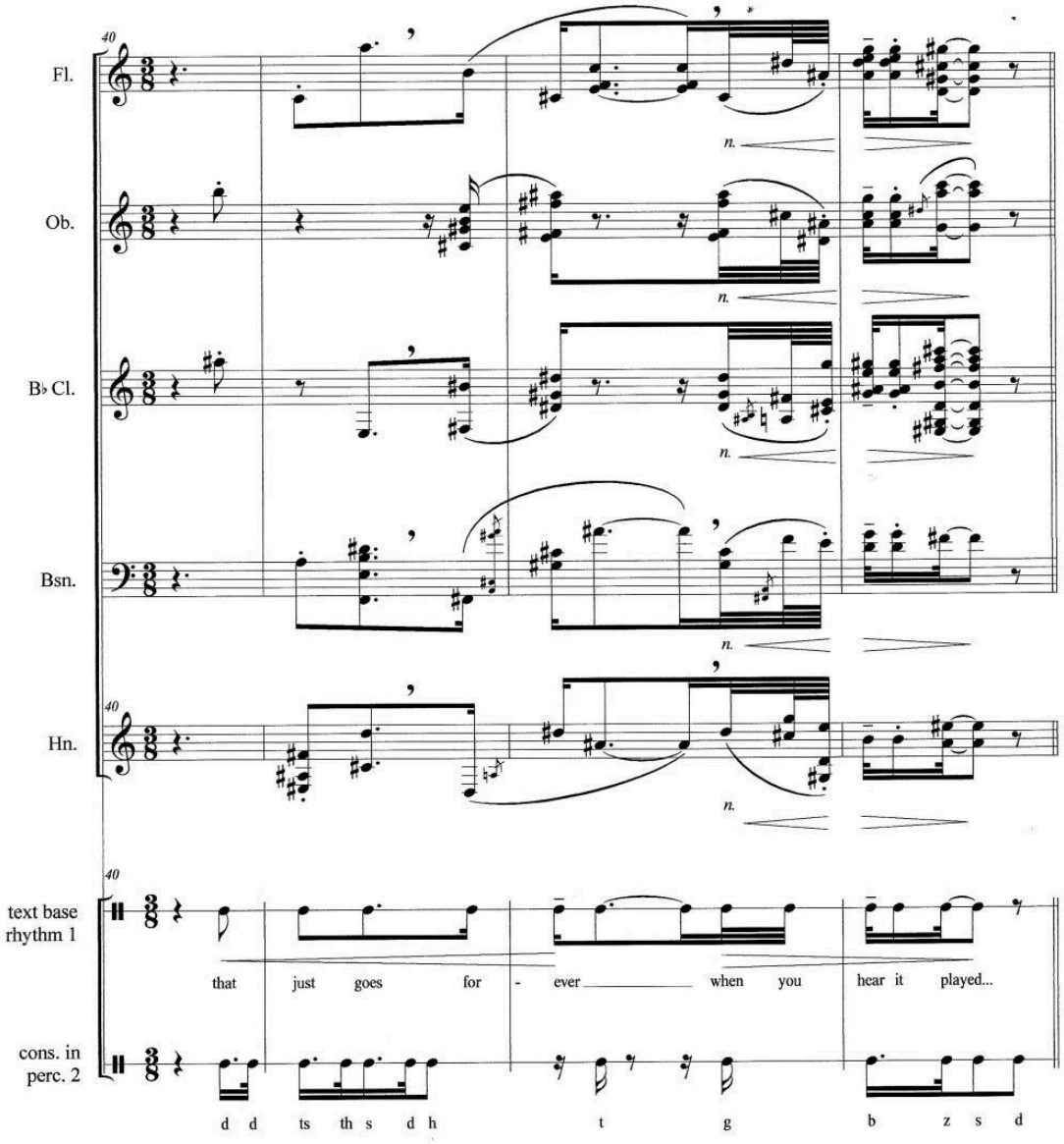
Using Mvt. II Phrase 13 “that just goes forever when you hear it played” as a point of example for the prescribed applications outlined above, the creation of a phrase using SCT will be detailed here. For a reference to the score for the steps described in this breakdown, please see Appendix B of this document, where a complete score can be found. This phrase is located at m. 228 of Mvt. II.

The pitches and rhythms used in Phrase 13 are based on those found in the Pitch and Rhythm Construct in Figure 4.7. This section of the piece was predetermined to render eight repetitions of the original phrase, using two to four instruments largely in alternation and tandem within each iteration. The iterations begin in consecutive fashion, then gradually become elided, finally to overlap. The iterations move in pairings in the following order: flute/piccolo, oboes, Bb clarinets and piccolo/flute, oboes and bassoon, piccolo/flute and bassoon, horns and clarinets, followed by all winds in groupings with iterations initiating every beat or two, finally ending in an ensemble sustained crescendo in mm. 246 and 247 aided by the strings and timpani.

The rhythmic construct is slightly faster than real time (4:3 durational ratio) as compared to the source recording. The pitches were derived from the Pitch and Rhythm Construct Chart (see Figure 4.7). In the first iteration at m. 228, the flute and piccolo begin with the word “just” in the text, moving from C4 in the flute, to A6 followed by B5 in the piccolo (adjusted for octave from the original of A5 and B4), then Db4 in the flute (C# in the chart) to C6 (C5 at 8va) in the piccolo. In m. 229, the flute begins with

C#4 in the flute (coinciding with “when” from the text), followed by a string of pitches in the piccolo (Eb and Bb of the diphthong ”you,” consecutive G’s from “hear it,” and a sustained Ab from “played”). The flute moves with the piccolo for the words “hear it played,” using pitches A5 and C#6 (transposed up the octave from the A4 and C#5

Figure 4.7 Mvt II Phrase 13 Pitch and Rhythm Construct



prescribed in the chart for octave proximity to the piccolo). The deviations in octave placement in this section were based on tessiture and artistic considerations.

The second iteration of the phrase begins in m. 231 voiced in the oboes. The pitches follow those prescribed in the chart. The third iteration begins at m. 234 in the piccolo, flute, and clarinets. The pitches follow those prescribed in the chart, with pairings selected at the discretion of the composer for aesthetic reasons. The third iteration begins in m. 237 in the piccolo, flute, oboes and bassoon; followed by the fourth iteration at m. 239, truncated to the portion representing "hear it played" in the piccolo, flute, and bassoon; followed by the fifth iteration at m. 240 in the clarinets and horns. Here the clarinets and horns render the melody in an 8:3 durational ratio to the original, halving the pace of change initiated by the theme at m. 228. This continues in mm. 242-245 in all voices, initiating iterations six and seven, while the flutes return to the former 4:3 durational ratio at m. 245, overlapping the same figure in the clarinets still in the 8:3 durational ratio. The bassoon ends with the final iteration, still truncated, at the 4:3 durational ratio to the original source recording.

The process is relatively simple and straightforward. Deciding on the durational scheme (ratios), the number of iterations, the voices to use in each iteration, and composer discretion to decide when each iteration occurs within the context of the overarching section, this phrasal representation creates a complex mixture of voices, rhythmic webs, and harmonic textures in a relatively fugal style. The final "chord" (vertical sonority) at mm. 246 and 247 helps to punctuate the end of the section and prepare for the introduction of Phrase 14 at m. 248. This cluster in the strings is simply

a sympathetic harmonic gesture reinforcing the winds, derived from the pitch chart. Each phrase rendering in Mvt. II. was constructed using a similar process.

In conclusion, Sound Color Technique features a unique set of characteristics within the methods of Phonetic Composition prescribed in this document. These characteristics include a predominance of solo and small ensemble orchestration, as well as a vowel-timbre oriented, pointillistic approach to melody. This approach, and the selection of the melodic construct, is based on pitch timbres which have been assigned to speech sounds determined via analysis to be similar in spectral signature. The process is intended to emulate colors of speech sounds, as well as their stress, duration, and flow, as they might in a choral work based on the same text.

Chapter 5: Phonetic-Event Technique: An Analysis of Mvt. III. No One Would Perish

Phonetic-Event Technique is based on a current technological trend in speech synthesis known as Text-to-Speech Synthesis (TTS Synthesis).[33] The goal of speech researchers when using TTS Synthesis is to allow computers the ability to "converse" with humans in a natural sounding way. To accomplish this task, speech components are dissected mid-phoneme, reassembled into all practical combinations, and adjusted to follow natural speech intonation and inflection patterns. This allows computers to reproduce less-than-common terms and names in relatively natural ways based on information it gathers form other similar words.

This is especially helpful when sounding out names, such as "Hubert," which can be assembled from phoneme combinations found in the more common words "human," "tube," "bird," and "hurt." Reassembled for transitory elements and adjusted for intonation tendencies, the resulting pronunciation can not only be relatively precise, it can express speech patterns in a way that represents natural human speech.

These same techniques were applied in the process of creating *Prologue to Sounden Horn*, mvt. III. 'No One Would Perish.' The steps followed in this process are detailed below. The process is much the same as current computer applications of diphonic (literally "two phonemes") speech synthesis, with the exception that their results are adapted for use in standard orchestration techniques.

The distinctive characteristics of Phonetic-Event Technique (PET) are its fluid connectivity between speech sounds, its consistency of phonemic representation adjusted for pitch context, its direct replication of speech patterns, and the emphasis of

[33] Peter Ladefoged, *Vowels and Consonants*, 75.

phonetic events over formants or timbres. PET shares a fundamental basis in spectral analysis, phonetic-event mapping, and formant plotting with both SMT and SCT.

The founding principal of PET lies in the basis of all speech. Speech patterns are nothing more than a series of hums, clicks, buzzes, hisses, and whistles which are interpreted by the human brain as language.[34] With enough repetition and consistency of execution, musical or sonic events, applying linguistic patterns, can be interpreted in the same way.

The major difference from SMT is PET's rare use of the full ensemble to carry the phonetic component, though this does occur a few times in Mvt. III when directly quoting real-time speech patterns, as opposed to artistic representations at slower tempi. Its contrast with SCT is that a much larger force, with a larger selection of instruments and pitches, is available for use in the phonetic component. These elements can be applied at any time without consideration of pitch-specific instrument timbre or the inclusion of the harmonic series. The primary formants are represented in Mvt. III, but they are not tied to the harmonic series of a fundamental pitch construct, as in Mvt. I. An F0 is used as the basis for Mvt. III, but its use is resourced for intonation (prosody) and rhythm only.

Mvt. III is in a modified rondo form, given that the final phrase of the excerpt selected from the source recording, Phrase 19, is used as a recurring theme, while all other phrases are alternated with it. The modification in this rondo form is that the ritornello never returns to the "home" key, each iteration shifting the ascending interval of a minor 3rd from the previous iteration. The term "rondo" is used loosely, as the material that is reused is the fundamental spectral and rhythmic framework, not the

[34] Remez website (accessed March 13, 2017)

pitch content. This ascension of minor 3rds is mediated by a progression of other phrases from the original source file, each expressing its subsequent section's dominant key area. This is possible given the paraphrase-nature of PET, wherein formant relationships and concatenative processes take precedence over specific pitches or timbre relationships.

The selection of this form and key structure was meant to satisfy forward momentum from one section to the next. The minor third is a primary interval throughout the composition, given that this is the interval expressed by the sounden horn. The dominant area preceding each reiteration of Phrase 19, serving as a sort of *ritornello*, serves the purpose of setting up each subsequent return of the theme via classic tonal progression. Within these key areas no functional tonality is intended and any occurrence thereof is coincidental due to formant relationships inherent in the spectral content of the phonemes represented.

A comparison with the other two Phonetic Composition techniques of Spectral Mass and Sound Color, along with a summation of the unique characteristics of Phonetic-Event Technique are as follows:

- Interpretive treatment of formant transitions into musical statements using spectrogram-based pitch ratios
- Like SMT and SCT, PET is dependent on the original speech source with respect to declamation and aspects such as rhythm, timing, phrasing, shape, consonant placement, and dynamic contrasts
- Unlike SMT, PET is less concerned with intelligible phoneme colors and more with formant transition relationships

- Unlike SMT and SCT, PET's interpretive representation of formant relationships is much more flexible with regard to transposition, accompaniment to the phonetic component, and pitch assignments

The following steps outline the processes used in the creation of Mvt. III, based on prescribed techniques of Phonetic Composition and adapted from previous methods employed in Mvts. I and II. The steps here are listed in chronological order of execution. They have been combined and summarized where appropriate for the interest of conciseness. Figures are inserted to represent the formulaic components of the process, which was used as the framework for this Phonetic-Event composition, upon which interpretive decisions were made.

The source text selected as the phonemic content of Mvt. III is derived from the same recording used in Mvts. I and II. The text is as follows, accompanied by its phonetic transcription:

15. uhh… And it was sort of like the bell cow
ə ænd ɪt wəz sɔrt əv lɑɪk ðə bɛł kɑu

16. everyone would orient to that sound,
ˈɛvrɪˌwʌn wud ˈorɪˌɛnt tu ðæt sɑund

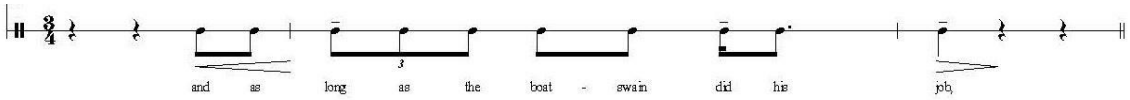
17. and as long as the boatswain did his job
ænd æz lɔŋ æz ðə ˈbotˌswen dɪd hɪz d͡ʒɔb

18. no one would parish.
no wʌn wud ˈperˌiʃ

19. And pretty much no one did parish.
ænd prɛtɪ mut͡ʃ no wʌn dɪd

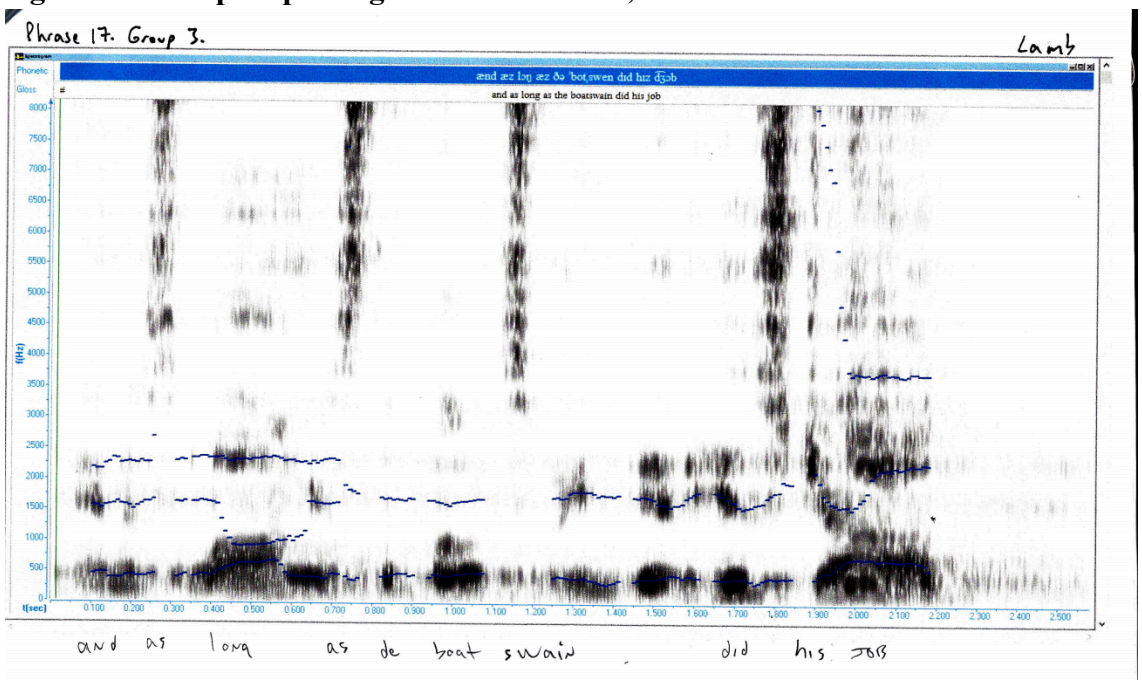
The rhythmic content was then established, as seen below in Figure 5.1.

Figure 5. 1 Sample Rhythmic Context from Mvt III Text Excerpt, Phrase 17



Spectrograms were created and analyzed for formant characteristics, shape, and phrasing. An example spectrogram can be seen in Figure 5.2 below.

Figure 5. 2 Sample Spectrogram from Mvt III, Phrase 17

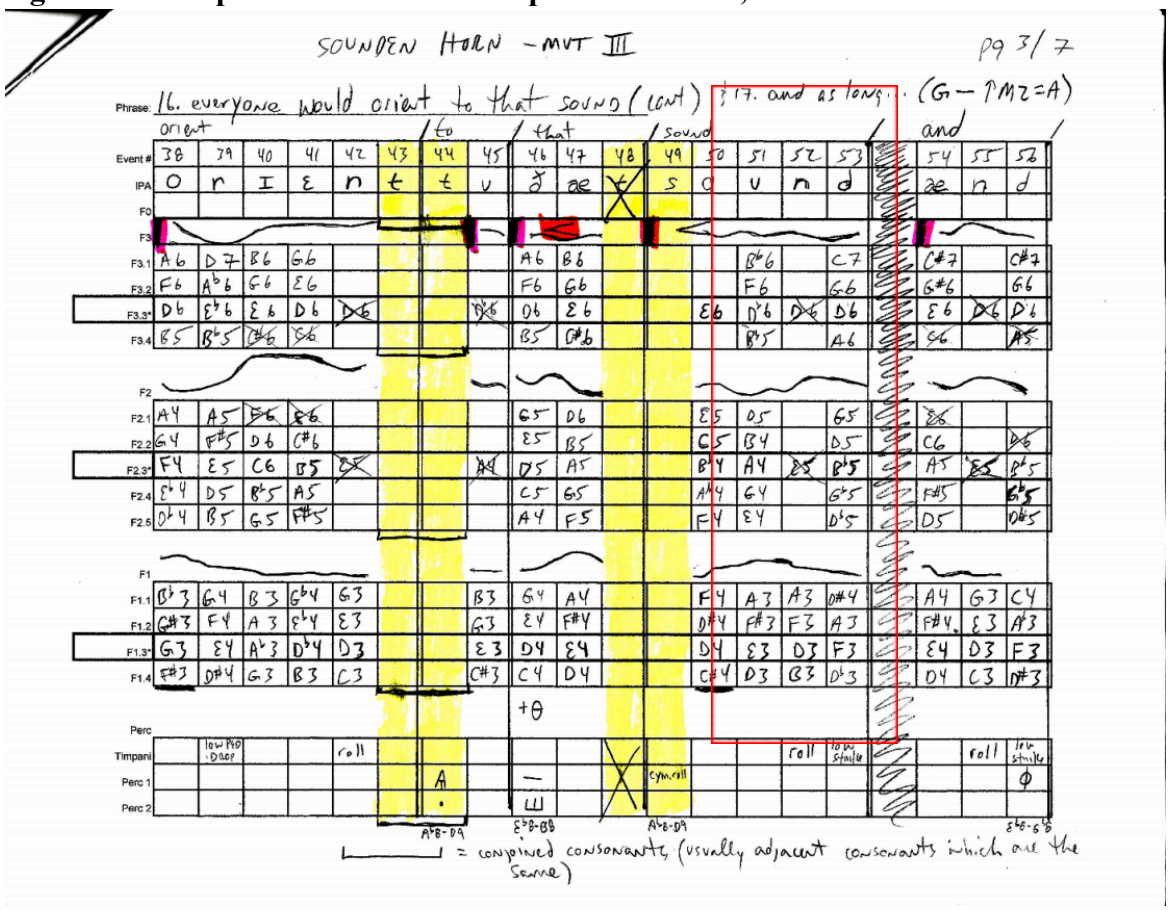


A phonetic map was created from the extraction of data and formant analysis based on the spectrograms created. A sample page from the phonetic-event map is found in Figures 5.3, 5.4, and 5.5 below.

The data extraction and analysis processes outlined above are fundamental to all three techniques of Phonetic Composition and have been described previously in some

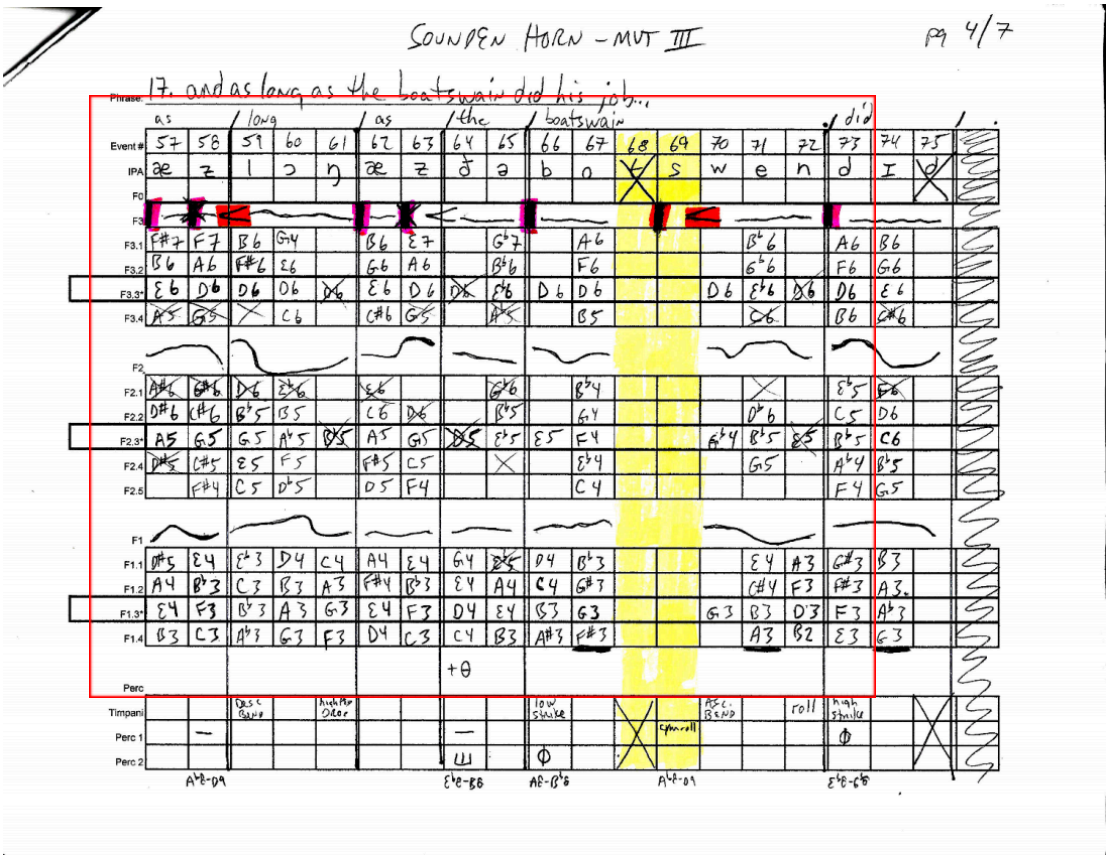
depth. For this reason, the samples of data extraction serve to establish a basic understanding of content used for the phonetic basis of Mvt. III. This outline is not intended to serve as an exhaustive resource of data collection for Mvt. III. The methods unique to Phonetic-Event Technique (PET) are discussed in more detail below.

Figure 5. 3 Sample Phonetic-Event Map from Mvt. III, Phrase 17



After creating the phonetic-event map, and in keeping with the concatenative processes characteristic of PET, a Phoneme Formant Pitch Relationship Chart was created to establish a consistent set of relationships between formant zones. These relationships acted as prototypes for phonemic content, which were then modified

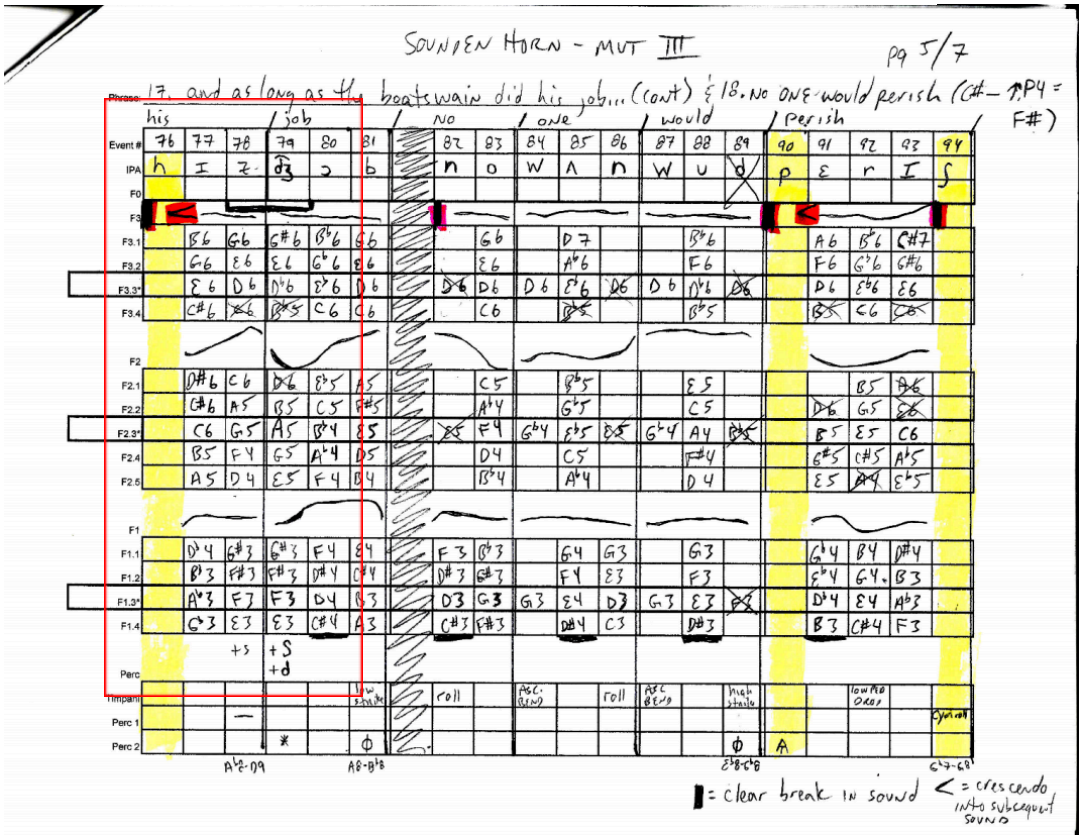
Figure 5. SEQ Figure_5. * ARABIC 4 Sample Phonetic-Event Map from Mvt III, Phrase 17



according to processes described by Ladefoged in the linguistic application of phonemes for the intelligibility of speech sounds (see Ladefoged's *Vowels and Consonants* in the bibliography). This chart can be seen in Figure 5.6 below.

The next step was to establish the overall form of the piece. For this step, a flow chart was created to provide a basic form scheme, including the succession of phrases, number of voices representing the phonetic component in each section, the general mood of each section, the prescribed duration of each section and the tempo and number of measures required to achieve that prescribed duration, and the transposition

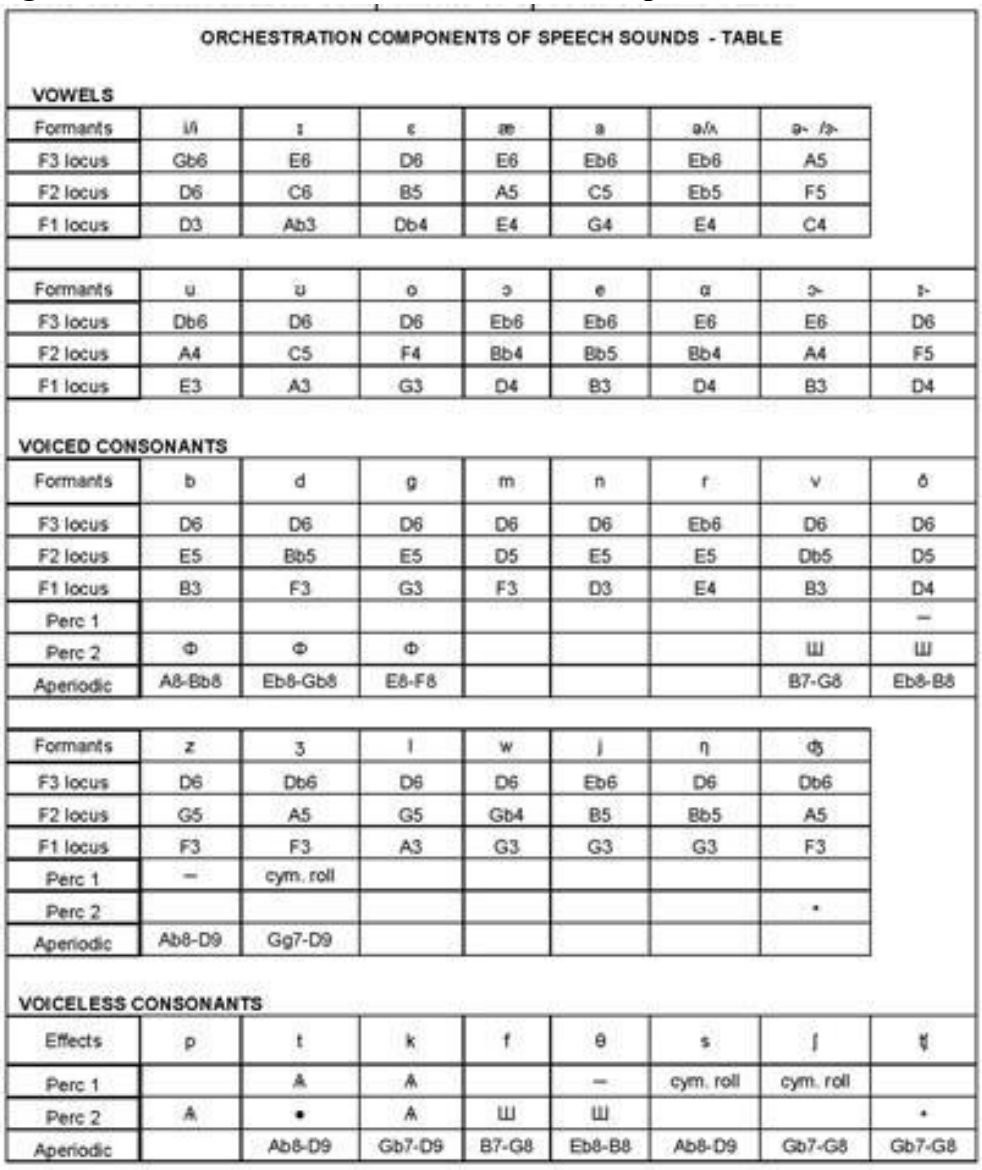
Figure 5. SEQ Figure_5. * ARABIC 5 Sample Phonetic-Event Map from Mvt III, Phrase 17



required to adjust tonal centers to fit the prescribed tonal scheme. The flow chart created for this purpose is seen in Figure 5.7 below.

To indicate the applications of the above data collection and analysis, the creation of Phrase 17 will be detailed here. Phrase 17 occupies section VI within the overall form of the movement and is located in the complete score (see Appendix B to this document) beginning at Mvt. III, m. 117.

Figure 5. 6 Phoneme Formant Pitch Relationship Chart



ORCHESTRATION COMPONENTS OF SPEECH SOUNDS - TABLE

VOWELS

Formants	i/ɪ	ɪ	ɛ	æ	a	ə/ʌ	ə- /ɔ-
F3 locus	Gb6	E6	D6	E6	Eb6	Eb6	A5
F2 locus	D6	C6	B5	A5	C5	Eb5	F5
F1 locus	D3	Ab3	Db4	E4	G4	E4	C4

Formants	u	ʊ	o	ɔ	e	ɑ	ɔ-	ɜ-
F3 locus	Db6	D6	D6	Eb6	Eb6	E6	E6	D6
F2 locus	A4	C5	F4	Bb4	Bb5	Bb4	A4	F5
F1 locus	E3	A3	G3	D4	B3	D4	B3	D4

VOICED CONSONANTS

Formants	b	d	g	m	n	r	v	ð
F3 locus	D6	D6	D6	D6	D6	Eb6	D6	D6
F2 locus	E5	Bb5	E5	D5	E5	E5	Db5	D5
F1 locus	B3	F3	G3	F3	D3	E4	B3	D4
Perc 1								–
Perc 2	Φ	Φ	Φ				Ш	Ш
Aperiodic	A8-Bb8	Eb8-Gb8	E8-F8				B7-G8	Eb8-B8

Formants	z	ʒ	l	w	j	ŋ	ʧ
F3 locus	D6	Db6	D6	D6	Eb6	D6	Db6
F2 locus	G5	A5	G5	Gb4	B5	Bb5	A5
F1 locus	F3	F3	A3	G3	G3	G3	F3
Perc 1	–	cym. roll					
Perc 2							•
Aperiodic	Ab8-D9	Gg7-D9					

VOICELESS CONSONANTS

Effects	p	t	k	f	θ	s	ʃ	ʧ
Perc 1		A	A		–	cym. roll	cym. roll	
Perc 2	A	•	A	Ш	Ш			•
Aperiodic		Ab8-D9	Gb7-D9	B7-G8	Eb8-B8	Ab8-D9	Gb7-G8	Gb7-G8

In Section VI of Mvt. III, five voices carry the phonetic component at any one time. Winds, strings, and percussion alternate vowels and consonants in adjacent phrase repetitions. The timpani serve as a motoring device. The overall mood is carried over from section V, designated in the flow chart as “optimistic yet authoritative.” This is accomplished via the use of wide harmonic intervals, homophonic rhythms within voices, and aggressive articulations.

Figure 5. 7 Flow Chart for Mvt III

Flow Chart for Mvt III.

Overall Form:

Phrase	# of vox	Mood	Dur.	Multi.	Tone Cntr	Transp.	Time Stamp
I. Phrase 19	1	Concern	45"	x4	Db	Tritone	:00 – :45
Note: app. 20 measures at m.m. = 108							
II. Phrase 15	2	hopeful	1'	x2	E	min. 3rd	:45 – 1:45
Note: app. 28 measures							
III. Phrase 19	4	determined	1' 15"	x8	A	Maj. 2nd	1:45 – 3:00
Note: app. 34 measures							
IV. Phrase 16	3	focused	45"	x4	Db	Tritone	3:00 – 3:45
Note: app. 20 measures							
V. Phrase 19	7	optimistic	1'	x12	Gb	min. 2nd	3:45 – 4:25
Note: app. 28 measures							
VI. Phrase 17	5	same	45"	x5	A	Maj. 2nd	4:25 – 5:10
Note: app. 20 measures							
VII. Phrase 19	11	heavy & burdened	2'	x16	D	Perf. 5th	5:10 – 7:10
Note: app. 56 measures							
VIII. Phrase 18	4	reprise hopeful	45"	x4	F#	Maj. 3rd	7:10 – 7:55
Note: app. 20 measures							
IX. Phrase 19	2	at peace	1'	x2	B	Maj. 3rd	7:55 – 8:55
Note: app. 28 measures							

Total of 254 measures

Temporal Feel breakdown:
Sections @ x8 or greater = 118 measures (4' 15")
Sections @ x6 or lesser = 136 measures (5')

The duration of the section is twenty measures in length, with the addition of four "flex" measures. "Flex" measures are used for transitions between sections, or to act as codettas to preceding material. The content of these transitions and codettas is based on the prescribed material of their respective sections, but treated with much more artistic license to help stream sections in a more fluid or musical manner. The

prescribed duration of the section is forty-five seconds, allowing for five repetitions of the original phrase, each instance beginning at m. 117, m. 121, m. 125, m. 129, and m.133 consecutively.

The prescribed tonal center is determined by the vowel contained in the most stressed syllable of each phrase. In this instance, the word “long” is the stressed syllable, and its F1 locus is the pitch G. Transposed up a major second, as prescribed in the Overall Form schematic, yields a tonal center of A. This tonal center is also reinforced throughout the phrase section via frequent repetitions of and doublings on this pitch to reinforce its presence as an acting tonal center. This “progression” of thirds with interjecting “dominant” intermediaries serves to propel the movement forward unlike either of the first two movements.

In the first iteration, at m. 117, the first horn and first clarinet carry the vowel sounds, representing Formants 1 and 2, while the strings, timpani, and remaining winds carry the consonant sounds. At m. 121, the low winds follow a counter-melodic line to the strings, based on the prescribed consonant locations. The strings carry the vowel formants, forming a melodic outline of those formant zones. At m. 125, the horn and clarinet resume the melodic vowel content, with the other voices resuming consonant functions. This alternates again, akin to the second iteration, wherein the low winds carry a counter-melodic line based on consonants, while the strings move homophonically, depicting the phonetic content. Measure 133, the final iteration of the phrase, returns once again to clarinet and horn carrying the vowels, while the remainder of the ensemble fills in the consonants. The section ends with an extension of the string line and harmony, drawing to a brief pause in the music at m. 141.

The pitch content used for this section is related to that prescribed in the Phoneme Formant Pitch Relationship Chart, seen above in Figure 5.6. The rhythms are as prescribed for phonemic durational augmentation in Figure 5.9 above. The rhythms within the context of any one phoneme were subject to composer discretion as was deemed to serve the aesthetics of the music. This means that rhythms within a prescribed phoneme duration could be of any character, given that the material representing each phoneme was executed within the bounds of the prescribed rhythm on the flow chart for that particular phoneme. For instance, the word "job" is represented in m. 136 as an eighth-, quarter-, eighth-note. Its prescribed rhythm in the flow chart (Figure 5.7) was a half note. The rhythms and pitches in representing "job" here were derived from the vowel and consonant pitches prescribed in Figure 5.6. It should be noted that the specific pitches used are not critical to the process of Phonetic-Event Technique, as was stated earlier, but the overall relationships of the formants one to another. With this in mind, adjustments were made as deemed aesthetically appropriate to the composer.

While sharing similar fundamental principles with Spectral Mass Technique and Sound Color Technique, such as spectral analysis, phonetic-event mapping, formant plots, and phonological applications, Phonetic-Event Technique is unique in its independence from harmonic and pitch-timbre constraints.

The concatenative approach to speech synthesis, which is used extensively in synthesized speech production using computer-aided technologies, is complex, versatile, and highly effective in its application to orchestration as well. The perceived

audible difference between sections representing direct quotes and those with more of an artistic, symbolic approach is at once compelling and effective.

The range of possibilities within this method of Phonetic Composition is vast. Choices determined for the expression of this method with regard to Mvt. III include the embedding of formant transition tendencies within each prescribed diphone and the exaggeration of time to explore those transitions in an interpretive and artistic fashion.

Chapter 6: Reflections on *Prologue to Sounden Horn*

The refinement of orchestration techniques in the processes applied to Phonetic Composition since the initiation of research in 2012 is witnessed in the composition, *Prologue to Sounden Horn*, which is discussed in this document. Eliminating unnecessary steps, sharpening those steps which yield positive effects, and gaining a more comprehensive outlook on the techniques employed has been a rewarding process of discovery. Mastering the orchestration tools brought to light from the study of phonetics and what I call the "dance of the formants" has allowed me to render more musically the data extracted from the scientific tools available at this time for speech research.

The true achievement of this research is the unveiling of orchestration techniques which reflect the beauty of language and its speech-sound components. A perspective on phonetics that has yet to be fully explored, the musical applications of speech science are vast. The Spectral Mass method uncovered techniques for orchestrating the timbres of vowels based on their harmonic structures. The Sound Color method explored individual spectral characteristics of instruments and their pitch colors. The Phonetic-Event method applied the most modern speech replication techniques in a fashion favorable to standard orchestration. Underlying them all was the development of phonetic-event mapping and formant scoring strategies.

The idea that spoken language could be represented to great effect using standard acoustic instruments by composers Ablinger and Barlow was inspiration for the research expressed in this document. The coupling of Ablinger's and Barlow's

research, and the advancement of speech replication methods, such as Text-to-Speech Synthesis technologies, formed the basis of Phonetic Composition.

Ablinger integrated spectral analysis with a mechanized piano to reproduce human speech patterns successfully. Much of the orchestrational technique used in *Prologue to Sounden Horn* is influenced by the piano renderings of Ablinger. Examples of this are consonant cluster treatments and short, high-range pitches to represent aperiodicity, which can be found throughout *Prologue to Sounden Horn*.

Barlow's writing served as a great inspiration as well. The successful results of his orchestrations served as proof that the art of Phonetic Orchestration was viable. The fact that the science of phonetics could be applied to standard orchestration with any degree of success was inspiration for further research, as demonstrated within this document. The excerpt which sparked my imagination to seriously pursue Phonetic Composition was Barlow's phrase "Why me? No money. My way," from *Orchidae Ordinariae*.[35] The audible affirmation of this orchestrated phrase, along with the initial orchestration experiments conducted in preparation for this document, piqued my curiosity and interest into this only slightly-explored area of composition.

The processes presented in this document differ from Ablinger and Barlow in their spectral extraction and applications. Barlow converted audio signals into MIDI note-values, which were then orchestrated expressly as dictated. Ablinger eliminated the need for performers altogether by use of a mechanized piano linked directly to the computer analyzing the sound source's spectral data. The phonetic component in *Prologue to Sounden Horn* was constructed based on a combination of spectral analysis and phonetic transcriptions of the text, yet with ultimate control for all orchestrational

[35] Bob Gilmore, "Clarence Barlow Interview" (accessed March 18, 2017).

decisions left to the composer. The technique of orchestration here was applied note-by-note, phrase-by-phrase in a subjective manner using the empirical data as a guide. The reason for this was to render an "artistic" interpretation of speech sounds, not simply a direct quote of those sounds.

The three techniques employed in the creation of *Prologue to Sounden Horn* – Spectral Mass, Sound Color, and Phonetic-Event – are all branches from the same tree, so to speak. All three methods use as their basis spectral analysis of speech sounds, phonetic-event mapping based on phonetic transcriptions, and formant properties. Of course, these elements were applied in differing degrees and in altered contexts, but their orchestration techniques were highly similar, the primary difference being their respective sources of selecting pitch constructs.

Regarding overall effectiveness, continued research, experimentation, and writing will yield more refined techniques for further development and advancement in this area. Speech research and synthesis techniques continue to evolve and grow, and with them their applications to Phonetic Composition. The ability to render intelligible speech sounds by virtue of standard orchestration methods may never be achieved to the degree that words, phrases, and text can be understood clearly by most or all listeners without external prompts, such as program notes or projected visual aids. However, the artistic settings of speech applications in music yields something new, relevant, and ambitious in contemporary music composition.

One significant source of enlightenment with regard to general compositional techniques is rooted in the application of consonantal characteristics. The importance of these components has proven to be essential to modern language, but when translating

them into a composition, they seem alien. This led me to the observation that most art music seems to be written as if all notes express vowel-like characteristics exclusively. Notes or note groups exhibiting consonant-like behavior, as determined by this study, are rarely seen in music compositions. The consideration of connective material between more sonorous moments was a revelation to my compositional perspective.

This awareness allowed me to see melodic and transitional construction in a new light. Melodies and counter-melodies, rhythms and sub-components of music are more than just isolated events with only an over-arching connective motivation. I believe more can be done in the transience of musical elements to enhance phrasing and forward motion by taking great care in the application of these elements.

The advances in Phonetic Orchestration in this document from those of previous composers, such as Ablinger and Barlow, are meaningful, in that they take advantage of current advances in speech research, computer analysis technologies, and a convergence of previous applications. Also addressed are the complete spectrum of consonants and aperiodic sources, which are not attempted by Barlow, and only by their periodic counterparts by Ablinger. The departure from literal representations of speech is an important development here, as the interpretive possibilities of phonetic applications have only begun to be explored.

References

List of sources referenced for research purposes in the development of Phonetic Composition and its theories, regardless of whether used for quotation, paraphrase, or content of any kind. It is acknowledged that the information contained in these sources may have influenced the foundations of Phonetic Composition to some degree.

Arrell, Chris. “Chris Arrell.” http://www.chrisarrell.com/index.html [accessed March 18, 2017].

Bickford, J. Albert, David Tuggy, ed. “The Principal Organs of Articulation.” Summer Language Arts Institute of Mexico. http://www.mexico.sil.org/es/ publicaciones /articulacion [accessed March 18, 2017].

Bunnell, H. Timothy. “Synthesis of Speech.” Applied Science and Engineering Laboratories, University of Delaware. http://www.asel.udel.edu/speech/ tutorials/synthesis/vspace2.html [accessed March 18, 2017].

Carr, Philip. *English Phonetics and Phonology: An Introduction*. Malden, MA: Blackwell Publishing, 1999.

Department of Computer Science, The. “Since-Wave Speech Cocktails.” University of Sheffield. http://staffwww.dcs.shef.ac.uk/people/M.Cooke/MAD/sws/sws.htm [accessed June 11, 2013].

Department of Electrical Engineering and Computer Science. “Acoustic Theory of Speech Production.” Massachusetts Institute of Technology. http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-345-automatic-speech-recognition-spring-2003/lecture-notes/lecture2.pdf [accessed March 18, 2017].

Division of Psychology and Language Sciences. “ Basics of Voicing.” University College London. http://www.phon.ucl.ac.uk/cgi-bin/wtutor?tutorial =siphtra/vb1.htm [accessed March 18, 2017].

Dodge, Charles. “Speech Songs.” Perfect Sounds Forever Online Magazine. http://www.furious.com/perfect/ohm/dode.html [accessed June 11, 2013].

Fineberg, Joshua. *Spectral Music: History and Techniques*. Denmark: The Overseas Publisher Association, 2000.

Grout, Donald J. and Claude V. Palisca. *A History of Western Music.* 4th ed. New York: W. W. Norton and Company, 1988.

Georgia State University. “Timbre.” Department of Physics. http://hyperphysics.phy-astr.gsu.edu/hbase/Sound/timbre.html [accessed March 18, 2017].

Georgia State University. “Vocal Sound Production.” Department of Physics. http://hyperphysics.phy-astr.gsu.edu/hbase/Music/voice.html [accessed March 18, 2017].

Gilmore, Bob. “Clarence Barlow Interview. “ Paris Transatlantic Magazine. http://www.paristransatlantic.com/magazine/interviews/barlow.html [accessed March 18, 2017].

Hanavan, Perry C. “Speech Perception.” Augustana College. http://www.coe.fau.edu/faculty/danesh/courses/documents/SPA4104.pdf [accessed March 18, 2017].

Harvey, Jonathan. “Jonathan Harvey.” http://vivosvoco.com/index.html [accessed March 18, 2017].

Hindemith, Paul. *The Craft of Music Composition, Book 1*. Mainz, Germany: Schott and Company, Ltd., 1970.

Hoffman, Aljoscha. “Peter Ablinger.” http://ablinger.mur.at/ [accessed March 18, 2017].

Indian Institute of Technology, Bombay. “Source-Filter Theory of Speech Production.” Indian Institute of Technology, Department of Humanities and Social Sciences. http://www.hss.iitb.ac.in/courses/HS435/Acousticslecture_files/frame.htm [accessed June 7, 2013].

Institute of Technology, Bombay, India. “Acoustic Phonetics.” http:www.hss.iitb.ac.in.courses/HS435/Acoustics%20lecture_files/frame.htm [accessed June 11, 2013].

International Phonetic Association. *Handbook of the International Phonetic Association*. Cambridge, MA: Cambridge University Press, 1999.

Janer, Jordi. “Syllabling on Instrument Imitation: Case Study and Computational Segmentation Method.” Music Technology Group, Pompeu Fabra University, Barcelona. http://www.uni-graz.at/~parncutt/cim07/CIM07%20Proceedings/CIM07_Janer-Penalba_Syllabling%20on%20instrument%20imitation.pdf [accessed June 11, 2013].

Johnson, Keith. *Acoustic and Auditory Phonetics*. 3rd ed. Malden, MA: Blackwell Publishers Ltd., 2012.

Jones, Daniel. “Cardinal Vowels.” Universitiet Utrecht. http://www.let.uu.nl/~audiufon/data/e_cardinal_vowels.html [accessed June 11, 2013].

Kent, R. D. “Vocal Tract Acoustics.” Journal of Voice, Texas A&M University. http://courses.cs.tamu.edu/rgutier/cpsc689_s07/kent1993vocalTractAcousticsSLIDES.pdf [accessed March 18, 2017].

Kenyon, John S. and Thomas A. Knott. *A Pronouncing Dictionary of American English*. Springfield, MA: Merriam-Webster Inc., publishers, 1953.

LaBouff, Kathryn. *Singing and Communicating in English: A Singers Guide to English Diction*. Oxford: Oxford University Press, 2008.

Ladefoged, Peter. *Elements of Acoustic Phonetics*. Chicago: The University of Chicago Press, 1996.

Ladefoged, Peter. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Malden, MA: Blackwell Publishing, 2003.

Ladefoged, Peter. *Vowels and Consonants*. Los Angeles: Blackwell Publishing, 2005.

Lamb, Marvin. “Sounden Horn.” Verbal program notes and original composition presented at the University of Oklahoma, Catlett Music Center, Sharp Performance Hall, October 2015.

Lorenzi, Antoint, and Benjamin Chaix. “Sound Representation.” NeurOrielle. http://www.cochlea.org.en/spe/sound-representation.html [accessed June 11, 2013].

Maidment, John. “Speech Internet Dictionary.” Division of Psychology and Language Sciences, University College London. http://www.phon.ucl.ac.uk/home/johnm/sid/sidhome.htm [accessed June 11, 2013].

Mannell, Robert. “Speech, Spectra, and Spectrograms.” Linguistics Department of Macquerie University. http://clas.mq.edu.au/acoustic/speech_spectra/index.html [accessed March 18, 2017].

Mannell, Robert. “TTS: An Overview of Concatenative Approaches to Speech Synthesis,” Macquarie University, http://clas.mq.edu.au/speech/synthesis/tts_concat/index.html (accessed March 27, 2017).

Massachusetts Institute of Technology. “Automatic Speech Recognition.” Department of Electrical Engineering and Computer Science. http://1337institute.com/c/62/automatic-speech-recognition [accessed March 18, 2017].

McGill, David. *Sound in Motion: A Performer’s Guide to Greater Musical Expression*. Bloomington: Indiana University Press, 2007.

McSperitt, Kenneth E. “Perigee and Apogee: A Phonetic Composition for Chamber Orchestra.” Masters thesis, Stephen F. Austin State University, 2013.

Medical Research Council Cognition and Brain Sciences Unit. “Sine-Wave Speech.” Cambridge University. http:www.mrc-cbu.cam.ac.uk/people/matt.davis/sine-wave-speech/ [accessed March 18, 2017].

Murail, Tristan. “Tristan Murail.” http://www.tristanmurail.com/en/index.html [accessed March 18, 2017].

Mustafa, Kamron and Ian C. Bruce. “Robust Formant Tracking for Continuous Speech with Speaker Variability,” *IEEE Transactions on Audio, Speech, and Language Processing 14*, no. 2 (March 2006).

National Center for Voice and Speech. “Models of Vocal Fold Oscillation.” http://www.ncvs.org/ncvs/tutorials/voiceprod/tutorial/model.html [accessed June 11, 2013].

Rabiner, Lawrence. “Digital Speech Processing – Lecture 3.” Rutgers University. http://soe.rutgers.edu/~Irr/digital_speech_processing_course/lectures/Lecture%203_fall_2010.6tp.pdf [accessed June 11, 2013].

Remez, Robert E. "Sine-wave Speech." Department of Psychology, Barnard College, Columbia University. http://www.scholarpedia.org/article/Sine-wave_speech [accessed March 18, 2017].

Rubin, Philip, Robert E. Remez, Jennifer Pardo. “Sinewave Synthesis.” Haskins Laboratories, Yale University. http://www.haskins.yale.edu/research/sws.html [accessed March 18, 2017].

Russell, Kevin. “Phonetics.” Department of Linguistics, University of Manitoba. http://home.cc.umanitoba.ca/~krussll/phonetics/index.html [accessed March 18, 2017].

Samson, Jim. *Music in Transition: A Study of Tonal Expansion and Atonality*. New York: W. W. Norton and Company, Inc., 1977.

Scavone, Gary P. “Vocal-Tract Influence during Saxophone Performance.” Music Technology, Schulich School of Music, McGill University. http://www.music.mcgill.ca/~gary/vti/index.html [accessed March 18, 2017].

Section de Linguistique. “Sound Production.” Universite de Lausanne. http://www/unil.ch/Jahia/site/ling/cache/offonce/pid/24434 [accessed June 11, 2013].

Slawson, Wayne. *Sound Color*. Talent, Oregon: Yank Gulch Music, 2007.

Story, Brad H. “Using Imaging and Modeling Techniques to Understand the Relation Between Vocal tract Shape to Acoustic Characteristics.” Department of Speech

and Sciences, University of Arizon. http://sal.shs.arizona.edu/~bstory/Story SMAC03.pdf [accessed June 11, 2013].

Tontechnik-Rechner – sengpielaudio. “Keyboard and Frequencies.” http://sengpielaudio.com/calculator-notenames.htm [accessed June 11, 2013]

Traube, Caroline. “Instrumental and Vocal Timbre Perception.” Laboratoire d’informatique, acoustique, Universite de Montreal. http://www.uni-graz.at/~parncutt/guests/2006/traube/Slides_CTraube_timbre.pdf [accessed June 11, 2013].

Widmayer, Sharon and Golly Gray. “Sounds of English.” http://www.soundsofenglish.org/index.html [acessed June 11, 2013].

Wolfe, Joe. “Speech and Music, Acoustics and Coding, and What Music Might Be ‘For’.” School of Physics, the University of South Wales. http://newt.phys.unsw.edu.au/jw/ICMPC.pdf [accessed March 18, 2017].

Appendix A: Annotated List of Notable Compositions

Annotated List of Notable Compositions since 1830 whose techniques influenced or are related to those found in *Prologue to Sounden Horn*
(chronologically listed by date of composition)

Hector Berlioz	*Symphonie Fantastique* (1830) – use of pointillistic techniques in the fourth movement, wherein the melody is shared between strings and winds (orchestra)
Claude Debussy	*Prelude a l'apres-midi d'un faune* (1894) – use of the color of the flute and harp as referential functions, as well as timbral considerations of other instruments for orchestrational effect (orchestra)
Claude Debussy	*Nuages* (1899) – first notable composition elevating timbre to a primary compositional status with pitch and rhythm (orchestra)
Bela Bartok	The works of Bela Bartok from 1904-1918, which were based on scales derived directly from resonances of wind instruments, inspired the development of spectrum-based composition
Arnold Schoenberg	*Farben* from *Five Pieces for Orchestra* (1912) – experiment with tone-color structure (orchestra)
Igor Stravinsky	*The Rite of Spring* (1913) – timbral elements treated with unprecedented importance (orchestra)
Charles Ives	*Universe Symphony* (1928) – timbre as a principle element of composition (orchestra)
Olivier Messiaen	*Quatuor pour la fin du temps* (1914) – composition employing color effects based on scale modes, known as 'synaesthesia' (piano, violin, cello, and clarinet)
Iannis Xenakis	*Metastisis* (1954) – use of orchestral sonic 'clouds' of pizzicato and glissandi effects (orchestra)
Gyorgy Ligeti	*Apparitions* (1959) – spectral composition employing sound-mass techniques and micropolyphony to evade a sense of surface texture (orchestra)

Karlheinz Stockhausen	*Kontakte* (1963) – compositional elements pitch and timbre are combined and separated, manipulated individually to show independence; produces an "aw"-like sound color (electronic)
Milton Babbitt	*Composition for Synthesizer* (1964) – sound color as a primary compositional element; use of sound color as contrapuntal tool; timbre dictates form (electronic)
Kenneth Gaburo	*Antiphony IV* (1967) – use of vocal timbres as primary compositional elements, forming the basis of pitch, rhythm, and form (three instruments and two-channel tape)
Morton Subotnick	*Silver Apples of the Moon* (1967) – use of sound color as the primary compositional element (electronic)
Karlheinz Stockhausen	*Hymnen* (1969) – developed a systemization of timbre for his compositions; applies color matching techniques of timbre manipulation to strongly contrasting sound sources; spectral representation of the German national anthem (electronic)
Karlheinz Stockhausen	*Telemusik* (1969) – timbre used as a sense of completion (electronic)
Charles Wourinen	*Time's Encomium* (1969) – the use of filters to alter individual aspects of timbre (electronic)
Karlheinz Stockhausen	*Stimmung* (1971) – based on vowel and consonant sets using the harmonic series (six amplified voices)
Jacob Druckman	*Synapse* (1971) – timbre invariance used as 'equivalence class' (electronic)
James Tenney	*Clang* (1972) – extensively based on the harmonic series (electronics)
Benjamin Boretz	*Group Variations* (1974) – elements of pitch and duration are set against timbre for contrapuntal effect; waveform held as invariant (electronic)
Gerard Grisey	*Le Espaces Acoustiques,* Mvt. III. *Partiels* (1975) – harmonic series of a fundamental represented in

	orchestration for standard orchestral instruments (chamber orchestra)
Charles Dodge	*Speech Songs* (1976) – analysis and resynthesis of speech sounds in composition (electronic)
Milton Babbitt	*Phenomena* (1977) – based on vowel and consonant sets (voice, tape or piano)
Hugues Dufourt	*Saturne* (1978) – spectral work exploring the nuances of duration (percussion, wind ensemble, electronics)
Hubert S. Howe	*Improvisation on the Overtone Series* (1980) – harmonic series as primary compositional element; referred to as a "timbre piece"; harmonics manipulated in ways similar to pitches in a twelve-tone row (electronic)
Tristan Murail	*Gondwana* (1980) – entire work based on representations of the sound spectrum and the harmonic series; use of electronic techniques to generate instrumental orchestration (orchestra)
Morton Subotnick	*A Sky of Cloudless Sulphur* (1980) – sounds replicating diphthongs used as primary compositional element (electronic)
Clarence Barlow	*Im Januar am Nil* (1981) – based on the speech sounds vowels, nasals, and laterals; set for standard orchestration (chamber orchestra)
Jonathan Harvey	*Mortuos Plango, Vivos Voco* (1981) – morphing of the vocal timbre of his tenor son and Winchester Cathedral's Great Bell (electronic)
Wayne Slawson	*Colors* (1981) – use of vocal timbre as primary compositional element; development and use of SYNTAL system of freeware for specifying and synthesizing speech-derived music (tape and electronics)
Paul Lansky	*Six Fantasies on a Poem by Thomas Campion* (1982) – use of speech analysis and resynthesis methods in composition (electronic)

Clarence Barlow	*Orchideae Ordinariae* (1989) – spectral analysis of speech sounds used to orchestrate short phrases for standard instrumentation (orchestra and electronics)
Chris Arrell	*Argot* (2000) – spectral work which integrates vocals with chamber orchestra (soprano and chamber orchestra)
Jonathan Harvey	*Speakings* (2007) – computer software used to orchestrate speech sounds for orchestra; speech spectrum superimposed over amplified orchestral instruments to produce speech-like sounds (electronics and orchestra)
Peter Ablinger	*Voices and Piano* (2009) – 'talking piano' controlled by computer, voicing spectral representations of speech superimposed over those of instruments for speech-like replication; use of screened text in synchronicity with musical speech sound representations (computer and mechanical piano)
Julian Anderson	*The Comedy of Change* (2009) – use of extended spectral techniques to represent environmental and biological changes (chamber orchestra)
Kenn McSperitt	*Perigee and Apogee* (2013) – spectral analysis, phonetics, and phonology used to orchestrate speech sounds for standard acoustic orchestral instruments based on the harmonic series, timbral elements, and