THE MUSICAL SEMIOTICS OF TIMBRE IN THE HUMAN VOICE

and

STATIC TAKES LOVE'S BODY

by

Smith Alexander Reed

B.Mus., B.A., The College of Wooster, 2001

M.A., University of Pittsburgh, 2003

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2005

UNIVERSITY OF PITTSBURGH

SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

S. Alexander Reed

It was defended on

April 21, 2005

and approved by

Peter Havholm

Eric Moe

Andrew Weintraub

Mathew Rosenblum
Committee Chairperson

Copyright by S. Alexander Reed
2005

Mathew Rosenblum

THE MUSICAL SEMIOTICS OF TIMBRE IN THE HUMAN VOICE
and
STATIC TAKES LOVE’S BODY

Smith Alexander Reed, Ph.D.

University of Pittsburgh, 2005

In exploring the semiotics of vocal timbre as a general phenomenon within music, theoretical engagement of the history of timbre and of musical meaning bolsters my illustrative analyses of Laurie Anderson and Louis Armstrong. I outline first its reliance on subtractive filtering imparted physically by the performer’s vocal tract, demonstrating that its signification is itself a subtractive process where meaning lies in the silent space between spectral formants. Citing Merleau-Ponty’s phenomenology and placing the body’s perceptual experience as the basis of existential reality, I then argue that the human voice offers self actualization in a way that other sensory categories cannot, because the voice gives us control over what and how we hear in a way that we cannot control, through our own bodies alone, our sight, touch, taste, and smell. This idea combines with a listener’s imagined performance of vocal music, in which I propose that because of our familiarity with the articulations of human sound, as we hear a voice we are able to imagine and mimic the choreography of the vocal tract, engaging a physical and bodily listening, thereby making not only performance but also listening a self-affirming bodily reflection on being. Finally I consider vocal timbre as internally lexical and externally bound by a linguistic context. Citing Peirce and Derrida, and incorporating previous points, I show vocal timbre as a canvas on which a linguistic and musical foreground is painted, all interpreted by the body. Accompanying theoretical discussions is a concerto addressing relevant compositional issues.

PREFACE

What began life as a hastily conceived whim has grown tremendously and in ways I could not have predicted at this project's outset. Considering its conceptual breadth, while an investigation like this may never be fully *done*, this dissertation now comfortably rests in a state vaguely resembling completion, and there are many who have encouraged, comforted, helped, taught, critiqued, or dragged me kicking and screaming to this point. To them I am forever grateful.

I thank my advisor, Mathew Rosenblum, for four years of unwavering positivity, solid expertise in theory, linguistics, and semiotics, and for writing more letters of recommendation on my behalf than my direst nightmares could behold. Thanks also goes to my committee — Eric Moe, Andrew Weintraub, and Peter Havholm— who respectively honed my composition, placed my thinking in a broader view of the world, and exposed (and inspired) me to perhaps the highest standard of classroom teaching I have ever known. I am especially fortunate because I can truly call every member of my committee a friend; I have valued their conversation and attitude as much as I have their particular knowledge and insight into my dissertation topic.

I spoke with a number of people who helped me shape and edit ideas as I formulated them for this project. I thank these unofficial advisors, all dear friends as well: Camille Peters, Des Harmon, Ned Kirby, Robin Hitchcock, Galen Brown, Rick Mook, Cara Milne, Phil Sandifer, Courtney White, Margie Clayman, Andria Poiarkoff, Ian Reyes, and most of all Stephanie Petrusz. Thanks in this category also to Elisabeth Le Guin, Philip Tagg, and Suzanne Cusick. Enormous gratitude also goes to Marie Agatha Ozah for her kindness and expertise in translation.

Other mentors along the way have been vital, if not directly to the writing of this study, to my acquiring the interests, skills, and desire for learning that are poured onto these pages. These people include Mary Lewis, Roger Zahab, Jack Gallagher, Peter Mowrey, Jim Cassaro, and Deane Root. More broadly, I fondly include Donna Butler, Jane Woods, Janet Borgerson, and J. B. Wilkins.

More personally, I thank my fellow musicians Aaron Fuleki and Jeremy Long for putting up with this continued occupation of my time and energy and with all that it brings. I thank my kittens Amelia and Pie May for respectively being my alarm clock and my lap warmer through many mornings and nights of writing, and I thank Meredith Collins for her endless patience, support, discussions, movie nights, and cocoa. Finally I am thankful to God and to my family, who proofread, sent cookies, and showed me by example how to find happiness and success in academia and in life.

TABLE OF CONTENTS

PREFACE iv

I THE MUSICAL SEMIOTICS OF TIMBRE IN THE HUMAN VOICE 1

  1 TIMBRE, MEANING, LANGUAGE, AND VOICE 1

  2 VOICE AND NEGATIVE FORMANT SPACE 29

  3 VOICE, BODY, AND BEING 45

  4 MEANING REVISITED 62

II STATIC TAKES LOVE'S BODY 85

BIBLIOGRAPHY 126

LIST OF FIGURES

Figure 1 Spectrograph of a street organ playing a melody 3

Figure 2 Spectrograph of the author singing “do-fa-ti-do” ascending on C-F-B-C 4

Figure 3 Spectrograph of a clay drum repeatedly hit while its head is stretched 4

Figure 4 Comparison of vocoded and unprocessed vowels in Laurie Anderson’s voice 40

Figure 5 Harmonic saturation at the conclusion of “O Superman” 42

Figure 6 Detail of birdsong’s frequency range in “O Superman” 43

Figure 7 Spectrograph of the refrain in Louis Armstrong’s “It Takes Two To Tango” 53

Figure 8 Detail of lower frequencies in “It Takes Two To Tango” 56

Figure 9 Louis Armstrong: [u] on F3, from “two” in “It Takes Two to Tango” 57

Figure 10 Louis Armstrong: [æ] on F3, from “tango” in “It Takes Two to Tango” 57

Figure 11 Louis Armstrong: [i] on A3, from “feeling” in “It Takes Two to Tango” 57

Figure 12 Louis Armstrong: [ʌ] on F3, from “love” in “It Takes Two to Tango” 58

Figure 13 The author: [u] on F3, from “two” 58

Figure 14 Pairing possibilities in a Peircen sign system 73

Figure 15 Stage setup for “static takes love’s body” 87

I. THE MUSICAL SEMIOTICS OF TIMBRE IN THE HUMAN VOICE

1. Timbre, Meaning, Language, and Voice

All sounds can be expressed as waveforms on a graph of amplitude over time. Though in daily life we encounter very few pure tones — individual sine waves — we are almost always surrounded by a multiplicity of discrete sine waves built into every periodic sound that we encounter. Periodicity, the phenomenon of a waveform — however complex — repeating itself, is necessary for any pitch, chord, or sound otherwise to hold constant for any amount of time longer than a single cycle of its wave's repetition. Though his research interest lay chiefly in thermodynamics, the French mathematician Jean Baptiste Joseph Fourier proved two hundred years ago that any constant sound could be exactly recreated with the sum of enough sine waves of specific frequencies and amplitudes. Therefore, when any sound vibrates, even if it changes over time, it is activating a (sometimes very large) number of frequencies that combine to make the sound's totality; these individual frequencies have their own durations and envelopes, and they come and go as a sound changes.

The lowest frequency at which a sound's overall periodicity uniquely repeats is called the fundamental frequency.[1] In most cases, the fundamental is a sound's most audible, and we thus identify it as the sound's pitch. Each individual frequency in a sound's whole is called a partial. Very often, a sound's many partials congregate at and around the multiples present in the harmonic series of the fundamental, meaning it is typical for a violin playing a 440 Hz A to produce significant amounts of amplitude at 880 Hz, 1320 Hz, 1760 Hz, 2200 Hz, and so on, though in general, the higher the multiple of the fundamental frequency, the quieter it is in the total sound.

---

[1] While the definition I give here is standard, in the third chapter I address subharmonics, which are in fact lower than fundamentals. However, it is ultimately the fundamental

Regardless of whether they occur at purely harmonic intervals, the loudest and most prominent non-fundamental frequencies in a sound are called its formants. The number, intensity, and position of a sound's formants are the most important characteristic in a dimension of sound that is neither pitch — which refers chiefly to the fundamentals, both in melody and harmony — nor rhythm; this dimension is timbre.

Murray Campbell, in the New Grove Dictionary of Music and Musicians, defines timbre as:

> A term describing the tonal quality of a sound; a clarinet and an oboe sounding the same note at the same loudness are said to produce different timbres. Timbre is a more complex attribute than pitch or loudness, which can each be represented by a one-dimensional scale (high–low for pitch, loud–soft for loudness); the perception of timbre is a synthesis of several factors, and in computer-generated music considerable effort has been devoted to the creation and exploration of multi-dimensional timbral spaces. The frequency spectrum of a sound, and in particular the ways in which different partials grow in amplitude during the starting transient, are of great importance in determining the timbre.[2]

---

from which the harmonic series is derived and which is practically always of greater amplitude than the subharmonic.

[2] *The New Grove Dictionary of Music and Musicians,* 2nd ed. S.v. "Timbre." Murray Campbell.

The starting transient, as Campbell calls it, is the "attack" of a sound in which its spectral content very frequently differs momentarily from the rest of the sound's resonating duration. This manifests in a variety of ways, from the brief scraping noise at the start of viola's marcato inflection to the quiet glottal noise that begins our pronunciation of a word like "all" or "elephantiasis."

Being the way something sounds, timbre is therefore how we are able to differentiate and characterize sound. In order to speak meaningfully about these issues of sound difference and character, we must have a way of measuring timbre. For my purposes, the most useful tool to accomplish this measurement is a spectrograph, which plots the spectral content of sound on a graph of frequency over time, where the brightness or intensity of a partial indicates its volume in the sound. The following are spectrographs' displays of a variety of sounds, purely to familiarize the reader with spectrographs' ability to show visually the characters of sounds. Here is a street organ playing a basic melody:

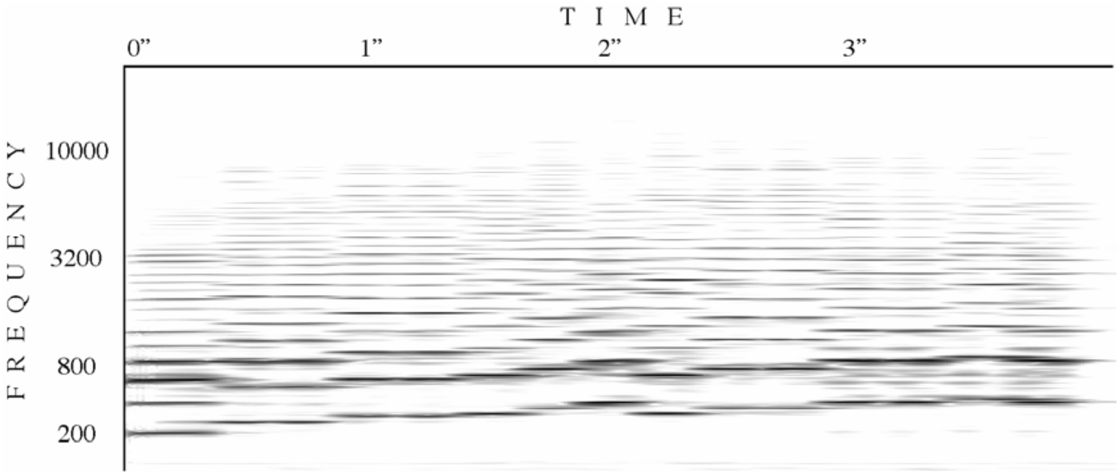


Figure 1: Spectrograph of a street organ playing a melody

This is the spectrograph of my singing "Do-Fa-Ti-Do," ascending on C-F-B-C:

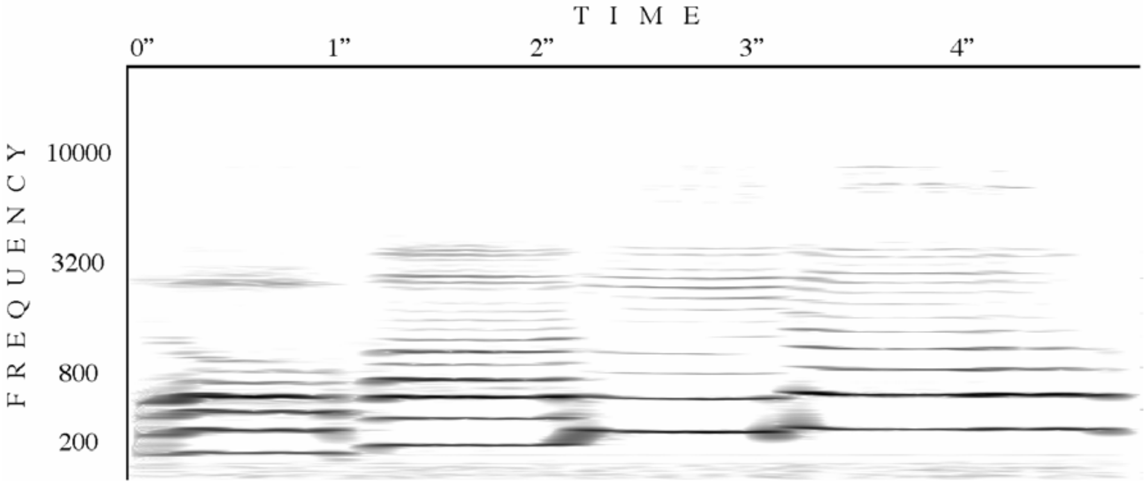


Figure 2: Spectrograph of the author singing “do-fa-ti-do” ascending on C-F-B-C

And as a last example, this is a clay drum being played while its head is stretched, changing certain frequencies within the sound:

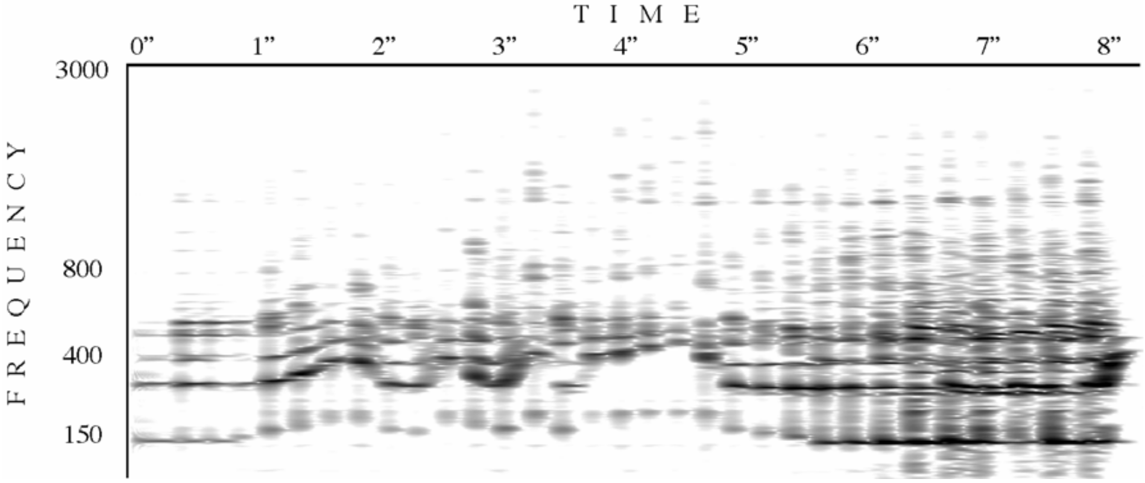


Figure 3: Spectrograph of a clay drum repeatedly hit while its head is stretched

One of the most important applications of spectrography has been in determining the actual difference between very similar sounds, such as an inexpensive starter violin and a Stradavarius, or between two different voices saying the same thing. With particular attention to the differences among voices, the influences on timbre owe largely to the individual physiology of speakers.

> The cavities of the vocal tract possess natural notes of resonance which "pass" or reinforce certain harmonics of the vocal chord tone. These emphasized harmonics form concentrations of acoustic energy at frequency regions on the spectrum corresponding to the natural notes of resonance of the cavities. The term *formant* refers to the selective resonance in a particular frequency which characterize the timbre, or color, of vowels, while the higher formants contribute mostly to the timbre, or quality of the individual voice.[3]

Spectrographs are therefore not merely the domain of music scholarship, but are an important tool in linguistics. Driven by the concern that different speakers of language have markedly different voices and yet are able to communicate with what people perceive as the same basic linguistic sounds, Roman Jakobson and the "Prague School" of linguists in the 1950s began looking to spectrography to determine the distinguishing features of linguistic units and linguistic continuity. Jakobson, Gunnar Fant, and Morris Halle explain the issue:

> Two speakers uttering the "same" vowel, have somewhat different formant frequencies depending on particular vocal tract dimensions. The spread of formant data may be specifically large if all the possible contextual variants of a phoneme as well as all possible speaker categories are taken into account. However, in a particular context it is to be expected that any speaker following the code of his language will produce phonemically different sounds by means of consistent distinctions in the formant pattern.[4]

---

[3] Howie, John and Pierre Delattre. "An Experimental Study of the Effect of Pitch on the Intelligibility of Vowels." *The Bulletin of National Association of Teachers of Singing* XVIII (May, 1962): 6-9. Quoted in William A. Hunt. "Spectrographic Analysis of the Acoustical Properties of Selected Vowels in Choral Sound." (Ph.D. diss., North Texas State University, 1970), 2.

[4] Jakobson, Roman, C.G.M. Fant, and M. Halle. "Preliminaries to Speech Analysis." *MIT Acoustic Laboratory Technical Report No. 13* (1952). Quoted in Andrew H. Harper Jr. "Spectrographic Comparison of Certain Vowels to Ascertain Differences Between

Many scholars in music, linguistics, and even ornithology have used spectrography to determine the distinctive features of instruments, vowels, and birdcalls. That our ears are presented with more information than is necessary to identify individual sounds is certainly a motivating factor in learning what timbral features are essential in the identification of general sound types, but for the purposes of this study, I wish to step beyond the argument of distinctive features. When we are only concerned with that which makes a piano a piano, or which gives the phoneme [e] its identity, then we lose the nuances that allow us to tell a Steinway grand from a dulled spinet, a computer program taught to say [e] from a Shakespearean actor's enunciation of it.

In his book, *Sound Color*, Wayne Slawson explores the perceptual characteristics of "steady-state portions of sound,"[5] which he recognizes as slightly different than timbre. His chief means of exploring and denoting sound color is the comparison of the first two formants of sound above a fundamental tone (F1 and F2), because "It is known that the higher resonances must be adjusted from their neutral position in certain vowels…. However, these adjustments are not ordinarily independent; they are almost always associated with particular values of F1 and F2."[6] Slawson too, then, limits severely the aspects of sounds to which we may pay attention. While it is true that the acoustic phenomenon of "masking" sometimes renders certain frequencies present

---

Solo and Choral Singing, Reinforced by Aural Comparison." (Ph.D. diss., Indiana University, 1967), 20.

[5] Slawson, Wayne. *Sound Color*. Berkeley, CA: University of California Press, 1985, 20.

[6] ibid, 53

in a sound (and thereby on a spectrograph) inaudible due to their being overwhelmed by other — usually lower — frequencies[7], the reduction of sound description in accordance only with F1 and F2 is one that, were I to adopt it, would be greatly detrimental to this inquiry.

Above the range of F2 — whose upper limit Slawson gives as 2000 Hz[8], but which according to Lawrence Rabiner and Ronald W. Schafer can reach 3600 Hz[9] — are the high partials of sound, where sibilance, breathiness, and "brightness" of timbre are communicated. This is where we receive a great deal of information regarding the location, clarity, and richness of a sound. Most linguistics writing about the spectra of speech are concerned with vowels, but many consonants — particularly fricatives — rely heavily on high frequency partials for their utterance. The range of human hearing extends roughly from 20 Hz. to 20,000 Hz; if we cap it at 2000 Hz, [f] can too easily blend into [s], and snare drums become slamming doors.

The uses of spectrography and the limitations of the dominant methodology and purpose thereof have applied both to speech and to music. In the voice, the general application of music is in singing, and one chief difference between speech and singing is that there is much greater movement of and stability between movements of the fundamental frequency in a singing voice. As for the timbral qualities, however, Harper notes that the formants obtained from spoken vowels are applicable to sung vowels[10] , and so singing, aside from the frequent presence of the “singers’ formant” — a bandwidth of enhanced amplitude around 2500 Hz. found commonly in singers of the Western classical tradition[11] — abides by the same laws of acoustics as speech.

---

[7] This article demonstrates the removal and isolation of masked tones in music, revealing that a tremendous amount of high and quiet frequency remains audible in many cases: Deutsch, W. A. & F. Födermayr. "Visualization of Multi-Part Music (Acoustics and Perception)." *Arbeitsberichte der Forschungsstelle für Schallforschung der ÖAW* 1 (1995), 1-19.

[8] Slawson, 41.

[9] Rabiner, Lawrence R. and Ronald W. Schafer, *Digital Processing of Speech Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1978, 43.

With an understanding of the acoustic nature of the voice, of sound and the nature of its measurement, I now wish to discuss more concrete applications of spectrography. I reiterate that my approach to this area of study chiefly favors a hearing of the total content of sound, rather than merely its distinctive features.

Robert Cogan, in *New Images of Musical Sound*, contributes to the field of music scholarship some of the most important spectrographic work we have. While Jakobson, Halle, and Noam Chomsky dedicated tremendous efforts to the classification of sound by lists of binary characteristics, Cogan is concerned with using spectral analysis in revealing the sonic fingerprints of composers, instruments, and performers. His analyses concern formal structure in music as dictated by its spectra rather than an intensive cataloguing of individual spectral moments, although he does conduct small analyses of such concerns, as a means to his end.

I do not want to make Jakobson and his colleagues a collective straw man, however. Cogan very significantly follows their lead in assembling a list of binary oppositions by which “sonic characters” — Cogan’s own phrase through whose use he can avoid speaking of timbre — may be quantified. That he uses these thirteen acoustic criteria for a broader reading of musical works sets him apart, however, from Jakobson and Halle's work, a main intent of which was simply to show that speech sounds could indeed be described and classified at all in a more positivistic manner than the articulatory descriptions of and prescriptions for phonemic generation of speech sounds that were standard in linguistics until the pair's 1956 publication of *Fundamentals of Language*.

---

[10] Harper, Andrew H. Jr. “Spectrographic Comparison of Certain Vowels to Ascertain Differences Between Solo and Choral Singing, Reinforced by Aural Comparison.” Ph.D. diss., Indiana University, 1967, 121.

[11] Cogan, Robert. *New Images of Musical Sound*. Cambridge, MA: Harvard University Press, 1984, 38.

Suggesting yet another comparison between language and music, there is a strong similarity between Jakobson and Halle's nine binary pairs and Cogan's thirteen. Here is the set of binaries proposed by Jakobson and Halle in their book's 1971 second edition:

1. Vocalic / Non-vocalic
2. Consonantal / Non-consonantal
3. Nasal / Oral
4. Compact / Diffuse
5. Abrupt / Continuant
6. Strident / Non-strident (mellow)
7. Checked / Unchecked
8. Voiced / Voiceless
9. Tense / Lax
10. Grave / Acute
11. Flat / Non-flat
12. Sharp / Non-sharp[12]

Cogan's list is as follows:

1. Grave / Acute
2. Centered / Extreme
3. Narrow / Wide
4. Compact / Diffuse
5. Non-spaced / Spaced
6. Sparse / Rich
7. Soft / Loud
8. Level / Oblique
9. Steady / Wavering
10. No-attack / Attack
11. Sustained / Clipped

---

[12] Jakobson, Roman and Morris Halle. *Fundamentals of Language*, 2nd revised ed. The Hague: Mouton, 1971, 40-44.

12. Beatless / Beating
13. Slow beats / Fast beats[13]

The specifics of these binary pairs aside, Cogan uses them to tally a score of sonic character — negative or positive — with which he compares different moments in musical works to one another in order to enhance our understanding of the works' individual structures.

This analysis is perhaps most revealing when Cogan's timbral reading of a piece is informed by the words sung in the performance analyzed. In *The Sounds of Song*, Cogan compares spectrographs of four recordings of Robert Schumann's "Ich hab' im Traum geweinet," a song he describes as "Always obsessive, but uncertain, lost — quite literally *traumatic*."[14] In analyzing one performance — Gerard Souzay's — he writes,

> The spectrograph presents a picture of almost perfectly egg-shaped tones, beginning softly, then gently swelling and receding. Every sustained sound is brightened by its *singers' formant* and colored by the most regular of *vibratos*. Is this not, however, a singing teacher's dream rather than a half-crazed lover's?
>
> There is not the slightest doubt that Souzay commands a beautiful (single) sound; we must ask, however, whether beautiful sound alone is the point here.[15]

In contrast, Cogan describes Charles Panzera's recording as "the very embodiment of obsession. [Panzera] limits the vocal resonances to registers 3-5; even the ring of the common *singers' formant* in register 7 is suppressed."[16] [17]

---

[13] Cogan 1984, 126.
[14] Cogan, Robert. *The Sounds of Song*. Cambridge, MA: Publication Contact International, 1999, 18.
[15] ibid, 19.
[16] ibid, 18.

Evident here is the use of spectrography as a means to an end rather than an end unto itself. By taking into account the whole spectrum of sound, and not just F1 and F2, Cogan is able to shed light on the way that the performance interacts with the song's components. This is a significant step in acoustic evaluation of music in that it hints at the process of meaning-making, though Cogan himself stops short of attempting to state or explore the meanings of vocal timbres in any rigorous way other than their bolstering of a structural or thematic notion already present in the lyrics or music.

To Cogan, timbral meaning is thereby subjugated to semantic meaning, but in flirting with structuralism through his analyses of Beethoven and Stravinksy, and in spectrographically viewing the purely quantitative ground where language and structure and performance and voice are laid bare, we cannot escape that there is, on several levels, an approach to structure that owes significantly to linguistics. In *New Images of Musical* Sound, he writes, "The experience of linguistics suggests that the essential sonic features of any musical instrument are... to be found in the sum total of its structural sonic contributions to musical contexts."[18] Here again, Cogan reminds us of his interest in distinctive features.

In my final chapter, I shall return to the philosophical underpinnings of this aspect of Cogan's work, but it is sufficient for now to have demonstrated the ways Cogan uses spectrographs, the terminology by which he analyzes them, and through its intellectual heritage, its concern with structure, and most significantly through its suggestion of meaning, the linguistic nature of his approach to timbre in music.

---

[17] The spectrographic comparison of Charles Panzera with Dietrich Fischer-Diskau would seem, to someone versed in Barthes, to be a direct answer to the question of Barthes's remarks about this exact pair. Oddly, Cogan makes no mention of this connection.

[18] Cogan 1984, 145.

Cogan is by no means the first to approach musical meaning by way of a linguistic model. To present an overview of the history of this single topic would require an entire book,[19] and so my coverage of it very quickly moves from the broad essentials to the research expressly suited to my scholarly aims. Though Bruno Nettl first proposed a linguistic model for assessing meaning in musical structure in order to clarify and systematize music[20] and though Leonard Bernstein popularized the notion in his Harvard lecture series, *The Unanswered Question*, which was, according to Bryce McCrary Johnson, "so rudimentary that the parallels he sought to establish were not and cannot be regarded with any modicum of seriousness,"[21] the serious discourse of music and linguistics begins with Jean-Jacques Nattiez, who delved into it in 1973 with an awareness of philosophy and structuralism that other musicologists had lacked in such endeavors.[22]

The framework on which Nattiez organizes a linguistic approach presupposes a kind of semiotic encoding of music based on an idea of Jean Molino's: Nattiez delineates three levels of a musical text[23] — the poietic, the neutral, and the esthetic — which refer respectively to the author's emotional intent, the music itself, and a listener's reaction thereto. Nattiez concerns himself chiefly with the purely structural neutral text, independent of author and of culture, whose roles are in encoding and decoding the text. In this respect, the model is not entirely unlike Ferdinand de Saussure's three part model of linguistic communication.

---

[19] Conveniently, such a book exists: Monelle, Raymond. *Linguistics and Semiotics in Music*. Philadelphia: Harwood Academic, 1992.

[20] Nettl, Bruno. "Some Linguistic Approaches to Musical Analysis." *Journal of the International Folk Music Council* 10 (1958), 37-41. Quoted in Raymond Monelle. *Linguistics and Semiotics in Music*. (Philadelphia: Harwood Academic, 1992): 28.

[21] Johnson, Bryce McCrary. "The Semiotics of Perception: Towards a Theory of the Minimal Units of Linguistic and Musical Sound." Ph.D. diss., Duke University, 2000. 1.

[22] Nattiez, Jean-Jacques. "Linguistics: A New Approach for Musical Analysis." *International Review of the Aesthetics and Sociology of Music* 4/1 (1973): 51-68.

[23] Throughout this dissertation, "text" means any work to be interpreted, and is not limited to words.

One cannot say that the neutral level is structurally neutral at all, however. Because it is encoded from and decoded into culturally dependent experiences, then its structure, created by humans within culture and not naturally occurring as a mathematical law or an element, is only purely formal within the tacitly understood bounds of our grammar. When we analyze music, or more importantly when we notate it, through culturally engrained filters we focus only on a very limited number of its characteristics, and we attempt to find form — usually linear and often teleological — in the favored aspects of the sound or the score. The neutral text is only very selectively concerned with structure. The structures it investigates are ones that are culturally created, such as dodecaphonic organization and regular pulse; attempts by composers and theorists such as Henry Cowell or Tristan Murail to view music in terms of relations other than those traditionally favored by culture have been, albeit fascinating and fruitful, considered difficult to understand and implement fully because they lie so far outside of people's cultural ethos of music. In addition to its cultural foundation, part of this is also physiologically based. Not only is human hearing limited both in volume and in frequency range, but the shape and size of the ear and its tympanum produce a selective pattern of masking tones, the natural Fletcher-Munson curve of frequency valence, and a particular level of ability — or inability — to combine or separate sounds by frequency stratification and directional triangulation. Furthermore, our voices are limited by our bodies and oriented at a very young age by our language; our stature and dexterity idiomatically shape the instruments which, in turn, shape our compositions; and our brains may have neurological limits to their attention spans, if because of nothing else than on one end, the intermittent necessities of eating and sleeping, and on the other end, the non-instantaneous reaction time of the nervous system. My point here is not to list exhaustively every trait that shapes our views of music, but I feel it is crucial in discussing Nattiez to recognize how our communities, languages, and bodies give favor to a finite subset of ways in which we can create and hear music. In short, the neutrality of the neutral text is very much a conditional one. I am not the first to raise objection in this realm. The music theorist David Lidov, in considering Molino's writing, similarly protests that it "disturbs" him, and that "Whatever its justifications in theory, the tripartition has been utilized in practice to stage a retreat from the problems of meaning."[24]

Nattiez is himself at least slightly wary of a text's neutral level, claiming its neutrality means "*that the poietic and esthetic dimensions of the object have been 'neutralized,'*"[25] placing it therefore simply in the space between the author and audience. But Nattiez nevertheless is reluctant to contextualize his tripartition in its entirety. He suggests something greater than this system by at times calling the neutral level the *trace*[26], implying some absence or other, but even this is problematic, as the term is already so heavily associated with Derrida's specific use of it, to which Nattiez gives no acknowledgment.

---

[24] Lidov, David. *Is Language A Music?* Bloomington, IN: Indiana University Press, 2005, 86.

[25] Nattiez, Jean-Jacques. *Musical Discourse: Toward a Semiology of Music*. Translated by Carolyn Abbate. Princeton, NJ: Princeton University Press, 1990, 13.

[26] ibid, 12.

Nattiez's foremost concern in music, however, is one of grammatical structure. This is an issue that inevitably pervades most writings on musical meaning, and so I address it here. One of his major developments in this respect is in the significant furthering of Nicolas Ruwet's paradigmatic analysis, which determines in a work of music the distribution patterns of small musical units — usually motivic figures. From these discoveries concerning what paradigms go where in a piece, Nattiez demonstrates that the analyst can set out descriptive rules for their use, just as a linguist would do with phonemes in the construction of morphemes, or with words in the construction of sentences. Though the rule-making process is not entirely dissimilar between these two levels of linguistic analysis, it is the latter, more macroscopic level to which Nattiez and countless music theorists since have been drawn. This is not terribly surprising, as many, if not most musical analyses have as their goal the deepening of insight into the pieces being analyzed, and as their terminology the common small-scale ingredients of notes and durations. The distinctive features, as it were, of traditionally analyzed musical works become more manifestly clear then with every degree to which one widens one's lens: the larger and more complicated the arrangement of small-scale ingredients, the less common it becomes to the musical repertory. Ray Jackendoff and Fred Lerdahl present significant further steps toward understanding musical grammar in *A Generative Theory of Tonal Music*.

I am, however, in search of neither linear grammatical structures of music nor distinctive features. I want to make it clear that my intent is to further an understanding of vocal timbre as a general phenomenon, rather than a situational truth about individual musical works. Though aware of Nattiez's contributions to the study of musical meaning, I cannot achieve this dissertation's goals via his theoretical trajectory alone. While his concern is in the neutral level, were I to assume a position within Nattiez's model, it would be to view the relationship between the neutral and esthetic levels. As Bryce McCrary Johnson — into whose work I shall shortly delve — writes, "Meaning does not reveal itself within language or music nor is it found within the psyche of the individual. Rather, meaning — any type of meaning — is generated in the space between."[27] I am therefore interested in actual voice acts, and more specifically how, conceptually, their quantifiable timbral content enacts meaning to an audience. Here I readily recognize that the generality of this study is subject to David Lidov's criticsm that it "fails to distinguish the meaning of one piece from another" and that it "talks about the meaning of 'music,' not 'pieces.'"[28] However, it is my hope that in musical vocal timbre's being clarified here as a whole, its role in creating meaning in individual works can be rigorously taken into account in future endeavors. To delineate the markers of difference in timbral meaning among pieces — especially as a general set of rules — would be a task well beyond the already wide scope of this study.

Nattiez himself does offer some limited commentary on timbre in his writings, where he places it in a binary category of noise and sound. In his view, timbre as a general phenomenon is entwined with other musical parameters and unlike notes, it alone is too monolithic to convey meaning and to allow "a distinction between poietic, neutral, and esthetic [to] reemerge."[29] Angelo Orcalli lambastes such a dismissal, asserting that Nattiez finds timbre-centric approaches to music "soporific" and that in refusing to look more deeply into timbral meaning, he “seems to be preoccupied with… a methodology more related to medieval scholarship which has its own methods of scientific research: locating that which verifies the theoretical premises he already assumes.”[30]

---

[27] Johnson, 279.
[28] Lidov, 90.
[29] Nattiez 1990, 82.

Two decades after his dismissal of timbre’s poetic capacity, Nattiez has begun to recognize its greater role in music, but his current consideration of it as a “secondary parameter”[31] still reveals an attitude incompatible, I believe, with the human voice, for which timbre, especially given its overlap with pitch, duration, and language, is essential at every level of meaning. As long as Nattiez approaches timbre with the assumption of its subservience to form, and on a level of linguistic consideration that does not engage its role in actually forming verbal language or musical idiolect, his writing will concern a set of problems outside the scope of a phenomenological semiotics of timbre expressly modeled for the voice.

While not solving problems specifically of timbre, one of the few documents to consider the meaning of minimal units of music and language as general phenomena is Bryce McCrary Johnson’s 2000 dissertation, *The Semiotics of Perception: Towards a Theory of the Minimal Units of Linguistic and Musical Sound*. Johnson lays claim to “a common constructive source for linguistic and musical minimal units,”[32] viewing the phoneme — the smallest indivisible unit of language, representable by a single character — and note as the bases of any connection between linguistics and music. Standing in opposition to the musical grammarians, he argues that “qualities of the phoneme and note show a greater affinity than the referential qualities of the word and any possible referential qualities that may be ascribed to more complex musical constructions."[33] This notion is based in their necessary universality in their respective systems, as well as their culturally constructed roles, evidenced by their confluence in systems of meaning such as Cantonese opera or rap music.

---

[30] Orcalli, Angelo. “La Ricerca Timbrica nela Semiologia della Music di Jean-Jacques Nattiez.” *Il saggiotore Musicale* v. 2, n. 2 (1995), 355. Translated by Marie Agatha Ozah.
[31] Nattiez, Jean-Jacques. “Le Timbre Est-Il un Paramètre Secondaire?” Presented at Conference on Interdisciplinary Musicology, March 2005, Montréal.
[32] Johnson, 275.

Johnson's dissertation is not without its problematic elements. He claims that while phonemes are determined by inherent features and thus take on identity regardless of context, notes only have prosodic features, meaning that it is only by their relation to one another that they can be identified.[34] A careful understanding of the relation of pitch to timbre (which Slawson, among others, addresses) will allow one to argue that in any actual application, this is not entirely true, for there exists a particular and unique range of harmonic characteristics for any note, the internal relations of which necessarily, if only very subtly, imbue particular pitches with particular timbres. For Johnson's claim to be fully true, the notes of which he speaks would need to be entirely theoretical and unuttered, thereby no longer "musical sound" as he purports in his title. Furthermore, his use of the phoneme, while it is indeed a minimal unit of language, betrays his attempt to locate meaning in "linguistic sound," for "we never pronounce a phoneme, only its allophones."[35] Given that Johnson is implicitly then working with two purely theoretical units, it is no wonder that by the end of his work he finds that "the issue of timbre remains somewhat problematic and further study of this topic is necessary."[36]

---

[33] ibid, 275-6.

[34] ibid, 47.

[35] O'Grady, William, Michael Dobrovolsky, and Mark Aronoff. *Contemporary Linguistics*. New York: St. Martin's Press, 1989, 61.

[36] Johnson, 276.

The strength of Johnson's insights, however, lies not in any of the topics he presents in his dissertation's title, but in his theoretical grounding. His very careful approach to semiotics will guide my ultimate conclusions at the end of this dissertation. Johnson's final revelations, which, robbed of their extensive basis and preamble seem humble, are that music and language, mediating to an individual in his or her surroundings, generate meaning in "a series of dynamic and multifaceted relationships."[37] Because the human voice is, in many ways, a significant — if not *the* significant — meeting point of music and language, it is important then to keep Johnson's methods and conclusions, if not his particular focal arguments and occasional oversights along the way, never too far from my own explorations of the voice's timbral meaning.

The final source I wish to explore in depth at this point is possibly the most often cited writing on the meaning of a voice's sound. Roland Barthes, in his landmark essay, "The Grain of the Voice," explores "*the encounter between a language and a voice*."[38] Barthes is heavily concerned with the metaphorical and real body from which the voice emanates, and in perhaps too freely moving between the physical and the metaphorical, he leaves us with a series of contradictions as to the nature of the grain. The *geno-song*, with which Barthes aligns grain, is

> the volume of the singing and speaking voice, the space where significations germinate 'from within language and its very materiality'; it forms a signifying play having nothing to do with communication, representation (of feelings), expression; it is that apex (or that depth) of production where the melody really works at the language — not at what it says, but the voluptuousness of its sounds-signifiers, of its letters — where melody explores how the language works and identifies with that

---

[37] ibid, 279.
[38] Barthes, Roland. *Image Music Text*. New York: Hill and Wang, 1977, 181.

> work. It is, in a very simple word but which must be taken seriously, the *diction* of the language.[39]

In creating an eloquent diagram of the voice's inadvertent and inescapable musical signature, Barthes points to tangible landmarks by which one identifies grain — volume, melody, letters, diction — but in establishing the concept of the grain, he is quick to decentralize its meaning, arguing "The 'grain' of the voice is not — or is not merely — its timbre; the *significance* it opens cannot better be defined, indeed, than by the very friction between the music and something else, which something else is the particular language (and nowise the message)."[40] Though linguistic analysis of a singer's phrasing could in fact help us to understand where in the sound the grain is encoded, and through timbral analyses can show the subtleties of inflection, Barthes's implication is that the grain is connected to a physicality that sound alone cannot portray: "The 'grain' is the body in the voice as it sings, the hand as it writes, the limb as it performs."[41] Despite being disconnected from its artifact — the performer's body no longer physically present in the recording, the hand in the penmanship, nor the limb in its effect — the grain is not an imagined characteristic, but something felt in the text's actual creation and thereafter in the resultant text itself perceived as a specific absence. These missing nuances ("nothing is left but the pheno-text"[42]) are "in no way 'subjective' (it is not the psychological 'subject' in me who is listening."[43]

---

[39] ibid, 182-3.
[40] ibid, 185.
[41] ibid, 188.
[42] ibid.
[43] ibid.

Being so concerned with “a different history of music from the one we know now,”[44] Barthes’s work, while a valuable resource to my own (and one to which I will later return), heads down a path whose eventual destinations lie more in our understanding of music as a new kind of text (particular with regard to “where” is music) than in understanding how we interpret it. That said, in his unscientific language, he outlines a series of ideas that I believe to be shockingly prescient when viewed in the light of this dissertation’s final chapter.

One clear implication of Barthes’s writing is that especially with regard to timbre and voice, the crucial moment of signification in musical semiotics occurs not at a score’s textual level of notation, nor in the idea of a piece abstractly pondered, but instead at the moment of performance or playback, for until then, music is only a possibility: a dormant neutral text is merely a roster of scheduled sonic events, and timbre is only the vaguest of potentialities in sound. From a perceptual point of view, music is not music until it is performed. This is in keeping with my earlier positioning of this project’s approach within Nattiez’s tripartition. No amount of notational intricacy can account for the extreme and relatively unpredictable complexity and flux in musical timbre, and from a semiotic standpoint, particularly in their capacity to connote the extramusical, the formant structures that create timbre, when compared to rhythm and pitch, can occupy as much — if not more — of the total semiotic package. If we allow our understanding of timbre to be one of the total content of sound waves in time, then in fact it completely subsumes all other attributes internal to a musical performance. Cogan explains, “spectral formations… *include* the melodic lines and registral fields of music’s deployment in space, its harmonic textures, its rhythmic patterns and proliferations in time, as well as its instrumental and vocal sounds, and their combinations."[45]

---

[44] ibid, 189.

That structures of sound and meaning interact and overlap in complex ways such as this is integral to a deep understanding of any text that involves signs, the standing of one thing for something else. Having already spoken of linguistic methods and analogies and of structuralism, I wish to prepare the reader in as efficient and painless a manner as possible for my maneuvers within and beyond this notion of structure itself having meaning.

Just as melody and rhythm are embedded within the timbral signal of music, and just as Chomsky argues for a generative grammar of language, many theorists in the mid 20th century began to conceive generally of content as a product of form, rather than form as the arrangement of content. This is the basic principle of poststructuralism, a method of interpretation that gives tremendous power of meaning to the architecture of a text, usually on a linguistic level, but also on any number of superimposed structures of meaning.

A musicological example of this is Kofi Agawu's book *Playing With Signs*, in which the author investigates works of the classical era (which he calls "Classic") on first melodic, chordal, and Schenkerian levels, then on levels of their full internality versus their signification of the world around them, and finally on the collective level of an entire historical era and style of music. He concludes his book with the seeds of a relating of the Classic era to the Romantic. This multi-tiered poststructuralist semiotic is an impressive testament to the usefulness and potential of viewing structure as a central node of meaning.

---

[45] Cogan 1984, 124.

Famously taking Saussure's structuralist ideas to their logical conclusions, Jacques Derrida in 1966 pointed out that if, like melody in timbre, all things could be viewed as structures within structures, then this central node of meaning is necessarily decentralized, and meaning is equally dispersed into the totality of interlocking systems.[46] Derrida called the investigation of meaning in this approach and the attempt to parse it out *deconstruction*.

Deconstruction is a potentially infinite unfolding of structures which shows that it is impossible to view, interpret, or otherwise encounter anything at all in an unmediated way, and that the explicit decoding of one system — music for example — necessarily gives way to another — frequently language. To explain any feeling or item or act is to present its existence in relief to all things that it is not, because its very system of existing is mediated by the existence of all other things, which by virtue of their being something else, allow this feeling, item, or act in question to be itself. This ultimate relativism gives particular value to the use of binary oppositions in the deconstructive process. Sound relies on silence, solidity relies on space, and black relies on white. In short, the meaning of any one thing depends on the meaning of all other things. Derrida calls this *différance*[47], a French pun that implies both that something's meaning is determined by its difference from all other things and that the heart of its actual meaning — what Derrida calls *trace* — defers its meaning to all other meanings on whose difference it depends as well as to yet another level of mediation. An example often given of deconstruction is that all words in a dictionary are defined in terms of other words that are defined elsewhere in the dictionary, and to look up the meaning of each word in each definition that one reads is an endless process that never brings the reader closer to the actual "meaning" of the words; it merely defers him or her *ad infinitum*. This does not imply that words or signs are meaningless, but instead that their meanings to us are always filtered through some kind of structure.

---

[46] Derrida, Jacques. "Structure, Sign, and Play in the Discourse of the Human Sciences." In *Criticism: Major Statements*, *Third Edition*, edited by Charles Kaplan and William Anderson. New York: St. Martin's Press, 1991.
[47] ibid, 534.

In critical practice, deconstruction is enacted in many ways. Offering one particularly simple example of deconstructive revelations in practice, Derrida cites in the discipline of ethnology, which

> could have been born as a science only at the moment when a de-centering had come about: at the moment when European culture — and, in consequence, the history of metaphysics and of its concepts — had been *dislocated*, driven from its locus, and forced to stop considering itself as the culture of reference…. [T]he ethnologist accepts into his discourse the premises of ethnocentrism at the very moment when he is employed in denouncing them.[48]

However, when applying deconstruction more intensively and concretely as a philosophical, linguistic, and literary technique, Derrida and his followers are keen on never quite explaining its precise process. This is because in their eyes, it is not a definable procedure at all; to systematize prescriptively is intrinsically against the deconstructive process, which, in her introduction to Derrida's *Disseminations*, Barbara Johnson calls

> an analysis that focuses on the grounds of [a] system's possibility… in order to show that these things have their history, their reasons for being the way they are, their effects on what follows from them, and that the

---

[48] ibid, 521.

> starting point is not a (natural) given but a (cultural) construct, usually blind to itself.[49]

It is easy to see already why deconstruction is considered by some to be a fruitless endeavor, a frustrating headache, or the pinnacle of navel-gazing. If we take it as an end rather than a means, this is understandable. However, in the years since it first came into academic vogue, deconstruction's political implications and subversive potential have receded, allowing for it to be viewed as a critical tool rather than a nihilistic panacea. Many agree that it is in such a respect that this extreme of poststructuralism is most rewarding. By revealing the mutual reliance of binary oppositions and the mutual embedding of semiotic structures, we can gain a better glimpse into the dynamic and multidimensional ways that texts can mean, if not into *what* they mean. To this end, while I do employ ideas of deconstruction and borrow its processes, this dissertation is in no way expressly deconstructionist. (I direct readers wishing to explore this approach deeply as an analytic basis for music to Marcel Cobussen's very thorough dissertation *Deconstruction In Music*.) Because my endeavor here however is to reach some semblance of at least temporary conclusion with regard to the topic at hand, my use of these methods is always toward an intellectual goal rather than as a celebration of meaning's elusiveness.

Having stated what I intend to accomplish in this dissertation, I wish to delineate briefly that which I will *not* accomplish and the methods I will *not* employ. My approach is philosophical at heart, rather than psychological. This means I do not offer experimental data nor have I conducted surveys of listeners' interpretations of vocal timbre in music. I also generally do not venture into questions of neurological perception in music. Furthermore, this is not an ethnomusicological study. While issues of culture are of course vital to our interpretation of timbre, they fall into a category of concerns I outline in this dissertation's final chapter as unpredictable in a general theory. Such codes of meaning are also peripherally addressed in my third chapter, but questions otherwise of politics, social context, community, or specific ethnography are best answered, I believe, in case studies of individual cultures' interpretations of vocal timbres.

---

[49] Derrida, Jacques. *Dissemination*. Translated by Barbara Johnson. Chicago: University Press, 1982, xiv-xv.

By outlining as culturally independent a model as possible, I recognize the generality and vagueness to which I surrender my conclusions, but it is my hope that the specificity sacrificed will be made up for in the model's potential to ground any number of insightful, specialized readings of musical works and audiences. Even the broadest reading of vocal timbre's musical meaning will, however, have limitations to its universality. My assumptions about the listeners and voices that interact within this model include the abilities of hearing, of recognizing human voices as human, of knowing the sound of one's own voice and the experience of using it. I assume a recognition of the body's existence on the part of the listener, and I assume a cultural context in which aural language is present and in which vowel and consonant sounds are phonologically meaningful. Furthermore, within the world-view of a listener, I take for granted the existence of silence and of sounds other than the voice. The vast majority of the world's inhabitants live by all these same assumptions. Of more situational concern but of less general import is the potential disagreement with certain assumptions of meaning in my musical examples, particularly in the second chapter. While the specific discussion therein addresses chiefly the structural moments of musical meaning and not any supposed truth of *what* the music means, thereby avoiding many potential pitfalls, recall again that my use of examples is solely to illustrate larger concepts; if my reading of the music's meaning is objectionable, it only affects the clarity of such an illustration, and not the points I seek to exemplify — ideas that stretch widely across genres and performances.

In the next two chapters I present spectrographic analyses of vocal timbre within short musical works. I first seek to demonstrate that meaning in vocal timbre lies heavily in the blank spaces between formants. Following that chapter, I illustrate how the timbre of the voice signifies the human body, and that by duly embodying a voice, a listener claims an active and actualized role in his or her phenomenological state of being. The analyses in these two chapters are conducted in preparation for my final arguments and explorations, in which I address the relation between vocal timbre and language, and in which I ultimately offer an explanation of the musical role that the human voice's timbre plays in meaning.

Before embarking on my analyses, I offer a brief technical note. Because my analyses are of recorded music, one might initially object to the lack of "authority" in any one performance of a musical work, and beyond that to the imperfections in both the performance and recording of a piece of music. While indeed microphone placement, room acoustics, mixing, editing, and mastering can all contribute as much to a recording's sound as the performers, instruments, conductors, and producers, to criticize the use of recordings is to miss completely the point I have repeatedly stated where I ascribe meaning to actual rather than potential experience. If we hear tape hiss in a sound, then we cannot entirely ignore its presence in the name of adhering to a Platonic ideal in which a piece's performance or hearing is merely a shadow on the cave wall. That said, the recordings I analyze do exhibit clarity of performance and engineering. In this dissertation, the "meaning" of timbre in the human voice is a gerund of a verb rather than a noun; it is a process into which I hope to gain a greater insight and not a destination I seek to reach. To this end, while I chiefly concentrate on issues of timbre, I do not rule out in my analyses the acknowledgment or use of music's melody, harmony, rhythm, production, or lyrical setting when these aspects contribute toward a stronger situational and universal understanding of vocal timbre. It is my hope that the insights illustrated in my analyses help to point ultimately to a musical semiotics of timbre in the human voice.[50]

---

[50] The spectrographs I use were created with Jonas Åström's *Frequency* software for the Macintosh, with the highest possible resolution and a noise floor of –72 dB. Additional analytic and plotting software included Seventh String's *Transcribe!* and Martin Hairer's *Amadeus II*.

2. Voice and Negative Formant Space

In this chapter, I propose that the process of meaning-making within vocal timbre is, at heart, a subtractive one in which the method and the contents of timbre's interpretation are determined at least as much by the frequencies *not* present in the voice from moment to moment as by those that are. To this end, I use the musical example of "O Superman," by Laurie Anderson, as an argumentative catalyst. My analysis of the song shows the ways it draws specific attention to registers of sound not present at specific moments in the human voice and how it aligns its own totality as a musical experience with those negative spaces. The purpose of this chapter is not to reveal insights about the song by way of spectrography and semiotics, but instead to use the song as a demonstration of insights about voice and meaning.

"O Superman" has been the focus of an intense amount of scrutiny and writing — both popular and academic — since its appearance in Anderson's 1980 epic performance art piece *United States*, its subsequent release as a single in 1981, and its inclusion on her 1982 *Big Science* album. Most of the commentary has focused on the cultural divide between aggressive postmodernism and popular appeal that, for eight minutes and twenty-seven seconds, the song bridges. Despite her Columbia University graduate education and *artiste* status, Anderson's song reached number two on the British pop charts, sold 800,000 copies worldwide, and was called by Robert Chistgau in *The Village Voice* “the most compelling pop event of the year.”[51]

Written during and with reference to the Iran hostage crisis, “O Superman” is lyrically oblique, and its musical delivery contributes to its sense of alienation — a word that arises over and over in the writing about the song. Though I do not wish to linger in extraneous details of the song, I reprint its lyrics here as a reference:

> O Superman. O judge. O Mom and Dad. Mom and Dad
>
> Hi. I'm not home right now
> But if you want to leave a message
> Just start talking at the sound of the tone
>
> Hello? This is your Mother
> Are you there? Are you coming home?
> Hello? Is anybody home?
>
> Well, you don't know me, but I know you.
> And I've got a message to give to you.
> Here come the planes.
> So you better get ready.
> Ready to go.
> You can come as you are, but pay as you go
>
> And I said: OK. Who is this really?
> And the voice said: This is the hand, the hand that takes
>
> Here come the planes
> They're American planes
> Made in America
> Smoking or non-smoking?
>
> And the voice said:
>   Neither snow nor rain nor gloom of night shall
>   stay these couriers from the swift completion
>   of their appointed rounds

---

[51] Christgau, Robert. “The Year the Rolling Stones Lost the Pennant.” *Village Voice* 27 Jan, 1982: 36-7. Quoted in Samuel Austin McBride, “Performing Laurie Anderson.” (Ph.D. diss., University of California Riverside, 1997), 204.

'Cause when love is gone, there's always justice
And when justice is gone, there's always force
And when force is gone, there's always Mom (Hi Mom!)

So hold me, Mom, in your long arms
In your automatic arms
Your electronic arms
In your arms
So hold me, Mom, in your long arms
Your petrochemical arms
Your military arms
In your electronic arms[52]

In "O Superman," Anderson engenders the exception to Barbara Johnson's claim of a text's cultural construct being "usually blind to itself." Immensely self-aware, Anderson, as she composes, plants the seeds of her work's own deconstruction by establishing, acknowledging, and then subverting binary oppositions. As Susan McClary asserts,

> Anderson's monologue causes us to map the alternations with certainty at first: Man/Machine, Home/Alienation, and so on. But then things become confused, as Mom becomes Machine, and the clichés of American patriotism become codes of totalitarian control.[53]

In revealing and playing with its own binaries, the song joins Barthes's cadre of "works in the course of which a system feigns self-interpretation."[54] Though not explicit, this self-interpretation places the voice at its crux; "O Superman" draws attention and gives primacy to the timbre of the voice, and through binary oppositions — in particular the "natural / technological" pair, which is manifested in several ways — the song illustrates how the voice's timbre, on its acoustic levels, behaves as a sign.

---

[52] Anderson, Laurie. "O Superman." *Big Science*. Warner Bros., 1982.
[53] McClary, Susan. *Feminine Endings: Music, Gender, and Sexuality*. Minneapolis: University of Minnesota Press, 2002, 143.
[54] Barthes, 179.

This “natural / technological” binary pair I do not call “man / machine” (as McClary does) for a few reasons. First, the presence of other natural sounds — non-human but nevertheless unmediated (save for their having been recorded) — are an important part of the song. Second, McClary and Kay Dickinson both heavily address the issue of gender with regard to Anderson’s voice, arguing that it is a central feature of the performance, and so to group her under the category of “man” is at best imprecise. I myself will not expressly delve into the question of gender in “O Superman,” though in making this choice I do not deny that it can be a part of the piece’s overall effect on a listener. Finally, I use “technological” to encompass both the electronic and the mechanical aspects of the song. Because the sounds this word describes are necessarily mediated by something inanimate — synthesizers, organs, flutes, saxophones — they cannot enter into the same realm of natural immediacy that the human voice or the calls of birds occupy.

This opposition of the natural and the technological is, from the song’s beginning, both highlighted and put under stress: a pure recording of Anderson singing “ha” once on middle C is repeated digitally in such a way that the unprocessed humanity of the sound is initially retained, but put under suspicion by its inhuman precision of repetition. This single unprocessed use of her voice becomes the basis of the song’s accompaniment, to which the foreground is Anderson’s lyrical and melodic delivery through extreme processing in the form of a vocoder. Already in the first seconds of the song, the standard of a natural foreground (voice) over a technological backing (instruments) is reversed. In relief against the natural “ha,” the vocoder’s technological sound, given its linguistic content, is a point at which “the situation becomes confused.”

What happens on a sonic level in “O Superman” is the contamination of vowel purity with the imported carrier frequency from the synthesizer as enabled by Anderson's vocoder. By introducing multiple bands of melodic sound to the spectrum of Anderson's voice at all times, regardless of vowel, she effectively limits the range of acoustic differentiation between speech sounds. In short, because her voice is consistently vocoded through the same carrier source in a limited tonal range, Anderson's words, relative to one another, are not as distinct as are words in regular speech.

While in spectrum analyses of other music containing clearly defined and unconfused vocal and instrumental sounds we can see an overlap of vowel-oriented frequencies between sound sources, in our hearing of them we are almost always able to differentiate clearly between separate sound sources, thus allowing us to parse the voice's frequency content without spectral homogenization. This is largely because the envelope of the instrumental spectrum (including attack, release, and dynamics) are in almost no piece of music exactly aligned with the voice's envelope, thus giving listeners a temporal and dynamic means of differentiation. Furthermore, both on recording and in live performance, our hearing is very likely to triangulate a special placement of the two or more sound sources in a way that makes clear the individuality of the voice.

Unlike an unprocessed voice act, vocoders necessarily conflate both the space and the envelope of natural vocal utterances with their external spectral contamination. Dickinson describes the effect of the vocoder as “superimposing a ghost… over an instrumental line,”[55] in opposition to which Marcel Cobussen’s aside, “But what is the presence of a ghost? The opposition presence-absence does not seem able to adequately define spectrality"[56] seems all too apt. A more complete explanation of the vocoder's mechanism than Dickinson's is as follows:

---

[55] Dickinson, Kay. “‘Believe’? Vocoders, Digitalised Female Identity and Camp.” *Popular Music* 20/3 (2001), 333-4.

A sound (speech, for example) is dynamically analyzed for frequency content across a spectrum. This spectral analysis can be thought of as a three-dimensional graph, with frequency on the x-axis, amplitude on the y-axis, and time on the z-axis. While this graph is the output of the initial sound — the modulator — it is then used as a processing filter for a new sound (synthesized strings, for example), which is called the carrier. This filter serves as a dynamic amplifier and attenuator of frequencies as imposed upon the native spectrum of the carrier, much like a graphic equalizer on a stereo system. By squeezing into the mold of the modulator's spectral shape but with its own pitch and harmonic content, the synthesized strings in this case would appear to speak the words of the modulator sound. The clarity of the modulator's spectral imprint on the carrier (in this case the clarity of the speech) depends both on the resolution of the harmonic increments by which the modulator is analyzed and by which the carrier is in turn filtered, and on the degree to which the filter is applied, that is, the value of the amplitude on the graph showing the analysis and filter. In many practical uses of vocoders, this amplitude by which the carrier is filtered is diminished in such a way as to give the illusion of a purely vocoded sound being combined to varying degrees with the original modulator (in popular music, almost always a human voice) to increase both intelligibility and the pitch of the modulator, which is expressed in its lower harmonic content and is very frequently otherwise masked by the carrier.

Therefore in vocoded sound, such as that in "O Superman," the envelopes of the voice (modulator) are precisely in line with the contaminating frequencies (carrier) because they are not the additive combination of sounds, but rather the *dynamically subtractive* mediation of one sound by another, in which the modulator's total silence in one or all frequency bands means the total subtraction of the carrier's content in the respective frequency band(s). This means that when either input in an unattenuated vocoder is silent, the output is necessarily zero. As Friedrich A. Kittler describes it, "one and the same controls one and the same: one acoustics controls the other."[57]

---

[56] Cobussen, Marcel. "Deconstruction in Music." Ph.D. diss., Erasmus University Rotterdam, 2001, 167-168 / IV, 38.

The vocoded sound is thus complicated not only on purely acoustic or culturally associative levels, but also on a structural Derridean level. Its very mechanism is grounded in difference, one signal relying on the other, and yet with the hegemony of the modulator signal's ability to convey linguistic meaning, we privilege it over the carrier, whose chief role is tonal. In Derrida's terms, language is far more transcendental a signifier than tone. The physiology of the vocal apparatus designates the linguistic sibilance and distinguishing features of speech sound to the mouth, whose movements do not so much create sound as filter it, while the larynx is the origin of the vocal sound.

Wayne Slawson provides us with the terminology that implicitly connects the purely acoustic human production of speech with the electronic model of sound generation that we find in analogue synthesizers: the source / filter model, in which "The independence of the source and filter and, at the same time, the modification of the source by the filter are the essential features," and where "Sound color is associated with the filter, not the source."[58] Slawson speaks of sources and filters being weakly- versus strongly-coupled. Sound color is independent of pitch in weakly-coupled systems, but strong coupling, which occurs between the source and filter in most pitched instruments, allows the filter to impose so strong a spectrum envelope upon the source sound that it changes the fundamental frequency. Cupping a hand to one's ear at different angles to change the sound is an example of a weakly-coupled system of filtering, while moving a trombone slide changes the instrument's pitch without necessarily altering the vibration of its players' lips. Despite this, even weakly-coupled systems are not without limits of their independence between pitch and timbre. Resonant frequencies brought about by electronic filters, or by the shape and size of the human vocal apparatus can force breaks between registers and subsequently the altering of timbre. This limit of weak coupling has been one of the grounds on which some have questioned the usefulness of the source / filter model, reverting instead to articulatory phonology and modeling speech as a complex system of task dynamics. However, in addition to the source / filter model being well-supported in the linguistics community, if one allows for a flexible continuum between strong and weak coupling, then the theory can be assumed for this study. This is because neither the human voice nor any sound system is purely weak, for weak coupling relies on the independence of upper versus fundamental frequencies, and when the fundamental rises into a high register, above its formants, the sound color, in Slawson's words, could not remain the same.

---

[57] Kittler, Friedrich A. *Gramophone, Film, Typewriter*. Translated by Geoffrey Winthrop-Young and Michael Wutz. Stanford, CA: Stanford University Press, 1999, 49.
[58] Slawson, 23.

Curtis Roads demonstrates further the technological implication of this terminology as he outlines the basis of subtractive synthesis in electronic music: "Subtractive synthesis implies the use of *filters* to shape the spectrum of a source sound. As the source signal passes through a filter, the filter boosts or attenuates selected regions of the frequency spectrum."[59]

By using technologically-oriented language to describe the human apparatus, Slawson suggests that the binary identities of nature and technology, or as Anderson puts it, Mom and her electronic arms, are in reality not so far apart. This precisely explains the blending of the analogue synthesizer and the voice — both subtractive source / filter models — into the vocoder, which is more complicated both semiotically and acoustically, by virtue of its dynamic (and thereby multidimensional) operation. Kittler extends the model, allowing us a broader view from which to consider "O Superman":

> Media facilitate all possible manipulations: acoustic signals control the optical signals of light consoles, electronic language controls the acoustic output of computer music, and in the case of the vocoder, one set of acoustic signals control another… Since machines have taken over the functions of the central nervous system, nobody can say whether the roaring comes from the blood or from the sirens, from the ears or from the sea goddess Amphitrite.[60]

Our experiential reliance on the source alone for the very existence of a sound — the filter merely shapes and resonates what is already vibrating — is confounded by the vocoder, and its presence empowers the filter explicitly in a way that we usually find only implicit. The centrality of vocoded sound in "O Superman" and the relative paucity of any sound at all other than Anderson's voice forces our attention on this issue of filtering. While filtering is, I propose, vitally important to timbral meaning in all music and sound, "O Superman" is an especially clear example of filtering's significance because through the song's juxtaposition of voice and vocoder (and the multiple layers of filtereing therein), its structure expressly highlights this timbral issue.

---

[59] Roads, Curtis. *The Computer Music Tutorial*. Cambridge, MA: The MIT Press, 1996, 184-5.

[60] Kittler, 49-51.

As I have mentioned, the sparse analogue synthesizer sounds that fill the later portions of “O Superman” and that provide the signal by which Anderson’s voice is vocoded, are created through subtractive synthesis. With speech and analogue synthesis alike, though particularly with their confluence in the vocoder, we are hearing the process of filtering — of subtracting — just as much as we hear the sound itself. That we privilege the filter means that in our search for meaning in vocoded sound — where, as is not the case with the human voice or analogue synthesizer alone, source relies upon filter to be heard — we dwell in the negative relations between sound substances. Recalling the binary of nature and technology, meaning is not found then in Barthes’s “grain of the voice” nor in Adorno’s “curves of the needle,” but in the differential space between them.

Although the vocoding filter in “O Superman” is attenuated so that we hear a small amount of Anderson's unaffected voice as she speaks and sings, the minimal units of her vocoded sound are less distinct than in regular speech, and so the semiotic field must, in order to glean meaning from Anderson’s voice, either zoom in, and by focusing on a smaller range of sound, devote its attention to a higher resolution of perceived sound difference, or it must zoom out and create meaning from Anderson’s voice not based on how it is specifically affected by vocoding, but instead on its having been vocoded at all, thus taking the process — rather than its effects — as sign.

To regard the vocal effect in this latter way from a structural standpoint creates a rather dull reading of the song’s main features: a voice with timbral complexity and consistency moves over a pulse of the same voice with less timbral complexity, and the nuances of this relative complexity are solely relegated to our associations with vocal processing. Therefore we must look closely at what the particular differences between the processed and unprocessed voice are.

Below are the spectrographs of Anderson's unprocessed speaking voice — taken from the introduction to her piece "Born, Never Asked" — and of her vocoded voice in "O Superman." I present the same four vowel sounds from each example so that differing vowels may be compared as well as the same vowel in its unprocessed versus vocoded signal.[61]

---

[61] These vowels' spectra may differ from other sources' illustrations of idealized pronunciations. This is because Anderson's vowels are not uttered in isolation, but are lifted from actual speech, where a certain fluidity of sound and an individual's speech patterns may, as here, present vowels that are eccentric from wordless sounds, but that are ultimately comprehensible.

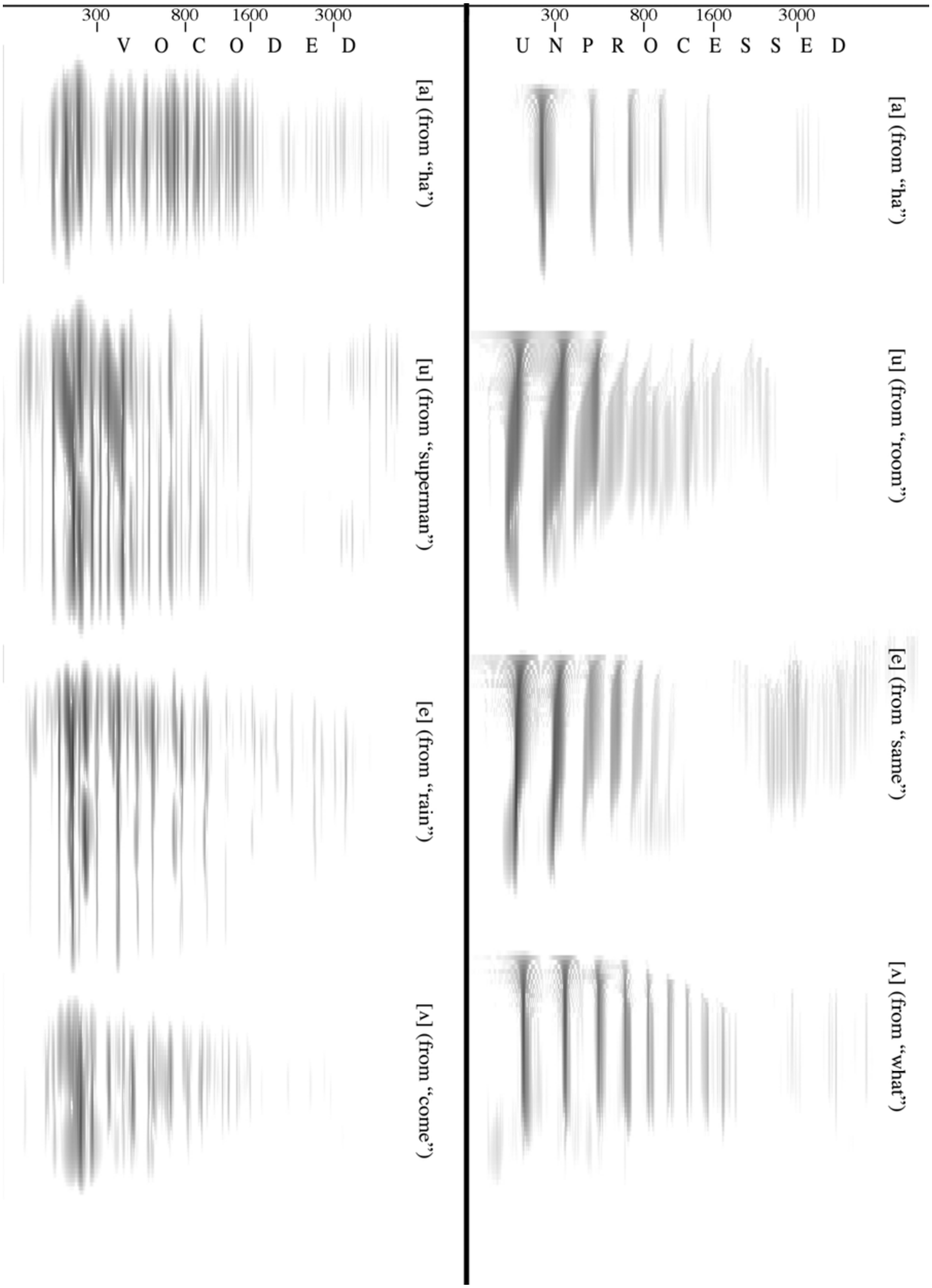


Figure 4: Comparison of vocoded and unprocessed vowels in Laurie Anderson's voice

In every case, the vocoded signal contains a greater number of discrete formants than the unprocessed voice. While the unprocessed formants' frequencies are generally separated from one another by differences at least equal to the fundamental frequency of the voice's pitch, the vocoded signal's formants are much more narrowly spaced by virtue of there being no single fundamental frequency by which to create a particular harmonic series; instead, the chords infused into the voice — always constructed from a {C E♭ F G A♭ B♭} set — overlay several harmonic series, the upper formants of which interact with, beat against, and mask one another in a complex system. That this occurs in the higher frequencies of the given spectra means that it overlaps with the upper harmonics whose chief function is in determining timbral (and phonetic) character. In short, simply by introducing additional tones to a sung speech-sound, not only is the distinct identity of individual formants weakened by a closing in and complication of these formants, but given that the timbrally determinant frequencies are obfuscated, the distinctness of the sounds from one another is lessened. This is most plainly visible in the spectrographs of the vocoded [u], [e], and [ʌ], which resemble one another more closely than their unprocessed counterparts. As I have suggested, the lack of relative clarity in formant structure among these vocoded sounds can directly lessen their capacity to create meanings that are individually different from one another.

Another critical effect of the vocoder's contaminations of Anderson's voice is that by filling in the spectral gaps between her naturally occurring formants, thereby giving more equal volume to a greater number of registers, it starts on a path toward harmonic saturation. Within the context of "O Superman" this can be defended by the song's culmination in harmonic saturation: the final ostinato, played on synthesizers and saxophones, swells registrally until it envelops the entire harmonic range of the song, holding it, as it were, in its electronic arms. With this harmonically dense goal fulfilling the spectral aggregate and thereby allowing the piece to end (following a brief and uneventful denouement), we must recognize the import of timbral range.

Below is the cloud of harmonic saturation that concludes and climaxes "O Superman":

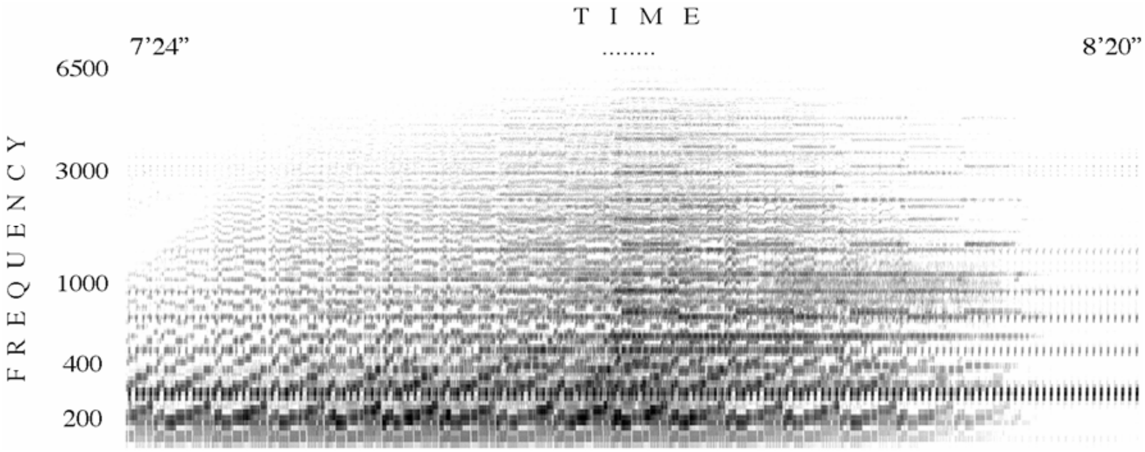


Figure 5: Harmonic saturation at the conclusion of "O Superman"

This usage of all registers is most explicitly foreshadowed by a striking combination of sounds. That the song ends with this unprecedented wash of sound pushes forward the argument that this fulfillment of the spectral aggregate is a goal to be achieved in the song.

The "ha" that permeates the piece has a spectral content such that all registers and harmonics from 261 Hz (middle C) through 1568 Hz (two and a half octaves higher) are activated, but for a full octave from 1568 Hz to 3136 Hz, all five natural formants in the harmonic series are at a negligible volume. Several frequencies immediately thereabove are strongly activated, giving the vowel its crisp aural character. Aside from "ha," only one sound in the entirety of "O Superman" is not electronically or mechanically generated (as the vocoder and woodwinds respectively are): at 1’23” into the piece, a three second loop of a birdsong enters the mix seven times, carrying with it no audible processing, but only its quintessentially “natural” character. The birdcall returns for another seven iterations at 7’02”, this time dovetailing into the final aggregate cloud of saturation. This birdcall is particularly remarkable in that its harmonic content dips and swoops in a region bound by 1600 Hz and 3200 Hz — the exact register left unsounded by “ha” (see figure 4), thereby creating, between the two natural sound elements in the piece, a spectral aggregate.

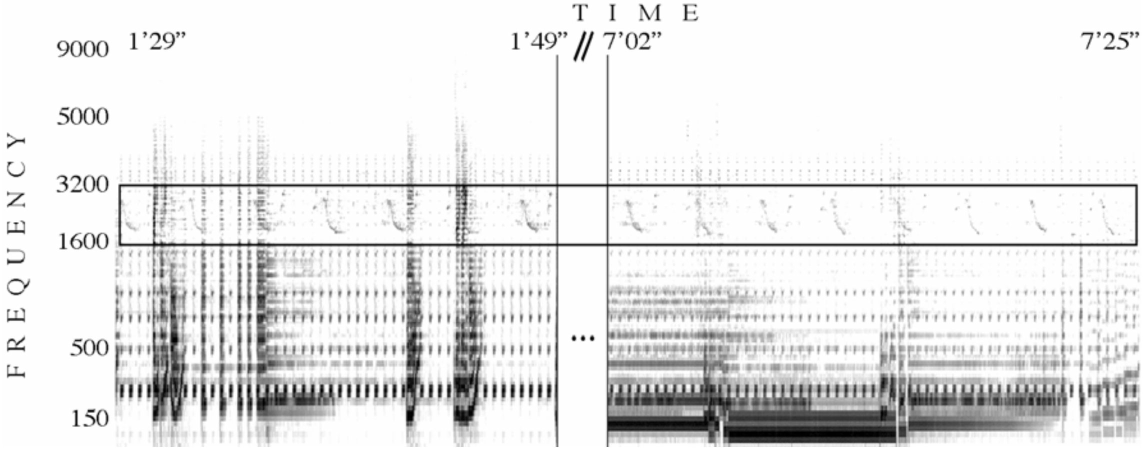


Figure 6: Detail of birdsong’s frequency range in “O Superman”

It cannot be a great surprise that the birdsongs appear first when the character of the mother starts speaking on the answering machine — the first lyrical confluence of human and machine — and stops on the line “Well you don’t know me,” when the mother is no longer speaking. Its second occurrence, immediately before the final cloud of saturation, comes with the invocation of “Your petrochemical arms, Your military arms, In your electronic arms,” where we again see the mechanization of the human.

Within the natural / technological binary, we are drawn to the piece's machine-based culmination by the suggestion of what is lacking from the human voice. Climax and emotional arrival are achieved by filling in the frequencies that the voice cannot fill in for itself, highlighted by the birdsong.

This serves as an active illustration of meaning in vocal timbre, highlighting the frequencies not present. Functioning subtractively as a source / filter instrument, the voice, in its formant structure from moment to moment, allows meaning via difference between sounds. Given that a pitch's source, vocal cords in this case, will activate the same harmonic formant structure from one sound to the next before filtering, thus allowing difference (within a given range of normal "chest voice" speech and singing) only in the frequency of the fundamental — whose value is phonologically immaterial in the vast majority of languages[62] — we absolutely cannot view meaning as constructed by the frequencies present, because their presence is a given; rather, timbral meaning lies in the absent formants and is therefore subtractive. "O Superman," through its highlighting of the voice and of difference, as well as through the ubiquity of its vocoder, continuously reminds us that the voice's naturalness — that which allows it to mean on its own — on which the vocoder preys and whose filters it contaminates with technology, dwells in the silent gaps between formants.

---

[62] On page 29 of *Studies In Musicology 1935-1975,* Charles Seeger writes, "In the reasoned mode of speech usage reliance is solely upon tonal density; for it is in terms of this resource or function alone that the consonances and vowels produced by the human voice are formed into the words that enable us to name and write about, as here, not only this tonal density but about all the other things we name or speak. It makes no difference whether the law of gravity is enunciated in a high or low voice, softly or loudly, rapidly or slowly, in even or uneven metrical feet."

3. Voice, Body, and Being

The musicologist Elisabeth Le Guin relates a story of an undergraduate class she taught where she played a record by U2 in which the singer, Bono, spends the better part of the song wailing loudly, straining his voice. After the students listened to the song, one commented, "That made my *throat* hurt."[63]

In this chapter, I wish to investigate how, timbrally speaking, a hearing of the human voice implies, signifies, and actualizes the human body. In response to Susan McClary's assertion that "By far the most difficult aspect of music to explain is its uncanny ability to make us experience our bodies in accordance with its gestures,"[64] it is my proposal that in listening to the human voice we might insert ourselves into a fantasized version of its performance. While a drummer may inadvertently tap a beat to a song on the radio, such engagement of the body into music need not be kinetically enacted. Rather, in the case of vocal music, there is a kind of meaning at work in the listener's plugged-in fantasized tactile experience of a singer's oral contortions and vocal straining and relaxing. Suzanne Cusick calls this "the listener-as-mental-performer,"[65] citing Edward Cone's work[66] as an influence on the idea's development. My own impetus for arguing such a hearing is rooted less in rigidly musicological writing, and more significantly in the work of the 20th century French philosopher Maurice Merleau-Ponty.[67]

---

[63] Le Guin, Elisabeth. Personal communications with the author on 11/19/04 and 3/4/05.
[64] McClary, 23.
[65] Cusick, Suzanne. "Feminist Theory, Music Theory, and the Mind/Body Problem." In *Music/Ideology: Resisting the Aesthetic*, edited by Adam Krims. Amsterdam: Overseas Publishers Association, 1998, 47.

With regard to the phenomenological approach I believe necessary for considering timbre, Merleau-Ponty eloquently articulates his compatibility with the same grounds of performance's and timbre's primacy in musical meaning from which I make my arguments:

> Musical meaning... is inseparable from the sounds which are its vehicle: before we have heard it no analysis enables us to anticipate it; once the performance is over, we shall, in our intellectual analyses of the music, be unable to do anything but carry ourselves back to the moment of experiencing it.[68]

Like Martin Heidegger, Merleau-Ponty was a follower of fellow phenomenologist Edmund Husserl, but Merleau-Ponty takes Husserl's incorporation of the body into the "lifeworld" (*Lebenwelt*) a step further, focusing on the recognition that all experiences are mediated through the body, even those — and especially those — that we cannot articulate, or intellectualize. Because the body and its senses are the means by which we perceive, act, and exist, Merleau-Ponty is expressly against the dualism of the Cartesian mind / body opposition that McClary asserts “has plagued Western culture for centuries”[69]:

---

[66] Cone, Edward. "The Authority of Music Criticism." *Journal of the American Musicological Society* 34 (1981): 1-18.

[67] One peripheral goal of my writing this chapter is to present Merleau-Ponty as a valuable resource to modern musicology, as his under-studied and under-cited writing is, I believe, highly relevant, particularly to scholars concerned with gender, sexuality, and gesture in music.

[68] Merleau-Ponty, Maurice. *Phenomenology of Perception*. Translated by Colin Smith. London: Routledge & Kegan Paul Ltd., 1981, 182.

> It is perceptual experience which gives us the passage from one moment to the next and thus realizes the unity of time. In this sense all consciousness is perceptual, even the consciousness of ourselves.[70]

And yet in spite of this, because “Perception is not a science of the world, it is not even an act, a deliberate taking up of a position,”[71] as he writes in the preface to *Phenomenology of Perception*, we spend passive the vast majority of our physical and sensible time. James B. Steeves, in *Imagining Bodies: Merleau-Ponty’s Philosophy of Imagination*, clarifies Merleau-Ponty’s notion of *anchorage* as “the experience of being in the world in a way that cannot be measured or defined.”[72] Steeves offers an example:

> When an individual looks at himself in a mirror, the most that he is able to see of his body is an image, or externalization, that fails to represent the experience of anchorage, or his particular experience and perspective from where he stands in the world. This experience of anchorage cannot be gauged or observed by others, either, since a person standing in the same room as the man in front of the mirror would only view his body as seen, and not as it would be experienced.[73]

This mode of being is, in Merleau-Ponty’s view, not reconcilable with the view of intellect and thinking being purely metaphysical processes. In light of this anchorage, he calls on us to reframe our notions of the mind:

---

[69] McClary, 151.

[70] Merleau–Ponty, Maurice. *The Primacy of Perception and Other Essays*. Translated by J. M. Edie. Chicago, Illinois: Northwestern University Press, 1964.

[71] Merleau-Ponty 1981, x-xi.

[72] Steeves, James B. *Imagining Bodies: Merleau-Ponty’s Philosophy of Imagination*. Pittsburgh: Duquesne University Press, 2004, 15.

[73] ibid

> We said earlier that it is the body which “understands” in the acquisition of habit. This way of putting it will appear absurd, if understanding is subsuming a sense datum under an idea, and if the body is an object. But the phenomenon of habit is just what prompts us to revise our notion of “understand” and our notion of the body. To understand is to experience harmony between what we aim at and what is given, between the intention and the performance - and the body is our anchorage in the world.[74]

This is a difficult challenge, however. When we remove “understanding” from the realm of the intellect and into the realm of bodily experience, we are asked, in a way, to downgrade it from an active thought process to a passive way of being.

I believe that the voice offers a solution to this problem, however, allowing one to experience physical being in an active rather than merely perceptual way. Because the senses are input mechanisms that receive without exuding, seeing one’s own body or touching one’s own body is at best an act of will only in the decision to look or to put one hand on another. We cannot control whether we appear or whether we are tangible, nor can we directly manipulate with immediacy and wide versatility the way we look and feel. We cannot control with precision whether we taste or smell, and we certainly cannot control by internal will alone how we taste and smell.

It is only in the sense of sound and hearing that we can both receive signals and have a dominant control over our output of them, namely by way of the voice. In this way, when we hear our own voice, it is not analogous to seeing our eyes or touching our hands. It is instead an act of our own will that we both control and bodily perceive, allowing us, by controlling our body and its perception, to control our own anchorage, for indeed “The body is the site for the institution of a meaning of Being.”[75] Through the voice, anchorage becomes active, and we seize control of our own being and we can understand “understanding.” As the philosopher Charles Taylor states, “Our understanding is itself embodied.”[76]

---

[74] Merleau-Ponty 1981, 144.

Merleau-Ponty gives special attention to tactile sensation in his work because it indicates the direct contact with the body.[77] The voice is especially potent in this capacity then because the process of making sound — of speaking or singing — is not merely one that we hear as we enact it, but it is felt: for a split second, our pharynx and diaphragm rise to attention, and then a buzzing spreads from our throat almost instantly in all directions, resonating high in our chest and gently on our lips; we lower a pitch and the epicenter of vibration sinks several inches down the front of our body; the vibration smooths into a sensation that is both faster and softer when we flex where our neck and mouth conjoin, pushing both this feeling and our pitch upwards; we can sculpt the sound with our tongue and restrain it with our teeth. All the while, just below the threshold of feeling, our voice feeds back into “the entire body, through the eardrums initially as they are made to vibrate by the sound, and eventually through the rest… as it responds” to its own signal.[78] These sensations are known both in the muscles that govern them and the nerves that receive them. It is no misnomer matter that singers speak of “chest voice” and “head voice.”

---

[75] Steeves, 156.
[76] Taylor, Charles. *Philosophical Arguments*. Cambridge, MA: Harvard University Press, 1995, 170.
[77] Taste also of course suggests direct contact, but on its own it tells us nothing *about* this contact. It also provides data about only certain things with which we have contact; to describe the taste of water, fabric, or stone inevitably surrenders to a tactile account of texture
[78] Steeves, 64-5.

Knowing this experience of our own voice so intimately in a more active and controllable way than when we experience our own appearance, for example, then we can relate more deeply to hearing a voice than can Steeves's external observer of the mirror-gazer relate to the narcissistic body in the room with him. Furthermore, it is not only the hearing of a voice to which we relate, but as Cornelia Fales observes,

> Confident of their auditory acuity, listeners feel themselves directly and aurally linked to a source in the acoustic world. So strong is the source orientation of environmental listening, that listeners project the fundamental premise of their auditory logic onto the data it is meant to interpret and subjective auditory sensation onto a world of sources until sound equals source. We say — I hear a *cricket;* not — I hear a *sound* that may indicate the presence of a cricket.[79]

Therefore when we hear the sound of a body as I have described, we are *hearing a body*, and beyond that, we know that we are hearing a body in the act of anchoring itself. We hear another's sense of touch, and are made aware, if only in our bodily subconscious, of our own, and thereby the possibility of our own actualization of anchorage. Hearing a voice is a kind of aural voyeurism: we are aroused to realize ourselves when we observe another doing so. In this capacity lies one important way that, as Dickinson says, "The expulsion of feeling through the voice, through visceral bodily vibrations, consequently bears the potential to trigger sentient responses within the listener."[80] The human voice, in its special primacy over being that other sounds and human traits lack, invites us into a self-aware mode of being through embodiment, not only of the singer or speaker's body or of our own, but also in a grander sense. This is the sort of hearing that Steeves means when he says that it "requires a virtual body that positions itself in the various dimensions of the sensible, exploring the timbre of certain sounds and the silences between them."[81] The practical result of this is the ability to plug ourselves not merely into the sound of the performance, but into the performer's bodily delivery of it, and on the level of muscular impulse we empathize with the singer's tensions, inflections, and physical articulations. Beneath the surface, we not only dance to the timbre, but we play a silent air guitar with our voice, and like bats in a cave, here our interpretation of the real is illuminated by the experiencing of our own echoes.

---

[79] Fales, Cornelia "The Paradox of Timbre." *Ethnomusicology* 46 no. 1 (Winter 2002), 63.

[80] Dickinson, 336.

Just as Steeves writes "The flesh is essentially the pure imagination, pure possibility, that does not await actualization but rather bears it as a mother bears her child,"[82] I believe that this sort of embodiment is not the mere will or choice of a listener, but is a necessary presence into which one is simply more or less consciously aware of his or her own tapping; different performers, listeners, and environments affect individual hearings, but the potential to actualize is always there. Merleau-Ponty describes this: "There is a human body when, between the seeing and the seen, between the touching and the touched, between one eye and the other, between hand and hand, a blending of some sort takes place."[83]

Though all vocal sounds that we identify as such have the capacity to invite this imitative embodiment, it is those with the greatest connectivity with the tactile that make most explicit this process. As if channeling Merleau-Ponty, Dickinson connects the bodily-ness of a voice with its capacity to enact meaning: "certain vocal conceits are cherished as exceptionally direct conduits to the core of the self, to some sort of emotive truth, with Bob Dylan's scratchiness, or James Brown's grunts winning more of these types of prizes than the smooth, non-grating and physically less aligned vocal offerings of the likes of ABBA."[84] The timbre of some voices such as Leonard Cohen's or Janis Joplin's suggest so prominent a physical vibration of the throat — either through extremely low fundamentals (hence slower and more individually felt vibrations) with noisy (unmusically breathy) upper partials or through a subharmonically modulated and spectrally focused voice that Ingo Titze calls *ventricular*[85] — that they are understood as bodily emanations as much as musical expressions. To make most clear this chapter's central argument, I investigate a voice well-known for its emotive capacity in this regard, that of Louis Armstrong.

---

[81] Steeves, 65.
[82] ibid, 151.
[83] Merleau-Ponty 1964, 163.

By the time "It Takes Two To Tango" was recorded with Sy Oliver's Orchestra in 1952, Armstrong was already long considered a jazz great, known for his exceptional trumpet playing and unique singing. For the purposes of my spectrographic analysis, it is fortunate that his voice was so cherished, for because of this, it is mixed at a noticeably higher relative volume to the orchestra than was typical for many vocal jazz records. The resulting spectrograph displays his singing with unambiguous clarity. Here a selection of half the song's refrain follows the lyrics it portrays:

> But it takes two to tango
> Two to tango
> Two to really get the feeling of romance[86]

---

[84] Dickinson, 335-6.
[85] Titze, Ingo R. "Definitions and nomenclature related to voice quality." In *Vocal Fold Physiology: Voice Quality Control*, edited by Osamu Fujimura and Minoru Hirano. San Diego: Singular, 1995.
[86] Armstrong, Louis. "It Takes Two To Tango." *Louis Armstrong's All Time Greatest Hits*. MCA Records, 1994.

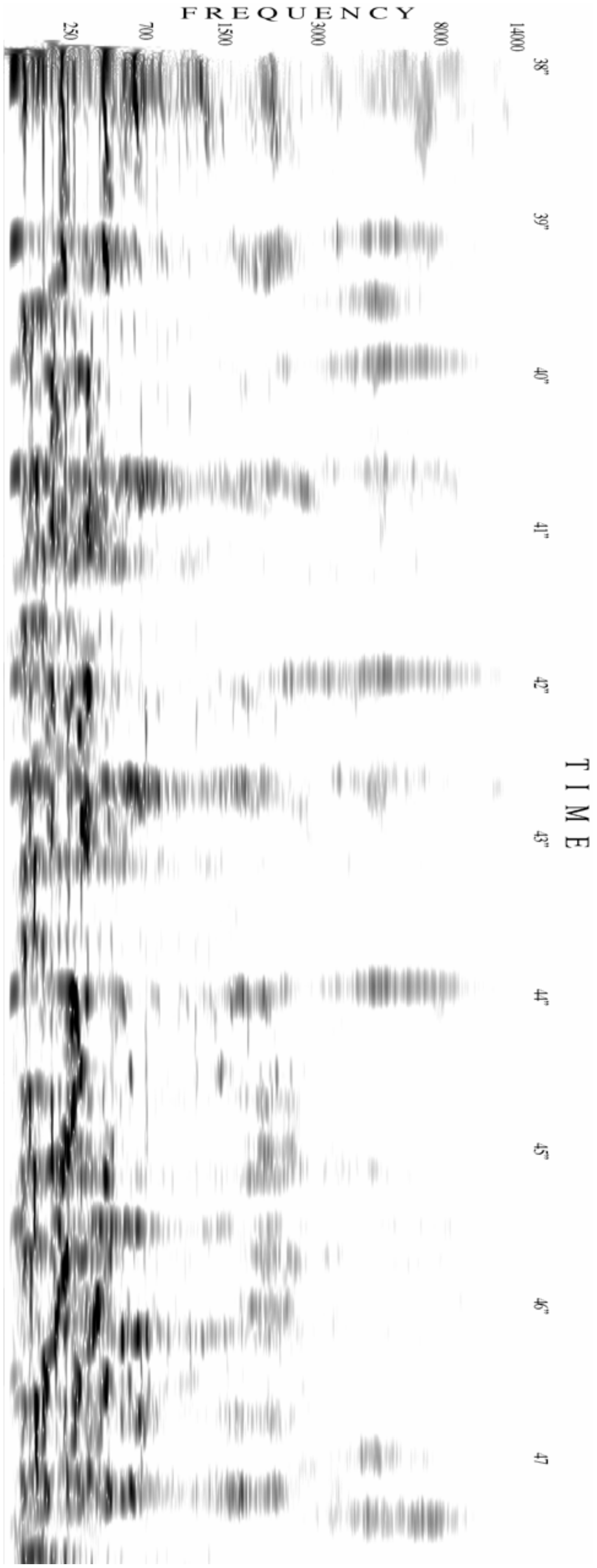


Figure 7: Spectrograph of the refrain in Louis Armstrong's "It Takes Two To Tango"

One of the immediately striking aspects of this performance is its tremendous richness of lower partials. As I shall show, it is these lower frequencies that suggest tactility most potently. This is because it is in these lower regions of vibrational frequency that we experience the strongest tactile sensation, which applies both to the direct contact with a vibrating body and to the physical experience of our own vocal cords oscillating. Handy Oey and Volker Mellert demonstrate this valence in sensitivity toward the bottom of the audible (and singable) spectrum. First they outline the different ways we feel vibration, explaining several nerve receptor types which all react to periodic movement:

> The Meissner receptors (RA-I) are responsible for perception of the velocity of the skin deformation, used to control the strength of pressure with which a certain part of the skin touches a surface or grabs an object. The Pacini Corpuscels (RA-II / PC) are responsible for accelerations in the skin deformation with highest sensitivity at about 100-200 Hz and serve for the perception of roughness (even when touching a surface with a tool)…. The RA-I and the PC sensitivity curves [across the frequency spectrum] are… narrow and quite steep. All receptors contribute to the sensation of vibration.[87]

Then they offer the conclusions of their experiments, in which they test tactile response to a broad spectrum of vibrational frequencies: "The PC system has its highest sensitivity in the range 125-160 Hz independent of probe diameter or hand location. The RA-I system is most sensitive at 40 to 80 Hz."[88]

---

[87] Oey, Handy and Volker Mellert. "Vibration Thresholds and Equal Vibration Levels at the Human Fingertip and Palm." Presented at The 18th International Congress on Acoustics, 2004, Kyoto, Japan, 1.
[88] ibid, 4.

In the following spectrograph, I show only frequencies under 500 Hz in hopes of achieving a greater clarity of the spectral content in the range that Oey and Mellert describe. I have outlined the spectral activity of Armstrong's voice in the 40-160 Hz range and will momentarily discuss it. The range's unoutlined activity is the orchestra's playing, and is easily distinguished by the ear as not germane our hearing of the voice.

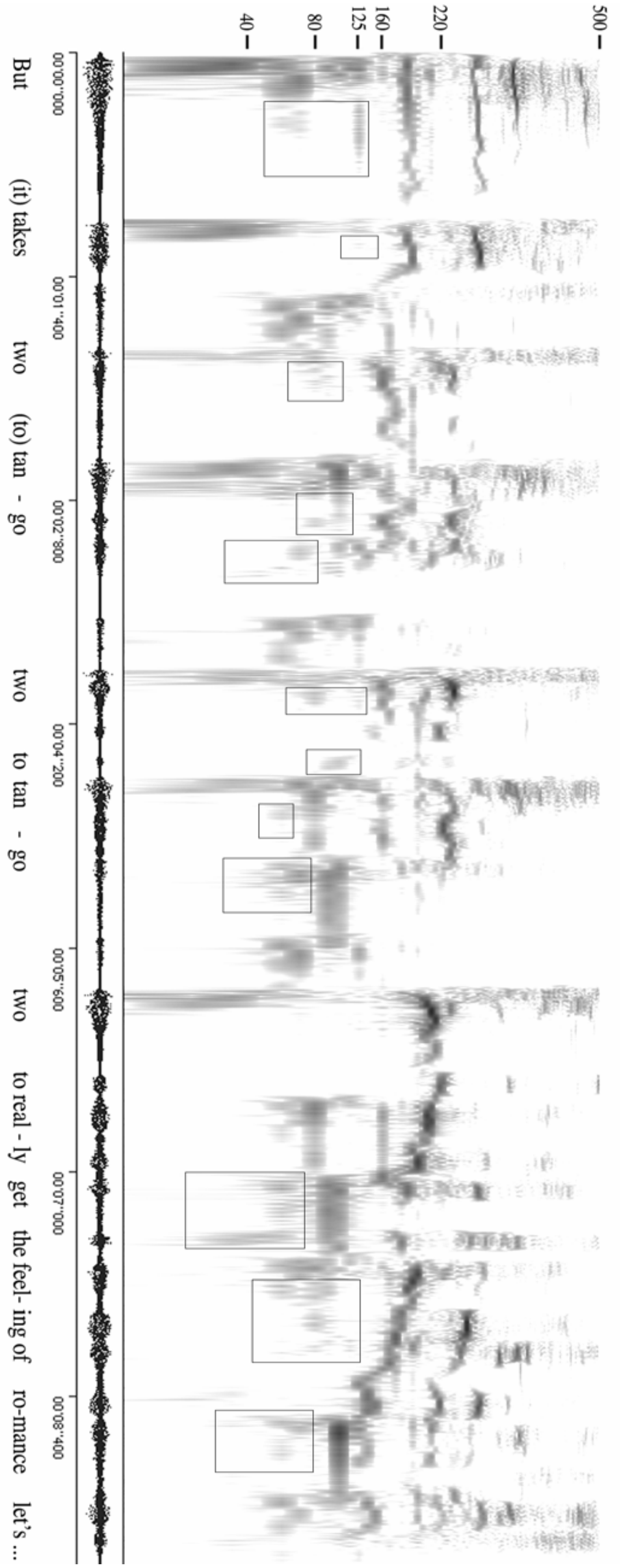


Figure 8: Detail of lower frequencies in "It Takes Two To Tango"

The selected areas on this spectrograph are generally not the fundamental pitches of notes in the melody. The low C and the D a ninth above it that outline the chorus offer a frequency range of 131 Hz to 294 Hz, and the dipping of fundamentals in the 125 Hz – 160 Hz range only occurs on the low C and D, which appear rarely in the melody. Instead, these tactile frequencies are occurring below the fundamentals, usually by an octave. This is more clearly observable on the following spectrum analyses of Armstrong's vowels in particular words during the song:

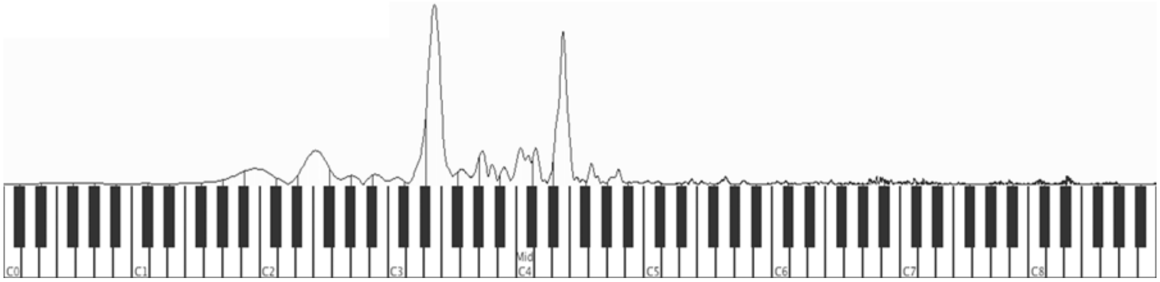


Figure 9: Louis Armstrong: [u] on F3, from "two" in "It Takes Two to Tango"

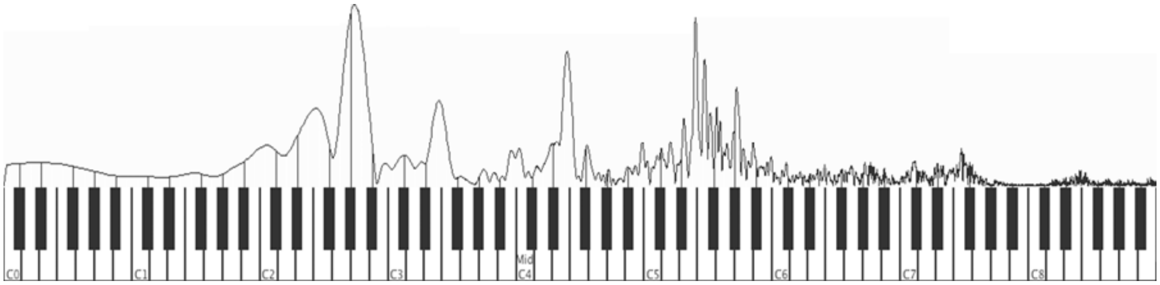


Figure 10: Louis Armstrong: [æ] on F3, from "tango" in "It Takes Two to Tango"

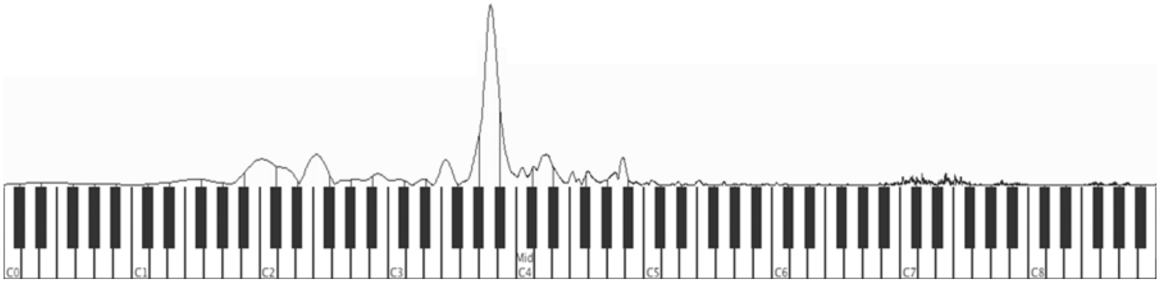


Figure 11: Louis Armstrong: [i] on A3, from "feeling" in "It Takes Two to Tango"

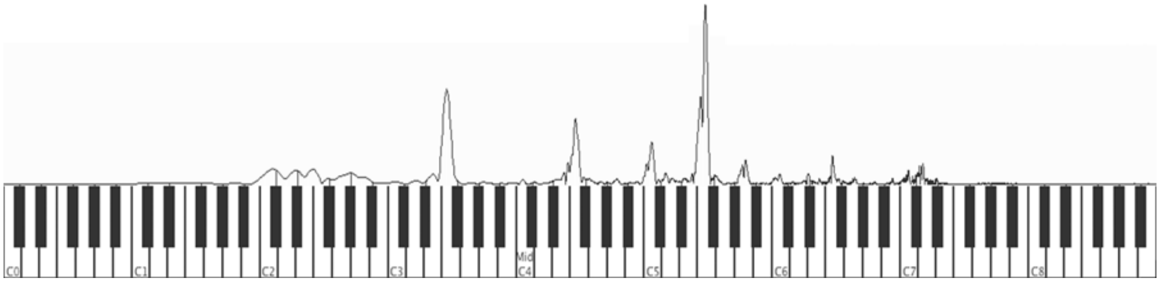


Figure 12: Louis Armstrong: [ʌ] on F3, from "love" in "It Takes Two to Tango"

These partials below the fundamental are called subharmonics, and they occur with frequencies equal to integer divisions of the fundamental (usually only noticeably occurring in the voice at fundamental/2). The vocal apparatus, like a wind instrument, is such that if the larynx's ventricular folds are even very slightly open, the resonations of the voice can pass back through the vocal folds, increasing the size of the resonating area and effectively multiplying the period of pitch, thereby adding a tone with a relatedly lower frequency. These are not difference tones nor are they the "real" fundamental of which the perceived pitch is merely an unusually strong overtone, but they are a performative imperfection that can arise to differing degrees depending on the throat and its situational physical behaviors in making sound. Below is the spectrum analysis of my own unaffected and unaccompanied singing of [u] on the same F3 (albeit slightly flat) on which Armstrong sings "two" (see above):

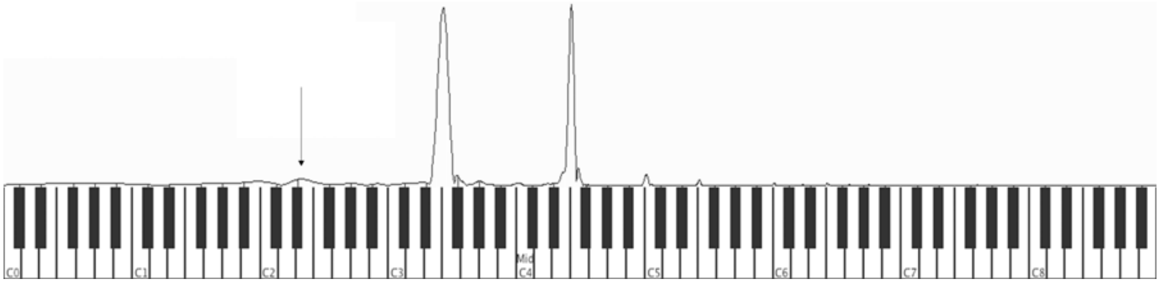


Figure 13: The author: [u] on F3, from "two"

Though much quieter, the subharmonic presence still arises as indicated by the arrow. My point is that while the character of subharmonics in Armstrong's voice is idiosyncratic, their presence in it is not unique, and with or without their assistance, the lower in pitch a voice moves, the more intimately close it is with this range of vibration to which our bodies are most physically sensitive. This is the resonating aspect of an embodied hearing of the voice.

In addition to this low frequency, in every spectrum analysis of his voice here, Louis Armstrong invokes an array of upper partials ranging from C7 to B8, and according to the first spectrograph shown, they extend upward even beyond that. The very high pitch and cluttered density of these signals suggest not an elegant harmonic ringing, but a bright noisy popping that sizzles atop the tangible rumble of his throat. In the model discussed thus far in this chapter, the interpretation of high noise in the voice is slightly different than that of low frequency content. In this case it is not the vibration that we are called upon to feel, but the muscular constriction of the mouth and throat that pinches air, finds resonant bands in whispered noise, and splays the high formants of our vowels across the spectrum. It is an articulatory reality that the highest formants to be found in vowels are those that require the greatest tensing of the mouth, and that the higher still frequencies in fricatives are the result of even more effort — "forming a nearly complete obstruction of the airstream so that when air passes through the small passage, turbulent airflow (i.e., frication) is produced."[89] In these frequencies, we then hear the unvoiced wind from Armstrong's lungs as it luffs into noise, the wetness rasping in his pharynx, and with it, the reliance of sound upon matter. Not only do we know the basic articulation of abstract phonemic speech sounds — Armstrong's voice contributes nothing unique to that — but through his timbre we know his throat's shape and the intensified motions of his mouth that exaggerate words in the song, like "romance." We can almost see his famous mouth widen on the word, but more to the point, we come even closer to feeling it. This is the articulatory aspect of an embodied hearing of the voice.

---

[89] Stewart, Thomas W. Jr. and Nathan Vaillette, eds. *Language Files*. 8th ed. Columbus: The Ohio State University Press, 2001, 492.

That I connect the potential to create meaning with certain timbral behaviors and frequencies present does not conflict my previous chapter's conclusions. Timbral meaning's occupation of the space between formants is a structural conclusion that owes to the nature of timbre, while this chapter's focus is on the relationship between humans and voice; here I look not abstractly nor theoretically, but concretely and experientially at this relationship, and so for the time being, my concern is not *where* meaning is, but *how* we understand it.

Between the aforementioned two basic hearings — the resonating and the articulatory — we bodily understand a voice through its production of sound and its shaping of it. The periodic vibrations of the low fundamentals and subharmonics ground the noise that hisses high above, contributing significantly toward what acousticians call *roughness*, even then linguistically connecting sound with tactility. It is from these real and identifiable aspects of timbre that, as Dickinson puts it, tactility's "meanings and consequences reverberate."[90]

These two hearings are available to us in every voice, no matter how smooth. While some, as earlier stated, present their body in greater tactile detail, the act of speech, of singing, of soundmaking is ultimately common to all people, and the sensations of the mouth, throat, skin, ears, and bone that it produces are irrevocably entwined with our mere recognition of the human voice. From a baby's cry to a death rattle, this component of the voice's semiotic is not beholden to culture[91] (although it doubtless affects culture), nor is its particular domain affected by the semantic content of the speech that employs it.

---

[90] Dickinson, 336.

When Merleau-Ponty writes of the "several ways for the human body to sing the world,"[92] he is discussing spoken language, but I am convinced that singing the world is possible without making a sound at all. When we listen to a voice enacting itself, seizing control over its body's perception and thereby making active Merleau-Ponty's anchorage, we cannot help but respond in a way beyond the scope of a simple mind / body divide: by taking in the sound's resonant vibration and articulatory muscularity, our own bodies become understanding, and we are aligned with and reverberated into the performance. This embodiment is an unavoidable possibility in the perceptual space between performer and listener, or as Louis Armstrong quips, "It takes two."

---

[91] Again, here I work under the assumptions presented at the end of the first chapter.
[92] Merleau-Ponty 1981, 187.

4. Conclusion: Meaning Revisited

There is no single process by which the human voice's timbre creates meaning in music, nor is there a single role it plays in our overall response to the music in which it appears. Certain meanings that vocal timbre enacts are directly related to one another and combine into larger semiotic machineries, while categories of others behave in ways so culturally and individually specific that we simply cannot speak with any authority about them; an individual's particular hearing of vocal timbre is not on any scale fully knowable. Vocal timbre still makes meaning, however. As with music itself, timbre is located at a planar dimensional intersection, one where we encounter the pitches, rhythms, and linguistic content it encodes. It is there that we can acknowledge its role as a common ground for a number of separate but interdependent levels of meaning-making.

This chapter's process is in three parts. First, I wish to argue for the necessity of a contextual view of vocal timbre, for indeed, this dissertation concerns its general musical semiotics, and not its full meaning. To this effect, it is important to view any musical voice act as a sign within larger systems. Building upon this, I will secondly outline this particular context of hearing and explain what we can safely assume about any encounter with timbre in the voice. Finally, I conclude this chapter, and with it the dissertation, by exploring the relation that the timbre of the voice has with the particular contextual space that surrounds it. I will incorporate the middle two chapters' conclusions into my final explanations, and will, I hope, be contributing a useful set of ideas to the field of musical scholarship both in timbre and in semiotics.

One cannot meaningfully study timbre, nor any sign at all, unto itself and without regard to those systems of concept and experience with which it is convolved. Indeed, one cannot even speak of timbre in total isolation; for it to be considered fully on its own is simply impossible. Throughout this study I have incorporated, when appropriate, elements of the deconstructive process. To deconstruct music is to reveal its systems, showing their individual self-referentiality and their mutually recursive relationships with one another, denying inherent “truth” or a simple and stable connection to reality. It relies on the exposure of conceptual oppositions in which the presence of one feature both suggests and relies upon the existence, textual or otherwise, of the other — a notion summed by Derrida's famous quip, “There is nothing outside the text.”[93] Deconstructing music must show not only that textual units exhibit *différance* with regard to one another, but it must show how this occurs and furthermore it must define those textual units so that we may agree upon the scale at which we speak of structural hierarchy, and in turn, the family of unstated text whose *trace* is suggested by the music at hand.

Derrida approaches deconstruction from a linguistic point of view, typically using words as his signifying system to manipulate and expose. This is a particularly easy choice because words, unlike phonemes (at one end) or paragraphs (at the other) have a less complicated reference to the reality they purport to signify than do minimal units or larger syntactical complexes.

In music, some systems such as pitch in the case of Western equal-tempered scales, are clearly enough defined — in this case, by twelve discrete pitch classes, organized into notes — that a deconstructive process might have a clear starting point. Other systems vital to music's meaning, however, such as timbre, and to a different degree, rhythm, have no such obvious organizational system. When composers such as Milton Babbit piece together *a priori* systems by which they create timbral, dynamic, and rhythmic structure, it is less a text to be deconstructed than it is “play,” because constructing such systems reveals a clear awareness of their self-referentiality. That the resultant music of this sort is so frequently avant-garde supports Christopher Butler's claim:

---

[93] Derrida, Jacques. *Of Grammatology.* Translated by Gayatri Chakravorty Spivak. Baltimore: Johns Hopkins University Press, 1976, 158.

> Postmodernist doctrines thus drew upon a great deal of philosophical, political, and sociological thought, which disseminated into the artistic avant-garde... The postmodern period is one of the extraordinary dominance of the work of academics over that of artists.[94]

Even when through reason we can assemble an *a posteriori* system of classifying into textual units whatever concepts we choose to analyze, we are still left with the difficulty that, excepting lyrics and direct mimicry of specific sounds from nonmusical settings, individual musical elements have no supposed precise correlation to a greater epistemology, and so meaning in music is a nebulous haze even before the winds of deconstruction blow wild on it.

This last point is of course nothing new. A great many musicologists have both bemoaned and celebrated music’s abstract nature. However, we must still be able to view music as a text to be interpreted and investigated; I hardly need say that simply because music’s meaning is encoded and unstable does not mean that music is meaningless, but that instead its meaning comes largely from the way that its parts relate to one another and to the structures surrounding the music. Implied in this approach is that this process of meaning making — and with it, all others — is part of the structure surrounding the music, and so by investigating the transference of meaning from within the music to outside of it, we of course are acting out its opposition where music contextualizes its externality, which in turn learns about itself by relating to a signifying text. This is the beginning of the process that ultimately relates the subtractive internal structure of timbre's meaning with the bodily external interpretation of it.

---

[94] Butler, Christopher *Postmodernism: A Very Short Introduction*. New York: Oxford University Press, 2002, 7.

The insight just reached through these deconstructive means is a valid one supported by a reading of musical timbre based on its structures of meaning and on our perceptions of them. The importance of structure within our reading of timbre is not only visible by way of deconstruction, which *de facto* displays the importance of structure in everything, but also through more traditional semiotics, the study of signs.

Semiotics has a long history, and while Aristotle gives us an early articulation of the notion of the sign in *On Rhetoric*, while Sextus Empiricus grants the subjective nature of the sign (i.e. not all signs signify the same things to all people) in *Against the Logicians*, and while contributions by Augustine and Locke significantly explore the classification and functions of signs, I will begin my own exploration with the work of the American philosopher Charles Sanders Peirce, filling in whatever crucial background information is needed along the way.

Peirce wrote a gigantic amount of semiotic theory from the mid 19th century until his death in 1914. Taken as a whole, his work is often self-contradictory, and its most saliently useful contributions lie scattered across the corpus. It is for this reason that I heavily bolster my reading and use of Peirce with the work of Bryce McCrary Johnson, Thomas Turino, and other scholars who have summarized and contextualized Peirce in ways that are particularly fruitful for this study.

Peirce's theory of semiotics is a universal one, dividing the signification process into three components — sign, object, and interpretant — each of which can occur on three levels, namely firstness, secondness, and thirdness. Momentarily, I shall give a deeper explanation of these sets of three, which are referred to as Peircean Trichotomies. Several further trichotomies are implied in the relations between the components of signification and their depth.

The circularity and self-reflexivity suggested in trichotomies categorizing one another and defining relationships between others is intentional; Peirce founds his model on the notion that

> Semiosis involves a type of chaining process through time in which the interpretant at one temporal stage becomes the sign for a new object at the next stage of semiosis, creating a new interpretant which becomes the next sign in the next instant, ad infinitum until that "train of thought" is interrupted by another chain of thought, or by arriving at a belief or conclusion.[95]

This approach of constant chaining achieves what a simple two-part sign / signified model such as Ferdinand De Saussure's cannot, for a two-part model lacks an explanation for the relational shift from signified to sign that a chain of thought necessarily involves. Having shown its worth in light of the possibilities in trichotomies, I will now describe the components of Peirce's model.

---

[95] Turino, Thomas. "Signs of Imagination, Identity, and Experience: A Peircean Semiotic Theory for Music." *Ethnomusicology* 43 no. 2 (Spring/Summer 1999), 223.

The sign is the basic name for something that stands for something else, whether that sign is the idea of the color blue, Ray Charles, or the arrangement of cards in a standard playing deck. Peirce sometimes uses the word *representamen* in place of “sign” for this meaning. Given that those secondary sources that do mention representamen mostly do so secondarily or parenthetically, I shall avoid the term here in lieu of “sign.” There are three levels of signs, *qualisign, sinsign,* and *legisign*, which relate to firstness, secondness, and thirdness, respectively. Proper definitions and examples of these sign types are contingent upon the remaining key vocabulary of Peirce’s model, and so my elaboration of these levels shall follow, after a few more terms are introduced.

The object is the thing for which a sign stands, and as with the sign, it can be anything from a duck to racism to the taste of cilantro. When we speak of object in Peircean terms, we speak not only of the thing represented, but its *relationship to the sign*. On the levels of firstness, secondness, and thirdness, the three respective classifications of objects and their relation to signs are *icon, index,* and *symbol*.

The interpretant is the effect of the sign on a perceiver; it is the way a sign is interpreted. It denotes a category of signification rather than the sort of thing being signified (which is the domain of the object). Within firstness, secondness, and thirdness, I call the three levels of the interpretant, respectively, *rheme, dicent*, and *argument*[96].

---

[96] It is in this category that both Peirce and his disciples are the least internally consistent. Turino offers that the three levels of interpretant on the gradations of firstness, secondness, and thirdness, are *rheme, dicent,* and *argument*, respectively. Though this is the terminology I shall use, I wish to imbed into it Johnson’s descriptors of the interpretant. Johnson takes two of Peirce’s less clearly positioned trichotomies — *immediate, dynamic,* and *final*; and *emotional, energetic,* and *logical* — and makes them parallel within the single category of the interpretant. In his flagship article on Peirce, Turino also mentions the emotional-energetic-logical complex, but only in an unconvincing and incidental way that ultimately comes across as tacked on, calling it the

With these three basic vertices of signification now in place, we must delve into the firstness, secondness, and thirdness of sign systems, for it is on these levels rather than the surface trichotomy I have just identified that the most important functions of meaning-making occur. In defining and exploring these levels, I shall also give application and more concrete meaning to the aforementioned trichotomies that are subsets of sign, object, and interpretant.

Firstness, associated with qualisigns, icons, and rhemes, is the realm of possibilities and abstractions. Peirce calls firstness "the mode of being of that which is such as it is, positively and without reference to anything else."[97] For example, colors, on their own and not in the context of actuality, are firsts, as are isolated phonemes, as is heat. Signs such as these are therefore qualisigns: qualities that are signs. The only objects signified in the realm of firstness are those to which signs bear a direct resemblance of mimicry; these are icons. Greenland is the icon of a map of Greenland; the sign of sampled birdsongs in "O Superman" is iconic to the actual birdsong. Blood would be an icon of the qualisign "redness." A rheme — the interpretant within firstness — is the act of recognizing a sign and in no way acting or intentionally pondering that sign, but instead

---

"dynamic interpretant," never acknowledging, much less exploring its three parts. Brendan Lalor argues that the immediate-dynamic-final set shows the interpretant in its more abstract structure, while the emotional-energetic-logical trichotomy relates it to the human experience of semiosis. Whatever the distinction, it cannot be known for certain, according to Edna Andrews, who notes on page 53 of *Markedness Theory*, "since Peirce himself does not evoke these two triads of interpretants together in his published works, a definitive statement is impossible to derive from him." Given the functional similarity of each of these trichotomies' levels in their respective pairing, I choose to use their similar definitions together and in tandem with the rheme, dicent, and argument components to deepen an understanding of firstness, secondness, and thirdness in the interpretant. The exact relation of the rheme-dicent-argument complex to these two trichotomies that Johnson explores is a very close one, but imperfectly defined.

[97] Peirce, Charles Sanders. *Collected Papers of Charles Sanders Peirce*, vol. 8. Edited by Arthur W. Burks. Dulles, VA: Thoemmes Press, 1998, 221.

experiencing the sign in an immediate and emotional way solely unto itself. The rheme does not explicitly relate the sign to the object, but merely perceives viscerally.

Secondness is the realm of actuality, simple referentiality, and action. It is where the first leaps in association and behavior occur within semiosis. The sign in secondness is the sinsign, which is an actual occurrence of something. A boat is a sinsign, as is a taco, as is the meow of a kitten. These things are not qualities, but full occurrences that we perceive in their entirety. A vast amount of the signs present in the world and in the Peircean chains we construct are sinsigns. An object that exhibits secondness is called an index. Signs represent indices "independently of any resemblance to them, only by virtue of real connections with them,"[98] and so danger would be an index of the sound of a gun firing (where the act of the gun firing would be icon). When one sees hotdogs cooking on a grill and wonders where the buns are, this is an indexical relationship, for frankfurters bear no inherent mimicry of buns, but are instead paired with them through experience. The dicent, which is the interpretant at the level of secondness, as I interpret it in this study, involves the cognition of a relationship between sign and object. Implied in this is the expenditure of effort in relating one to the other. This active relating of the two can manifest kinetically in dancing to music, whose index would be movement, or mentally in the physical attraction to a stranger, the index of whose face and body might be sex. This is also the process that governs the connection between voice and the body: a voice's sound connotes the feeling of our own bodies in the act of vocalizing.

Thirdness, the most complex of levels in Peirce's model, is largely ignored in many applications of the model to music. Focusing almost exclusively on firstness and secondness, Turino writes that

---

[98] Peirce, Charles Sanders. *The Essential Peirce*, vol. 2. Edited by Nathan Houser *et al*. Bloomington, IN: Indiana University Press, 1998, 461.

> listeners do not have symbols to rationalize or "domesticate" musical events at the level of Thirdness, that is, to understand them as a part of generalizable, predictable, mediated categories and processes…. Ultimately, it is the more ambiguous nature of indices and the fact that icons and indices are not organized in a distributional-grammatical system that allows them to be juxtaposed within sign complexes that compound the polysemy.[99]

Thirdness is the domain of mediation, of organization, and of lawmaking. A sign that exhibits thirdness is a legisign, which Raymond Monelle explains "is best understood as a type of class of which a sinsign is token. A word or sentence of language is a legisign, for it may occur in many different places yet is always the same."[100] A legisign, on its own, is an experientially or culturally encoded sign. As it signifies on its lower levels, a legisign like the written exclamation "AAAUUGH!!" conveys a relationship with experience, in this case secondness, and hence "AAAUUGH!!" is a legisign-index: through the arbitrary code system of writing, a generic human reaction to a number of stimuli suggests an index of shock, fear, or surprise, even though there is nothing imitatively shocking, fearful, or surprising in the sign "AAAUUGH!!" Most legisigns that we recognize as such in our daily lives are interpreted into the object of thirdness, the symbol. The symbol has a fixed meaning (unlike an exclamation) that is decided by experience, and it is the basic object type that allows meaningful communication to occur between people. A legisign-symbol requires mental thought and mediation through the given class of which Monelle speaks, which means that without symbols, legisign words cannot be processed as anything other than non-systematic utterances, and thus either foreign or nonverbal. Most legisigns that we recognize function symbolically, expressly by virtue of our recognizing them. Excepting onomatopoeia, elements of organizational systems such as language or musical notation are chiefly arbitrarily constructed with regard to their meaning[101], and thus there is nothing intrinsically chocolate about the word "chocolate," nor does Middle C have an epistemologically determined reason for its home on the first ledger line below the treble clef, and yet when we learn the system of words and meanings and of musical notation and its sound, then we can correctly relate the legisign words, sounds, or images to their meanings, collectively determined by the common experience of the community who uses these signs. The legisign is the word, sound, or image itself, whereas the symbol is the meaning of word, sound, or image. To a common eye, a printout of the many thousands of lines of programming code that make up the word processing software with which this document was written bears no intrinsic word processing-ness and would be nothing but a dangling and incompletely interpreted legisign, but when these data are processed by a specific sort of machine, we realize that the symbol is the key to decode that experientially or culturally encoded sign. It is the symbol that allows "Chocolate," Middle C, and Microsoft Word — all legisigns on their own — to mean anything at all to us other than a conglomeration of letter images and phonemic sounds. Finally, recalling the ways that we relate signs to objects, we address thirdness in the interpretant category. Interpretants exhibiting thirdness are arguments, and an argument is the construction of or adherence to a law or specific system in the process of creating meaning out of the legisign-symbol connection. This means, according to Thorkild Theleffsen, that

---

[99] Turino, 48-9.
[100] Monelle, 196.

---

[101] Steven Pinker, in chapter six of *The Language Instinct*, explores briefly the arbitrary nature of meanings and their word assignments, hinting at a possible relation between articulatory movement and meaning.

> something is being stated about the sign. An example of an Argument could be whole passages of text, i.e. meaningful links of Dicent Signs. I emphasize this interpretation and state that Arguments could very well be knowledge domains, cultures, societies.[102]

which means that mentally determined conclusions (embodied in the last levels of the immediate-dynamic-final and emotional-energetic-logical trichotomies) such as mathematical or philosophical proofs, or even the knowledge that the recipe and instructions for making brownies will, when followed, in fact produce brownies, are arguments. Furthermore, an acknowledgment of the system itself by which meaning is conveyed is an argument, so the conscious operation of language or explicit revelation of culture enacted are arguments as well. With argument comes the awareness of the processes in which one is invested.

We must understand that these three semiotic levels are cumulative; no sign system can exhibit secondness without containing firstness, nor can a sign system bear thirdness without both secondness and firstness. Beyond that, it is possible — and indeed very common — for sign, object, and interpretant to function not all on the same level at the same time with regard to firstness, secondness, and thirdness. Peirce makes clear that a sign system cannot step up in complexity along the chronology by which semiosis occurs, being sign, object, and interpretant.[103] The following diagram of the limit on combinations of components' complexities demonstrates the possible paths of Peircean semiotic systems:

---

[102] Theleffsen Thorkild. “Firstness and Thirdness Displacement: Epistemology of Peirce's Sign Trichotomies.” *Applied Semiotics / Sémiotique Appliquée* no.10 (2000), article 2.

[103] Many interpret this stipulation as an extension of the cumulative nature of complexity in sign systems, arguing by induction that sign, object, and interpretant are themselves aligned respectively as first, second, and third. This last idea is one that Turino and many others support, but it is not universally accepted on account of the multidimensionality

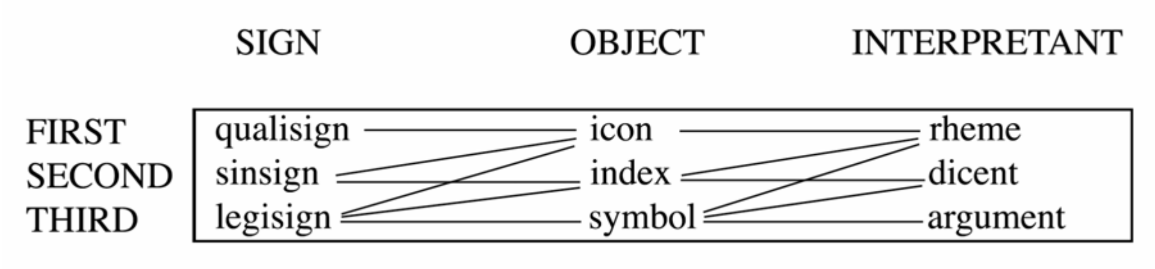


Figure 14: Pairing possibilities in a Peircean sign system

Furthermore, within sign, object and interpretant, levels of secondness contain the component levels of firstness, and thirdness, in turn, contains secondness and firstness, meaning for example that icon is part of index, and both sinsign and qualisign are part of legisign.

This set of rules produces ten possible sign systems, allowing for a very broad range of semiotic nuance. Note that only one possible sign system involves the argument. All other components, however, are relatively common among sign systems (and the qualisign, icon, and rheme are always part of every sign system). My point in going through all of this explanation is not to enable the categorization of every aspect of sound into a specific and rigid sign system, but instead to show the many varieties of sign systems that are nested in one another, and given the possibility of associative chaining whereby an object can become a sign, once interpreted, the extreme fluidity of form and complexity possible in a signification process.

Shifting focus now back to the issue of timbre, I wish to spotlight, in a Peircean view, the sign systems available to us when we hear vocal timbre as a sign. Peirce

---

foisted upon the first-second-third trichotomy by this interpretation, as well as the resultant nesting of all levels within all other levels when in fact Peirce only allows for nesting of lower levels in higher levels.

himself makes brief mention of this topic in his “Principles of Philosophy,” saying that timbre is “itself highly complex” but that it is “only known to us by extraneous experience” and not “heard in the sound.”[104] That is, we do not typically hear sound as naturally dissected into its constituent parts, but instead take it in as a total sign. This is framed in an argument for the firstness — the purely qualitative nature — of an “eternally sounding and unvarying” signal, which “should constitute the entire universe” and thus function without context. Turino speaks from a similar standpoint, but grants that this “extraneous experience” nevertheless is a component in the process of meaning-making, thus that not only are timbres firsts, but they “can and often do function as discrete icons, indices, rhemes, and dicent signs.”[105]

I propose that in the case of vocal timbre, however, we must recognize a very fundamental complication of this issue — one that is entirely central to the musical semiotics of timbre in the human voice. On the purest structural level, we might wish to look for meaning solely in the acoustic content of a voice, first at a given moment and then in the context of moments’ positivistic relationship to one another. Wayne Slawson, in summarizing Pierre Schaeffer's notion of “reduced hearing” describes such purely acoustic hearing: “The common mode of listening, in which we respond to a sound by identifying its source... must be distinguished, according to Schaeffer, from another mode, in which we purposely... divorce what we hear from its source, concentrating instead on the properties of the sound itself. This kind of objectification or 'reduction' of sound is required for a sonic event to be heard as a ‘sound object.’”[106] However, in practice, such a hearing of the voice is, I believe, effectively impossible. In first identifying a series of audible signs as being part of a common system, in then recognizing that this system has a single common source, and finally in identifying that source, a listener has already stepped beyond the allowable limits of a Schaeffer’s “reduced hearing”, or what in an evolutionary context David Burrows calls “protosemiosis.”[107] Combined with our extreme reliability in distinguishing a human voice from other sound (or as Cornelia Fales says, “ordinary listeners with no special training possess an extraordinary amount of knowledge about sound and its sources.”[108]), this chain of contextualizing sound signs will in all but the most extreme examples lead us to an interpretation of the sounds in question that is mediated through our experiential impressions of how the voice communicates and what it is capable of communicating. It is for these reasons that in the previous two chapters I have framed my analyses of vocal timbre in terms of its context within a voice act’s formant content and within the human experience of voice, respectively.

---

[104] Peirce, Charles Sanders. *Collected Papers of Charles Sanders Peirce,* vol. 1. Edited by Charles Hartshorne and Paul Weiss. Cambridge, MA: The Belknap Press of Harvard University Press, 1960, 151.
[105] Turino, 236-7.

Complicating, if not totally derailing an entirely isolated and quantifiable acoustic approach is the inextricability of voice and language from one another. Language, the most explicit and concrete ground on which we can speak of signs and their meanings, is the chief domain of the voice, and voice, in turn, is the conduit through which language is most chiefly enacted. In music, language is theoretical without voice and as Burrows puts it, "singing without words we feel impotent and incomplete, as though deprived of the use of our hands."[109]

---

[106] Slawson, 6.
[107] Burrows, David. *Sound, Speech, and Music*. Amherst, MA: The University of Massachusetts Press, 1990, 71-2.
[108] Fales, 59.

And yet despite the entwined relation between voice and language, virtually no concepts or terms translate directly between the two. In linguistics, when one speaks of specific formant parameters, one does so with reference to phonemes. While phonemes, unlike morphemes or words, are singularly describable because they are minimal units of speech that do not exhibit change in sound, the phenomena they describe are not interchangeable with the idea of timbre. Though they are importantly connected, their difference is crucial.

Phonemes are not particular occurrences of sound, but instead they are discrete classifications. Jay Lemke writes

> The articulatory apparatus of human speech production is capable of producing a continuous spectrum of vowels… and many continuous dimensions of difference in other respects, but the linguistically significant distinctions are categorical: a word belongs to our language only if it is assimilated to one of a finite number of discrete possible sequences of phonemes; a harmonically coloured, shouted, or choked vowel does not a new word make. Language reduces the continuum of the acoustic spectrum to a finite set of discrete contrasts.[110]

A given phoneme — [u], for example — is the representation of all pronunciations of [u] that differ in a linguistically meaningful way from other sounds. A phoneme remains the same whether James Earl Jones or Tiny Tim speaks it, though the timbres of two speakers uttering the phoneme may differ wildly. It is this abstraction of possible speech sounds that allows Roman Jakobson to weigh in on their semiotic value:

---

[109] Burrows, 85.

[110] Lemke, Jay. "Topological Semiosis and the Evolution of Meaning." Presented at the Conference of the Washington Evolutionary Systems Society, May 1998, Georgetown University.

> The linguistic value… of any phoneme in any language whatever is *only its power to distinguish the word containing this phoneme from any words which, similar in all other respects, contain some other phoneme*… What corresponds to the difference between two phonemes is solely the *fact* of a difference in meaning, whereas the *content* of these different meanings varies from one word to another… Only the phoneme is a purely differential and contentless sign. The phoneme's sole linguistic content, or more generally its sole semiotic content, is its dissimilarity from all other phonemes of the given system. A phoneme signifies something different from another phoneme in the same position; this is its sole value.[111]

From a linguistic standpoint, a phoneme is a qualisign, and its enactment — the speech that brings it into secondness — must be noted as *the speaking of a phoneme* and not as the phoneme itself. A phoneme therefore always retains its total firstness.

In the second chapter of this dissertation, I demonstrated that vocal timbre's meaning lies in what it lacks, spectrally speaking, with regard to the semiotic system in which it occurs. In showing that its meaning lies in *how* it differs from other vocal timbres, I underscore the difference between vocal timbre and phonemes, whose meanings lie in *that* they differ from other phonemes. While the fact of phonemic difference does not contain meaning, save for itself, the means of timbral difference does.

From a timbral viewpoint, and more specifically, from the information that a filter subtractively leaves in a musical human voice, I propose that, in order of Peircean complexity and in order of occurrence, a listener will predictably make four crucial judgments.. The first is the mere recognition of the sound, where we acknowledge the simple signal in the most basic sensory way; this is a qualisign-icon-rheme. The second is in identifying that sound as a human voice. Assuming a normal familiarity with the voice and a normal ability to recognize it, a listener will process this as a sinsign-icon rheme, thus branching into the realm of secondness provided by one's understanding of a specific occurrence within experience. The third judgment is the determination of whose voice one hears. From timbre, regardless of linguistic content, one inadvertently attempts to determine, to varying degrees, the sex, age, and size of the person whose voice is one hears. It would be a most inhuman listening to hear a speech or an aria and to have no interest in or acknowledgement of who — in the broadest terms — delivers it. This step is very closely related to the previous chapter's focus on the body, for the sound of the voice indicates the kind of body with which we align our own. This determining process is a sinsign-index-dicent, bound wholly in secondness because it actively relates an actual event abstractly and experientially to a real thing. Whether one succeeds in identifying whose voice one hears does not affect the reality of this third judgment being enacted. The final judgment on the listener's part is the determination of the linguistic content of the voice. Stated otherwise, the listener asks, "What is the voice saying?" In doing so, one first listens for words in order to determine whether the voice speaks a familiar language. If this is not the case, then the effort becomes not one of recognition, per se, but of eking out whether the vocal sound belongs to a linguistic system or whether it is in fact nonsense. And even within nonsense, a skilled scat singer, for example, will perform a variety of timbres and sounds over the course of a solo to inject variety — difference — into an otherwise wordless solo for an audience to whom such timbral variation is a feature to be closely and comprehendingly followed. That the voice and language are so connected, as I previously explained, makes sense when placed alongside the fact that most vocal music contains language, and therefore most acts of this fourth judgment in hearing the voice will result in recognizing that a semantic hearing is possible, and in the case of a listener's familiarity with the voice's language, engaging in this semantic hearing to make sense of the text. Nonlinguistic voice acts, as found in music like scat or in pieces by Diamanda Galás and Claude Debussy, are therefore marked by our hearing because of their distinct lack of language, and ultimately they are still subject to the fourth judgment, even if the situational answer to "What is the voice saying?" is "Nothing." The variety of nuances in the fourth judgment cannot all fit into a single Peircean triad, for once a semantic hearing begins, its semiosis is almost entirely textually dependent. The one clear statement that I can confidently make is that this judgment is one of thirdness. The relation between the voice and words is mediated through the structure of language, and with each structure of linguistic meaning one perceives, another is revealed. Phonological semiosis begets semantic semiosis, which in turn begets syntactical semiosis, and in parallel, the possibility of language — its inherent cumulative firstness, a legisign rheme — can, in linguistically recognizable music, become the actuality of language— its inherent cumulative secondness, a legisign dicent — which, in some texts can even beget the thirdness of thirdness — language's mediated negotiation, or in Peirce's terms, an argument. These four judgments are the predictable, even necessary steps in any active listening to a single voice.[112]

---

[111] Jakobson, Roman. *On Language*. Edited by Linda R. Waugh and Monique Monville-Burston. Cambridge, MA: Harvard University Press, 1990, 230.

Worthy of their own category are the thoughts that are not predictable in a larger sense, but that are dependent on the listener, the specific vocal timbres, the context in which they are heard, and if applicable, the specific words they form. These are individually and culturally dependent responses, and are therefore not predictable as a general rule. A broadly–scoped study such as this cannot account for the implications of "purity" suggested by a specific timbral imprint in a choirboy's singing of "O Holy Night," the mental relating of two voices with similar characteristics, such as those of Louis Armstrong and Tom Waits, or for the annoyance that one might feel at what is timbrally marked as a child's voice singing a Toys 'R' Us jingle.

---

[112] While it is possible to "tune someone out," doing so is frequently an active choice made in relation to — and typically only affecting— the fourth judgment.

For the predictable set of judgments, we can define the system in which they occur — one of sound (whether embodied or not) and a listener, who can in all but the most extreme cases have an assumed familiarity with the voice as a sound. Other judgments cannot be bound by a knowable system. Built into this opposition of predictable and unpredictable judgments is their respective initial alignment with Roman Jakobson's notions of introversive and extroversive semiosis, as explained by both Agawu and Nattiez.

Since I am concerned with dissecting the predictable listening, it must be broken into its smaller units — its timbral moments. Though Cogan notes that "in musical contexts there can exist no single, unchanging, meaningful instant of time that can be unambiguously isolated and analyzed,"[113] if we define a timbral moment in music as one whose component frequencies' periodicities are unchanging, particularly with relation to one another, then we start to understand the levels at which a sign's reading branches outside of that sign. A timbral moment relates outside of itself to all other timbres in the piece, and yet when the piece is taken as a whole sign, this relation of minimal units to one another can be viewed as an internal set, just as a song is relative to a cycle, and an opus is relative to a corpus. It is in studying the relationship between vocal timbres and the levels of structure surrounding them that the theoretical portion of this dissertation now enters its final campaign.

---

[113] Cogan 1984, 142.

Concerning the notion of individual moments, the field to which they must first be compared is the set of other timbral moments within the voice act or recording (for timbre relies on real enactment to exist). If music is imagined in the dimensions of time and in the dimensions of frequencies present therein — just as a spectrograph shows — then its edges, in the most literal sense, define it. The totality of a voice act then becomes the system by which its actual (rather than theoretical) minimal units are to be evaluated.[114] Because of the explicitly delimited acoustic content of a voice act, each timbre in the piece is a knowable and interpretable unit whose capacity to differ from other sounds and to defer meaning to them is bound entirely by the act. In this way, even without words or scales, a structuralist reading of sound units within a piece reveals that as the voice act dictates its own content, and consequently its systemic rubric, the constituent parts dynamically plot themselves through real time and retroactively in relation to one another and to the whole, constructing their idiolect even as they speak it; music becomes its own linguistic system of thirdness. I want to be careful not to call music a language in the traditional sense of the word, but its vocal timbres do nevertheless constitute a lexicon, and even if syntactic or semantic statements cannot be derived from this performative idiolect, there is still at the most basic level of repetition and difference a lexical behavior. This idea does not seem so revolutionary when viewed as a more fleshed out paraphrasing of Edward Cone's belief that any piece of music will reveal its own means of analysis. This is not to equate listening with analysis; listening can embody most any sign system, while analysis seeks out Peirce's argument, but in this respect, because argument cumulatively contains all other sign systems, analysis becomes the ultimate listening, and implicitly, the goal of hearing. Not wishing to contradict my earlier use of embodiment and sensual anchorage in perceiving the musical voice, I stipulate that analysis, in the wake of the mind and body duality's dissolution, can approach the same argumentative landmarks without resorting, in Cusick's words, to a body-denying "traditional music theory [that] consists more of answers — descriptions of practices which are understood to be objective and true."[115] Instead, a fluid approach "more of questions, or of hypotheses around which to frame questions"[116] can still guide us to these perceptual edges of music, for it is the wondering at the sensation of secondness — embodiment — that both allows and is included in the progression toward thirdness.

---

[114] This is the level of discretization that Nattiez describes in *Music and Discourse,* 80-1.

That a piece of music is bound in such a frame as I have described — another idea not dissimilar to Cone's theories — means that this boundary is a tangible line over which the internal and the external struggle. It makes sense then that a musical act's beginning and end are the moments at which we are most aware of this pressure between the work and the outside world. This is the pressure that Peirce speaks of with regard to thirdness, and is also, incidentally, a compelling reason that Laurie Anderson's "ha" in "O Superman" begins and concludes the piece in near total humanity, but takes on the feel of a mechanical component for the greater part of the work's middle body.

---

[115] Cusick, 46.
[116] ibid

The outside world with which a voice act exerts mutual pressure encompasses the act's immediate actual context, our assumptions about the voice, about performances, about music, as well as about sound, and it also ultimately encompasses everything that is not the act of music. Here I cite Cogan, who cites Roman Jakobson and Linda Waugh, who cite Peirce, saying "A thing without oppositions *ipso facto* does not exist."[117] It goes without saying that Derrida would agree with this.

Given that human language, in its traditional sense, is part of this outside world, then its involvement in any voice act is the invasion of one structure into another. But the contamination is mutual: the meaning of an aria, a chorale, or a rap song is by no means solely determined by its linguistic text, but nor can the language in these musics mean what it otherwise would on its own. More generally, "The center and the periphery exert influence on one another and as a result, the structure of the centre takes on characteristics of the periphery and vice versa."[118] The linguistic presences that timbres cast on human speech and singing are, in Barthes's words then, the "voices within the voice"[119] — the center within the periphery and the periphery within the center. The metaphorical *parole* of words is paired with a foreign *langue* of a voice act, and the sounds of the voice are the common translating feature to this disparate pairing.

In Derridean terms, the meaning of vocal timbre in music lies in the *trace* of that in opposition to which it stands. Recalling from the second chapter the location of meaning in vocal timbre as between the formants where the source is sculpted by the performer's physiological filter, then the *trace* intersects this same space of meaning.

---

[117] Cogan 1984, 125-6.
[118] Johnson, 269-70.
[119] Barthes, 184.

Between formants we encounter silence, and so in the confluence of meaning, a voice act's outside world must intersect with and become this silence. The externality of a vocal performance is not based in pitch or periodicity; rather, the relation of timbre to meaning is then one of grayness — a surface with no foreground. When timbres change and thereby dynamically change their relations to one another, this is still not a creating a semiotic foreground within a voice act, but instead as it presents new structural points in a voice's music-making, adds new elements to the timbral idiolect, and introduces new filtered gaps in the spectrum, it thereby suggests, if not presents, an all-encompassing aggregate of backgrounds.[120] Stretched onto a frame of external experience, vocal timbres are canvases upon which language, pitch, and rhythm — the contents borne of timbre's depth— are painted in shades of presence and absence.

Recalling the role of timbre in anchoring our experience of will and perception, the body then, finally, both centralizes and is central to all this. In making perception active, the voice incites the body to enact silently the timbres it hears, and so as we ponder the surface of the voice's timbre, we internalize it. We allow the meaning of words and of music to take form on this canvas within, and to color our being.

---

[120] To be clear, these surfaces that timbres lay out are not "grounds" in the Peircean sense.

II. STATIC TAKES LOVE'S BODY

Having explored vocal timbre from a broad theoretical viewpoint, I wish to apply my creatively its capacity to create meaning. In my piano concerto "static takes love's body," by juxtaposing oppositions of human / instrumental sound, ensemble / solo sound, pitched / unpitched sound, and language / nonsense, and by placing a choir in the role traditionally reserved for the orchestra — the accompaniment to a solo instrument's concerto — I bring to light issues very expressly related to the meaning of timbre in the human voice. The piece's literary quote is from William Gibson's short story "Fragments of a Hologram Rose" and is used with the kind permission of the author and his publisher. Herein I present the piece's score.

Performance Notes

Accidentals are held through the measure.

It is recommended that the choir have a pitch pipe planted in each section to facilitate quietly the preparation for its entrances at measures 54, 142, and wherever else the conductor deems it necessary.

Pianist must have a plectrum with which to scrape a single low string lengthwise at rehearsal letter **G** (m. 172) and beyond.

In the vocalized and physicalized percussion section beginning at rehearsal letter **H** (m. 215), “dvh” is articulated by clenching teeth, shaping mouth into an “o”, and then forcefully from the diaphragm making an affricate “dv” sound. It should vaguely resemble the sound of a bass drum. The “gzh” sound is articulated by clenching teeth, shaping mouth into an “ih”, and then forcefully from the diaphragm making an affricate “gzh” sound (where “zh” is the sound in “luge”). It should vaguely resemble the sound of a drum machine’s tom-tom.

For bodily percussion, hollow square noteheads are stomps, hollow triangle noteheads are claps, and black upside-down triangle noteheads are thigh slaps (with two hands).

Stage setup for performance features the piano and choir grouped separately on opposite ends of the stage, both gently angled inward, as shown in the following figure.

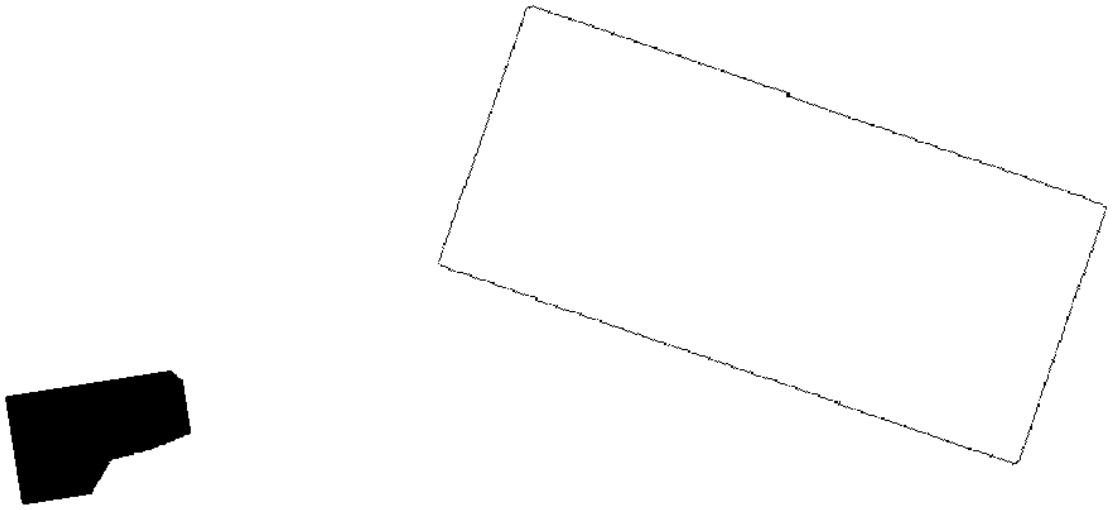


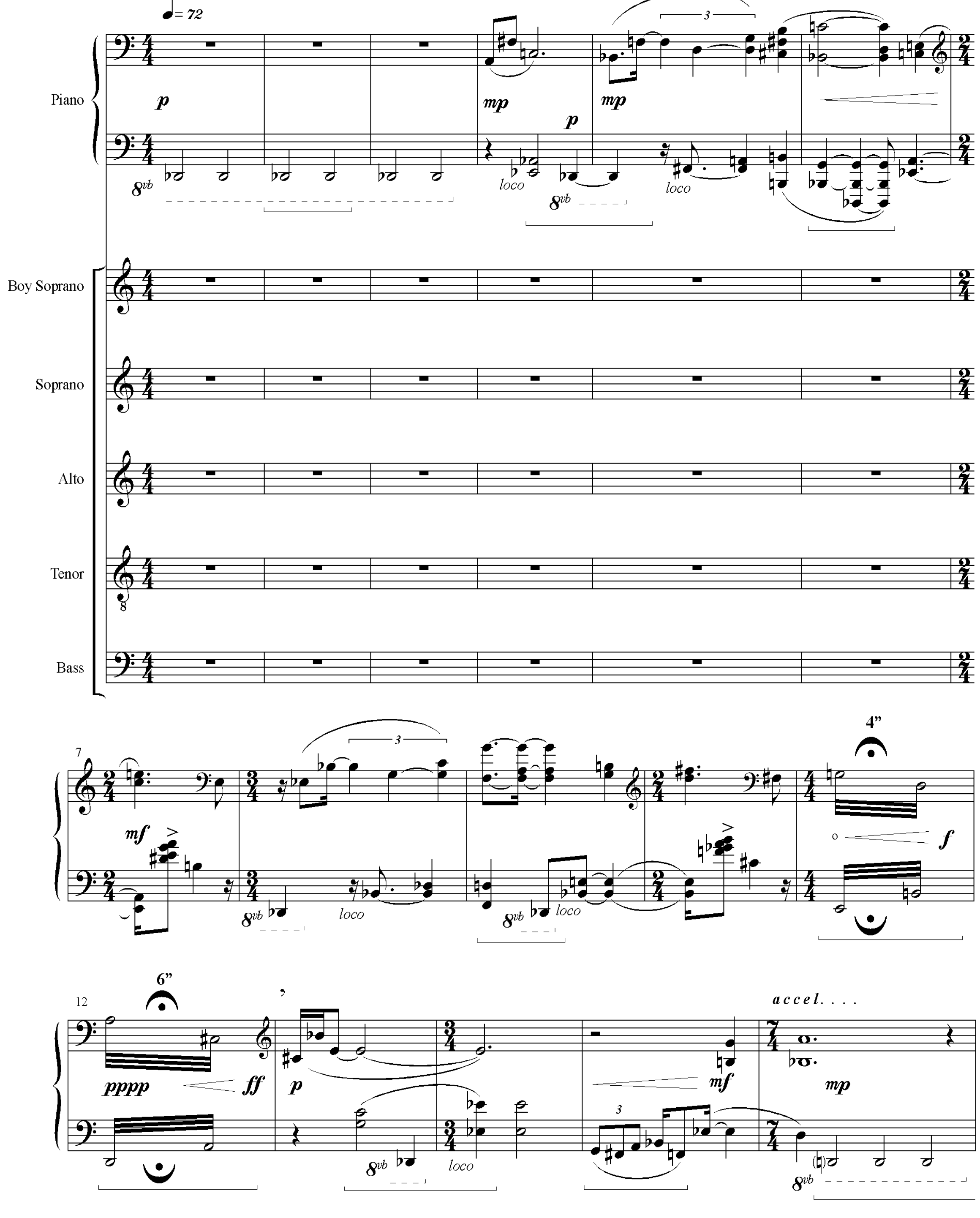
Figure 15: Stage setup for "static takes love's body"

The piece's approximate duration is nineteen minutes.

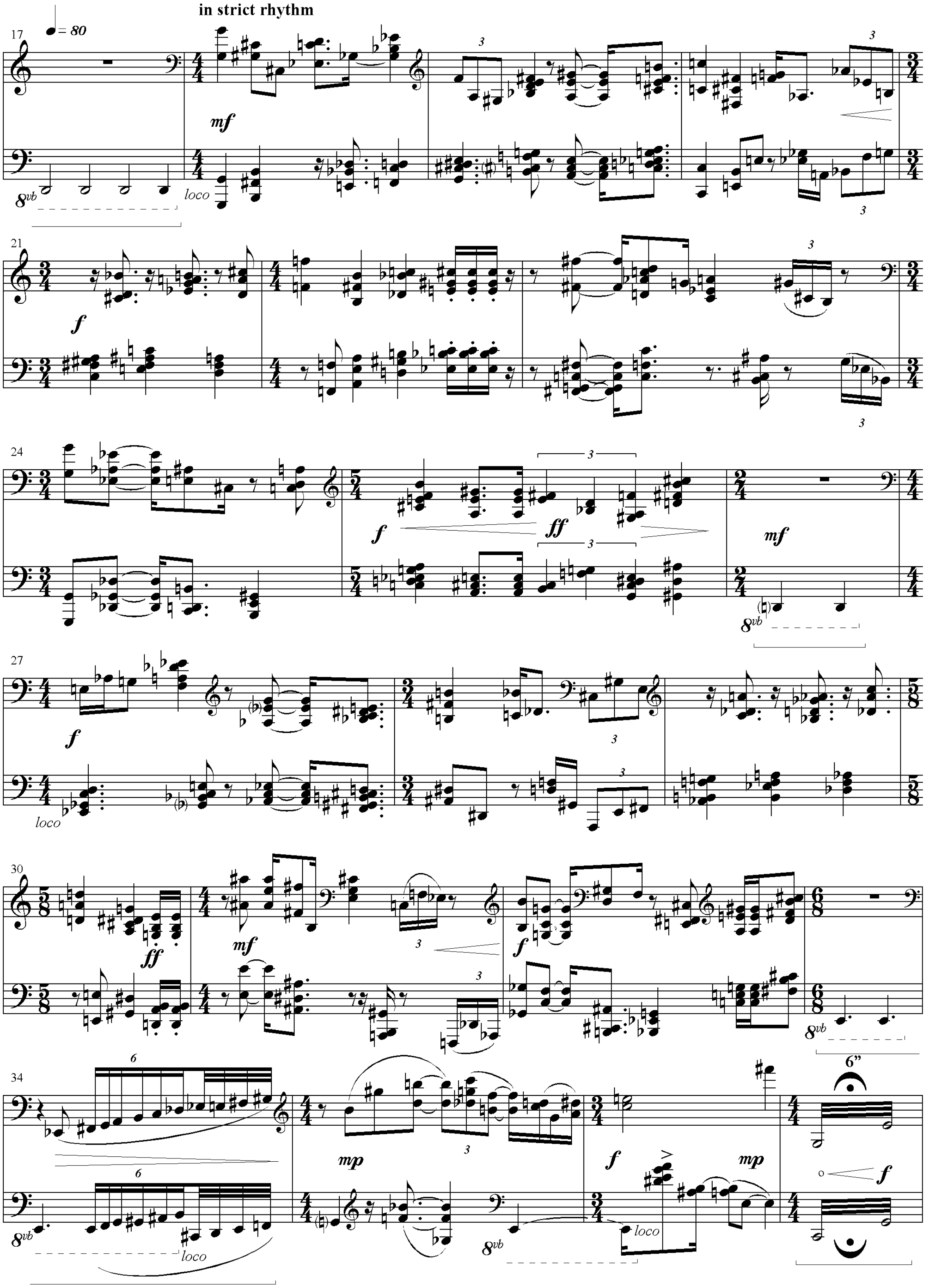
static takes love's body

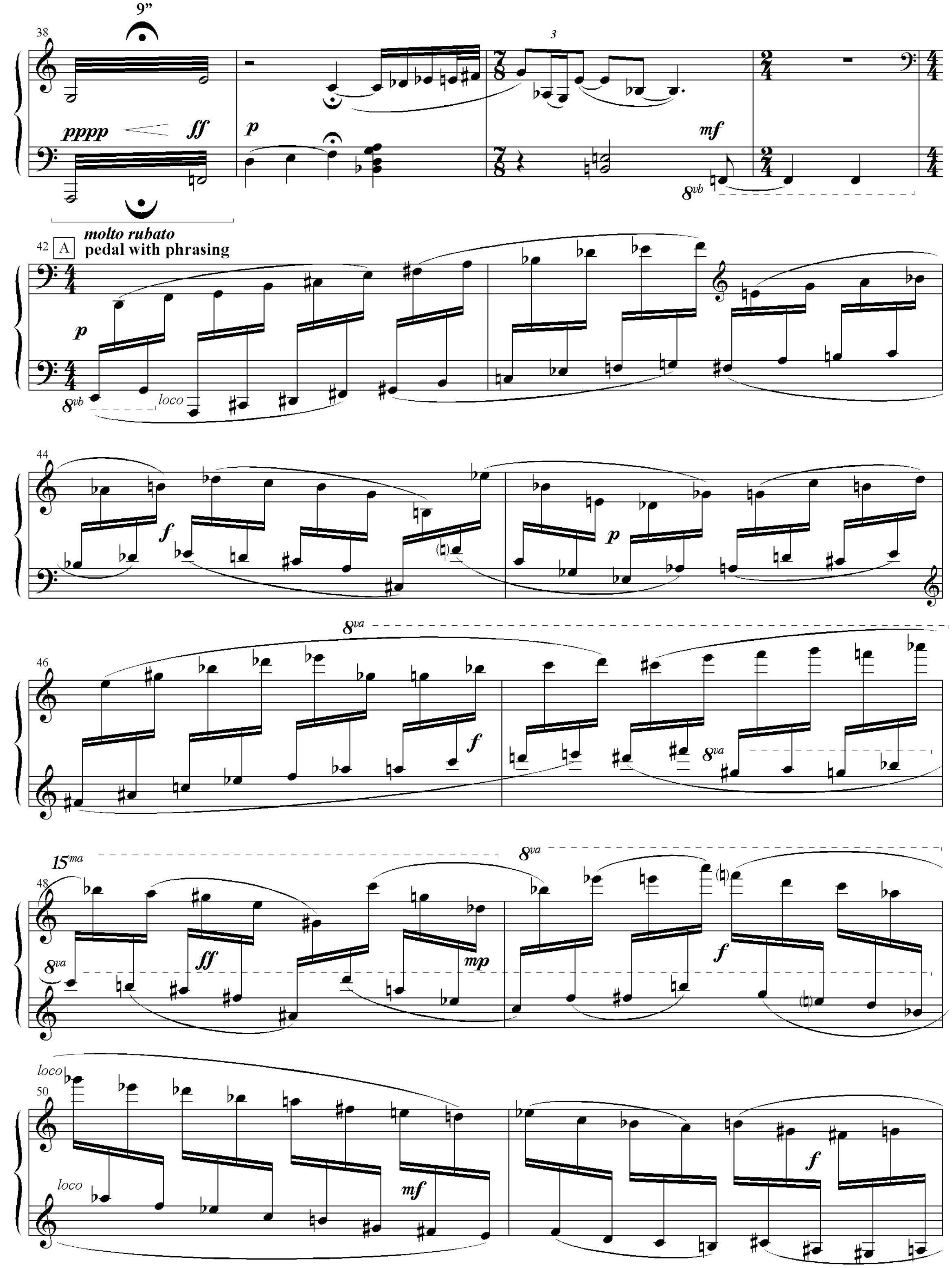
concerto for piano with large choir

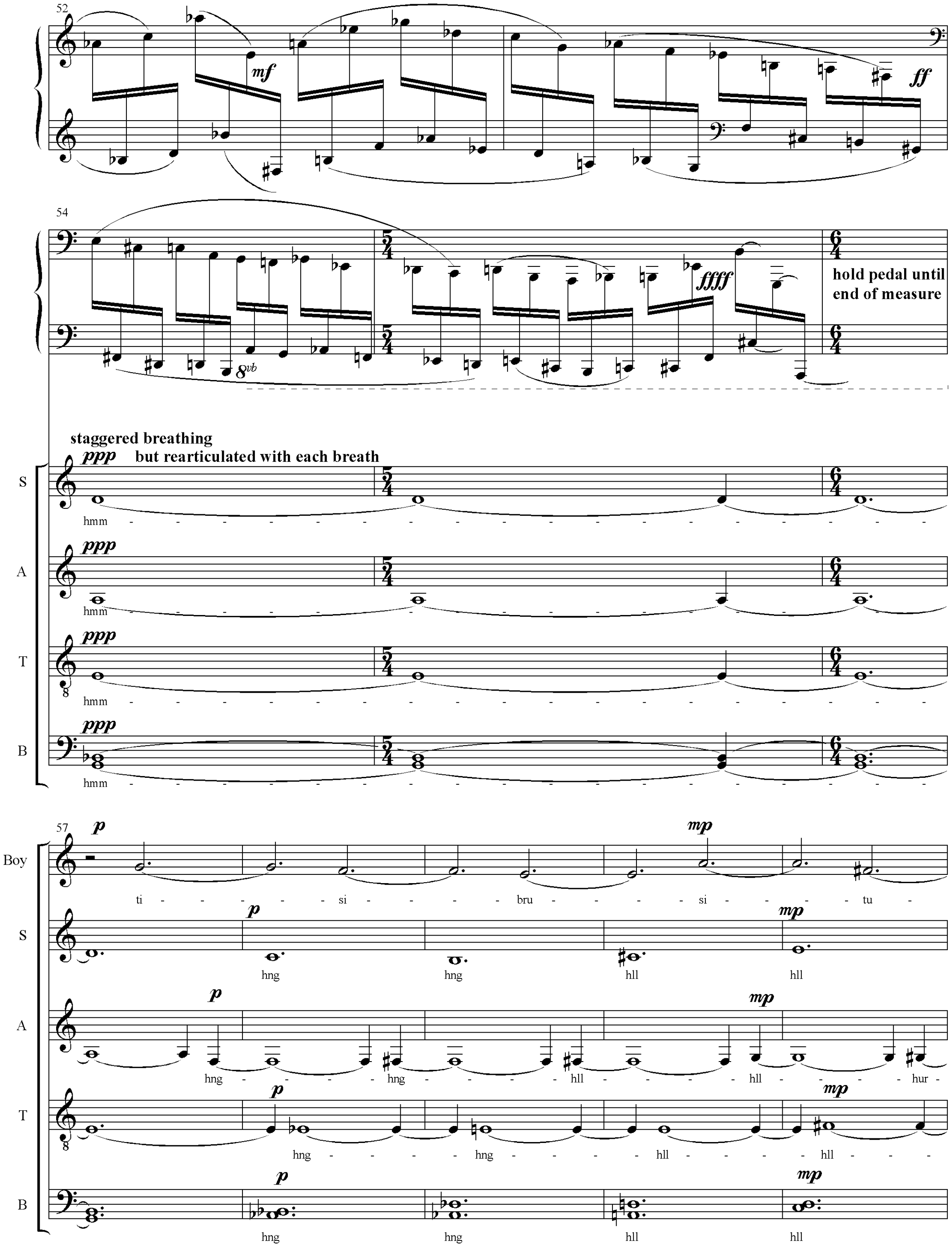
S. Alexander Reed

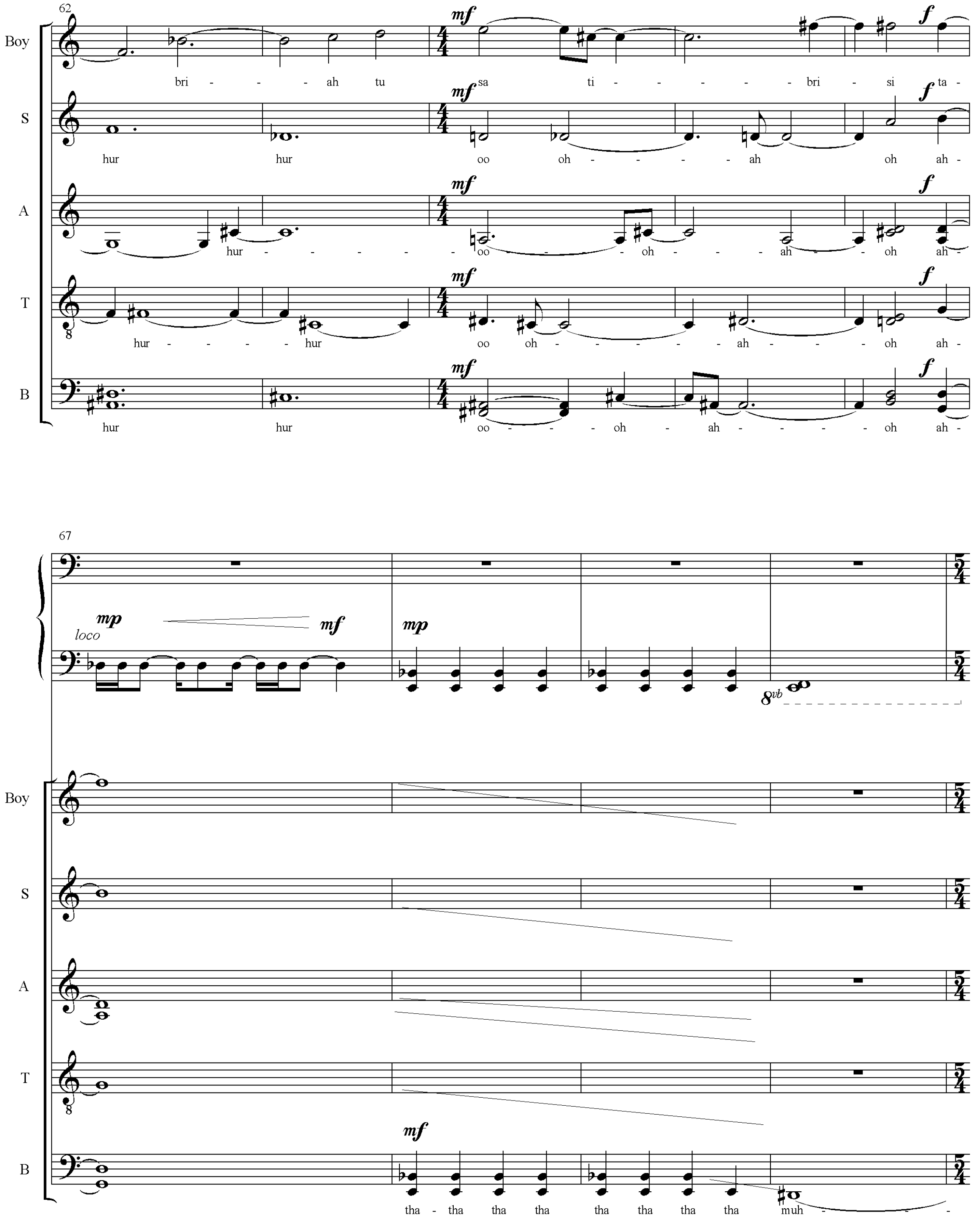


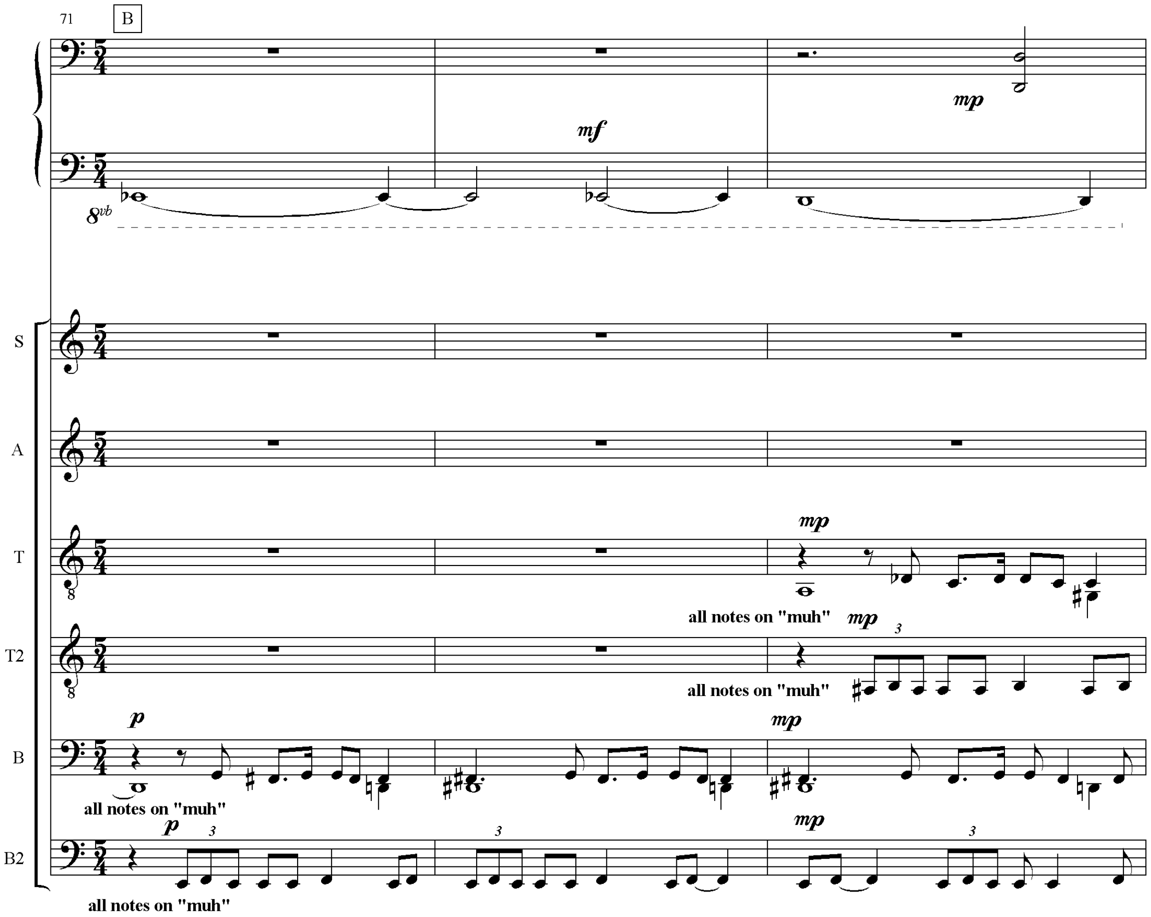
♩= 72
Piano
Boy Soprano
Soprano
Alto
Tenor
Bass
p
mp
p
mp
mf
f
pppp
ff
mp
8vb
loco
3
4"
6"
7
12
accel. . . .

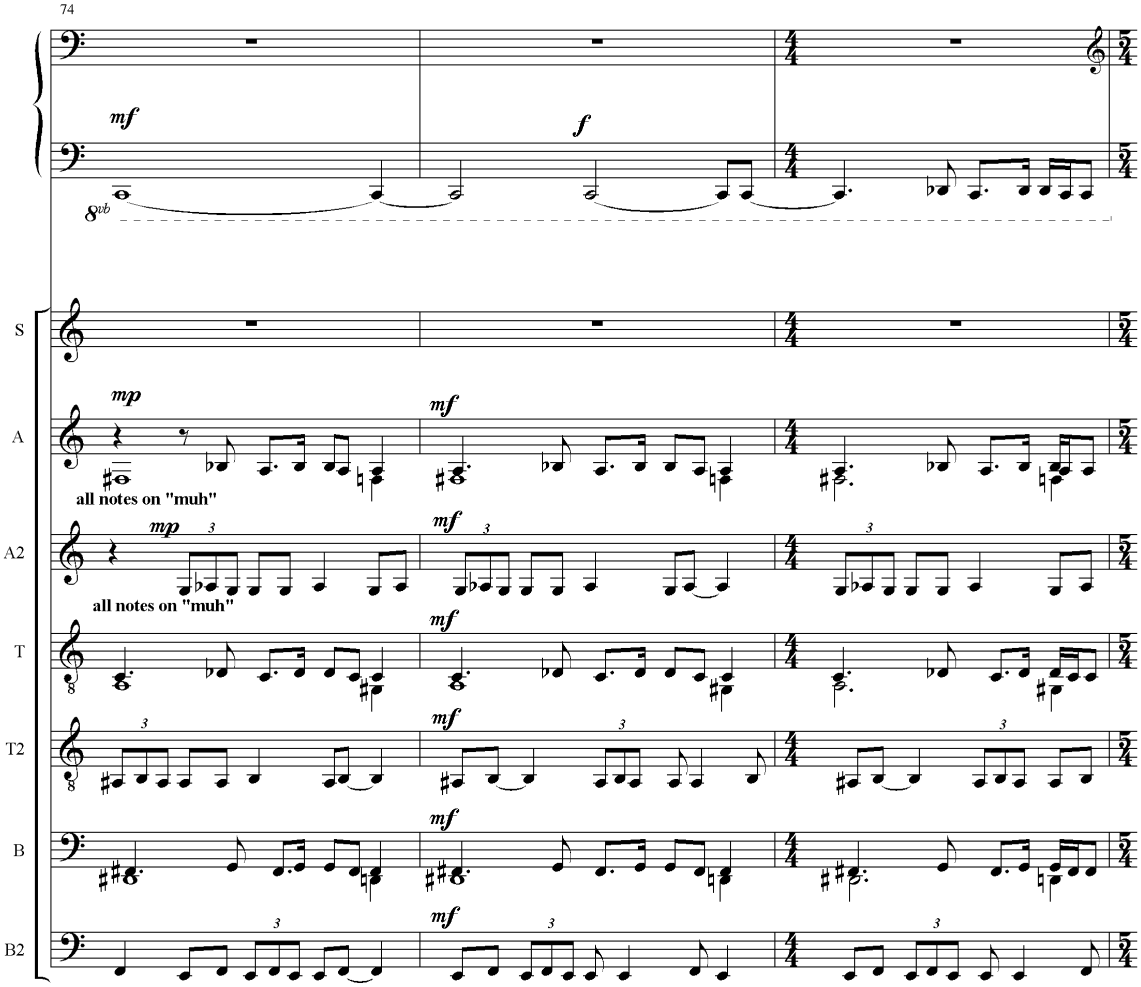

in strict rhythm
♩= 80
17
mf
8vb
loco
21
f
24
f
ff
mf
8vb
27
f
loco
30
ff
mf
f
8vb
34
6
mp
8vb
loco
f
mp
6"
f

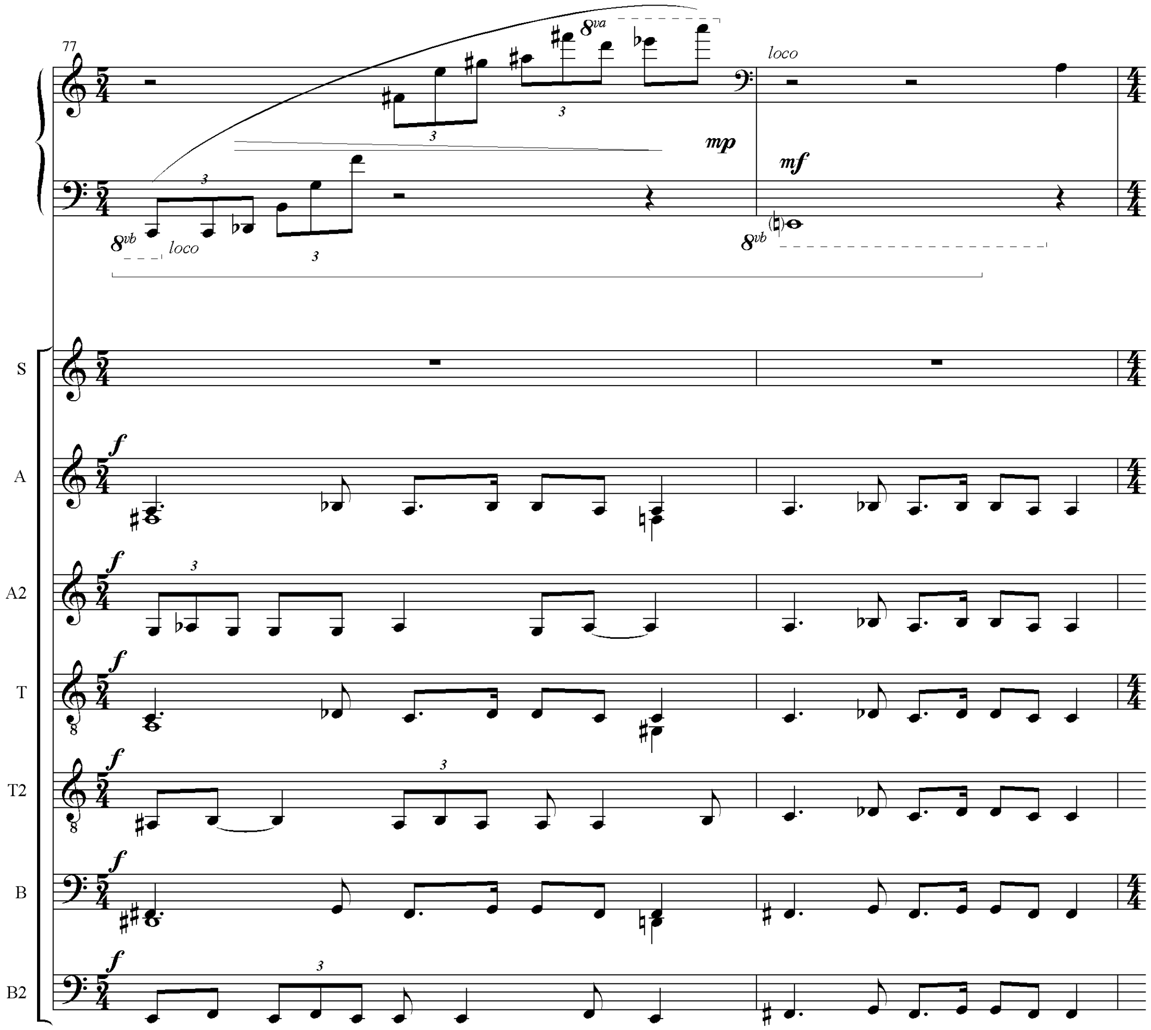

38
9”
pppp
ff
p
3
mf
8vb
42
A
molto rubato
pedal with phrasing
p
8vb
loco
44
f
p
46
8va
f
8va
15ma
48
8va
ff
mp
8va
f
loco
50
loco
mf
f

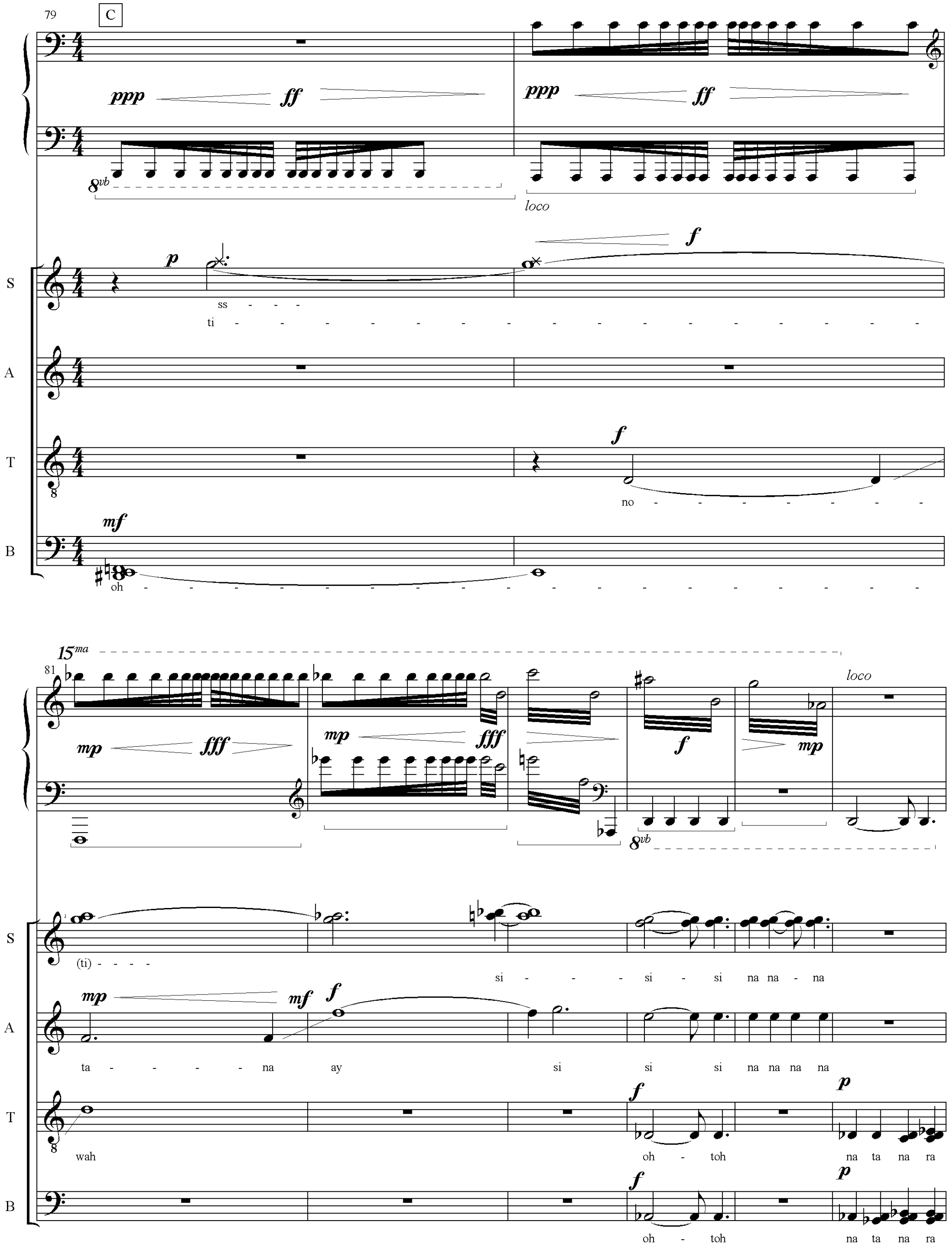

52
mf
ff
54
ffff
8vb
hold pedal until end of measure
staggered breathing
ppp but rearticulated with each breath
S
A
T
B
hmm
57
Boy
p
mp
ti - si - bru - si - tu -
hng
hng
hll
hll
hng - hng - hll - hll - hur -
hng - hng - hll - hll -

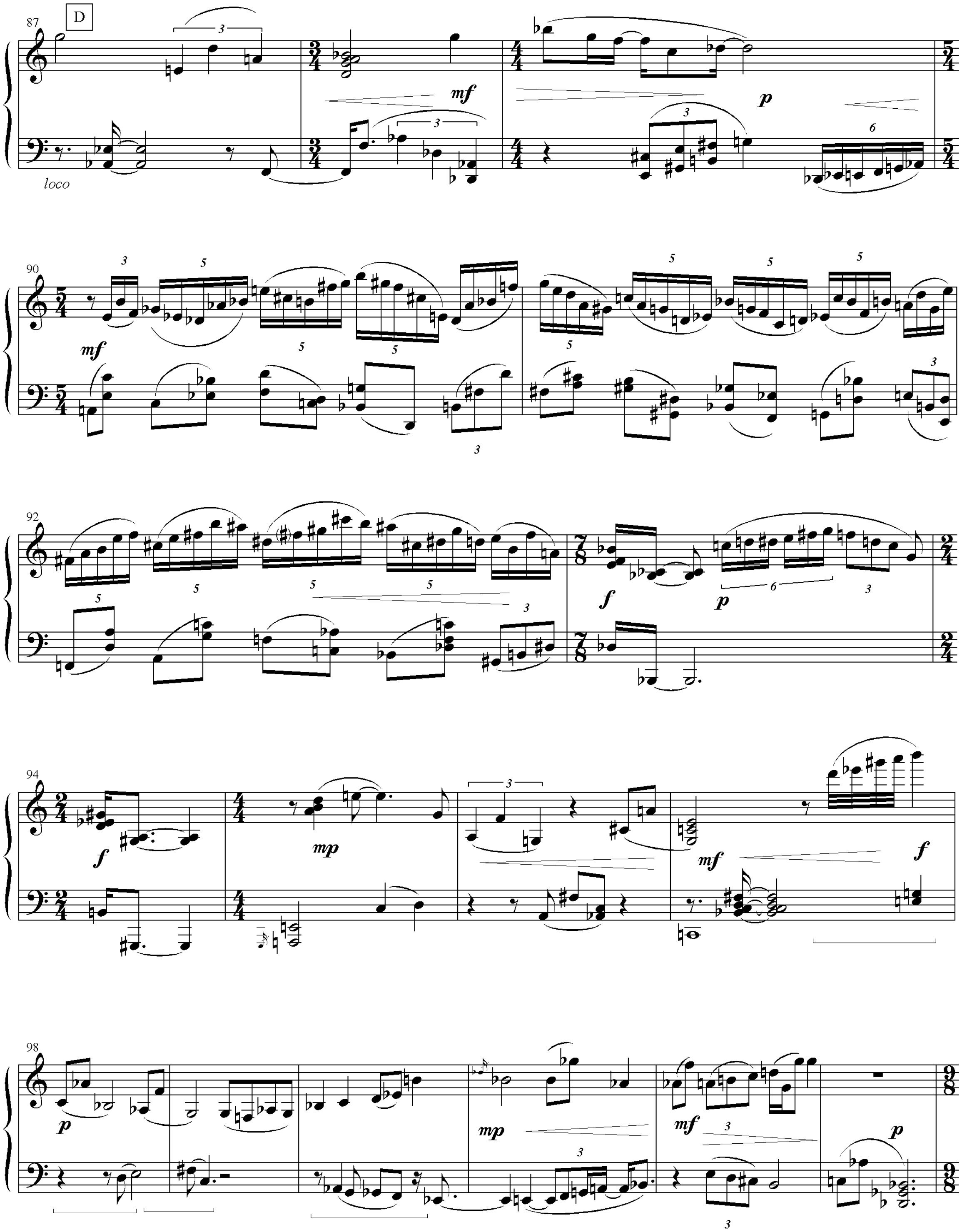

62
Boy
S
A
T
B
mf
f
bri - - - - ah tu sa ti - - - - - - bri - si ta -
hur hur oo oh - - - - - - ah oh ah -
hur - - - - - - oo - - - oh - - - - ah - - oh ah -
hur - - - - - hur oo oh - - - - - - - ah - - - - - oh ah -
hur hur oo - - - oh - ah - - - - - - oh ah -
67
loco
mp
mf
8vb
tha - tha tha tha tha tha tha tha muh - - -

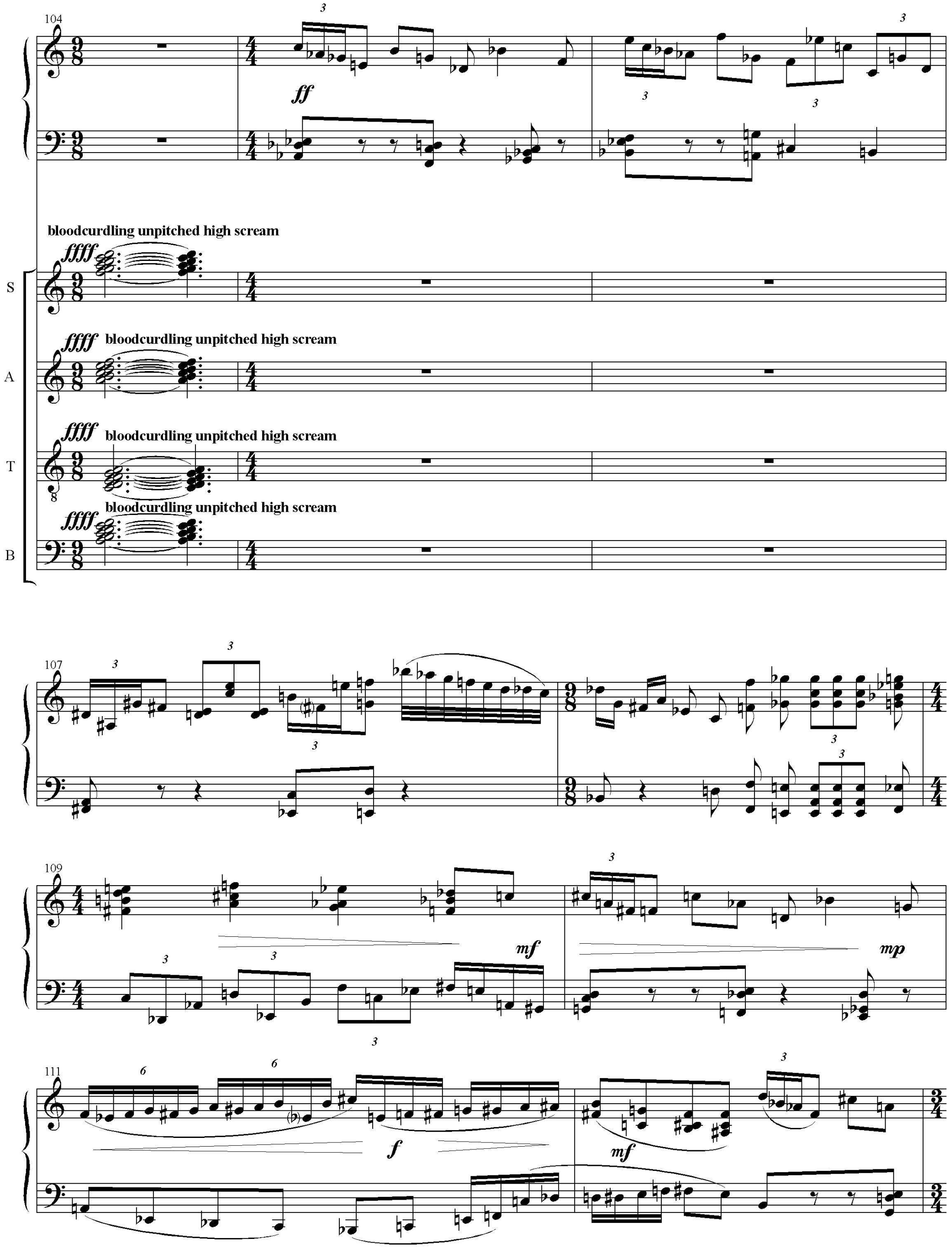

71
B
mf
mp
8vb
S
A
T
T2
B
B2
mp
all notes on "muh"
mp
3
all notes on "muh"
p
mp
all notes on "muh"
p
3
3
mp
3
all notes on "muh"

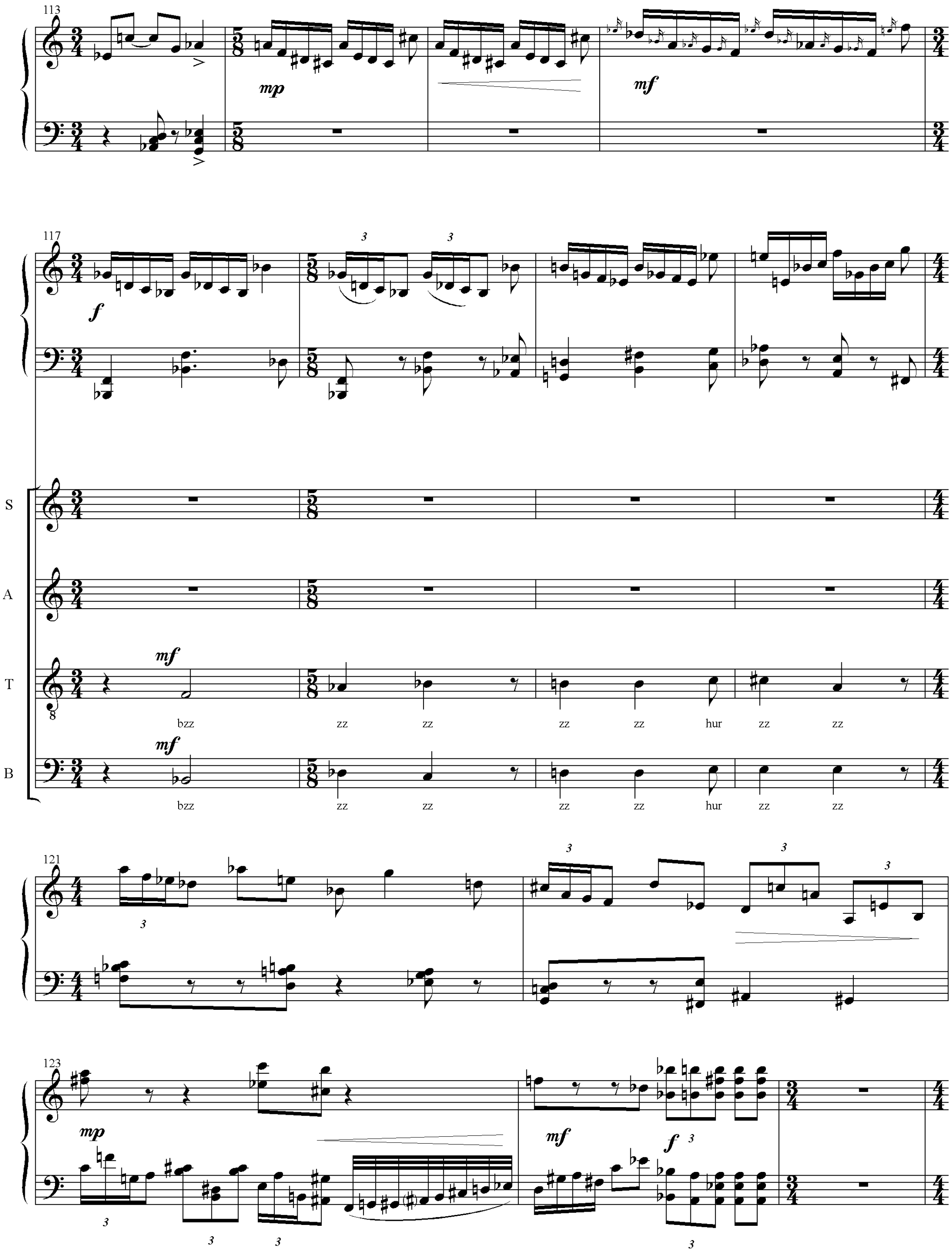

74
mf
f
8vb
S
A
mp
mf
all notes on "muh"
A2
mp
3
mf
all notes on "muh"
T
mf
T2
3
mf
B
mf
B2
mf
3

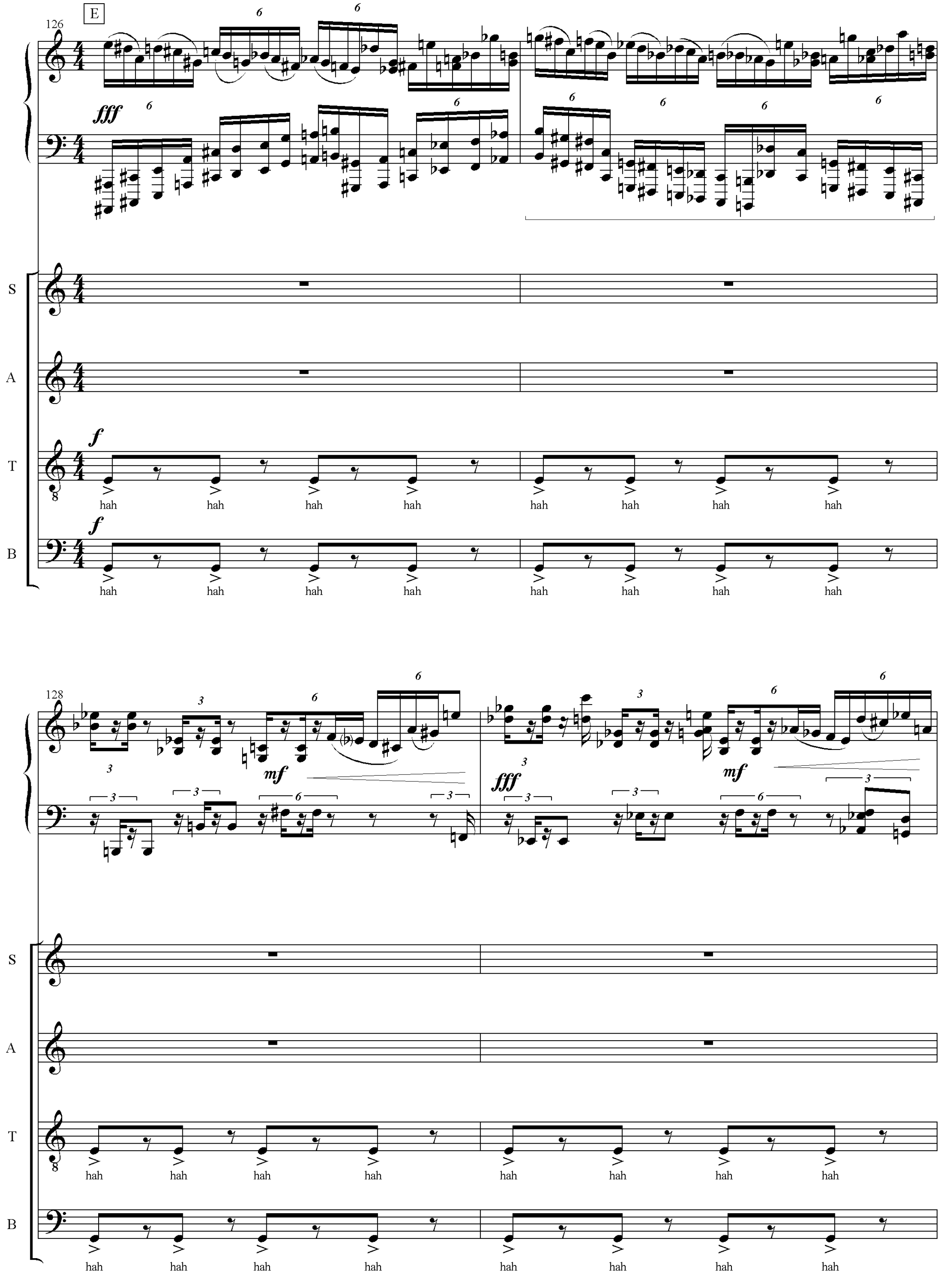

77
8va
loco
mp
mf
8vb
loco
8vb
S
A
A2
T
T2
B
B2
f
3

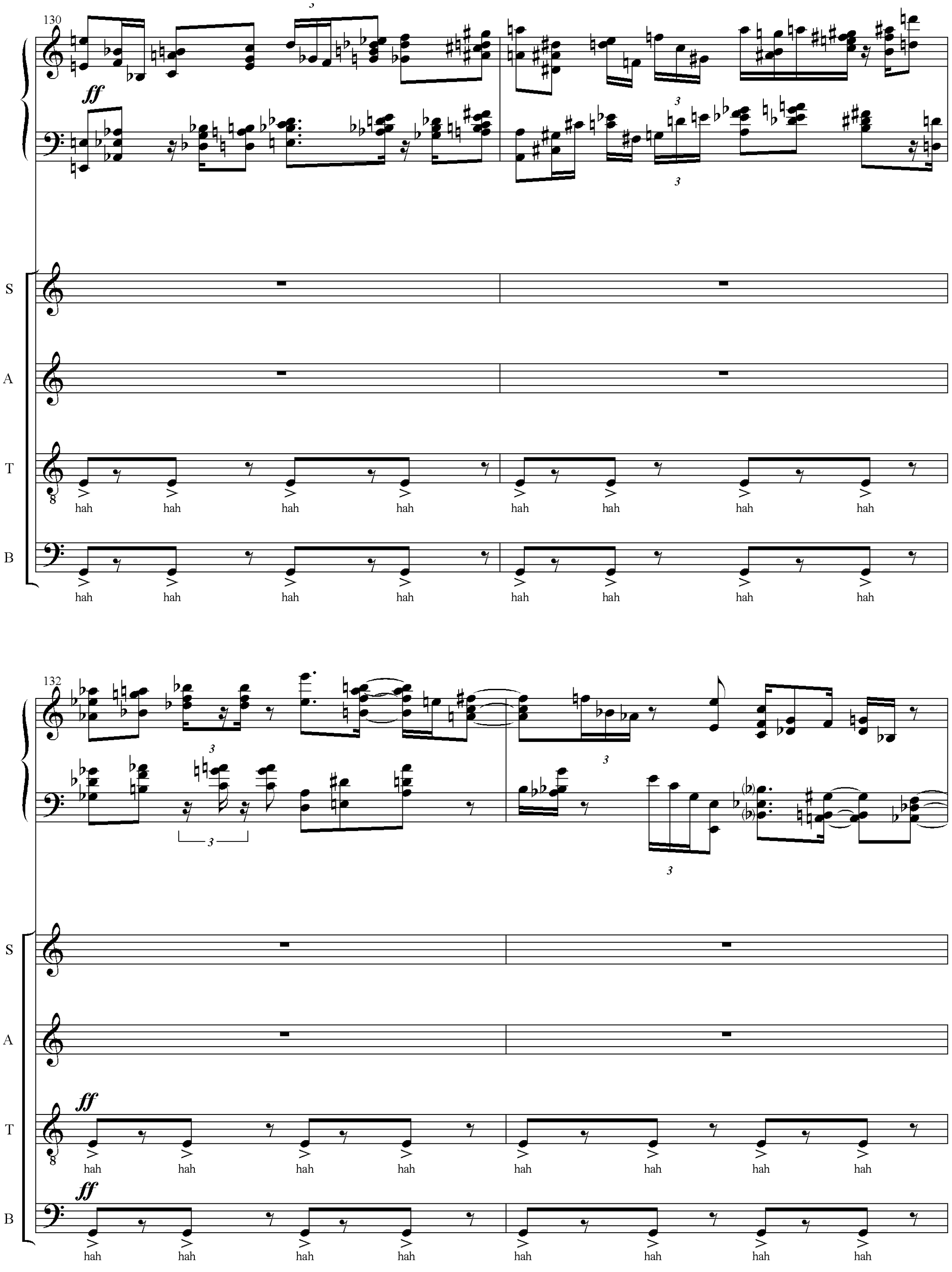

79
C
ppp
ff
8vb
loco
p
f
S
ss - - -
ti
A
T
no
B
mf
oh
15ma
81
mp
fff
f
mp
loco
8vb
(ti)
si - - - si - si na na - na
mp
mf
ta - - - - - na
ay
si
si
si na na na na
wah
oh - toh
p
na ta na ra
oh - toh
na ta na ra

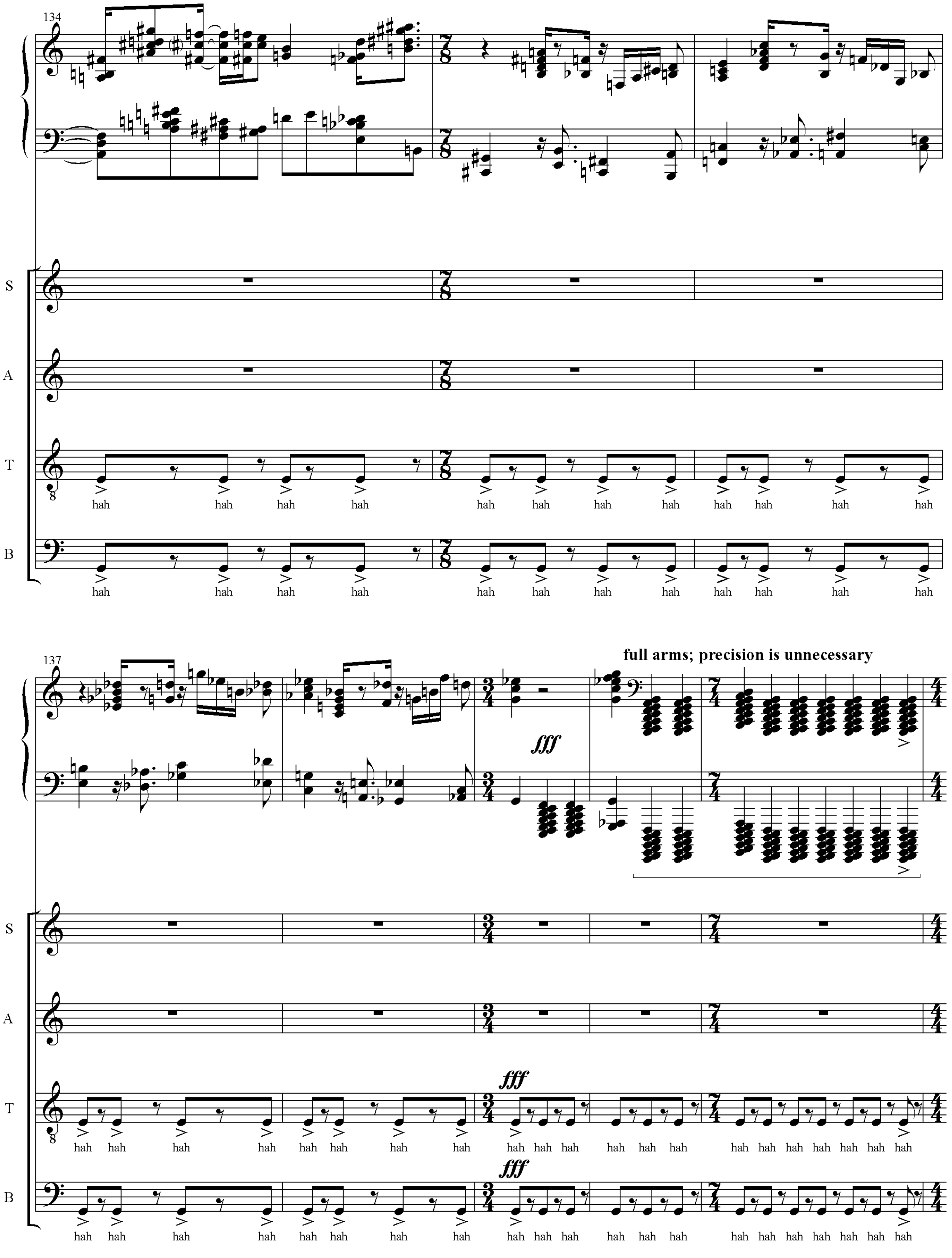

87
D
loco
mf
p
90
mf
92
f
p
94
f
mp
mf
f
98
p
mp
mf
p

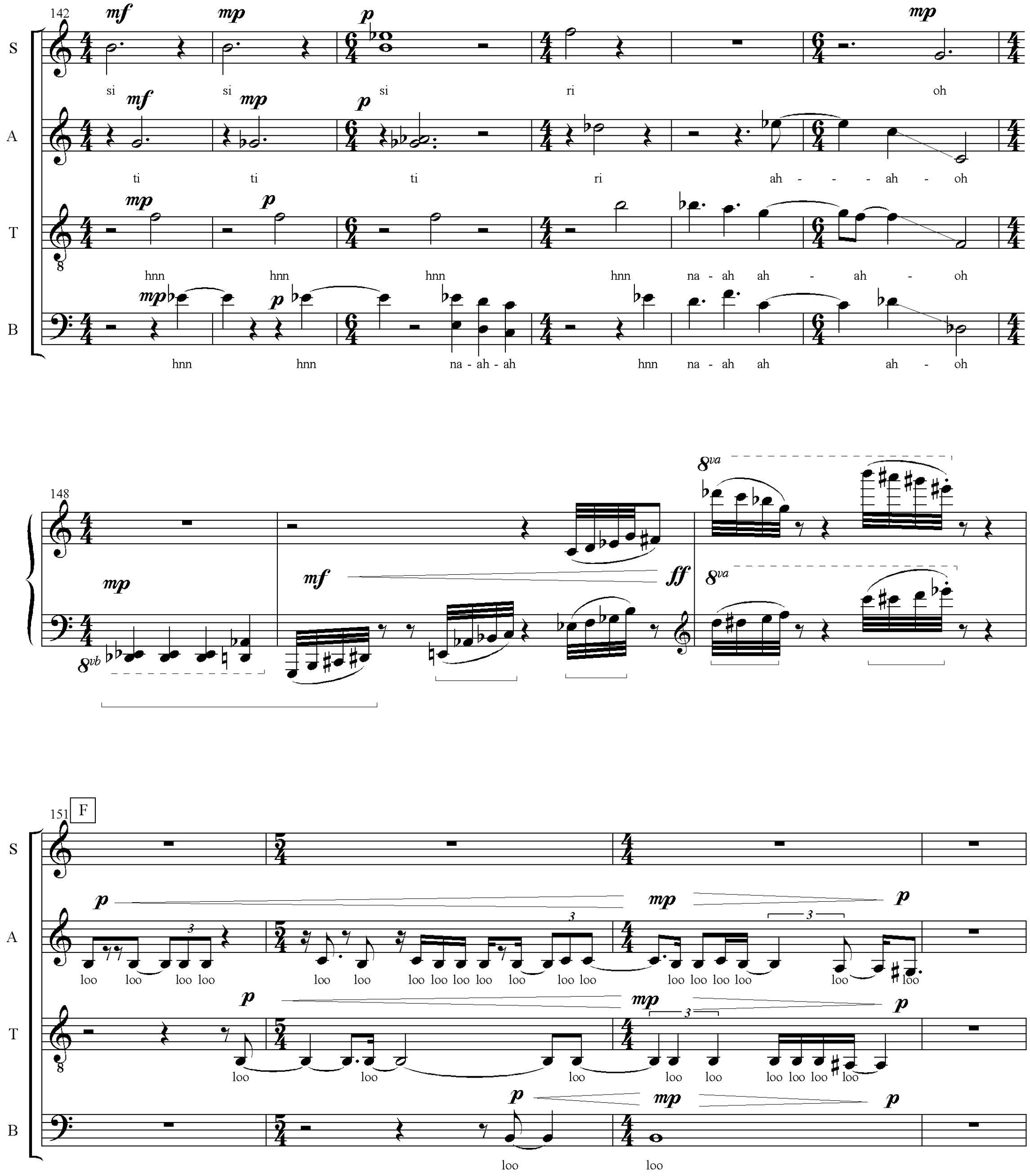

104
ff
bloodcurdling unpitched high scream
ffff
S
ffff bloodcurdling unpitched high scream
A
ffff bloodcurdling unpitched high scream
T
ffff bloodcurdling unpitched high scream
B
107
109
mf
mp
111
f
mf


113
mp
mf
117
f
S
A
T
B
mf
bzz
zz
zz
zz
zz
hur
zz
zz
mf
bzz
zz
zz
zz
zz
hur
zz
zz
121
123
mp
mf
f

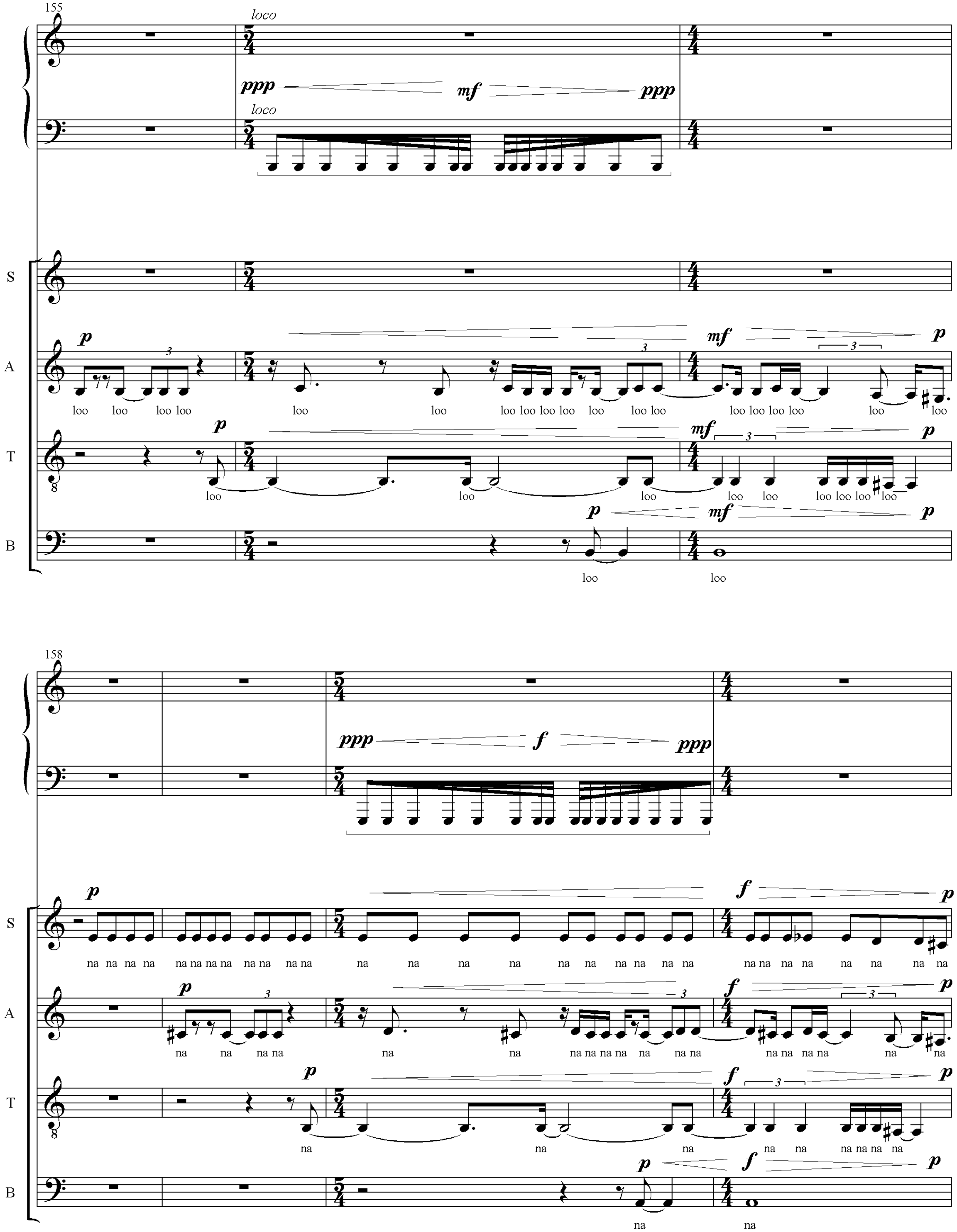

126
E
6
fff
S
A
T
f
hah hah hah hah hah hah hah hah
B
f
hah hah hah hah hah hah hah hah
128
3
mf
fff
S
A
T
hah hah hah hah hah hah hah hah
B
hah hah hah hah hah hah hah hah

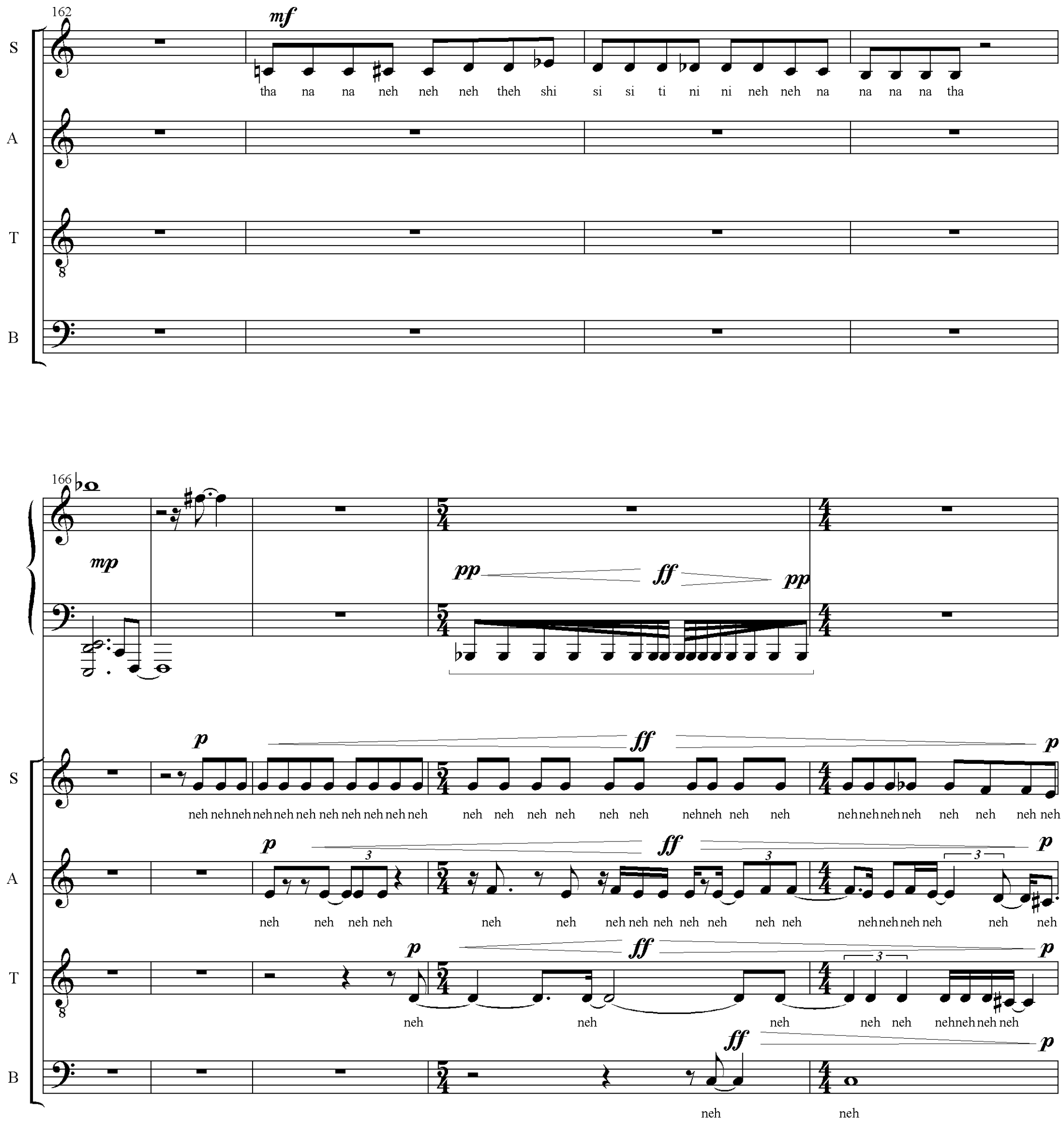

130
ff
3
3
3
S
A
T
B
hah hah hah hah hah hah hah hah
hah hah hah hah hah hah hah hah
132
3
3
3
(♭)
(♭)
ff
ff
S
A
T
B
hah hah hah hah hah hah hah hah
hah hah hah hah hah hah hah hah

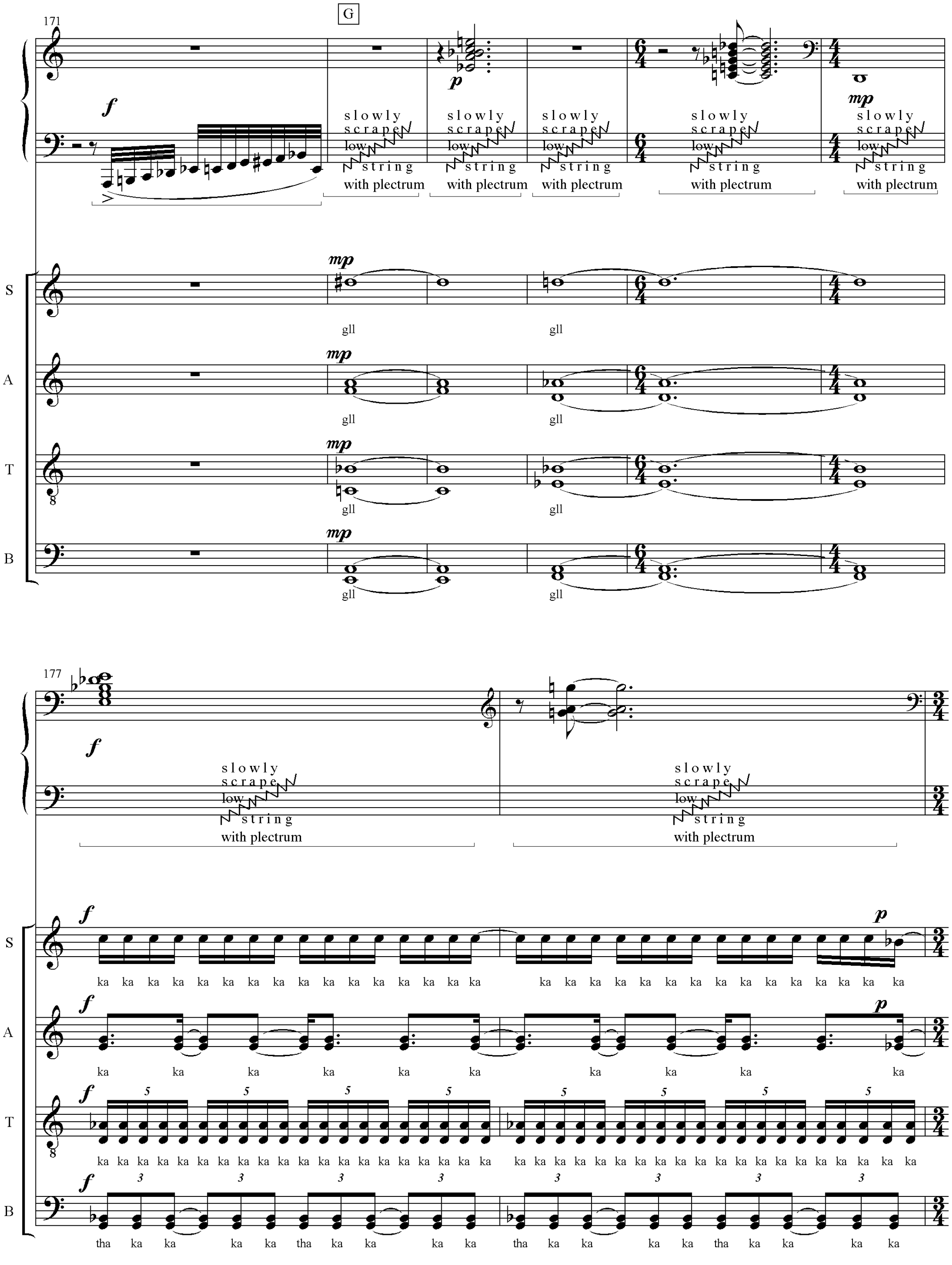

134
S
A
T
B
hah hah hah hah hah hah hah hah hah hah hah hah
137
full arms; precision is unnecessary
fff
hah hah hah hah hah hah hah hah hah hah hah hah hah hah hah hah hah hah hah hah hah hah
fff
fff

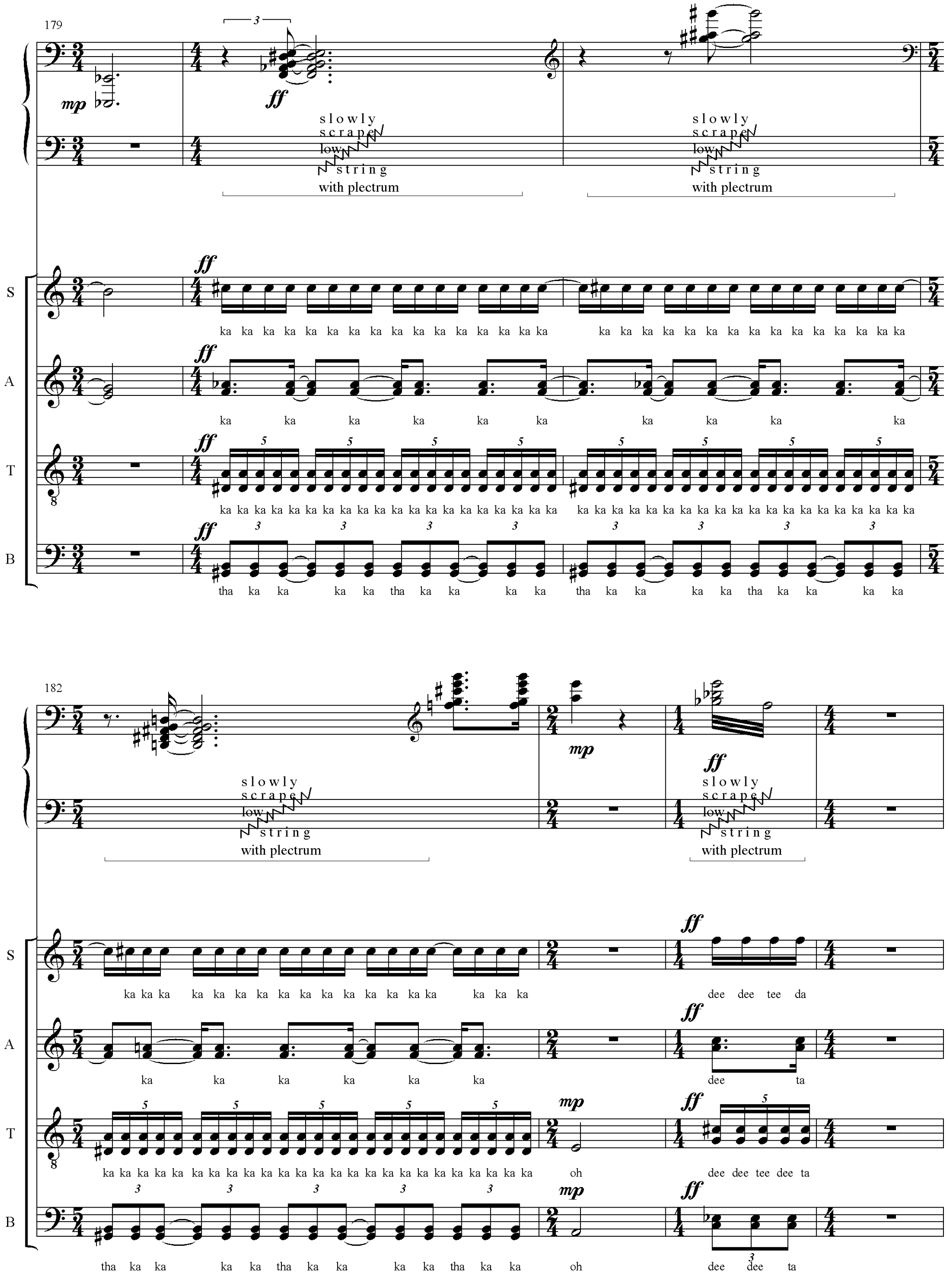


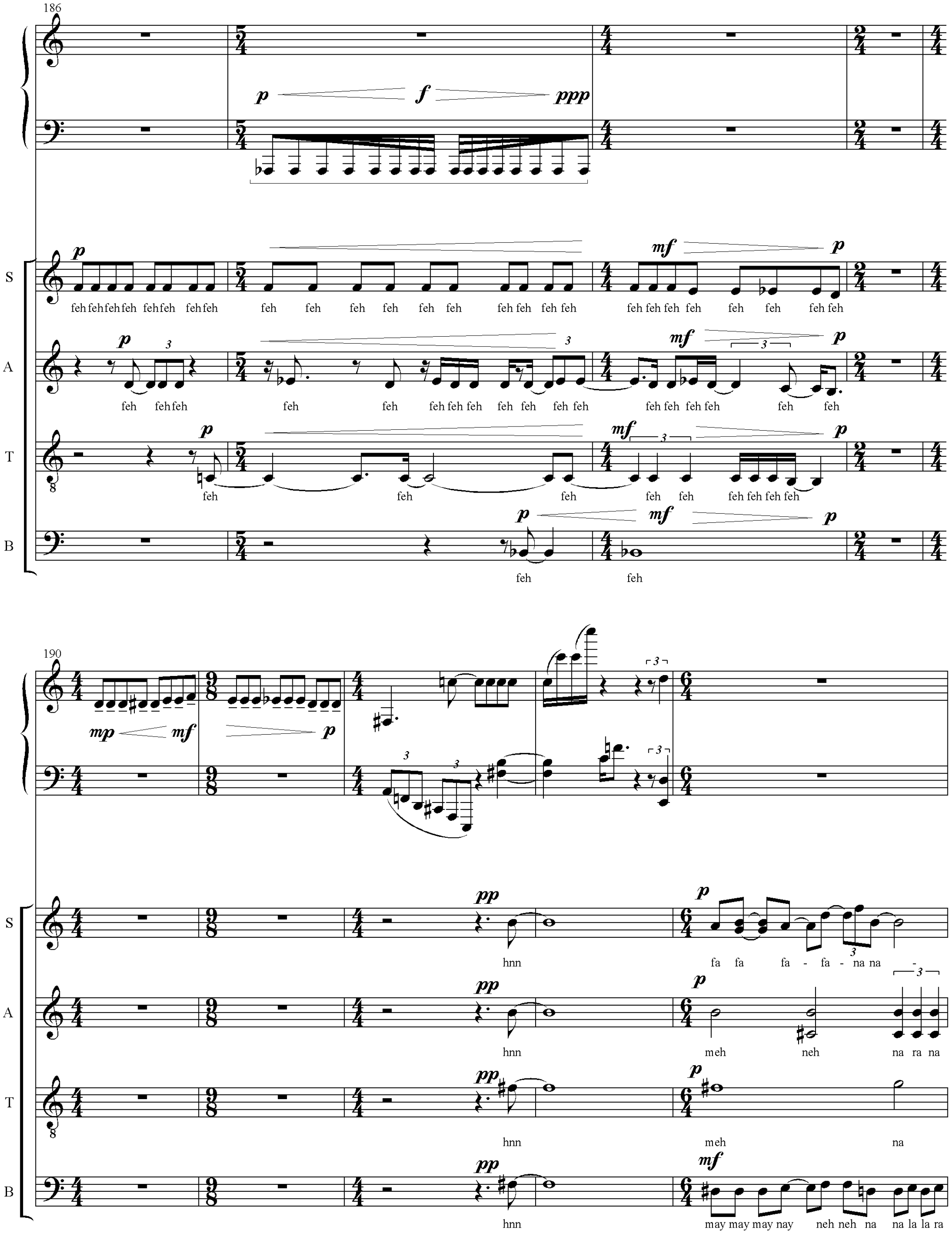
142
S
A
T
B
mf
mp
p
si si si ri oh
ti ti ti ri ah - - - ah - oh
hnn hnn hnn hnn na - ah ah - ah - oh
hnn hnn na - ah - ah hnn na - ah ah ah - oh
148
8va
8vb
mp
mf
ff
151
F
loo loo loo loo loo loo loo loo loo loo loo loo loo loo loo loo loo loo loo loo
loo loo loo loo loo loo loo loo loo
loo loo
3

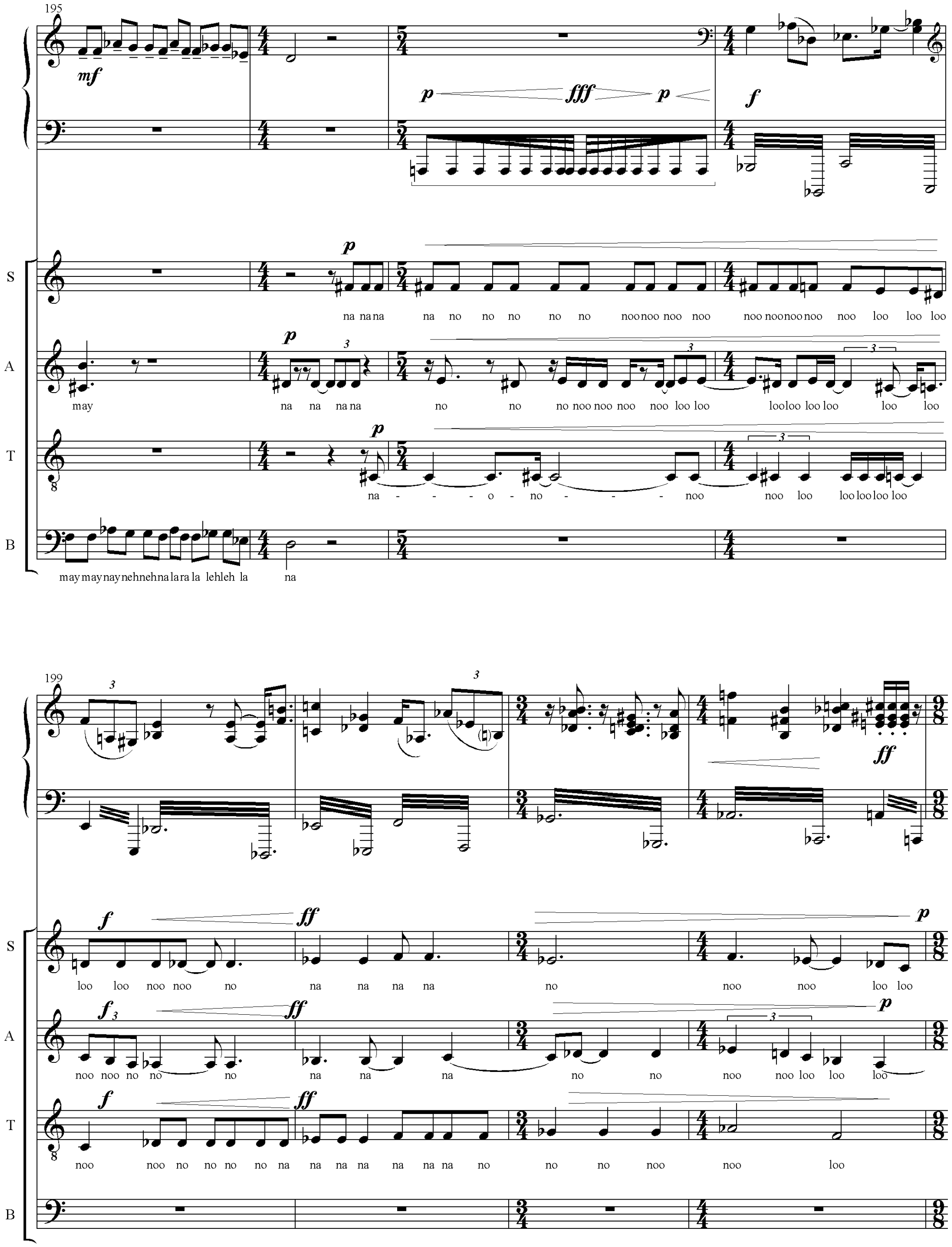

155
loco
ppp
mf
ppp
loco
S
A
T
B
p
mf
p
loo loo loo loo loo loo loo loo loo loo loo loo loo loo loo loo loo loo loo
loo loo loo loo loo loo loo loo
loo loo
158
ppp
f
ppp
na na na na na na na na na na na na na na na na na na na na na na na na na na na na
na na na na na na na na na na na na na na na
na na na na na na na na na
na na
3

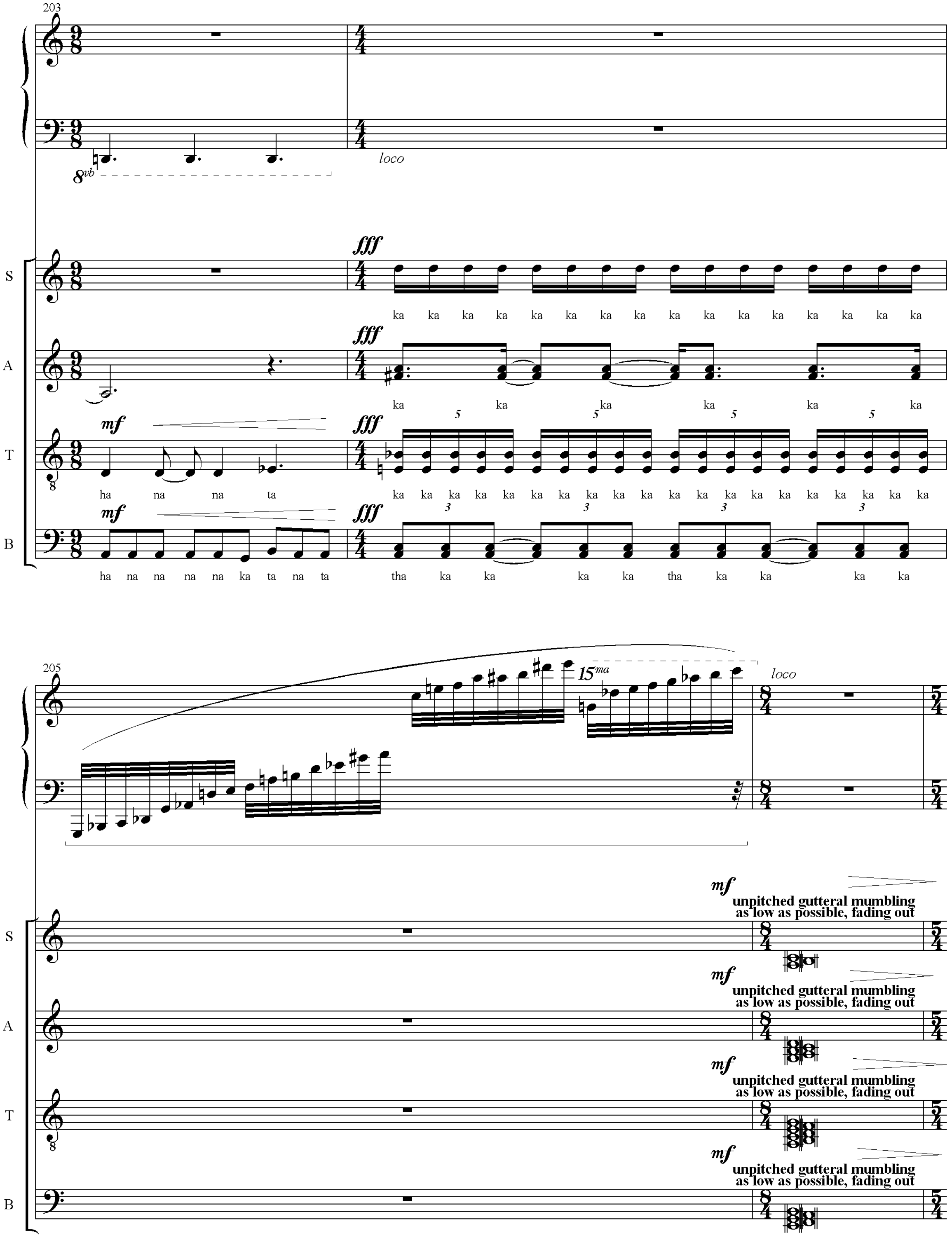

162
S
A
T
B
mf
tha na na neh neh neh theh shi si si ti ni ni neh neh na na na na tha
166
mp
pp
ff
pp
p
neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh
p
3
ff
p
neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh neh
p
ff
3
p
neh neh neh neh neh neh neh neh neh neh
ff
p
neh neh
5
4
4
4

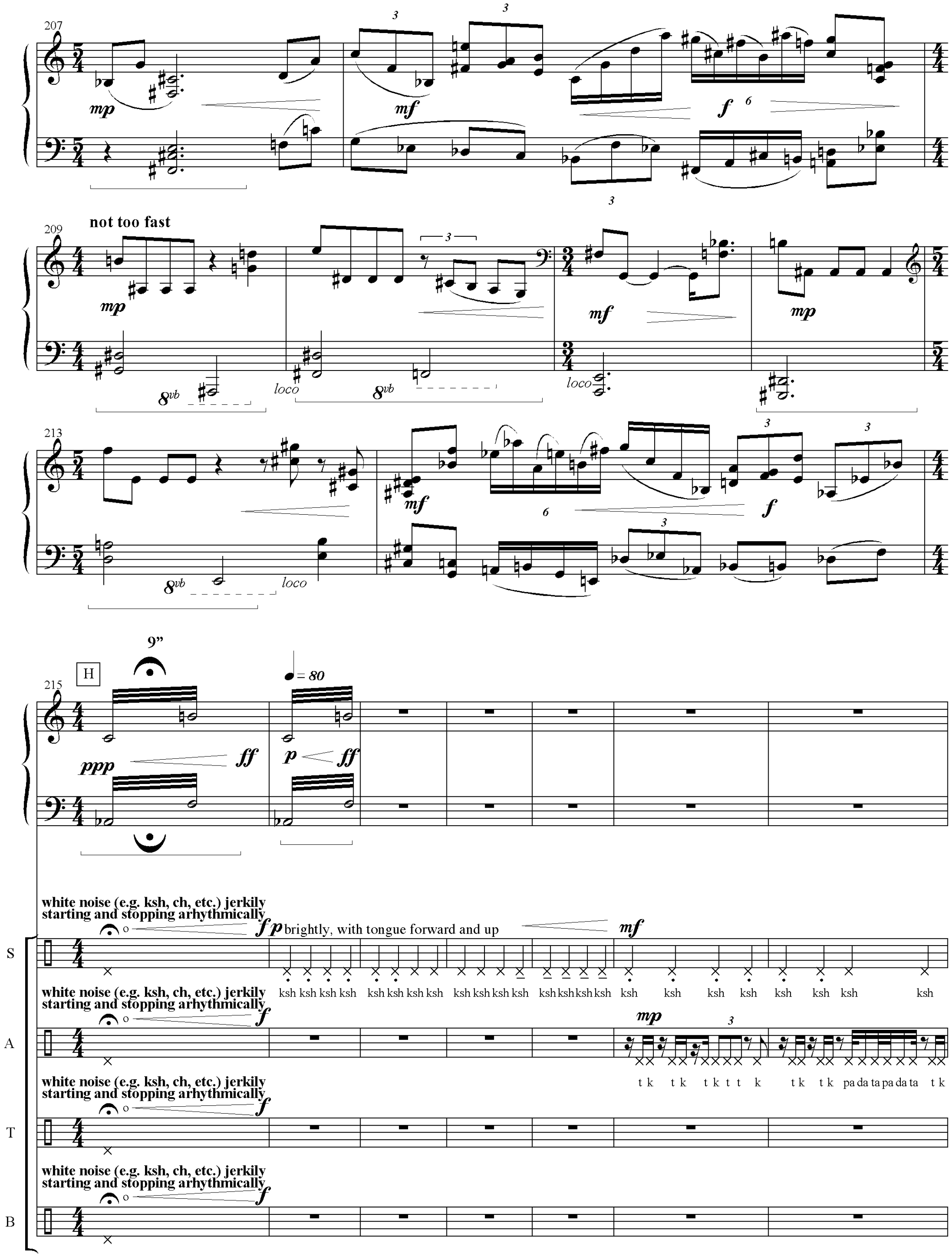

171
G
f
p
mp
slowly scrape low string with plectrum
S
A
T
B
gll
177
f
slowly scrape low string with plectrum
ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka
ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka
p
5
3
tha ka ka ka ka tha ka ka ka ka

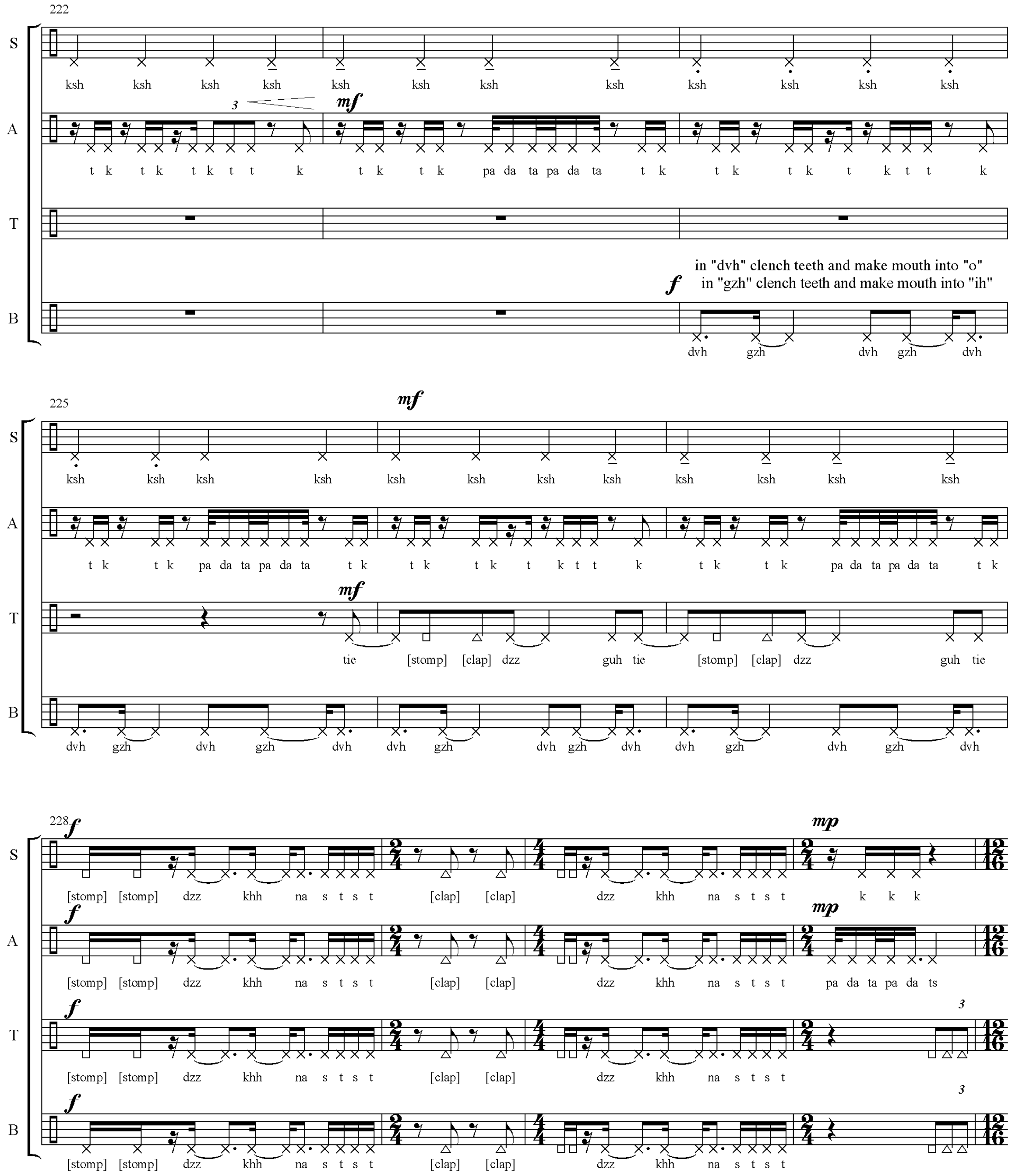

179
mp
ff
slowly scrape low string with plectrum
slowly scrape low string with plectrum
S
A
T
B
ff
ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka
ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka
ka ka ka ka ka ka
ka ka ka ka
ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka
ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka
tha ka ka ka ka tha ka ka ka ka
tha ka ka ka ka tha ka ka ka ka
182
slowly scrape low string with plectrum
mp
ff
slowly scrape low string with plectrum
ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka
dee dee tee da
ka ka ka ka ka ka
dee ta
ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka
oh
dee dee tee dee ta
tha ka ka ka ka tha ka ka ka ka tha ka ka
oh
dee dee ta

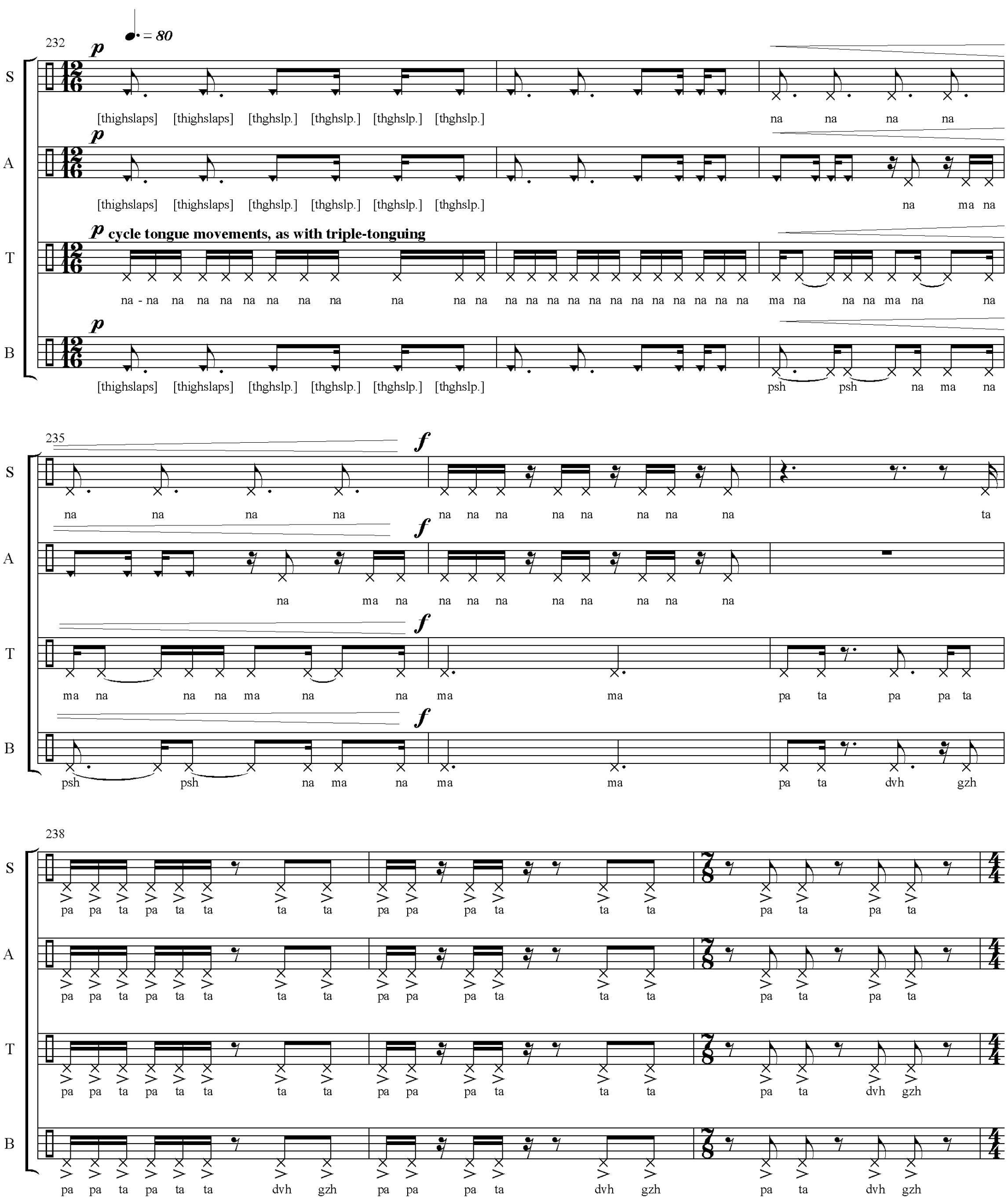

186
p f ppp
S
A
T
B
p mf p
feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh
feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh feh
feh feh feh feh feh feh feh feh feh
feh feh
190
mp mf p
pp
hnn
fa fa fa - fa - na na -
meh neh na ra na
meh na
may may may nay neh neh na na la la ra
mf

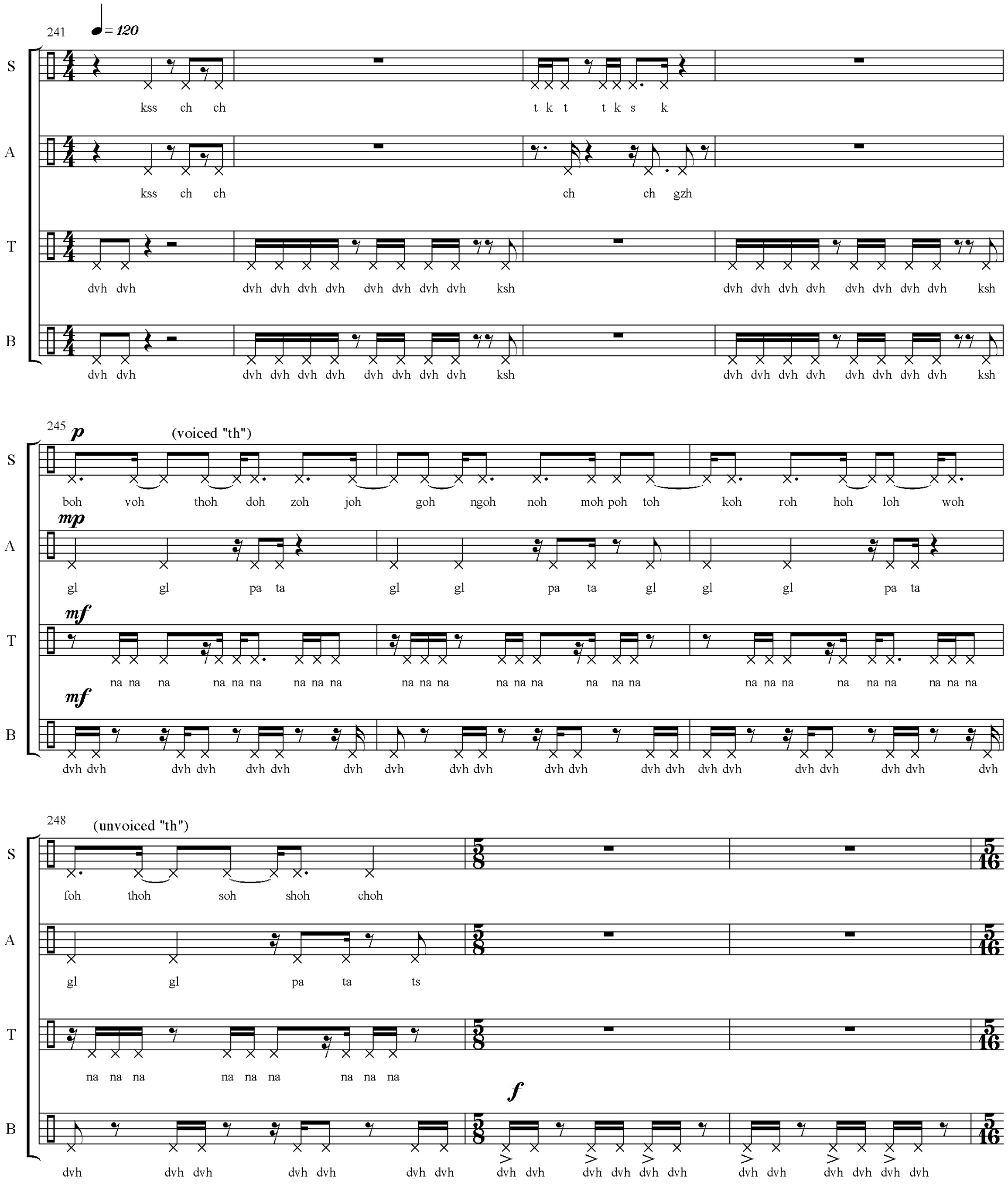

195
mf
p fff p
f
S
A
T
B
p
na na na na no no no no no noo noo noo noo noo noo noo noo noo loo loo loo
may
na na na na no no no noo noo noo noo loo loo loo loo loo loo loo loo loo
na - - - o - no - - - noo noo loo loo loo loo loo
may may nay neh neh na la ra la leh leh la na
199
ff
f ff p
loo loo noo noo no na na na na no noo noo loo loo
noo noo no no no na na na no no noo noo loo loo loo
noo noo no no no no na na na na na na na no no no noo noo loo

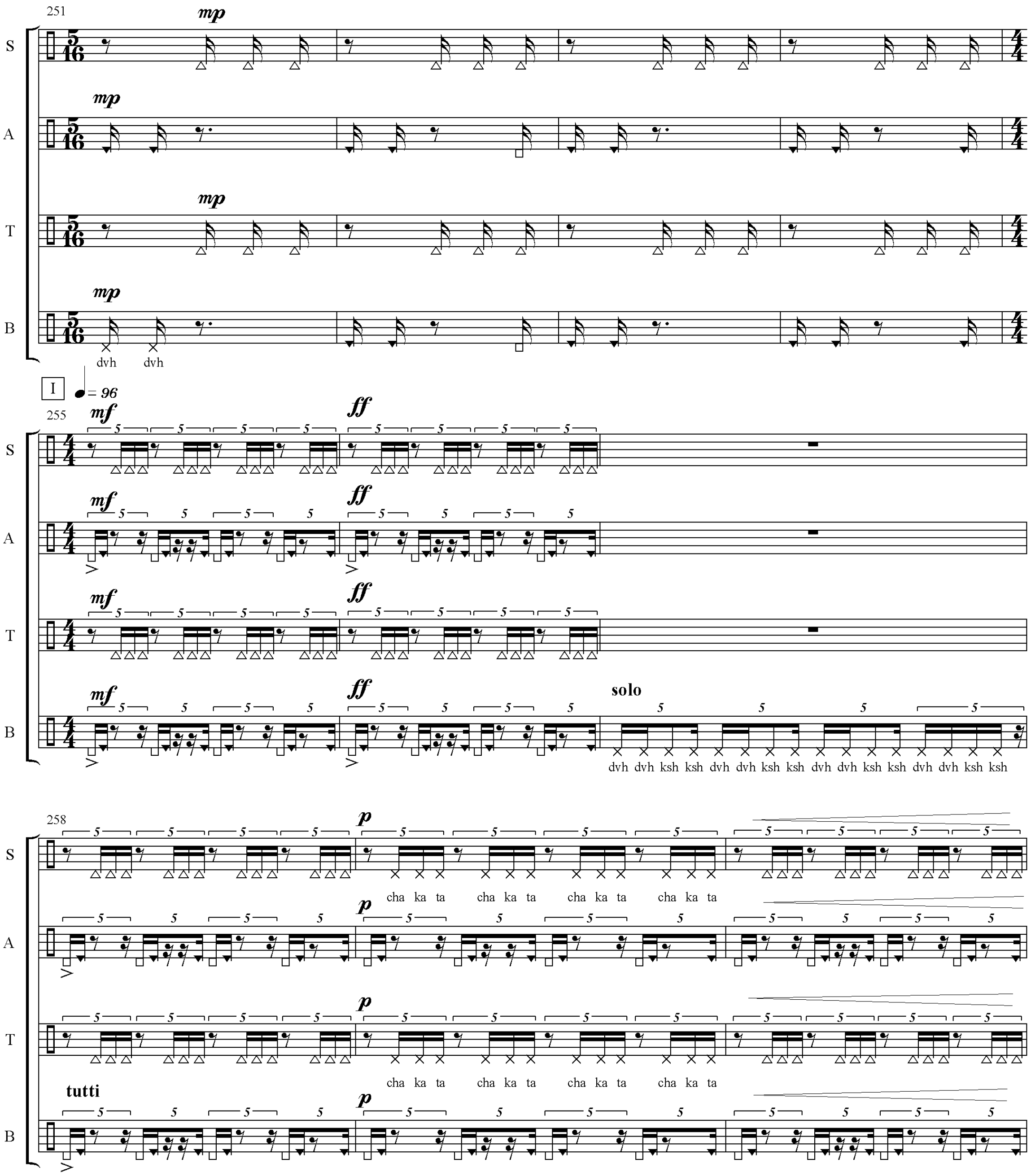

203
8vb
loco
S
A
T
B
fff
mf
ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka
ka ka ka ka ka ka
5
ha na na ta
ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka ka
3
ha na na na na ka ta na ta
tha ka ka ka ka tha ka ka ka ka
205
15ma
loco
unpitched gutteral mumbling
as low as possible, fading out

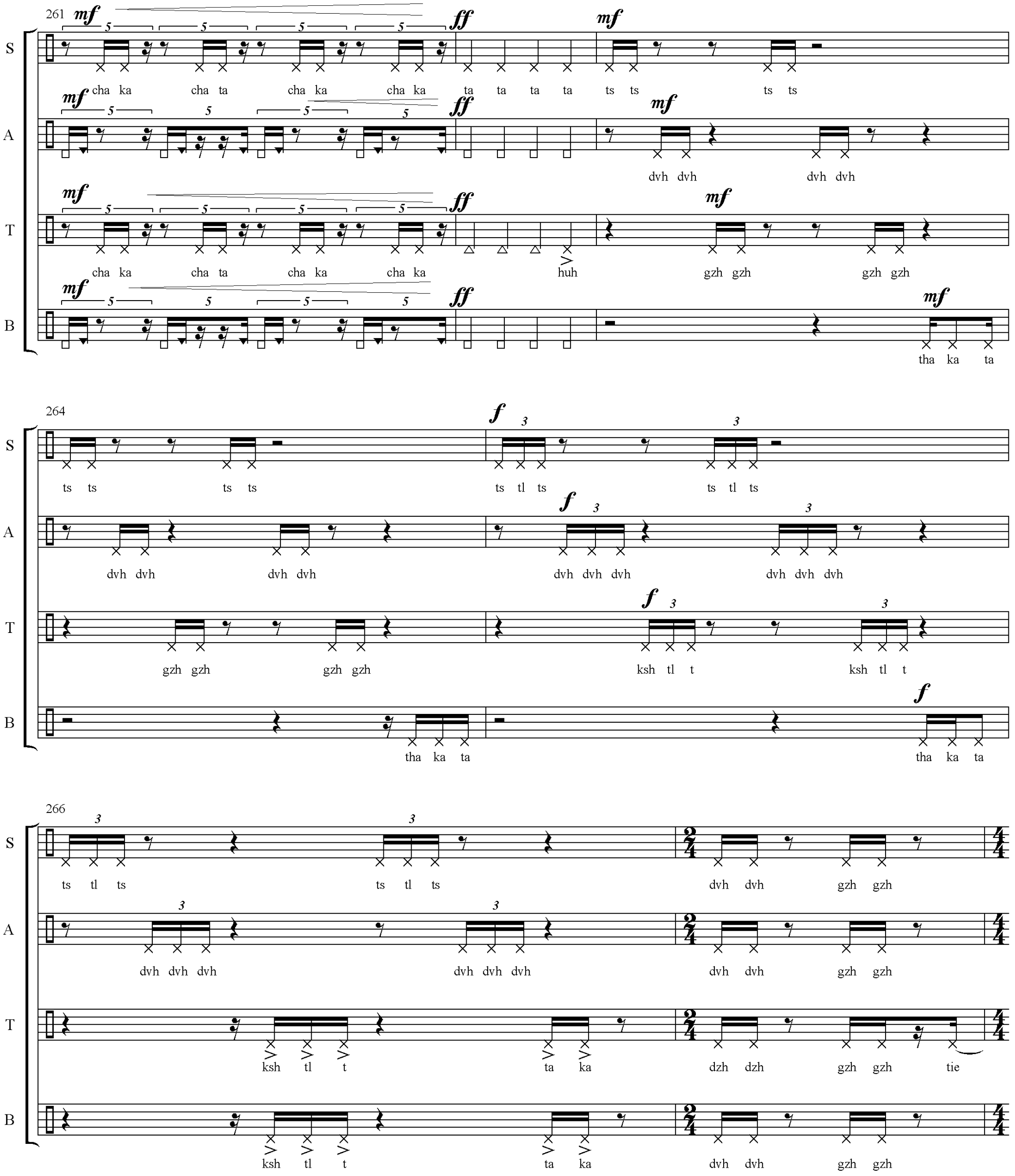

207
mp
mf
f
6
3
209
not too fast
mp
8vb
loco
mf
213
8vb
loco
H
215
9”
♩ = 80
ppp
ff
p
ff
S
A
T
B
white noise (e.g. ksh, ch, etc.) jerkily starting and stopping arhythmically
fp
brightly, with tongue forward and up
mf
ksh ksh ksh ksh ksh ksh ksh ksh ksh ksh ksh ksh kshkshkshksh ksh ksh ksh ksh ksh ksh ksh ksh
mp
t k t k t k t t k t k t k pa da ta pa da ta t k
f

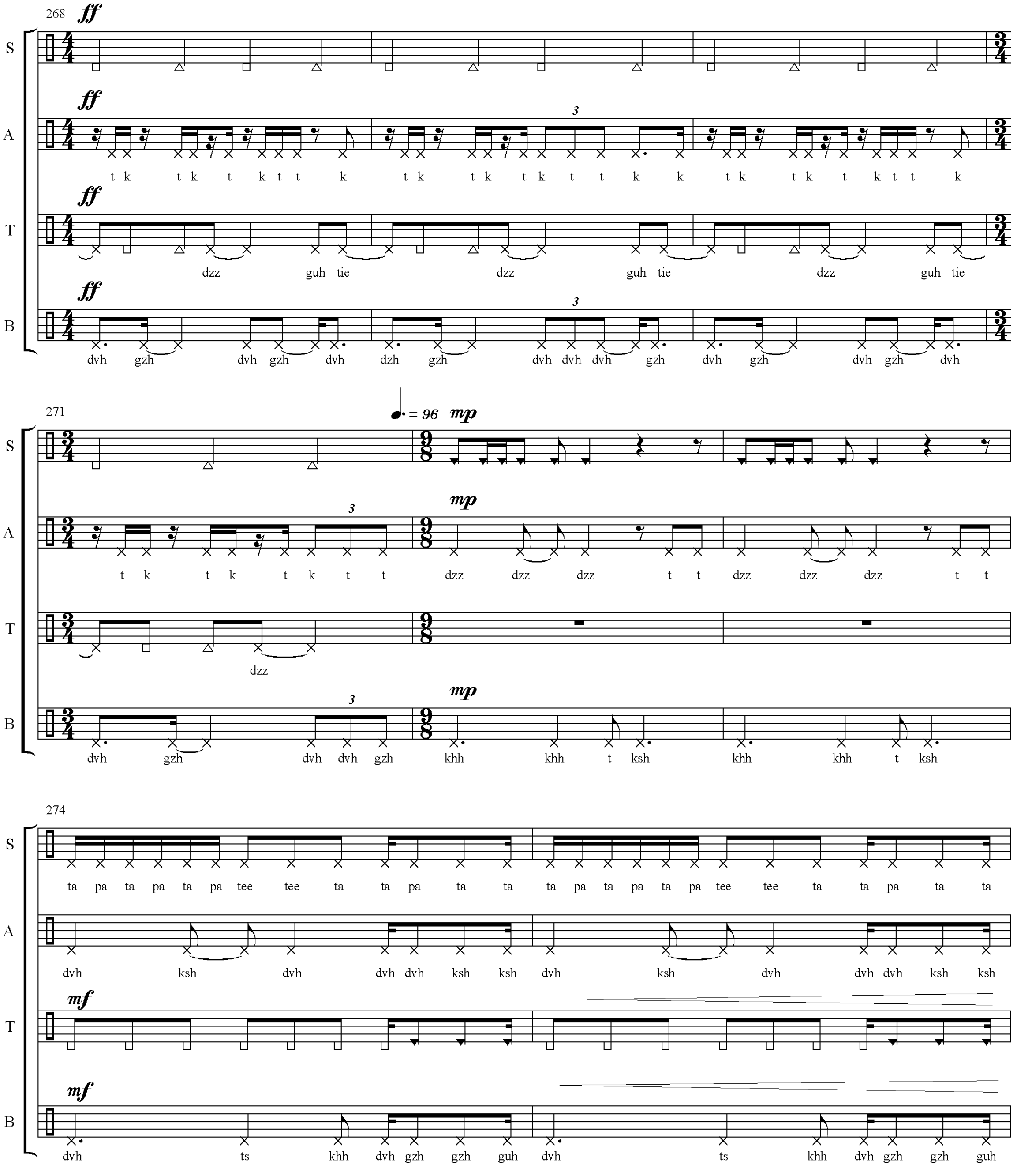

222
S
A
T
B
ksh ksh ksh ksh ksh ksh ksh ksh ksh ksh ksh ksh
3
mf
t k t k t k t t k t k t k pa da ta pa da ta t k t k t k t k t t k
f
in "dvh" clench teeth and make mouth into "o"
in "gzh" clench teeth and make mouth into "ih"
dvh gzh dvh gzh dvh
225
mf
ksh ksh ksh ksh ksh ksh ksh ksh ksh ksh ksh ksh
t k t k pa da ta pa da ta t k t k t k t k t t k t k t k pa da ta pa da ta t k
mf
tie [stomp] [clap] dzz guh tie [stomp] [clap] dzz guh tie
dvh gzh dvh gzh dvh dvh gzh dvh gzh dvh dvh gzh dvh gzh dvh
228
f
[stomp] [stomp] dzz khh na s t s t [clap] [clap] dzz khh na s t s t k k k
mp
[stomp] [stomp] dzz khh na s t s t [clap] [clap] dzz khh na s t s t pa da ta pa da ts
3
[stomp] [stomp] dzz khh na s t s t [clap] [clap] dzz khh na s t s t
[stomp] [stomp] dzz khh na s t s t [clap] [clap] dzz khh na s t s t
2/4
4/4
12/16

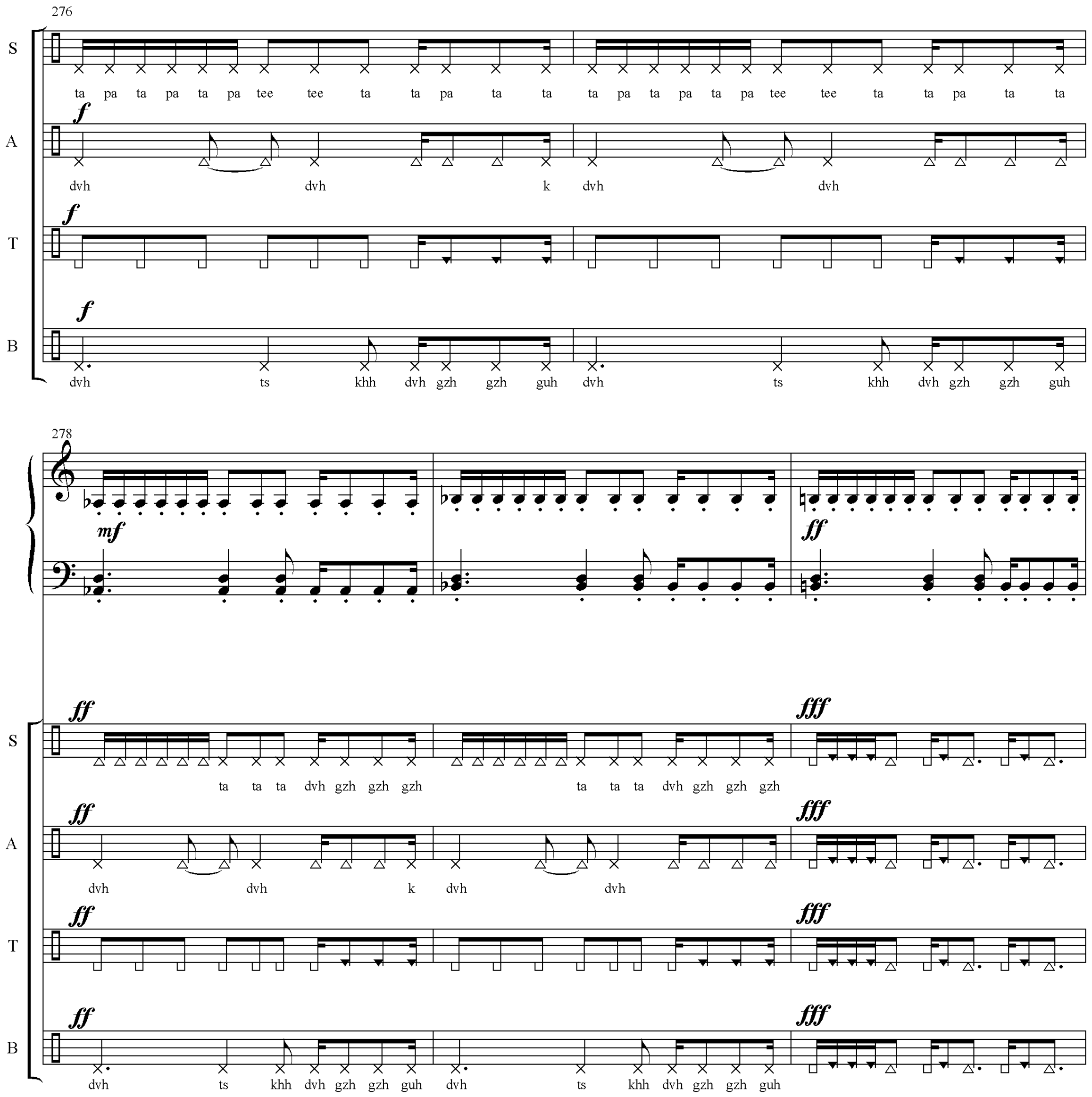

232
♩. = 80
S
A
T
B
p
[thighslaps] [thighslaps] [thghslp.] [thghslp.] [thghslp.] [thghslp.]
na na na na
[thighslaps] [thighslaps] [thghslp.] [thghslp.] [thghslp.] [thghslp.]
na ma na
p cycle tongue movements, as with triple-tonguing
na - na na na na na na na na na na na na na na na na na na na na na na na
ma na na na ma na na
[thighslaps] [thighslaps] [thghslp.] [thghslp.] [thghslp.] [thghslp.]
psh psh na ma na
235
f
na na na na na na na na na na na na ta
na ma na na na na na na na na na na
ma na na na ma na na ma ma pa ta pa pa ta
psh psh na ma na ma ma pa ta dvh gzh
238
pa pa ta pa ta ta ta ta pa pa pa ta ta ta pa ta pa ta
pa pa ta pa ta ta ta ta pa pa pa ta ta ta pa ta pa ta
pa pa ta pa ta ta ta ta pa pa pa ta ta ta pa ta dvh gzh
pa pa ta pa ta ta dvh gzh pa pa pa ta dvh gzh pa ta dvh gzh

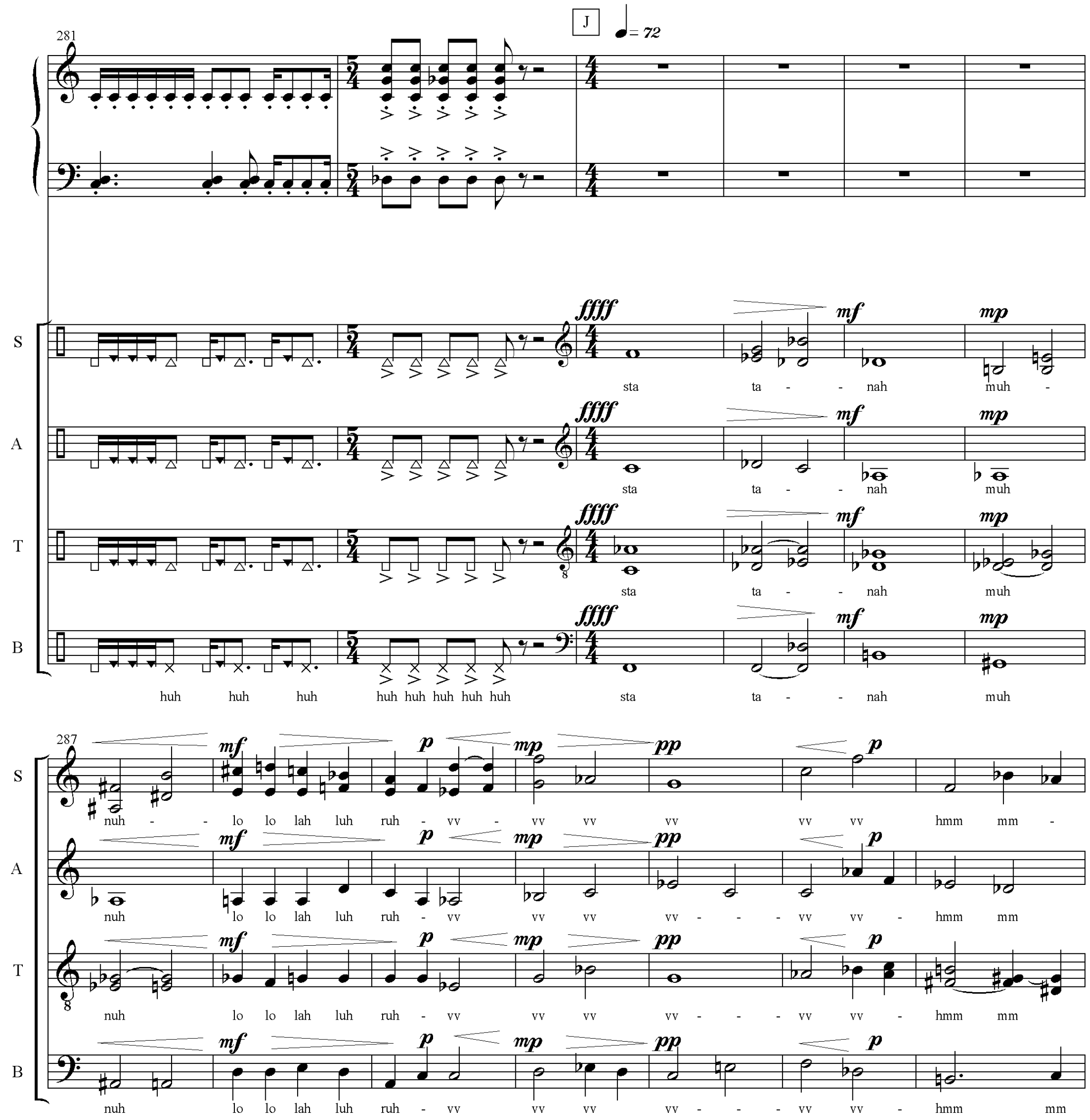

241
♩= 120
S
A
T
B
kss ch ch
t k t t k s k
kss ch ch
ch ch gzh
dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh ksh dvh dvh dvh dvh dvh dvh dvh dvh ksh
dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh ksh dvh dvh dvh dvh dvh dvh dvh dvh ksh
245
p
(voiced "th")
boh voh thoh doh zoh joh goh ngoh noh moh poh toh koh roh hoh loh woh
mp
gl gl pa ta gl gl pa ta gl gl gl pa ta
mf
na na na na na na na na na na na na na na na na na na na na na na na na na na
mf
dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh
248
(unvoiced "th")
foh thoh soh shoh choh
gl gl pa ta ts
na na na na na na na na na
f
dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh dvh

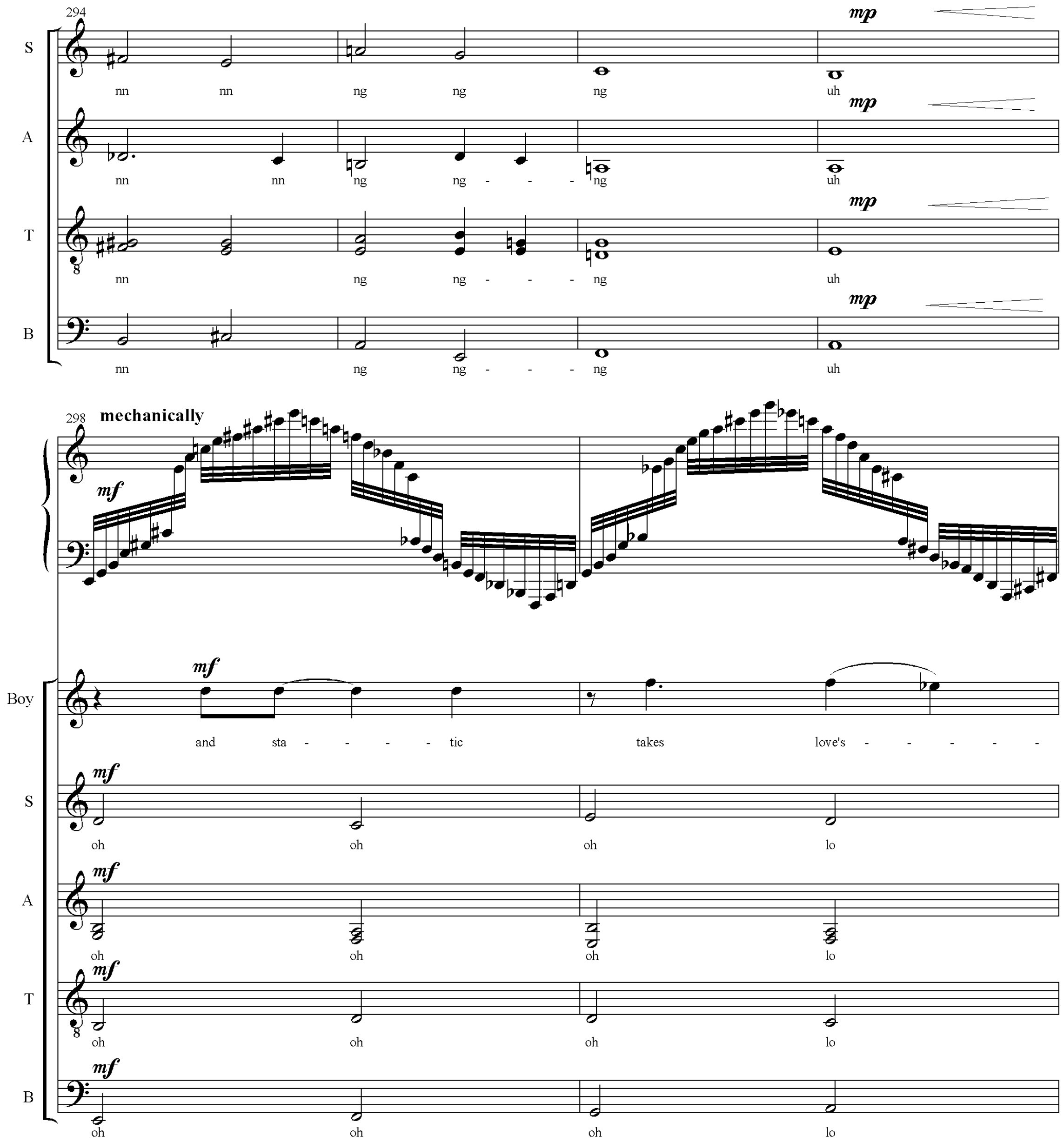

251
S
A
T
B
mp
dvh dvh
I
♩= 96
255
mf
ff
5
solo
dvh dvh ksh ksh dvh dvh ksh ksh dvh dvh ksh ksh dvh dvh ksh ksh
258
tutti
p
cha ka ta cha ka ta cha ka ta cha ka ta

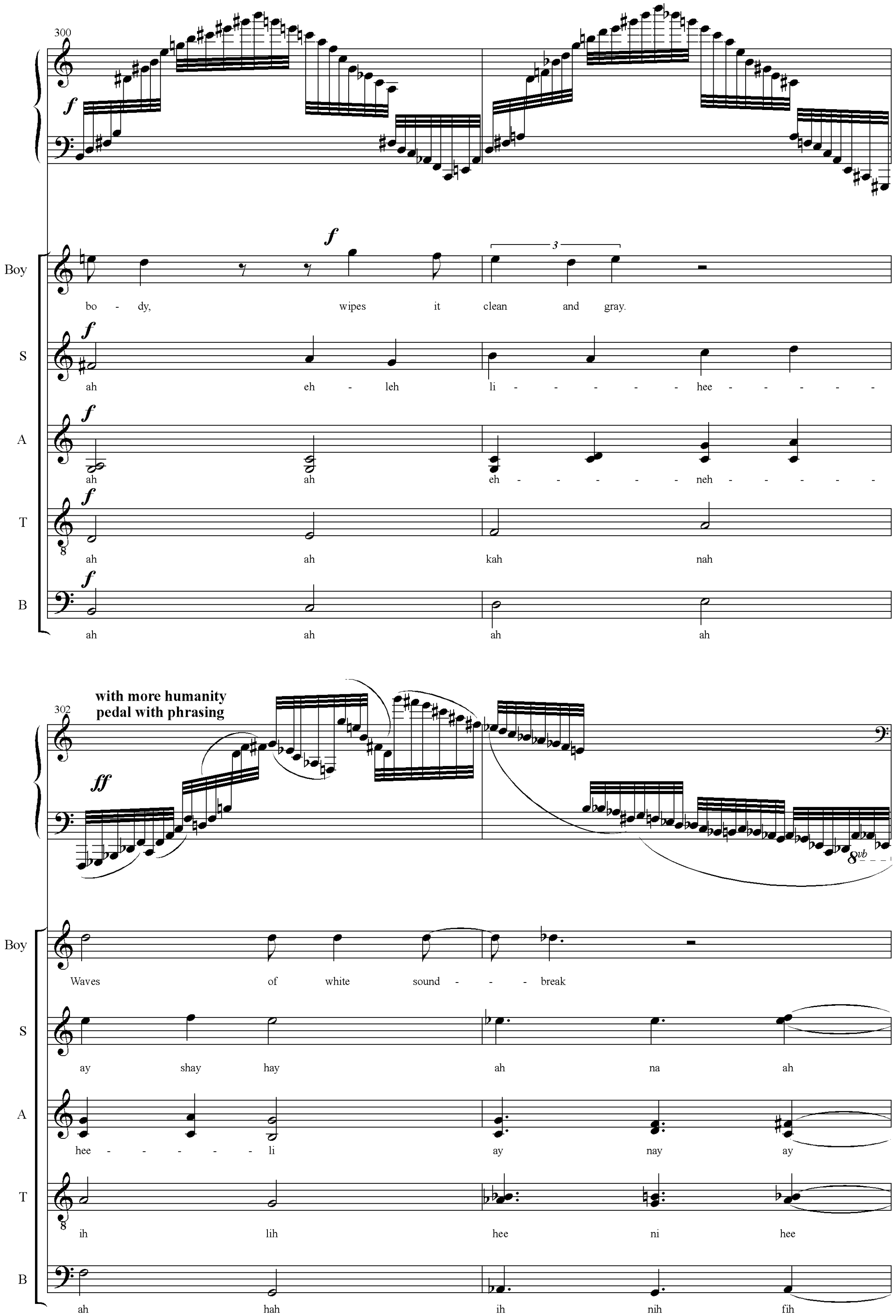

261
S
A
T
B
mf
5
ff
mf
cha ka cha ta cha ka cha ka
ta ta ta ta
ts ts ts ts
dvh dvh dvh dvh
cha ka cha ta cha ka cha ka
huh
gzh gzh gzh gzh
tha ka ta
264
f
3
ts ts ts ts
ts tl ts ts tl ts
dvh dvh dvh dvh
dvh dvh dvh dvh dvh dvh
gzh gzh gzh gzh
ksh tl t ksh tl t
tha ka ta tha ka ta
266
ts tl ts ts tl ts
dvh dvh gzh gzh
dvh dvh dvh dvh dvh dvh
dvh dvh gzh gzh
ksh tl t ta ka
dzh dzh gzh gzh tie
ksh tl t ta ka
dvh dvh gzh gzh
2
4
4
4

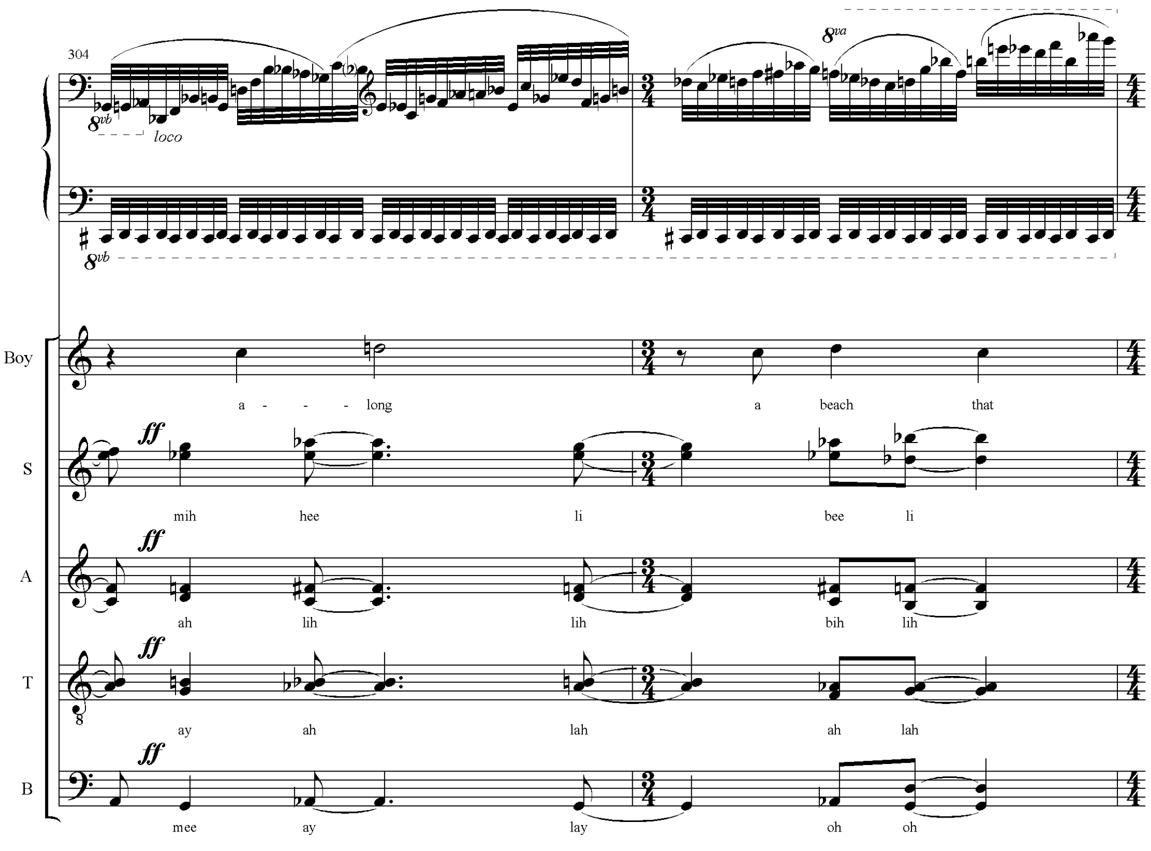

268
S
A
T
B
ff
3
t k t k t k t t k t k t k t k t t k k t k t k t k t t k
dzz guh tie dzz guh tie dzz guh tie
dvh gzh dvh gzh dvh dzh gzh dvh dvh dvh gzh dvh gzh dvh gzh dvh
271
♩. = 96
mp
t k t k t k t t dzz dzz dzz t t dzz dzz dzz t t
dzz
dvh gzh dvh dvh gzh khh khh t ksh khh khh t ksh
274
mf
ta pa ta pa ta pa tee tee ta ta pa ta ta ta pa ta pa ta pa tee tee ta ta pa ta ta
dvh ksh dvh dvh dvh ksh ksh dvh ksh dvh dvh dvh ksh ksh
dvh ts khh dvh gzh gzh guh dvh ts khh dvh gzh gzh guh

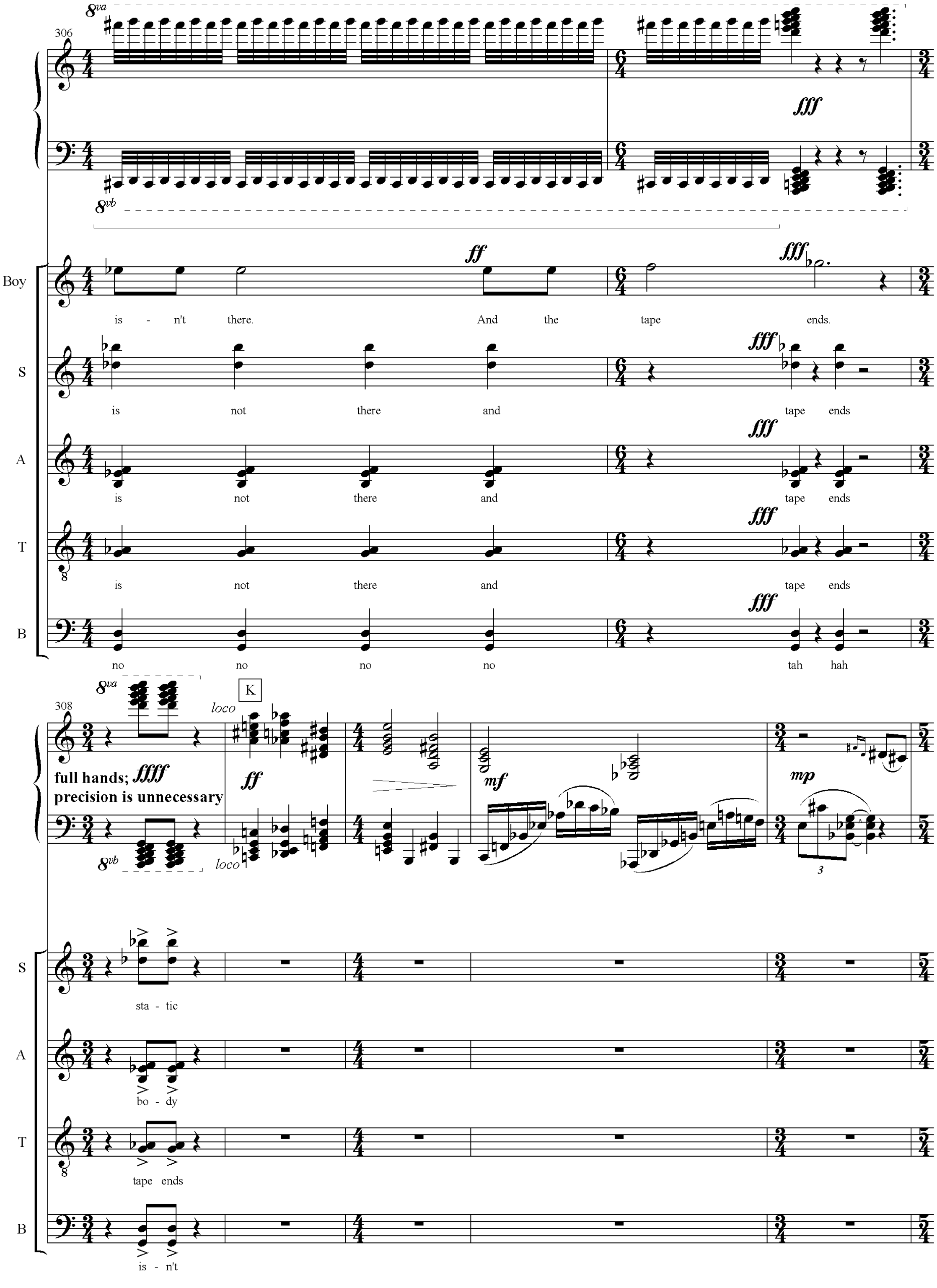

276
S
A
T
B
f
ta pa ta pa ta pa tee tee ta ta pa ta ta ta pa ta pa ta pa tee tee ta ta pa ta ta
dvh dvh k dvh dvh
dvh ts khh dvh gzh gzh guh dvh ts khh dvh gzh gzh guh
278
mf
ff
fff
ta ta ta dvh gzh gzh gzh ta ta ta dvh gzh gzh gzh
dvh dvh k dvh dvh
dvh ts khh dvh gzh gzh guh dvh ts khh dvh gzh gzh guh

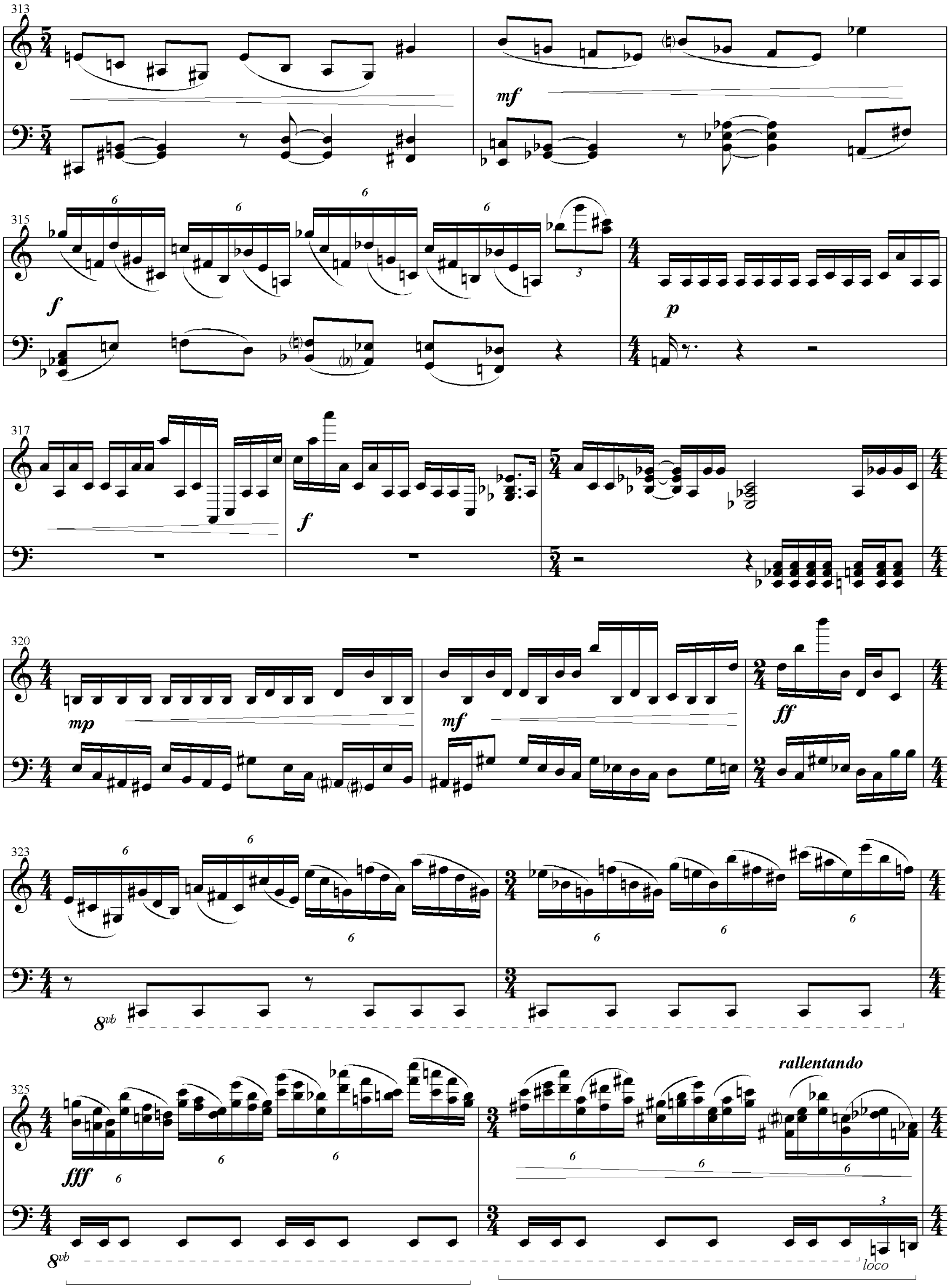

281
J
♩= 72
S
A
T
B
ffff
mf
mp
huh huh huh huh huh huh huh huh
sta ta - - nah muh -
sta ta - - nah muh
sta ta - - nah muh
sta ta - - nah muh
287
mf
p
mp
pp
nuh - - lo lo lah luh ruh - vv - vv vv vv vv vv hmm mm -
nuh lo lo lah luh ruh - vv vv vv vv - - - vv vv - hmm mm
nuh lo lo lah luh ruh - vv vv vv vv - - - vv vv - hmm mm
nuh lo lo lah luh ruh - vv vv vv vv - - - vv vv - hmm mm

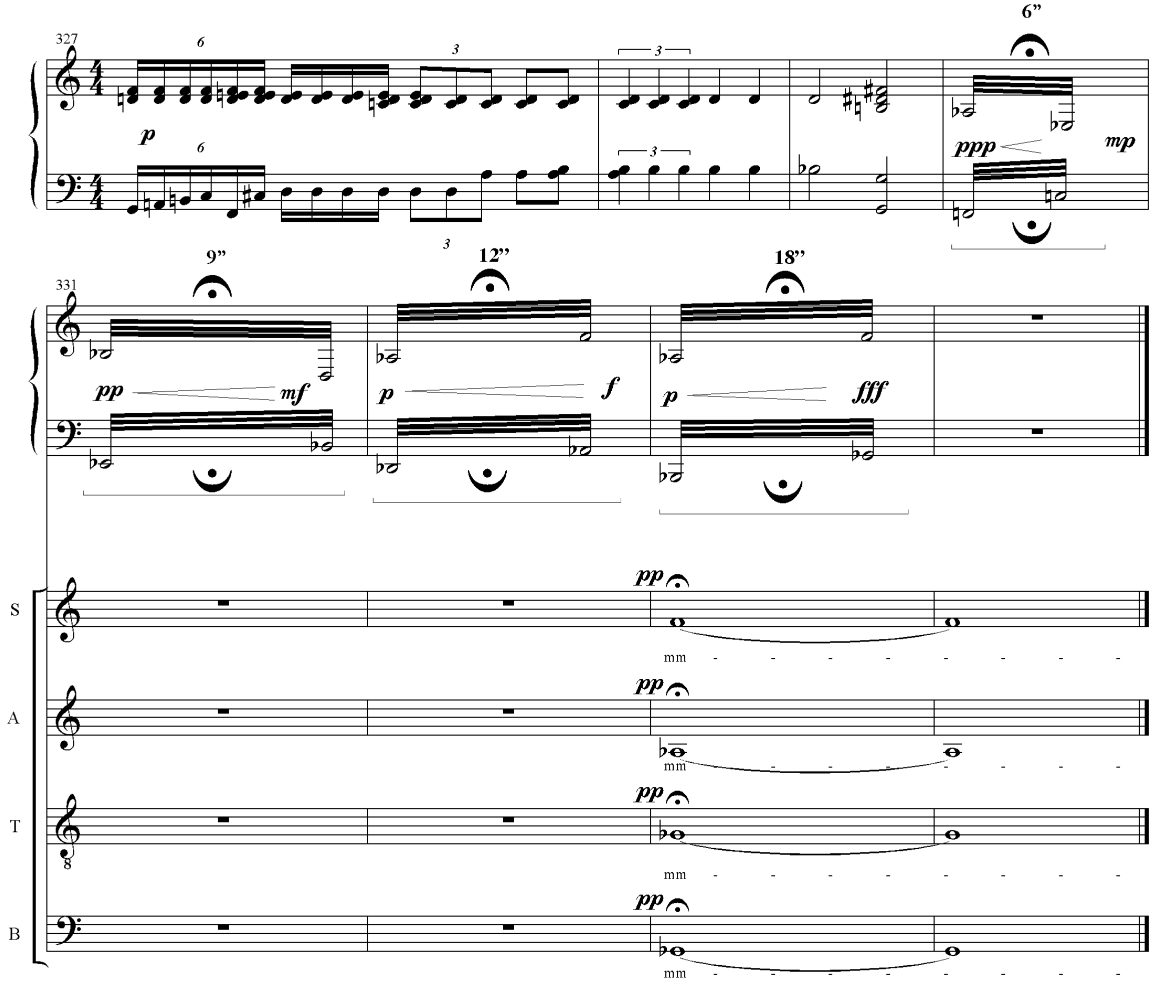

294
S
A
T
B
mp
nn nn ng ng ng uh
nn nn ng ng - - - ng uh
nn ng ng - - - ng uh
nn ng ng - - - ng uh
298
mechanically
mf
Boy
and sta - - - - - tic takes love's - - - - - -
oh oh oh lo
oh oh oh lo
oh oh oh lo
oh oh oh lo


300
f
Boy
bo - dy, wipes it clean and gray.
S
ah eh - leh li - - - - - - hee - - - - - -
A
ah ah eh - - - - - - neh - - - - - -
T
ah ah kah nah
B
ah ah ah ah
302
with more humanity
pedal with phrasing
ff
8vb
Boy
Waves of white sound - - - break
S
ay shay hay ah na ah
A
hee - - - - - li ay nay ay
T
ih lih hee ni hee
B
ah hah ih nih fih


304
8vb
loco
8va
8vb
Boy
a - - - long
a beach that
S
ff
mih hee li bee li
A
ff
ah lih lih bih lih
T
ff
ay ah lah ah lah
B
ff
mee ay lay oh oh


306
8va
8vb
fff
Boy
ff
fff
is - n't there. And the tape ends.
S
is not there and tape ends
A
is not there and tape ends
T
is not there and tape ends
B
no no no no tah hah
308
8va
K
loco
full hands; ffff
precision is unnecessary
8vb
loco
ff
mf
mp
3
S
sta - tic
A
bo - dy
T
tape ends
B
is - n't


313
mf
315
f
6
3
p
317
f
320
mp
mf
ff
323
6
8vb
325
fff
rallentando
loco


327
6
p
6
3
3
3
6”
ppp
mp
331
9”
3
12”
18”
pp
mf
p
f
p
fff
S
A
T
B
pp
pp
pp
pp
8
mm
mm
mm
mm

BIBLIOGRAPHY

Adorno, Theodore. *Essays On Music*. Translated by Susan H. Gillespie. Berkeley, CA: University of California Press, 2002.

Agawu, Kofi. *Playing With Signs*. Princeton, NJ: Princeton University Press, 1991.

Anderson, Laurie. “O Superman.” *Big Science*. Warner Bros., 1982.

Andrews, Edna. *Markedness Theory*. Durham, NC: Duke University Press, 1990.

Aristotle. *On Rhetoric*. Translated by George A. Kennedy. New York: Oxford University Press, 1991.

Armstrong, Louis. “It Takes Two To Tango.” *Louis Armstrong’s All Time Greatest Hits*. MCA Records, 1994.

Barthes, Roland. *Image Music Text*. New York: Hill and Wang, 1977.

Bernstein, Leonard. *The Unanswered Question: Six Talks at Harvard*. Cambridge, MA: Harvard University Press, 1976.

Burrows, David. *Sound, Speech, and Music*. Amherst, MA: The University of Massachusetts Press, 1990.

Butler, Christopher *Postmodernism: A Very Short Introduction*. New York: Oxford University Press, 2002.

Christgau, Robert. “The Year the Rolling Stones Lost the Pennant.” *Village Voice* 27 Jan, 1982: 36-7. Quoted in Samuel Austin McBride, “Performing Laurie Anderson.” (Ph.D. diss., University of California Riverside, 1997): 204.

Cobussen, Marcel. “Deconstruction in Music.” Ph.D. diss., Erasmus University Rotterdam, 2001.

Cogan, Robert. *New Images of Musical Sound*. Cambridge, MA: Harvard University Press, 1984.

_______. *The Sounds of Song*. Cambridge, MA: Publication Contact International, 1999.

Cone, Edward. “The Authority of Music Criticism.” *Journal of the American Musicological Society* 34 (1981): 1-18.

Cusick, Suzanne. “Feminist Theory, Music Theory, and the Mind/Body Problem.” In *Music/Ideology: Resisting the Aesthetic*, edited by Adam Krims. Amsterdam: Overseas Publishers Association, 1998.

Derrida, Jacques. *Of Grammatology.* Translated by Gayatri Chakravorty Spivak. Baltimore: Johns Hopkins University Press, 1976.

_______. *Dissemination*. Translated by Barbara Johnson. Chicago: University Press, 1982.

_______. “Structure, Sign, and Play in the Discourse of the Human Sciences.” In *Criticism: Major Statements, Third Edition*, edited by Charles Kaplan and William Anderson. New York: St. Martin’s Press, 1991.

Deutsch, W. A. & F. Födermayr. “Visualization of Multi-Part Music (Acoustics and Perception).” *Arbeitsberichte der Forschungsstelle für Schallforschung der ÖAW* 1 (1995): 1-19.

Dickinson, Kay. “‘Believe’? Vocoders, Digitalised Female Identity and Camp.” *Popular Music* 20/3 (2001): 333-347.

Fales, Cornelia “The Paradox of Timbre.” *Ethnomusicology* 46 no. 1 (Winter 2002): 56-95.

Gibson, William. *Burning Chrome*. New York: Arbor House, 1986.

Harper, Andrew H. Jr. “Spectrographic Comparison of Certain Vowels to Ascertain Differences Between Solo and Choral Singing, Reinforced by Aural Comparison.” Ph.D. diss., Indiana University, 1967.

Howie, John and Pierre Delattre. “An Experimental Study of the Effect of Pitch on the Intelligibility of Vowels.” *The Bulletin of National Association of Teachers of Singing* XVIII (May, 1962): 6-9. Quoted in William A. Hunt. “Spectrographic Analysis of the Acoustical Properties of Selected Vowels in Choral Sound.” (Ph.D. diss., North Texas State University, 1970), 2.

Jakobson, Roman, C.G.M. Fant, and M. Halle. “Preliminaries to Speech Analysis.” *MIT Acoustic Laboratory Technical Report No. 13* (1952). Quoted in Andrew H. Harper Jr. “Spectrographic Comparison of Certain Vowels to Ascertain Differences Between Solo and Choral Singing, Reinforced by Aural Comparison.” (Ph.D. diss., Indiana University, 1967), 20.

Jakobson, Roman and Morris Halle. *Fundamentals of Language*, 2nd revised ed. The Hague: Mouton, 1971.

Jakobson, Roman. *On Language*. Edited by Linda R. Waugh and Monique Monville-Burston. Cambridge, MA: Harvard University Press, 1990.

Johnson, Bryce McCrary. "The Semiotics of Perception: Towards a Theory of the Minimal Units of Linguistic and Musical Sound." Ph.D. diss., Duke University, 2000.

Kittler, Friedrich A. *Gramophone, Film, Typewriter*. Translated by Geoffrey Winthrop-Young and Michael Wutz. Stanford, CA: Stanford University Press, 1999.

Jackendoff, Ray and Fred Lerdahl. *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press, 1982.

Lalor, Brendan. "The Classification of Peirce's Interpretants." *Semiotica* 114-1/2 (1997): 31-40.

Le Guin, Elisabeth. Personal communications with the author on 11/19/04 and 3/4/05.

Lemke, Jay. "Topological Semiosis and the Evolution of Meaning." Presented at the Conference of the Washington Evolutionary Systems Society, May 1998, Georgetown University.

Lidov, David. *Is Language A Music?* Bloomington, IN: Indiana University Press, 2005.

McClary, Susan. *Feminine Endings: Music, Gender, and Sexuality*. Minneapolis: University of Minnesota Press, 2002.

Merleau–Ponty, Maurice. *The Primacy of Perception and Other Essays*. Translated by J. M. Edie. Chicago, Illinois: Northwestern University Press, 1964.

_______. *Phenomenology of Perception*. Translated by Colin Smith. London: Routledge & Kegan Paul Ltd: 1981.

Monelle, Raymond. *Linguistics and Semiotics in Music*. Philadelphia: Harwood Academic, 1992.

Nattiez, Jean-Jacques. "Linguistics: A New Approach for Musical Analysis." *International Review of the Aesthetics and Sociology of Music* 4/1 (1973): 51-68.

_______. *Music and Discourse: Toward a Semiology of Music*. Translated by Carolyn Abbate. Princeton, NJ: Princeton University Press, 1990.

_______. "Le Timbre Est-Il un Paramètre Secondaire?" Presented at Conference on Interdisciplinary Musicology, March 2005, Montréal.

Nettl, Bruno. “Some Linguistic Approaches to Musical Analysis.” *Journal of the International Folk Music Council* 10 (1958), 37-41. Quoted in Raymond Monelle. *Linguistics and Semiotics in Music*. (Philadelphia: Harwood Academic, 1992): 28.

The New Grove Dictionary of Music and Musicians, 2nd ed. S.v. “Timbre.” Murray Campbell.

O’Grady, William, Michael Dobrovolsky, and Mark Aronoff. *Contemporary Linguistics*. New York: St. Martin’s Press, 1989.

Oey, Handy and Volker Mellert. “Vibration Thresholds and Equal Vibration Levels at the Human Fingertip and Palm.” Presented at The 18th International Congress on Acoustics, 2004, Kyoto, Japan.

Orcalli, Angelo. “La Ricerca Timbrica nela Semiologia della Music di Jean-Jacques Nattiez.” *Il saggiotore Musicale* v. 2, n. 2 (1995): 349-356. Translated by Marie Agatha Ozah.

Peirce, Charles Sanders. *Collected Papers of Charles Sanders Peirce*, vol. 1. Edited by Charles Hartshorne and Paul Weiss. Cambridge, MA: The Belknap Press of Harvard University Press, 1960.

_______. *Collected Papers of Charles Sanders Peirce*, vol. 8. Edited by Arthur W. Burks. Dulles, VA: Thoemmes Press, 1998.

_______. *The Essential Peirce*, vol. 2. Edited by Nathan Houser *et al*. Bloomington, IN: Indiana University Press, 1998.

Pinker, Steven. *The Language Instinct*. New York: Harper-Collins, 1994.

Rabiner, Lawrence R. and Ronald W. Schafer, *Digital Processing of Speech Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1978

Roads, Curtis. *The Computer Music Tutorial*. Cambridge, MA: The MIT Press, 1996.

Seeger, Charles. *Studies In Musicology 1935-1975*. Berkeley, CA: University of California Press, 1977.

Sextus Empiricus. *Against the Logicians*. Translated by R. G. Bury. Cambridge, MA: Harvard University Press, 1935.

Slawson, Wayne. *Sound Color*. Berkeley, CA: University of California Press, 1985.

Steeves, James B. *Imagining Bodies: Merleau-Ponty’s Philosophy of Imagination*. Pittsburgh: Duquesne University Press, 2004.

Stewart, Thomas W. Jr. and Nathan Vaillette, eds. *Language Files*. 8th ed. Columbus: The Ohio State University Press, 2001.

Taylor, Charles. *Philosophical Arguments*. Cambridge, MA: Harvard University Press, 1995.

Theleffsen Thorkild. “Firstness and Thirdness Displacement: Epistemology of Pierce's Sign Trichotomies.” *Applied Semiotics / Sémiotique Appliquée* no.10 (2000), article 2.

Titze, Ingo R. “Definitions and Nomenclature Related to Voice Quality.” In *Vocal Fold Physiology: Voice Quality Control*, edited by Osamu Fujimura and Minoru Hirano. San Diego: Singular, 1995.

Turino, Thomas. “Signs of Imagination, Identity, and Experience: A Peircean Semiotic Theory for Music.” *Ethnomusicology* 43 no. 2 (Spring/Summer 1999): 221-253.