UNIVERSITY OF CALIFORNIA
SANTA CRUZ

Wordless Syntax and Holarchic Flux in *Terrain*

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF MUSICAL ARTS

in

MUSIC COMPOSITION

by

Andrew C. Smith

June 2020

The Dissertation of Andrew C. Smith is approved:

Professor Larry Polansky, Chair

Professor David Dunn

Professor Carla Harryman

Quentin Williams
Acting Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

Abstract

Andrew C. Smith, "Wordless Syntax and Holarchic Flux in *Terrain*"

Central to the comparison between language and music is what Theodor Adorno termed a "wordless syntax," such that the elements of a work of art might be held in relation to one another, rather than fixed relative to a static hierarchical form. In music, this syntax creates a tension between the expressive, voice-like aspects of melody, and the abstract, syntactical structures of harmony. The use of linguistic material further complicates the musical syntax, as there are multiple competing hierarchical structures at play. As an alternative to the word "hierarchy" in relation to musical form, the composer James Tenney offered "holarchy" as a replacement which sets elements in relation to one another, rather than to a fixed structure. The friction and instability created by multiple syntactical systems and the material nature of the music generates what I call in this dissertation *holarchic flux*.

In addition to a theoretical overview of holarchic flux and its relation to language and music, this essay outlines a novel algorithm for interval and chord rationalization, using Tenney's metric of harmonic distance. This essay also describes a number of compositions of mine, including a solo voice work *How to Get There From Here* and ensemble compositions *We remember not the word, but the sound of the word*, and *Terrain*, with regard to the concepts of holarchic flux and musical syntax.

Acknowledgements

I would like to thank my committee, Larry Polansky, David Dunn, and Carla Harryman for their support in developing these ideas over the course of my graduate studies, and for making links between disparate disciplines, as well as Nina Treadwell for the seminars during which many of these connections to historical performance practice were made. To the musicians who patiently worked through multiple revisions of the compositions described: Séverine Ballon, Ghost Ensemble, and wasteLAnd. And to the friends and colleagues who discussed these ideas with me throughout the course of my studies, particularly Madison Heying, David Kant, Jon Myers, and Beau Sievers. Finally, to my family, for their ongoing support and encouragement.

1. Introduction: A theory of language and harmony

> Artworks move toward the idea of a language of things only by way of their own language, through the organization of their disparate elements; the more they are syntactically articulated in themselves, the more eloquent they become in all their elements.[1]

> Such time lengths—in order to be perceived as "parts"—must be *articulated* by some other means, and these means may or may not include the specifically "harmonic" devices of cadence, modulation, etc.[2]

A language—whether verbal or musical—exists simultaneously as a sensible stream and as a set of symbolic materials: phonemes, words, harmonic intervals, chords. The question of how to apprehend the material stream as symbols requires drawing boundaries, and therefore determining how those boundaries are articulated by the material and its context. In verbal language, similar sequences of material sound might be parsed into words in different ways. The phrases "to hear" and "together" both begin with "to," pronounced roughly identically, yet the syllable "to" occupies entirely different syntactical functions in both instances. In language, the segmentation of the sensible stream into discrete "parts"—to use the composer James Tenney's word from the epigraph—depends on our recognition

[1] Theodor Adorno, *Aesthetic Theory*, ed. Gretel Adorno and Rolf Tiedemann, trans. Robert Hullot-Kenter, 2nd ed. (London: Continuum, 2002), 140.
[2] James Tenney, "John Cage and the Theory of Harmony," in *From Scratch: Writings in Music Theory*, ed. Larry Polansky et al. (Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015), 287.

of the parts as words; "tohear" is not a word, and so we reject that grouping, until perhaps we hear it enough times that we alter our language to accommodate this new word.

In a work of art, the boundaries and articulations fluctuate. The "parts" created by drawing these boundaries therefore fluctuate as well. If we take Adorno's above statement that the "disparate elements" of a work of art "move toward the idea of a language of things" to mean that these elements of any work of art are analogous to words in the context of language, then the elements develop syntactical functions through repeated use inside the artwork and across multiple works of art throughout related genres and cultures. Upon continued exposure to the same parts, we come to expect a syntactical consistency. To use an example from tonal music: when a listener hears a dominant chord in a given key, they will likely expect it to return to the major or minor tonic. The sense of expectation results from the conditioning of the listener and the context of the composition, but this expectation is contingent on the ability of the listener to recognize a certain sonority as a dominant chord in the first place. The listener must hear in multiple stages: they must identify the chord as major, then must identify a particular member of the chord as the root, and finally must identify the relationship of the root to the tonic of the overall composition. Just as in language the same sound might be "heard" in many different ways, dependent on context—"there," "they're," and "their"; or "two," "too," and "to"—the syntactical function of music is dependent on both the sound itself and the context in which it is heard. Were the listener not *listening for* a harmonic cadence (i.e., not listening within the context of a tonal composition) this particular sequence of sounds would not have the same syntactical impact. In this way, the language of harmony, the parsing of its material, and the syntax of its arrangement are all contingent on the cultural presuppositions of the listening context as well as the features of the material itself.

In the absence of the culturally given markers of "cadence, modulation, etc."—what Tenney categorized as the "subjective set,"[3] following gestalt psychology—we must find new ways of describing the internal articulations of the musical work, and discussing the tensions between these material articulations and the cultural expectations. For both Adorno and Tenney, the articulation of parts within the whole occurs not just in the work itself, but in the facilities that the listener brings to the work. Tenney devoted a significant part of the later period of his life to developing a theory that might describe the articulations of harmony independent of cultural context, and yet he repeatedly frames his theoretical project in its relation to Cage as an acknowledgement and then a "renunciation"[4] of functional harmony. In a conversation after his lecture at the Darmstadt Institute, Tenney recognizes the influence of culture but in responding to a question from Gertrud Meyer-Denkmann about how his conception of form relates to Adorno's states that, "There is no harmonic function other than what we choose. […] [T]here is no

[3] James Tenney, "Meta Hodos," in *From Scratch: Writings in Music Theory*, ed. Larry Polansky et al. (Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015), 13–96.
[4] Tenney, "John Cage and the Theory of Harmony," 285.

function given. There are relationships given. There is a nature there."[5] He acknowledges the cultural history of harmony, but responds with negativity. He responds to the idea of a fixed form, with functions that are "given," with a focus on *relationships*.

Barrett Watten, a poet and critic whose work is strongly associated with the "language writing" movement, writes about negativity with regard to John Cage's solo piano work *Etudes australes*, created by combining star charts with a list of chords that were possible to play with both hands by pianist Greta Sultan, for whom the work was written.

> [T]he resulting performance is 'incompressible'—it cannot be reduced to the information value of its initial premise (stars) or the decisions (rules) by which chance was applied to them.[6]

The negativity of the work is its resistance to closure, and furthermore its resistance to clear harmonic syntax. The "incompressible" nature of Cage's composition recalls Tenney's *ergodic form*, where the statistical properties of some set of lower-level parameters are constant throughout the composition;[7] this excludes the possibility of a complete description that is not simply the work itself. Note that the word "negativity" as used by Adorno and Watten is distinct from "negation" as used by Tenney:

> Cage's "renunciations" since 1951 should therefore not be seen as "negations" at all but rather as efforts to *give up the old habits of negation*—the old *exclusions* of things

[5] James Tenney, "Darmstadt Lecture," in *From Scratch: Writings in Music Theory*, ed. Larry Polansky et al. (Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015), 362.
[6] Barrett Watten, *Questions of Poetics: Language Writing and Consequences* (Iowa City: University of Iowa Press, 2016), 182.
[7] James Tenney, "Form in Twentieth-Century Music," in *From Scratch: Writings in Music Theory*, ed. Larry Polansky et al. (Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015), 157.

> from the realm of aesthetic validity, the old *limitations* imposed on musical imagination, the old *boundaries* circumscribing the "art of music."[8]

The core belief of Tenney's theory of harmony is a description of the musical work as an autonomous object, independent of its creator, and subject to the perception of the listener. In this, all material, including that which evokes tonal harmonic syntax, is available for use. Yet, Tenney's theory of harmony is a theory of *relationships*, rather than of the *hierarchy* of tonality. This network of relationships independent of a fixed hierarchy sets the elements against one another, rather than against a "tonic" key. It theorizes what Adorno called "the development of the bindingness of [the work of art's] elements, a wordless syntax,"[9] and opens the door to a conception of harmony that resists closure. The hierarchy—or *holarchy*, as Tenney would revise his own thinking—grows from the materials themselves, and their relationships which are constantly at play. In the work of art, these holarchies resist stasis; they are in *flux*.

Syntactical articulation

Adorno describes syntactical articulation as the division of a whole into parts, and the further arrangement of those parts, connecting articulation with the "linguistic quality" of the work of art. He calls this linkage of elements a "wordless syntax,"[10] referencing the fundamental symbolic element of language—the word—and its organizing principle. A

---

[8] Tenney, "John Cage and the Theory of Harmony."
[9] Adorno, *Aesthetic Theory*, 184.
[10] Adorno, 184.

syntax requires elements to be categorized and treated in terms of classes, such that we might deal with categories of verb, noun, participle, etc., rather than with *specific* elements; these classes are called parts of speech. Adorno's term "wordless syntax" empties the specifics out of the syntactical units, dealing only with *classes* of elements rather than discrete, denotative symbols in the form of words. The work of art is linguistic in *form* due to its syntactical articulations, but that linguisticality does not extend to denotative meaning, which is analogous to *content*.

In his book *Total Syntax*, Barrett Watten initially defines syntax as "the relation of total sense to the order of elements in a language; that is, the way words make sense by means of their sequence in time."[11] That is, there is a temporal nature to syntax, but that temporality is an ordering of symbols, rather than an ordering of the properties of the materials. Yet, the articulations of the material, below the level of the word, are what fix the material into classes of symbols. We might categorize a heard sound as the spoken word "word," even as every actual instance of "word" is a different sonic experience—depending on intonation, speaker, speed, etc. Our syntax deals with the symbolic category of "word," not with the features of its material nature. Likewise, a major triad in any inversion is heard as a particular class of sonic event which, depending on the context and the listener, might have a particular syntactical implication. When we hear a major triad, the identification of

---

[11] Barrett Watten, *Total Syntax* (Carbondale and Edwardsville, IL: Southern Illinois University Press, 1985), 65.

that sound *as a member of the class of major triads* is then separate from the material nature of the sensation, and our identification of the sound in terms of its class becomes an abstraction of the sound itself. Furthermore, the sound's syntactical position informs every other instance of that class of major triads, invariant over instrumentation, dynamic level, speed, and so forth.

Form and content

Tenney and Adorno both reject the universals of "forms," while also rejecting the nominalist viewpoint of the particularity and uniqueness of specific instances of form. Adorno writes,

> The nominalistic artwork should become an artwork by being organized from below to above, not by having principles of organization foisted on it. But no artwork left blindly to itself possesses the power of organization that would set up binding boundaries for itself: Investing the work with such a power would in fact be fetishistic.[12]

Adorno observes that while there are not Platonic universal forms—that is, there is no such thing as an ideal "sonata form," independent of the works themselves—this statement does not entail that each art work is fundamentally *unique* in terms of form. Rather, form in works of art is created by corresponding relationships between the articulated parts, and these articulated parts are themselves contingent on the listener's understanding, expectations, and prior experiences with similar forms. The artwork does not exist in

[12] Adorno, *Aesthetic Theory*, 220.

isolation as an entirely self-organizing object, bound at its edges. The listener brings their cultural history to the listening experience. The listener finds articulations in the content, generating the form of the piece, based not just on the material reality of the articulations, but also on their prior experience with similar forms they expect to perceive in a given situation. Form and content therefore potentially interfere with one another, and are held in tension in the understanding of the listener.

Through this tension, the syntactical articulations of content generate form, but that form is perpetually in flux, as the articulations are subject to cultural assumptions as well as material phenomena. Tenney, in an essay entitled "Form in Twentieth-Century Music," writes that "Actually, the 'thing-in-itself' doesn't even exist in music apart from our perception of it."[13] The form-content dichotomy is one of scope and of the level of analysis: "what we do finally call 'content' is the result of 'forms' at a level below the first one we have decided to deal with formally."[14] Music's linguistic capacity is its capacity to transform potentially "meaningless" sound into symbols, through internal articulations and the usage or manipulation of the cultural presuppositions of the listener—the received language, as well as the internal language of the work. It is able to organize itself at hierarchical levels, and to create its own set of classes and of symbols that bear relation to one another. Music is not just "a language," but is rather *generative* of language in that it demands that the

---

[13] Tenney, "Form in Twentieth-Century Music," 151.
[14] Tenney, 151.

listener infer an abstract syntax from articulations that cohere into a collection of symbols. Its capacity for meaning is in its internal relationships of material, the cultural frame that surrounds its selection of symbols, and its own ability to hold these two forces in tension.

Holarchy

Throughout his early theoretical work *Meta + Hodos*, Tenney uses the word "hierarchy" to describe the ways in which the perception of groupings of sounds—what he calls *temporal gestalts*—are organized into a nested structure. In this hierarchy, the largest unit is the whole piece, and the smallest unit might be a single note, with intermediate groupings called *clangs* and *sequences*. Tenney initially describes "form" as a structure of nested hierarchies, but he later revised his own terminology to substitute "holarchy" for "hierarchy" throughout his theoretical work. In an interview with Brian Belet discussing his monumental computer-composed work *Changes: Sixty-Four Studies for Six Harps,* Tenney reassesses his vocabulary:

> Incidentally, in my writings, the word "hierarchy" is often used—I don't like it any more. I want to replace it with the word "holarchy." Because, in looking up the etymology of hierarchy, it has to do with orderings of power and value. And I don't perceive what I was calling these hierarchical formal structures as having to do with power or value. They're simply hierarchies of inclusion. So, a better term would be "holarchy," which means an organization of wholes; an organization of gestalt units.[15]

[15] Brian Belet and James Tenney, "An Interview with James Tenney," *Perspectives of New Music* 25, no. 1/2 (1987): 464.

Furthermore, Tenney's assessment of form is located entirely in the perceiving and analytical capacity of the listener. There is no "form" imposed from outside, and there is no true division of form from content, as he stated in his Darmstadt lecture:

> Well, in the most general sense I view material as form on a microstructural level. So they are basically the same thing at different hierarchical levels. So there is no dualism, there is not form and content. There is form at all levels, or you might say there is content at all levels. And one can choose to deal with them at any level.[16]

In other words, "form" is a matter of the frame of perception, and the level of analysis, and every "level" is potentially considered to be either "form" or "content," depending on the vantage point of the perceiver. It follows, then, that "the whole piece" might even be considered to be "content" at the level of the evening's performance, and that the evening's performance might be a unit of content within the whole context of society.

This revision of the 1980s carries with it an undertone of social critique and the intertwining of form and society, in the vein of John Cage, Pauline Oliveros, or Christian Wolff; the implication given by the use of the term "holarchy," and especially the charged words of "power and value," is that form and society are related. The term *holon*—from which *holarchy* is derived—originated in the work of Arthur Koestler. He describes the etymology of the word as "from the Greek *holos* = whole, with the suffix *on,* which, as in proton or neutron, suggests a particle or a part."[17] Holons are dependent on a taxonomy

[16] Tenney, "Darmstadt Lecture," 362.
[17] Arthur Koestler, *Janus: A Summing Up* (London: Hutchinson & Co., 1978), 33.

of values in which groupings higher in the chain give general direction to their subordinates, coupled with dynamic, complex interrelations between units that are simultaneously "parts" and independent "wholes," or "sub-wholes." Koestler frequently uses the metaphor of the human body, in the way that the body's organs seem to function semi-independently of the body as a whole; for example, the fact that an organ might be transplanted from a donor to another patient in distress. Koestler frequently uses the word "hierarchy" in his work, similarly to Tenney, but acknowledges that the word hierarchy "is loaded with military and ecclesiastic associations, or evokes the 'pecking hierarchy' of the barnyard, and thus conveys the impression of a rigid, authoritarian structure,"[18] and so he admits the use of *holarchy* as a replacement.

[18] Koestler, 34.

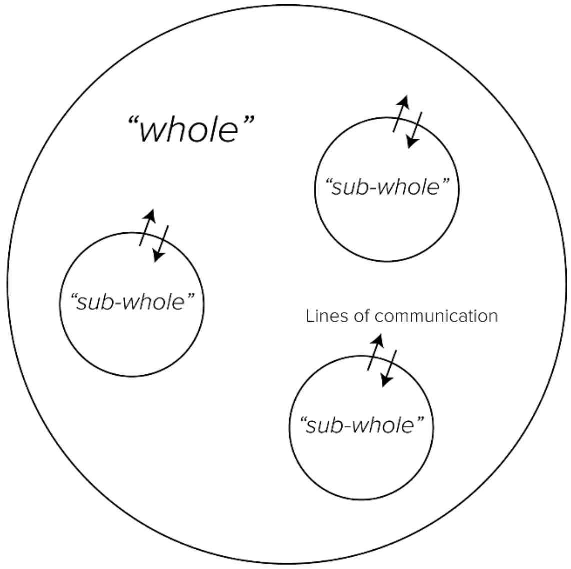

Figure 1: Diagram demonstrating holarchical organization

The term *holarchy* is further developed in the work of mathematician Ralph Abraham, who describes "*complex dynamics* or *holarchic dynamics*,"[19] in an "manifesto" appearing in the *International Synergy Journal* the same year that Tenney stated his revised thinking on the topic of hierarchy and holarchy. Abraham describes holarchical systems (and dynamical systems theory in general) as systems that evolve in time through a set of holons. These complex dynamics give rise to unpredictable behavior largely because the mechanisms of control are not top-down, but rather multilateral and coupled; the interrelated behaviors of the smaller parts affects the behavior of the larger "wholes" just as those larger units influence the behavior of the smaller units they contain. Abraham

[19] Ralph Abraham, "Mathematics and Evolution: A Manifesto," ed. David Dunn, *International Synergy Journal*, no. 3 (1986): 14.

proposed that holarchic dynamics be taken up specifically as a tool to model society, and potentially to influence the behavior of groups of individuals through the feedback mechanism of mass media. This cybernetic approach—what Abraham called "*action mechanics*,"[20] involving the mechanistic modeling of complex systems through the construction of coupled self-contained units—was an integral part of an analytical method that focused on repeated refinement and construction of models to simulate features of fundamentally unpredictable, complex behavior.

Abraham observed that dynamical systems are *complex*, and therefore difficult to predict, but they are not *random.* In his 1986 "manifesto," he wrote:

> The trajectory of an evolving dynamical system approaching a chaotic attractor is very erratic, and casual observation of such a time series may give the impression of a random, unpredictable process. The long-run behavior is neither settling to rest, nor is it approximately periodic. Nevertheless more careful observation using recently developed, computer-based graphic methods (macroscopy) reveals a highly ordered geometric pattern, a chaotic attractor.[21]

Abraham's discussion of dynamical systems is framed by his term *holarchic dynamics*, which he calls "the mathematics of metapatterns, patterns of patterns."[22] Abraham described how to determine where patterns lie—how to find an "ordered geometric pattern" in a complex system that is not even "approximately periodic"—and used the concept of holarchy to model these patterns. Fundamental to the concept of holarchy is that the constituent

---

[20] Ralph Abraham, "Mathematics and Evolution: A Proposal," ed. David Dunn, no. 5 (1987): 29.
[21] Abraham, "Mathematics and Evolution: A Manifesto," 17.
[22] Abraham, 15.

holons (or "sub-wholes") can affect one another through relationships and behaviors outside of the top-down hierarchical structure. Abraham proposed that these feedback mechanisms be used to model the possibility of nuclear winter, the depletion of the ozone layer, and global warming, taking into account not just the mathematical modeling of the system itself, but the way in which individuals and by extension society as a whole might respond.[23] Furthermore, he proposed that holarchic dynamics be used to actually *affect* society, through "social intervention" as a "low-level injection," including broadcasting "political weather reports" through mass media as a means for avoiding war.[24] Abraham's holarchical vision is radical in its break with the hierarchy of independent nation-states; his repeated calls for global projects that affect the behavior of individuals without necessitating hierarchical mechanisms of coordination imply that our ideas about form reflect our ideas about society.

The use of holon to designate both "whole" and "part" implies that while the viewer's frame of reference defines the top-level "whole," that frame of reference might be changed at any moment. If we take "form and content" as analogous to "whole and part," then we arrive at a working analogy between musical form and holarchic organization; just as there is "no dualism" between form and content in Tenney's conception of form, in holarchy there is no dualism between the whole and the parts—or, "sub-wholes." In this,

[23] Abraham, 21.
[24] Abraham, "Mathematics and Evolution: A Proposal," 39.

Abraham's use of holarchy to discuss society parallels Tenney's uncharacteristically charged revision of "hierarchy" into "holarchy," with Tenney's mention of "power and value." These discussions of holarchy and form have a counterpart at the close of Adorno's discussion of art and society:

> The liberation of form, which genuinely new art desires, holds enciphered within it above all the liberation of society, for form—the social nexus of everything particular—represents the social relation in the artwork; this is why liberated form is anathema to the status quo.[25]

Yet, what Tenney and Abraham proposed is not a new form in particular, but rather a new *conception* of form: a metaform. The concept of holarchy in music is a way of thinking about form that breaks from genre, and even from the particularity of the work, to theorize the mechanism that determines the way relationships are structured. In this way, it is a radical break in thought from the hierarchies of pre-20th century music genres such as sonata form. Tenney's theory describes the way things relate *to one another*, not to a pre-existing structure or form. This theory is valid at the level of articulation within the musical work, of larger structures of form across multi-movement works, and even when evaluating the form of works in comparison to one another.[26] It follows then that the form of relations in a holarchy is the "social relation" of the musical work.

[25] Adorno, *Aesthetic Theory*, 255.

[26] Though it is beyond the scope of the limited discussion in this essay, a detailed discussion of the topic is in Tenney, "Form in Twentieth-Century Music."

Holarchic flux

The mental model of hierarchical or holarchical thinking in linguistic syntax creates parallels between higher-level units of language—clauses and sentences—and the "holons" of holarchy. These higher-level units are situated in their context, but can transcend that context through a syllogistic movement. Poet and critic Ron Silliman's theorized a "new sentence" in San Francisco Bay Area experimental poetry, which decouples the paragraph from the unit of argument and instead transforms it into a quantity. Silliman described a precursor to the new sentence in the work of Gertrude Stein. He writes that Stein, in her work *Tender Button*, is:

> [Equating] clauses, which divide as indicated into dependent and independent, with sentences. Anything as high up the chain of language as a clause is already partially a kind of sentence. It can move syllogistically as a sentence in itself to a higher order of meaning.[27]

Silliman is describing a process in Stein's writing in which the clause breaks free of its status as a part of a larger whole: the sentence, the paragraph, and the complete text. Silliman is thinking in hierarchical terms about sentence structure in Stein's writing. A clause is "high up the chain of language," and can move to "a higher order of meaning." He implies that these articulated units exist as a set of nested entities, and furthermore that Stein is deliberately blurring these hierarchical divisions. Silliman points out that the earliest

[27] Ron Silliman, *The New Sentence* (New York: Roof Books, 2003), 86.

definition of a sentence is the material that exists between two "full stops," or periods.[28] This period mark articulates a clear boundary between two segments, over which the concept of syntax does not extend. Syntax stops at the sentence's edge; the sentence itself is a self-contained holon. Stein's work deliberately blurs this sentence boundary through undermining the stability of the syntax of the material between two periods.

Yet, this ambiguity is not just horizontal, but vertical as well. It blurs the very definition of a sentence, and even calls into question whether a sentence is actually a sentence, or whether it is merely a clause (i.e., a lower level holarchic entity) that happens to have a period at the end. Lyn Hejinian, a critic and experimental writer in the San Francisco Bay Area associated with Language writing, describes the concept of the "open text," which "rejects the authority of the writer over the reader and thus, by analogy, the authority implicit in other (social, economic, cultural) hierarchies."[29] In parallels to the cybernetic discourse of Koestler, Gregory Bateson, and others, Hejinian writes that in this juxtaposition and rejection of stable hierarchies, "the poem *is* a mind."[30] The combination of syllogistic movement between holarchic levels and the openness of the text to the reader or listener—the text's rejection of its own authority—create what might be described as *holarchic flux*.

[28] Silliman, 64.
[29] Lyn Hejinian, "The Rejection of Closure," in *A Guide to Poetics Journal: Writing in the Expanded Fields 1982-1998*, ed. Lyn Hejinian and Barrett Watten (Middletown, CT: Wesleyan University Press, 2013), 88.
[30] Hejinian, 89.

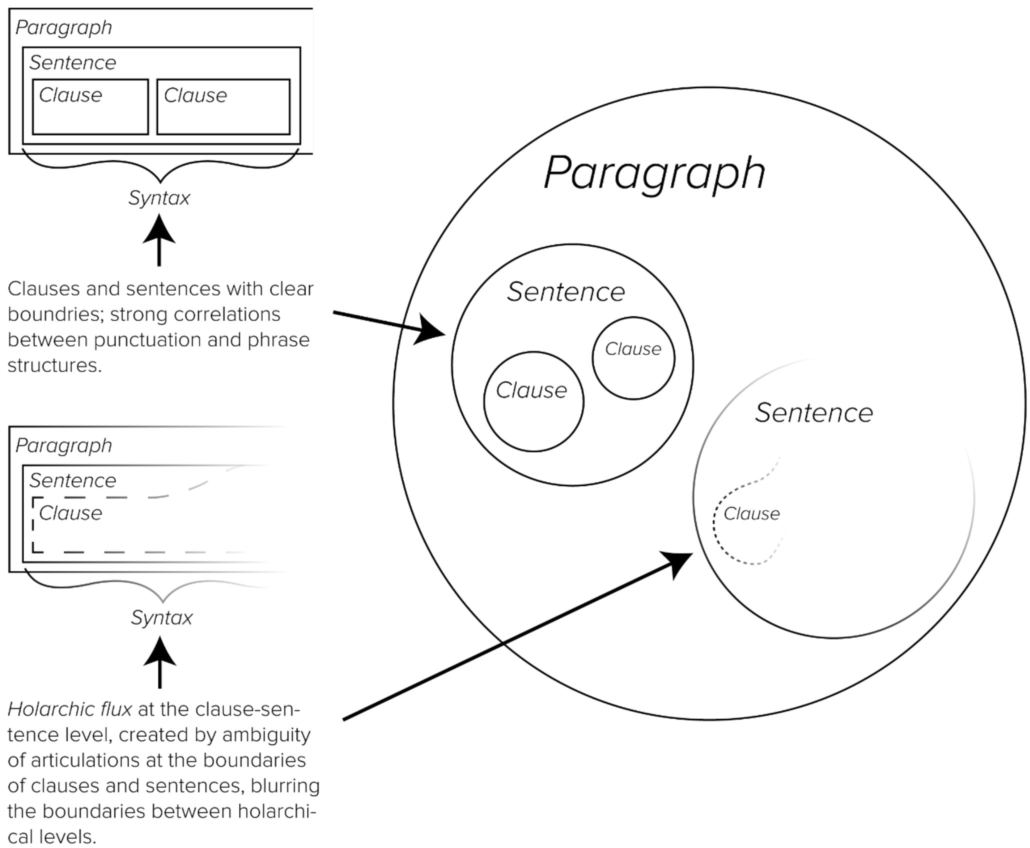

Figure 2: Diagram of holarchic flux as it relates to language

Articulation as the "linguistic quality" of works of art is intertwined with a relational way of thinking; linguistics implies syntax, which implies arrangements of form, empty of content. Adorno's description of harmony is this: a collection of symbols, related only by syntax, without the burden of denotative meaning. The linguistic quality of art is one in which symbol and syntax are paramount. Yet, *material* qualities always interfere with the stability of the symbol, and therefore of the stability of the syntax. The destabilization of syntax is a destabilization of form through the materiality of the work. If we take Adorno's analogy seriously—that of the "liberation of form" to the "liberation of society"—then the destabilization of syntax through the material of the artwork is analogous to the destabilization of society. Abraham and Koestler, in using the term "holarchy," drew clear relations between formal structures in mathematics, biology, and society; Tenney's use of holarchy in turn brought music into this relation. Furthermore, we must not overlook Tenney's discussion of the "subjective set," and the extensive dialectical discussion of musical history in his essay "John Cage and the Theory of Harmony." Tenney's technical methods are methods of *material,* and perception, but Tenney's historical thinking is one rooted in the progression of form, and the historical development of musical ideas and syntax. Tenney's invocation of holarchy indicates that there is always one more "meta" level to consider—and that level might be subject to holarchic flux. The linguistic quality of art becomes one in which every level of language is always *situated* in a context, and that context is always permeable through modifications of the material of language or music itself.

Chapter outline

This dissertation examines a number of disparate activities of musical production through the lens of language, syntax, articulation, and holarchical organization. Moreover, it examines how the material nature of the work of art defines its own articulations, particularly with regard to harmony. Throughout the dissertation, my own compositional work is discussed in an effort to bring into relief the issues and problems created by the use of linguistic material, linguistic structures, and finally harmonic syntax, in the construction of a composition.

Chapter 2 outlines a selected survey of sound poetry, text-sound compositions, and poet-musician improvisatory projects, as an introduction to a discussion of my work for solo voice, *How to Get There From Here*. I then discuss the relation of this work to the concept of "holarchic flux," and the effect of the competing syntactical structures of linguistic and musical material concerns.

Chapter 3 discusses the history of speech-music transcription projects, and examines the ways in which four composers have derived musical notation (or computer-performed actions) from language. This examination looks in particular at the intersections between writing and information. My own work, *We remember not the word, but the sound of the word*, is given as an example of this practice, and the nature of its performance practice and composition is discussed.

Chapter 4 explores the concept of "harmonic identity," through a discussion of interval tolerance and chord rationalization. A novel algorithm is introduced to rationalize arbitrary sets of frequencies using continuous-valued functions and an iterative gradient descent procedure. This algorithm is a necessary precursor to a discussion of the syntactical nature of harmony, and its analytical and generative potentials are described.

Chapter 5 describes my chamber music composition *Terrain* and its methods of construction. This construction incorporates practical uses for the chord and scale rationalization algorithms given in Chapter 4. It further discusses the concepts of holarchic organization and holarchic flux, and relates holarchy to Cage's conception of indeterminacy.

2. Sound poetry and the material of language

The 20th century saw a reoriented focus toward the material signifier in poetry. The sound poetry movement, growing from the Dadaist work of Hugo Ball, Kurt Schwitters, and others, sought to put the materiality of language at the forefront of the work of art, and to allow the arrangement of linguistic materials to be driven by the sound of the spoken word, rather than by semantic meaning. Musicologist and curator Nancy Perloff described the beginnings of the sound poetry movement as tied to the experiments in musical material in the early 20th century, situating the movement in the context of "a common interest in incorporating sounds produced by new kinds of musical instruments and by virtuosic vocal techniques."[31] One of the most well-known early sound poems, the *Ursonate* by Schwitters, adheres primarily to conventional musical forms from the prior century, with a multi-movement work including the standard musical forms of rondo, largo, scherzo-trio, and a "presto finale that includes a lively cadenza."[32]

The correspondences between writers and musicians extend more deeply than the surface-level material, however, to include formal experiments as well. Compositional constraints and indeterminacy are common techniques to both writers and musicians.

---

[31] Nancy Perloff, "Sound Poetry and the Musical Avant-Garde," in *The Sound of Poetry / The Poetry of Sound*, ed. Marjorie Perloff and Craig Dworkin (Chicago and London: University of Chicago Press, 2009), 97.

[32] Perloff, 108.

Christian Bök's *Eunoia* consists of a five chapters, each containing only words with one vowel each (the first chapter containing no vowels except "a," the second no vowels except "e," and so on).[33] While not a "process" *per se*, the mechanism of compositional constraint in Bök's work parallels constraints in works such as, e.g., Christian Wolff's *Duo for Violins*, which contains only three pitches separated by a semitone;[34] or the tone rows of Anton Webern containing all twelve chromatic pitches but focusing primarily on minor ninths and major sevenths.[35] Jackson Mac Low's poems from his "Asymmetries" series use chance operations similar to those employed by John Cage to control not just their visual representations and composition on the printed page, but also the method of reading them aloud in performance.[36] One of the clearest uses of language as musical material appears in composer Kenneth Gaburo's work *Maledetto*, which consists entirely of transformations, phonetic fragments, definitions, and scatological asides on the word "screw."[37]

Apart from compositional practices prior to performance, there are performative analogues between musical and text-based disciplines. The Four Horsemen, a Canadian

---

[33] Christian Bök, *Eunoia* (Toronto: Coach House Books, 2001). Bök's reading style is highly performative, highlighting the sonic characteristics of the work in a way that recalls Jaap Blonk's performances of Kurt Schwitters's *Ursonate*. *Christian Bök Reads from Eunoia*, accessed May 20, 2020, https://www.youtube.com/watch?v=zhQjfr8b9Wg. *Blonk Performs Ursonate with Real-Time Typography*, accessed May 20, 2020, https://www.youtube.com/watch?v=rs0yapSIRmM.

[34] Christian Wolff, *Duo for Violins* (Edition Peters, 1950).

[35] As Henri Pousseur wrote in an issue of *die Reihe* dedicated to the composer, "Webern breathes a new spirit into chromaticism […] he seems finally to have restored its primal innocence." Herbert Eimert and Karlheinz Stockhausen, eds., *Anton Webern*, Die Reihe 2 (Bryn Mawr, Penn.: Theodore Presser Co., 1958), 52.

[36] Jackson MacLow, "Methods for Reading Asymmetries," in *An Anthology of Chance Operations*, ed. La Monte Young (New York: La Monte Young & Jackson Mac Low, 1963).

[37] Kenneth Gaburo, *Maledetto* (La Jolla, Calif.: Lingua Press, 1967).

sound poetry performance group, performs in such a way that Dick Higgins notes "tropes the style of rock and roll to the point of listening to each other take riffs and solos and playing off each other as any tight rock group would."[38] The Australian group Machine for Making Sense includes writers Chris Mann and Amanda Stewart along with Rik Rue, Jim Denley, and Stevie Wishart. Amanda Stewart describes the ensemble's practice as a collaborative process:

> In our "Machine" performances, distinctions between text and speech, improvisation and composition, speech and music are blurred. No one element can ever be assumed to be accompanying another; rather we run parallel to each other, inhabiting and forming a shared space with totally different reference points.[39]

Stewart describes an aesthetic in which the material of language and the material of music are considered part of the same strata—potentially, stripping the language of its denotative component. What remains is the structure, and the parallel gestures between the two modes of sound production. More recently, Carla Harryman's *Open Box (Improvisations)* project combines reading with freely improvised music in a collaboration with saxophonist and experimental musician Jon Raskin; this is also a collaborative, improvisatory work, published both in audio (as a CD) and in book format.[40] It is clear, then, that a range of compositional techniques with regard to form and process might be applied to the

[38] Dick Higgins, *Intermedia, Fluxus, and the Something Else Press*, ed. Steve Clay and Ken Friedman (Catskill, New York: Siglio, 2018), 320.
[39] Brenda Hutchinson et al., "Contributors' Notes," *Leonardo Music Journal* 3 (1993): 76, https://doi.org/10.2307/1513274.
[40] Carla Harryman, *Open Box (Improvisations)* (New York: Belladonna, 2007).

linguistic signifier (whether written or spoken) and that various strata of meaning might emerge from the intersection of music and language.

These experiments articulate a concern for ways of generating poetic language, whether fixed as a written text or created on-the-spot in performance, that focuses on material and process rather than semantic meaning. Moreover, particularly in the chance operations of Mac Low and the improvisations of Harryman, the work is "done," in a sense, in its first iteration; neither the mode of improvisation or chance operations affords the revision of particulars, outside of revising the process as a whole. The "work itself" is the construction of a system, or the embodiment of an activity, rather than a particular instance of an objectified result. In this way, it is akin to the aesthetic concerns of computer music (or "algorithmic composition") as considered in contrast to the note-by-note composition of a piece.

The abstraction of meaning—what we might call the author's argument—is in the system or process itself, but the individual moments will inevitably create individual meanings in the understanding of each individual listener. This is in part because the cultural history of linguistic materials is so strong. We cannot avoid hearing words—at least fragments of words—when listening to a work consisting primarily of linguistic material. Articulations occur in the perception of the listener, even when words themselves are not part of the construction of the work, because the listener arrives at the work with presuppositions about where word boundaries might lie. This flux between sonic processes and word boundaries is a critical component of my own compositional work, in which the unintended semantic content of language seeps into a work that is constructed without concern for semantics.

How to Get There From Here

Throughout this section, I intend to give a technical overview of the methods of construction for each section of *How To Get There From Here*. This will be accompanied by a discussion of the ways in which the individual sections use the tension between linguistic semantics and their methods of construction to create a constantly fluctuating sense of meaning-making. This discussion intersects with an overview of the performance practice of this piece, developed through a series of composer-performer appearances from 2016-2018.

My piece *How To Get There From Here,* for solo speaking voice, is focused on the emergence of meaning that occurs when the material of language is the primary concern for its composition. This work uses "we remember not the word but the sound of the word" as a "source phrase," from which the four sections are constructed. In each section, this source phrase is transformed with regard to a particular element of "content." This procedure recalls Tenney's statement that the separation of form and content is a matter of our frame of reference.

The work has four sections, including transformations on the phoneme, letter, context, and semantics. In each section, the aim is to find a mechanism of constructing the piece of writing in which the author is not dealing with moment-to-moment details, but which still opens itself to meaning at the level of the line or word through the listener's perception. The construction of this work followed a sequence of procedures through each section, ranging from transformations on the spoken text (analyzed and recombined through computational processes), to transformations of the spelling of the words themselves, to a collection of contextual references, and finally to a traversal of semantics and concepts using the WordNet database,[41] with quotations drawn from the *Oxford English Dictionary*.

Section 1 scrambles and reconstructs a recording of Section 2, using methods of stochastic composition to reassemble the phonemes of Section 2 with regard to their sonic properties, rather than their semantic meanings. There were two major steps involved in creating this section, once the source recording was made. First, I segmented the stream of audio into a fuzzy collection of "phonemes" using an automated process drawing on the measure of information (or, entropy) at a given moment. The phoneme parser accepts a single block of audio, and draws boundaries based on the "information" occurring at those boundaries; the assumption is that a moment containing new information is likely a phoneme boundary. This process creates what I called a "dictionary" of phonemes, which

[41] Princeton University, "About WordNet" (Princeton University, 2010), https://wordnet.princeton.edu/.

are then available to be used as compositional material through the stochastic scrambling process.

The automated phoneme parser begins by calculating the MFCC vector of each frame of sound, with a bin size of 2048 samples, or roughly 50 milliseconds at a 44.1 kHz sample rate. These MFCC vectors are then clustered into 26 discrete symbols, using a Gaussian Mixture Model clustering technique with covariance normalization.[42] Next, the boundary entropy and frequency of appearance of each n-gram were calculated and stored. Finally, the symbolized representation of the stream of speech is segmented using the Voting Experts[43] algorithm. This sequence of steps results in a dictionary of segmented phoneme samples, which I transcribed and tagged manually using the International Phonetic Alphabet (IPA). Because of the nature of the automated and unsupervised process, the phoneme separation is necessarily messy; the goal of this process, however, was not to achieve exact accuracy, but rather to remove myself from the process as much as possible, relying only on the material component of the spoken text and its statistical properties to inform the composition.

[42] This procedure is implemented in the Rust Programming Language, for use in real-time synthesis as well as offline algorithmic composition. The source code for this procedure can be found at https://github.com/andrewcsmith/soundsym, current as of commit ae12bd27a909269319d324c3300c9f0690e109c5.

[43] Paul Cohen, Niall Adams, and Brent Heeringa, "Voting Experts: An Unsupervised Algorithm for Segmenting Sequences," *Intelligent Data Analysis* 11, no. 6 (December 2007): 607–25, https://doi.org/10.3233/IDA-2007-11603. My implementation of the algorithm is available at https://github.com/andrewcsmith/voting_experts, current as of commit 022b5a34f9a387f1e2f9ee530dd891c3f777caa3.

Once this dictionary is completed, these phonemes are arranged in a series of 180 lines through a stochastic procedure, such that the statistical parameters of the whole section adheres to a given parametric curve. Each phoneme was analyzed for its RMS power (a measure of the overall "loudness" of the sound) and the pitch confidence—the inverse of how "noisy" a sound is—which is calculated through autocorrelation.[44] Then, for each of the 180 lines, the mean parameters of sonic power and target pitch confidence were set to a sinusoidal curve. For each line, six phonemes were selected to approach this target power and pitch confidence, using Charles Ames's statistical feedback algorithm[45] to ensure that the statistical properties of the lines adhered to their respective targets. Lastly, the transcriptions of these phonemes were collected into a score, performed at roughly the pace of 2-3 seconds per line.

The overall sonic impression of this section is a sequence of sounds which, at first, sounds like nonsense; it is too unordered and "random" sounding to create coherent meaning. Initially, the most salient aspect is the target power parametric curve, as it is clear when the phonemes transition from quiet to loud and back to quiet. However, the pitch salience becomes evident particularly in the second half of the piece, as the soft voiced vowels of the first section transition to unvoiced consonants at the end of the composition.

---

[44] Both of these algorithms are translated into Rust from the Praat C++ source code, and are part of my own open-source vox_box library. https://github.com/andrewcsmith/vox_box.rs.

[45] Charles Ames, "Statistics and Compositional Balance," *Perspectives of New Music* 28, no. 1 (1990): 80–111, https://doi.org/10.2307/833345.

Throughout this process, it is clear that the words spoken are *not nonsense*, but rather a jumbling of some other text; the phoneme dictionary is too restricted to be entirely nonsensical, and therefore must come from a text with similarly restricted sonic properties. The question, then, is as to what semantic meaning might the listener hear *through* the nonsense, based solely on a reordering of phonemes drawn from these sonic properties.

Section 2 was constructed from the spelling of the words themselves, such that each successive line only modifies one letter, and such that each line consists entirely of words present in a spelling dictionary file.[46] Each word in the source phrase is morphed into the next ("we" into "remember" into "not," etc.) using the A* pathfinding algorithm to find a shortest path from one word to the next.[47] This A* pathfinding algorithm—frequently used in pathfinding for video games to determine the routes that computer-controlled characters travel—finds the shortest path between two points, allowing for the possibility of "unacceptable" terrain. The algorithm requires the user to define a heuristic measure of distance, for which I used the Levenshtein edit distance between the current line and the target destination line.

[46] This section used the spelling dictionary provided with Linux distributions and macOS, consisting of a newline-delimited list of words. This file is normally located in the folder /usr/share/dict/words, but any newline-delimited list of words is sufficient. This words file is frequently used in spellcheck programs on Linux.

[47] Peter E. Hart, Nils J. Nilsson, and Bertram Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," *IEEE Transactions on Systems Science and Cybernetics* 4, no. 2 (July 1968): 100–107, https://doi.org/10.1109/TSSC.1968.300136.

The sequence of word-by-word transformations is then expanded into two-word pairs ("we remember" into "not the") and so on until the entire phrase is transformed ("we remember not the word" into "but the sound of the" into "word"). Each line is performed at the same rate, such that the delivery increases in density throughout the course of the piece as the number of words in each line increases. The repetition of this section, though highly rhythmic, retains some kind of semantic impression at moments—such as "not the / not tue / not tut / not but," recalling the repetitious nature of Zaum poetry[48]—and these repetitive structures slip in and out of sense-making.

The section highlights the fracture between the material signifier as it exists on the page and as it is spoken aloud. Take for example the following sequence, from the middle of the composition:

> the word but
> the word out
> the word oft
> the word of
> the wold of
> the would of
> the wound of
> the sound of
> the soud of
> the soud f [49]

[48] Perloff, "Sound Poetry and the Musical Avant-Garde," 100.
[49] "Section II: The Word," from Andrew C. Smith, *How To Get There From Here* (Santa Cruz, Calif.: Indexical, 2016), ll. 99-108.

In this sequence of lines, the visual effect is clear, as each line is graphically a slight alteration of the line above. However, the sonic effect is constantly in flux, as in the pair of lines, "the wound of / the sound of," which shifts the pronunciation of the written pair of letters "ou" as a byproduct of changing a completely different letter in the word.

This section moves *through* segments of text that make sense, semantically speaking, as a list of words preceded by "the word." Yet, the construction of the piece has nothing to do with reciting a list of words; it might recall formal methods of composition in the mind of the listener, but the formal methods of list-making have little to do with its own construction. The section also contains common words, such as "word," and "would," but moves through them with obscure connecting words like "wold," referring to the rural English countryside, which perhaps sounds like nonsense to many modern listeners. As such, the listener is continually oscillating between listening to the "sound itself," and attempting to make sense of the sounds through some kind of semantic meaning.

Section 3 consists of contextual quotations, where each word in the source phrase begins a line, followed by some number of intervening words found in a corpus of English literature including Chaucer, Shakespeare, Austen, Melville, Joyce, and a number of other authors. These lines gradually expand, with additional words on each intervening line, until the length of 180 lines is reached. When a potential match is not found in the entire corpus of texts selected, the line remains blank. Similar to Section 2, each line is performed at the same rate such that the delivery speed approaches incomprehensibility due to the density of words and the jarring effect of changing phrase structure mid-sentence.

Section 4, the final section in the piece, is generated by exploring the WordNet database,[50] which organizes words into terms that are more general ("hypernyms") and less general ("hyponyms"). These relations take the form of a hierarchical structure, and the section as a whole moves through a series of hypernyms, and again through hyponyms, throughout its trajectory. This is a kind of "pathfinding" in semantic space—rather than the physical spelling-space of Section 2—in its transition from "memory" to "word," intersecting at their common hypernym "abstraction." This section consists entirely of quotations given in the *Oxford English Dictionary* for each word along this path, selected and formatted manually into a sequence of 180 lines.

The effect of this *semantic* transformation, particularly after a series of purely material transformations on the text itself, is an imposition of a form of meaning-making that has not been in place earlier in the piece. Furthermore, semantic parsing while listening to live musical performance is an infrequent mode of interpretation in musical settings; generally speaking, musicians listen for sonic articulations, form, and structure, and perhaps occasionally for mimetic content, but denotative meaning itself is generally not present.

---

[50] Princeton University, "About WordNet."

Holarchic flux in practice

How To Get There From Here exhibits holarchic flux in the way that the ambiguity of the articulations creates further uncertainty about the hierarchical level of any given segment. Recalling Tenney's statement that "there is not form and content," but rather that the distinction is a matter of the analytical frame, I would argue that a work that exhibits holarchic flux blurs the boundaries between holarchical levels themselves. In *How To Get There From Here,* this is demonstrated most clearly by the mid-sentence pivots of "Section III: The Context." These pivots become more confounding as the sentences grow longer, and the grammar becomes more fluid—the longer lines allow more time to develop a thought, and for that thought to begin to make sense, before the context of the sentence is interrupted. This blurs the line between clause, sentence, and paragraph.

> not quite like
> the modern slang
> word to none
> but also of
> the trumpet shall
> sound of chariots
> of life of
> the most immodest
> word
>
> […]
>
> sound of the peal of the hour of the night by the chime
> of the world let us have a bit of fun first God help
> the name of those exercises he bought me one of those new some

word [51]

The use of a "pivot word" between one line and the next results in a stream of language where the boundaries between lines and phrases are blurred; the line-by-line articulation is in flux, and the parsing of individual "sentences" (or sentence-like units) is ambiguous. While I have not added periods or any other performative markings to the text (aside from empty lines, where no contextual match could be found) there are plenty of moments of partial sentences recalling Gertrude Stein's description of roast potatoes in *Tender Buttons*—"Roast potatoes for."—such that, as the writer and critic Carl Peters interprets in his reading of the text, only one half of a clearly relational whole is given.[52] This interruption forms the basis of the ambiguity generated by holarchic flux. In Figure 3, the highlighted yellow words on the left are spoken at regular intervals at approximately a tempo of 2 ½ seconds per line, creating a form of musical structure and articulation. However, the syntax of the clauses bleeds over from one line to the next, demonstrated with the greyed-out text and boxed boundaries. This disjunction between musical-holarchical organization and linguistic-holarchical organization generates holarchic flux.

[51] "Section III: The Context," from Smith, *How To Get There From Here*, lines 25-33; 140-143.
[52] Carl Peters, *Studies in Description* (Vancouver, BC: Talonbooks, 2016), 234.

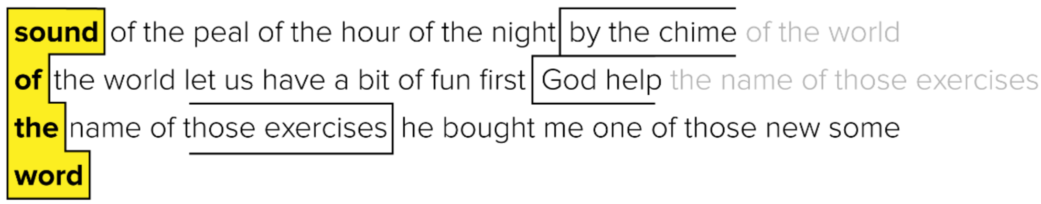

Figure 3: A diagram of lines 25-33 of "Section III: The Context"
from *How to Get There From Here*

Holarchic flux plays a crucial role in the suspension of meaning in the work, even beyond the clear sentence-level ambiguities in Section III. In Section I, there is very little indication of where boundaries begin and end; there is no syntax or "sentence" to speak of, yet the process of segmentation is still held in suspension for the listener. It is not merely a mass of sound. Below is the phonetic transcription used for a score in performance:

o o ɛɹd u bə bo
θə wu wu oe ɛɹ o
oɹ wu wa θə o θəs
o ɹe wuɹɛmɪm oe wa æ[53]

This segment can be translated to the following text, mapping these phonemes to where they likely occurred in the source recording of Section II, with letters that would alter the pronunciation in square brackets:

o o [w]ord u bu[t] bo[rn]
the wu wu oe [w]or[d] o
or wu wo[d] the o thus

[53] Smith, *How To Get There From Here*, "Section I: The Phoneme," lines 122-125.

o re we remem[ber] oe wo[d] ae

It is clear to see, once translated into a more familiar form of writing, that the inventory of much of this first section draws on a very restricted pool of language. Furthermore, the bracketed (unheard) segments at moments such as "we remem[ber]" are conspicuous in their absence in performance; they are an interruption at the level of the phoneme/word holarchy, rather than the clause/sentence holarchy. This holarchic flux only exists because of the preexisting expectations of the listener, through which the listener hears glimpses of phonetic material that is not "just sound," but that reminds them that the sound exists in the context of a particular language. This is less of a blurring of temporal gestalts and hierarchical levels than an *overriding*, or an *interfering with* the expectations of the listener with regard to syntax and articulation. The use of linguistic material makes this interference accessible and perhaps more easily perceived than musical material.

Development of performance practice

Throughout the past four years, I have performed this piece over a dozen times in various settings, primarily in "new music" venues or festivals. The piece existed first as a computer-generated text, such that the performance direction came *after* the text. In determining how to perform this work, I was interested in how the musical factors of density, amplitude, pacing, and so forth could affect the sense of the piece as something in-between speech and music. This is tied to the articulation and the boundaries of phrases, in that the articulations given by semantic meaning would occasionally be at odds with those given by the musical parameters of the performance. A few aspects of performance practice came out of this that I consider to be important to the piece's interpretation and eventual effect.

First, the line-by-line pacing is critical to maintain, but the placement of words within those lines does not necessarily need to be "rhythmic." That is, the first word or phoneme of a line should occur at a regular rate, while the rest of the line can more or less freely float in between those line beginnings. This technique is informed by the operas and work of composer Robert Ashley, particularly the published libretto-score of *Improvement (Don Leaves Linda)* in which each line is performed in three "beats," marked by accented syllables, while the intermediate words are freely placed.[54] There is a parallel here, as well, to the *sprezzatura* of the early opera of Jacopo Peri, who calls for *unwritten* rhythmic variation, described as, "Subtle and often spontaneous modifications of tone—faster or slower—which reflect the conceits of the soul."[55] In my work, this rhythmic freedom suspends the performance between music and speech, such that the delivery is not strictly metrical, but is also distinct from the vocal delivery of most in-person readings of either poetry or prose with its approach to the metrical regularity of the lines. This partial freedom of rhythmic delivery goes hand-in-hand with my approach to vocal intonation, in that the

[54] John Cage et al., *The Guests Go in to Supper* (Oakland, Calif: Burning Books, 1986).
[55] Barbara Russano Hanning, *Of Poetry and Music's Power: Humanism and the Creation of Opera*, Studies in Musicology ; No. 13 (Ann Arbor, Mich: UMI Research Press, 1980), 78.

rhythmic pauses that indicate the end of a "sentence" are frequently accompanied by a descent in vocal pitch.

Second, I prefer to perform this piece in a relatively small chamber music setting using a condenser microphone at a very high volume. This allows for microscopic vocal detail in the first movement in particular, and allows the unvoiced phonemes to be better heard. At the Center for New Music in San Francisco, this amplification led to the occasional near-feedback at the resonant frequency of the room itself, which I then found myself using as a central pitch for my vocal delivery. The piece then takes on a heightened musical quality, informed naturally by Alvin Lucier's *I Am Sitting in a Room*—a comparison which is particularly relevant given my occasionally present stutter, a holdover from a childhood diagnosis of Tourette's Syndrome—and the intimacy created by the amplification and setting connects the presentation of the work to that of the earliest operas. The musicologist H. Wiley Hitchcock, in his introduction to Giulio Caccini's vocal compendium *Le Nuove Musiche*, writes that:

> When in 1637 the Medici court chronicler for the wedding festivities of Ferdinando II spoke of the early days of the "new music," he made an interesting comment: in that time, he wrote, when Giulio Romano had invented *musica recitativa* … "from the beginning they chose the smaller rooms (for its performance) since they considered that in the large ones one could not enjoy the sweetness of the style."[56]

[56] Giulio Caccini and H. Wiley Hitchcock, *Le Nuove Musiche*, 2nd edition, Recent Researches in the Music of the Baroque Era 9 (Middleton, Wis.: A-R Editions, 2009), 10.

The intimacy of the performance situation allows a level of detailed perception of speech that does not exist in a larger venue.

3. Speech-music transcription

A process used frequently in my compositions is the "transcription" of spoken language into notation to be performed by acoustic instruments. More accurately, we might describe it as a translation or, following the composer and theorist Clarence Barlow, a *derivation* of music from language. Throughout this section, I aim to give an overview of concerns that arise when deriving acoustic music from linguistic material, and then discuss my own work in the context of others' engagement with the process of speech-music transcription processes. Music based on speech must grapple with two competing concerns that I will call "modes of understanding."

Continuous understandings of sounds are those such as a change in the inflection (pitch) of a sound, or how we might hear the vowel sound /i/ as "brighter" than the vowel sound /u/. These could be considered "parametric," following language used by James Tenney, in the sense that they exist as continuous values, and can be operated on using quantitative numerical methods.

Categorical understandings of sound are there where specific *qualities* are evoked that enable the definition of the sound as a member of a set. This includes specific vowel sounds, or the harmonic relationships of frequency ratios—the perfect fifth, major third, octave, etc. These are frequently defined by regions with boundaries in some parametric space, but with allowances for contextual information that complicate a purely parametric categorization.

These two modes can potentially work in conjunction with one another; the categorical mode serves as "signposts" for our conception of the continuous parametric alterations. In the context of musical composition with language-derived materials, composer David Evan Jones contrasts echoic memory, which "retains a *recording* of a sound," with phonetic memory, which "retains a coded *representation* of a sound."[57] Jones cites the timbral structures of vowels as common categories that might be employed in spectrally-informed electro-acoustic composition, in that just as we can identify a perfect fifth as a harmonic category, we can also identify the vowel sound /u/ as a timbral category. Yet, using timbral structures informed by vowels runs the risk of creating a sequence of sounds with *linguistic* meaning, blurring the line between the semantics of language and the syntactical structures of music. The potential ambiguities are compounded when dealing with human language in music, because of the potential for cross-domain interference between musical and linguistic interpretations of what is heard. As linguistically-derived sounds in music approach linguistic meaning, the listener is constantly evaluating both the categorical and continuous aspects of their perception, and these modes frequently come into conflict with one another.

What we are forced to do as listeners is to continuously oscillate between these two modes of understanding, altering every categorical signpost with a continuous adjustment.

---

[57] David Evan Jones, "Compositional Control of Phonetic/Nonphonetic Perception," *Perspectives of New Music* 25, no. 1/2 (1987): 143.

The flux that exists between these two modes is productive in the course of the composition; our reference categories change how we "hear" a sound, because they change what we are listening for, and what we are listening in relation to. Furthermore, the syntactical component of language—and therefore its position within the holarchical structure of syntax—operates on the categorical mode of understanding. Blurring the lines between categories through continuous parametric alterations therefore creates ambiguity in the syntactical units at any given level; this is a form of flux.

Information and writing

The difference between communicating in a musical and linguistic setting can be attributed, at least in part, to the difference in where the *information* has historically been contained. The term "information" is used here in the lineage of information theory and cybernetics described by Norbert Wiener and Claude Shannon as a measure of entropy.[58] This definition of information is related to the categorical mode of understanding in that it requires a set of symbolic terms (not parametric values) to perform the calculation. In the speech of non-tonal languages, these categories are primarily located in spectral, or timbral, features carried by the changes in vocal resonances and voiced/unvoiced patterns.

---

[58] Norbert Wiener, *Cybernetics; or, Control and Communication in the Animal and the Machine.*, Second edition. (Cambridge, Mass: MITPress, 1961); Claude Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, no. 27 (October 1948): 379–423.

Cyberneticist Colin Cherry and linguist Roman Jakobson formalized this criterion for evaluating information in spoken language, writing that,

> We must determine the minimum set of such features [binary oppositions such as voiced/unvoiced, etc.] that the listener needs in order to recognize and distinguish all except homonymic morphemes [two identical sequences of letters that are pronounced differently], without help from context or situation.[59]

The inventory of binary pairs for distinguishing Russian phonemes by their sonic features was tied to the ability of the listener to translate the spoken language of Russian into writing. As Cherry and Jakobson presupposed, information and writing have a deep correspondence with one another. Phonetic writing is symbolic, in the sense that it can clearly be broken into morphemes, words, and so on; this hierarchy of symbolic classes enables an analysis of the information content of the writing. The authors did not consider change in pitch to be a carrier of information; pitch is irrelevant, as far as writing is concerned.

The domain of music is no exception to this correspondence between writing and information, but the carrier of information is shifted from timbre to pitch. Pitch and rhythmic relationships have, until the 20th century, been the primary features of written Western music—even dynamics, tempo, and to some extent instrumentation had been left off the page before. This is perhaps why Cage opened his essay "Indeterminacy" by

[59] E. Colin Cherry, Morris Halle, and Roman Jakobson, "Toward the Logical Description of Languages in Their Phonemic Aspect," *Language* 29, no. 1 (1953): 34, https://doi.org/10.2307/410451.

connecting *The Art of the Fugue* with Stockhausen's *Klavierstuck XI,* as two compositions which are "indeterminate with respect to [their] performance,"[60] in different ways. While Bach's hierarchical pitch relationships over long periods of time were important, Stockhausen placed emphasis on the specific dynamics and timbre of the sound. As the nature of notation and writing changed in the 20th century, so did the locus of information. In particular, music places great importance on the *harmonic* component of pitch relationships which are invariant to absolute pitch-height. This harmonic component relates to the categorical mode of understanding and is frequently referred to as a "language," perhaps most famously in the title of Olivier Messiaen's treatise,[61] but more elaborately by Jackendoff and Lerdahl in their study of tonal music in its relation to linguistic structures.[62] When an artist breaks from the strict conveyance of "information"—that is, breaks from the written notation—that act is tied to *expression* and presence; this is contrary to information. Critic and writer Nathaniel Mackey calls attention to Amiri Baraka's association of articulating *imprecise* categories of pitch, including quarter-tones and glissando attacks, with personal liberatory expression.[63] The composer and instrument-builder Harry Partch bemoans the restrictions of categories when transcribing a reading by

---

[60] John Cage, *Silence* (Middletown, CT: Wesleyan University Press, 2011), 35.
[61] Olivier Messiaen, *The Technique of My Musical Language. 1st Volume, Text*, Bibliotheque Leduc (Paris: Alphonse Leduc, 1944).
[62] Fred Lerdahl, *A Generative Theory of Tonal Music*, MIT Press Series on Cognitive Theory and Mental Representation (Cambridge, Mass: MIT Press, 1983).
[63] Nathaniel Mackey, "The Changing Same: Black Music in the Poetry of Amiri Baraka," in *Discrepant Engagement* (Cambridge: Cambridge University Press, 1993), 22–48.

W.B. Yeats: "I made diagrams of his inflections, but my memory of his vibrant tones is more accurate than my marks."[64] But Partch's situatedness in the tradition of Western music held; he famously devised a system of writing and an instrumentarium to realize this writing, to further discretize these "vibrant tones" into ever-more-specific categories.

The presuppositions of the cyberneticists—that writing is a transcription of speech—were critiqued by the philosopher Jacques Derrida, in that they belied a "phonocentrism" that was inextricable from the metaphysics of presence and theology of *parousia*. Derrida's critique is in essence an argument that cybernetics should go further in redefining the "message," to cease evaluating information in terms of the relationship between writing and speech, and to instead break with the metaphysics of phonetic writing that center the *logos*.

> But beyond theoretical mathematics, the development of the *practical methods* of information retrieval extends the possibilities of the 'message' vastly, to the point where it is no longer the 'written' translation of a language, the transporting of a signified which could remain spoken in its integrity.[65]

Derrida recognized that cybernetics might extend its conception of information retrieval *beyond* the correspondence of spoken and written language, into a message that is independent of the historical dependencies of one on the other; it potentially opens the door to a writing that exists independent of speech, and therefore to a new metaphysics

[64] Harry Partch, *Bitter Music: Collected Journals, Essays, Introductions, and Librettos*, Music in American Life (Urbana: University of Illinois Press, 1991), 167.
[65] Jacques Derrida, *Of Grammatology*, trans. Gayatri Chakravorty Spivak (Baltimore: Johns Hopkins University Press, 1997), 10. Emphasis original.

and theology. Derrida continues his critique of the phonocentrism of writing through a discussion of "accentual marks" in written French, quoting from Jean Jacques Rousseau's *Essay on the Origin of Languages* to demonstrate the correspondence between accentual marks and presence:

> When the subject is no longer there, force, intonation, and accent are lost in the concept. Then one writes, one "substitutes" in vain "accentual marks" for "accent," one bows to the generality of the law: "In writing, one is forced to use all the words according to their conventional voice, determining them as one pleases. Being less constrained to clarity, one can be more forceful. And it is not possible for language that is written to retain its vitality as long as one that is only spoken" [*Essay,* pp. 21-22].
>
> Thus writing is always atonal. The place of the subject is there taken by another, it is concealed. The spoken sentence, which is valuable only once and remains "proper only to the place where it is," loses its place and its proper meaning as soon as it is written down.[66]

Just as we locate information in writing, the "accentual marks" of writing that Rousseau calls a "substitute" necessarily translate expression previously located in the spoken word onto certain *parametric features* of the word's delivery—the intonation, prosody, etc.—which then become symbolic carriers of information and those their expression. The "accentual marks" transform what was previously a continuous and parametric feature into a set of discrete categories. In Derrida's critique, expression marks the presence of a speaker, and writing therefore signifies absence, or "death." The "place and proper meaning" that is lost in the act of transcription is the meaning of *expression,* not of information, and is

[66] Derrida, 315. The citation is original, and is a quotation of Rousseau's *Essay on the Origin of Languages,* as cited in Spivak's translation.

tied to the presence of the speaker. The act of transforming continuous values, or specific inflections, into discrete categories of symbols is the death signified by writing.

In a musical context, we frequently call expression "interpretation," such that a performer would "interpret" *The Art of the Fugue* by adding dynamics and specifying instrumentation. Yet, there is no interpretive room left in a performance of *Klavierstuck XI* for modifying the precisely written dynamics. Adorno notes that "To interpret language means to understand language; to interpret music means to make music."[67] In both senses of "interpret," there is a translation *from writing* into something that recalls a present speaker. The "understanding" of language is roughly the reanimation of the author's "meaning"—or recalling the *presence* of the author—while the "making" of music is the bringing-to-life, or more specifically *making-present*, the intentions of the composer mixed with the "expression" of the performer. It is in the sensible activity within which we locate presence, yet it is in the notation—the categories, the symbols—that we locate information.

I discuss these concerns of writing, information, and speech, because the question of speech-music transcription is necessarily a question of writing, and therefore of determining the locus of information when creating a symbolic representation of a continuous phenomenon. Systems of writing reflect features of perception, and moreover

[67] Theodor W. Adorno, "Music, Language, and Composition," trans. Susan Gillespie, *The Musical Quarterly* 77, no. 3 (1993): 403, https://doi.org/10.1093/mq/77.3.401.

of *intention,* and an analysis of these systems of writing might reveal the priorities of their creators. Music derived from speech necessarily must grapple with writing—and, furthermore, the structural aspects of language itself—because it is a translation of speech through writing into sound. Because speech contains linguistic meaning, composers must decide what it is they will *do* with that meaning and with the voice of the source speaker. Composers engaging in speech-music derivation must decide where to locate *information as writing*, and their choices of writing reflect a philosophy of language—even an implicit philosophy. Is speech pure sound, and must it be grappled with as sound, or can the analysis take place through writing? What philosophies of language and information might we unravel from a study of attempts at speech-music transcription?

Overview of historical antecedents

Four compositions provide direct precedent for the boundaries of my compositional project, and are representative of a range of techniques for speech-music transcription. Composed from 1979 to 2008 by James Tenney, Clarence Barlow, Jonathan Harvey, and Peter Ablinger, these works put into practice distinctive methods for speech-music transcription that reflect not only technological concerns, but also aesthetic concerns of the composers themselves, including the contexts in which their works operate. In addition, each of these exhibit some degree of *construction*, in that their processes are not entirely controlled on a micro-level by the composer. While some of the works use models of the human vocal tract to inform their compositional techniques, others treat acoustic speech as "pure sound," no different from if they were executing a transcription of city traffic,[68] or analyzing a low tone on a trombone as in Gerard Grisey's *Partiels*. Speech fundamentally conveys information through categories of timbre, and so all of these works can be said to be "spectral" in the sense that they involve at least considering the spectrum of a recorded sound, often (but not always) generated through a Fast Fourier Transform analysis in the process of composition.[69] This criterion for "spectral music" as based on concepts and technical features is drawn from scholar and composer Robert Wannamaker, whose work on James Tenney's spectral methods has proven invaluable.[70] Furthermore, these works bring to bear different attitudes about the question of "intelligibility" in speech-music transcriptions, and other cross-domain concerns of working with speech that do not necessarily appear when working with non-linguistic sound sources.

James Tenney, *Three Indigenous Songs* (1979)

Tenney's *Three Indigenous Songs* (1979), composed without the use of a computer, but rather through transcription of vowels by ear, places instrumental sounds within the

---

[68] G. Douglas Barrett's work "Sixteen Hollywood Silence" is a prime example of the technique of spectral analysis and orchestration of a cityscape. https://gdouglasbarrett.com/sixteen_hollywood_silence.html, Accessed Feb. 8, 2020.

[69] I avoid a capital-letter "Spectral Music" label, in that spectral music is "a question of attitude," according to composer Tristan Murail, considered one of the primary exponents of the methods of spectral music. Joshua Fineberg, "Spectral Music," *Contemporary Music Review* 19, no. 2 (March 2000): 1, https://doi.org/10.1080/07494460000640221.

[70] Robert A. Wannamaker, "The Spectral Music of James Tenney," *Contemporary Music Review* 27, no. 1 (February 2008): 91–130, https://doi.org/10.1080/07494460701671558.

various formant regions of the sung words he transcribed manually.[71] In the first of these, "No More Good Water," on a song by Jaybird Coleman, Tenney alternates manually composed sections derived by ear from the harmonica interludes with homophonic orchestrations of the voice as sung by Coleman for each stanza. In the orchestrated voice sections, each vertical sonority is drawn from a subset of the harmonic series on the fundamental frequency of the F0 pitch of the bassoon or tuba (the lowest instrument in the ensemble). According to Wannamaker, Tenney's process of transcription involved consulting frequency charts of the fundamental formants for each vowel sound from the acoustical literature, determining the range of possible center frequencies, and assigning a harmonic of the F0 pitch to each instrument based on the formant height.[72] Later compositions by Tenney employed computers, as he refined this technique throughout the 1990s.[73]

Wannamaker writes that the speech-derived works by Tenney "decisively stake out for music a perceptually and conceptually engaging domain of sonic organization not previously regarded as proper to it[.]"[74] This sense of sonic organization, however, indicates a concern for speech as *structure,* rather than as sound or tool of communication. That is, while a performance of the piece is not intelligible as speech, aspects of repetition in the

71 Wannamaker, 109.
72 Wannamaker, 110.
73 Wannamaker, 111.
74 Wannamaker, 111.

vowel structure of the source material are translated to repeated events in the resulting orchestration. Larry Polansky, in program notes for a performance of the piece, calls attention to this as a "composition that is solely determined by the microstructure of speech acoustics,"[75] yet this determination is also necessarily a result of the decoupling of pitch and timbre. Repeated sequences of vowels will likely be perceived as discrete temporal gestalt units, to use Tenney's terminology, due to the similar parametric morphologies of their formant structures.

Tenney's subjectivity vanishes in the construction of his composition, as subjectivity in his work is a construction of process, rather than of particulars. Through the process of construction, Tenney articulates what spoken language is, and more importantly *how it means.* It is an articulation of the mechanism, rather than of the particulars. His title indicates that language is indigenous to human experience, as is song, and the repetition of perceptually salient features of speech (formants, pitch-height) creates a sequence of symbols through repeated structures of perceivable parameters. Tenney's argument, therefore, is that particular sonic sequences are perceptual constants, and will be perceived as units of information even as their relationship to the source material is obscured and unintelligible. Using Tenney's framework, if one were to map the temporal-gestalt units in both the source material (Jaybird Coleman's performance) and the resulting

[75] Larry Polansky, "Program Notes for Three Indigneous Songs," *San Francisco Symphony*, November 1984, http://eamusic.dartmouth.edu/~larry/misc_writings/out_of_print/tenney_symphony.pdf.

chamber ensemble performance, one should find the same temporal-gestalt units in the same locations in time. In other words, Tenney's work is reflexive toward its material, and its subjective argument is about its material, and therefore about the human perception of the song and speech of its material.

Tenney's methodology places the translation of phonemes from the sonic source to a written transcription as an intermediary step between the source and the resulting notation. This transcription step reflects a categorical understanding of vowel sounds, and it is that sequence of categories—the "microstructure of speech acoustics"—that is then reflected in the final composition. The same vowel, on the same pitch, should have the same harmony. But the concern is not only with each phoneme as a discrete symbol, because Tenney's method does not necessarily produce the same rational harmonic relationships for the same vowel sound. The formant frequencies F1-3 of a vowel would be in slightly different positions depending on the F0 frequency (though at approximately the same pitch-height) due to differences in the harmonic series for each F0 that would cause the selection of different pitches. The vertical harmony in a Western sense (even disregarding pitch height, and focusing only on pitch class) is therefore *not* the category of information that is conveyed in the transcription process, but is resultant from the spectral transcription. A given phoneme might create any number of different harmonies, depending on the F0 of the melody. This is most clearly given in the word "fishin'," in m. 29, in which the dominant seventh harmony over C moves to an open ninth chord (with no third) over D. In instances like this, the received "syntax" of harmony is deliberately eschewed in favor of the syntax of language in the formant analysis of the phonemes. Yet, the music is not exactly "atonal," in that the F0 of tuba/bassoon is essentially performing a blues tune. Its similarity to the sonority of tonal composition makes these particular moments meaningful, as the quirky chords (C7 to D9) are recognizable as analogous to *words* even as they make no sense as a chord progression in the received syntax of tonality.

Clarence Barlow, *Im Januar am Nil* (1984)

Clarence Barlow's *Im Januar am Nil* (1984) is a transcription for strings and winds of meaningless, but syntactically coherent, German phrases consisting entirely of vowels and voiced consonants. Barlow's technique—which he calls "synthrumentation," combining synthesis and instrumentation—aims to employ acoustic instruments as building blocks for additive synthesis. In an analysis of *Im Januar am Nil,* critic and composer Tom Pollar notes that Barlow's technique "has the potential to be recognizable as speech,"[76] though the difference between recognizability and intelligibility should be kept in mind. Unlike Tenney's work, in which the F0 pitch of the instrumental ensemble corresponded directly to the F0 of the source material, in Barlow's work the speech content only serves to inform which harmonics from a harmonic series are chosen—the timbral characteristics of the ensemble, constructed through additive synthesis—while the melody and rhythm of the

[76] Tom Rojo Poller, "Clarence Barlow's Technique of 'Synthrumentation' and Its Use in Im Januar Am Nil," *Tempo* 69, no. 271 (2015): 7–23.

composition are determined by non-speech-derived means. This generation of source material essentially is a constraint imposed by the subsequent process and treatment of the material; in a footnote, Poller calls attention to the similar constraint-based writing practiced by Oulipo authors such as Georges Perec.[77] While Tenney conceptually applies the source-filter model of the vocal tract, Barlow takes this a step further by modulating the source fundamental independently of the filters, through a "spiral melody" pattern. Barlow's decoupling of the single sound into independent parts likewise separates the expressive aspect of the voice (the vocal intonation) from the "linguistic" aspect of the voice—that is, the formant structure, which is the carrier of information, and furthermore which *identifies* the sound as a vocal sound.[78] As Poller writes, "for [Barlow's] works in general, the process of synthrumentation is never employed in its pure form […] rather, it is only just one aspect of an explicitly musical context[.]"[79] In decoupling the source (F0) from the filter, Barlow is treating the voice not as pure sound, but as an object with dual functions: a fundamental pitch, which he ties to the harmonic structure of an algorithmic process, and a set of frequency resonances which are the primary carriers of linguistic information.

---

77 Poller, 12.

78 For this, I refer to David Evan Jones's notion of vowel sounds as immediately salient and memorable; the "hearing" of a vowel timbre determines whether or not a sound is recognizable as speech even if, as I have discussed, it is not intelligible.

79 Poller, "Clarence Barlow's Technique of 'Synthrumentation' and Its Use in Im Januar Am Nil," 22.

Barlow and Tenney both model the human voice by decomposing the voice into an F0 fundamental and a series of formant filters. In this model, each instrument is assigned a formant; the combination of instruments is not necessarily taken into account, nor is there any acoustic or timbral information about the instruments themselves used in the model. This method used by Barlow and Tenney requires a vocal source; it is not a generically "spectral" approach, but rather one that relies on a voice-specific model of an impulse and filters. The symbolic model is necessary for Barlow's process, as his composition *decouples* the F0 fundamental of the voice from the formant frequencies.

In the last half of *Im Januar am Nil,* the style abruptly shifts to a piece *Six Quatrains,* "in which 600 years of music history continuously roll back in less than a minute (25 years in each of 24 bars),"[80] beginning with explicitly atonal music, and ending with organum. This insertion is clearly an intervention into the *understanding* of the structure of what we are hearing. Barlow has presumably prepped the audience with program notes to understand that the orchestration of *Im Januar am Nil* is guided by the formant structures of language. Yet, after 13 minutes of music, none of which is actually meaningful or even intelligible as language, the audience is likely to slip into a way of listening that treats all the sounds as merely sound, without linguistic meaning. Barlow's intervention with *Six Quatrains* calls attention to the shift in the languages of Western

[80] From the program notes to "Im Januar am Nil" included with the album Clarence Barlow, *Musica Algorithmica*, CD (World Edition, 2018), 6.

music used throughout the past 600 years, and recontextualizes the remainder of the piece. Most of all, this insertion pulls us back from a notion of music as "pure sound," concerned only with the surface, to recognize music as fundamentally constitutive of symbolic, categorical structures.

Jonathan Harvey, *Speakings* (2008)

Jonathan Harvey's work *Speakings* takes a very different approach in its implementation of electronics to derive music from a linguistic source; his work is for a full orchestra with soloists, by far the largest ensemble employed.[81] Harvey also employs an analysis/resynthesis technique on the acoustic instruments themselves, sampling and processing the orchestra's soloists and filtering their sound using a process analogous to Linear Predictive Coding (LPC), recalling the earliest techniques for speech synthesis.[82] In contrast to all other implementations, Harvey's technical collaborators Charpentier and Bresson created a database based on acoustic features analyzed from recordings of instruments. The authors note that their automated orchestration engine Orchidée "embeds a *timbre model* that efficiently estimates the joint features of any instrument sound mixture." This database is target-sound-agnostic, and could work just as well for simulating a car horn as for a speaking voice. The orchestration engine Orchidée contains a "symbol"

[81] Nouno Gilbert et al., "Making an Orchestra Speak" (SMC, Porto, Portugal, 2009).
[82] B.S. Atal and Suzanne L. Hanauer, "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave," *The Journal of the Acoustical Society of America* 50 (1971): 637.

and a "sound" space. The "symbol" space serves as a set of instrumental constraints within which an orchestration might be found—note, dynamics, playing styles—and an optimal solution within the "sound" space is found that satisfies those constraints. Therefore, the end-user is able to control certain musical parameters to categorically exclude certain possibilities.

In *Speakings,* the first salient speech-like elements are those emulating a baby learning to speak, primarily in the first movement of the three-movement piece (although these elements make a return at the very end). Harvey writes in the program note that:

> The vowel and consonant spectra-shapes flicker in the rapid rhythms and colours of speech across the orchestral textures. A process of 'shape vocoding', taking advantage of speech's fascinating complexities, is the main idea of this work.[83]

This "shape vocoding" is transparent in many sections, such as rehearsal mark G of the first movement (the measures are not numbered). In this section, the solo oboe line is in rhythmic unison with an electronic sample triggered by a MIDI keyboard. The sample, titled "Unstretched Baby Scream" in the score, is meant to be balanced with the oboe soloist through amplification in the hall. Additionally, a subset of the strings—seven players in all—play material in homophony with the oboe, frequently doubling the oboe at the octave or fifth. (In these moments, the texture approaches that of Tenney's *Three Indigenous Songs*, albeit considerably more dense.) The sample playback of the baby serves

[83] Jonathan Harvey, *Speakings* (London: Faber & Faber, 2008).

as an index of the source of the orchestration; it justifies a logic behind the orchestration, while also helping the gesture register as mimesis. The aim of this section is striking as a cry *against* construction, even in the midst of a compositional process tied so intimately to computation.

These gestural, spatialized samples of the sound of the baby's cry strike the listener as a kind of grasp for expression in an imagined pre-technological history. Harvey connects this aspect of his work to "a baby with its mother, or like first man," referencing the "original, pure speech"[84] of Buddhist mythology. Yet, the "speech" segments are fundamentally mimetic in nature. They dissolve the sense of musical form through their non-relevance to any harmonic or otherwise syntactical element, and generally dissolve any sense of metered time. Furthermore, the speech registers as speech primarily due to the dense glissando textures in the pitch contours of whichever instrument is performing a solo at the given moment. In juxtaposing the linguistic elements of the work of art and the mimetic impulse, Adorno writes:

> Everything in artworks that resembles language originates in form and is thus transformed into the antithesis of form, the mimetic impulse. Form seeks to bring the particular to speech through the whole.[85]

While the distinction between form and content is one of a matter of the level at which they are viewed (to paraphrase Tenney) mimesis generates a unitary element that cannot,

[84] Harvey.
[85] Adorno, *Aesthetic Theory*, 144.

logically, be broken into parts. The mimetic element is irreducible. That is, if Harvey had split a single "speech-like" glissando segment into two segments, each segment would fundamentally convey the same amount of information; the information conveyed is a registering of the segment *as speech,* not an evaluation of the "content" of the speech-segment itself.

In Harvey's work, we can locate the opposition of information—roughly corresponding to form—and expression in the fixedness of pitch and metrical time. The clearest example of this is Movement III, rehearsal mark L, at which point the clarinet soloist is given the instruction "accel. independently,"[86] and the performer of the live electronics is meant to spatialize the sound of the solo clarinet. This occurs simultaneously with strict eighth-note figures in nearly all other instruments; Harvey calls this texture a "hymn," and it is filled with quasi-tonal elements including imitative entrances. In the program note, Harvey describes this section as "close to Gregorian chant." It is striking, that even in the section of the work that most resembles what Adorno would describe as the linguistic qualities of music—as distinguished from the "speech-like" aspects of this particular musical work—these passages take on a mimetic quality due to their program notes and context within a work nearly entirely devoid of them. Their mimetic nature registers them as gesture, rather than as musical syntax; they are a *reference* to a prior period where musical syntax and speech were intertwined, but they do not enact that intertwined

[86] Harvey, *Speakings*, 102.

quality. The hymn becomes a reification of an earlier form, rather than a living embodiment of that form.

Peter Ablinger, *A Letter from Schoenberg* (2004)

Peter Ablinger's *A Letter from Schoenberg,* is from his "Quadraturen" series of works which all involve "transcribing" found sounds. In this work, Ablinger synthesizes speech with an acoustic instrument, but an inhuman performer. His speech synthesis project results in something intelligible—though not without the power of suggestion, as the audience is instructed to read the text simultaneously with the listening—as it is transcribed and performed by a software-controlled player piano, through the analysis of speech.[87] This technique for transcription models neither the source nor the mouthpiece, but rather considers *both* to be neutral. The sonic source of the voice is then removed even from specialized models unique to it; it is treated no differently from any other sound, and given no special consideration through speech-specific models like those used by Tenney and Barlow. Likewise, the use of a piano is "neutral" in the sense that it is a ubiquitous presence in Western concert music, and is absent of any consideration of timbre; the piano reduction of a score is considered a score without orchestration—even as it is, in fact, a *re-*orchestration for the piano.

[87] Ablinger requires the audience to read the "source text" simultaneously with hearing the piece. Peter Ablinger, "Quadraturen," accessed March 16, 2020, https://ablinger.mur.at/docu11.html#principles.

Ablinger's piece operates at the threshold of intelligibility. He writes that his "main concern is not the literal reproduction of speech itself but precisely this border-zone between abstract musical structure and the sudden shift into recognition—the relationship between musical qualities and 'phonorealism': the observation of 'reality' via 'music'."[88] The power of suggestion, in the form of written prompts given to the audience, is necessary for the success of the composition, as the tension between intelligibility (as language) and sonic material is the core function of the piece. Furthermore, the quoted words "reality" and "music" call into question the stability of those terms; both reality and music are subject to our perception and interpretation, and derive their categorization through such interpretive mechanisms. The composer, critic, and curator G. Douglas Barrett—who has composed many "transcription" works himself—writes of the piece's "potential for constant oscillation: one could hear the mass of sound as piano music (extremely loud and fragmented piano music), and then upon recognition of a word, for instance, be immediately thwarted into intelligibility."[89] This *play*, the constant in-motion state of the understanding of an artistic work, destabilizes the mechanism of interpretation in the listener, and engages multiple simultaneous understandings—that is, an understanding of the material as either *reality* or *music*—that keep the parsing of the material in flux. Listening to *A Letter From Schoenberg,* we hear continuous aspects of the sound as musical

---

[88] Ablinger.

[89] G. Douglas Barrett, "Between Noise and Language: The Sound Installations and Music of Peter Ablinger," *Mosaic: An Interdisciplinary Critical Journal* 42, no. 4 (2009): 160–61.

events (pitch-height, density, and overall volume in the room), while the categorical mode of understanding is *entirely non-musical*—there is no “harmonic series” harmonic language to speak of, as meaning is purely linguistic, and only that through prompting. In this, Ablinger seems to be indicating that “information” is located not in any particular set of physical aspects (whether written or heard) but rather that information is a co-production with the listener.

Speech-music techniques in my work

A number of my compositions are concerned with the transcription of spoken language into notated music. However, I do not generally aim to render the music “intelligible” in terms of speech. Rather, I am interested in how the continuous parameters of speech might be reliably mapped to certain parameters of notated instrumental music. In speech, these repeated parametric morphologies become “symbols,” through their repetition, even as humans are able to separate the carriers of “information” (the formants, and spectral properties of the signal) from the aspects of speech such as cadence and pitch inflection, which are not, strictly speaking, translated into writing. In this collection of pieces, I am interested in the possibility of music that takes on some aspect of symbolic repetition, while maintaining distance from the semantic meaning of language.

In the sense of transparence

An early piece to undertake this project was *In the sense of transparence,* for solo cello, in which the continuum of "darker" to "brighter" vowel sounds were mapped to the position of the cellist's bow from the fingerboard to the bridge; the sense of "dark" or "bright" was taken from the position of the second formant, after some experimentation with various options with the performer. To simulate diphthongs, a slanted line was drawn across the full duration of the vowel to indicate the changing position of the bow. The slope of this line was derived from the best-fit line through the second formant within each vowel. This diphthong slope changed throughout the course of the piece, such that when the melodic lines became more intricate (such that the pitch-contour of the speech was more accurate) the slope of the line was flattened to a global mean for each sounding syllable. Two excerpts from the score are given here as examples, from repetitions 1 and 16 of the same excerpt of text.

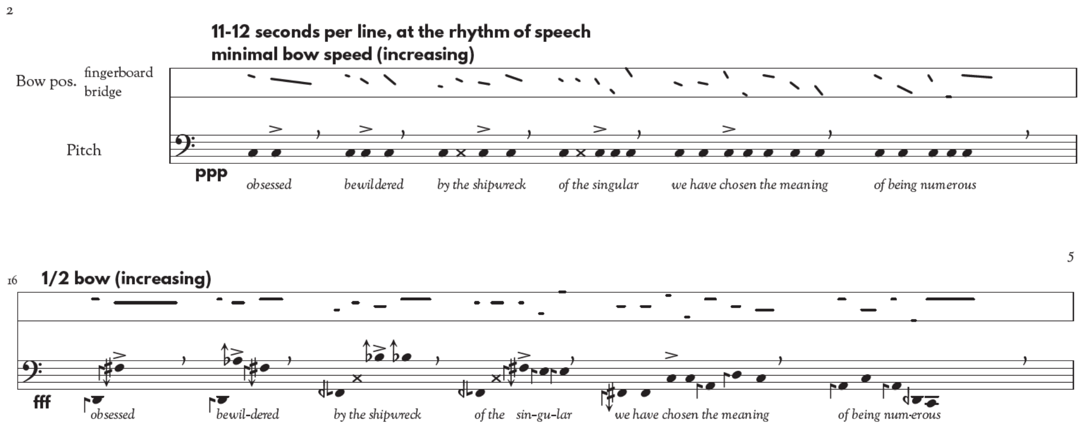

Figure 4: Two excerpts from *In the sense of transparence*, an earlier piece for solo cello

The results of this experiment led me to draw two conclusions about this continuous, parametric approach to speech-music transcription. First, there was *interference* between the discrete pitch content heard as "music" and the identification of the sound as "speech." As the contour of the speech line became more intricate, the focus of both listener and performer turned to the pitched nature of the melodic line, and the line sounded more like a melody than like a text "spoken" through an instrument. Second, the rhythmic interpretation also became more regular (i.e., more musical and less speech-like) as the intervallic leaps became wider; perhaps this is because the ease of performing a single pitch leaves the performer more room to focus on the emulation of speech through rhythm and articulation. This piece was performed a few times by the cellist Séverine Ballon, and was developed through the course of a week-long rehearsal and workshop process.

We remember not the word, but the sound of the word

My efforts at speech-music transcription were most fully realized with a piece for mixed chamber music sextet titled *We remember not the word, but the sound of the word.* In this piece, I wrote an original text, which I recorded myself speaking for the source material of the transcription. The orchestration follows a process similar to Tenney's *Three Indigenous Songs*. For each vowel phoneme, the harp plays the fundamental pitch of a section, while the rest of the ensemble (contrabass, viola, oboe, and flute) play one of the four vowel formants. These are quantized to the 32nd note with slurs over the pitches, so that instrumentalists are frequently changing pitch mid-phoneme. In the orchestration, each instrumentalist is responsible for sounding the harmonics nearest to the center frequency of an assigned formant as analyzed by the computer process. To facilitate the performance of natural harmonics over an arbitrary fundamental, the contrabass and viola both tuned to a scordatura g-c-G-C, and used moveable rubber "floating bridges" between the strings and fingerboard. This allowed for music consisting entirely of natural harmonics without a constantly changing scordatura. These bridges were moved throughout the piece during specified "tuning breaks." The oboe and flute generally found alternate fingerings, or adjusted their embouchures to find the harmonic series pitches, and the "tuning breaks" helped the winds find their pitches for each proceeding section.

The piece is structured in 13 sections, matching the 13 syllables of the source phrase "We remember not the word, but the sound of the word." This source phrase ends every section and also informs the harmonic structure of the work as a whole. Each section consists entirely of harmonics of a given fundamental, with the harp playing the fundamental at the start of each word. The pitch contour of my recording of the source phrase was quantized and mapped to a just intonation 12-note system, such that a handful of intersecting harmonics would overlap from section to section. Because of the natural prosody of English declamatory statements, the piece "peaks" with a fundamental on a high C about two thirds of the way through, and then descends to a low C (open string) fundamental at the end.

A summary of the compositional process follows:

1. Compose and record the text, as the source material for the composition.
2. Segment every vowel and consonant manually into separate phonemes.
3. Analyze vowels and consonants separately:
    a. Analyze the F0 pitch and formant frequencies of each vowel, at each frame, and assign each formant F1-4 to an instrument.
    b. Match each consonant sound to a percussion gesture (recorded in collaboration with composer-percussionist Brian Baumbusch) using the cosine distance metric on the MFCC vector.
4. Transcribe all vowels into pitches, using the formant curves (both center frequency and formant filter bandwidth) as stochastic distributions for harmonics of a fundamental pitch that is different for each section.

5. Quantize vowels (pitched ensemble) and consonants (percussion) to 32nd notes.
6. Determine time signature changes, based on the beginnings of words and ideal conductor cues.

As the entire structure of the composition is given (essentially) by the structure of the text, this composition was an experiment in drawing out the aural effect of repetitive verbal statements—the "microstructure of speech acoustics," to use Polansky's description of Tenney's work—even as the actual *instances* of that repetition change. For example, the first three sections of the piece begin with the phrase "After light, light," provided here as excerpts from mm. 2-3, 14-15, 26-27. The first two formants, in contrabass and viola, follow approximately the same general contour in the same general range for each repetition, yet they do not repeat precisely. This messy near-repetition is the key point of interest in my project, as the repetition of spoken language is similarly imprecise.

Figure 5: *We remember not the word, but the sound of the word*, mm. 2, 14, 26

Central to the composition, as well, is a sense of information which is "hidden" from the listener. The ensemble's music is derived from speech, but the text itself is not a part of the ensemble's parts (although it is present in the conductor's score), and furthermore the text is never given to the audience. An initial series of performances and even an entire recording session was entirely instrumental, but in determining the aesthetic for a final recording and series of later performances, I decided to add my own voice with vocals "low in the mix," buried under the instruments. The ensemble and I replicated this aesthetic in performance, through four dates in California in May 2018, during which I sat in the back of the ensemble, following the conductor, and read the text at the level of audibility, but (usually) below the level of intelligibility.

Furthermore, the process of preparing for and finally recording the composition meant that the piece took on a different life as a studio construction, such that it could be recorded in sections and could be spliced to realize tuning shifts that are impossible in performance. There is a gap between the performance of the composition as a live event, and the artificial studio construction; yet, the studio construction is recognizable as something that is very nearly a live event. They are, in a sense, very different compositions, as the studio processing and overdubbing enables new techniques and semantic shifts in mixing and audio processing that are not possible in a live setting.

This opens two doors for critical inquiry: how does the presence (on stage) of a composer-performer alter the audience's perception of the composition? And, second, what mental mechanism is employed when hearing an ensemble *as speech* versus hearing it as purely instrumental sound? Following Barrett's analysis of Ablinger, and Adorno's sense-making apparatus existing at the threshold of a clear conceptual whole, I aim in my work to keep the listener on edge, always oscillating between competing modes of understanding and parsing between language and music. The presence of the composer foregrounds the *process of composition itself*, and furthermore presents the composition as an object that is the result of a determined process. The understanding (or at least awareness) of the process, for the audience member, can potentially confound the sense of authorial intention. If the specific note-to-note details are not "composed" by the individual sitting on stage talking, but rather the source material remains hidden beneath the ensemble, the sense of where the "actual" composition exists is blurred. The harmonic results clearly exist, but they are a result of a process rather than precisely placed.

4. Rationalization and the syntactical articulation of harmony

James Tenney described articulation in music as dependent on the gestalt psychology factors of similarity and proximity. For parameters such as pitch height and loudness the factor of similarity is relatively straightforward. A collection of pitches within a few steps of C4 will be similar to one another in register, and likewise a sequence of fairly quiet pitches will tend to cluster as a single grouping. Pitch-height and loudness are clearly unidimensional axes, with "up-down" or "soft-loud" relationships, which lend themselves to simple statistical measures of similarity such as measuring the difference from a mean value in a proposed grouping. However, harmonic similarity is complex and multidimensional, informed by cultural factors as well as physical ones, and recalling concepts of syntax and even a "harmonic language." Segmentation and articulation creates the linguistic quality in musical works, through its ability to transform specific sequences of sounds into "symbols" that construct syntactic relationships. Furthermore, these collections of symbols create a hierarchy (or holarchy), dependent on the horizontal articulations (i.e., the boundaries of the temporal gestalts), and the expected syntax of nested groupings. Tenney recognized the necessity of developing an alternative to "the specifically 'harmonic' devices of cadence, modulation, etc.,"[90] as the 20th century largely

[90] Tenney, "John Cage and the Theory of Harmony," 287.

turned away from the received musical syntax of tonality. Just as the syntax of a verbal language operates on a set of symbols organized into classes, the syntax of a harmonic language is similarly symbolic; a "major triad" is recognizable as a member of a particular class of harmony, regardless of its register or orchestration.

Tenney used the term *harmonic space* to describe the placement of rational harmonic interval relationships as a multidimensional *lattice*.[91] In harmonic space, each pitch occupies a *point*, and an interval between two pitches can be described by an integer vector listing the exponents of the prime factors that generate the frequency ratio of the interval. For example, the interval of a perfect fifth (3:2) can be described by the interval vector [-1, 1], as it describes the exponents of successive prime numbers in the expression $2^{-1} \times 3^{1}$. In harmonic space, each new prime number increases the dimensionality of the space by one. Therefore, a chord containing only the prime numbers 2 and 3 could be described using a 2-dimensional space; 2, 3, and 5 would require a 3-dimensional space; and so on, adding a new dimension for each new prime number. This conception of harmonic space allows for the creation of a similarity metric to inform temporal gestalt formation, which Tenney called *harmonic distance*. Points that are "close" in harmonic space have related harmonies, and a lower harmonic distance from one another; colloquially, we might think of them as "in the same key," although that term is susceptible to cultural implications in recalling the hierarchies that Tenney avoids. By contrast,

[91] Tenney, "John Cage and the Theory of Harmony."

collections of pitches that are distant from one another will tend to form their own groupings. Just as we might hear two pitches a half step apart as "close" by virtue of their similarity in pitch-space, we also may hear pitches related by an octave (one step on the 2-axis) or a perfect fifth (on the 3-axis) as "close" through their similarity in harmonic space. Evaluating harmonic distance therefore becomes critical for a theoretical project which seeks to determine the articulations of a musical work in which harmonic relationships play a role.

The syntactical associations that particular harmonies develop throughout the course of a musical culture are far more intricate than can be explained by physics alone. Yet, there are some physical and perceptual features of simple harmonic intervals—e.g., the phenomenon of octave equivalency,[92] described by Roger Shepard—that provide some justification for a measurement of harmony based on ratios using small prime number integers. However, many of the actual "heard intervals" in music are not exact ratios, but rather slightly altered, or what we might colloquially call "out-of-tune." The evaluation of these inexact ratios in terms of a harmonic distance metric based on small prime numbers requires a *rationalization* of these intervals; in other words, irrational intervals must be considered in terms of the rational intervals they are presumed to approximate, in order to utilize harmonic distance as a quantitative tool in evaluating temporal gestalt formation.

---

[92] Roger N. Shepard, "Circularity in Judgments of Relative Pitch," *The Journal of the Acoustical Society of America* 36, no. 12 (December 1, 1964): 2346–53, https://doi.org/10.1121/1.1919362.

A method of interval rationalization forms the basis of this chapter, followed by a novel algorithm for *interval aggregate* (i.e., chord) rationalization in higher-dimensional spaces. These methods are both analytical and generative. They might be used to evaluate temporal gestalt formation in music based on tempered tuning systems, but they also may be used to create sets of pitches and scales for use in generative algorithmic musical composition. A discussion of their usage as a generative tool in my composition *Terrain* continues in the following chapter.

Tuning systems, perceptual tolerance, and interval rationalization

Of the many culturally contingent aspects of music, the *tuning system* is one of the most ubiquitous and influential. A tuning system encompasses a set of pitches, from which smaller subsets may be selected. In a given culture, these subsets may be called *scales* (though "scale" often carries other structural implications as well) or *modes,* or any number of other culturally specific terms. Unlike many aspects of music that might be altered through developing a new performance practice—rhythm or tempo, for example—the tuning system of a body of music is overwhelmingly influenced by the physical instrumentation available to musicians and composers; the tuning system and corresponding instrumentation is influenced by industry as well as culture.[93]

A theoretical framework proposed by Polansky et al. models tuning systems as an optimization problem over a set of parameters, and theorizes that a given tuning system is generated through the negotiation of five primary concerns: number of pitches, repeat factor, target intervals, hierarchy of intervals, and hierarchy of keys.[94] If these concerns are in part mutually exclusive (e.g., if one scale degree must approximate two different target intervals in two different contexts) then a *tempered* tuning system is generated, in which an interval in the system may split the difference between two potential target intervals; this is called a *tempered interval*. The tempered interval is then a syntactical stand-in for either of the target intervals, depending on the context. There is a trade-off in this process between economy and accuracy; a culture that produces pianos with 24 keys to the octave would be able to more accurately perform music that approximates harmonies using higher harmonic partials such as the 7th and 11th harmonics, but it will have doubled the amount of piano wire necessary to string the instrument.

[93] It should be noted that "influenced" does not mean "controlled." See Nathaniel Mackey's comment on Amiri Baraka, mentioned earlier, for a statement on "imprecise" pitch as a proxy for protest. Mackey, "The Changing Same: Black Music in the Poetry of Amiri Baraka."

[94] Larry Polansky et al., "A Mathematical Model for Optimal Tuning Systems," *Perspectives of New Music*, 2009, 69–110.

Tolerance and harmonic syntax

In unequal tempered tuning systems, where the intervals between adjacent degrees are not identical as in 12-tone equal temperament, there might be multiple different tempered intervals that are all considered to "stand in" for some target interval, with varying degrees of accuracy. One example is the temperament "Werckmeister III," by Andreas Werckmeister. In this temperament, there are many different versions of the perfect fifth, but each of these fifths is understood syntactically as a perfect fifth. This enables the functional harmony of Bach's *Well-Tempered Clavier*, as the fugue entrance which starts on the fifth is always heard as the dominant as long as the fifth scale degree of the key of the fugue approximates the interval 3:2. The range within which a tempered interval might be *syntactically understood* as a rational interval is considered the *tolerance* of the rational interval.

The term *tolerance* has taken on two distinct but related meanings corresponding to both practical and perceptual concerns, and therefore the historical terminology must be clarified. In the Early Modern and Baroque eras of Western music, "tolerance" related to the practical concern of tuning a harpsichord, and to the sound's unpleasantness. This sense of the word is closer to the related word "tolerable," and can be thought of as the opposite of "intolerable." Baroque-era theorist and composer Johann Kirnberger said that while a well-tuned keyboard instrument must ensure that all octaves are tuned exactly to an ideal 2:1 frequency ratio, and that the perfect fifth should be fairly close to the ideal of

3:2, the more complex interval of the major third of 5:4 has a higher tolerance for error before it becomes unpleasant (or even unrecognizable). Kirnberger's statement that "[T]hirds tolerate still greater deviation from purity before becoming unpleasant"[95] indicates that there is a hierarchy of intervals (octaves, fifths, thirds) and that this hierarchy corresponds to the complexity of the prime number factors that constitute the ideal ratios of the intervals: 2:1, 3:2, and 5:4.

Tenney's use of "tolerance," however, is directly tied to a syntactical and perceptual argument about harmony. Tenney argues that there is a region around a simple interval ratio where a more complex interval will be syntactically understood as the simpler interval. Tenney defines this as the "tolerance range" of an interval:

> [S]ince our perception of pitch intervals involves some degree of approximation, these frequency ratios must be understood to represent pitches with a certain *tolerance range*—i.e., a range of relative frequencies within which some slight mistuning is possible without altering the harmonic identity of an interval.[96]

Tenney's notion is not one of unpleasantness (or "intolerability") but rather a more aesthetically neutral statement of equivalence—or "harmonic identity," in Tenney's words. His use of "understood," in particular, places the function of rationalization in the mind of the listener; it is a mental drawing of boundaries and demarcations upon the physical world. Tenney's intertwining of tolerance and approximation with "harmonic identity"

---

[95] Johann Kirnberger, *The Art of Strict Musical Composition* (New Haven: Yale University Press, 1982), 22.
[96] Tenney, "John Cage and the Theory of Harmony," 294.

concerns the mechanism by which sounds are grouped into symbols according to the perception and understanding of the listener. As the translation of sounds into symbols is necessary for the creation of syntax, interval tolerance is therefore an essential concept for the development of a *harmonic syntax* in the context of a tempered tuning system.

Melodic variety and the human voice

Questions of tolerance inform limits on the interval variety of a tuning system. If each interval is bounded by a region of tolerance, then the harmonic identities of the simpler intervals will eventually subsume the more complex intervals in their vicinity. As I will demonstrate later in this chapter, this feature of tolerance creates an "upper limit" on interval variety, at which the pitch-space is entirely saturated; this is acknowledged by Tenney as a potential perceptual limit on the dimensionality of harmonic space.[97] An upper limit on interval variety is a compositional constraint, as it points to a perceptual limit on the intervallic resources available. If all intervals within a tolerance range of 20 cents of 3:2 are understood syntactically as the perfect fifth 3:2, and if this follows for every simple ratio within an octave, the octave will eventually reach a "saturation point" at which no further harmonic detail is possible. For example, intervals such as 63:32 (very nearly the octave 64:32, which reduces to 2:1) will likely be understood as an "out-of-tune octave," rather than a highly complex harmonic interval that is close in pitch-height to an

[97] Tenney, 295.

octave. This potential limit on harmonic resources was clearly a historical concern. Andreas Werckmeister, whose tuning Werckmeister III is a classic Baroque "well-temperament," in fact proposed a number of other, less-widely-adopted tuning systems that occasionally accounted for variation. Werckmeister wrote about his tuning Werckmeister IV that, "The thirds C#, F; G#, c; and F#, Bb are indeed somewhat harsher [than Werckmeister III], but in the whole concept and change of harmony, completely pleasing, for *one perceives some variation.*"[98] Werckmeister, perhaps reflecting concerns of his era, notes that a number of thirds (which he spells as diminished fourths) are "somewhat harsher," but he makes an accommodation due to a secondary criterion of variety.

Variety is frequently a concern for the *melodic* resources for the composer. This is often tied to a consideration of the human voice, and to the tension between melodic (horizontal) and harmonic (vertical) concerns. A key figure in the use of microtonality in the 20th century, the composer Harry Partch used a complex, 43-tone-to-the-octave set of rational pitches in his transcriptions of voices, with the explicit intention of better approximating human speech. Partch even makes a language-specific claim, citing that, "Rousseau prophesied that one day some French composer would realize a 'recitative appropriate to the simplicity and clarity of our language,' that this recitative would 'proceed by very small intervals,' that it would have 'few sustained notes [tones],' and that it would

[98] Andreas Werckmeister, *Musicalische Temperatur*, trans. Elizabeth Hehr (Oberlin, OH: Oberlin University Press, 1974), 114.

be "nothing that resembles song."[99] In a recollection of a speech-transcription project by W.B. Yeats, Partch wrote that "I made diagrams of his inflections, but my memory of his vibrant tones is more accurate than my marks."[100] Partch's theoretical framework was aimed in part at reaching maximum melodic variety in a tuning system, in order to enable more precise transcription of spoken language at a finer resolution. In this, Partch echoes Rousseau:

> [Harmony] shackles melody, draining it of energy and expressiveness. It wipes out passionate accent, replacing it with the harmonic interval. It is restricted to only two types of songs, within which its possibilities are determined by the number of oral tones.[101]

Melodic variety—correlated with the irrational—is seen as tied to expressiveness, to a liberatory and authentic mode of communication in music. This tension between harmony and melody, traced back to Rousseau, pits the "linguistic quality" of harmonic intervals—that is, its syntax, and its hierarchical generation—against the speech-like characteristics of melodic intervals; it pits information and structure against expression and individuality. Partch made a point of promoting his own system for what he described as its affective and subtle qualities; he called his aesthetic "corporeality." He wrote, "Intricacy and subtlety are

[99] Partch is likely restating a quotation given in Lockspeiser's biography of Debussy, which he attributes to the composer Maurice Emmanuel, and which echoes claims made in Rousseau's "Essay on the Origin of Languages." Edward Lockspeiser, *Debussy*, Fifth edition, Master Musicians Series (London: Dent, 1980); Jean-Jacques Rousseau and Johann Gottfried Herder, *On the Origin of Language* (Chicago and London: University of Chicago Press, 1966); Harry Partch, *Genesis of a Music; an Account of a Creative Work, Its Roosats and Its Fulfillments.*, Second edition, enlarged. (New York: Da Capo Press, 1974), 27.
[100] Partch, *Bitter Music*, 167.
[101] Rousseau and Herder, *On the Origin of Language*, 57–58.

inherent in a musical system of numbers."[102] It is clear, then, that Partch believed that rationalization and expression were not categorically *opposed*, but that greater interval variety overall was necessary to bring both the syntactical, linguistic qualities of harmony and expressive, speech-like qualities of melody under the auspices of a single theoretical system. For Partch, "variety" becomes the link that enables a connection between syntax and speech.

Continuous harmonic space and interval rationalization

One of the primary applications of interval tolerance is the process of interval rationalization. Broadly speaking, each rational interval is given a *region of tolerance*, such that any more complex interval that falls within this region is considered to be *rationalized* to the harmonic identity of the given rational interval. Interval rationalization asks the question: when we hear a given frequency ratio of arbitrary complexity, what harmonic identity do we assign to it? This method can be thought of as the inverse of tolerance, in that instead of finding the region of tolerance around a given interval ratio, it seeks to find the interval ratio whose tolerance range encloses a given tempered pitch. Barlow describes this process as "a more or less subconscious bending of the notes to an optimal tuning within the mind's ear of the listener."[103] Tenney notes that it is necessary to have some

[102] Partch, *Bitter Music*, 113.
[103] Clarence Barlow, "Two Essays on Theory," *Computer Music Journal* 11, no. 1 (Spring 1987): 48.

sense of tolerance for any tempered system in which rational harmony plays a role.[104] Both composers utilize interval rationalization to a great degree in their works for tempered tuning systems, and have discussed the theoretical underpinnings of this process in detail. In Barlow's case, rationalization featured prominently in his work for retuned piano *Çoğluotobüsişletmesi*,[105] and for Tenney most fully in his 72-tone work for six harps, *Changes*.[106]

For both Tenney and Barlow, it is assumed that the region of tolerance of an interval is some window around a simple ratio in logarithmic pitch space. Tenney expresses this assumption with the formula: [107]

$$r = k\log_2(q/p)$$

Equation 1: Tenney's tolerance formula

In Tenney's formula, the value *k* describes the size of this region of tolerance around a given rationalized pitch; he suggests that this value could be set such that the region of tolerance is about 17 cents, to accommodate Western tonal harmony in which the third is "mistuned" in 12-tone equal temperament by approximately 17 cents.[108] Barlow, in a

[104] Tenney, "John Cage and the Theory of Harmony," 294.
[105] Clarence Barlow, "Bus Journey to Parametron," *Feedback Papers* 21–23 (1980): 135.
[106] James Tenney, "About Changes: Sixty-Four Studies for Six Harps," in *From Scratch: Writings in Music Theory*, ed. Larry Polansky et al. (Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015), 327–49.
[107] James Tenney, "The Structure of Harmonic Series Aggregates," in *From Scratch: Writings in Music Theory*, ed. Larry Polansky et al. (Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015), 246.
[108] Tenney, 247.

slightly different approach, draws a Gaussian curve around each rational interval, "damping" the strength of the interval by some factor incorporating the pitch distance and the harmonic strength (what he calls *harmonicity*) of the interval, and calls the "nominal tolerance" the width at which the damping factor of the Gaussian curve equals a factor of 20.[109] The method of interval rationalization outlined in this section follows the work of Barlow and Tenney, in considering pitch proximity and harmonic distance (or harmonicity) in the rationalization of a given interval.

Throughout this process, I use Tenney's harmonic distance function (rather than Barlow's harmonicity) as the core metric. There are three major reasons for this. First, Tenney's harmonic distance metric is computationally simpler and is better suited for computing large data sets on the hardware at hand. Second, Barlow's harmonicity metric steeply penalizes higher prime numbers. To my ear, the ordering of intervals given by Tenney's metric more accurately represents the Western classical harmonic language (e.g., 7:5 is significantly simpler than 45:32, as 7:5 is the salient interval of the dominant seventh chord). Finally, Tenney makes clear that his harmonic distance metric is rooted in a physiological basis concerning the intersection of simultaneous tones, relating directly to the logarithmic metric of pitch distance.[110] This direct connection between pitch distance and harmonic distance opens the door to fruitful mathematical relationships that are

[109] Barlow, "Two Essays on Theory," 48.
[110] Tenney, "The Structure of Harmonic Series Aggregates," 258.

unavailable with other metrics. Below is a subset of these rational pitches over one octave (Figure 6) with the harmonic distance of each interval given as the *y* value. In this graph, the lower y-values refer to simpler interval ratios. I refer to the annotated ratios as *potential rationalizations.*

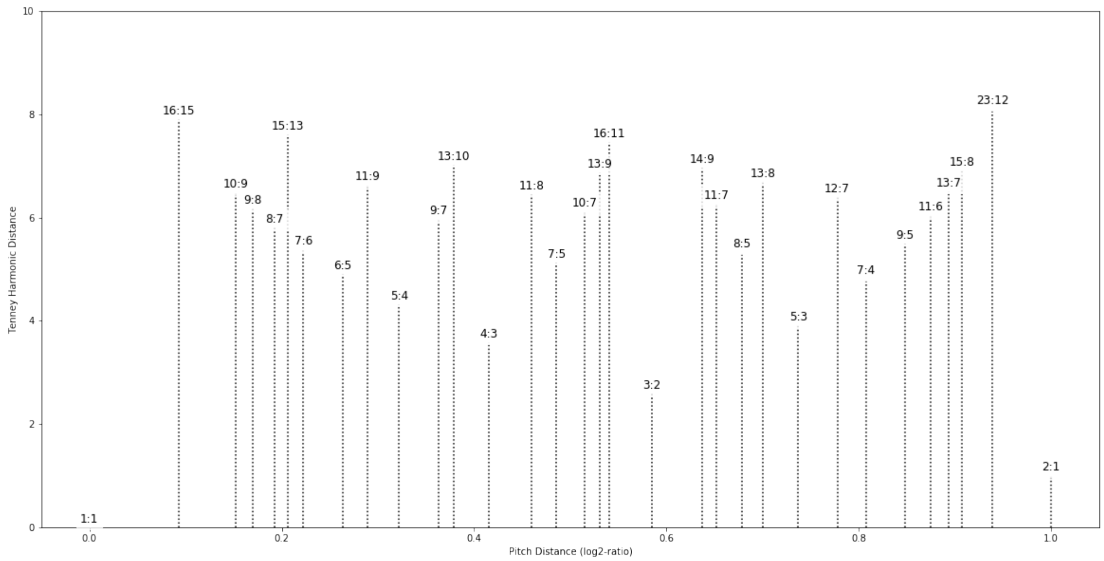

Figure 6: A subset of the potential rationalizations within one octave

I then calculate a parabola around each of these potential rationalizations, following a Gaussian curve determined by a coefficient which corresponds to the variance of the parabola. The parabola function used, with relevant variables, is given here:

$$P(d) = \frac{d^2}{c}$$

Equation 2: Parabola function for potential rationalizations

d : pitch distance from potential rationalization to target interval
c : variance of parabolic curve

This parabolic-scaled pitch distance is then scaled by the harmonic distance of each rational interval as an exponent of base 2,[111] and then added to the original harmonic distance. I call this the "scaled distance," to indicate that it is scaled relative to the harmonic distance and pitch distance deviation from a rational interval.

$$SD(x, p, h) = ((2^h) \times P(x - p)) + h$$

Equation 3: Scaled distance of exact pitch relative to secondary interval

x : pitch distance of irrational interval
p : pitch distance of potential rationalization
h : harmonic distance of potential rationalization

This formula is conceptually similar to Barlow's parabolic weighted frequency window of a Gaussian "damping" curve,[112] with the distinction that it uses Tenney's harmonic distance metric instead of Barlow's harmonicity. This scaled distance is

[111] The exponential raising is necessary so the unison 1:1, which has a harmonic distance of 0, still scales based on the parabolic curve.
[112] Barlow writes, "I preferred this method [of parabolic dampening] to a nonweighted frequency window, which seems less differentiated." Barlow, "Two Essays on Theory," 48.

calculated relative to every potential rationalization in a set along the entire scale. My implementation allows the user to further specify a maximum and minimum number of steps in each dimension of harmonic space, and to restrict these potential rationalizations by given pitch boundaries. For any point along the x-axis, the scaled distance is calculated relative to every potential rationalization in the set; the minimum value calculated is then taken to be the scaled distance of the irrational interval. Within a given domain, the scaled distance can then be calculated for entire length of the x-axis, and can be plotted along the y-axis as shown. This graph shows a number of local minima, or "troughs," which each correspond to one of the possible rationalizations, while the parabolic slopes surrounding each trough contain that local minimum's region of tolerance. We can see here that certain values, such as 15:8, exceed the parabolic curve of other values (2:1, in this case) and do not appear as troughs with the value c=0.0048 shown. Therefore, this method with this given variance parameter considers 15:8 to rationalize to the octave 2:1.

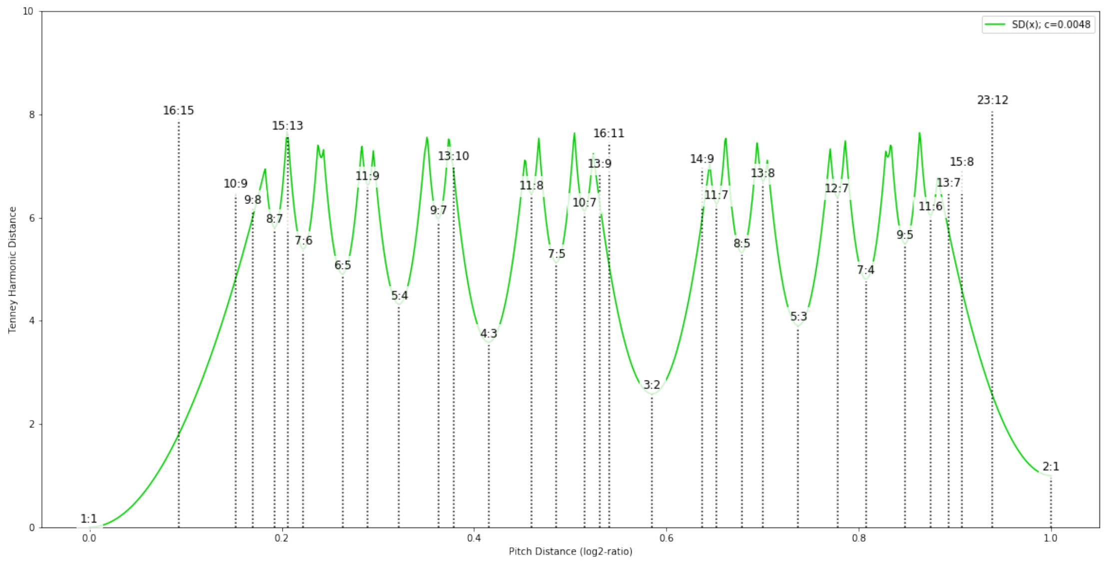

Figure 7: Scaled distance of *x* overlaid on potential rationalizations

The number of troughs—and therefore the number of "tuneable intervals" in an octave—is directly controlled by the variance *c*.[113] To illustrate, a series of lines is given below with varying values of *c,* in which the number of local minima decreases as the value of *c* increases.

[113] Barlow calls this the "damping factor." Barlow, 48.

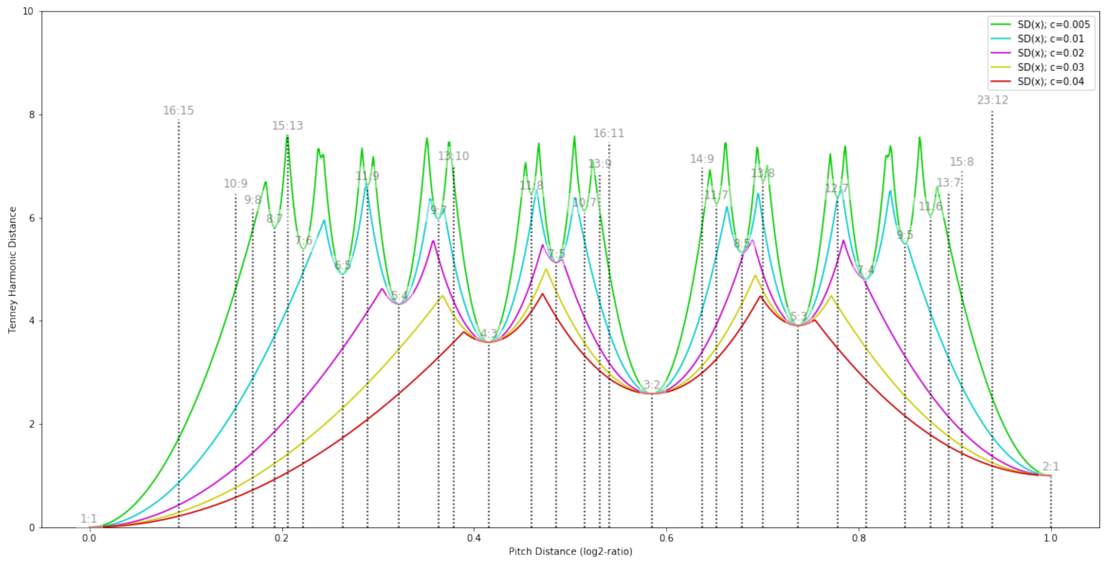

Figure 8: The effects of varying values of the parabolic width coefficient *c*

One method of setting the value of *c* is to examine the critical band phenomenon, considered as a *physiological constant* of the tuning tolerance for the unison, and to set the value of *c* such that the interval 8:7 is the narrowest possible interval available interval in an octave. Any value of *c* between 0.0048 and 0.0063 achieves this result. This selection of 8:7 as the boundary of the critical band phenomenon and the smallest tuneable interval is informed by both Tenney's composition *Critical Band,* as well as the work of Marc Sabat, who documented all of the just intonation intervals he and his collaborators could reliably tune by ear both above A3 and below A4.[114] It is worth noting that the c=0.0048 value gives a list of local minima within one octave corresponding to local minima that is almost identical to the chart given by Sabat; exceptions are a handful of intervals marked "very

[114] Marc Sabat, "23-Limit Tuneable Intervals above and below A," 2005, http://www.marcsabat.com/pdfs/TIab.pdf.

difficult" by Sabat which are near 3:2 and 2:1, and are subsumed in the parabolic curve around those intervals.[115] On the whole, most of the ratios available to Sabat but not to this algorithm involve higher prime factors; perhaps there is a modification of the parabolic curve that could allow for easier accommodation of higher primes.

In computational terms, the method described above transforms Tenney's discrete harmonic distance function, which operates only on integer ratio representations of intervals, into a continuous function operating on floating-point frequency ratios. This continuous function is made up entirely of convex parabolic regions, as shown in the graphs above, and is therefore well-suited to gradient descent optimization; we can think of gradient descent as traveling "downhill" from a point until we find the bottom of a trough. If we consider each local minimum to be a possible interval, then the task of rationalizing a floating-point frequency ratio simply requires optimizing that floating-point value using this naïve form of gradient descent. Framing this problem in terms of gradient descent opens doors to methods of computation that leverage recent architectures and code libraries intended for large-scale machine learning applications; the training of neural network models generally uses some form of gradient descent.[116] Furthermore, the iterative procedures of gradient descent (and its variants, such as Adam, or adaptive

---

[115] Intervals identified by Sabat but not our algorithm include 13:10, 13:9, 16:11, 14:9, 15:8, and 23:12. Intervals identified by our algorithm but not Sabat include 13:11 and 16:9.

[116] At the time of writing this section (February 2020) my implementation is programmed in Python and uses the Google machine learning library Tensorflow 2.1 to run gradient descent on a GPU. Martín Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, 2015, https://www.tensorflow.org/.

gradient descent with momentum) provide interesting results when mapped to musical parameters. Further speculation about this potential for gradient descent as a musical resource is given toward the end of this chapter.

Extension to multiple dimensions

So far in this chapter, we have dealt with single intervals, in one dimension of pitch-space. In general terms, a single interval could be considered to be a *pitch aggregate* of size two, which can be represented in a one-dimensional space. To generalize further, a pitch aggregate of size n can be represented as a point in n-1 dimensional space, where the distance from pitch 1 to each other pitch in the aggregate is represented on each axis in the space. For example, a three-pitch aggregate (such as a major triad) can be represented in a two-dimensional space, such that the distance from pitch 1 to pitch 2 is the x-axis, and the distance from pitch 1 to pitch 3 is the y-axis. It is possible as well to calculate the sum of the harmonic distances (HD-sum) of a pitch aggregate, by summing the harmonic distances of all combinations of intervals in an aggregate, and dividing the total HD-sum by n-1. This scaling is necessary to ensure that the individual axes in the higher-dimensional space have the same values as the original one-dimensional plot.[117] Next, a set of potential

[117] The implementation of this combinatorial harmonic distance computation uses an optimization in which basis space matrices are used to calculate the HD-sum from a collection of prime number vector exponents. This is indebted to recent work of Larry Polansky and David Kant on the calculation of combinatorial contour metrics. The full implementation of this optimized algorithm is available in the accompanying code.

rationalizations must be generated from the collection of all combinations (of size *n*) of potential ratios. The minimum SD value for each point in this space is calculated in the same way as above, with each potential rationalization representing a local minimum in the two-dimensional space. A contour plot of this space for a three-note pitch aggregate is shown below.

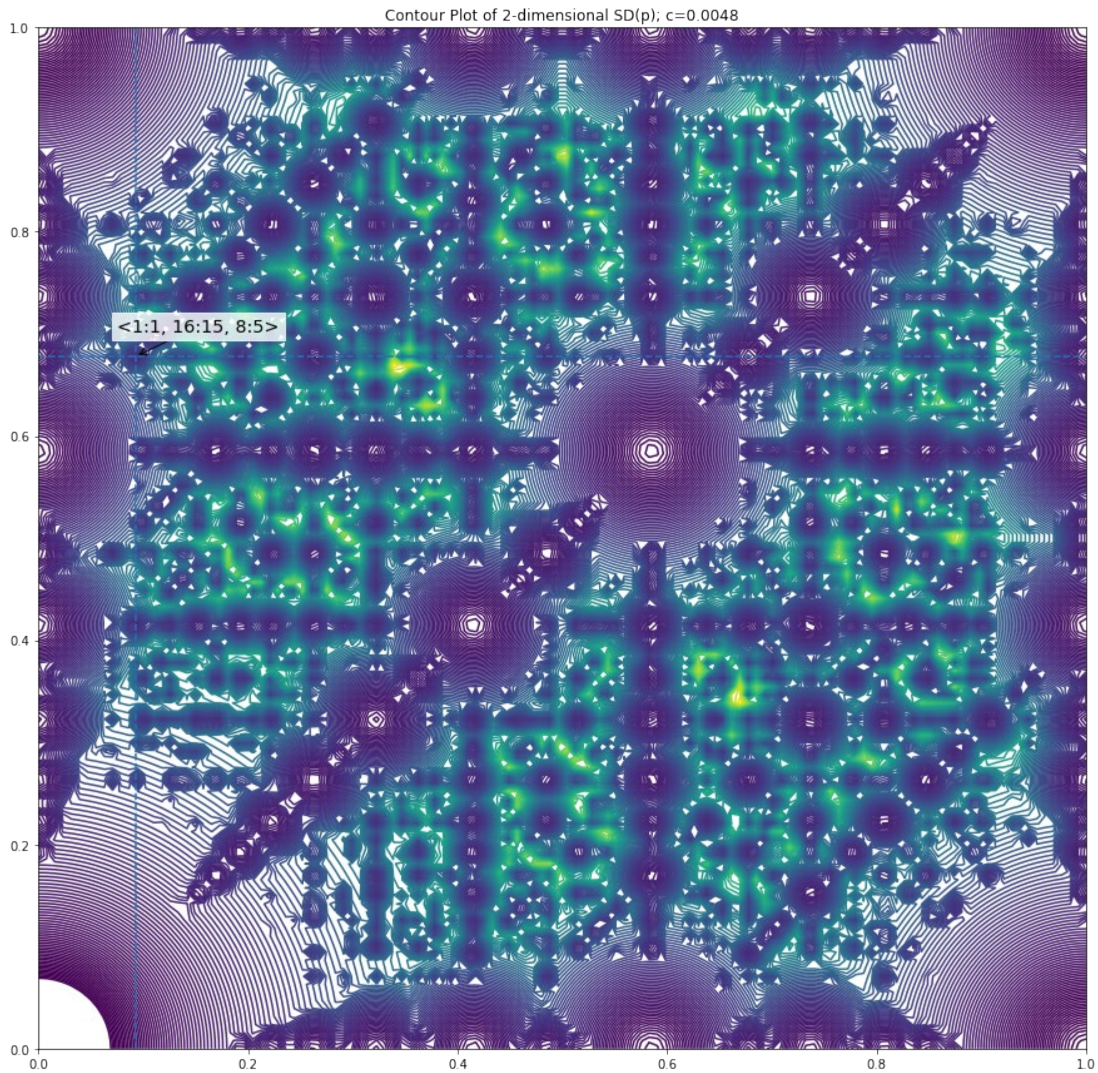

Figure 9: Contour plot of a two-dimensional space, representing the SD-value of a three-note pitch aggregate

The contour plot of this space exhibits a few features that conform to intuitive musical expectations. First, it contains large local minima along the diagonal *x*=*y*. These local minima correspond to instances in which the three-note aggregate contains two pitches in unison with each other. Second, this plot shows that local minima where *x*≠*y* generally appear along *at least one* of the coordinates corresponding to local minima along one of the axes. Intuitively, this means that a tuneable chord must contain at least one interval originally tuneable to 1:1. Finally, this contour plot is symmetrical on the diagonal, conforming to the musical intuition that chords with the same pitch content in a different ranking have the same HD-sum. For example, the coordinates <1:1, 16:15, 8:5> and <1:1, 8:5, 16:15> both have the same SD value, since they have the same combinatorial interval content in different orderings.

This two-dimensional space allows us to extend our use of gradient descent to find a collection of *tuneable chords,* which are visible in the contour plot above as dark purple local minima (sometimes surrounding large white areas). The collection of tuneable chords when *c*=0.0048 is found by extending a 128 by 128 point grid (roughly a resolution of 10 cents) along the space, keeping only the local minima within that resolution, and then optimizing each of these points according to the gradient descent function. This ensures that each point is as close to the local minimum as possible, while avoiding unnecessary computation by removing points that will optimize to the same local minimum. This method found a total of 857 tuneable triads within the span of one octave.

The two-dimensional space contains many local minima whose *x* or *y* values do not appear in the one-dimensional space outlined earlier in this chapter. For example, the first local minimum in one dimension is 8:7. Likewise, the first local minimum along either axis of the two-dimensional space is also 8:7. However, the two-dimensional space also contains a local minimum at the coordinates <16:15, 8:5>. While the interval 8:5 appears as a local minimum in the one-dimensional space, 16:15 does not; this is a newly available frequency ratio made possible through its simple relationship to the frequency ratio 8:5.[118] The contour plot shows the local minimum at <16:15, 8:5> with a vertical line at 16:15 to demonstrate that the 16:15 interval does not appear as a local minimum along the *x*-axis.

In the two-dimensional plot, the diagonal at $x=y$ is longer by a factor of the square root of two, measured with a Euclidean distance metric, than either axis itself. The two-dimensional parabolas used to calculate the SD value do not, however, take into account this additional distance. Therefore, the diagonal of $x=y$ contains more local minima than exist along either axis. Intuitively, we would expect this not to be the case; this implies that two voices moving in unison have a larger number of available pitches than one voice moving alone. One solution to this problem is to normalize the point <1., 1.> to the unit circle, such that the local minima along the diagonal are not spread any further apart than

[118] Marc Sabat's study of tuneable intervals by ear includes a list of intervals tuneable with the aid of a third interval, in which 16:15 appears.

the local minima along the axes. Essentially, this transforms the graph into something resembling a polar plot. The graph of this plot is given below.

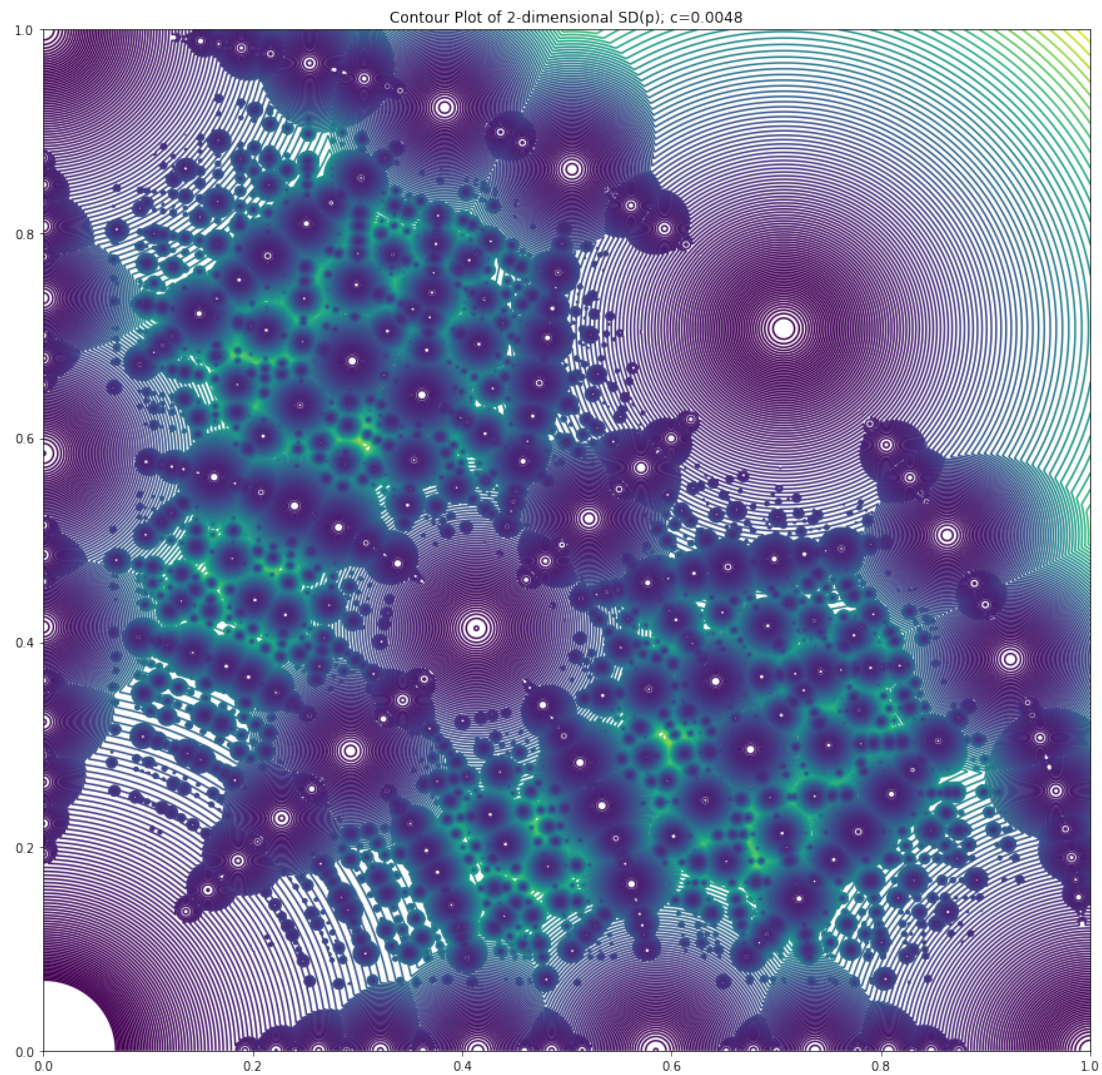

Figure 10: Contour plot of SD-value of a two-pitch aggregate, transformed to normalize the point at a unison octave to a magnitude of 1

After this polar transform of the original space, we see that points with identical combinatorial interval structures have identical magnitudes (distances) from the origin. For example, the large region representing a local minimum in the top-right of the graph represents the chord <1:1, 2:1, 2:1>, now has the same magnitude from the origin as <1:1, 1:1, 2:1>. This makes intuitive sense, as the combinatorial interval structures of these two points are the same. Additionally, the diagonal along $x=y$ now has the same contour as either of the axes. This is particularly critical in applications to tolerance and rationalization when considering a pitch aggregate. The distance between two points in the two-dimensional Cartesian projection is an error value (or measurement of "mistuning") for one chord relative to another. It is also worth noting that the polar method results in finding *far* fewer tuneable chords in the given pitch space: a total of 431 unique chords, symmetrical across the diagonal $x=y$, with 15 tuneable chords along the diagonal.

Scale generation and a "slice" of higher-dimensional space

The contour plots of the two-dimensional spaces above show clear arrangements of local minima. In the cartesian plot, "lines" of minima appear, while the polar plot has "rings" of minima. Both these "lines" and "rings" intersect with large local minima on the x and y axes, demonstrating that there are a substantial number of local minima where one of the values has a strong relationship with the origin. It is possible to extract the contour along that polar plotted ring as a one-dimensional "slice," revealing a collection of local minima with SD-values relative to *two* pitches, rather than only to an origin 1:1. In general, this one-dimensional slice in two-dimensional space reveals a contour with many more local minima.

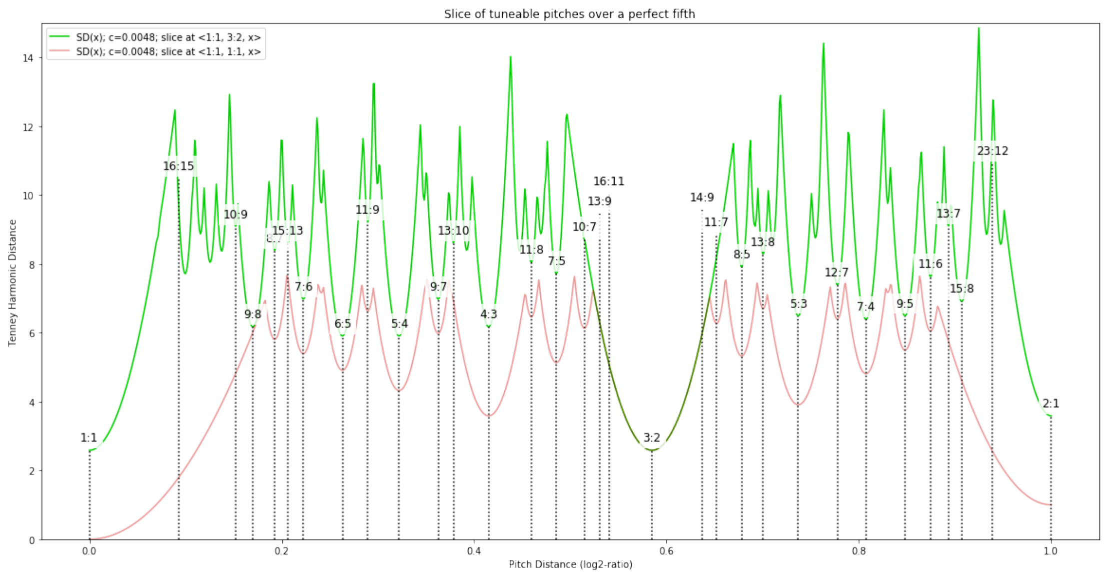

Figure 11: Demonstration of slices of intervals over a perfect fifth, 3:2

The chart above contains two measurements: the semi-transparent red line is a slice along the *y*-axis of the two-dimensional Cartesian plot chart measuring SD(*x*) along <1:1, 1:1, *x*>, while the green line measures the values of SD(*x*) along the chord <1:1, 3:2, *x*>.[119] Note that the contour of the red line is identical to that of the one-dimensional chart of single intervals above 1:1 given earlier in this chapter. However, the slice at 3:2 has considerably more local minima close to the origin 1:1. This indicates that the presence of a second pitch reduces the influence of the unison 1:1. Additionally, the values of 1:1 and

[119] I will note here that the values along the slice <1:1, 1:1, *x*> and <1:1, *x,* 1:1> are identical. Logically, the elements of a chord should be unordered; they occur simultaneously. This was one of the musical criteria that I used to evaluate the potential accuracy of the projections of the space.

3:2 in this slice are identical; this follows from the intuition that both <1:1, 3:2, 3:2> and <1:1, 1:1, 3:2> have the same combinatorial interval content. Conceptually, if the one-dimensional chart intended to show all regions of tolerance tuned to a single 1:1 pitch, this two-dimensional slice could be considered to show all regions of tolerance in relation to *two pitches*—in the case graphed above, the tolerance regions relative to a perfect fifth. In colloquial musical terms, this method of slice generation finds a *scale* which harmonizes with a *chord*.

Analytical and generative applications

In addition to theoretical research, this study of tolerance, pitch rationalization, and enumeration of local minima enables numerous generative musical applications. First, the method of pitch rationalization within variable regions of tolerance allows for a more nuanced and microtonal application of categorizing arbitrary collections of frequencies as rationally tuned intervals. The relatively large regions of tolerance around 1:1, 3:2, and 2:1 mean that rationalizing a time-series of arbitrary frequencies is likely to predominantly result in pitches on stronger intervals such as 1:1, 3:2, and 2:1, with more dissonant pitches as ornaments throughout the sequence. This musical phenomenon clearly echoes composers of early recitative such as Jacopo Peri and Giulio Caccini, whose melodic lines tend to rest on "strong" pitches, after moving through intermediate dissonances.[120] Furthermore, the continuously variable parameter of *c,* denoting the width of the parabolic curves around the regions of tolerance, allows for controlling the total number of available frequency ratios.

On the functional workings of the algorithm itself, using gradient descent to tune an arbitrary collection of frequencies to a local minimum is a clear and deterministic process that could be generative of entire musical gestures. In my composition *Terrain,* which I will discuss in the following chapter, the two trombone players repeatedly glissando from highly complex frequency ratios "downhill" to local minima, finding a three-note chord in relation to the reference point 1:1 played by the bass. In electronic works, the slope of gradient descent leads to highly interesting sonic effects, when each iteration of the gradient descent algorithm is given the same time-frame; since gradient descent initially moves very quickly downhill, we hear a quick slide followed by a very slow glissando into a tuned just intonation chord.

The method of "slicing" a higher-dimensional space along a given axis to find a collection of local minima that can be tuned relative to a chord is essentially a method for generating scales of ratios that harmonize with a sounding chord. This is useful in large part as a tool to generate compositional material; the local minima found along a horizontal

[120] In Caccini's preface to "Le nuove musiche," he writes of a "certain noble negligence [*sprezzatura*] of song, sometimes transgressing by [allowing] several dissonances while still maintaining the bass note[.]" Caccini and Hitchcock, *Le Nuove Musiche*, 44.

or vertical line can be used as melodic scales, and will contain microtonal inflections that are dependent on their harmonic contexts. My composition *Terrain* uses this method to generate collections of pitches for the string players to mimic the contours of speech, while harmonizing with the sounding chord in the brass and contrabass.

From an analytical perspective, perhaps the most novel feature of these methods is demonstrating the shifting harmonic features—the "harmonic identities," to use Tenney's term—of tempered pitches that occur even when the sounding pitches do not change. This is a step toward rationalizing pitches while taking into account their harmonic contexts. To give one simple example, the equal-tempered tritone of 600 cents rationalizes to the interval 7:5, or 583 cents, when presented as a simple interval. Adding two dimensions to complete the seventh chord harmony (with the 12-tone equal-tempered pitch vector <0, 3, 6, 8>, or a first-inversion dominant seventh chord) makes this rationalization even clearer. However, changing the other dimensions to the pitch vector <0, 2, 6, 9>, or a third-inversion seventh chord, results in pitch-class 6 reaching a rationalization of 10:7, or 617 cents. Although the tritone of <0, 6> remains the same, the contextual information given by the other pitches results in a different harmonic identity for pitch-class 6. This shift demonstrates, in essence, the shift in harmonic identity that occurs in tritone substitutions, like those commonly used by players in reharmonizations of jazz standards. There are numerous examples of these shifts of syntax, dependent on context, and it is clear to see how this method might be used to describe the contextual results of interval rationalization.

5. About *Terrain*

This section intends to be an in-depth explanation of the compositional process and musical context for my piece *Terrain,* composed for a nine-piece mixed ensemble of strings, winds, and two trombones. The work contains a number of compositional mechanisms, generating circular structures in overlapping but interdependent layers, such that the mechanisms follow some combination of their own logic and the contextual information of the rest of the ensemble. The process is one of a "machine" set in motion, where each moment is primarily contingent on the previous moment, and where the activities of instrumental groupings control one another. The orchestration itself synthesizes previous experiments in deriving musical features from extracts of spoken language, specifically my solo cello work *In the sense of transparence,* and my ensemble work *We remember not the word, but the sound of the word,* which both use methods for deriving musical parameters from speech. The rhythmic repetition of the "speaking" strings, coupled with the gradual parametric changes and overlapping structures in the various orchestration groups, are meant to diminish the sense of discrete sections of the composition, and instead give the sense of a process that is set in motion, rather than a product controlled moment-to-moment by the composer.

The title of the work references two distinct elements that informed the composition. The first is simply related to the repeated line, from a poem by George Oppen: "There are things we live among, and to see them is to know ourselves." The title

Terrain, and the use of Oppen's opening line, points to a notion of reality that is grounded in our surroundings, and in the sense that an investigation of our surroundings is in turn an investigation of our inner lives. In my work, this is tied to a belief that certain elements of music are dependent on our physical bodies, and that a theorizing and understanding of music must be grounded in physiology and psychological processes. I consider much of my work to be a construction—and an object—that exists separately from my individual expression, and believe that Oppen's outward-looking line points to a similar understanding of language.

Second, the title obliquely refers to the "terrain" that is generated when overlaying the height measurements of harmonic space onto a two-dimensional representation of pitch space; in the contour plot diagrams of harmonic space given in the previous chapter, we see a collection of peaks and valleys, and maybe a series of canyons along the diagonal and each axis. The mechanisms in this piece traverse that continuous terrain—related to the discrete lattice structure of harmonic space—in an effort to explore its possibilities. As I will discuss below, each moment of the piece is a window into a certain way of viewing this continuous terrain, whether through the one-dimensional topography of the string scales or the two-dimensional navigations of the bass and trombones. I therefore hope that this piece conveys a sense not of expression, but of exploration.

Instrumental groupings in *Terrain*

Terrain is organized as three instrumental groupings that each operate on their own internal logic, coupled with interdependencies among the groupings that determine certain aspects of the piece as time progresses. The overall parametric trajectory of the piece is then controlled by a handful of high-level algorithmic values which change throughout the course of the composition. I use the word "holarchical" to refer to these interdependent but distinct layers within the instrumental groupings; these layers determine their own music at any given moment by referencing a harmonic motion that is not determined by an organized hierarchical plan, but rather is emergent from the iterative process involved in generating the composition. The "parametric" element of the compositional structure indicates that there are fixed processes that happen throughout the course of the piece, but that the specifics of these processes are controlled by a few values that control the trajectory of the composition.[121]

There are three instrumental groupings that inform both the orchestration and structure of *Terrain*. (1) contrabass, trombone 1, trombone 2, (2) violin, viola, cello, (3) recorder, flute, and bass clarinet. Broadly speaking, group 1 moves through a series of harmonic progressions demarcated by V-I (dominant-tonic) cadences, group 2 articulates

---

[121] This gradual transition from one state to another, or from one value to another, is similar to the process that occurs in Tenney's *Spectrum* pieces, or to the parameter states drawn from hexagrams in his series of studies for six harps, *Changes*. Tenney, "About Changes: Sixty-Four Studies for Six Harps," 329. The use of parametric limits—in this case the registral limits of the trombones—to inform the beginning and ending of the piece is noted by Tenney as well. Tenney, "Form in Twentieth-Century Music," 156.

the pitch contour of Oppen's text every measure, and group 3 articulates rationalizations of the formant frequencies in Oppen's text at gradually increasing speeds.

Group 1 (harmonic structure) follows a set of iterative steps, and does not contain a global parameter shift. The two trombones start in their upper register of the 12^{th} harmonic, and after each cadence one of them creates a split-tone and moves one partial lower. After the first trombone reaches the bottom of its range, the second repeats the process until both are at the lower end of their range. These chord progressions determine the circular harmonic structure of the composition, and are used by groups 2 and 3 to inform their melodic material.

Group 2 (melodic chants) is globally controlled by three parameters which change throughout the course of the piece, and which affect both the precision of the pitches of the reconstructed speech contour and the scaling of the speech contour itself. The first two parameters control the scaling of the speech contour itself, in a multiply-add operation. Throughout the piece, the magnitude of the speech contour is gradually scaled from monotonic (flat contour) in the lowest possible register, to a three-octave range, and then again monotonic at the highest possible register. The third parameter is the width c of the parabola used in generating scales of possible pitches for each measure. The scales are generated using the "slice" method described in the previous chapter, with a slice of a three-dimensional space in which the first two dimensions are fixed to the values of the two trombones and bass. In musical terms, the strings draw their pitches from a scale intended to harmonize with the sounding pitches from group 1. Throughout the piece, the value of the coefficient *c*, controlling the width of the parabolic curve described in the previous chapter, decreases linearly from 0.5 to 0.01. The result is that initially the "resolution" of the strings is very coarse, but the gaps gradually fill in to form a full gamut of pitches.

Group 3 (formant chorale) plays the center formant frequencies of the vowels, as derived from a phonetic analysis described below. This succession of sixteen frequencies is stretched in time and presented longest-to-shortest in 10 iterations, in a reverse Fibonacci series, so that iteration 1 is 55 measures, iteration 2 is 34 measures, iteration 3 is 21 measures, and so on until iterations 9 and 10 of 1 measure length each. Each of these iterations then contains 16 onsets (one for each vowel syllable), with the group of three instruments in rhythmic unison. Furthermore, these formant frequencies are quantized to pitches drawn from the same scale as the strings; in this way, they also begin with a relatively coarse resolution, and end with a much finer grain of pitch. Rhythms throughout are quantized to the nearest sixteenth note and have no relationship to the measured time, with the exception of the final two iterations (one measure each) which are accompanied by text and performed in rhythmic unison with the strings.

Contrabass and brass: harmonic structure

The harmonic structure and cadence of *Terrain* is tied to the physical properties of the trombones and contrabass. Through the piece, the following sequence occurs:

1. The contrabass plays a pitch, with the piece starting on a low C.

2. The two trombone players glissando (informed by gradient descent) to bring their pitches "in-tune" with the contrabass.
3. The contrabass moves in one of three ways:
    a. If the contrabass is playing a G, it moves to a low C in a V-I cadence.
    b. If the contrabass is less than a whole step away from G, it moves to a G.
    c. Otherwise, the contrabass moves to the "most harmonic" pitch that is "downhill" from the current pitch.
4. There is then a second condition:
    a. If the contrabass is sounding a C (and just played the V-I cadence) one trombone plays a split tone crescendo with the partial below their current partial, and the ensemble continues with that trombone on the lower partial, starting again from Step 2.
5. Steps 2-4 are repeated for each measure.

This circular structure meanders, approaches the V by step or half-step, and then repeats the V-I cadence, in a pattern that is continuous throughout the whole piece. The result is a three-note chord that gradually morphs throughout the piece, moving by mostly microtonal deviations in the trombones to keep in-tune with the overall harmony.

There are a few moments of decisions relating to specific steps above. In Step 2, the trombone players' glissando occurs simultaneously in two dimensions (one dimension for each pitch offset from the contrabass's 1:1), using the gradient descent optimization technique described in the previous chapter. At times, one trombone will be static, but with particularly odd harmonic root movements we see that both trombones glissando into neighboring chords. An example from mm. 19-20, early in the piece, is given below, in which the trombones' perfect fifth on F-C slides into the 7th and 11th harmonics of the G in the contrabass. This figure in particular occurs repeatedly, as the contrabass frequently moves from the inflected G-flat (64:45) to the G-natural (3:2).

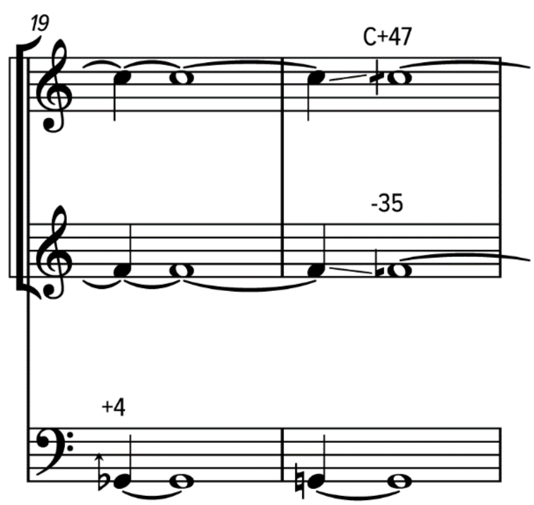

Figure 12: *Terrain*, mm. 19-20

In Step 3, the program evaluates the HD-sum of the two trombones in combination with every possible pitch from a subset of intervals available for a root-progression. This subset of intervals only includes intervals with ratio representations *p/q,* where the value of *q* is greater than the value of *p* when the prime number 2 is factored out. It is a pitch-class projection space of harmonic space (ignoring octaves),[122] in which the root motion is constantly moving toward values that are simpler in this pitch-class projection space. Available ratios are 4:3, 8:7, 6:5, and others in which the "higher primes" tend to be in the denominator; unavailable movements for the contrabass are 3:2, 5:4, 9:8, and others in which the numerator is more complex.[123] The term "downhill" is intended to convey the sense that root movement tends to move in a particular direction, toward simpler ratios, and was proposed by Tenney in an interview with composer Brian Belet in 1987.[124] Considering each tone as made up of harmonic partials, moving in an "uphill" direction does not introduce any new odd partials; it is merely moving to a higher partial of the current tone. However, downhill motion, such as from C to F, moves "downward" in the harmonic series, introducing new harmonic content.[125] The repetition of this procedure results in a series of very small intonational adjustments, with demarcated boundaries by cadences in the contrabass, that anchor the harmony of the piece.

---

122 Tenney describes this term (his term) as reflective of "a *categorical* difference—a difference in *kind,* not just in degree—between the octave-relation and all other harmonic relations." Tenney, "The Structure of Harmonic Series Aggregates," 266. See also Diana Deutsch's work on octave equivalency and pitch-height perception, particularly with regards to timbre, in Diana Deutsch, Kevin Dooley, and Trevor Henthorn, "Pitch Circularity from Tones Comprising Full Harmonic Series," *The Journal of the Acoustical Society of America* 124, no. 1 (July 1, 2008): 589–97, https://doi.org/10.1121/1.2931957.
123 In Barlow's harmonicity metric, "downhill" would correspond to the set of negative harmonicity values.
124 Belet and Tenney, "An Interview with James Tenney," 463.
125 It is worth noting that, although Tenney's *Changes* does include 3:2 as a potential root progression, the probability of a 4:3 root progression occurring is three times as high: 0.3 for the ratio 4:3, versus 0.1 for the ratio 3:2.

Strings: melodic chants

Throughout *Terrain*, the string players repeat a melodic line to the rhythm of the phrase, "There are things we live among, and to see them is to know ourselves," from the beginning of George Oppen's poem *Of Being Numerous,* as he read it on-air on KPFA Berkeley on August 22, 1968.[126] The particularities of the melodic line change throughout the course of the piece in a number of ways, but the contour of the line is either the same or flattened from the original speech sample.[127] The performers are asked to "breathe together," and to deviate slightly from the conductor in the accents of their speech-like rhythms.

After the harmonic structure and total duration of the composition is generated through the bass and trombones, each measure then has a specific three-note harmony that is sustained by group 1. The program then generates a "scale" for the strings and winds based on the factors of (1) the sustained chord in the bass and trombones, and (2) the current value of the parameter *c,* which decreases linearly from 0.5 to 0.01 throughout the course of the piece. The scale generation method used is described in the previous chapter; the two-dimensional contour plot that describes the trombones and bass is extended to three dimensions, and a "slice" is taken from only that third dimension. The result of the gradual change in the parameter of *c* in the strings is that the strings' melodic resources

[126] "PennSound: George Oppen," accessed January 30, 2020, https://writing.upenn.edu/pennsound/x/Oppen.php.
[127] "Contour" here is used to indicate successive up-down-same relations of the melodic line, as in Larry Polansky, "Morphological Metrics," *Journal of New Music Research* 25, no. 4 (December 1, 1996): 289–368, https://doi.org/10.1080/09298219608570710.

begin with a very coarse scale—initially, mostly fifths and octaves, relative to one of the pitches in the three-note sustained harmony—and gradually move toward a scale full of microtonal inflections.

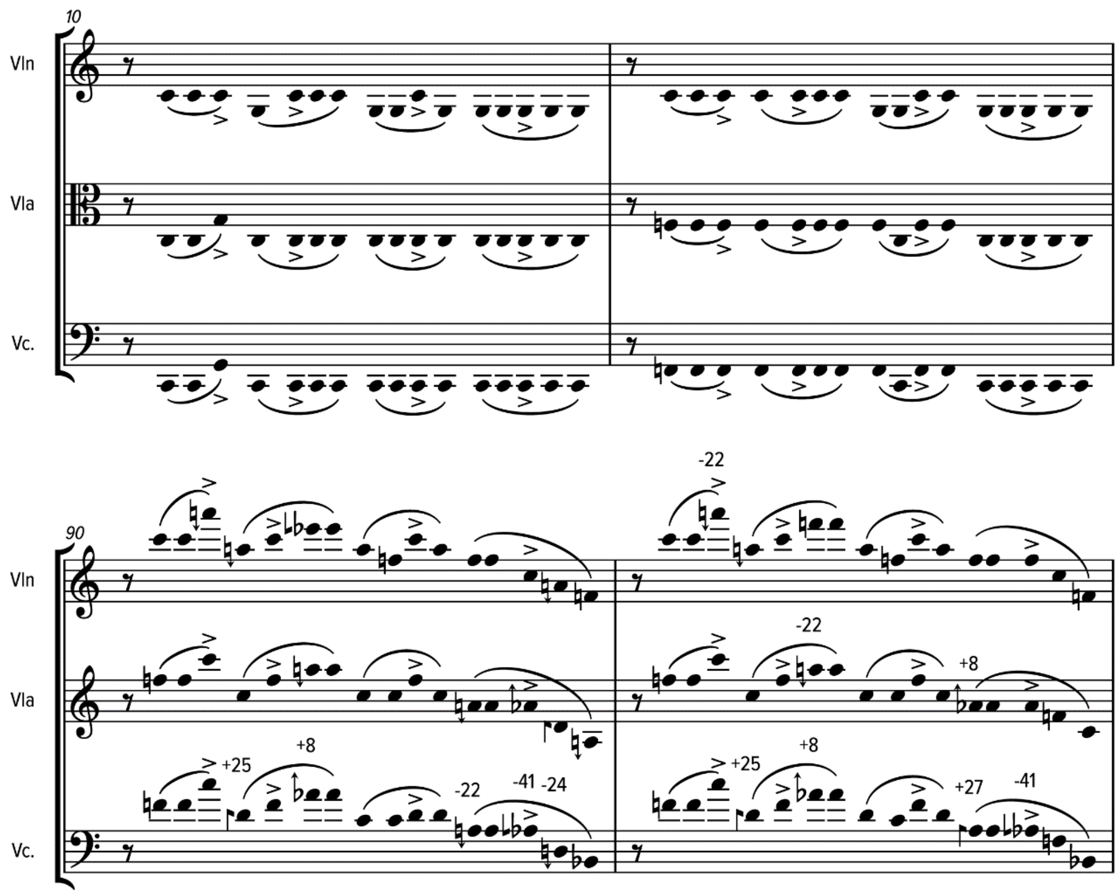

Figure 13: Excerpts from the string parts of *Terrain* at mm. 10-11, mm. 90-91

This parametric progression in the strings draws on my past work, *In the sense of transparence* for solo cello, in its gradual shift from "monotonic" lines to higher-resolution realizations of the speech contour. I found in that work that the cellist, Séverine Ballon, tended to perform the monotonic lines with more rhythmic freedom—that is, the lines had less musical "baggage," and likewise put less mental and physical strain on the performer—and that these more fluid performances of the repeated, monotonic lines tended to sound more "speech-like." In *Terrain,* as in *In the sense of transparence,* I aimed to give a gradual transition from that which could be perceived as speech—tied to a human body, and connected to expression—to that which might be considered abstract melody based on ratios.

Although the scale as a whole shifts every measure, sometimes into obscure harmonic territory, there tend to be many repeated notes from measure to measure due to the propensity for the bass and trombones to stay, generally, in a single harmonic region. However, the occasional movement by fourth or fifth can generate almost "tonal" harmonic sequences. An example is given below, where in mm. 68-70 the bass shifts by fourth from C to F, and the trombones move from octaves on C to what jazz harmony would describe as a first inversion D-flat major seventh chord with no fifth. The strings then move from mostly a harmonic series on C to a collection of pitches harmonizing with the new three-note chord. As a result, the violin's E (5:4 relative to the bass) in m. 69 resolves to an F (2:1 relative to the bass) in m. 70. These resolutions are a function of the "strong" pull of the harmonic root in any given section.

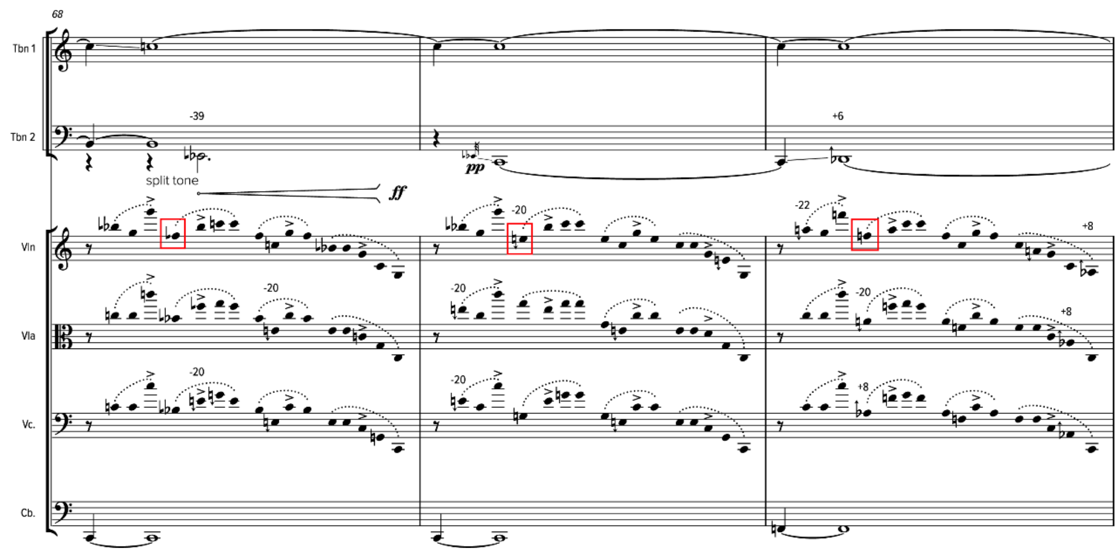

Figure 14: *Terrain*, mm. 68-70, demonstrating the slight shifts of intonation and pitch as the grounding harmony changes

Winds: formant chorale

The final element of the ensemble orchestration is what I refer to as the "chorale" in the winds. The wind players in *Terrain* perform music derived from the same George Oppen voice sample as the string players, but drawn from the vowel formant frequencies of each syllable rather than from the fundamental frequency (F0) of the vocal intonation. This method of orchestration was simpler than that used in *We remember not the word, but the sound of the word,* in that I assigned frequencies directly based on the closest center frequency of the target formant, rather than using statistical feedback to determine harmonics over a fundamental frequency, in addition to a few other slight implementation changes.[128]

In *Terrain,* as in my previous piece, the wind players are each assigned one of the first three formants in the range of the instrument in the order of bass clarinet, flute, and sopranino recorder. These instruments were chosen from the possible doublings of the ensemble because of their relatively pure harmonic content (especially at lower dynamic levels) and were therefore treated as though they were single frequencies.[129] This mode of speech-derived orchestration is informed by "sine wave speech" experiments, in which as few as three sine waves directly articulating the formant center frequencies can be heard as speech with vowels identified by a listener.[130] This method is distinct from the source-filter model used in linear predictive coding algorithms for electronic speech synthesis, which inspired the techniques used in my composition *We remember not the word, but the sound of the word*.

---

[128] This implementation used the Python library Parselmouth, created between the composition of *We remember not the word, but the sound of the word,* and the composition of *Terrain*, which greatly simplified the code. Parselmouth calls the C/C++ source code of the program Praat directly, which avoided using my own partial implementation of the formant tracking algorithms used by Praat. This allowed me to more quickly prototype and graph my parameter fine-tuning in Praat, and then copy those parameters directly to Python code to run on larger batches of sound files. Y. Jadoul, B. Thompson, and B. de Doer, "Introducing Parselmouth: A Python Interface to Praat," *Journal of Phonetics* 71 (2018): 1–15, https://doi.org/10.1016/j.wocn.2018.07.001.

[129] This assumption follows other speech-derived orchestration methods described in the previous section, especially those by Tenney. Wannamaker, "The Spectral Music of James Tenney."

[130] R. E. Remez et al., "Speech Perception without Traditional Speech Cues," *Science* 212 (1981): 947–50.

The instruments' ideal target pitches, derived from the formant analysis, were then rationalized to the same scales used in the rationalization of the string melodies for each measure, and this was repeated for each repetition of the chorale throughout the piece with the linearly decreasing value of *c* given from the string scales. The first three entrances of the winds are given below. Note that the entrances are relatively close to one another in pitch space, but that as the piece progresses the microtonal inflections allow for more precise realizations of the intended frequencies. For example, the bass clarinet and recorder play something close to a C for the first two entrances (mm. 1 and 54), but the finer grain later in the piece allows for microtonal adjustments during their entrance on m. 87. For this third entrance, the bass clarinet adjusts the pitch flat to a B (5:4 above G), and the recorder adjusts to a D that is over a quarter-tone flat (7:6 above the B in the bass clarinet). The three notes in this entrance of the winds are related to one another by the relationship <1:1, 7:6, 4:3>, but the 1:1 in this case is 15:8 in relation to the bass. Harmonically, moments like this afford the sense of multiple distinct sub-groupings, as the winds fuse as a unit in a different harmonic zone from the harmony of much of the rest of the ensemble. In addition, the wind players are asked to sustain these pitches, once tuned, for subsequent measures even when the rest of the ensemble has moved on to new harmonies, including the string players drawing on the same scale. A more robust implementation of this chorale would involve executing the gradient descent algorithm to tune the entire wind chord, after "fixing" the two dimensions of the trombones. This would, however, require calculating a five-dimensional space, which is computationally prohibitive in the current implementation of the tuning algorithm with the resources at hand; perhaps future refinement of the implementation will allow for this possibility through further optimization.

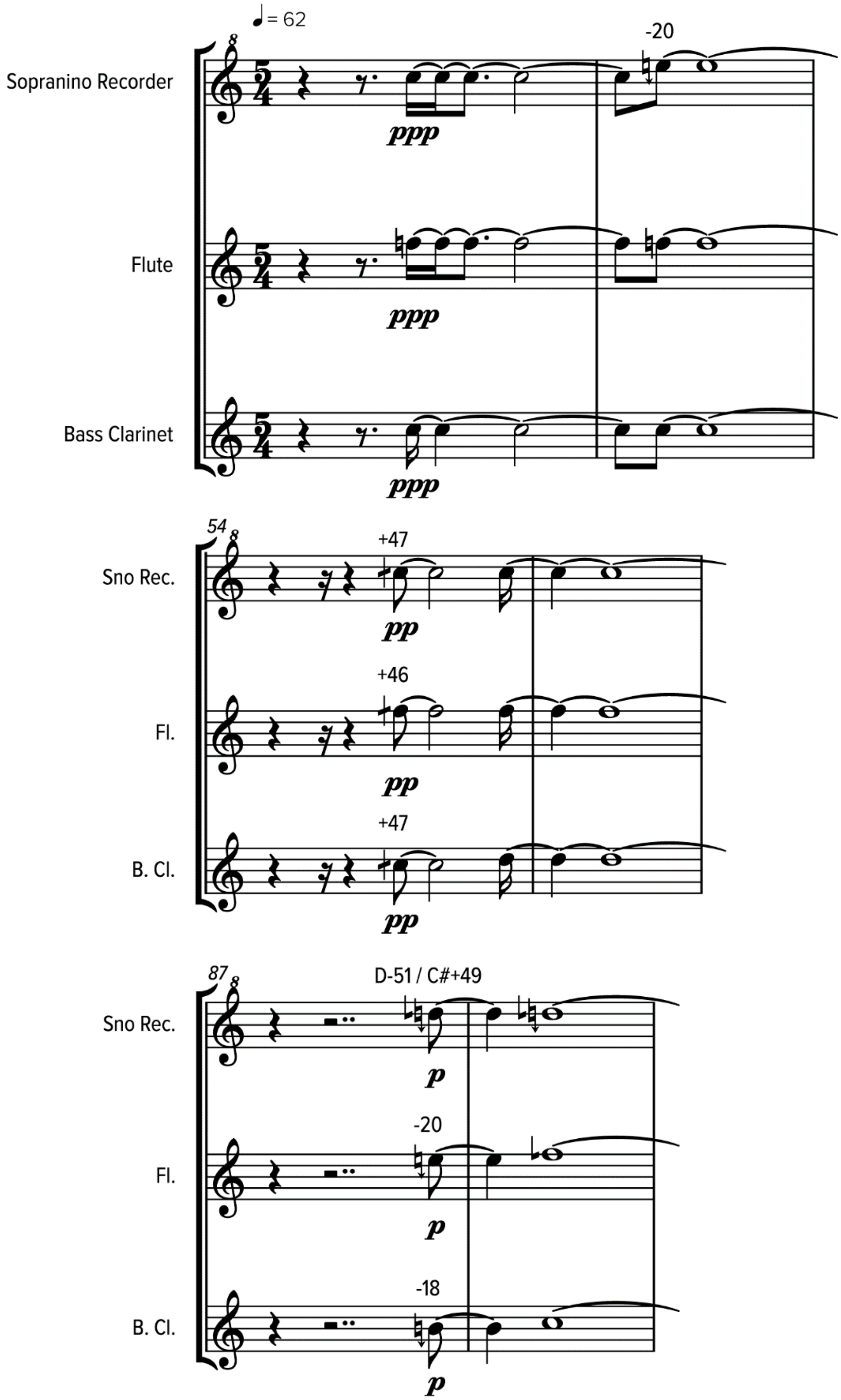

Figure 15: Wind entrances at mm. 1, 54, 87 of *Terrain*

The rhythmic aspect of the winds' formant chorale is simple: the onsets of each vowel in the sample of the George Oppen sample are translated directly to onsets in the winds, and quantized to the nearest sixteenth note. These durations, given by the vowel onsets, are stretched to various lengths before quantization, at the sectional entrance durations above, until finally they are presented as freely "speech-like" notation in the final two measures. At this ending of the composition, the articulation of the melodic lines as vertical harmonies is much less relevant; furthermore, the string players are all playing monophonic lines at the top of their ranges, and the ensemble as a whole is at a fortissimo dynamic level. The vertical harmonies in the winds, especially the sopranino recorder, are discordant, essentially freely microtonal melodic lines that bear little relation to the sounding harmony.

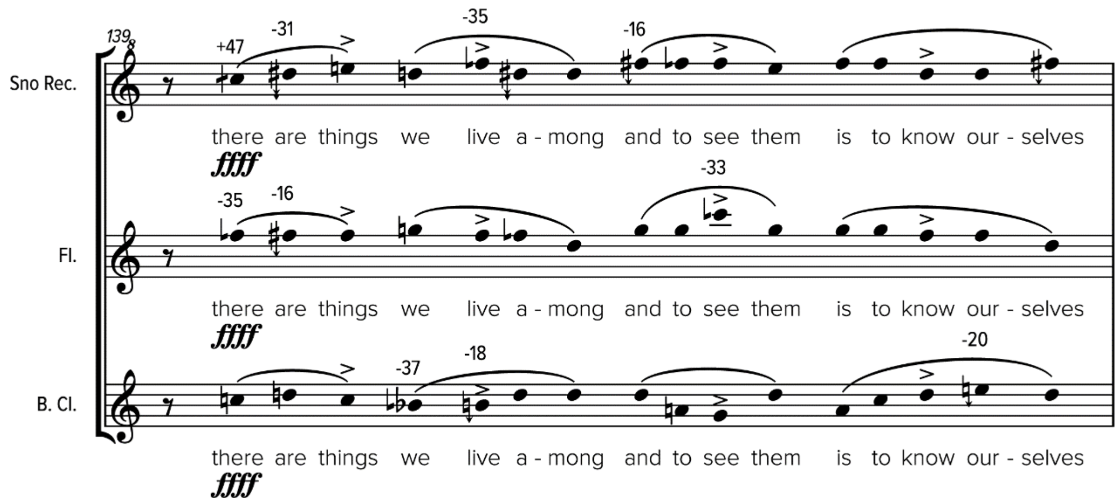

Figure 16: *Terrain*, m. 139 in the winds to show essentially free microtonal lines

Holarchic organization and flux

The concept of holarchy manifests in *Terrain* through its structure of organization and orchestration. Furthermore, the term *holarchic flux*, which aims to synthesize the linguistic aspects of music with the cybernetic concepts of feedback and information, can be used to describe the perception of the composition. I use the word "holarchy" to describe my work—recognizing that it might be, in part, an aspirational description for a composition involving a conductor—because holarchies are *complex* but not *random*.[131] That is, the results are difficult to predict due to the interactions between independent actors, but the behavior itself is fundamentally deterministic. It is important to note that the term holarchy has been used in a few different ways throughout this discussion. The first is the mechanistic metaphor—that is, holarchy describes the orchestration, and the players, who operate in nested groupings—while the second is the perceptual nature of holarchy, as used by Tenney to describe the way in which musical form occurs as a collection of nested groupings. Tenney primarily uses the second sense in his writing, for example his description of "holarchical perceptual levels"[132] in a discussion of his composition *Changes*, yet the feedback-oriented nature of the procedural mechanisms in *Changes* evokes the first sense as well.

[131] This sense of analysis is taken from Abraham, "Mathematics and Evolution: A Manifesto."
[132] Tenney, "About Changes: Sixty-Four Studies for Six Harps," 327.

In *Terrain*, there is a perceptual interference between the relatively clear and static divisions in the harmonic syntax—that is, the repeated V-I cadences and relatively circular motion in the contrabass—and the perception of the harmony in the strings. Throughout the piece, there is a sense of flux between the articulation of the harmony and the "speech-like" expressive mode as it occurs in the strings, through a tension between the horizontal nature of the expressive line and the vertical nature of the harmonic content.

Rationalization and articulation

Terrain utilized the chord and scale rationalization algorithm described in the previous chapter, and worked to create an emergent harmonic syntax through these quantitative (as opposed to symbolic) mathematical models of harmony. First, the harmonic algorithm driving the brass and contrabass directly encodes a fundamental feature of harmonic syntax through the repeated V-I cadences, in order to examine the potential for an emergence of syntactical harmonic indicators such as the leading tone, given here in mm. 48-49. The nature of the gradient descent tuning algorithm meant that the trombones would always find the nearest local minimum in pitch space for each cadence. Loosely, this models the intuition that harmony might be driven by voice leading toward minimizing harmonic distance, rather than by the contrapuntal concerns of tonal syntax.

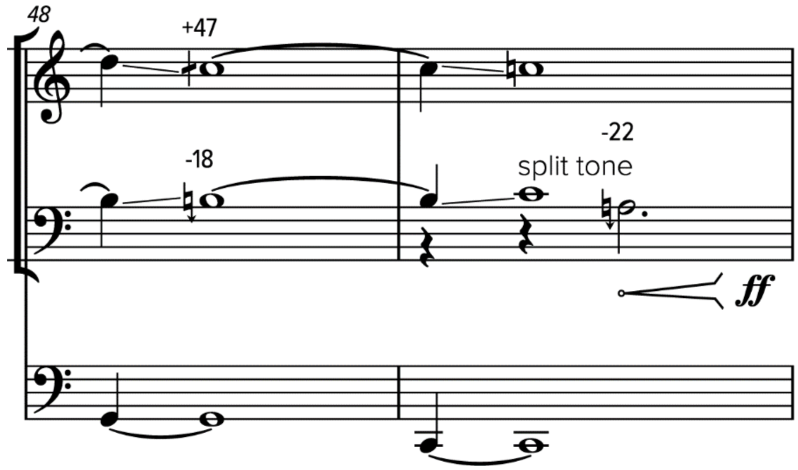

Figure 17: *Terrain*, mm. 48-49, showing brass and contrabass leading tone motion

The occasional emergence of these tonal syntactical articulations creates an interesting effect in the understanding of a listener accustomed to tonal music; it reinforces and sometimes creates boundaries and articulations. A mechanism that can *evoke* these syntactical associations without actually encoding them in the logic of the composition helps to set apart the composition as an autonomous object—it distances the composition itself from the specific desires of the composer—and it reduces the sense of composer-listener "communication" in the work. Tenney notes in his discussion of *Changes* that "the 'random walk' character of the series of root-progressions should gradually be 'focused' in such a way that each study would end with a dominant-to-tonic progression to the same root PC [pitch class],"[133] and also notes that he encodes these exact rationalizations of the V-I progression in the logic of the program.[134] In the future, I am interested in further

[133] Tenney, 335.
[134] Tenney, 340.

developing a more complex consideration of harmonic motion that does not code the V-I motion through conditional logic, but rather refines this notion of a "pull" in some harmonic space to create the repeated V-I cadences. I hope to avoid encoding these progressions as absolute logic, and rather to develop a sense of root-motion that draws on my recent research. This could open the door to the creation of cybernetic improvisation devices, or any number of other intelligent musical systems that potentially interpolate between disparate systems of harmonic syntax, rather than encoding specific systems as a form of conditional logic.

Melodic lines and harmonic syntax

In each measure the strings and winds draw their pitches from a scale tuned to the sounding harmony of the contrabass and trombones; this uses the "slice" method described at the end of the previous section, with more pitches permitted throughout the piece through a continually decreasing value of the parabolic width coefficient c, also described in the previous section. After each V-I cadence, this sounding scale harmony is relatively similar and, as a result, these measures can be compared to one another to evaluate the general "resolution" of the sounding pitches. These sounding pitches are generally close to a harmonic series on the low C of the contrabass, with a few exceptions for instances when the trombones are playing pitches other than C or G. To illustrate this increasing resolution, I have extracted three instances of the complete scale available to the strings and winds throughout the piece. These scales generally have wider steps in the lower octaves, and smaller steps in the upper octaves, roughly corresponding to the distribution of the harmonic series. Early in the piece, when the strings primarily arpeggiate triads and fifths, there is a clear sense of harmonic progression from measure to measure. However, as the piece progresses and the scales become intensely microtonal in nature—sometimes straying far from the harmonic series on the C of the contrabass—the melodic lines begin to lose their "harmonic" nature, and start to approach a kind of "speech," toward the primacy of the horizontal motion, and finally reaching a kind of scream in the upper registers.

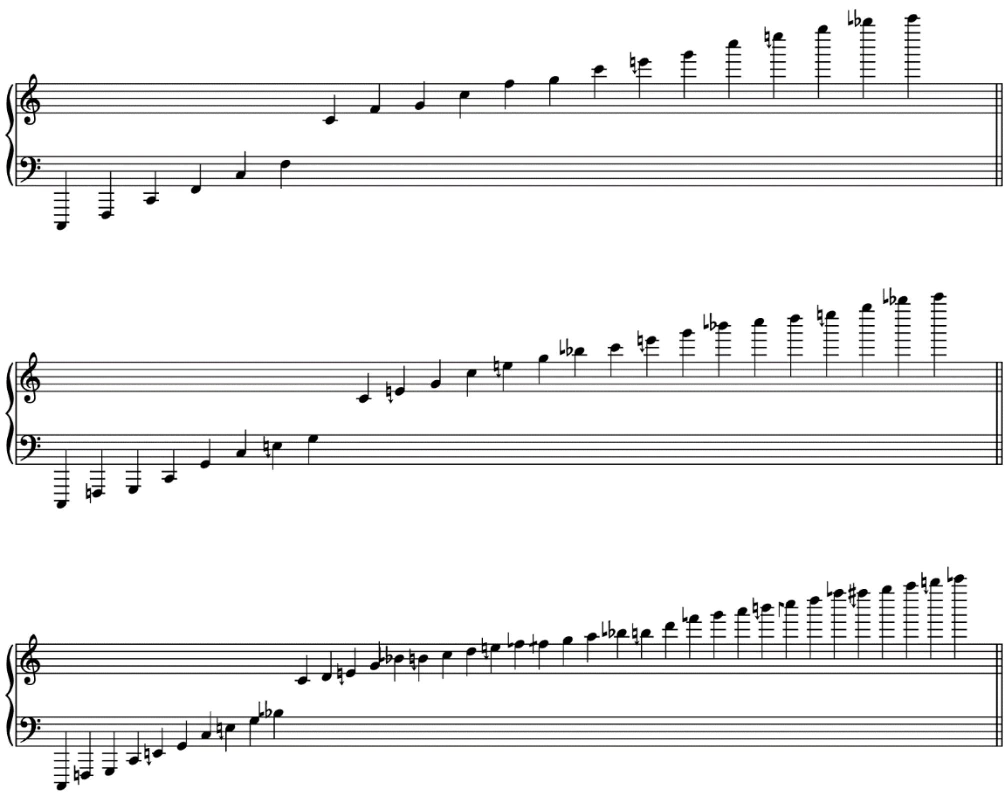

Figure 18: *Terrain*, scale rationalizations for potential pitches in strings and winds at mm. 3, 69, and 117

The string lines at each of these measures used as examples produce highly complex vertical harmonies, such that as the harmonies increase in complexity the voices themselves tend to separate from one another and become independent from the harmony rather than subservient to it. Broadly speaking, as the horizontal lines increase in intricacy, and as the harmony increases in complexity, these lines become less "musical" and more "speech-like." The result is that harmonic progression exists on the level of the measure, rather than the note, and emerges from the repeated articulations of the strings in conjunction with the contrabass and brass. There is not necessarily any "counterpoint" to speak of in these repetitions, but it still maintains a tension between the musical-syntactical structure of harmony and the expressive structure of speech.

Adorno writes about the syntactical nature of harmony in his essay "Schoenberg and Progress" that:

> What remains insufficient in Schoenberg's twelve-tone music is harmony—the opposite of the problem in Bach, where the harmonic schema sets limits to the independence of the voices, limits that are transcended only in the speculation of *The Art of the Fugue*.[135]

This statement occurs in the context of a discussion of harmonic and melodic independence in twelve-tone music, but might equally describe the nature of transcribing the pitch contours of the spoken voice for instruments moving through harmonic territory.

[135] Adorno Adorno Theodor, *Philosophy of New Music*, trans. Robert Hullot-Kenter (Minneapolis: University of Minnesota Press, 2006), 72.

Adorno's observation of Bach is that the vertical fusion provided by Bach's harmonic schema (also at the level of the measure, or at least the beat) led to a higher-level syntax, while lessening the independence of the voices—that is, the voices are not entirely independent expressive entities but are subject to a musical syntax. Schoenberg's music, which is tied intimately to expression, does not have the higher-level syntax of tonal harmony and therefore reverts to what Adorno terms "pre-critical" forms, articulating these forms through rhythmic repetition. He writes of the repeated rhythmic figures in twelve-tone composition that, "Whenever those rhythmical formulae make an appearance, they herald correspondingly formed components, and it is these correspondences that raise the specter of precritical forms—but certainly, only the specters."[136] Yet, the concept of independence of the voices is one of complexity and degree; it is possible to hold the harmonic syntax and speech-like expression in flux, and to continually oscillate between the two.

[136] Adorno, 75.

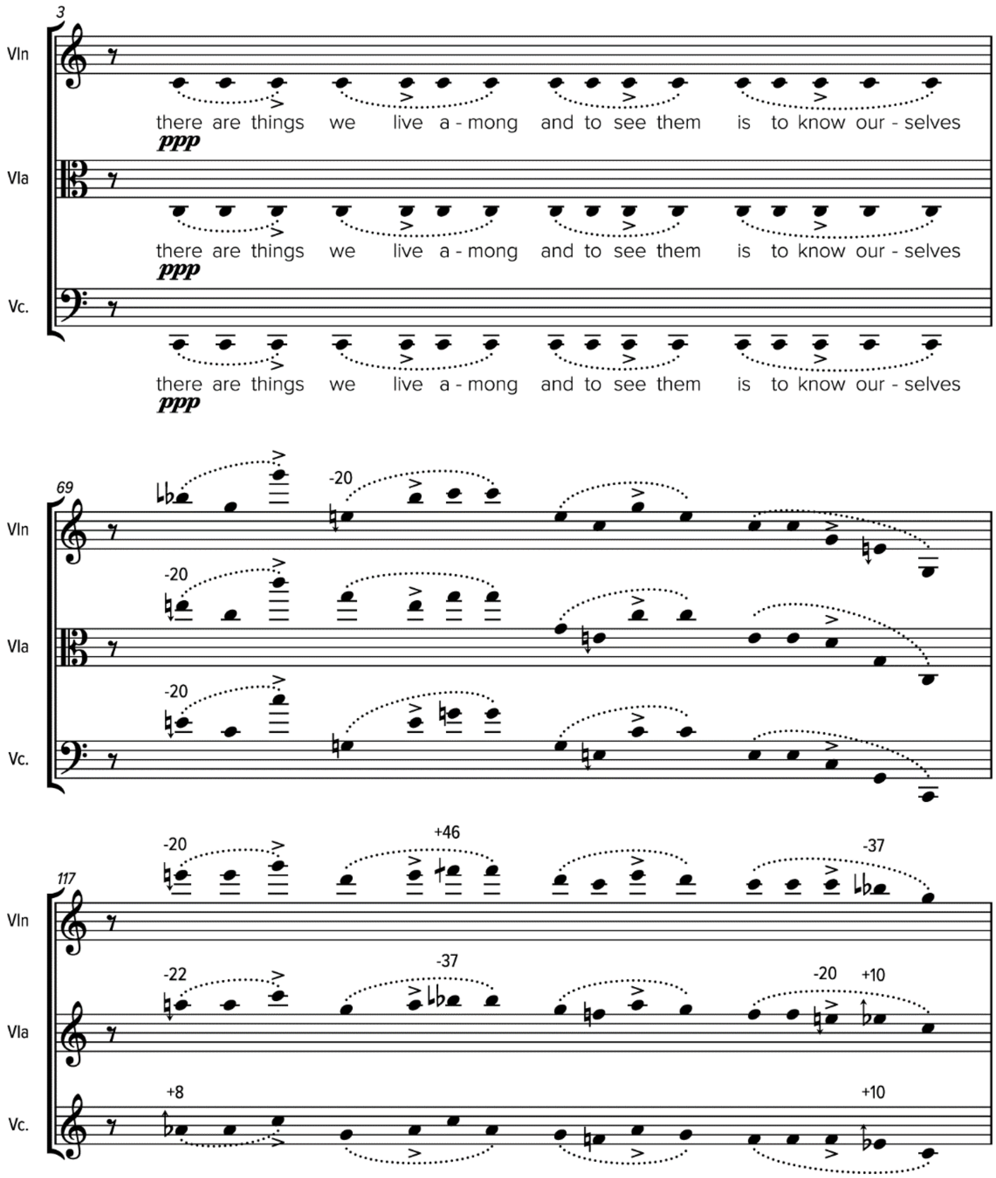

Figure 19: *Terrain*, mm. 3, 69, and 117, showing the string articulations of the harmonies over the low C in the bass, after the V-I cadence

Holarchy and indeterminacy

John Cage described music that is "indeterminate with respect to its performance," in contrast to indeterminacy with respect to composition.[137] Yet Cage's notion of indeterminacy in both the compositional and performance processes is fundamentally tied to the ability of the composer to predict the outcome. In other words, there is a distinction to be made between that which is *deterministic* and that which is *foreseeable;* this is clarified by Cage's intention of removing his own authority from the work, often at both the levels of composition and performance. Framing indeterminacy in terms of the composer's foresight or moment-to-moment intention opens the door for methods of composition that might be considered "indeterminate" through an understanding of complexity, rather than stochasticity. In this sense, indeterminacy becomes an epistemological concern.

This epistemological framing is my understanding of indeterminacy in my own compositional process: the construction of unpredictable and complex interlocking systems, through which the composer cannot foresee the outcome. I approach the notion of holarchy in music with a granularity similar to Cage's approach of indeterminacy; there is a difference between a work that is holarchic with respect to its composition, and one that is holarchic with respect to its performance. Tenney's notion of holarchy can be framed in many of his acoustic works as one that is holarchic with respect to their composition; holarchy with respect to performance is perhaps a fruitful way of understanding Kenneth

[137] Cage, *Silence*, 35–40.

Gaburo's *Twenty Sensing Compositions* or a number of Pauline Oliveros's *Sonic Meditations,* in which individuals are treated as interlocking—rather than essentially independent—performing agents. Furthermore, examples exist of human-machine holarchic relationships in work such as Larry Polansky's *B'reysheet,* in which the performers of the voice and computer parts essentially execute independent parametric changes in live performance, while still listening to one another.

Terrain is holarchic with respect to its composition, as the agents in *Terrain* are not the individuals as performers—rehearsal time constraints and the concert setting of the occasion would make such an arrangement highly difficult[138]—but rather the various sub-groupings and mechanisms used in the process of composition. It is holarchic with respect to its composition in large part because of the interlocking groups that operate independently with their own internal parametric changes throughout the piece, while still interacting with the results of one another's processes. It is a composition created beginning-to-end by interlocking machines, rather than top-to-bottom with nested hierarchies and sections.

[138] I explored this method of interlocking systems with respect to performance in a previous piece, *Topology A/A*, composed for wandering violist and four electric guitars, before my graduate studies.

6. Conclusion

A work of music is a form of symbolic communication—but it is not just that. Each musical work creates its language, immanent within the work, but with other existing historical works as referents. In the case of musical work that directly uses *linguistic material* (whether structurally, in its compositional process, or verbally in its performance) that linguistic material brings its own history and syntax to bear on any encounter with the work. Yet, the music itself carries with it a form of eloquence that is based in a syntax internal to the work; this syntax relies on a translation of the material of the work into a symbolic form. The theoretical thread of James Tenney's writing on harmony—that of harmonic space, and the question of the harmonic identity of tempered intervals and chords—can be considered as a necessary precursor to a discussion of the linguistic element of harmony in music. Furthermore, Tenney's quantitative methods lend themselves to something not fully explored in his theoretical work: the potential for the articulations and holarchical groupings of harmony to be *constantly in flux*.

The quantitative theoretical considerations outlined in this dissertation point the way toward a consideration of harmonic identity and chord rationalization based on continuous values; these algorithms have a "gray area" built into their construction. The philosophical discussion in this essay seeks to orient the development of these quantitative methods toward an idea of *holarchic flux* in the syntax of the work. A quantitative method for describing holarchic flux as it pertains to harmony is a future project, but the quantitative, continuously-valued calculation of harmonic identity given in this dissertation is a necessary precursor for that future work.

I intend for these compositions to be *enactments* of the theoretical and aesthetic concepts outlined in this work; they are instances and experiments, rather than fleshed-out theories and formulae. The work itself—like language—finds its autonomy in the ability of the listener to regard it as something separate from its creator. The message, in an aesthetic sense, is not a medium for communication but rather material that a listener interacts with in a specific historical and interpretive context. The flux inherent in this material, and the way in which the material blurs its own articulatory boundaries, is constantly generative of new meanings. These compositions aim for the space between potential poles of interpretation; they are objects created through a process, but irreducible to that process. The flux of their meaning exists in their interactions with the cultural context and the historical presuppositions of the listener. They are constructions, existing within and acknowledging the history of their related languages even as they hold their own language in suspension.

Supplemental Files

1. Score: *How to Get There From Here*

2. Score: *We remember not the word, but the sound of the word*

3. Score: *Terrain*

Bibliography

Abadi, Martín, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, 2015. https://www.tensorflow.org/.

Ablinger, Peter. "Quadraturen." Accessed March 16, 2020. https://ablinger.mur.at/docu11.html#principles.

Abraham, Ralph. "Mathematics and Evolution: A Manifesto." Edited by David Dunn. *International Synergy Journal*, no. 3 (1986): 14–27.

———. "Mathematics and Evolution: A Proposal." Edited by David Dunn, no. 5 (1987): 27–45.

Adorno, Adorno, Theodor. *Philosophy of New Music*. Translated by Robert Hullot-Kenter. Minneapolis: University of Minnesota Press, 2006.

Adorno, Theodor. *Aesthetic Theory*. Edited by Gretel Adorno and Rolf Tiedemann. Translated by Robert Hullot-Kenter. 2nd ed. London: Continuum, 2002.

Adorno, Theodor W. "Music, Language, and Composition." Translated by Susan Gillespie. *The Musical Quarterly* 77, no. 3 (1993): 401–14. https://doi.org/10.1093/mq/77.3.401.

Ames, Charles. "Statistics and Compositional Balance." *Perspectives of New Music* 28, no. 1 (1990): 80–111. https://doi.org/10.2307/833345.

Atal, B.S., and Suzanne L. Hanauer. "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave." *The Journal of the Acoustical Society of America* 50 (1971): 637.

Barlow, Clarence. "Bus Journey to Parametron." *Feedback Papers* 21–23 (1980): 135.

———. *Musica Algorithmica*. CD. World Edition, 2018.

———. "Two Essays on Theory." *Computer Music Journal* 11, no. 1 (Spring 1987): 44–60.

Barrett, G. Douglas. “Between Noise and Language: The Sound Installations and Music of Peter Ablinger.” *Mosaic: An Interdisciplinary Critical Journal* 42, no. 4 (2009): 147–164.

Belet, Brian, and James Tenney. “An Interview with James Tenney.” *Perspectives of New Music* 25, no. 1/2 (1987): 459–66.

Blonk Performs Ursonate with Real-Time Typography. Accessed May 20, 2020. https://www.youtube.com/watch?v=rs0yapSIRmM.

Bök, Christian. *Eunoia*. Toronto: Coach House Books, 2001.

Caccini, Giulio, and H. Wiley Hitchcock. *Le Nuove Musiche*. 2nd edition. Recent Researches in the Music of the Baroque Era 9. Middleton, Wis.: A-R Editions, 2009.

Cage, John. *Silence*. Middletown, CT: Wesleyan University Press, 2011.

Cage, John, Melody Sumner, Kathleen Burch, and Michael Sumner. *The Guests Go in to Supper*. Oakland, Calif: Burning Books, 1986.

Cherry, E. Colin, Morris Halle, and Roman Jakobson. “Toward the Logical Description of Languages in Their Phonemic Aspect.” *Language* 29, no. 1 (1953): 34–46. https://doi.org/10.2307/410451.

Christian Bök Reads from Eunoia. Accessed May 20, 2020. https://www.youtube.com/watch?v=zhQjfr8b9Wg.

Cohen, Paul, Niall Adams, and Brent Heeringa. “Voting Experts: An Unsupervised Algorithm for Segmenting Sequences.” *Intelligent Data Analysis* 11, no. 6 (December 2007): 607–25. https://doi.org/10.3233/IDA-2007-11603.

Derrida, Jacques. *Of Grammatology*. Translated by Gayatri Chakravorty Spivak. Baltimore: Johns Hopkins University Press, 1997.

Deutsch, Diana, Kevin Dooley, and Trevor Henthorn. “Pitch Circularity from Tones Comprising Full Harmonic Series.” *The Journal of the Acoustical Society of America* 124, no. 1 (July 1, 2008): 589–97. https://doi.org/10.1121/1.2931957.

Eimert, Herbert, and Karlheinz Stockhausen, eds. *Anton Webern*. Die Reihe 2. Bryn Mawr, Penn.: Theodore Presser Co., 1958.

Fineberg, Joshua. “Spectral Music.” *Contemporary Music Review* 19, no. 2 (March 2000): 1. https://doi.org/10.1080/07494460000640221.

Gaburo, Kenneth. *Maledetto*. La Jolla, Calif.: Lingua Press, 1967.

Gilbert, Nouno, Arshia Cont, Grégoire Charpenier, and Jonathan Harvey. “Making an Orchestra Speak.” Porto, Portugal, 2009.

Hanning, Barbara Russano. *Of Poetry and Music's Power: Humanism and the Creation of Opera*. Studies in Musicology ; No. 13. Ann Arbor, Mich: UMI Research Press, 1980.

Harryman, Carla. *Open Box (Improvisations)*. New York: Belladonna, 2007.

Hart, Peter E., Nils J. Nilsson, and Bertram Raphael. “A Formal Basis for the Heuristic Determination of Minimum Cost Paths.” *IEEE Transactions on Systems Science and Cybernetics* 4, no. 2 (July 1968): 100–107. https://doi.org/10.1109/TSSC.1968.300136.

Harvey, Jonathan. *Speakings*. London: Faber & Faber, 2008.

Hejinian, Lyn. “The Rejection of Closure.” In *A Guide to Poetics Journal: Writing in the Expanded Fields 1982-1998*, edited by Lyn Hejinian and Barrett Watten, 87–97. Middletown, CT: Wesleyan University Press, 2013.

Higgins, Dick. *Intermedia, Fluxus, and the Something Else Press*. Edited by Steve Clay and Ken Friedman. Catskill, New York: Siglio, 2018.

Hutchinson, Brenda, Paul Dutton, Valeri Scherstjanoi, Amanda Stewart, Trevor Wishart, Henri Chopin, David Moss, and Libor Zajicek. “Contributors' Notes.” *Leonardo Music Journal* 3 (1993): 73–79. https://doi.org/10.2307/1513274.

Jadoul, Y., B. Thompson, and B. de Doer. “Introducing Parselmouth: A Python Interface to Praat.” *Journal of Phonetics* 71 (2018): 1–15. https://doi.org/10.1016/j.wocn.2018.07.001.

Jones, David Evan. “Compositional Control of Phonetic/Nonphonetic Perception.” *Perspectives of New Music* 25, no. 1/2 (1987): 19.

Kirnberger, Johann. *The Art of Strict Musical Composition*. New Haven: Yale University Press, 1982.

Koestler, Arthur. *Janus: A Summing Up*. London: Hutchinson & Co., 1978.

Lerdahl, Fred. *A Generative Theory of Tonal Music*. MIT Press Series on Cognitive Theory and Mental Representation. Cambridge, Mass: MIT Press, 1983.

Lockspeiser, Edward. *Debussy*. Fifth edition. Master Musicians Series. London: Dent, 1980.

Mackey, Nathaniel. "The Changing Same: Black Music in the Poetry of Amiri Baraka." In *Discrepant Engagement*, 22–48. Cambridge: Cambridge University Press, 1993.

MacLow, Jackson. "Methods for Reading Asymmetries." In *An Anthology of Chance Operations*, edited by La Monte Young. New York: La Monte Young & Jackson Mac Low, 1963.

Messiaen, Olivier. *The Technique of My Musical Language. 1st Volume, Text*. Bibliotheque Leduc. Paris: Alphonse Leduc, 1944.

Partch, Harry. *Bitter Music: Collected Journals, Essays, Introductions, and Librettos*. Music in American Life. Urbana: University of Illinois Press, 1991.

———. *Genesis of a Music; an Account of a Creative Work, Its Roosats and Its Fulfillments.* Second edition, Enlarged. New York: Da Capo Press, 1974.

"PennSound: George Oppen." Accessed January 30, 2020. https://writing.upenn.edu/pennsound/x/Oppen.php.

Perloff, Nancy. "Sound Poetry and the Musical Avant-Garde." In *The Sound of Poetry / The Poetry of Sound*, edited by Marjorie Perloff and Craig Dworkin, 97–117. Chicago and London: University of Chicago Press, 2009.

Peters, Carl. *Studies in Description*. Vancouver, BC: Talonbooks, 2016.

Polansky, Larry. "Morphological Metrics." *Journal of New Music Research* 25, no. 4 (December 1, 1996): 289–368. https://doi.org/10.1080/09298219608570710.

———. "Program Notes for Three Indigneous Songs." *San Francisco Symphony*, November 1984. http://eamusic.dartmouth.edu/~larry/misc_writings/out_of_print/tenney_symphony.pdf.

Polansky, Larry, Daniel Rockmore, Micah K Johnson, Douglas Repetto, and Wei Pan. "A Mathematical Model for Optimal Tuning Systems." *Perspectives of New Music*, 2009, 69–110.

Poller, Tom Rojo. "Clarence Barlow's Technique of 'Synthrumentation' and Its Use in Im Januar Am Nil." *Tempo* 69, no. 271 (2015): 7–23.

Princeton University. "About WordNet." Princeton University, 2010. https://wordnet.princeton.edu/.

Remez, R. E., P. E. Rubin, D. B. Pisoni, and T. D. Carrell. "Speech Perception without Traditional Speech Cues." *Science* 212 (1981): 947–50.

Rousseau, Jean-Jacques, and Johann Gottfried Herder. *On the Origin of Language*. Chicago and London: University of Chicago Press, 1966.

Sabat, Marc. "23-Limit Tuneable Intervals above and below A," 2005. http://www.marcsabat.com/pdfs/TIab.pdf.

Shannon, Claude. "A Mathematical Theory of Communication." *The Bell System Technical Journal*, no. 27 (October 1948): 379–423.

Shepard, Roger N. "Circularity in Judgments of Relative Pitch." *The Journal of the Acoustical Society of America* 36, no. 12 (December 1, 1964): 2346–53. https://doi.org/10.1121/1.1919362.

Silliman, Ron. *The New Sentence*. New York: Roof Books, 2003.

Smith, Andrew C. *How To Get There From Here*. Santa Cruz, Calif.: Indexical, 2016.

Tenney, James. "About Changes: Sixty-Four Studies for Six Harps." In *From Scratch: Writings in Music Theory*, edited by Larry Polansky, Lauren Pratt, Robert Wannamaker, and Michael Winter, 327–49. Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015.

———. "Darmstadt Lecture." In *From Scratch: Writings in Music Theory*, edited by Larry Polansky, Lauren Pratt, Robert Wannamaker, and Michael Winter, 350–67. Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015.

———. "Form in Twentieth-Century Music." In *From Scratch: Writings in Music Theory*, edited by Larry Polansky, Lauren Pratt, Robert Wannamaker, and

Michael Winter, 150–65. Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015.

———. "John Cage and the Theory of Harmony." In *From Scratch: Writings in Music Theory*, edited by Larry Polansky, Lauren Pratt, Robert Wannamaker, and Michael Winter, 280–304. Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015.

———. "Meta Hodos." In *From Scratch: Writings in Music Theory*, edited by Larry Polansky, Lauren Pratt, Robert Wannamaker, and Michael Winter, 13–96. Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015.

———. "The Structure of Harmonic Series Aggregates." In *From Scratch: Writings in Music Theory*, edited by Larry Polansky, Lauren Pratt, Robert Wannamaker, and Michael Winter, 240–79. Urbana, Chicago, and Springfield, IL: University of Illinois Press, 2015.

Wannamaker, Robert A. "The Spectral Music of James Tenney." *Contemporary Music Review* 27, no. 1 (February 2008): 91–130. https://doi.org/10.1080/07494460701671558.

Watten, Barrett. *Questions of Poetics: Language Writing and Consequences*. Iowa City: University of Iowa Press, 2016.

———. *Total Syntax*. Carbondale and Edwardsville, IL: Southern Illinois University Press, 1985.

Werckmeister, Andreas. *Musicalische Temperatur*. Translated by Elizabeth Hehr. Oberlin, OH: Oberlin University Press, 1974.

Wiener, Norbert. *Cybernetics; or, Control and Communication in the Animal and the Machine.* Second edition. Cambridge, Mass: MITPress, 1961.

Wolff, Christian. *Duo for Violins*. Edition Peters, 1950.