

## Music of Poetry and Poetry of Song: Expressivity and Grammar in Vocal Performance

TYLER BICKFORD / Columbia University

While poetry attempts to convey something beyond what can be conveyed in prose rhythms, it remains, all the same, one person talking to another; and this is just as true if you sing it, for singing is another way of talking.

—T. S. Eliot

In the poetry of the folk-song, language is strained to its utmost that it may *imitate music* . . .

—Friedrich Nietzsche

### Introduction

Following T. S. Eliot (1942:16) and Friedrich Nietzsche ([1872] 1995:18), this article develops the claims that singing is (only) discourse—not the superlative site of expression that romantic idealism would have it; and singing is more than (just) language—like verse it organizes utterances in forms not native to language. The boundaries of speech and song have long been among the important topics of ethnomusicology, notably addressed in George List's foundational article (1963), in George Herzog's early explorations of the relationship between music and text (1934, 1942, 1950), in John Blacking's account of musical "discourse" (1982), and in linguistically informed work such as that by Leanne Hinton (1980, 1984), Laura Graham (1984, 1987), Charles Briggs (1993), Jean-Jacques Nattiez (1999), and Aaron Fox (1992, 2004). This dialogue with linguistics, prominent in ethnomusicology through the 1980s, has largely receded from view since.<sup>1</sup> In this article I revisit a topic of early interest to our discipline by investigating the patterns of sound and gesture that structure acts of speaking and singing as intelligible and interpretable utterances. Recognizing, as Bakhtin notes, that "discourse is a social phenomenon . . . throughout its entire range and in

each and every of its factors, from the sound image to the furthest reaches of abstract meaning" ([1975] 1981:259), I examine the expressive role of phonology and phonetics in vocal performance in order to articulate a closer perspective on the place of singing in discourse, and thereby to contribute to an understanding of discursive practice generally.<sup>2</sup>

Where the boundaries of musical forms are often imprecise or unstable, the sounds of language are always part of a regular system of meaning. The distinctive features of language are the substantive basis of all the levels of systematic, hierarchical, and productive regularity that linguists have found in syntax, morphology, and phonology (Jakobson 1978). Therefore language "is functionally unique among the phenomena of culture" (Silverstein 1976:12). Following Silverstein (1976) and Urban (1991), who explore other "phenomena of culture" by considering their interrelations with the functionally grounded forms of language, I propose that the substantive regularity of linguistic grammar constitutes a useful analytical context in the study of song. Because the sounds of language are already parsed into functionally and perceptually relevant units, the sounds of language in song provide a material grounding that can help constrain the slippery acoustic bounds of musical forms.

Elements of phonology such as stress, phrasing, and syllable structure, to name a few, have real bearing on singing, and their functional (i.e., grammatical) presence in the linguistic context of a song's lyrics plays an important role in the expressive forms of singing. My aim is not so much to view the sounds of poetry as themselves musical (Lerdahl 2001) as it is to view the sounds of language *in* music (Feld and Fox 1994) as a site of expressive interaction between systems of aural meaning and form.

In pursuit of this goal, I present an analysis of Bob Dylan's recorded performance of "Down the Highway," from *The Freewheelin' Bob Dylan* ([1963] 2004). "Down the Highway" is an easy and competent blues; its sparse guitar accompaniment facilitates close attention to Dylan's idiosyncratic singing style, which highlights the pronunciation and inflection of words and phrases. An extensive literature explores the biographical, literary, political, and commercial contexts of Dylan's music, while interrogating his position as a curator and interpreter of American vernacular music (e.g., Gill 1998; Ricks 2003; Marcus 1997, 2005; Shank 2002; Gray 2000; Boucher and Browning 2004; Corcoran 2002). I focus here on the micro-level interrelations of language and music in Dylan's singing in order to address issues of vocal performance more generally.

The following section provides a review of relevant literature, arguing for a grammatical, phenomenological, and semiotic understanding of the ways language sounds are properly musical in singing. In the third and fourth sections, I present a detailed analysis of the poetics and music of Dylan's

singing and argue that the expressive form of “Down the Highway” cannot be understood without an integrated approach to music and language in the voice. Next I acknowledge and account for certain contexts of the recording’s meaning, performance, and reception. The conclusion situates these analytical positions within a broader approach to the layering of meaning and form in discourse.

### **Background: Verbal Art in Linguistics and Anthropology**

This article draws on and contributes to two frequently opposed approaches to the language of verbal art. First, asserting the ubiquitous presence of linguistic grammar as a structuring force in expressive singing puts the study of vocal music in dialogue with formal linguistics and with a school of musicology that looks to linguistic theory as a framework for understanding music (e.g., Nattiez 1973, 1990; Lerdahl and Jackendoff 1983; Halle and Lerdahl 1993). I hope to avoid the pitfalls that may come with such an approach: especially that of applying linguistic formalisms to music without interrogating the assumptions that underly the motivating “language-music” analogy (Feld 1974). Rather, singing in particular must be understood in dialogue with linguistics because singing is full of language; this is not to say that singing *is* language, or that transformational linguistics should provide more than passing insight into musical form. Expressive singing is intimately bound by the grammar of language because the words it includes are so bound; therefore linguistic grammar must be given some attention in an approach to expressive singing.

By the same move—asserting the relevance of linguistic forms in song—this article contributes to an approach to verbal art that has been fruitfully developed by linguistically inclined anthropologists and folklorists. Urban argues that the material substance of “actually occurring instances of discourse” (1991:1) constitutes the principal site of “culture, as a collection or history of publicly accessible sign vehicles” (ibid.:19). Literature in this tradition claims that the formal properties of discourse are clues to the function of sign vehicles in social life (Silverstein 1976; Hymes 1981; Tedlock 1983; Urban 1985; Sherzer 1987; Briggs 1988). The fact that discourse, and verbal art in particular, is so often structured in regular and sensible forms provides a foothold from which to explore the unconscious ways in which social life structures embodied practices of thought and behavior (Urban 1991). As Woodbury (1985) demonstrates that the structures of discourse are pervasively involved with patterns of syntax and phonology, and as Jakobson ([1960] 1987) argues for understanding grammar and poetics as related and interacting functions in verbal art, I develop the claim that singing is structured by the interrelations of different levels of discourse. By asserting

the relevance of language sounds to singing, I make a claim about the phenomenology of song: the linguistic systems of sound and meaning mediate the sensory and semiotic modes of expressivity in singing.

### **Linguistic Approaches to Singing**

Formal linguistics, especially in the work of Roman Jakobson (e.g., 1960, [1960] 1987, [1970] 1987; Caton 1987), has considered verbal art mostly in its engagement with poetry. In particular, the study of meter in poetry has provided a rich site of inquiry for linguists investigating phonological phrase structure and prosody.<sup>3</sup> Bruce Hayes and Abigail Kaun extend linguistic perspectives on metrics to the problem of “textsetting,” which involves the rhythmic performance of lines of metrical poetry, and they argue for situating singing and chanting more prominently in studies of metrics:

The poetry found in most cultures is not the kind usually studied by generative metrists. As far as we know, the default state of metrical poetry is to be *sung*, or at least rhythmically chanted, rather than spoken. Homer, the Serbo-Croatian *guslars*, rap artists, and the singers of Appalachian folk ballads are all exponents of poetry in what is arguably its most natural form. (Hayes and Kaun 1996:244)

Textsetting involves productive interaction between the different rhythmic systems of music and language. Hayes and Kaun work with laboratory data of settings of 670 English folksong lines elicited from ten informants. Their statistical analysis reveals notable regularities in the relationship between the prosody of the text and its rhythmic setting. In addition to preserving stress contours and syllable length (mapping these onto musical beats or pulses), Hayes and Kaun find that textsetting is significantly constrained by phonological phrasing. The edges of clitic groups, phonological, and intonational phrases (phonological structures of increasing order) condition processes like “metrical inversion”—where, in singing, the spoken stress contour of a word or metrical foot is reversed at a prosodic phrase boundary (e.g., a strong-weak foot may be rendered weak-strong at the end of a line). Phrase structure in many cases is favored over syllable stress.

In response to studies by Halle and Lerdahl (1993) and Halle (1999), Hayes (Forthcoming b) argues for understanding textsetting in terms of Optimality Theory (Prince and Smolensky [1993] 2002), which identifies rankable “constraints” on well-formedness that “mark” against (identify as potentially ill-formed) phonological structures. Hayes reformulates his and Kaun’s 1996 conclusions as a ranked list of constraints on English textsetting. In particular, English textsetting is seen to be structured by the following: (1) constraints against associating strong syllables with weak beats and vice versa, (2) a constraint against overlong pauses between syllables, (3) constraints that condition the realization of the phonetic length of syllables as rhythmic du-

ration, and (4) constraints that line up prosodic phrase edges with musical phrase edges.<sup>4</sup>

These authors all delimit their studies in two important ways. First, they emphasize their avoidance of questions of expressivity, bracketing “spontaneous artistic decision[s]” (Hayes and Kaun 1996:248) and “impulses [that] may be preferred by composers precisely because they are violations and hence are striking and unexpected” (Halle and Lerdahl 1993:9). While such a narrow focus has its uses for the development of systematic representations of these musical processes, it is not clear with regard to normal musical practice (i.e., beyond the standardizing context of the linguistics laboratory) that these analyses develop our understanding of singers’ competence at setting lines in novel ways. Nonetheless, these studies might provide insights about “artistic” singing (which I take to be the norm rather than the exception) by formalizing an implicit ground against which variations in performance are contextualized and understood. Second, these authors choose not to examine pitch and melisma, considering only rhythm in their analyses of textsetting. I find that in Dylan’s singing both melody and rhythm interact with phonological elements in ways that specifically emphasize the singer’s expressive moves.

### **Discourse-Centered Approaches to Verbal Art**

Ethnographic studies of verbal art by folklorists and anthropologists have contributed broad cross-cultural knowledge about the mediations of language and culture (Sherzer 1987; Hymes 1972b). In particular, by examining the formal structures of verbal art in the course of careful ethnographic attention to society and culture, linguistic anthropologists have argued that verbal art across cultures is structured by various types of parallelism (the repetition of segments of discourse from one instance to another), which takes place at the level of the line and verse (Hymes 1977, 1980), in timing and inflection (Tedlock 1983), and among the interrelations of syntax and phonology (Woodbury 1985), indexing and shaping genre and style (Urban 1985). Broadly defined, parallelism mediates between discourse and culture, instantiating relations of similarity and difference between segments of discourse of varying proximity and thereby structuring cultural notions of historical change or continuity and of the boundaries and limitations of social roles (Urban 1991). Music might be incorporated easily into a framework that looks to parallelism:

The formal parallelism and informational redundancy of musical structures, musically structured song texts, performances, and musically structured kinesic forms such as dance . . . exemplify and expand what Jakobson [1960] called the poetic and metalinguistic functions in language . . . In this view, musically structured

communication suppresses verbal referentiality in order to reveal the formal and pragmatic ordering of messages, codes, and communicative contexts. Music's formal redundancy and auto-referentiality heighten poetic texts and produce a musical metalanguage. (Feld and Fox 1994:27)

Parallelism in verse depends on categorical equivalences between linguistic elements produced through rhyme, assonance, and stress:

Equivalence [in poetry] is promoted as the constitutive device of the sequence. In poetry one syllable is equalized with any other syllable of the same sequence; word stress is assumed to equal word stress, as unstress equals unstress; prosodic long is matched with long, and short with short; word boundary equals word boundary, no boundary equals no boundary; syntactic pause equals syntactic pause, no pause equals no pause. Syllables are converted into units of measure, and so are morae or stresses. (Jakobson 1960:358)

Phonology is often thought to be unable to count higher than two. That is, unlike phonetics, which continuously differentiates gestural and acoustical scales, the grammatical patterns of language sounds only make binary distinctions: distinctive features are either privative or binary, phrase structure distinguishes only stress and unstress at any hierarchical level, prosodic feet are maximally bimoraic or bisyllabic, etc. Therefore, Jakobson's description of poetic equivalences extends a basic principle of phonological measure to poetics.

As Feld and Fox note, music has much "formal redundancy" and therefore parallelism, but musical progressions also often depend on gradient articulations of sound, a notion that cannot be expressed well in terms of parallelism. Music is not apparently constrained by categorical equivalences as in phonology. Rather, like phonetics, musical parameters often seem to operate along continuous, gradient scales of pitch, timbre, rhythm, and dynamics. Phonology, and therefore poetics, has no means to express structures that depend upon continuous progressive change, like such canonically musical forms as "crescendo," "accelerando," or melodic sequencing.<sup>5</sup> Where formal linguistics takes as its starting point the categorical distinctions features make between acoustically continuous sounds (Jakobson 1978), there is simply no equivalent of the distinctive feature in music. Musical forms are related or distinguished based on shifting and never universally salient segmentations of gradient acoustic scales in different ways in different environments.

Connecting phonetics and music through "gradience" in an analytical approach to language in song contributes a recognition that the featural qualities of linguistic segments can have distinctly musical value in the right contexts. Two recent studies in particular, both inquiries into speech about music, demonstrate the potential musicality of language: Thomas Porcello's ethnographic account of a Texas recording studio, "Talk about Timbre in the Recording Studio" (Feld, Fox, Porcello, and Samuels 2004:323-28), and Caro-

line Traube's dissertation, "An Interdisciplinary Study of the Timbre of the Classical Guitar" (2004).

Porcello examines the conversations of musicians and recording engineers as they discuss the sound qualities of instruments. Among the many discursive strategies studio participants use to communicate about sound, "spoken/sung 'vocables' are used in an attempt to iconically mimic in vocalizations the timbral features of the musical sound(s) under consideration" (Feld, Fox, Porcello, and Samuels 2004:324). These "vocables" comprise linguistic segments rendered in what are often grammatically questionable syllables (e.g., [ŋ:], [dʒ:], [ŋ:ts], etc.; see the appendix for a guide to phonetic symbols used in this article). The segments in such forms do their usual phonemic work distinguishing between structures, but here the forms themselves are ad hoc and unconventionally "morphological" representations of musical sounds. The sounds of language are put to use in a decidedly linguistic context, but as sounds they are newly inflected with a metaphorical musicality.

Traube's findings are similar, though she disregards the pragmatic, expressive function of onomatopoeia in favor of systematic investigation of the perceptual relation between musical acoustics and phonetics. Examining the musical categorization of classical guitar timbres through informants' phonetic representations of sound color, Traube collected agreed-upon timbral categories (e.g., "brassy," "round") that appeared to correspond consistently with acoustical characteristics and performance techniques. Her study found highly productive correlations between timbral categories and the consonant-vowel phonetic gestures (i.e., syllables) that she elicited from her informants to represent the attack and release of guitar notes ([tê], [to], [gA], etc.). Traube thus provides real-world, discourse-based links between linguistically contextualized utterances that are prosodified into syllables—CV(C) strings of segments that are accountable to grammar—and the musical function and meaning of these utterances in performers' understanding of the sounds of the guitar. Language as language, here, is seen to have musical meaning for Traube's informants.

The tools of language that mediate sound and meaning (distinctive features) intervene where otherwise the semiotic relation between musically meaningful structures and purely acoustic sound is either opaque or discontinuous. These studies of onomatopoeia suggest a semiotic relation between linguistic and musical sounds that has distinct relevance to more general questions of the role of language in singing.

### **Sound and Meaning**

Investigations of musical form should look beyond the physicality of sound to the complex layering of meaning as the interaction of both refer-

ential and formal semiotic systems. Determining the physical correlates of musical values—“match[ing] . . . perceptual and acoustic systems” (Fales and Berger 2005)—has less relevance for questions of musical form and meaning than analysis that looks for structure in the complex interplay between various semiotic systems and between individual levels within one system. Analysis should be grounded “between sensuous concreteness and deductively organized abstraction” (Urban 1991:28), in ever more detailed levels of discourse. This framework expands outward through “the formal constituents of performance, the relationships that tie poetic form to social function and interpretive meaning, and the broader role of performance in the conduct of social life” (Bauman 2002:94).

If we seek something like this approach to culture and discourse for the study of music—Feld (1984) proposed essentially this under the rubric “sociomusicology”—then the formal elements of music must be better understood. The mediation of culture by discourse occurs in most cases just outside direct consciousness, between “sensibility and intelligibility” (Urban 1985:28; also Silverstein 1976), and it is the presence of analytically and practically salient forms at those levels that allows analysis its insights. The multiple mediations of sound and meaning in song have phenomenal resonance in their embodied practices of performance and reception. Knowing how song structures itself—how culture grabs hold of the singing voice to elicit particular sounds and forms (Urban 1991:132)—is a first step in understanding how the micro- and macro-structures of singing reflect and subtly shape the micro- and macro-structures of their reception, just as language mediates the phenomenological reception of sounds, where both music and language exist in a matrix of social and semiotic systems.

The following analysis is an attempt at such a project, treating language as an integral mediator between form and meaning. The aural characteristics of musical form are seen to be linguistic values: vowels, syllables, feet, and phrases in combination with melody, rhythm, timbre, and dynamics. In singing, language—already a constant mediator between sound and meaning—intervenes in music as part of a specifically vocal mediation of sound and meaning.

### **Bob Dylan’s “Down the Highway”**

Dylan’s style of singing is unusual, even considering his position at the intersections of several music and verbal art traditions. He places irregular stress on words, sings with a strained nasality, and often pronounces lyrics in unexpected ways. I am concerned with instances where Dylan’s performance runs against (my) expectations, in particular by altering conventional rhyth-

mic settings and by singing with pronounced oral constriction. To identify such tropes as “impulses . . . preferred by composers precisely because they are violations and hence are striking and unexpected” (Halle and Lerdahl 1993:9) is only half the story. The expressive, artistic, musical, or poetic framework in which these tropes exist—that motivate and contextualize the “striking and unexpected” phrases or gestures—should be accountable through analysis.

Theories of expressive form allow us to understand the apparent unpredictability of Dylan’s performance: in terms of poetic parallelism and musical progression, Dylan’s singing makes sense. Locally ungrammatical utterances in “Down the Highway” are in fact consistent with and motivated by larger-scale forms. In the following section, for example, the inversion of the stress contour of the word “suitcase” in the first verse is seen to highlight the syllable “-case” in a way that underscores one interpretation of the poetic context: “-case” shares an important assonance with several other prominent words, and it is repeated later in the song as part of a larger parallelism between the first and the fifth verses. In the section on “sonority,” an orally constricted syllable ([ŋ] of “ocean”) that is “marked” as unlikely in a prominently sustained and melismatic melodic figure is in fact performed in such an environment as part of a larger progression over the course of several verses toward orally constricted syllables. It resolves in the following verse both through a return to orally open syllables in prominent positions and through the repetition of this marked syllable in an expected, unstressed position. Therefore, expressive styles, by virtue of the imposition of larger formal structures onto small-scale or local utterances, can be seen to allow and even motivate otherwise marked structures.<sup>6</sup>

**Rhythm and Meter**

The opening line of “Down the Highway” is sung twice, with two different rhythmic settings. “Well I’m walking down the highway with my suitcase in my hand”<sup>7</sup> is metrically straightforward. As verse, not set to music, it is comprised of simple trochees.

x		x		x		x		x
x	x	x	x	x	x	x	x	x
well	I’m	wálk-	ing	dówn	the	hígh-	wà	y
x		x		x		x		x
x	x	x	x	x	x	x	x	x
with	my	súit-	càse	in	my	há	nd	

I use grid notation to represent prosodic and musical rhythm in the manner of Hayes and Kaun (1996), and I represent strong and weak underlying syl-

lable stress with acute and grave accent marks, respectively.<sup>8</sup> This setting, like any, is imperfect. It does not express the (weak) stress of the second syllable of “suitcase,” setting it in a weaker position than the following unstressed function word, “in.” There is no good way of getting around this problem, and this evenly trochaic setting is the clear “poetic” parse.

In singing, unlike in English verse, duration is relevant. Hayes and Kaun (1996) find that, all else being equal, their informants provide settings that account for the phonetic length of syllables. They also find that edges of phrases, especially phrase endings, condition some durational extensions that improve the correspondence between textual and musical phrases. Dylan’s performance initially seems to do this, giving two beats to each (long) syllable of “highway.”

				x						
				x				x		
	x			x		x		x		x
	x	x	x	x	x	x	x	x	x	x
	well	I’m	wálk-	ing	dówn	the	high-	—	wàý	—

(Em-dashes denote syllables sustained from the position to the immediate left. Ascending levels on the grid represent eighth, quarter, half, and whole notes, respectively, as in Lerdahl and Jackendoff 1983.) In what follows, I will refer to the first instance of this line as the “statement” and the second as the “repeat.” The diphthong in “down” is not expressed as extra duration in this setting, but the need for diphthongs and long vowels to have relatively long duration stems from a low-ranked (all-else-being-equal) constraint. Lengthening it to two beats as in “high” would require also placing unstressed “the” in a stressed position, a more fatal violation. It is clear from Hayes’s and Kaun’s (1996) data that phonetic length is salient in textsettings only if it does not conflict with other constraints, such as faithfully mapping stress contours to rhythmic contours.

A similar environment is found at the word “suitcase” in the second line, “with my suitcase in my hand.” Here the length of both “suit” and “-case” can easily be expressed while preserving the word’s descending stress contour.

				x						
				x				x		
	x			x		x		x		x
	x	x	x	x	x	x	x	x	x	x
	with	my	súit-	—	càse	—	in	my	hánd	—

Dylan’s actual phrasing is close to this, but he articulates “-case” a bit early. The song’s meter is rather flexible, so it is not totally clear to what extent Dylan’s rendering constitutes true rhythmic syncopation or just imprecise phrasing, but it is clearly audible that “-case” is stretched somewhat.<sup>9</sup>



larly returns throughout the song as an important aural focus. From verse to verse the poetic form centers on these three words and their rhymes (discussed in detail below). The stressing of “-case,” therefore, is accounted for by the imposition of larger-scale structures onto phrase-level textsettings. “Optimal” or standard settings can be violated if doing so improves the setting with respect to the musical or poetic framework of the performance.

### Enunciation As Poetic or Musical Focus

This account of the emphasis on “-case” does not explain the parallel lengthening of “my.” The extension of “my” may be a relatively natural precipitate of the focus on “-case.” Its underlying diphthong, as in “-case,” is realized rather than reduced. And, by pushing into the durational allotment of “suit,” it increases the contrast between “suit” and “-case”—helping to focus the vowel of “-case” as part of a poetic structure. Though “my” is the only syllable that is rhythmically repositioned so drastically, the entire repeat sees the realization of otherwise dropped or reduced syllable structures.

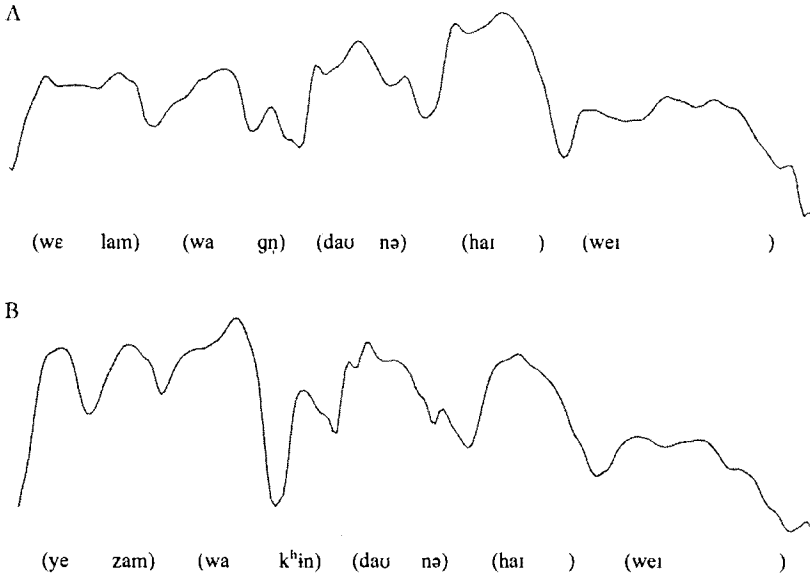
In the statement, “walking” is sung [wagŋ], with the second syllable reduced and the *k* voiced and unaspirated. But in the repeat it is heard as “walking” [wa:kʰin], with a CVC shape and aspirated *k* in the second syllable. The same thing happens on “with”: in the statement its final *θ* is deleted entirely and its vowel reduced: “with my” [wima]. In the repeat a coda is audible, assimilated to the following *m*, and the vowel is unreduced: “with my” [wivma:].

wɛ.lam.wa.gŋ.dau.nə.haɪ.wɛɪ||wi.ma.sʊʔ.kʰe::zɪn.ma.hæ  
 yɛ.zam.wa.kʰin.dau.nə.haɪ.wɛɪ||wiv.ma::suʔ.kʰe::zɪn.ma.hæ

The aspirated *k* of “walking” [wa:kʰin] is forced and ungrammatical, just as stressing “my” is inconsistent with the lyrics’ prosodic phrasing. Each seems to be *over*pronounced. If we compare the dynamic contours of the statement and the repeat as in Figure 1, we see that, whereas in the statement syllables run together (especially within poetic feet), in the repeat they have clearly defined boundaries. Even the substitution of “yes” for “well” at the beginning of the line increases the contrast between vowels and consonants (“well I’m” [welam] versus “yes I’m” [yezam]).

The differences between the statement and repeat have the effect of shifting the aural emphasis from musical to linguistic structure. In the statement, the voiced consonants allow the melodic pitch to continue, uninterrupted by consonantal stops. In effect, the melody is brought into focus. But in the repeat, the unreduced vowels and better-pronounced consonants “enunciate” the lyrical structures. Insofar as there is necessarily some inherent ambiguity in an acoustic signal that contains both musical and phonetic information, it seems reasonable that singers would be able to vary their performances

**Figure 1. Dynamic contours (db) of “I’m walking down the highway.” A. Statement B. Repeat.**



in ways that set on one in relief against the other. And because the musical (rhythmic) setting of this line serves to emphasize the *poetic* function of the syllable “-case,” the concurrent aural emphasis on linguistic over musical information follows.

Such processes suggest clear empirical questions about the patterning of music and language in singing. How do singers shift focus between musical and linguistic structures, if in fact they do? If musical and poetic structures have different roles in singing, then an emphasis on poetic forms (as with the large-scale assonance of “-case”) might condition processes like the “enunciation” in the repeat, whereas focusing musical forms should condition consonant voicing between syllables or other processes that emphasize melodic or rhythmic information. Of course, some ambiguity is likely: musical “staccato” would be sung with short vowels and plosive stops, for instance. List’s (1963:9) diagram of the several parameters that differentiate speech from singing might easily be expanded to account for these prosodic structures.

Graham (1984) identifies a similar process in Native South American Shavante vocal expression that differentiates speech styles from one another by their relative emphasis on linguistic or musical parameters. The differences

between ritual wailing, collective singing, and political oratory—all marked speech genres among the Shavante—suggest a continuum from “melody” to “semanticity.” Ritual wailing, which is characterized by sustained and fully pitched vowels, a relative lack of contrastive consonants, highly mobile pitch, and few if any discernible words, is at the far end of the spectrum from political oratory, which though vocally heightened is characterized by indeterminately pitched vowels and fully realized consonants and comprises only words; communal singing falls somewhere in between. If the markers of melody and language can be put to such productive use in distinguishing genres of vocal expression, it seems reasonable that a singer might have similar access to the same verbal and musical parameters within a single performance to distinguish one phrase from another, as Dylan seems to do here by shifting the emphasis from continuous pitch to fully articulated segments.

### Sonority

With the exception of one prominent recurring melody, the lyrics of this song are set syllabically, with each syllable realized as a single musical note. Each verse has a melodic climax at the end of its penultimate line, in which a single weak syllable is sung in an upper register and melismatically. I will call this motive the “theme” (Figure 2).

In the first verse, the theme is sung at the end of the word “baby.” Held out over three notes (D-F-D), the final vowel of “baby” is diphthongized, so vowel changes accompany the pitch changes: [-be.i.i:] (periods separate notes rather than syllables in this transcription). It is easy to imagine articulatory or phonological reasons for the diphthongized pronunciation. For instance, notes and syllables might be thought of as corresponding musical and linguistic constituents and so should align; breaking the diphthong at the pitch change effectively correlates at least two syllables with different pitches.<sup>10</sup> We might similarly imagine a desirable result in the alignment of pitch height with vowel height—low-to-high pitch motion is coarticulated with grave-to-acute vowel changes.

An additional motivation for the diphthongized pronunciation of “baby” is apparent in higher-level structure: when it is pronounced [beibeɪ], “baby”

Figure 2. “Theme” motive of “Down the Highway.”



rhymes better with “highway.” Its metrical position in parallel with “highway” motivates the alternation in its pronunciation. In this way, the diphthongized pronunciation of “baby” is a process similar to the lengthening of “-case” discussed earlier, in which both effects emphasize the relatedness of certain words and word sounds in the quality of their vowels and their parallel metrical positions. “Baby” is not consistently pronounced this way; later in the song it recurs without the diphthong. Whatever the specifically poetic reasons for the diphthongization of “baby,” it is prominent within the verse because of its melodic setting as the highest pitched and most active melody so far. Each verse repeats the same structure, in which the fifth (penultimate) line ends with the high, melismatic “theme,” whose dominant harmony is resolved in the final line.

The fifth lines are set apart poetically as well as melodically. In particular, although the pronunciation of “baby” in the first verse establishes a clear ABABAB rhyme scheme, the subsequent verses do not fully adhere to that scheme (Table 1). The fifth lines, the final words of which are set to the melodic theme, do not rhyme as we expect them to: the rhyme schemes of the

**Table 1. Rhyme scheme in verses 2–4 of “Down the Highway.” Underlined words are sung to the “theme” (Figure 2).**

---

Your streets are getting empty	A
And your highways getting filled	B
Your streets are getting empty	A
And your highways getting filled	B
The way I love that <u>woman</u>	C
I swear it’s bound to get me killed	B
I been gambling so long	A
I ain’t a got much more to lose	B
I been gambling so long	A
I ain’t a got much more to lose	B
Right now I’m having <u>trouble</u>	C
Please don’t take away my highway shoes	B
I’m bound to get lucky baby	A
Or I’m bound to die trying	B
I’m bound to get lucky baby	A
Or I’m bound to die trying	B
Meet me in the middle of the <u>ocean</u>	C
[ <i>unintelligible</i> ] leave this old highway behind	B

---

second, third, and fourth verses are ABABCB. Despite their failure to rhyme, the final syllables of the alternating lines do correspond to a repeating stress template: the A lines (1, 3, and 5 of each verse) have weak-final endings, while the B lines have strong-final endings.<sup>11</sup> So the first verse sets weak-final “highway” and “baby” on lines 1, 3, and 5, and it sets strong-final “hand” and “land” on lines 2, 4, and 6. The apparent exceptions to this pattern, “trying” and “behind” in verse 4, are pronounced as perfectly rhymed, strong-final feet: [tráĩ] and [bəháĩ]. So while the ABABAB rhyme scheme is not strictly regular, a variety of parallelisms collaborate to structure the poetic form.

To summarize, the endings of the fifth lines of each verse are related through several parallel structures: in the penultimate line in the metrical framework, they are the exceptions to the rhyme scheme; they are harmonically prominent (the “dominant” harmony on which they are sung is conventionally followed by a resolution to the “tonic” harmony); and they are sung on the theme, at a higher pitch and with the only instance of melisma in the song. Poetically and musically, the fifth line of each verse is the exceptional line; that this local “exceptionality” repeats from verse to verse incorporates it into a framework of parallelism.

Parallelism “projects the principle of equivalence from the axis of selection into the axis of combination” (Jakobson 1960:358). In rhyme, for instance, words in paradigmatic relation with each other by their shared sounds are put in syntagmatic relation at the ends of lines—the axis of selection (words with certain sound-shapes) projected onto the axis of combination (words grammatically appropriate at the end of specific lines). In this case, metrical position, stress contour, failure to rhyme, and melodic repetition create a strong sense of parallelism between the successive fifth-line endings. Further, verse-internal parallelism is effectively neutralized: these words distinctly do not rhyme with the lines we expect them to rhyme with. The fifth line, in effect, draws the strongest parallels between verses; where verses are internally structured by rhyme and meter, the fifth lines are structures that parallel each other more than they parallel elements in the same verse.

I emphasize the extent of the connections between the fifth line endings because I want to draw out a structure that is otherwise obscured by layers of semantics, metrics, and phonology. So many things are identical about the successive themes on “baby,” “woman,” “trouble,” and “ocean”—repeating melodies, nearly exact rhythms, comparable positions in the rhyme scheme, equivalent word-stress templates, and shared metrical placement in the penultimate line of each verse—that the otherwise minor differences between syllables are set in relief. When prosody, phrase structure, melody, and rhythm are all basically the same, what remains besides the semantic reference of the words is segmental quality, that is, the different sounds of the “vowels” in the syllables “-by,” “-man,” “-ble,” and “-cean.” In particular, Dylan sings the successive

themes with greater and greater oral constriction, creating a progression of decreasingly sonorous syllables.<sup>12</sup>

“Sonority” describes a gradient phonetic scale that represents the relative acoustic prominence of segments (Lacy 2002); low vowels such as [a] are most sonorous and unvoiced stops such as [t] are least sonorous. The relevant phonological environments for our purposes are syllable nuclei (also called “heads”), the vowels (though this category is not strictly limited to vowels) to which consonants “attach” to form syllables. Syllable nuclei are more sonorous, and thus louder and less orally constricted, than neighboring consonants. Across cultures and certainly in European and English-language musical practice, syllable nuclei are the segments that are pitched, lengthened, and otherwise constituted in singing as musical “notes.”

Prince and Smolensky distinguish between sonority as a phonetic property and its phonological or grammatical implications in terms of the Optimality Theoretic concept of “harmony.” “A higher sonority *nucleus* is more harmonic than one of lower sonority,” but “*segments* of high sonority are not more harmonic than those of lower sonority” ([1993] 2002:17, emphasis added). “Harmony” refers to structural function (syllable nucleus), whereas “sonority” is simply a gradient feature of individual segments: “It is only when segments are contemplated in a structural context that the issue of well-formedness arises” (ibid.). Nuclear harmony in Prince and Smolensky’s account is divided into twelve distinct constraints, each of which categorically marks against the presence of certain segments in syllable nuclei ([1993] 2002, see ch. 2 and ch. 8). I will refer throughout to “sonority” rather than “harmony” for two reasons: (1) despite an intuitively plausible relationship between sonority and the well-formedness of sung notes, it is not clear that formal linguistic accounts of such phenomena are appropriate, especially if progressive forms are seen to result from its effect, and (2) the musical context would render the term “harmony” hopelessly ambiguous.

Sonority decreases from low vowels to high vowels to liquids to nasals to voiced fricatives to unvoiced stops.<sup>13</sup> It is represented as a single feature with eleven positions, so [i], a high vowel, is [ooooooooxxSonority] and [s], an unvoiced fricative, is [oaxxxxxxxxxSonority] (Lacy 2002:48). Table 2 shows the sonority of each theme vowel. Downward arrows between lines indicate decreasing sonority and upward arrows indicate increasing sonority from theme to theme. No arrow indicates equal sonority.

While (English) textsetting is apparently structured by processes highly consistent with “grammatical” analysis (groups of speakers/singers largely agree upon the best possible rhythmic setting of a line and can identify settings that are unmetrical), it is not as clear that unsonorous singing is so regularly accountable to systematic rules of well-formedness. Nonetheless, when Dylan sings “trouble” [trʌb:] and “ocean” [oʃn:], to me his voice sounds

**Table 2. Sonority of vowels in the “theme” of each verse of “Down the Highway.”**

Line ending with “theme”	As sung	Sonority
Lord I really miss my <u>ba</u> by	[-be.i.i]	[oooooooooax]
...	↓	↓
Well the way I love that wo <u>ma</u> n	[-mə.ə.ən]	[oooooooooaxx]
...	↓	↓
Right now I’m having <u>trou</u> ble	[-b].l.[]]	[oooooooooaxxx]
...	↓	↓
Meet me in the middle of the <u>oce</u> an	[-]n.ŋ.ŋ]	[oooooooooaxxx]
...	↑	↑
She packed it up in a <u>sui</u> tcase	[-ke.i.is]	[oooooooooax]
...		
Yes I’m walking down your <u>high</u> way	[-we.i.i]	[oooooooooax]
...	↑	↑
All the way to the statue of <u>liber</u> ty	[ti.i.i]	[oooooooooaxx]

Credit. Bob Dylan, © Copyright 1963, renewed 1991 by Special Rider Music. Reprinted with permission.

strained, reedy, and constricted. His pronunciation of “ocean,” especially, brings this progression of decreasing sonority into relief. Whether musically or phonologically, the stressed, long pronunciation of a syllabic *n* is without question highly marked. We would expect a singer at least to realize the underlying vowels of normally reduced syllables, singing “trouble” [trʌbʌ:l] and “ocean” [oʃʌ:n].

It would not be unreasonable to posit a constraint against unsonorous “note heads” (on analogy to syllable heads, or nuclei). Though a style like humming provides some counterevidence, singing is more stringent than speech in requiring more volume, consistent laryngeal activity to support continuous pitched voicing, and durational extension of vowels. Hinton (1984), for instance, finds that Havasupai singing differs phonologically from Havasupai speech: it includes more sonorous glides and nasals, it conditions changes in pronunciation to increase the sonority of consonants (lenition), and it prevents vowel reduction or inserts vowels in positions that increase the sonority of the phonetic material being sung.<sup>14</sup> In a narrowly conceived singing context such as Hayes’s and Kaun’s language laboratory, we might expect that just as stress and phrasing map onto rhythm, so should stress, length, and syllable prominence (sonority) have salience in the vowels or segments on which notes are set. That is, the class of acceptable segments for note heads ought to be a more sonorous subset of the class of acceptable segments for syllable nuclei.

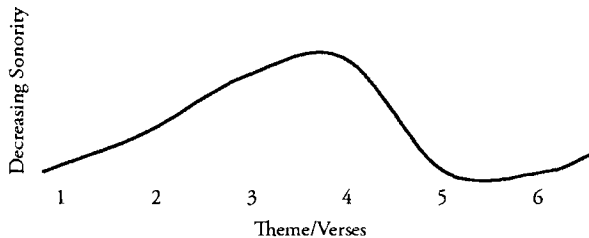
In speech, to pronounce “ocean” [oʃʌn] is phonologically ill-formed (English phonology avoids stressing reduced vowels and final short syllables, and

it absolutely forbids stressed syllabic liquids and nasals). When such marked forms arise—in speech or any other context—we assume they are somehow motivated, not simply random variation (Urban 1985:327). Because it can claim no possible phonological justification, the fact that this pronunciation occurs in a musical/poetic context suggests we look to expressive forms as motivation for the presence of these marked structures.

The low sonority of the final syllable of “ocean” is the particular characteristic that marks its pronunciation as problematic. Therefore, if we intend to contextualize it in terms of expressive form, we need to consider sonority on a larger scale. The parallelistic relations between “ocean” and the other words of the theme, “baby,” “woman,” “trouble,” and later “suitcase,” give a good sense of how to account for this markedness in the fourth verse: looking at the contour in Figure 3, we see that “ocean” in fact sits at the lowest sonority point of the series of connected themes. That Figure 3 depicts continuous progression toward low sonority in the theme and a clear return to more sonorous vowels after the fourth verse suggests a less than random relation between markedness and large-scale expressive forms.

There is further evidence for the non-randomness of these unsonorous, stressed vowels in prominent phrasal position. In the second verse Dylan performs “woman” on the theme as [wʊmə:n] and not \*[wʊmɪ:n], which would be consistent with his rendering of “trouble” and “ocean.” In weak position, the segments *l* and *n* are often syllabic, that is, pronounced as syllable nuclei (without an intervening schwa), as in “button” [bʌt.ɪ] and “bottle” [bɑ.rɪ]. It is “easier” to articulate *l* or *n* with the tip of the tongue if the preceding segment is also articulated in the same manner with the tip of the tongue, so neither “woman,” “trouble,” nor “ocean” provides an unambiguous environment for syllabic sonorants in their weak second syllables. The *ʃ* in “ocean” better conditions a syllabic *n* than the labial consonants of “woman” and “trouble,” but it is still ambiguous ([ʃ] is articulated with the tip of the tongue on the hard palate rather than the alveolar ridge, farther back in the mouth than [n]). English dictionaries distinguish between these environments in their

Figure 3. Sonority “contour” of the “theme” vowels of “Down the Highway.”



pronunciations, suggesting a meaningful and regular variation.<sup>15</sup> Whatever the status of the intervening vowel in the standard pronunciation of “woman” [wʊmən]—whether an underlying vowel or the phonetic manifestation of a non-alignment of consonant gestures (Gafos 2002)—Dylan’s variation from one prominent, melismatic environment to another suggests a contrastive function between different elements of the series. That these variations in pronunciation line up to create a progressive contour suggests variation that is functional, not random or trivial.

Finally, this progression-return contour occurs in complementary environments that reinforce the sonority progression; the “return” in the fifth verse also occurs lexically and through rhyme. When the theme of the fifth verse returns on a sonorous vowel on the word “suitcase,” shown above to be rhythmically marked by its setting, it also returns on the specific diphthong emphasized in the first verse, and on a word that previously stood out as remarkable. Musical, poetic, and lexical forms intersect in the fifth verse.

### **Rhyme and Assonance**

When the increasingly unsonorous series of theme syllables resolves in the fifth verse, it does so on a word that has already been set apart as both lexically and prosodically important: “suitcase.” The return of the theme on sonorous, less marked vowels, then, occurs simultaneously with a lexical return to a poetically important word. This establishes a rhyming relation between the first and fifth verses, reemphasizing the vowel [eɪ] that is important in the opening lines. The rhythmic marking of “suitcase” in the first verse that emphasizes its assonance with “highway” and “baby,” therefore, sets it up to act as an anchor in the fifth verse, when several processes begin to resolve (Table 3).

While the first instance of “ocean” is marked by its pronunciation, its repetition in the fifth verse is set with normal prosody. “Ocean” here is set line-internally, with its weak second syllable correspondingly rhythmically weak, not accented and extended as in the previous verse.

“Baby” recurs as well in the fifth verse, in a position to rhyme with “suitcase” as it previously rhymed with “highway” in the first verse. But “baby” here is at the end of the first and third lines and not set to the theme; therefore, it does not diphthongize to rhyme with “suitcase.” The off-rhyme that is “fixed” by diphthongization in the first verse is here left as an off-rhyme. In the first verse, it was unclear whether the best account for diphthongizing “baby” was its melismatic setting or its rhyming relation with “highway.” It appears that, when not set melodically, “baby” does not have sufficient motivation for its alternate pronunciation.

At the end of the song, “liberty” is sung on the theme like “baby” in the first verse, and it does not diphthongize to [eɪ] either. A subtle shift has occurred over the course of the song from [eɪ] to [i]. In fact, except for “suit-

**Table 3. Verses 1, 5, and 6 of “Down the Highway.”**

Well I'm walking down the <b>highway</b>	A	
With my <b>suitcase</b> in my hand	B	
Yes I'm walking down the <b>highway</b>	A	
With my <b>suitcase</b> in my hand	B	
Lord I really miss my <b>baby</b>	A	[beɪ.be.i.i:]
She's in some foreign land	B	
...		
Well the <b>ocean</b> took my <b>baby</b>	A' (?)	[beɪ.bi]
My <b>baby</b> took my heart from me	A'	
Yeah the <b>ocean</b> took my <b>baby</b>	A'	[beɪ.bi]
My <b>baby</b> took my heart from me	A'	
She packed it up in a <b>suitcase</b>	A	[sʊt.ke.i.i:]
Lord she took it away to Italy, Italy	A'	
So I'm walking down your <b>highway</b>	A	
Just as far as my poor eyes can see	A'	
Yes I'm walking down your <b>highway</b>	A	[haɪ.we.i.i:]
Just as far as my eyes can see	A'	
From the Golden Gate Bridge	C	
All the way to the Statue of Liberty	A'(A?)	[li.bi.ti.i.i:]

Underlined syllables are sung to the “theme” and lexical repetitions between early and late verses are set in boldface.

Bob Dylan, © Copyright 1963, renewed 1991 by Special Rider Music.  
Reprinted with permission.

case” and “highway,” every line in the final two verses ends in [i], as in Table 3. The “return” in the fifth verse does not mean the song has ended. Rather, the final two verses subtly revisit the off-rhyme of the first verse, which ends up in a different position with regard to the “highway”-“baby” rhyme than it started (the slight upward turn of the contour in Figure 3 is because of this heightening of the vowel from “highway” to “liberty”).

Finally, the last verse is altered to include two instances of the theme. The first, occurring on the third line of the verse, is set to “highway,” a return to the initial rhyme, a word of titular focus, and something of a completion: of the three rhyming/assonant trochees in the first verse (“highway,” “baby,” and “suitcase”), each has now been set to the theme. But “highway” is also the *melodic* climax of the piece. Dylan sings its theme in falsetto, with a sweet, round tone. And though the fifth verse returns the theme to a sonorous vowel and reincorporates “ocean” into normal trochaic stress, the nasality of the falsetto “highway” might be said to complete the process of resolution from “ocean,” recontextualizing the nasal melodic segment within a more normal singing environment. Including the theme twice in one verse changes its effect. Now, instead of being (just) a climax in each verse that must resolve

in the final line, it becomes a point of resolution itself, and Dylan ends the verse and the song on the theme, in a lower register.

### Summary

A contrast between [eɪ] and [i] that is underlyingly present from the theme of the first verse, on “baby,” shifts subtly over the course of the song. The first verse establishes [eɪ] as a characteristic rhyme/assonance that is emphasized by the rhythmic emergence of the second syllable of “suitcase” and the resyllabification and diphthongization of “baby” in the theme. From verse to verse, the series of themes progresses through four verses toward highly marked syllable heads, climaxing on a syllabic nasal that is sung melodically in an upper register. The fifth verse “resolves” the progression of the first four verses, returning the theme on the sonorous [eɪ], while lexically returning to strongly established words, “baby” and “suitcase.” Though “suitcase” is an assonant return, the surrounding lines include [i] rather than [eɪ] in their final syllables, especially in the pronunciation of “baby,” which is repeated here in a new metrical position. This distinction is continued, and more fully incorporated into the overall rhyme scheme, in the final verse as the song ends on “liberty.” Like “baby,” and unlike “highway” and “suitcase,” “liberty” has a weak final syllable. It is sung to the theme on the single segment [i], not diphthongized like “baby” in the first verse.

An overall shape emerges from the interaction of musical, poetic, and sung elements. Rhythm and melody emphasize and redefine prosodically ambiguous syllables: “suitcase,” “baby,” “woman,” “trouble,” and “ocean” all receive rhythmic emphasis on their second syllables, contrary to the trochees of the underlying text and, with the exception of “-case,” in contrast to their lack of phonological stress. Such emphasis, however, produces a prosodic structure that is not found in the text and a musical structure that is not found in the melody. A small but important shift in the performance of rhyming words redefines the dominant sound in an off-rhyme that persists from the first verse (“highway”-“baby”). And a rising and falling shape is laid over the repetitive melody of the verses by a series of increasingly less sonorous syllables sung on the melodically prominent theme. This progression resolves as part of a lexical return in the fifth verse, which sets the theme to “suitcase,” whose final [eɪ] is both sonorous and a parallel link to the first verse.

Marked forms occur in prominent positions: metrical inversions of certain words and stressed placement of unsonorous segments. These forms do not arise out of random variation or frivolous rule breaking but are motivated by their position in poetic and musical frameworks that structure the song over its entirety.

## **Music and Language in the Voice**

To identify the expressive contours I outline in this analysis requires examining the singing voice in performance. Traditional musicological or poetic analyses that look to either melody and rhythm or prosody and meter would necessarily overlook the important structures that develop out of the interactions between these elements.<sup>16</sup> We might, for instance, examine a staff transcription of “Down the Highway,” with the goal of discovering the important elements of its musical form. At the small scale, the musical form within verses is clear (a melodic phrase repeats, sequenced a diatonic step lower, and the final phrase has a melismatic climax at its halfway point before resolving downward to a tonic pitch). But from verse to verse, only a minor set of elements changes, and we are left with a musical shape that is repetitive for five verses before ending with the insertion of an extra, falsetto theme in the final verse. Such a perspective presents a picture that is jagged and uneven and that does not, I think, adequately represent this music, in which the falsetto rendering of theme on “highway” is a surprise, to be sure, but not jarring or out of context. Though there is no a priori reason to expect development, climax, and resolution from a performance,<sup>17</sup> it is unmistakable to me that the fourth verse of “Down the Highway” is lexically, poetically, musically, and syllabically exceptional (climactic), in ways that are contextualized by parallelism and progressions through the preceding several verses. Dylan’s performance emphasizes musical and prosodic elements of the melody, rhythm, and lyrics that structure the interrelations of these parameters with a uniquely musico-poetic shape. The cross-cultural salience of parallelism in verbal art suggests that, even in additive stanza- or strophe-based genres, large-scale forms link and structure the material of verbal art across potentially great spans of time.

An analysis of the page-bound text alone would miss the primary assonant relationship in the song, between [eɪ] and [i]. The text sets “highway” and “baby” in an off-rhyming relation in the first verse. But in the performance, that relation is not an off-rhyme but a perfect rhyme that changes over time into an off-rhyme. The text as such is certainly relevant, especially as it reflects the contrast in stress between the second syllables of “baby” and “highway” that apparently motivates the different pronunciation of “baby” in its various metrical positions. But the poetics of the text in Dylan’s pronunciation are very different than in the text on the page: the page can only represent the underlying off-rhyme, while the performed version has access to a great variety of possible pronunciations. That the shift from [eɪ] to [i] occurs both in the frequency of words that contain these vowels and in the pronunciations of individual words is evidence that singing shapes and is shaped by poetic assonance. Further, the metrical similarities between the words in which this

rhyme-shift occurs, “liberty” and “baby” (with unstressed final syllables, in contrast to “highway” and “suitcase”), and their prominence at the beginning and end of the song suggest that this shift from [eɪ] to [i] is not (just) random or phonetic, but is in fact part of the overall structure of the song.

Separate poetic and musical analyses might account for some forms. The rhythmic lengthening of “-case” in the first verse links poetic parallelism—shared assonance—with musical rhythm in a way that might be seen by overlaying a musical analysis with a poetic one. But to posit such a linkage requires assuming that musical form follows poetic form: rhythms accommodate poetics. It is not clear whether such an account can respond to musically conditioned poetic structures.

The musical conditioning of poetics is most clear in the series of theme syllables. The unstressed final syllables of a handful of lines (“baby,” “woman,” “trouble,” and “ocean”) have little more than a statistical relationship in the poetic text. They are each the weak-final end of the fifth line of successive verses, and all but one is conspicuously unrhymed. If in poetry “unstressed equals unstress” (Jakobson 1960:358), then the gradient sonority relation between the final syllables of this series is obscured behind categorical stress distinctions. Why not rhyme these lines? What is it about the line-final placement of “woman,” “trouble,” and “ocean” that is so important that it takes precedence over the otherwise consistent rhyme scheme? It is only in the context of the musical setting, and particularly the song’s performance, that we can understand the motivation for the poetic structure.

Musical parallelism frees unstressed syllables from their equivalence relations. The theme in each verse sets these syllables in prominent positions, and the repeating melody projects, not the principle of equivalence, but a principle of gradient similarity, from the axis of selection (the final syllables of “baby,” “woman,” “trouble,” and “ocean”) onto the axis of combination (the end of the penultimate line of a repeating musical and prosodic verse structure). The sameness of the repeating melody begs the question of what is different. Dylan’s performance frames the answer by rearticulating the lyrics’ prosody (performing reduced, weak-final lines as fully pronounced, strong-final lines) to draw out elements of comparison unrealized in the lines of verse.

It is uniquely a phenomenon of singing’s mediation of music and poetry that Dylan preserves something hidden in the text—the quality of unstressed vowels to which there is no good reason to be faithful—in his pronunciation. To sing “trouble” [trʌbɫ] and “ocean” [oʃn:] is to violate any principles of syllable harmony in singing that we might propose, which would certainly require replacing [ɫ] and [n:] with vowels, [ʌ:l] and [ʌ:n]. Dylan’s performance asserts the primacy of the voice in this expressive relation: the voice has full authority over the musical and poetic texts and their interrelations. An analysis that does not account for the containment of music and poetry

within the domain of the voice misses a remarkably clear example of a distinctly musico-poetic progression that builds large formal structures out of tiny musical and poetic elements.

### **“Hesitations Between Sound and Sense”: Genre, Intertextuality, and Listening**

This analysis focuses on the sounds of language in musical form. Syntactic or lexical values are included only indirectly: syntactic phrasing influences phonological phrasing (Inkelas and Zec 1995; Selkirk 1995), which plays a significant role in the musical phrase; also, certain words are seen as placeholders (“suitcase,” “highway,” “baby”), markers that refer structurally back to their original occurrence in the song, but attention to their referential content is limited to that parallelistic function. Nonetheless, language sounds function in speech primarily to facilitate communication, and I certainly do not claim that form is independent of meaning or semantic content. Rather, analysis that truly seeks to understand the totality of musical forms and their reception must include the semantic meaning of a song’s text.<sup>18</sup> Jakobson makes this argument with regard to the sounds of language in verse:

No doubt, verse is primarily a recurrent “figure of sound.” Primarily, always, but never uniquely. Any attempts to confine such poetic conventions as meters, alliteration, or rhyme to the sound level are speculative reasonings without any empirical justification. The projection of the equational principle into the sequence has a much deeper and wider significance. Valéry’s view of poetry as “hesitation between the sound and the sense” [Valéry 1958] is much more realistic and scientific than any bias of phonetic isolationism. (1960:367)

“Down the Highway” is quite evocative, and its imagery of a failed, gambling, brokenhearted, walking traveler with a distant destination and even more distant lover limns topological territories that call on real and imagined traditions in American folklore, hillbilly music, and African American blues, representing a perspective on authentic verbal expression that is historically tied to the early-1960s folk revival (Gill 1998; Hajdu 2001; Dylan 2004; Weissman 2005). Many of the most salient phonetic and phonological structures I identify as part of the song’s expressive structure—notably syncopation and nasality—are themselves vocal tropes that index complex and highly charged relationships of race, class, and place in mid-century American culture. In the matter of syncopation, the analysis of textsetting “grammar” on which I base my own analysis of Dylan’s rhythms (Hayes and Kaun 1996; Hayes Forthcoming b) works from an admittedly small data set. African American music, and the blues in particular, were in all likelihood not well represented in the musical backgrounds of Hayes and Kaun’s respondents. Informants fluent in other English-language song traditions

would probably have produced different ideal settings and therefore different “grammars” of textsetting.

In the structures of expressive performance, genre, in which style and form become markers of musical communities of practice, provides a context that complements grammar. The recording analyzed in this article is the product of an ambitious young singer’s self-conscious and unremitting acquisition of performance styles. Dylan’s vocal technique combines styles of talking and singing from his native Midwest, from a mediated and commodified version of American “hillbilly” music, from ballads collected and recorded by academic folklorists, and from an idealization of African American blues. Without taking on the overwhelming task of sorting out the influences on his singing, we can briefly note that Dylan admired and learned from singers such as Hank Williams, Woody Guthrie, and Dave Van Ronk. Williams’s striking nasality and Guthrie’s metrical flexibility are models Dylan consciously imitated, and Van Ronk provided a model for stylistic fluidity (Dylan 2004:15). Ballad and blues song forms were models for Dylan’s early compositions, and his provocative and unforgiving singing stylized vocal tropes already familiar in these genres. In a environment saturated with country, folk, and blues, Dylan’s oral constriction is not only “ungrammatical,” it is also intertextual, and Dylan’s vocal performance is as much a complex and changing index of genres of American vernacular music as it is an elegant play on the grammatical expectations of Standard English. The character and scene evoked by Dylan’s lyrics complements the performance traditions evoked by his singing. Genre and grammar, therefore, demand mutual understanding, and a recognition of the role of verbal grammar in singing requires collapsing some of the strict boundaries that separate “underlying” linguistic structures from contextualized and responsive performance.

While it is important to identify the social, historical, and intertextual contexts that give rise to a specific performance, it is equally important to note that, in a musical culture that depends upon recording and circulation of commodified songs, the sites of listening and consumption are potentially independent of the social and historical contexts that give rise to performances.<sup>19</sup> A compelling example of such mediated reception comes from my experience working as an elementary and middle school teacher in a rural school in southern Vermont. Early in the term in the fall of 2002, I presented “Down the Highway” as the subject of a lesson that explored the relationship between text and performance and between poetry and song. I asked fifth- to eighth-grade students to examine a sheet with the lyrics to “Down the Highway” on it, and I instructed them to determine whether the words represented a poem or a song. My students responded enthusiastically, offering plausible explanations for either designation based on meter, rhyme, the length of lines and of the entire piece, and the song’s content. (“It’s a

love poem,” said one, to which another responded “why can’t it be a love song?” and a third interjected, “it’s not a love song, it’s about his broken heart, look here—”.) Several students volunteered to improvise sung and spoken performances to illustrate their positions. In accord with the claims of this article, the lyrics on the page did not effectively guide their performances to something like Dylan’s. Nonetheless, I strongly suspected that as my students grew older and were exposed, even in passing, to more and more music, many of them would gain the knowledge of song forms and genre that immediately specify the phrase structure of “Down the Highway” as “blues.”

As the students finally tired of debating the point and asked me to give them the answer, I put on the CD. After the first verse of “Down the Highway,” it took five minutes to bring the class back to order. My students immediately forgot the motivating problem of genre and began to snarl, groan, and whine along with the recordings. They were compelled to get out of their seats and run around, to lift their heads and howl (several in protest at having to listen to this “terrible” music). Dylan’s grunts and snarls were so far removed from the expectations of the page, from any of my students’ attempts at translating the printed words into song, that the sounds required notice and participation. Engaged in a challenging if traditionally structured exercise in textual analysis, the class anticipated the sort of logical resolution of the written word and its problems that their school experience had taught them to expect. Instead they encountered what they must have heard as incoherent wailing—barely language—that inverted and transformed their foundational and daily process of translating from page to speech. For the rest of the term students would ask me to play this song, and each time I did they would excitedly reenact the pandemonium of their first experience, jumping up, singing and squealing along. Several students developed passable, if caricatured, impersonations sung entirely through the nose. These performances seemed necessarily to include an awkwardly extended neck, twisted shoulders, and a big sneer, suggesting an extraordinary figure called to mind by Dylan’s voice.

My students did not need to hear the structure of the whole song to experience it as marked; they reacted with such disorder before the recording ever reached the constricted syllables at “trouble” and “ocean” upon which much of my analysis depends. While for educated listeners Dylan’s vocal style may index a complex network of musical influence, for my middle school students it signified decontextualized “weirdness.” Something about Dylan’s voice, perhaps in relief against the rendering they imagined from the page, represented a prodding, powerful strangeness to which my students felt compelled to add their voices.

As they isolated and mimicked particular vocal tropes that stood out to them as affectively powerful and stylistically characteristic of the performance—nasality, constriction, vocal strain, syncopation—my students’ recep-

tion is exemplary of what Adorno called “atomistic listening” ([1938] 1991). Habits of listening that focus on musical style—arrangement and orchestration, timbre, “the memorability of disconnected parts, thanks to climaxes and repetitions” (ibid.:41)—rather than internal structure are characteristic of commodity listening, which “snatches the reified bits and pieces out of their context and sets them up as a potpourri. It destroys the multilevel unity of the whole work and brings forward only isolated popular passages” (ibid.). While I explore what might be called the “multilevel unity of the whole work” of “Down the Highway,” I am aware that this sort of coherent musical form may not necessarily inform reception. Neither “form” nor “grammar” is independently capable of accounting for these real-world responses to a recording, and neither category has the capacity to represent the breadth and complexity of this music’s historically rich intertextuality in style and content. But in reacting to Dylan’s nasality and syncopation, my students responded to some of the same tropes of vocal performance that position Dylan’s style in its cultural and musical contexts and that stretch the bounds of phonology and phrasing. Style (e.g., a culturally significant proclivity to oral constriction) can motivate structure (the parametrization of oral constriction to develop a performance’s “shape”)—and vice versa. The boundaries of “grammar,” “music,” “poetry,” “genre,” and “style” are invariably blurred in practice, but together they contribute to a vocal gestalt that can have practical social power to produce movement, interaction, and even more expressive vocalization.

## Conclusion

Language’s role in singing remains a largely unexplored but necessary site of analysis if we are to understand the relationships between discursive fields like music and language and their embeddedness in cultural practices. Any vocal music is built upon features, segments, syllables, and feet—forms that articulate both structure and style. The presence of language sounds in music integrates verse and song in ways that cut across the near-totality of each, exposing assumptions and yielding insights about the boundaries between speech and song. If singing is itself a layer of discourse built on top of language, this analysis develops hypotheses about the interrelations of these levels, consistent with the findings of previous scholars of expressivity and discourse. In the essay that inspires this article’s title, Jakobson ([1960] 1987) explores the role of grammar in poetic forms, noting the integration of hierarchically separated linguistic levels in poetry: grammatical parallelism and poetic parallelism occur simultaneously and in complementary ways. Anthony Woodbury updates Jakobson’s approach to interacting levels of linguistic hierarchies, finding in Central Alaskan Yupik talk that the “rhetori-

cal structure” of large segments of discourse brings “together separate but strongly interacting components, *viz.* a prosodic component, a syntactic component, and others as required by the language in question” (1985:155), such that when syntax and intonation pattern together, their interrelation provides important markers for the structure of discourse above the level of the utterance.

Woodbury argues that “by regarding prosody and syntactic constituency as interacting components of a larger rhetorical structure, it is possible to accurately reflect the exact degree of formal and functional overlap they display” (1985:160). The relevance of such a framework for the relations of music and language are clear: by regarding musical and poetic constituency as interacting components of a larger song structure, the intersections and boundaries between the two components are revealed. These intersections appear to involve the intuitive relationships between musical and prosodic rhythm, the sound quality of notes and syllables as an integration of timbre and assonance, and a tension between gradient and categorical relations in the structures of artistic form.<sup>20</sup> Poetic and musical structures should be understood in parallel with “rhetorical structure,” as large-scale forms that have top-down effects on the utterances they comprise. Optional or expressive frameworks into which language and music are fitted must be understood in terms of large-scale hierarchies of verbal discourse.

The sense of the term “competence” as it is used by anthropologists and folklorists in reference to the social and pragmatic uses of language in performance (Hymes 1972a; Bauman [1977] 1984; Briggs 1988) can be reconciled in part with the term’s formal linguistic (Chomskian) usage against which Hymes and others push. A substantial amount of competence in expressive performance lies in the ability to work within forms—to limit possible utterances to those that are acceptable within the verse forms and musical frameworks in which verbal art often is found. Whether these forms might be fully incorporated into a systematic grammar is an open question (I am inclined to doubt it), and it is clear that much performative competence lies in the ability to communicate effectively in particular situations with particular people—not, except trivially, a function of grammar. But verbal art is also always engaged with grammar, whether by stretching limits or enforcing new patterns and judgments. In this way, communicative competence is a matter of both performance and grammaticality: singers/speakers must fit their utterances into expressive frameworks that may be developed on the spot, but that are nonetheless “formal” in that they constrain those utterances by making reference to word order, prosody, segmental quality, and other linguistic constituents.

In the end, any analysis must be accountable to experience: to histories of reception, to the judgments of audiences, and to phenomenologies of

listening—the socially mediated but individually felt experiences of a song, a performance, or a recording in time and in memory. In part, my appeal to grammar in this article is motivated by a need to hold formal analysis accountable to experience. Grammar has an inherent phenomenal resonance in discourse and practice. Though professional linguists might object that the structures of grammar are “unconscious,” their discipline has at its empirical and theoretical foundation a feedback loop that runs through explicit judgments by speakers of the “grammaticality” of utterances. So while the grammatical structures of everyday language may rarely come to the surface of consciousness—a claim undermined by the constant presence in discourse of jokes, word play, and lines of song and poetry (Kirshenblatt-Gimblett and Sherzer 1976)—linguistics necessarily assumes that the “ungrammaticality” of an utterance will always be immediately perceptible to a competent speaker.

Therefore markedness, judgments about which mediate evaluations of grammaticality and ungrammaticality, is crucial to the phenomenological salience of grammar. It is for this reason that marked structures like metrically inverted feet and orally constricted syllables play such a prominent role in the musico-poetic forms of Dylan’s singing. Markedness does not (just) rely on cognitive or perceptual systems, so divorced from discourse’s concrete field of action. Rather, grammar and markedness—whatever their cognitive basis—are fundamentally *semiotic*, and as such they form a vital part of the cultural and phenomenological mediations of daily life (Mertz and Parmentier 1985).

The conflict between the sensible and sensual markedness of Dylan’s singing and the powerfully expressive normalization of that markedness into Dylan’s recognizable style produces the “felt ‘naturalness’” (Feld 1988) of his musical aesthetics and the musical experience of the recording. Feld’s phrase is intended to suggest not the inherent qualities of a given style, but the cultural practices that negotiate and frame experience to produce the feeling of an unmediated iconicity between aesthetic theory and practice. The unconscious awareness of grammar by speakers gives linguistic variation its expressive power: never explicitly objectified by discourse, “stylistic” idiosyncrasies by singers remain at once perfectly “natural” and phenomenologically unavoidable.

This space between the forms and structures that govern musical or linguistic utterances and their never-perfect realizations in performance consists of what Charles Keil has called “participatory discrepancies” (1987, 1995). I think it is appropriate here to extend Keil’s theory beyond musical rhythm and texture to the nuances of Dylan’s phonetic realization of phonologically governed utterances and to the experience of such nuance by listeners who are attuned to linguistic variability while largely unaware of it. Keil notes that it is this semiconscious character of such performance tropes that mark them

as having “originary or active participatory power” (1987:275, citing Barfield 1965). By taking seriously the scientific and decontextualized findings of generative linguistics in pursuit of a better understanding of situated discourse, performance, and expressivity, we might gain deeply human insight into the phenomenology of language: the participatory discrepancies of grammar in performance mediate the sensible and the sensual experience of language for its users. So as Dylan’s performance finds expressive effect in the tactical stretching of grammatical bounds, he participates in what Michel de Certeau saw in discourse as a struggle between *parole* and *langue*: the “appropriation, or reappropriation, of language by its speakers” (1984:xiii)—and, we might add, by its singers.

### Appendix. Guide to symbols.

Characters not included in this chart are pronounced as in English.

#### Vowels

a	f <u>a</u> ther
ɛ	b <u>e</u> t
i	b <u>e</u> e <u>t</u>
ɪ	b <u>i</u> t
ɪ	st <u>a</u> t <u>i</u> c
ə	ab <u>o</u> ut
ʌ	b <u>u</u> t
u	b <u>oo</u> k

#### Consonants

ʃ	sh <u>ar</u> p
ɹ	r <u>e</u> d
g	g <u>e</u> t
ʤ	ju <u>dg</u> e
ʔ	uh_oh (glottal stop)
θ	w <u>it</u> h
r	ed <u>it</u> ed

#### Diacritics

k <sup>h</sup>	aspirated	(k <u>i</u> te)
i:	long	(b <u>e</u> e <u>t</u> )
ã	nasal	(r <u>a</u> pp <u>ro</u> ch <u>e</u> m <u>e</u> n <u>t</u> )
ŋ	syllabic	(b <u>u</u> tt <u>o</u> n)
ú	primary stress	(s <u>u</u> it <u>ˈ</u> case)
à	secondary stress	(sui <u>ˈ</u> t <u>ˌ</u> case)
.	syllable break	(po.ny)

## Diphthongs

aɪ	<u>ti</u> e
eɪ	<u>da</u> y
au	<u>no</u> w

## Acknowledgements

I am especially grateful to Maria Gouskova for generously sharing her time and expertise in phonology. Many thanks to John Myers, Jamie Hutchinson, and Larry Wallach for their guidance in the early stages of this project, and to Aaron Fox, Fred Lerdahl, Ellen Gray, Susan Boynton, Karen Hiles, and Maria Sonevtsky for their helpful comments. The generosity of my teachers, friends, and colleagues is a gift; they are not responsible for any failures or shortcomings in this work. Thanks to Special Rider Music for permission to reprint sections of the lyrics to "Down the Highway." I welcome correspondence regarding this article: tb2139@columbia.edu.

## Notes

1. Despite a decrease in ethnomusicological interest, investigations of related issues, such as the relationship between sound and meaning in vocal performance, continue to be undertaken by linguists (e.g., Hinton, Nichols, and Ohala 1994).

2. My use of the term "discourse" refers to linguistic and communicative practices such as speech, singing, and performance, following the "discourse-centered approach to language and culture" (Sherzer 1987; Urban 1991) and in the tradition of the ethnography of speaking (Hymes 1962; Bauman and Sherzer 1975). This anthropological approach to the discursive construction and mediation of social institutions is not incompatible with the Foucauldian sense of "discourse" as fields of knowledge and power.

3. Halle and Keyser (1966, 1971) first laid out metrics as a problem for generative linguistics. See Hayes (Forthcoming a) for a brief overview of the theoretical developments in linguistic metrics since the 1970s.

4. Syllable length is commonly associated with rhythmic duration in many singing traditions, such as Hill and Podstavsky (1976) find in Hausa praise-singing, and Bright (1957, 1963) and Powers (1980) identify in South Indian song. The constraint associating phrase edges can be seen as a formal statement of Herzog's claim that in European folk song "text line and musical line almost invariably coincide" (1950:1040).

5. This is not to say that gradience or progressive forms are never present in verbal art, and music, of course, is not just pitch, timbre, rhythm, and dynamics, nor do "crescendo," "accelerando," and "sequencing" necessarily represent universally salient forms. My argument is that the poetic function of language specifically—the "project[ion] of the principle of equivalence from the axis of selection into the axis of combination" (Jakobson 1960:358)—follows phonology in not making gradient distinctions.

6. In phonology, an Optimality Theoretic (Prince and Smolensky [1993] 2002) approach that formalizes constraints on text settings would account for such observations with some ease: just as top-down effects are seen in utterance-level phonology all the time (such that phrase structure affects syllable or foot structure), it is a simple matter to express a notion of large-scale poetic or musical "well-formedness" as a constraint on large-scale structures that dominates constraints on smaller-scale structures. The problem is that poetic and musical forms tend to be context specific. It is not clear that at the super-utterance level there are regular poetic or musical forms with consistent constituents. I leave the question of the universality of expressive forms to further research and for now I will merely note the specific ways such forms structure utterances in this song.

7. "Down the Highway" by Bob Dylan. © Copyright 1963, renewed 1991 by Special Rider Music. Reprinted with permission.

8. Distinct grids for periodic musical rhythm and relative metrical stress on syllables would be more precise, but in this case mostly redundant. Grid notation itself is properly attributed to Liberman (1975) and Lerdahl and Jackendoff (1977).

9. The syncopation of "suitcase" here is probably better represented as a dotted eighth followed by a sixteenth-note-early attack for "-case." The metrical flexibility of the performance makes such precise distinctions difficult to ascertain. "-Case" (0.45 sec) is fifty percent longer than "suit" (0.3 sec), not three times as long. The diphthong [ei] in "-case," however, is nearly three times as long as the [u] in "suit" because the [s] is weirdly lengthened in the first syllable. The larger values are justified by the absolute durations: the total duration of "suitcase" (.75 sec) is basically equal to that of "with my" (0.3 sec) plus "in my" (approx 0.4 sec, depending on where the [z] at the end of "-case" associates). In the conclusion to this article I appeal to Keil's theory of "participatory discrepancies" (1987, 1995), in which small variations such as these, not so substantial as to alter the structural representation but still phenomenally salient, are seen as central to music's processual "power."

10. "Note" designates a much less precise domain than "syllable." The combination of pitch and duration into bounded constituents (notes) is not an entirely regular occurrence and in many cases depends upon inflectional articulation of rhythmic boundaries and relatively imprecise segmentations of pitch scales. For instance, one of the cues of melodic movement in Dylan's melismatic performance of "baby" is the change in vowel quality from [e] to [i]. Without overemphasizing the indeterminacy of musical constituents, it is worth taking account of their frequent non-irreducibility. The apparent "psychological reality" of the musical note to Western ears remains empirically questionable as a universal, and syllable-note alignment may well be culturally specific. On the other hand, the syllable may be the best cross-cultural diagnostic of the "note" available, as in Charles Seeger's "working hypothesis" about singing that "the smallest unit of form is the 'sung syllable'" (1958:8-10). Porcello notes, "monosyllables are employed by speakers as optimal mappings of musically articulated sounds; just as a musical sound consists of an envelope comprising an attack, a sustain, and a decay, so does the monosyllabic utterance" (Feld, Fox, Porcello, and Samuels 2004:326), and his findings, along with Traube's (2004), are suggestive of the importance of CV(C)-syllable templates to Western conceptions of musical sounds. Also relevant is the extensive ethnomusicological literature on "vocables" (especially in Native American music), meaningless syllables that are musically significant and phonologically conventional (e.g., Nettle 1953, 1954; Hymes 1965; Frisbie 1980; Hinton 1980, 1984).

11. "Weak-final" and "strong-final" in this usage refer to poetic feet, rather than phonological stress, and so do not distinguish between the stress of "highway" and that of "baby."

12. Note the relevance of this argument to the music-theoretic concept of "prolongation." Metrical, melodic, and poetic relationships at the level of the verse set apart and draw connections between the fifth line endings. The processes that set these syllables in parallel are effectively prolongations of the larger-scale progressions from verse to verse. Thanks to Fred Lerdahl for noting this connection (p.c., 25 August 2006). In his recent analyses of Robert Frost's short poem "Nothing Gold Can Stay" (2001, 2004), Lerdahl interprets the prosodic hierarchy as representing phrase structures analogous to those of his Generative Theory of Tonal Music (Lerdahl and Jackendoff 1983). Through an approach to "timbral prolongation" in poetic parallelism, Lerdahl identifies sound structures in Frost's poem that are remarkably similar to the "sonority contour" that I identify in this section.

13. Low vowels: [a]; mid vowels: [e], [o]; high vowels: [i], [u]; liquids: [l], [r]; nasals: [m], [n]; voiced fricatives: [z], [v]; unvoiced fricatives: [s], [f]; voiced stops: [b], [d]; unvoiced stops: [p], [t].

14. In fact, Hinton's study identifies phonological restrictions on singing that are independent of those on normal speech. For example, Hinton (1980) notes that while the total vowel inventory of Havasupai speech and singing are the same—*a, e, i, o, u*—the relative frequency of mid and high vowels is reversed. In speech, the vowels *a, i, and u* occur more than twice as

frequently, consistent with a common cross-linguistic prevalence of “peripheral” vowels. But in Havasupai singing, the frequencies are reversed almost exactly, such that *e* and *o* replace *i* and *u*. Increased sonority is apparently of greater importance in Havasupai singing than peripherality. Sundberg (1969) similarly finds in a laboratory context that sung vowels are consistently lower than spoken vowels.

15. *American Heritage Dictionary*, 4th ed., s.vv. “woman,” “ocean,” “trouble,” “button,” “bottle.”

16. This argument is more or less consistent with Betsy Bowden’s ([1982] 2001) early emphasis on Dylan’s singing performance in her literary analyses.

17. Peter Manuel comments that “formal structures seeking development, climax, and closure are distinctively modern bourgeois creations” (2006:98).

18. Fox contributes such an analysis (Feld, Fox, Porcello, and Samuels 2004:328–32), convincingly correlating musical and performative gestures with the narrative of a song text.

19. Whether musical recording and commercial circulation constitute “schizophonia” (Feld 1994) or simply represent another mode of de- and re-contextualization of musical performance and the performance of listening (Bauman and Briggs 1990; Silverstein and Urban 1996), any account of a recorded music’s “context” cannot exclude commerce and technology. My own interest in this music is as disconnected from Dylan’s cultural milieu as that of the children I will describe here.

20. These conclusions are in agreement with those drawn by Lerdahl (2001, 2004), who proposes that the areas of similarity and difference in musical and linguistic structures suggest overlapping cognitive and brain functions.

## References

- Adorno, Theodor W. [1938] 1991. “On the Fetish Character of Music and the Regression of Listening.” In *The Culture Industry: Selected Essays on Mass Culture*, edited by J. M. Berstein, 29–60. New York: Routledge Classics.
- Bakhtin, Mikhail M. [1975] 1981. “Discourse in the Novel.” In *The Dialogic Imagination: Four Essays by M. M. Bakhtin*, edited by Michael Holquist, 259–422. Translated by Caryl Emerson and Michael Holquist. Austin: University of Texas Press.
- Barfield, Owen. 1965. *Saving the Appearances: A Study in Idolatry*. New York: Harcourt Brace Jovanovich.
- Bauman, Richard. [1977] 1984. *Verbal Art as Performance*. Prospect Heights, Ill.: Waveland Press.
- . 2002. “Disciplinary, Reflexivity, and Power in *Verbal Art as Performance*: A Response.” *Journal of American Folklore* 115(455):92–98.
- Bauman, Richard, and Charles L. Briggs. 1990. “Poetics and Performance as Critical Perspectives on Language and Social Life.” *Annual Review of Anthropology* 19:59–88.
- Bauman, Richard, and Joel Sherzer. 1975. “The Ethnography of Speaking.” *Annual Review of Anthropology* 4:95–199.
- Blacking, John. 1982. “The Structure of Musical Discourse: The Problem of the Song Text.” *Yearbook for Traditional Music* 14:15–23.
- Boucher, David, and Gary Browning, eds. 2004. *The Political Art of Bob Dylan*. New York: Palgrave Macmillan.
- Bowden, Betsy. [1982] 2001. *Performed Literature: Words and Music by Bob Dylan*. Lanham, Md.: University Press of America.
- Briggs, Charles L. 1988. *Competence in Performance: The Creativity of Tradition in Mexicano Verbal Art*. Philadelphia: University of Pennsylvania Press.
- . 1993. “Personal Sentiments and Polyphonic Voices in Warao Women’s Ritual Wailing: Music and Poetics in a Critical and Collective Discourse.” *American Anthropologist* 95(4):929–57.

- Bright, William. 1957. "Singing in Lushai." *Indian Linguistics* 17:24-28.
- . 1963. "Language and Music: Areas for Cooperation." *Ethnomusicology* 7(1):26-32.
- Caton, Steven C. 1987. "Contributions of Roman Jakobson." *Annual Review of Anthropology* 16:223-60.
- de Certeau, Michel. 1984. *The Practice of Everyday Life*. Translated by Steven Rendall. Berkeley: University of California Press.
- Corcoran, Neil, ed. 2002. *"Do Your, Mr. Jones?": Bob Dylan with the Poets and Professors*. London: Chatto and Windus.
- Dylan, Bob. 2004. *Chronicles: Volume One*. New York: Simon and Schuster.
- Eliot, T. S. 1942. *The Music of Poetry*. Glasgow: Jackson, Son, and Company.
- Fales, Cornelia, and Harris Berger. 2005. "'Heaviness' in the Perception of Heavy Metal Guitar Timbres: The Match of Perceptual and Acoustic Features." In *Wired for Sound: Engineering and Technologies in Sonic Cultures*, edited by Thomas Porcello and Paul D. Greene, 181-97. Middletown, Conn.: UPNE/Wesleyan University Press.
- Feld, Steven. 1974. "Linguistic Models in Ethnomusicology." *Ethnomusicology* 18(2):197-217.
- . 1984. "Sound Structure as Social Structure." *Ethnomusicology* 28(3):383-409.
- . 1988. "Aesthetics As Iconicity of Style, or 'Lift-up-over Sounding': Getting Into the Kaluli Groove." *Yearbook for Traditional Music* 20:74-113.
- . 1994. "From Schizophrenia to Schismogenesis: On the Discourses and Commodification Practices of 'World Music' and 'World Beat.'" In *Music Grooves: Essays and Dialogues*, by Charles Keil and Steven Feld, 257-89. Chicago: University of Chicago Press.
- Feld, Steven, and Aaron A. Fox. 1994. "Music and Language." *Annual Review of Anthropology* 23:25-53.
- Feld, Steven, Aaron A. Fox, Thomas Porcello, and David Samuels. 2004. "Vocal Anthropology: From the Music of Language to the Language of Song." In *A Companion to Linguistic Anthropology*, edited by Alessandro Duranti, 321-45. Malden, Mass.: Blackwell.
- Fox, Aaron A. 1992. "The Jukebox of History: Narratives of Loss and Desire in the Discourse of Country Music." *Popular Music* 11(1):53-72.
- . 2004. *Real Country: Music and Language in Working-Class Culture*. Durham, N.C.: Duke University Press.
- Frisbie, Charlotte J. 1980. "Vocables in Navajo Ceremonial Music." *Ethnomusicology* 24(3):347-92.
- Gafos, Adamantios I. 2002. "A Grammar of Gestural Coordination." *Natural Language and Linguistic Theory* 20(2):269-337.
- Gill, Andy. 1998. *Don't Think Twice, It's All Right: Bob Dylan, The Early Years*. New York: Thunder's Mouth Press.
- Graham, Laura. 1984. "Semanticity and Melody: Parameters of Contrast in Shavante Vocal Expression." *Latin American Music Review* 5(2):161-85.
- . 1987. "Three Modes of Shavante Vocal Expression: Wailing, Collective Singing, and Political Oratory." In *Native South American Discourse*, edited by Joel Sherzer and Greg Urban, 83-118. Berlin; New York: Mouton de Gruyter.
- Gray, Michael. 2000. *Song and Dance Man III: The Art of Bob Dylan*. London; New York: Continuum.
- Hajdu, David. 2001. *Positively 4th Street: The Lives and Times of Joan Baez, Bob Dylan, Mimi Baez Fariña, and Richard Fariña*. New York: Farrar, Strauss, and Giroux.
- Halle, John. 1999. "A Grammar of Improvised Textsetting." Ph.D. dissertation, Columbia University.
- Halle, John, and Fred Lerdahl. 1993. "A Generative Textsetting Model." *Current Musicology* 55:3-23.
- Halle, Morris, and Samuel Jay Keyser. 1966. "Chaucer and the Study of Prosody." *College English* 28(3):187-219.
- . 1971. *English Stress: Its Form, Its Growth, and Its Role in Verse*. New York: Harper and Row.

- Hayes, Bruce. Forthcoming a. "Faithfulness and Componentiality in Metrics." In *The Nature of the Word: Essays in Honor of Paul Kiparsky*, edited by Karen Hanson and Sharon Inkelas. Cambridge, Mass.: MIT Press.
- . Forthcoming b. "Textsetting as Constraint Conflict." In *Towards a Typology of Poetic Forms*, edited by Jean-Louis Aroui and Andy Arleo. Amsterdam: Elsevier.
- Hayes, Bruce, and Abigail Kaun. 1996. "The Role of Phonological Phrasing in Sung and Chanted Verse." *The Linguistic Review* 13:243-303.
- Herzog, George. 1934. "Speech Melody and Primitive Music." *Musical Quarterly* 20(4):452-66.
- . 1942. "Text and Melody in Primitive Music." *Bulletin of the American Musicological Society* 6:10-11.
- . 1950. "Song." In *Funk and Wagnalls Standard Dictionary of Folklore, Mythology and Legend*. Vol. 2, edited by Maria Leach, 1032-50. New York: Funk and Wagnalls.
- Hill, Clifford, and Sviatoslov Podstavsky. 1976. "The Interface of Language and Music in Hausa Praise-Singing." *Ethnomusicology* 20(3):535-40.
- Hinton, Leanne. 1980. "Vocables in Havasupai Song." In *Southwestern Indian Ritual Drama*, edited by Charlotte J. Frisbie, 275-305. Albuquerque: University of New Mexico Press.
- . 1984. *Havasupai Songs: A Linguistic Perspective*. Turbingen: Gunter Narr Verlag.
- Hinton, Leanne, Johanna Nichols, and John J. Ohala, eds. 1994. *Sound Symbolism*. Cambridge; New York: Cambridge University Press.
- Hymes, Dell. 1962. "The Ethnography of Speaking." In *Anthropology and Human Behavior*, edited by Thomas Gladwin and William C. Sturtevant, 13-53. Washington, D.C.: Anthropological Society of Washington.
- . 1965. "Some North Pacific Coast Poems: A Problem in Anthropological Philology." *American Anthropologist* 67(2):316-41.
- . 1972a. "On Communicative Competence." In *Sociolinguistics: Selected Readings*, edited by J. B. Pride and Janet Holmes, 269-93. Harmondsworth, U.K.: Penguin Books.
- . 1972b. "The Contribution of Folklore to Sociolinguistic Research." In *Toward New Perspectives in Folklore*, edited by Américo Paredes and Richard Bauman, 42-50. Austin: University of Texas Press.
- . 1977. "Discovering Oral Performance and Measured Verse in American Indian Narrative." *New Literary History* 8(3):431-57.
- . 1980. "Verse Analysis of a Wasco Text: Hiram Smith's 'At'Unaqa.'" *International Journal of American Linguistics* 46(2):65-77.
- . 1981. *In vain I tried to tell you": Essays in Native American Ethnopoetics*. Philadelphia: University of Pennsylvania Press.
- Inkelas, Sharon, and Draga Zec. 1995. "Syntax-Phonology Interface." In *The Handbook of Phonological Theory*, edited by John A. Goldsmith, 535-49. Cambridge, Mass.: Blackwell.
- Jakobson, Roman. 1960. "Closing Statement: Linguistics and Poetics." In *Style in Language*, edited by Thomas A. Sebeok, 350-77. New York: Technical Press of MIT; John Wiley and Sons.
- . [1960] 1987. "Poetry of Grammar and Grammar of Poetry." In *Language in Literature*, edited by Krystyna Pomorska and Stephen Rudy, 121-44. Cambridge, Mass.: Belknap Press of Harvard University Press.
- . [1970] 1987. "Subliminal Verbal Patterning in Poetry." In *Language in Literature*, edited by Krystyna Pomorska and Stephen Rudy, 250-66. Cambridge, Mass.: Belknap Press of Harvard University Press.
- . 1978. *Six Lectures on Sound and Meaning*. Translated by John Mephram. Cambridge, Mass.: MIT Press.
- Keil, Charles. 1987. "Participatory Discrepancies and the Power of Music." *Cultural Anthropology* 2(3):275-83.
- . 1995. "The Theory of Participatory Discrepancies: A Progress Report." *Ethnomusicology* 39(1):1-19.
- Kirshenblatt-Gimblett, Barbara, and Joel Sherzer. 1976. "Introduction." In *Speech Play*, edited

- by Barbara Kirshenblatt-Gimblett and Joel Sherzer, 1–18. Philadelphia: University of Pennsylvania Press.
- Lacy, Paul de. 2002. "The Formal Expression of Markedness." Ph.D. dissertation, University of Massachusetts, Amherst.
- Lerdahl, Fred. 2001. "The Sounds of Poetry Viewed as Music." *Annals of the New York Academy of Sciences* 930:337–54.
- . 2004. "A Music-Theoretic Approach to the Sounds of Poetry." Unpublished manuscript.
- Lerdahl, Fred, and Ray Jackendoff. 1977. "Toward a Formal Theory of Tonal Music." *Journal of Music Theory* 21(1):111–71.
- . 1983. *A Generative Theory of Tonal Music*. Cambridge, Mass.: MIT Press.
- Lieberman, Mark. 1975. "The Intonational Systems of English." Ph.D. dissertation, MIT.
- List, George. 1963. "The Boundaries of Speech and Song." *Ethnomusicology* 7(1):1–16.
- Marcus, Greil. 1997. *Invisible Republic: Bob Dylan's Basement Tapes*. New York: Henry Holt.
- . 2005. *Like a Rolling Stone: Bob Dylan at the Crossroads*. New York: Public Affairs.
- Manuel, Peter. 2006. "Flamenco in Focus: An Analysis of a Performance of Soleares." In *Analytical Studies in World Music*, edited by Michael Tenzer, 92–119. New York: Oxford University Press.
- Mertz, Elizabeth, and Richard J. Parmentier, eds. 1985. *Semiotic Mediation: Sociocultural and Psychological Perspectives*. Orlando, Fla.: Academic Press.
- Nattiez, Jean-Jacques. 1973. "Linguistics: A New Approach for Music Analysis?" *International Review of the Aesthetics and Sociology of Music* 4(1):51–68.
- . 1990. *Music and Discourse: Toward a Semiology of Music*. Translated by Carolyn Abbate. Princeton, N.J.: Princeton University Press.
- . 1999. "Inuit Throat-Games and Siberian Throat Singing: A Comparative, Historical, and Semiological Approach." *Ethnomusicology* 43(3):399–418.
- Nettl, Bruno. 1953. "Observations on Meaningless Peyote Song Texts." *Journal of American Folklore* 66(260):161–64.
- . 1954. "Text-Music Relationships in Arapaho Songs." *Southwestern Journal of Anthropology* 10(2):192–99.
- Nietzsche, Friedrich. [1872] 1995. *The Birth of Tragedy*. Translated by Clifton P. Fadiman. New York: Dover.
- Powers, Harold S. 1980. "Language Models and Music Analysis." *Ethnomusicology* 24(1):1–60.
- Prince, Alan, and Paul Smolensky. [1993] 2002. *Optimality Theory: Constraint Interaction in Generative Grammar*. ROA-537, Rutgers Optimality Archive, <http://roa.rutgers.edu/>.
- Ricks, Christopher. 2003. *Dylan's Visions of Sin*. London: Viking.
- Seeger, Charles. 1958. "Singing Style." *Western Folklore* 17(1):3–11.
- Selkirk, Elisabeth. 1995. "Sentence Prosody: Intonation, Stress, and Phrasing." In *The Handbook of Phonological Theory*, edited by John A. Goldsmith, 550–69. Cambridge, Mass.: Blackwell.
- Shank, Barry. 2002. "'That Wild Mercury Sound': Bob Dylan and the Illusion of American Culture." *boundary 2* 29(1):97–123.
- Sherzer, Joel. 1987. "A Discourse-Centered Approach to Language and Culture." *American Anthropologist* 89:295–309.
- Silverstein, Michael. 1976. "Shifters, Linguistic Categories, and Cultural Description." In *Meaning in Anthropology*, edited by Keith Basso and Henry A. Selby, 11–55. Albuquerque: University of New Mexico Press.
- Silverstein, Michael and Greg Urban. 1996. "The Natural History of Discourse." In *Natural Histories of Discourse*, edited by Michael Silverstein and Greg Urban, 1–17. Chicago: University of Chicago Press.
- Sundberg, Johan. 1969. "Articulatory Differences Between Spoken and Sung Vowels in Singers." In *Speech Transmission Laboratory Quarterly Progress and Status Report*, 33–46. Stockholm: Royal Institute of Technology.
- Tedlock, Dennis. 1983. *The Spoken Word and the Work of Interpretation*. Philadelphia: University of Pennsylvania Press.

- Traube, Caroline. 2004. "An Interdisciplinary Study of the Timbre of the Classical Guitar." Ph.D. dissertation, McGill University.
- Urban, Greg. 1985. "The Semiotics of Two Speech Styles in Shokleng." In *Semiotic Mediation: Sociocultural and Psychological Perspectives*, edited by Elizabeth Mertz and Richard J. Parmentier, 311-29. Orlando, Fla.: Academic Press.
- . 1991. *A Discourse-Centered Approach to Culture: Native South American Myths and Rituals*. Austin: University of Texas Press.
- Valéry, Paul. 1958. *The Art of Poetry*. Translated by Denise Folliot. Vol. 7 of *Collected Works*, edited by Jackson Matthews. New York: Pantheon Books.
- Weissman, Dick. 2005. *Which Side Are You On? An Inside History of the Folk Music Revival in America*. New York: Continuum.
- Woodbury, Anthony C. 1985. "The Functions of Rhetorical Structure: A Study of Central Alaskan Yupik Eskimo Discourse." *Language in Society* 14(2):153-90.

## Discography

- Dylan, Bob. [1963] 2004. *The Freewheelin' Bob Dylan*. Columbia/Legacy 90321. Compact disc.