

Blackwell Companions to Anthropology

The Blackwell Companions to Anthropology offer a series of comprehensive syntheses of the traditional subdisciplines, primary subjects, and geographic areas of inquiry for the field. Taken together, the series represents both a contemporary survey of anthropology and a cutting edge guide to the emerging research and intellectual trends in the field as a whole.

1 *A Companion to Linguistic Anthropology*

Edited by Alessandro Duranti

2 *A Companion to the Anthropology of Politics*

Edited by David Nugent and Joan Vincent

Forthcoming

A Companion to Psychological Anthropology

Edited by Conerly Casey and Robert B. Edgerton

A Companion to the Anthropology of Japan

Edited by Jennifer Robertson

A Companion to the Anthropology of American Indians

Edited by Thomas Biolsi

A Companion to Linguistic Anthropology

Edited by Alessandro Duranti

 Blackwell
Publishing

2004

- Yaasiin C. Keenadiid (1976). *Qamuuska Af-Soomaaliga*. Xamar: Wasaaradda Hiddaha iyo Tacliinta Sare iyo Akademiyada Dhaqanka.
- Yaasiin C. Keenadiid (1984). *Ina Cabdille Xasan e la sua attività letteraria*. Naples: Istituto Universitario Orientale.
- Yip, M. (1980). The Metrical Structure of Regulated Verse. *Journal of Chinese Linguistics* 8: 107-124.
- Yoder, P. B. (1972). Biblical Hebrew. In W. K. Wimsatt (ed.), *Versification: Major Language Types* (pp. 52-65). New York: New York University Press.

CHAPTER 14 Vocal Anthropology: From the Music of Language to the Language of Song

*Steven Feld, Aaron A. Fox,
Thomas Porcello, and David
Samuels*

I INTRODUCTION

There is a considerable history to research exploring relations between music and language. We begin with some of this intellectual background to better locate the research questions taken up in the body of this chapter. While these questions – about the linguistic mediation of musical and especially timbral discourse, and the connections between the singing voice and place, class, ethnicity, and identity – are very contemporary ones, they are clearly prefigured historically, in the overlapping legacies of Franz Boas, Edward Sapir, and Roman Jakobson, all of whom made a place for music in their programs for the study of the mental, semiotic, communicative, expressive, and discursive roles of language. In ethnomusicology it was their student and contemporary George Herzog, also trained as a linguistic anthropologist, who drew out the musical implications of their ideas and elaborated the importance of bringing linguistic sophistication to the social analysis of music.

A review of the key themes in the language and music literature during the formative periods of twentieth-century anthropology, linguistics, and ethnomusicology (Feld and Fox 1994) indicates how those early programs developed into four principal conversations. The first, and perhaps best known of these, is the general and abstract consideration of *music as a language*. This is the perspective that produced formal linguistic models of music, models meant to use linguistic theory to advance the analysis of music. These models have been based on analogies between the

distributional character of pitch systems and phonetic inventories, as well as analogies between syntactic structures and the harmonic, metrical, and motival organization of musical pieces. Looking at the products of these efforts, one recognizes critical differences between linguistic approaches to syntax and semantics and musical approaches to form and meaning. From a semiotic point of view, music seems far more syntactically redundant and overdetermined when compared to language. At the same time music is semantically far more diffuse and ambiguous than language. In other words, forms of repetition, cyclicity, and predictable recursiveness dominate musical structure more than language structure. In this sense music may seem a simpler formal system to describe with logical rules. But at the same time, meaning in music is notoriously more complex to formally characterize when compared to the semantic structures of language. Nonetheless, linguistically motivated models for the description and analysis of music have, since the early 1980s, produced a number of intellectually productive developments toward a cognitive science of music (Lerdahl and Jackendoff 1983; Lerdahl 2001).

A second, less theoretical and more strictly empirical conversation focuses on *music about language*, that is, of speech surrogates. As a semiotic for musicking language, speech surrogates involve the transposition of linguistic tonal and temporal contours to surrogate articulatory modes and media. These are principally instrumental, like the phenomena of "talking drums," or involve other secondary signaling systems, for example, "whistled speech" (Umiker-Sebeok and Sebeok 1976).

The substance of this chapter sidesteps these two areas, as they represent both the most general and the most specific poles of all language/music discourses. Instead we concentrate on developments in two other research conversations, the locations where the fully social implications of the language/music issues now most completely emerge. The first of these is devoted to considerations of *language about music*, namely, the intertwining of verbal and musical discourse. This research starts with the simple observation that the dominant Western conceptualization of music as a mental construct and a performance practice obscures one of its most significant social facts. Namely, music is a ubiquitous topic for discourse. Musicians and listeners everywhere spend a great deal of time and productive social energy talking about music (Feld 1994a). Three empirical domains for social analysis have unfolded from this observation. First is research on the relations between musical terminology, local theories of music, and the metaphoric basis of language about music (Feld 1990). Second is research on the intertwining of speaking and musicking as a site of social interaction among musicians. Finally, engaging debates originating in the philosophy of musical aesthetics, there is research investigating the social location of evaluative, critical, and interpretive musical discourse (Feld 1994a).

In the first case study that follows, Thomas Porcello brings together a number of these themes in his research on the management of talk about timbre, a term often glossed by the synaesthetic metaphor of "tone color." This is talk about sound qualities, an area of acoustically and socially complex verbal practices critical to the production of music. Often imagined as an "unspeakable" realm of music, where words are either imprecise or unnecessary, timbre turns out, in Porcello's empirical research, to occupy a far more central location in verbal interactions among musicians and engineers.

The final set of conversations also focus on the centrality of vocal sound to verbal significance. They link the study of *music in language* and *language in music* through a focus on song and on texted vocalization. Here we begin where poetics meets performance, namely, in the voice. Language's musicality – its tonal, timbral, prosodic, and gradient dynamic qualities – highlights the role of vocal performance for linguistic meaning. Music's language – the texted dimensions of songs and other sung poetic genres – highlights verbal art as vocal art. Critical social issues arise in these areas, namely the account of voice and its relationship to social agency, difference, social imagination, and identity. These topics are addressed in two additional case studies, below, from the Texas country song research of Aaron A. Fox, and the Apache country song research of David Samuels.

2 TALK ABOUT TIMBRE IN THE RECORDING STUDIO

Sound recording studios provide a rich locale in which music is the predominant subject of, and reason for, discourse. Contemporary popular music recording sessions involve the intertwining of musical performance and speech – often performative speech – about music (Porcello 1998). All aspects of a musical work, from its melodic, harmonic, and rhythmic structure to its instrumental arrangement and the aesthetics of its performance, are open to discussion. But it is talk about sound quality, or timbre, that receives particular emphasis.

Technological changes since the advent of sound recording have been driven by the twin goals of increased sonic fidelity between live and recorded sounds, and increased control over the manipulation of recorded sounds (Théberge 1994). As the ability to isolate and craft musical sounds has grown, so too has the amount of time and attention devoted to working on the timbral dimension of musical performances and sound recordings. As a result, talk about popular music, especially in its sites and processes of production, is saturated with vocal descriptions and depictions of musical timbre. Among musicians, engineers, producers, and other recording professionals, this timbral dimension is generally glossed by the phrase "the sound," or simply the term "sound," both of which are conceptually separated from considerations such as pitch, performance, and technique.

The prominence of talk about timbre within the sound recording industry is in direct contrast to Western academic and critical discourses about music that emphasize the formal plane of music's harmonic, tonal, and rhythmic dimensions. Musicologists often characterize discussions of timbre as mere verbal imitation or impressionistic metaphor. Alternatively, timbre is imagined as a domain specific to the scientific discourse of acousticians.¹ Such portrayals are clearly problematized in recording sessions by naturally occurring speech, used to actively negotiate the social salience of timbre to music-makers and listeners (present and eventual). In this workplace context, a structured, technical lexicon of timbre circulates publicly; knowledgeable deployment and interpretation of this lexicon provide the essential framework for playing, evaluating, recording, and mixing music. To be ignorant of this lexicon and its rules for deployment is to be seriously disadvantaged as a participant in the music-making process.

2.1 Discursive strategies

Participants in recording sessions rely on numerous strategies for talking about timbre. A careful examination of field tapes I made in studios in the southwestern United States in the mid-1990s (Porcello 1996), as well as extensive review of professional magazines, reveal five common discursive strategies used for talk about timbre during recording sessions (see also the transcriptions in Meintjes 2003, whose work in South African studios shows a remarkably similar set of strategies).

(1) Spoken/sung "vocables" are used in an attempt to iconically mimic in vocalization the timbral features of the musical sound(s) under consideration. Such sound icons include, for example, /dz:::/ to represent the sound of the snare beads vibrating against the lower head of a snare drum, /s::: ts ts/ for a triple-meter maracas pattern, and so forth. These vocables are almost always enunciated with the speaker literally performing the musical part – its pitch and rhythmic dimensions included – as accurately as possible.²

(2) Lexical onomatopoeic metaphors, or, lexical items that bear a phonological resemblance to the sound they are describing based upon acoustic properties. Examples of this large class of timbral descriptors include *click*, *buzz*, and *hiss*, which clearly operate on principles of both iconicity and sound symbolism.

(3) "Pure" metaphor, in which timbral features are invoked not via sonic iconicity but with reference, generally, to other sensorial domains. Examples include *wet*, *dry*, *deep*, *bright*, *round*, each a synaesthetic metaphor. The majority of pure metaphors used to describe musical timbre are synaesthetic to touch and sight, the latter set almost always invoking spatial relationships. Many of the pure metaphors are highly codified, especially among sound engineers who recognize, for example, that a sound described as "too boxy" (a negative timbral trait) can regularly be corrected by dampening all frequencies between 250 and 500 Hz.

(4) Association, a strategy involving the citation of other musicians, recordings, sounds, technologies, time periods, etc. (see Feld 1994a). A typical use of association can be seen in this conversation between a producer (DE) and a drum tuner (JM), for example, recorded at the very beginning of a studio session near Austin, Texas:

69 DE: I like . . . I . . . you see, snare-wise . . . actually all the way around drum rise..drum wise..I like the sounds that were probably created but not captured in the '60s . . . a lot of the like..or even early '70s..a lot of the like Earth, Wind & Fire and uh . . .

70 JM: Tony Williams type stuff? . . . do you know . . .

71 DE: I don't know Tony..no.

72 JM: Oh, the drummer yeah..that sort of . . . what..Zeppelin sound? Real . . . flat? . . . like that?

73 DE: Nn:::.....

74 JM: Yeah cause I mean..John Bonham was a good drummer but . . .

75 DE: Yeah . . . ((skeptically))

76 JM: Personally I don't like his drums..but that has nothing to do with this job so . . .

77 DE: Yeah . . . ((thinking))

78 JM: But I mean you're talking about early '70s type sound? is what you're partial to?

79 DE: Yeah, some of those . . .

80 JM: . . . and that's what you want?

81 DE: No.. not so much the rock stuff but the funkier stuff.

(Porcello 1996)

Rather than attempt to describe the sounds directly, associations index frames of reference external to the session itself and in so doing strongly link musical timbre to linguistically mediated forms of sociability, as association succeeds or fails on the basis of shared cultural knowledge.

(5) Evaluation, in which rather than attempt a description of a musical sound, an assessment of its merit is offered (see, e.g., line 74 in the above transcription, assessing John Bonham's drum sound). Evaluation plays a key role in the microsocial relations of studio practice; it is often used to determine the boundaries of shared stylistic solidarity among session participants, to establish sociable working relationships, and to mark territory of shared or divergent aesthetics.

In extended discussions of specific musical sounds, interlocutors will often utilize a combination of these strategies. In a single turn at talk built around a pure metaphor, for instance, it is not unusual for the speaker to clarify by referencing a particular recording that embodies the timbral feature evoked by the metaphor (e.g., "a clipped snare, you know, like Phil Collins"). Similarly, across turns at talk in a dyadic exchange Speaker B may respond to Speaker A's use of a pure metaphor by supplying an association for purposes of checking Speaker A's intended meaning (e.g., Speaker A: "A clipped snare, you know?" Speaker B: "Like Phil Collins, you mean?") or to forcefully reorient Speaker A's approach (e.g., Speaker B: "You obviously have a sound in mind, but if you could say who it's by I'd have a better idea of what you mean"). Analysis of the moves among these strategies suggests that a great deal of the ability to communicate meaningfully about timbre lies in the metadiscursive work done by interlocutors.

2.2 Onomatopoeic form and semantics

Analysis of a corpus of the one hundred most frequently occurring vocabalic and lexical onomatopoeic metaphor timbral descriptors from recording sessions I taped suggests the utility of a modified version of Rhodes' typology of linguistic sound images (1994). Rhodes' tripartite continuum of linguistic strategies for mapping sound shape includes non-lexical forms (vocables such as /dz:::/ from Strategy (1) above), onomatopoeic forms – phonetic imitations that are nonetheless lexemes (Strategy (2) above) – and fully arbitrary forms (Strategy (3) above). Among onomatopoeic items, Rhodes suggests a further, two-part distinction between "wild" and "tame" forms. "Wild" forms are closer to simple phonetic imitations, while their "tame" counterparts strongly adhere to the phonological rules of the language in which they are uttered. For items applied to musical sounds in this corpus there is one additional significant point on this continuum linking imitation to arbitrariness: "motivated" arbitrary forms. Motivated arbitrary forms mix partial arbitrariness with partial onomatopoeia (e.g., lexemes *thin*, *tinny*, *sibilant*, in which the initial consonant–vowel cluster functions onomatopoeically but the balance of the lexeme resists iconicity). These are, in ways, reminiscent of Bolinger's characterization of

phonesthemic (submorphemic sounds or sound clusters with regularly associated meanings) rimes and assonances (1950). As speakers move from vocables to fully arbitrary forms, the tendency to perform the utterance disappears. (See figure 14.1.)

Strategy (2) – use of lexical onomatopoeic metaphors – is of particular interest because the majority of these items, like their pure metaphorical counterparts, are highly codified in studio discourse, deployed and interpreted by sound engineers not as impressionistic but as a technically precise metalanguage that glosses specific physical properties of sound waves. Part of the technical vocabulary of sound engineers and studio musicians, they are a primary means by which musical, engineering, and sonic competence are negotiated in recording sessions (see Bartlett and Bartlett 1994). As such, they powerfully link iconicity to reference as well as to the social roles of those who work in recording studios. Here the limitations of a linguistic ideology in which iconicity is rarely examined in contexts other than poetics or child language development (an ideology critiqued by Nuckolls 1999 and Wescott 1977) are evident; iconicity serves as a crucial, systematic professional linguistic resource in the world of audio production (see also Feld 1981 on the systematicity of metaphors in discourse about music; see Jakobson and Waugh 1979: ch. 4 on the systematicity of sound symbolism in general).

The vast majority of onomatopoeic lexemes will be deployed in adjectival form (e.g., in a syntactic frame such as “that sound is too _____”) but derive from a monosyllabic nominal or (more rarely) verbal form. Common examples are *boom*, *hiss*, *grunge*, *click*, *thump*, *honk*, which become *boomy*, *hissy*, *grungy*, *clicky*, *thumpy*, *honky*, etc. In these cases, the noun or verb is treated as the basic analytical unit. The corpus includes a limited number of disyllabic noun or verb forms that remain disyllabic in adjectival form; in each case the second syllable of the nominal form is built around an unstressed syllabic [l]. Thus *rumble*, *crackle*, *muffle* become *rumbly*, *crackly*, *muffled*. The preponderance of monosyllabic onomatopoeic forms in this corpus suggests that 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.

Non-lexical forms	Onomatopoeic forms	Motivated arbitrariness	Arbitrary words
 not “words” often sung/ spoken 	<i>Wild</i> often performed, may violate phonological rules	<i>Tame</i> not performed, whole word adheres to phonological rules	 one + individual segment(s) partakes of onomatopoeia
 [m:::] [dz:::] [s:::] etc.	 [ʔwæk + nasal vowel] [hɪs::]	 [kwæk] [hɪs]	 thin (θ and ɪ) tinny (t and ɪ) sibilant (s) etc.
			 <i>deep</i> <i>bright</i> round etc.

Figure 14.1 Continuum of linguistic descriptors of musical sound (modified version of Rhodes 1994)

If one accepts that onomatopoeia involves a fairly direct mapping between the timbral features of the musical sound and the phonetic/phonemic features of the linguistic item that replicates it, then the principles of phonesthemic patterning, particularly within the monosyllabic items, may be summarized as in figure 14.2.³

No onomatopoeic items in the corpus are vowel-initial or vowel-final, suggesting that the details of the acoustic properties comprising the onset and decay of musical sounds are just as crucial to speakers as those of the sustain in describing and evaluating musical timbre.

Aside from phonological concerns, certain semantic features appear particularly salient to engineers and musicians working in a Western recording studio context that highly values sonic purity, cleanliness, and the transparent accuracy of recorded sounds. Here a large number of pure metaphors signify undesirable diffuseness of sound. Some are generally applied solely to individual instruments or voices (e.g.,

Principles of phonesthemic patterning

- General structure is CVC (all begin and end with consonants as pronounced)
- Initial and final sounds may be individual consonant or consonant clusters (*Cr-*, *Cl-*, *spl-*, *-nɛ*, etc.)

A. Syllable-initial sounds (onset) → attack

1. Abrupt/explosive
 - stop only /p, t, k, b, d, g/ (*puff, boom, tick*, etc.)
 - stop + liquid /bl, kl, kr/ (*blat, click, crisp*, etc.)
2. More gradual
 - fricatives /h, s, f/ (*hiss, sizzle, fuzz*, etc.)
 - affricates /tʃ/ (*chunk*, etc.)

B. Syllable middle (medial vowels) → sustain, pitch

1. High /ɪ, i/ (*click, crisp, jingle*, etc.)
2. Neutral /æ/ (*clack, smack*, etc.)
3. Low /ʌ, u/ (*clunk, thump, boom*, etc.)

C. Final sounds → decay

1. Abrupt
 - a. stops (*snap, crack, smack*, etc.)
2. Gradual
 - a. non-resonant
white noise
– fricatives, affricates (*crash, splash, smash, hiss*, etc.)
 - b. resonant
fading to infinity
– nasals /m, ŋ/ (*boom, ring, hum*, etc.)
brief fading until a moment of truncation
– nasal + stop (*clunk, thump*, etc.)

D. Disyllabic words ending with unstressed syllable with /l/ as nucleus

1. Sustained decay with a point of noticeable fall-off
 - a. resonant (*rumble*, etc.)
 - b. non-resonant (*crackle*, etc.)

Figure 14.2 Structure of onomatopoeic forms

breathy) but the majority can be used to refer to the mix as a whole (e.g., *blurred*, *smearred*, *distant*, *puffy*, *muddy*, *muffled*). A similarly large set of terms concerns the extent to which the performed sound (as heard by a co-present ear) has been modified in the recording and reproduction process, a semantic domain perhaps best labeled "accuracy." The terms *blanketed*, *boxy*, *brittle*, and *thin*, for example, imply that in the recording process something has induced a modification to an original sound as emitted from its source.⁴ The large number of terms for accuracy and diffuseness suggest phenomena that are of overriding concern in studio production and which must therefore have a well-developed vocabulary to pinpoint particular sonic features in need of modification or correction.⁵

Musical timbre arguably provides the most important structure of hearing in the process of music-making in sound recording studios. It is therefore incumbent on session participants to develop elaborated, yet precise, discursive means of communicating about a phenomenon notoriously resistant to essentializing statements. This takes place through identifiable and significant principles that structure ways of speaking about timbre; these include the phonetic and phonological structure of onomatopoeic forms, or the semantics of metaphors. This focus on vocality, on the sonic material of articulation, constitutes a critical bridge between the musicality of language and verbal discourse about sound. It further connects to a perspective on singing and song that links music in language to language in music.

3 WORKING-CLASS "COUNTRY"

In South-Central Texas, the art of singing "country" music is highly valued and carefully cultivated, as are critical and aesthetic discourses about singing (Fox 1995). Country singers are musical specialists, responsible – like all folk artists – to their local communities for a wide range of performance skills (Bauman 1977); but they are especially respected for their mastery over (and creative extensions of) a canonical catalogue of vocal techniques. This case study focuses on several of those techniques, in order to consider the larger significance of singing as a cultural practice that presents ramifications for a general theory of music and language.

Many articulatory possibilities are available to working-class country singers. Fully pitched, metrically regularized, and amplitudinally shaped singing (with an expansive and exploitable range of voice qualities and vibratos and phonetic indexes of different modes, registers, genres, and dialects of "ordinary" speech) is of course a predominant "default" modality. Singers also produce heightened speech, which they call "recitation," and which mimics the less rigid meter of speech while retaining a song's original versification (similar to some types of operatic "recitative" and other quasi-spoken idioms in many musical drama traditions).

Country singing style also involves frequent importation of fragments of metrically non-regular and non-pitched speech into song performance (e.g., bits of dialogue between a singer and a "picker" in the band that are performatively spoken against the background of a song). Of course, such "fully spoken" discourse also shapes the boundaries around song performances. Stage patter, bandstand talk, and interactions with audience members (using intonationally and dynamically heightened speech) can also be used to move between the spoken frame of song performance and full

singing. Specific transitional gestures between speech and song include "count-offs," which establish the meter of the song, introductory words summoning full attention from audiences (for example, many singers begin songs with the formulaic expression "I said . . ."), and stylized formulae for marking reported speech, including elaborate techniques of voice imitation, reflecting a strong preference among Southern working-class Americans for direct discourse over indirect discourse.

Competent singers are often quite consciously aware of this range of articulations. They may refer to a corresponding set of ethnotheoretical concepts to describe specific formulae, articulations, and voice qualities. More commonly, both among singers and among listeners, metamusical discourse proceeds by analogy (usually comparing one singer's style to another well-known singer's example), or through a metaphorical vocabulary that emphasizes qualities of "hardness," "sweetness," "sadness," "volume," "power," "precision," "ordinariness," and appropriateness of singing style to textual content and sociomusical context.

All such evaluations are oriented toward the master-tropes of musical evaluation in South Texas, the sociomusical categories of "feeling" (as noun and verb) and "relating" (as a mode of telling and a mode of social and aesthetic engagement). Singers who perform "with feeling" are said to "relate to" their material, their traditions, and their communities. Likewise, audiences say they can "relate to" singers who sing "with feeling" (or "feel"). In turn, such evaluative tropes are further summarized through strongly inflected tropes of genre identification, typically in phrases like "that's *country*," or (most assertively), "that's *real country*." Such tropes resonate with a matrix of cultural sensibilities and practices for which musical style is a summarizing symbol. Evaluations of song performance thus extend pervasively into assertions of rural, working-class, Southern social identity and cultural continuity.

Vocal style in Texas country singing can be approached analytically through a fine-grained description of particular techniques employed by singers. Many properties of vocal sound (e.g., vibrato, amplitude, articulatory noise, melodic shape, etc.) can be measured and correlated with the communicative and expressive features of song texts (e.g., explicit affective verbs, specific references to feeling states, canonical moods of particular poetic (sub)genres). Similar correlations, considered quite important by most Texas country singers, can be made between vocal techniques and the language-structural properties of song texts (many vocal inflections, such as "cry breaks," appear to be conditioned equally by phonological environments and by affective connotations of the referential text).

A continuum of intonational markedness, ranging from unpitched, metrically irregular, but intonationally heightened speech to full pitched, metrically regular singing, is routinely employed to mark structural divisions in song texts, changes of point of view and narrative voice, contrasting affects, and degrees of expressive engagement by the singer. Timbral quality deserves special attention; Texas singers evoke important affects and moods through distinctive changes in the site and manner of voice production. A pharyngealized tone, for example, can be iconic of the ravaged voice of a character textually narrated explicitly or implicitly as "crying." "Crying" itself can be iconically represented with specific inflections known categorically as "cry breaks" – sharp deformations of the melodic line effected through intermittent falsetto or nasalization, pulsing articulations achieved through glottal

stops or diaphragmatic tensing, or the addition of articulatory vocal “noise” to an otherwise “smooth” tone. An enormous range of rhythmic and metric possibilities are suggested by the intertwining of musical and linguistic form in country singing as well, ranging from a relatively “natural” delivery that mimics the variable meters of ordinary speech, to the rigidly metrical delivery characteristic of other locally familiar oral genres such as auctioneering.

Working-class singers (and listeners) themselves often attend to these various dimensions of singing style as analytically and technically distinct. But of course the art of singing well involves a less consciously considered blending of these techniques into expressive syntheses that can be drawn upon for specific aesthetic purposes. Most typically, musicians and listeners use a shorthand terminology for such *gestalts* that refers to the names of canonical singers like George Jones, Merle Haggard, Patsy Cline, Johnny Cash, and Marty Robbins.

3.1 “He Stopped Loving Her Today”

Since much of the repertoire performed by local singers in Texas comprises “covers” of the classic country canon, the default local practice is to apply (or imitate) the style of the original recording artist associated with each particular song. This is especially true for less accomplished singers, and for those still learning their craft. As Texas singers become more competent and develop individual styles, and as they acquire a repertoire of original songs or covers by obscure artists whose styles are not familiar, they increasingly apply their own distinctive stylistic signature to everything they perform. Eventually they may be said to have “made (even a well-known ‘cover’) song their own” by fully restyling a canonical song, sometimes with dramatic modifications in affect and meaning from the canonical recording.

Typically, however, such singers retain a special affinity for one or two major stars’ styles (and repertoire), and these styles can be instantly recognized as “influences” on their personal styles. Thus, a singer might be classified as singing “in a Marty Robbins (or Merle Haggard, etc.) style.” This would not necessarily imply that the singer in question was simply competent in a derivative sense. A rich cultural conception of voice and orality, reminiscent of Bakhtin’s theory of voicing in literary discourse (1981), is the basis for such tropes of classification, resonating with the pervasive elaboration in working-class Texas discourse of direct and quasi-direct discourse and extensive voice-imitation in reported speech constructions.

Movements within this range of possible articulations are ubiquitous expressive resources. A careful listening to George Jones’s 1980 recording “He Stopped Loving Her Today,”⁶ for example, reveals why Jones is widely regarded as the greatest vocalist in the history of country music, especially by his fellow Texans, and why this particular performance is almost universally acclaimed by working-class country fans as among the most important in the history of the genre. What follows is a description of Jones’s complex vocal stylization of this song, from beginning to end. Interested readers should have little difficulty locating an audio recording of this legendary song (first released on Jones’s 1980 Epic recording “I Am What I Am,” but subsequently reissued on many compilations and greatest hits recordings) to supplement the verbal descriptions below.

Jones begins the narration, which is introduced as reported speech (“*He said* ‘I’ll love you ’til I die’ . . .”) by singing lightly, at times coming close to a spoken articulation, glancing off each word with a breathy tone and very delicate vibrato. Subtle changes in pitch quality and metric feel mark the boundaries between verbs of speaking and directly reported discourse. A subtly pharyngealized tone, in which one can hear the ravaging effects of crying on the narrator’s voice, conveys the ethos of elegiac sadness that dominates this song. The stressed vowel in the word “slowly” is mimetically elongated as Jones’s narrator describes the passage of years during which the protagonist cannot forget his beloved.

Over the course of the song’s unusually long sequence of four verses prior to the first refrain, Jones uses modulations in voice quality to intensify the abject poetic scene. These become ever more claustrophobic as the narrator gradually reveals the depth of the male protagonist’s obsession with a lost love, as he lives out his “half crazy” life surrounded by objects that serve as shrines to the memory of his beloved (her picture on his wall; her letters, which he keeps by his bed with every “I love you” underlined in red). The vocal line gradually acquires more intensity, with increasingly sustained and amplitudinally shaped notes (i.e., the amplitudinal envelopes of sustained vowels become more elaborate). Jones gradually develops a richer timbre and a broader vibrato as these verses progress. The song modulates up a half-step at the third verse, after Jones reaches for a high note on the desperate line in which we learn the protagonist never stopped “hoping she’d come back again.” Jones also gradually increases the use of his trademark melismas, and the entire sequence of verses is characterized by steadily increasing amplitude. These effects combine to create extreme tension and anticipation of a remarkable textual and performative *denouement*.

Finally, in the fourth verse we find out that the narrator has gone to see the protagonist as he is “all dressed up to go away,” and finally smiling (marked by the most melismatic articulation yet in the performance) for the first time “in years.” At this point, the anticipation of an explosive release of tension is literally unbearable for many listeners.

The long-anticipated refrain, from which the song takes its title, arrives in full-throated, elaborately melismatic, sustained, and vibrato-rich tones as Jones at last reveals the reason why the protagonist “has stopped loving her today”: he has died. He is dressed up and smiling because he is at his own funeral. At last, the somber, elegiac tone of the previous melodramatic verses is narratively justified. However, the song contains an additional poetic relaxation and release, in the following verse and refrain.

Withdrawing, after the refrain, to observe the arrival of the protagonist’s long-lost beloved at the funeral, Jones suddenly reverts, with an eerie effect, to what country singers call “recitation” – a loosely metric spoken articulation (breathy, with elongated vowels and heightened intonational movement) which preserves the poetic structure of the composition (rhyme, line breaks, and scansion). This sudden distancing of the narrator from the poetic *mise en scène* demands, as it were, a third vocal persona to emerge from Jones’s performance: that of the wry observer, personally uninvolved in the immediate emotional drama of the song. Behind this recitation, in which Jones describes the funeral scene, a wordless female voice laments in an operatic melodic descant, layering the most extreme form of singing (fully pitched, wordless, rich vibrato) behind Jones’s move into a quasi-spoken recitation, a

juxtaposition of non-referential pure song and song straining toward referential speech that is almost didactic in its representation of the speech/song continuum as it is understood in this musical tradition.

This recitation moves, in a manner typical of the genre, from a description of the scene to the first reported inner speech of the narrator himself (who has otherwise been merely a relatively dispassionate reporter of events). This thought is couched as wordplay – indeed, as reported inner speech – and it delivers the final refrain on the heels of a vertiginous joke that plays on a trite romantic cliché: “This time, he’s over her for good.” But just as this comic moment seems to break the morbid spell of the tale, the refrain returns, one last time, more fully and powerfully sung than previously, in an all-out quasi-operatic apotheosis in which Jones sweeps grandly through his entire vocal range until relaxing on the final line.

Such masterful examples of vocal style lie at the center of an extremely dense network of vocal practices and expressive ideologies that constitute the discursive and experiential world of working-class Texas communities extending from the most casual forms of talk through a rich range of verbal art genres to song. This network is focused and attended to explicitly in singing practice, but it extends well beyond the boundaries of song *per se*. An analysis of the techniques of Texas country singers reveals the dense relationships between song and speech in this culture. When the focus is expanded to incorporate verbal art, ordinary talk, and the soundscapes of everyday working-class life, one sees, or rather hears, a vivid acoustic refraction of a particular form of sociality materialized in every act of vocalization, and every act of listening. Singing is especially privileged in this community because it allows for a ritualized, explicit consideration (both by community members and by analysts) of the voice as the material embodiment of social ideology and experience. Song stands in an explicitly critical and denaturalizing relationship to “ordinary” speech in rural Texas, as in many other societies.

The parameters and microstructural details that emerge from an analysis of Texas singers’ technical practice represent only a small portion of the “voice consciousness” of working-class Texans. And this consciousness, in turn, represents only one perspective, either local or ethnographic, on the social life of these communities. But it is a locally privileged perspective, one that is highly cultivated and deeply valued by community members. This suggests that theorists of the speech/song relationship should engage more fully with local understandings of vocal practice across a wide range of expressive genres. It suggests also that close attention to vocal practice and to the speech/song relationship can provide vital insights into the life of language in human society.

4 VOICING APACHE COUNTRY

Like the rural Texans just discussed, many Apaches living on the San Carlos reservation in southeastern Arizona also love country music.⁷ There are songs, such as Johnny Horton’s “North to Alaska,” that put people in mind of their community’s history. There are others, like Merle Haggard’s “Silver Wings,” that everyone seems to know by heart. Others, like Hal Ketchum’s “Past the Point of Rescue,” pull people of all ages out onto the dance floor. And still others, like Charlie Daniels’s “Long

Haired Country Boy,” or George Jones’s “One Woman Man,” are so closely associated with particular singers in the community as to act almost like fingerprints.

Local bands in San Carlos, Peridot, and Bylas (the main communities of the San Carlos reservation) have played country music since at least the 1950s. Local songwriters have added to the local repertoire, composing country and rockabilly songs – some with English lyrics, some Western Apache – that sing about the places, histories, and experiences that are important to people in the reservation’s communities.

There is no doubt that country is a powerfully evocative genre of song on the reservation (Samuels 1998). Yet, when San Carlos Apaches sing country, they do not sing it the way commercial country singers do on discs, cassette tapes, and the radio. In fact, country singers on the reservation assiduously avoid some of the key vocal gestures that are most diagnostic of country singing, reserving them only for joking performances of white singers. This joking linguistic practice is reminiscent of Keith Basso’s analysis of joking imitations of “the Whiteman” in the Western Apache community of Cibecue, which is about 100 miles north of San Carlos, on the Fort Apache reservation (Basso 1979). Of what significance are these differences in the manner of sung vocalization between mainstream commercial country singers and local San Carlos performances of country songs?

4.1 Phonation

The body acts as a resonating chamber in the performance of both speech and song. As with any musical instrument, the acoustical qualities of sung or spoken utterances – their tone and timbre – are partly determined by the physical shape and resonance of the cavity through which air passes during vocalization. The subtle and naturalized control of lungs, diaphragm, larynx, pharynx, tongue, sinuses, lips, and teeth, in the production of sung or spoken vocalization, is the end result of conscious or unconscious discipline and socialization. The apparatus of phonation, especially the mouth and the vocal tract, are crucial bodily sites of hegemonic contestation over the indexical and iconic modalities of both language and music (speech and song).

The study of those variations in vocal practices is one of the key areas in which the interests of sociolinguists overlap strongly with those of ethnomusicologists. That study includes the means by which varieties of spoken and sung phonation are produced, the indexical and iconic links associated with them, and the cultural ideologies surrounding the acceptable or unacceptable performance of these vocalizations. On the one hand, the core insights of sociolinguists demonstrate the numerous and complex ways in which language varieties and dialects act as indexical signs of sociopolitical position. That is, sonic features of vocal speech production, such as deletion of postvocalic /r/ in New York City (Labov 1966), deployment of monophthongal /ai/ in the southern United States (Bernstein, Nunnally, and Sabino 1997), and other articulatory phenomena such as nasality (McMillan 1939), rising intonation at the ends of declarative sentences (McConnell-Ginet 1983), or fronting of back vowels (Hinton et al. 1987; Luthin 1987), act as indexes of regional, class, ethnic, gender, or generational group membership or identity.

On the other hand, ethnomusicology has often posited an indexical relationship between singing style and social order. This has included such statistically driven work

as Alan Lomax's Cantometrics project (Lomax 1968), but also includes more subtle and open-ended interpretive research, such as Feld's work on Kaluli "lift-up-over sounding" (1994b), Nattiez's explorations of Inuit and Siberian throat-singing (Nattiez 1983, 1999), and Hugo Zemp's comprehensive recorded anthology, *Voices of the World* (1996).

Both of these insights – the sociolinguistic and the ethnomusicological – claim the voice as the central locus in the production of social and cultural being. They are both concerned with the shape of the mouth as well as the placement of sounds within the vocal tract during processes of phonation, and both interpret these differences as enunciations of ideology, or as markers of social and political identity.

Unlike most musical instruments, the size and shape of the vibrating vocal tract in speech and song is not fixed. As the linguist Peter Ladefoged has written, "Different vowels are like different instruments" (2000: 118). Trumpets, guitars, clarinets, and cellos have distinctive timbral characteristics based partly on the size and shape of the chamber in which sound-making air vibrates to produce a note. But the size and shape of the resonant chamber used in speaking and singing – the mouth – is constantly being changed in the articulation of different sounds. The difference between pronouncing the words "hoot" and "heat," for example, involves a shift in the position of the tongue. This positional change in turn changes the shape of the resonant space of the mouth, which causes different clusters of overtones – called "formants" – to be emphasized for different vowel sounds.

The shape of the mouth when speaking is reflected in the diagram of the "vowel triangle" (see figure 14.3). As the tongue moves from being raised in the front to being raised in the back, the vowel sound produced changes from the high-front /i/, as in "heat," to the high-back /u/, as in "hoot." As the jaw is opened and the tongue lowered, the vowel sound produced shifts from the schwa /ə/ to the low-mid /a/. These sonic differences carry social differences, as well, and also culturally meaningful senses of contextual propriety and proper style. The shape of the mouth in singing also implicates numerous cultural values of proper style. The open and rounded mouth of the boy soprano in a church choir, for example, is an almost immediately identifiable visual marker of this vocal quality and this style of singing. These cultural values about proper singing style circulate and are perpetuated through the training and socialization of singers.

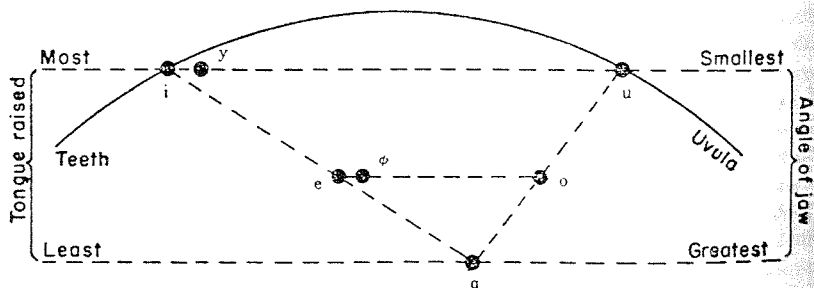


Figure 14.3 Viëtor's vowel triangle, showing the relationship of tongue position and jaw angle to vowel quality (from Heffner 1950: 84). © 1950. Reprinted with permission of the University of Wisconsin Press.

In Western classical forms, for instance, the buttery sound of the operatic soprano is the result of years of training and discipline. Operatic training produces singers who are able to eliminate the upper overtones from their sung performances – like turning the treble on your stereo all the way down. Figure 14.4 is a spectrographic depiction of operatic sopranos Renata Tebaldi and Mirella Freni. A spectrogram is a visual representation of the concentrations of acoustic energy in a spoken or sung utterance, from the lowest to the highest harmonic resonances. Operatic sopranos are able to configure their vocal tract in such a way as to cut most or all acoustic concentrations above around 3,000 cycles per second (Hz). By comparison, figure 14.5 is a spectrographic depiction of legendary Broadway belter Ethel Merman (singing "Bye bye baby, stop your yawnin'/So long baby, day will be dawnin'"). Quite obviously, in distinction to opera singing, Broadway show singing reveals concentrations of resonance well above 10,000 Hz.

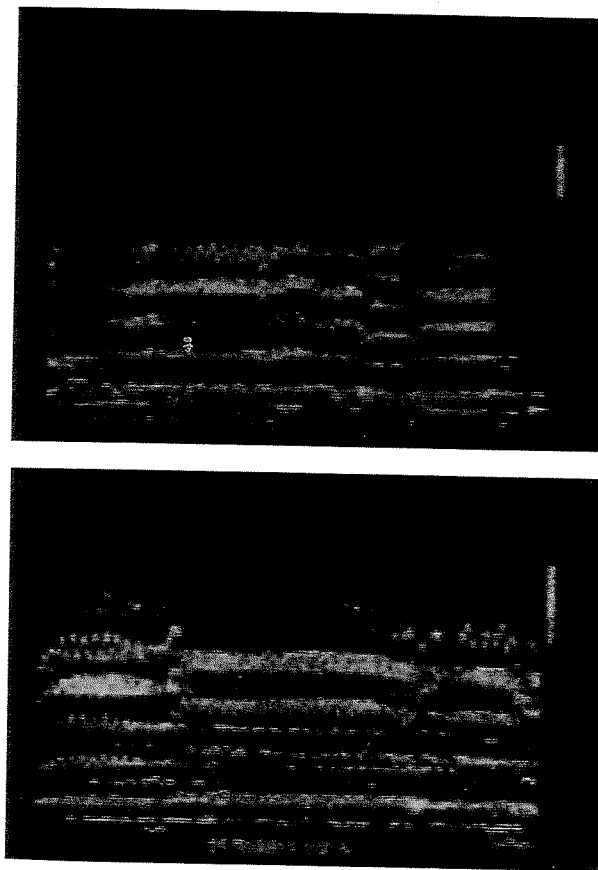


Figure 14.4 Spectrograms of Metropolitan Opera stars Renata Tebaldi and Mirella Freni, showing the extreme tapering of upper partials in the vocal timbre of operatic sopranos (from Miller 2001: 73). Reprinted by permission of the author.

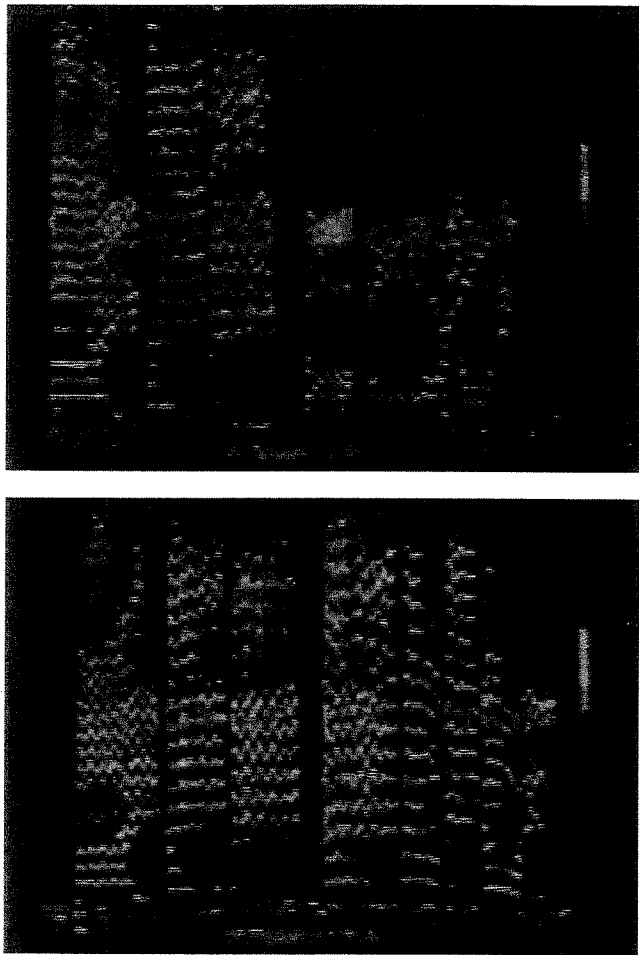


Figure 14.5 Spectrogram of Broadway star Ethel Merman, showing the concentrations of acoustic resonance, to more than 10,000 Hz, that marked the timbre of this showtune belter (from Miller 2001: 74). Reprinted by permission of the author.

4.2 Country vocalicity

Commercial country singing incorporates a number of distinctive phonological features of Appalachian American English that deviate from a received “standard” dialect (Wolfram and Christian 1976). Among the most important of these, for the present discussion, are a cluster of vowel-glide phenomena. These include long off-glides in diphthongs, found in the “Southern Drawl” (Bailey and Tillery 1996), as well as various forms of diphthongization. “Diphthongization” refers to a process by

which diphthongs, or vowel glides, are inserted into phonological segments which, in “standard” dialects, are “pure,” or monophthongal, vowels (/hɛɪd/ for /hɛd/, /hɛɪf/ for /hæf/, etc.). In country singing, these vowel glides often take the form of what one could call “twang” – the insertion of a tensing of the vowel followed by an offglide to the lax form (/bɛɪəd/ for /bɛd/, /mɛɪən/ for /mæn/, etc.).⁸ Country-western singing includes a network of phonological features, shared by numerous performers, all of which orbit around a collection of Appalachian American English dialectal pronunciations. These dialect features are in turn centered around questions of vowel location, and more specifically around vowel glides, diphthongs, and diphthongization as embodiments of what Jack Temple Kirby (1995) calls the “counter-cultural values” of southern redneck identity.⁹

The play of vowel quality and dialect in country singing creates opportunities for songwriters, arrangers, producers, and performers to use these Appalachian dialectal features for poetic and artistic effect. Singers often emphasize the offglides of diphthongs, especially when those diphthongs occur in line-final position. In the Lorrie Morgan song “Watch Me,” for example, Morgan consistently stresses the offglides of the line-final words “today,” “stay,” and “away,” creating a great deal of stylistic energy for the textual hook of the song. The songwriter, Gary Burr, conveniently distributes the two parts of the diphthong over a two-note melodic sequence.

Popular singers such as Faith Hill, Andy Griggs, and Reba McEntire consistently employ twanging to index their “country”-ness. Hill, in her hit song “This Kiss,” gently modifies the monophthongal vowel /ɪ/ in the word “kiss” in the chorus, into a diphthong /iə/. Griggs, in another example, diphthongizes the second syllable of the word “lonely” in the title song of his debut CD, “You Won’t Ever Be Lonely,” from monophthongal /i/ to the diphthong /eɪ/. As for Reba McEntire, her stardom and her backwoods identity are due in no small part to the way her vowel segments float in her mouth. One of the hallmarks of McEntire’s style is the way she sings long notes, drawing out not only their musical, but also their vocalic potential. In the song “One Honest Heart,” for example, Reba performs full-on twangs on the line-final words “sound” and “found,” remolding the /au/ diphthongs of Standard English into /æiə/s.

Of course, a good deal of Reba’s ability to vocalize in this manner is related to her Oklahoma upbringing. But these vocal qualities are identified with a commercial genre of music, and not simply a geographic region. It’s not necessary to be from Oklahoma to have twang as part of your vocal style if you’re a country singer. The strength of those indexical associations with a musical style give this “regional” dialect a social life that extends well beyond its Appalachian/Ozark regional boundaries. This kind of appropriation and assimilation is not limited to country singing. As Peter Trudgill pointed out in a well-known discussion of British rock singers appropriating “American” accents (Trudgill 1983), the mediating technologies of modernity have made multiple styles of singing available – and appealing as a means of expressing identity – to singers around the world. For example, the Canadian country trio the Wilkinsons, Sweden’s Inger Nordstrom and the Rhinestone Band, and Raebekah Roycroft from Australia, all feature vocalization styles in which the singers have disciplined their voices to produce the kinds of vowel glide phenomena discussed here, even though these twanging forms are not part of their everyday spoken discourse.

4.3 Real Apaches don't diphthongize

Although singers on the San Carlos reservation love country music deeply, they never (in my experience) diphthongize or twang, except in joking imitations of white singers. Keep in mind the earlier discussion of the prominence of twang in country singing style. These Appalachian features of country singing are so prominent, and so identified with the musical style, that one of linguist Dennis Preston's respondents to his "folk dialectology" project circled the area around North Carolina, Kentucky, and Tennessee, and wrote "Country, As In Music" (Preston 1997). The refusal of San Carlos Apache country singers to discipline their voices in this way is therefore quite radical, given the love for country music exhibited by many in the reservation community.

The Pacers, the band I did the bulk of my fieldwork with in San Carlos, played a number of songs that one might call "signature tunes." The lead singer, Marshall, was known for his renditions of these particular songs, and whenever the band played, or whenever Marshall was present with a guitar, people would always request songs from this special group.

Among those songs was one by Steve Earle, called "Copperhead Road." The story of three generations of bootleggers, sung in the first person, the song is something of an anthem of hardcore, hardscrabble, backwoods, redneck identity. In Marshall's performance, the song became somewhat gentler, more like a two-step than an all-out rock song. But the fact that the song was an expression of marginal resistance, and not a love song, was extremely important to everyone in the band.¹⁰ During my fieldwork, the band members had asked me not to pay them an hourly fee for interviews and so forth. Instead, they wanted me to do something else at the end of my time in San Carlos: to buy studio time for the Pacers so that they could make a cassette tape that the band could sell. And, as the time approached, it was an easy decision to include "Copperhead Road" on the cassette.

Marshall took this all very seriously, and one of the tasks he assigned to me was to make sure that all the words in the songs the Pacers were going to record were right. In the process of transcribing song lyrics from recorded versions, Marshall included a number of "oronyms," or shifting of phonemes so that words change while some kind of sense remains.¹¹ In live performances this hardly mattered, because the chances were small that anybody would attend to the differences. But for the purposes of making a permanent record, Marshall contracted a bad case of what one might call "lexico-mania." Every word had to be right. As the resident native English-speaking scholar, I was enlisted to double-check Marshall's transcriptions of the lyrics for every song the band was about to record. And in undertaking this task for Marshall, I heard the original Steve Earle version of "Copperhead Road" for the first time. Until that point I had only heard Marshall's rendition of it.

The third verse of the song is about the singer's decision to shift the bootlegging business over from liquor to marijuana, and his decision to rig Copperhead Road with booby traps he had learned about from his experiences fighting the Viet Cong. In Steve Earle's recording of his song, one rhyme in that third verse depends on inserting a particularly strong twang into the first line of the couplet.

I done two tours of duty in Viet Nam (/næiəm/)
 I came home with a brand new plan (/plæn/)

When Marshall sang this same couplet, he pronounced the end of the first line /naim/, sacrificing the rhyme – and also, in a sense, the cultural identification of the character. The words were right, but the pronunciation was wrong. So when the Pacers went into the studio, thinking that Marshall had wanted everything to be "right," I tried to convince him to sing *Viet Nayum* instead of *Viet Nahm*. In the final take, Marshall was willing to raise his /a/ to an /æ/, in order to make the rhyme scan, but he would go no further. He never explicitly objected, but clearly he was reticent to make his voice twang. On the drive back to San Carlos from the studio in Tucson, Marshall and Pat, the lead guitarist in the band, teased me mercilessly about how I had really made Marshall "sound like a redneck." I could never get Marshall or Pat to talk about this in any detail. They knew what I was talking about when I asked, but the only thing Marshall would say was, "well, you *know* we don't get along with them." And Pat corrected him, saying, "I think it's *them* that doesn't get along with *us*." But that was all.¹²

People in San Carlos love country music. There is no secret to this, and no magical explanation. Music has always been a key institutional expression of assimilation projects on the reservation, from Christian hymns to school bands to Elvis Presley movies. People in San Carlos have been playing guitars at least since the first Sears catalogues circulated in the community, and have been listening to "Hillbilly" and country music since the first radios arrived. And, to be sure, people in this hardscrabble community have much in common, in terms of the history of economic and social marginalization, with the rednecks in the next town over who also love country music.

But here we can make a distinction between *common* history and *shared* history. The parallax trajectories that have brought "rednecks" and "Indians" to occupy overlapping positions in the American social landscape fill that overlapping space with irony and ambiguity. Country music's sonic and textual evocations of loss, of place, of memories that refuse to recede into the past, and of broken hearts are strong markers of its connections to marginalized, white, redneck social and political existence. That these markers of rural, working-class whiteness can be made iconic with a feelingful experience of Apache social history makes country music an ironic and ambiguous aesthetic form on the San Carlos reservation. But San Carlos is an ironic and ambiguous place. And the historical and cultural ironies of social being in San Carlos saturate vocal expression down to the level of phonetic and phonological production.

In their refusal to twang, people in San Carlos insist that their love of country music grows out of an Apache, and not a "white trash," history of exploitation and marginalization. Through this local style of country singing, San Carlos Apaches perform identities that attempt to maintain an existence outside the American class system at the same time that they are being inexorably pulled into the American socioeconomic system in a class position. One possible question that could frame this vocal practice might be to ask, "What do you do if you love country but you hate rednecks? If you love the music, but you cæin't stæind the sæiund?" And one possible answer might be to sing country music as Marshall and so many others in San Carlos

do – in what Mikhail Bakhtin called a “polyphonic” style (Bakhtin 1981). Polyphony in this sense is not musical polyphony, or multi-voiced harmony. Rather, it is a means by which individual speakers, or singers, are able to layer varying attitudes within a single utterance. Through the manipulation of vocal qualities speakers can layer their speech with, for example, sarcasm or irony, making the sense of their words richer and more complex. Through such practices, Apache country singers are able to layer their attitudes about “rednecks” through the way they sing their music.

5 CONCLUSION

Important general works like David Burrows’ *Sound, Speech and Music* (1990) or Theo Van Leeuwen’s *Speech, Music, Sound* (1999) discuss the historical differences between language and music as sounding modalities, and the semiotic importance of approaching them in a unified framework of sound. The kind of vocal anthropology elaborated here takes the argument to the next level, working at the precise ways that music and language are phenomenally intertwined and socially dialogic. Music and language are fundamentally interrelated domains of expressive culture and human behavior and experience. The case studies presented here indicate some of the richness of verbal discourses about musical timbre and musical meaning, and of voice as the embodied site of both musical and linguistic expressivity, and of social distinction. These ethnographic studies reveal the micropolitics of emplaced, embodied, and voiced identity in particular local lifeworlds. They signal renewed attention to understanding how social identities are indexed and expressed in the intertwining of musical and verbal practices. Above all they demonstrate the post-structuralist commitment to linking the materiality of sound to the sociality of vocal practice.

Thomas Porcello’s case study of talk about timbre in the recording studio demonstrates why the musicality of language, the place where sound symbolism and iconicity meets lexical semantics, is critical to understanding the pragmatic role of sound knowledge in language use. We see in this research a detailed application of one of Roman Jakobson’s claims about linguistic iconicity (Jakobson and Waugh 1979). Namely, at the level of the lexicon, verbal iconicity is always both partial and systematic in the ways sound might relate to sense. “The vile vomit of the vicious vituperative villain” suggests one systematic relationship between the /v/ sound and a semantic field. But when that phrase is juxtaposed with another, “the voluptuous viscosity of the vivacious vamp,” the very partial nature of that suggested relationship is revealed.

When it comes to sound and sense at the lexical and morphological levels, Porcello’s analysis underscores Jakobson’s profound point, namely, that the whole of language is a field of potentially consummated iconicities, only a small segment of which are ever consummated. But when and where they are consummated they bring sensuous immediacy to the experience of sonic materiality. And that is what matters at the level of their practical and affecting use in discourse, particularly evident in the examples of morpho-iconic coherence, linking the audio envelope of attack/sustain/decay, and a pattern of [c + v] in a distinctive feature package.

Metalinguages of timbre, and talk about timbre in studio shop talk, indicate that timbre’s life is as much social as it is acoustic. In other words, it is public and circulates as both general and specialized components of a technical and poetic lexicon. Most critically, this lexicon indexes how knowledge and its practical use is central to the ordinary work that makes music happen – tuning, playing, recording, mixing, listening, commenting, evaluating. All of these dimensions of “getting the sound” can be understood as multiply dialogic relations among musicians and engineers, embedded in a distinct industrial habitus.

The perspectives presented in Fox’s case study of rural working-class Texas song and Samuels’ case study of Apache country also bring social specificity to questions of timbre and voice, dialogue and practice. Fox’s research makes clear that the heightened presence of the singing voice produces a site where speech and song intertwine to produce timbral socialities. The ones he particularly reveals are those of recognized and recognizable agency – artful mastery and its intervocal location in the history of song styles. David Samuels likewise takes up timbre at a crucial bodily site of the voice, where vowel placement is an embodied attitude saturated with a powerful sociality.

These arguments call to mind Roland Barthes’s celebrated essay “The Grain of the Voice” (1977). Barthes opposed the structures of language and musical style to what he called “the grain of the voice,” by which he means “the materiality of the body” that comes through voice and affects the listener at the level of personal pleasure. Barthes’s claim, that this grain is the body in the voice as it sings, has become canonical among humanists and theorists of difference (see Koestenbaum 1993). But the kind of vocal anthropology evident in Fox’s and Samuels’ essays makes clear the social critique both of the arch humanist position in Barthes, and the arch speech science position on this acoustic materiality as represented by Johan Sundberg’s *The Science of the Singing Voice* (1987): namely, that it is always the body social that is enunciated in and through the voice. In both the settings described by Fox and Samuels, the sonorosity of the voice indexes a clear social agency and a sense of place. Practice-oriented approaches to the phenomenological intertwining of language and music make clear that voice itself is a way of writing against essentialization, a way of writing for performativity, and creative agency. This in no way denies the body or pleasure, any more than it denies acoustics or physiology; rather it insists that these are materially social sites as much as anything else. In other words, the physical grain of the voice has a fundamentally social life.

Voice is the embodied locus of spoken and sung performance, the site where language and music have received closest ethnographic scrutiny. But voice has a more familiar articulation in contemporary anthropology, having also become a metaphor for difference, a key representational trope for identity, power, conflict, social position, and agency. Vocality, in this light, is a social practice that is everywhere locally understood as an implicit index of authority, evidence, and experiential truth.

If voice and vocality constitute a particularly significant site for the articulation of opposition, it may be argued that the root of this articulation is social ontogeny: voice is among the body’s first mechanisms of difference. The ability to differentiate one voice from another, the ability to recognize that each and every voice is different, the ability to hear oneself at the same time as hearing others, the ability to silently hear oneself within, the ability to auditorally imagine the voice of another in the absence of

their immediate vocalic presence – these are all fundamental human capacities (Idhe 1976; Appelbaum 1990; Feld 2000). It is therefore not surprising that phrases like “giving voice,” “taking voice,” “having voice” are now routinely linked to the politics of identity, to the production of difference, to the ability of the subaltern to speak, to the ability of indigeneity movements to “talk back” and for class, gender, and race politics to “back talk” the dominant. As feminist historians Leslie Dunn and Nancy Jones write, introducing *Embodied Voices*, their edited volume on female vocalicity in Western culture: “. . . ‘voice’ has become a metaphor for textual authority, and alludes to the efforts of women to reclaim their own experience through writing (‘having a voice’) or to the specific qualities of their literary and cultural self-expression (‘in a different voice’)” (1994: 1; also see Salvaggio 1999). Linking the histories of *vox populi* to “lift every voice and sing,” vocalicity has become the site where linguistic and musical anthropology most strikingly conjoin a poetics and politics of culture.

NOTES

- 1 Timbre, unlike melody, harmony, and rhythm, is a non-segmentable feature of music. An important parallel thus exists between its decidedly marginal status within music theory and the secondary status of work on suprasegmentals and acoustic phonetics in much of formal linguistic theory.
- 2 This is clearly close to the approach used by scat singers in jazz, but the vocalisms of Bobby McFerrin, especially in his evocation of particular musical instruments, are even closer to what is being described here (see especially his CD *The Voice*, 1988).
- 3 In comparing /Cl/ to /Cr/ in word-initial position (section A1, “stop + liquid” in figure 14.2), the former items all imply re- or deflection of a force off an object without rupturing the object’s form, while the latter items all suggest that rupture of the object’s form does occur. This makes these terms especially useful for describing certain signal-processing effects, particularly those involving electronic and mechanical distortion, as well as idio-phonetic instruments.
- 4 Methodologically, the patterning of meaning within these semantic fields can be pinpointed by examining the particular syntactic frames in which items appear. The most useful frames for the studio corpus involve adverb–adjective pairs such as “too ADJ” or “not enough ADJ,” as well as imperatives or request structures such as “Make it more ADJ” or “Can you reduce the NOUN?” For music genres that highly value accuracy, as were most of the sessions from which this data was gathered, for instance, there is something nonsensical about saying **Make it more constricted*, or **That sound is not muffled enough*, or **The kick drum needs to be more boxy*. In other musical styles, however, those may be desirable sonic characteristics.
- 5 The semantic field of diffuseness is also marked by a comparatively large number of antonymic pairs (blurred/focused, smeared/crisp) that provide explicit contrastive and evaluative information about desired sound quality.
- 6 “He Stopped Loving Her Today” by Bobby Braddock and Curly Putnam. Copyright 1978, 1980 by American Music. Copyright renewed, assigned to Unichappell Music, Inc. (Rightsong Music, publisher). I choose this example because it is a widely known and easily accessible performance, regarded by my ethnographic consultants as a masterpiece of country singing style. Jones developed his style in Southeast Texas, and his internationally renowned work has in turn strongly influenced vocal practice in Texas.

- 7 Country and Western are obviously not the only contemporary popular genres that people in San Carlos listen to. Rock, oldies, reggae, chicken scratch, and hip-hop are some others. There are generational differences, certainly, at work in some of the differences in individual taste. However, many people are hardly exclusive in their preferences, moving from one genre to another with ease.
- 8 “Tense” and “lax” are phonetician’s terms – /i/ [beat] is the tense form of /I/ [bit], /u/ [pull] the lax form of /u/ [pool] – and don’t necessarily imply that the muscles are any more tense during the production of that sound.
- 9 These vowel glides may be even more prominent now than in the past, as country music’s commercial success and growing pop music sensibility make the indexical rustic associations of the singer’s accent a more prominent part of the rootedness of the song to the country.
- 10 Diverging somewhat from the commercial category of “country and western,” the band members made a distinction between “country” music and “western” music, preferring the former for its rootedness and its lack of saccharine sweetness. This distinction is akin to that between “Hard Shell” and “Soft Core” country that Richard Peterson (1997) discusses in his history of the commodification of authenticity in Nashville.
- 11 Rock music is actually fairly famous for these kinds of interpretive errors, making someone who works with language and music sometimes wonder if the lyrics matter at all. One of the more famous oronyms is contained in the Beatles’ “Lucy in the Sky with Diamonds,” in which the line “the girl with kaleidoscope eyes” is often heard as “the girl with colitis goes by.”
- 12 The band Apache Spirit, from Whiteriver on the Fort Apache reservation, has also recorded “Copperhead Road.” In their version, as well, the lead singer, Midnite Ethelbah, does not perform the twang in the word “Viet Nam.”

REFERENCES

- Appelbaum, D. (1990). *Voice*. Albany: State University of New York Press.
- Bailey, G., and Tillery, J. (1996). The Persistence of Southern American English. *Journal of English Linguistics* 24(4): 308–321.
- Bakhtin, M. (1981). *The Dialogic Imagination: Four Essays* (ed. M. Holquist, trans. C. Emerson and M. Holquist). Austin: University of Texas Press.
- Barthes, R. (1977). The Grain of the Voice. In *Image, Music, Text* (pp. 179–189). London: Fontana.
- Bartlett, B., and Bartlett, J. (1994). Engineer’s Guide to Studio Jargon. *EQ* 5(2): 36–41.
- Basso, K. (1979). *Portraits of the Whiteman*. Cambridge: Cambridge University Press.
- Bauman, R. (1977). *Verbal Art as Performance*. Rowley, MA: Newbury House.
- Bernstein, C., Nunnally, T., and Sabino, R. (eds.) (1997). *Language Variety in the South Revisited*. Tuscaloosa: University of Alabama Press.
- Bolinger, D. (1950). Rime, Assonance, and Morpheme Analysis. *Word* 6: 117–136.
- Burrows, D. (1990). *Sound, Speech and Music*. Amherst: University of Massachusetts Press.
- Dunn, L., and Jones, N. A. (eds.) (1994). *Embodied Voices: Representing Female Vocality in Western Culture*. New York: Cambridge University Press.
- Fabb, N. (1997). *Linguistics and Literature: Language in the Verbal Arts of the World*. Oxford: Blackwell.
- Feld, S. (1981). “Flow like a Waterfall”: The Metaphors of Kaluli Musical Theory. *Yearbook for Traditional Music* 13: 22–47.
- Feld, S. (1990). *Sound and Sentiment: Birds, Weeping, Poetics, and Song in Kaluli Expression* (2nd edn). Philadelphia: University of Pennsylvania Press.

- Feld, S. (1994a [1984]). Communication, Music, and Speech about Music. In C. Keil and S. Feld, *Music Grooves* (pp. 77–95). Chicago: University of Chicago Press.
- Feld, S. (1994b [1984]). Aesthetics as Iconicity of Style, or “Lift-up-over Sounding”: Getting into the Kaluli Groove. In C. Keil and S. Feld, *Music Grooves* (pp. 109–150). Chicago: University of Chicago Press.
- Feld, S. (2000). They Repeatedly Lick Their Own Things. In L. Berlant (ed.), *Intimacy* (pp. 165–192). Chicago: University of Chicago Press.
- Feld, S., and Fox, A. (1994). Music and Language. *Annual Review of Anthropology* 23: 25–53.
- Fox, A. (1995). “Out the Country:” Speech, Song and Feeling in Texas Rural Working-class Culture. PhD dissertation, Department of Anthropology, University of Texas at Austin.
- Heffner, R.-M. S. (1950). *General Phonetics*. Madison: University of Wisconsin Press.
- Hinton, L., Moonwomon, B., Bremner, S., Luthin, H., Van Clay, M., Lerner, J., and Corcoran, H. (1987). It’s Not Just the Valley Girls: A Study of California English. In J. Aske, N. Beery, L. Michaelis, and H. Filip (eds.), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society* (pp. 117–128). Berkeley: Berkeley Linguistics Society.
- Idhe, D. (1976). *Listening and Voice: A Phenomenology of Sound*. Athens: Ohio University Press.
- Jakobson, R., and Waugh, L. R. (1979). *The Sound Shape of Language*. Bloomington: Indiana University Press.
- Kirby, J. T. (1995). *The Countercultural South*. Athens: University of Georgia Press.
- Koestenbaum, W. (1993). *The Queen’s Throat: Opera, Homosexuality and the Mystery of Desire*. New York: Vintage.
- Labov, W. (1966). *The Social Stratification of English in New York City*. Arlington: Center for Applied Linguistics.
- Ladefoged, P. (2000). The Sounds of Speech. In P. Kruth and H. Stobart (eds.), *Sound* (pp. 112–132). New York: Cambridge University Press.
- Lerdahl, F. (2001). *Tonal Pitch Space*. New York: Oxford University Press.
- Lerdahl, F., and Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.
- Lomax, A. (1968). *Folk Song Style and Culture*. Washington, DC: American Association for the Advancement of Science.
- Luthin, H. W. (1987). The Story of California (ow): The Coming-of-Age of English in California. In K. M. Denning, S. Inkelas, F. C. McNair-Knox, and J. R. Rickford (eds.), *Variation in Language: NWAV-XV at Stanford* (pp. 312–324). Stanford: Department of Linguistics, Stanford University.
- McConnell-Ginet, S. (1983). Intonation in a Man’s World. In B. Thorne, C. Kramarac, and N. Henley (eds.), *Language, Gender, and Society* (pp. 69–88). New York: Newbury House.
- McFerrin, B. (1988). *The Voice*. WEA/Elektra Entertainment. CD.
- McMillan, J. B. (1939). Vowel Nasality as a Sandhi form of the Morphemes *nt* and *ing* in Southern American. *American Speech* 14: 120–123.
- Meintjes, L. (2003). *Sound of Africa! Making Music Zulu in a South African Studio*. Durham, NC: Duke University Press.
- Miller, R. (2001). *Training Soprano Voices*. New York: Oxford University Press.
- Nattiez, J.-J. (1983). Some Aspects of Inuit Vocal Games. *Ethnomusicology* 27(3): 457–475.
- Nattiez, J.-J. (1999). Inuit Throat Games and Siberian Throat Singing: A Comparative, Historical, and Semiological Approach. *Ethnomusicology* 43(3): 399–418.
- Nuckolls, J. B. (1999). The Case for Sound Symbolism. *Annual Review of Anthropology* 28: 225–252.
- Peterson, R. A. (1997). *Creating Country Music: Fabricating Authenticity*. Chicago: University of Chicago Press.

- Porcello, T. (1996). *Sonic Artistry: Music, Discourse, and Technology in the Sound Recording Studio*. PhD dissertation, Department of Anthropology, University of Texas at Austin.
- Porcello, T. (1998). “Tails Out”: Social Phenomenology and the Ethnographic Representation of Technology in Music-making. *Ethnomusicology* 42(3): 485–510.
- Preston, D. (1997). The South: The Touchstone. In C. Bernstein, T. Nunnally, and R. Sabino (eds.), *Language Variety in the South Revisited*. Tuscaloosa: University of Alabama Press.
- Rhodes, R. (1994). Aural Images. In L. Hinton, J. Nichols, and J. J. Ohala (eds.), *Sound Symbolism* (pp. 276–292). New York: Cambridge University Press.
- Salvaggio, R. (1999). *The Sounds of Feminist Theory*. Albany: State University of New York Press.
- Samuels, D. W. (1998). *A Sense of the Past: Music, Place and History in San Carlos and Bylas*. PhD dissertation, Department of Anthropology, University of Texas at Austin.
- Sundberg, J. (1987). *The Science of the Singing Voice*. Dekalb: Northern Illinois University Press.
- Théberge, P. (1994). *Any Sound You Can Imagine: Making Music/Consuming Technology*. Hanover, NH: Wesleyan University Press.
- Trudgill, P. (1983). Acts of Conflicting Identity: The Sociolinguistics of British Pop-song Pronunciation. In P. Trudgill, *On Dialect: Social and Geographical Perspectives* (pp. 141–160). Oxford: Blackwell.
- Umiker-Sebeok, D. J., and Sebeok, T. A. (eds.) (1976). *Speech Surrogates: Drum and Whistle Languages*, vols. 1 and 2. The Hague: Mouton.
- Van Leeuwen, T. (1999). *Speech, Music, Sound*. New York: St Martins Press.
- Wescott, R. (1977). Ideophones in Bini and English. *Forum Linguisticum* 2(1): 1–13.
- Wolfram, W., and Christian, D. (1976). *Appalachian Speech*. Arlington: Center for Applied Linguistics.
- Zemp, H. (ed.) (1996). *Voices of the World: An Anthology of Vocal Expression*. Book with 3 CDs. Paris: Le Chant du Monde.