JOINT SPEECH: A MUSICAL PERSPECTIVE (OR, IN PRAISE OF SINGING)

JEREMY DAY-O’CONNELL
Skidmore College

In charting the “territory between speech and song,” Cummins (2020) identifies various forms of vocal behavior involving “multiple people uttering the same thing at the same time”—what the author calls “joint speech.” I interrogate this conceptual framework in light of specific examples both musical and linguistic, which suggest a usefully expanded category: “collaborative vocality.” At the same time, I also propose a distinctively musical account of joint speech that ultimately affirms the conventional separation between music and language. Central to this account is an analysis of the unique character of singing and an insistence on the centrality of pleasure in musical experience.

Received: December 20, 2019, accepted January 31, 2020.

Key words: music and language, singing, voice, social processes, pleasure

In Act I, Scene 5 of Meredith Wilson’s The Music Man (1958), a gaggle of petty, querulous bureaucrats experiences something tantamount to a religious conversion. The source of their unlikely transformation: a simple, glorious major triad. Thanks to an impromptu singing lesson from con-man Harold Hill, the four men discover the miracle and wonder of vocal harmony, which instantly turns their strife to unity and joy. Hill’s prophecy—“From now on, you’ll never see one of those men without the other three”—speaks to the power of music not only to shape human relationships but to model them. Harmony (of the musical kind) begets harmony (of the social kind). But the whole episode also expresses a profound fact that is as obvious as it is elusive: the sheer irresistibility of music-making and of musical sound.

Wilson’s scenario is of course fictitious, even cartoonish. And yet, it is splendidly real. The singular magic of vocal harmony, indeed of music in general, truly comprises one of life’s great gifts—and not only in some abstract sense, but as an utterly visceral, immediate experience of delight.

Cummins’ (2020) insightful and thought-provoking essay foregrounds a number of diverse and often underappreciated modes of communal vocalizing. But in surveying his newly charted territory of “joint speech,” I found myself wondering where melody and harmony might be located—and more so, wondering whether and how musical pleasure might figure in the analysis.

***

Cummins’ essay is offered as a corrective to the narrow, provincial views of both linguists and (ethno)musicologists, who too often overlooked “relevant degrees of variation” in vocal behavior. To that end, he proposes joint speech as an “organizing frame” defined by four straightforward criteria: 1) “multiple people,” 2) “uttering,” 3) “the same thing,” 4) “at the same time.” While the first criterion (“multiple people”) is specific and unambiguous, the second (“uttering”) embraces the utmost breadth, alluded to in the essay’s title, and underscored by the author’s gloss: “vocalization . . . without committing to an interpretation of it as either speaking or singing.” I will return to this ecumenical orientation momentarily and will advocate a distinctly musical perspective on the “territory between speech and song.”

First though, and in the spirit of Cummins’ own deft critiques, I find it necessary to interrogate his two other criteria—“the same thing,” “at the same time”—which strike me as possibly too restrictive. Loosening the insistence on sameness and simultaneity brings into focus other areas of vocal behavior that seem highly relevant here, especially insofar as they demonstrate varying approaches to two dimensions that Cummins found neglected in previous accounts: rhythm and social context. Speech, after all, typically takes shape as different things uttered by multiple people in close succession—i.e., conversation, that most basic social act. A large and growing body of research has shown how conversation adheres to elaborate principles of inter-subjective coordination, including along broadly musical dimensions, like pitch, tempo, and rhythm (Szczepek Reed, 2007). This interactive mode of participation also animates many musical situations, such as various forms of call-and-response, the “Amens” and other improvised
interjections of Gospel worship music, and comparable participatory utterances in a multitude of musical styles, each of which likewise hinges on a certain temporal flexibility. Such a conversational approach assumes a more structured form in imitative counterpoint, which could easily be dismissed as musical esoterica, were it not for the eminently accessible practice of canonic singing—i.e., the *same thing* in *close succession*. Meanwhile, vocal harmony involves multiple people singing *related*—but *pointedly different*—things at the *same time*, and again comprising examples both rehearsed and spontaneous. In short, there are important vocal practices just beyond the declared borders of “joint speech” that seem to belong in Cummins’ analysis. Conversation, antiphony, and certain forms of improvisation, harmony, and counterpoint all potentially connect with Cummins’ interest in the “manner in which collective identities are enacted” through the voice. We might therefore recognize *collaborative vocality* as a useful, broader “organizing frame,” of which joint speech is a special case.

In any event, let us now return to Cummins’ second, more provocative criterion, which conflates speech and song under the rubric of “uttering.” Actually, Cummins himself displays some ambivalence about the matter, claiming that joint speech “erases any principled boundary between language and music” even as he consistently appeals to those very categories as heuristics throughout his essay. What’s more, his attention is repeatedly drawn to those specimens of joint speech that are, as it were, unsung (literally and figuratively). Indeed, his very choice of the term “joint speech” betrays the same preference. This is perhaps understandable; after all, speech-like examples are the ones that stand to most enrich and reorganize our received ideas about the boundaries of “music.” Cummins’ vision refreshingly emphasizes human musicality at its most elemental, highlighting such crucial and ubiquitous principles as orality, spontaneity, sociality, solidarity, and entrainment—all of which contrast with an elitist conception of music that favors mind over body, artifact over process, individual over community.¹

On the one hand, Cummins is quite right to portray vocal behavior as a continuum rather than a polarity. Speech itself is more musical than we usually acknowledge. Its fundamental sonic resources—the phonetic distinctions that produce vowels and consonants—derive from nuances of spectral information, equivalent to what musicians would call timbre. Moreover, speech contains important rhythmic and melodic features that not only contribute to meaning but also, tellingly, expose the most fluent second-language learner as non-native. (See Patel, 2008, for instance.) Even a clear case of speech can, through the mere process of repetition, turn into something inextricably perceived as sung (Deutsch, Henthorn, & Lapidis, 2011). And spoken communication boasts certain idiosyncratic prosodic formulas that can be singery in the extreme (Day-O’Connell, 2013).

On the other hand, it behooves us to ponder the uniqueness of music and the ways in which “songs influence, amuse, and engage us”—a crucial matter that Cummins raises early on but leaves unaddressed. Is singing not eminently more gripping and meaningful than other forms of communal “uttering”? Consider worshippers singing “Amazing Grace,” versus professing the Nicene Creed; protesters singing “We Shall Overcome,” versus chanting responses to “What do we want?” and “When do we want it?”; revelers singing “Happy Birthday,” versus shouting “Hip-hip-hooray”; immigrants singing a national anthem, versus reciting a pledge of allegiance. For each of these pairs of “utterances,” there are surely profound differences in the communal effects of singing versus speaking, even though the ostensible purposes and social contexts of the two acts might be quite similar. Cummins begins to tease out some possible responses to these issues, invoking such factors as “formality,” “participation,” “responsibility,” “authorship,” and “intentionality.” But these feel incomplete to me without a fuller account of the power of music.

Such an account might well begin with the special qualities and import of vocal melody. If the voice is a sonic extension of the self, the singing voice is profoundly so, since it so vividly manifests the singer’s bodily state. Compared to speech, a sung utterance seems to entail a greater degree of subjective investment. At the same time, singing leverages the “defamiliarizing” impulse of art, transporting the voice beyond the realm of the personal or the human: speech is rational and finite, while singing is poetic and transcendental. Singing acquires special layers of meaningfulness through its sheer unlikeliness, its molding of acoustic chaos into order. Pitched vocalizing thereby elevates the voice to the numinous, or at least enfold it in an aura of other-worldliness.

These provisional musings concern song-like pitch in vocal production, but what about music *per se*? Musical patterning—through rhythm, melody, and harmony

---
¹ In this, Cummins’ work resonates with that of others who likewise promote an expansive view of human musicality and of music in the fabric of culture. In addition to those cited in his essay, see also Blacking (1974), Small (1998), Agawu (2016), and Honing (2018).
(and whether sung or otherwise)—capture the ear and the imagination like no spoken utterance can. The roles of such musical elements in joint speech are alluded to by Cummins at several points but, with the exception of rhythm, are never explicated. I will offer only the most tentative proposal here: just as rhythm represents and enables the coordination of action in a group (something Cummins describes well), melody might represent the coordination of affect, and harmony, the coordination of difference. (And “harmony” here should be understood broadly as the integration of disparate elements into a larger unified whole—hence including not only tonal principles, but analogous rhythmic and textural principles such as layering and “groove.”) But surely each of these proposed functional principles, lofty though they are, would amount to nothing, were pleasure not also in play. Musical pleasure adds a crucial dimension to any consideration of “joint speech,” as it is an ideal engine for the maintenance of social relations and values.

The irresistibility of music-making and of musical sound: this is a central fact of human existence. I join Meredith Wilson in celebrating that fact. And mindful that the sources of musical pleasure remain mysterious and elusive, I also join Cummins in the conviction that “there is work to be done.”

**Author Note**

*Correspondence concerning this article should be addressed to Jeremy Day-O’Connell, Music Department, Skidmore College, 815 N. Broadway, Saratoga Springs, NY 12866. E-mail: jdayocon@skidmore.edu*

**References**


