

5

Music, Language and Linguaging

Fred Cummins

Introduction: Interrogating the Category of “Music”

Some forms of music making practice seem to be an inevitable feature of human social organization. Although there have been some rare ideological prohibitions on music, especially as it pertains to communal activity and dance, the centrality of such practices to all human societies at all times provides sufficient reason to consider music making as a foundational feature of human sociality. But in trying to elaborate upon such broad claims, questions must arise as to how to properly understand the many varieties of music making and the contexts within which they occur. To a readership accustomed to Western art music, and to the treatment of music as a commodity that can be captured in recordings, traded, and even asserted as intellectual property, it may be difficult to identify a single concept that can also extend to include the strongly participatory activities that sit at the heart of ritual and less formal activities in social contexts far removed from the “music industry.” Yet both the Western framework, and also the ethnomusicological study of participatory musical activities, presumably, still differentiate their objects of study from speech and language.

It would be a simple matter, if somewhat arbitrary, to pick out many music making activities specific to individual cultures that are not well aligned with the everyday use of the term “music” in a Western context. We might call upon the role of drumming in West Africa, on the use of chant in ritual and spiritual activities in Brazil, or on any number of indigenous and local practices to broaden the term, but to do so would be to risk adopting the voyeuristic gaze of the Western Anthropologist

witnessing, without understanding, the activities of the remote and strange. Perhaps we might pose the same questions with greater effect by looking at some musical activities more familiar to a Western audience that might lead us to ask whether we can, or should, draw the boundaries of the concept of “music” in one way rather than another.

A first example close to home worth discussing is the simple ritual of singing Happy Birthday among families, friends and colleagues. This is a barely noticed slice of ritual, far removed from any uncomfortable associations with dogma or authority, in which most of us will regularly participate. The basic form is repeated countless times every day: A motley group of participants joins in in singing the familiar verse. Someone only needs to enunciate the first syllable, often drawn out as an invitation to join in: “Haaaaaaaaa...py birthday to...”¹ Participation is not usually begrudged. Indeed, to fail to join in, by at least appearing to mouth the words, would be an act of social protest, a violation of the unspoken commitment to a ritual that celebrates an individual as a distinguished person. We would not normally go so far as to say that participation is mandatory, for the ritual is enacted by the ad hoc community of those who happen to be present, but we are certainly aware that participation is expected, and withholding it may appear as an expression of protest or disquiet. We may find ourselves joining in even to celebrate the birthday of people we do not know, as in a restaurant.

“Happy Birthday” is a song, of a sort. Indeed, there was even a somewhat successful attempt to assert copyright over the text, a peculiar state of affairs that finally ended with legal decisions in the US (2016) and the EU (2017) that concluded that royalties are not payable for this particular song. But is it music? Listening to the average production of Happy Birthday, one could be forgiven for assigning the inharmonious, infelicitous and joyous cacophony that results to a different category altogether. Nobody, I suspect, will pop on a CD of their favorite recordings of Happy Birthday and settle back to listen attentively (Marilyn Monroe’s famous rendition maybe excepted).

A second example, drawn a little further afield, concerns the manner in which unison chant is treated within the more austere branches of Wahabi Islam, a religious and social force that has frequently condemned the enjoyment of music and its practice, condoning the destruction of musical instruments and forbidding gatherings devoted to music and dance. Even within this severe worldview, the *adhan*, or chanted call to

prayer, still issues from the minarets at prayer time. A telling distinction is illustrated by activities associated with the now largely defunct self-declared caliphate of the Islamic State (ISIS). ISIS conducted a hugely effective propaganda campaign through the use of video hosting sites such as YouTube, and the distribution of material over social media. While they publicly denounced hedonistic music making practices, official propaganda from the organization almost always featured unison chant (*nasheed*) over the arousing and frequently horrific images that were broadcast. The category of music must be interrogated anew here if such distinctions are to make sense.

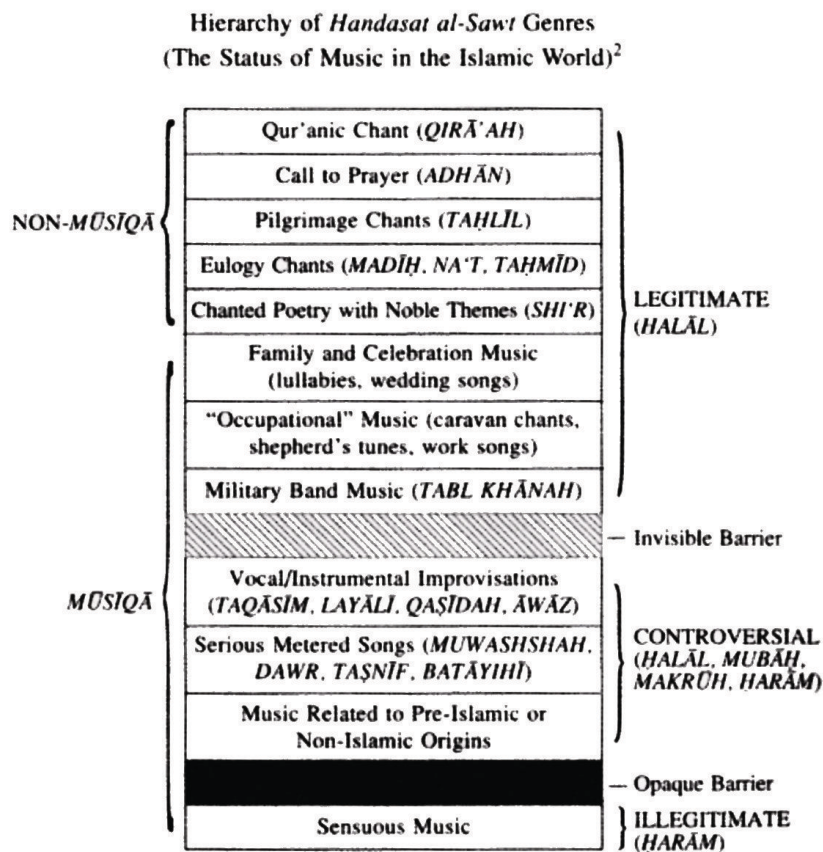


Figure 1: Category distinctions within Islam with respect to music. All such distinctions should be assumed to be contested. Reproduced from Quershi (1997)

Figure 1 reproduces a number of subdivisions within Islamic scholarship that distinguish between forms of musical activity that are to be tolerated or prohibited to varying degrees (Quershi, 1997). The specific divisions are presumed hotly contested, and for our present purposes it suffices to say that the category "music" fractures under this kind of analysis, revealing important distinctions that a single superordinate label

will collapse. We might also notice that those forms furthest removed from hedonic activities (“sensuous music”) all make use of the English term “chant” which is usefully ambiguous with respect to any division between speech and music.

Happy Birthday and sacred chant admit of a different form of characterization that is neutral with respect any such division. They are both examples of *joint speech*, which can be given a practical empirical definition as “words uttered by multiple people in synchrony.” We will explore this below. As we consider how the category of music may be expanded to include many importantly different kinds of activity, it will be first worth interrogating a second distinction concerning the relation of speech to language, or more broadly, to languaging.

Language and Languaging

Defining music is difficult. Defining language is no simpler. As a topic of scientific study, language has largely been framed as the passing of messages in one form or another from speakers to listeners, or from readers to writers. This framing has underpinned both the structuralist (De Saussure, 2011) and the generative schools of linguistics (Chomsky, 2002), and in so doing, has led to the articulation of a discipline, linguistics, that reveals much about such message passing and the encodings used, but that misses a great deal that, it might be argued, might also lay claim to being relevant to language more broadly considered. In an influential paper, for example, Hauser, Chomsky and Fitch characterize language as a faculty which is “hierarchical, generative, recursive, and virtually limitless with respect to its scope of expression” (Hauser et al., 2002). This approach describes its target (language) in terms that are agnostic with respect to the medium used (writing, voice). They make the creativity of sentence construction a central feature of language so conceived. The description of sentence structure also abstracts away from the situated uttering of any set of words, in which words are spoken within a unique and specific context. This particular leap to abstraction is a feature common to both Saussure's approach, describing *langue* rather than *parole*, and to Chomsky's, which describes notional *competence* rather than *performance*. One could put it more starkly: The theories describe what theoreticians think is going on, interpreted in terms of an underlying abstract system, rather than describing what it is that people actually do, in specific meaningful contexts.

The structuralist and generative approaches to language have greatly shaped the way we think about communication. But there is an argument to be made that the features identified as relevant in this way of thinking about language may obscure other features, equally relevant, or even more so, in considering what it was that happened to our species over the last 5 million years since the last common ancestor of the genus *pan* and the genus *homo*. In this time, a relatively small amount of biological change due to evolution has led to a vast disparity among the two branches. The record is incomplete, and we must rely on indirect evidence of cultural development such as the signature traces of music making, ritual, and fire tending to garner hints as to when language appeared. But such traces speak of more than language in the sense of message passing. They speak of coordinated consensual activity among affiliates, of cultural development more broadly, and of participatory activities that might ground social order of some sort.

It is useful, at this point, to bring joint speech to the fore, to provide a rather different view of how vocal signaling, social order, and collective enactment of a common world might be thought of. Joint speech is found whenever multiple people utter the same words at the same time (Cummins, 2019). This simple empirical definition singles out some important domains of human activity. Joint speech is a central part of practices of prayer and ritual in all cultures. It is also a reliable feature of social unrest and protest, where collective utterance serves to signal common purpose and affiliation. Joint speech is a common accompaniment to sporting and team activities generally, where its tribal nature may be stylized as in the impressive Haka chants of the Maori. Joint speech is frequently employed in teaching the young, both to ensure common attention, and to pass on culturally specific patterns ranging from Quranic chanting to learning the multiplication tables. It seems that every social order, society, or people, assembled transiently or persisting over millennia, will have activities involving joint speech at their center. And yet joint speech has hitherto been ignored as a foundational activity, masking its potential to shine light on the genesis of human sociality and identity.

In joint speech, there is no obvious distinction between speakers and listeners, as everyone is both. In joint speech, the words uttered are not generated creatively on the fly, but are authored elsewhere. The curation of the words to be uttered is a serious matter. While there may be a certain degree of improvisation at a political rally, there is assuredly no such indeterminateness in the text of the Nicene creed, first crafted

in 325 C.E. Joint speech does not have to be intelligible; indeed, ambient circumstances such as the fervor of a protest demonstration or the reverberation of a temple, frequently make it impossible to pick out individual words.² Although the words may be indistinct, it is the participation in collective uttering that seems to be of importance to those who take part. The participatory nature is underscored by the common feature of repetition that attends joint speaking. Whether iterating multiple decades of the rosary, whereby the same short text of the Hail Mary prayer is repeated 50 or more times, or chanting "lock her up" over and over at a political rally, repetition and joint speech go hand in hand. Furthermore, joint speech does not belong on the page. It is a vocal, not written, activity that is strongly tied to the situation of occurrence and its social significance. For all these reasons, joint speech has not been articulated as a distinguished form of language, and indeed has hardly been studied empirically at all.

Definitions are difficult things. While they make it easier to categorize the world, the simple drawing of a distinction in one way rather than another can greatly change the subsequent finer distinctions and categories that may be built on the first. For this and other reasons, the term "linguaging" is beginning to find some use among those studying language from a broader point of view, whose primary concern is with understanding a range of coordinative activities, serving many purposes, and involving the whole body (Thibault, 2011; Steffensen, 2011; Jensen, 2014). Perhaps it might serve us here as we undo the crisp distinctions that arise when language is regarded as the transmission of encoded messages, in order to come at the description of our coordinative activities anew. If so, there might yet be a case for using the term "musicing" (De Jaegher & Schiavio, 2017) or "musics" (Cross, 2003), to rejuvenate the way in which many sound and dance producing activities are understood. Alternatively, we may arrive at the view that music making is simply not categorically distinct from linguaging, and in this way we may better understand how such activities have played a central part in the joint enactment of shared human worlds.

Musical Features of Joint Speech

The simple definition of joint speech provided above was careful to couch the vocal activity as "uttering" rather than "speaking," in order not to prejudice an examination of the many areas of overlap and inter-

mingling between what we might conventionally regard as speech, on the one hand, and song, on the other. For where we find joint speech, musical elements are usually not far away. In discussing such speech, we might note that the sounds of speech have conventionally been divided into contrast-bearing phones or segments on the one hand, and supra-segmental (or sub-segmental) elements such as rhythm, melody, dynamics and voice quality, collectively referred to as “prosody,” on the other. The latter group is typically referred to as encompassing the more “musical” elements of the voice, precisely because they do not serve the purpose of encoding categorical distinctions. In the present discussion, we will pay particular attention to prosody as an index of the form of musicality that is involved.

We might begin by considering the swearing of a public oath of allegiance.³ This may be done once only, as when rag-tag groups of paramilitaries pledge allegiance to the Caliphate of the Islamic State, or when aspiring citizens of a country are required to pledge their new allegiance during a ceremony of naturalization. It may also be done repeatedly, as in the performative recitation of the Pledge of Allegiance found in many American schools. The differences between the two cases are instructive. When the act of swearing is purely instrumental, such that it must be done once, but only once, there is usually nothing musical about the joint speech. Its prosody resembles that of read speech, as, in fact, the text is frequently read from a sheet of paper. This public act is a performative in the sense provided by Austin (1962), in that its intrinsic purpose is to mark a transition, for example, from non-citizen to citizen, and to achieve this purpose it must be uttered aloud. This contrasts with the prosody of the daily pledge in American schools, where the text is recited with a strong beat and musical phrasing: “i PLEDGE al-LEG-iance TO the FLAG of the u-NI-ted STATES of a-MER-i-ca.” The intervals between the onsets of the stressed syllables may be experienced as roughly equal, or isochronous; the alternation between strong and weak elements is exaggerated, and the division into phrases is quite different than it would be if the text were met with fresh eyes and simply read aloud. In the repeated uttering of the Pledge of Allegiance, there is no associated change in the status of those speaking, and one might venture that Austin's account of performatives might be in need of considerable extension when we begin to consider joint speech as a form of languaging.

Repetition, of course, is the norm when it comes to joint speech, and the instrumental use of an oath in the above fashion is somewhat exceptional in being non-repetitive. The repetition from day to day of the Pledge of Allegiance induces a limited form of musicality in the resulting speech, but this pales in comparison with the enormous amount of repetition found in both prayer and protest. In prayer, we frequently find practices in which a small text is repeated over and over again. Internal repetition of this sort is associated with the mantra-using traditions of India and with the role of the Jesus Prayer in Eastern Orthodox churches, and is necessarily individual. Overt collective repetition is even more common, as evidenced by the use of prayer beads, or mala, in very many religious traditions, including all the Abrahamic faiths, Hinduism and Buddhism. Short texts subjected to this kind of repetition acquire an invariance of form, both articulatory and acoustic, and a concomitant change in, or dilution of, the sense of the words. The altered relation of the sound to meaning is a phenomenon known in the psycholinguistic literature as semantic satiation (Smith & Klein, 1990). Words function differently under such conditions, and the voice is being used here for purposes that bear little relation to the use of the voice in dialogical exchange. Analysis of speech produced in this repetitive fashion shows that there is a great deal of invariance from one repetition to another, but that such speech is frequently unintelligible as a result (Cummins, 2014). Intelligibility is, of course, not a prime concern in prayer repetition, while participation in the practice is.

Repetition is a reliable feature of joint speech in many other domains as well. In political protest, individual slogans are frequently repeated with some form of rhythmic intensification, that may be provided by the use of drums, of clapping, or of associated hand and arm gestures, e.g. fist-pumping. The embodied and repeated production enhances strong down beats. Chanting done while marching will interact with the rhythmic pattern of marching itself, giving rise to repeated metrical frameworks, even when the words still appear to be “spoken,” as opposed to being sung. However sung chants are entirely continuous with spoken chants, in both protest and in the activity of sports fans, further erasing any notional boundary between speech and song (Cummins, 2018). Military cadences (e.g., “I don’t know, but I’ve been told...”) integrate the rhythm of marching with the rhythm of chanting in a transparent manner.

Repetition may induce a further perceptual effect. When a short phrase is heard over and over, the intonational contour that is initially perceived as spoken may switch to being perceived as sung, even if the physical waveform is unchanged. This has become known as the speech-to-song effect in laboratory experiments (Deutsch et al., 2008). In the context of collective chanting, there is undoubtedly a tendency to amplify this perceptual effect by exaggerating or intensifying the melodic contours of such a phrase. Thus, in repeating short phrases, both rhythmic and melodic characteristics of the words uttered are exaggerated, and we find ourselves in territory that is ambiguously related to the categories of speech and song.

Many forms of joint speech are more overtly musical, with melodic elements that are explicitly composed, rather than emerging from the activity of repetition. Here we will encounter the many sacred and ritual chant forms that find extravagant cultural elaboration, including the Gregorian chant of the Western Christian traditions, the *nasheed* traditions of Sunni Islam, the many and rich varieties of chanting found in Sufi communities, and more besides. Although each such tradition bears the stamp of local cultural identity, and may even serve as a potent index of such distinctness, many common features may be observed. In many forms of a cappella singing, the length of individual phrases is derived from the length of prose sentences, so that no metrical hierarchy is present. Instead of a well-developed meter, a sustained beat is present, that may also be accompanied by simple percussion, drone, or unison instrumental parts. Instrumentation is typically sparse, and polyphony is also absent. Such chant traditions may be very long lasting. Gregorian chant dates back to the first centuries of Christianity, while Vedic chanting has been the authoritative form of transmission of the Vedic scriptures for over three thousand years. As a testament to the persistence of such traditions, Vedic chanting practices contain elaborate error correction methods, involving the precise sequencing of syllables both forwards and backwards, separating sound from sense and making any idiosyncratic deviation by one individual obvious when such chanting is done communally. Oral transmission is thus not intrinsically any less reliable than written transmission under these circumstances.

Uttering and Commitment

Joint speech lives in the performance, not on the page. It is in the act of recitation, of chanting, that its significance inheres. It is not enough that

the words be known, they must be uttered, and the uttering is collective. It is worth drawing out some well known contrasts between written and spoken words, contrasts that are lost when language is regarded as a code that may be implemented indifferent in either modality. Walter Ong (Ong, 1982) and David Olson (Olson, 1996) have teased out many of the differences between words spoken and words written — differences that affect the entire world within which interpersonal co-being is negotiated and sustained in many ways. Spoken words are necessarily spoken by someone in a specific context; with that, the speaker is responsible for the commitments that arise in the uttering in a manner that is quite different from the way writing works. A written sentence may quickly lose its necessary association with its author, and the sense of the words can be debated in a free-floating manner, without obligatory reference to the circumstances in which they were penned. Spoken words are tied to the context of their uttering, and they commit the utterer in a direct fashion.

The use of joint speech to swear public oaths of allegiance illustrates the necessary link between uttering and commitment. Likewise, lapsed Catholics who attend a church service, will typically join in the various kinds of postural synchronization — standing, kneeling, sitting along with everyone else — but joining in the prayers, and the overt vocalization of faith of the *Credo* is, for many, a bridge too far. Both of these examples make use of joint speech that is more speech-like than musical. There is no beat, much less meter, to the words uttered. Participation is helped by either having the words written on a sheet (as in the oath at naturalization ceremonies) or the words are very well rehearsed and known to everyone, as in the *Credo* and certain other key prayers.

In joint speech forms in which a more regular beat is present, it is somewhat easier to join in, simply because the temporal predictability of the beat facilitates synchronous uttering. As we survey a range of forms of joint speech, from the least to the most musical, it seems to me that the more overtly musical features of the activity have two contrasting effects: they facilitate participation, and they reduce the perceived cost of joining in. It is easier to participate as the musical elements make the overall Gestalt of the utterance more predictable. A beat, or especially a strong meter, makes not just the articulation of words more regular, but it facilitates the whole-body entrainment found in dancing, marching and vigorous gesturing. But as it becomes easier to participate, so the commitment associated with joining in seems to go down. As we consider more musical examples, we find that mouthing lyrics to a song

does not bring a strong commitment to the sense of those words. Some songs may be exceptional in this regard. Participation in the singing of a national anthem is fraught with significance, precisely because it becomes an overt signal of allegiance to a specific group or ideology. But most songs are relatively inconsequential in this manner.

In considering the consequences of joining with others in a highly significant collective activity, it is worth considering the type of speech act that is at home in joint speech. Utterances have several kinds of significance and consequence, as first singled out by Austin (1962) in his discussion of the varieties of meaning of specific utterances performed under specific conditions, and as elaborated by his student John Searle in speech act theory (Searle, 1969). Such theorizing has hitherto considered utterances performed by a single speaker within a dialogical situation. If we extend the broad idea to joint speech, we find that performativity is a far more central aspect of collective uttering than when speakers and listeners converse. While relatively few sentences uttered in dialogue have overt performative functions, almost all joint speech utterances seem to have identifiable and readily perceptible functions. In joint speech, we demand, we assert, and we insist. We may request, beseech, and venerate. And as we do so, we do so in contexts that make the performative function of the act perfectly clear. While the sounds of individual words in joint speech are frequently indistinct, even to the point of unintelligibility, one might argue that joint speech demands a different sort of interpretation. Joint speech is intelligible precisely when its performative function is overt and unambiguous. As long as the anger, the piety, the passion, or the joy is immediately obvious, joint speech makes sense.

One intriguing feature of joint speech bears mention here. When two people engage in dialogue, it is entirely possible for one participant to intentionally deceive the other. Lies are a necessary possibility in conversation. But in joint speech, the structural elements that make lying possible are absent. One could enjoin a bunch of speakers to chant “The Earth is flat!” something they all know is demonstrably false, but that would still not seem to constitute a lie, as there is nobody to deceive, and no speaking subject who might withhold, distort, or misrepresent their private views. Joint speech seems to rejuvenate many legacy issues in the philosophy of language.

Is it Language? Is it Music?

Within the various Pentecostal protestant traditions, speaking in tongues, or glossolalia, is sometimes found. To those within the tradition, this behavior arises when the person becomes a vehicle for the Holy Spirit to speak through them. The sounds produced are meaningless syllables, though the experience of uttering them may be of great significance for the speaker. Such practices are relatively rare, and for those who engage in them, they occur only occasionally, usually within specific collective church services. Most of us will never speak in tongues.

But every person who can speak will have spoken in synchrony with others. Most of us do it frequently. Some of this joint speaking takes place in churches, but it takes place also in schoolrooms, courthouses, the public street, the sports stadium, and more besides. In this case too, the occasions where such speaking takes place are of great significance to the speakers, but the occasions are vastly greater in number, and the speakers who partake are all speakers, not just a few.

And yet there seems to be a wealth of empirical research on glossolalia, and none, or virtually none, on joint speaking. One of the first questions one might have about glossolalia is whether the syllables are drawn from any extant or extinct language, such as ancient Aramaic. This is the kind of question that can be easily addressed. The answer is an unequivocal “no.” One can characterize the phonetics of speech spoken in tongues. There are questions that are relatively easy to ask about the psychological makeup of those who practice glossolalia. One can ask about the cultural transmission of such practices, their geographic distribution, the relevance of the behavior for theology, and so on. Glossolalia, as a research topic, seems to be rather well addressed. Google scholar provides roughly 11,000 references for publications relevant to its study. On joint speech, and its synonyms, chant, unison speech, synchronous speech, choral speech, there is almost nothing. What little there is seeks to elaborate upon culturally specific forms within equally specific traditions. The domain-transcending features that serve to link the protester, the football fan, and the monk, have not been thematised as an object of study. Why this enormous disparity?

If a question can be well posed, it can be addressed, and in all probability it can be answered. Where we have well-formulated questions in matters to do with language or with music, there are scholars working to arrive at answers. If joint speech is as ubiquitous and as important as I

have suggested, the problem would seem to lie with the posing of the right question, rather than with the challenge of addressing a well-formed one.

Some of the difficulty involved was hinted at when we asked above: "Is the singing of Happy Birthday music?" There is no clear answer to this, and at first there is also no obvious reason to insist on one. The ritual associated with Happy Birthday is pedestrian, unthreatening, and seems at first blush to be entirely unassociated with the big questions of philosophy and science. Yet as we grant the question some attention, paying heed to the consequences of the behavior, to the unwritten rules that affect the behavior of all those present, we begin to see that the ritual that seemed so inconsequential turns out to be masked by its very familiarity. We fail to notice that which is closest to home. And the practices in which joint speech is found are usually of this very familiar nature. They support and structure the social processes in which we exist. The rituals and practices using joint speech are those that establish common ground among participants, orienting them in a unified fashion towards the unknown and the other. Whether we adopt the identity of a Manchester United fan or of a Hare Krishna devotee, in each case the rituals, emblems, gestures, and paraphernalia are familiar to us, and they form part of that which unifies our group, our tribe, our people. We comment on the odd and unusual. We tend not to focus on the expected, the reassuring, and that which we take for granted. Furthermore, such practices do not create static identities. Rituals must be undertaken, chanting must be done, to sustain these identities. There is no difficulty with being both a Manchester United fan *and* a Hare Krishna devotee. One need only take part in both sets of rituals, chanting as a soccer fan now, and as a monk later.

When we get away from the entrenched idea that language—narrowly conceived—is about propositional content, describing a pre-existing world, we find that languaging—broadly conceived—is central to the practices that create and sustain specific meaningful worlds in the first place. And having arrived at this point, we have every reason to suspect that the received division between language and music has blinded us to their commonalities, their mutual involvement in the practices and behaviors that bring specific shared worlds into being. The lifeworld of a group, of a people, is grounded in the rituals observed by all without thinking, and in the commonplace and everyday way of doing things. Such commonalities provide a shared orientation towards selves and others. This is the ground from which we go forth and find things to discuss. We rarely look down to see where we stand.

Notes

- 1 The manner in which the singing is initiated itself varies across communities. In Columbia, I'm told, clapping is used to get things going.
- 2 The sustained existence of liturgical languages such as Coptic or Ge'ez that are spoken in no context other than ritual provides a stark illustration of the way in which the words of joint speech matter in a manner very differently from conversational speech.
- 3 Many examples of joint speech illustrating these and other characteristics are available at jointspeech.ucd.ie.

References

- Austin, J. L. (1962). *How to do things with words*. Clarendon Press.
- Chomsky, N. (2002). *Syntactic structures*. Walter de Gruyter.
- Cross, I. (2003). Music and biocultural evolution. In Clayton, M., Herbert, T., and Middleton, R., editors, *The Cultural Study of Music* (pp. 39–49.) Routledge.
- Cummins, F. (2014). The remarkable unremarkableness of joint speech. In *Proceedings of the 10th International Seminar on Speech Production* (pp. 73–77). Cologne, DE.
- Cummins, F. (2018). The territory between speech and song: A joint speech perspective. *Current Anthropology*. Under review.
- Cummins, F. (2019). *The ground from which we speak: Joint speech and the collective subject*. Cambridge Scholars. Forthcoming.
- De Jaegher, H., & Schiavio, A. (2017). Participatory sense-making in joint musical practice. In *The Routledge Companion to Embodied Music Interaction* (pp. 31–39). Routledge.
- De Saussure, F. (2011). *Course in general linguistics*. Columbia University Press.
- Deutsch, D., Lapidis, R., & Henthorn, T. (2008). The speech-to-song illusion. *J. Acoust. Soc. Am*, 124(2471), 10–1121.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569–1579.
- Jensen, T. W. (2014). Emotion in languaging: languaging as affective, adaptive, and flexible behavior in social interaction. *Frontiers in Psychology*, 5, 720.
- Olson, D. R. (1996). *The world on paper: The conceptual and cognitive implications of writing and reading*. Cambridge University Press.
- Ong, W. J. (1982). *Orality and literacy*. Methuen & Co. Ltd.
- Qureshi, R. B. (1997). Sounding the word: Music in the life of Islam. In Sullivan, L. E. (ed.), *Enchanting powers: Music in the world's religions* (pp. 263–298). Harvard University Press.

- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge university press.
- Smith, L., & Klein, R. (1990). Evidence for semantic satiation: Repeating a category slows subsequent semantic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(5), 852. 10
- Steffensen, S. V. (2011). Beyond mind. In Cowley, S. (ed.), *Distributed Language* (pp. 185–210), John Benjamins Publishing.
- Thibault, P. J. (2011). First-order languaging dynamics and second-order language: the distributed language view. *Ecological Psychology*, 23(3), 210–245.