Learning Jazz Language by Aural Imitation: A Usage-Based Communicative Jazz Theory (Part 1)

MATTIAS SOLLI, ERLING AKSDAL, JOHN PÅL INDERBERG

Abstract. *How can imitation lead to free musical expression? This article explores the role of auditory imitation in jazz. Even though many renowned jazz musicians have assessed the method of imitating recorded music, no systematic study has hitherto explored how the method prepares for aural jazz improvisation. The article picks up an assumption presented by Berliner (1994), suggesting that learning jazz by aural imitation is "just like" learning a mother tongue. The article studies three potential stages in the method, comparing with imitative, rhythmic, multimodal, and protosymbolic behavior of infant perception (building on the works of Stern, Trevarthen, and Merleau-Ponty). The demonstrations of the aural-imitation method draw on pedagogic experiences accumulated since 1979 in the Jazz Program at the Norwegian University of Science and Technology. By analyzing structures of behavior suggested by the method, the article indicates*

Mattias Solli, PhD, is a Norwegian philosopher and former jazz student and jazz musician currently working as a researcher at the Department of Philosophy and Religious Studies, Norwegian University of Science and Technology (NTNU). In addition to jazz music, his research interests are phenomenology, hermeneutics, aesthetics, and ethics. Solli is co-editor (with Ståle Finke), of *Oppløsningen av det estetiske: Kunstteori og estetisk praksis* (2021; Dissolution of the Aesthetic Phenomenon: Art Theory and Artistic Practice.)

Erling Aksdal is a Norwegian jazz pianist and composer. As a pianist, he has performed and recorded with the likes of Chet Baker, Elvin Jones, Warne Marsh, and Lee Konitz, as well as many notable Norwegian musicians. He has served as a council member of the Association Européenne des Conservatoires (AEC) and was a founding member of AEC's Pop Jazz Platform. Aksdal is currently head of the Jazz Program at the Department of Music, NTNU.

John Pål Inderberg is a Norwegian jazz musician and bandleader and a professor emeritus of the Department of Music, NTNU. As a saxophonist, he has performed and recorded with Warne Marsh, Lee Konitz, Chet Baker, Bob Brookmeyer, Gil Evans, Chick Corea, and Pat Metheny, as well as with many notable Norwegian jazz musicians. Inderberg taught the saxophone to several generations of outstanding Norwegian jazz musicians.

Journal of Aesthetic Education, Vol. 55, No. 4, Winter 2021

key traits that render aural jazz improvisation possible, such as a fundamental sense of rhythm, formation of symbolic behavior, joint musical attention, and the facility to "hear via the other." In conclusion, we critically address a frequent theoretical model describing musical improvisation as a synthesis of discrete elements or building blocks.

> Keep listening. Never become so self-important that you can't listen to other players.
>
> —John Coltrane

1. Jazz and Aural Imitation

Jazz improvisers generally agree that listening is of vital importance.[1] Jazz music is *ear music*, as Berliner puts it.[2] Though most professional musicians today can read music fluently, improvisation primarily unfolds through aural interaction. A sensitive, rapid, and creative ear is an enabling condition for being in the business. The core of this skill-based communication evolves in the facility to understand and generate music as heard—not intellectually but aurally, musically, and spontaneously.

The aural orientation of contemporary jazz is not accidental. Jazz originated in an oral tradition, namely, the African American tradition.[3] The same culture also developed a method for ear training based on aural imitation. Autobiographies and interviews with pivotal artists such as Charlie Parker, Miles Davis, John Coltrane, and Lennie Tristano,[4] or subsequent names like Cecil Taylor, Chris Potter, and Jan Garbarek,[5] describe how they began by imitating recorded music. They played the grooves off the records (and neighbors into despair), singing along with rhythmic figures, melodies, chords, and solos, gradually reproducing what they heard on their instruments. "You copy and imitate, and try to sound 'like that' as much as possible," as Galper puts it.[6] Berliner summarizes the imitative processes thus: "Cultivating an aural grasp of a solo before its reproduction with musical instruments . . . trains the voices of students and gives them a grounding in what are for improvisers essential linkages among voice, ear, and instrument."[7]

While originating in informal institutions, the aural imitation method is today part of academic jazz programs around the globe, for example, the American Berklee College of Music and New School of Music,[8] the Jazz Program at the Norwegian University of Science and Technology (NTNU),[9] and others.[10]

1.1 Berliner's Assumption: Communicative Musicality

But how can imitation prepare for musical improvisation? Isn't improvisation precisely the opposite of imitation? Berliner picks up a frequent

comparison among jazz musicians, stating that imitating recorded music is like learning a language.[11] More precisely, it is like learning a mother tongue.

Call this *Berliner's assumption*, although the idea probably stems from the interviewed musicians: "Just as children begin to learn to speak their native language by imitating older, competent speakers, so young musicians learn to speak jazz by imitating seasoned improvisers."[12] Imitating recordings is to get the "vocabulary" straight, just like the process wherein children acquire the language of the people around them. The imitator incorporates accent, articulation, pitch, phrasings, and dynamics of a musical style, gradually making it part of his or her expressive power. The auditory dimension unfolds in contexts of use. The sounds make sense within specific frameworks of human communication. Gradually, playing music becomes as natural for the player as using lips, tongue, and vocal cords when speaking. Just like words "come to" the fluent speaker, rhythmic, harmonic, and melodic features "come to" the musician. Finally, the music is as available for the aural musician as language is for the competent speaker. It fills itself in, so to say, supporting free expression.

Given the agreement among musicians, teachers, and students, it seems beyond doubt that the aural imitation method works. Moreover, granted the substantial idiomatic variations among players who use this way of learning, it looks correct to detach the aural imitation method from a specific jazz idiom, say, bebop. And, crucially, the fact that the method originated in the African American oral culture makes it plausible that the potential learning process has something in common with linguistic behavior. In the African and African American contexts, musical and verbal forms have traditionally always been neatly associated.[13] Ways of making music and ways of talking tend to blend. Scat singing in jazz, early blues, or, more recently, rap music are typical examples.

Against the background of cultural origin, Berliner's assumption seems to reflect an insight investigated elsewhere by Trevarthen, Malloch,[14] and others,[15] suggesting that music and language are interwoven phenomena. Without saying that music and language are the same, human communicative musicality renders both phenomena possible. "Musicality is a communicated talent, and it is a talent for communicating in live company," writes Trevarthen.[16] Music and language are cultural, expressive, and communicative phenomena. They are organized means and media of communication, and they are profoundly personal *and* intersubjective at the same time.

According to Trevarthen (and Malloch), the talent to generate, manipulate, and express oneself in music or language is inseparable from an innate desire for imitation and cultural learning. "Even though few in any society may be known as musicians, professional story-tellers in sound, all of us are . . . 'musical' from birth."[17] We are born with a drive, need, and intuitive readiness to imitate and move rhythmically with others and to engage

in endless varieties of dramatic temporal narratives.[18] "The core of every human consciousness appears to be an immediate, unrational, unverbalized, conceptless, totally atheoretical potential for rapport of the self with another's mind," Trevarthen states.[19] In this framework, this core is neither a Cartesian consciousness located "behind" the humdrum reality of interhuman relations nor something irrational as such. It is instead a communicative structure of behavior spontaneously unfolding between communicating human beings. It is *un*rational, *a*theoreatical, and *non*directive. It follows its course and logic of action, unfolding in the imitation and synchronization of people's behavior.

With Trevarthen and Malloch's concept of communicative musicality, Berliner's comparison between aural imitation of recordings and children's language acquisition seems to make sense. Both learning processes seem to animate and cultivate this fundamental resource in life. Both seem to draw on and develop the human capacity to "swing" with others and do so within the possibilities and constraints of shared media of expression. Both reveal the human ability to generate meaningful material *with*, *against*, and *in accordance with other*s, in real-time, communicative interaction. Left unexplored, however, are the many questions regarding how the aural imitation method works, notably,

- What, more specifically, is implied in Berliner's assumption or the frequent association between the learning of jazz by aural imitation and first-language acquisition?
- How can the imitation of recordings prepare for musical fluency and artistic autonomy within musical collectives?
- How can imitation ground the essential linkage among voice, ear, and instrument, as Berliner claims,[20] and what do *grounding* and *linkage* mean anyway?
- Granted that all humans possess musicality, as Trevarthen and Malloch hold, how does aural imitation help the aspiring jazz musician to develop his/her potential into the specialized skills of aural jazz improvisation?

Berliner leaves these questions hanging, and no systematic study within phenomenology, psychology, or pedagogy has picked them up. While philosophers have pondered the relationship between music and language for ages[21] and while a vast and multidisciplinary literature is produced on jazz improvisation, the intriguing phenomenon of learning how to improvise music by aural imitation has hitherto been unvisited. Few things are more common within the jazz literature than to mention briefly that jazz musicians play by ear, without pursuing propaedeutic and behavioral implications. Words like *listening*, *inner ear*, or the German/Scandinavian *Gehör* are tossed around as if the meaning of these words were self-evident or too trivial to analyze. At other times, the aural skill is left unmentioned

altogether, in favor of all kinds of other aspects of jazz improvisation.[22] Even jazz musicians who demonstrate superior fluency in aural communication seem to forget to elaborate on how they once obtained their aural facilities when asked to reflect upon how they improvise.[23] In effect, to the best of our knowledge, the theoretical literature on jazz improvisation misses a critical perspective. If aural facilities are conditions of possibility for jazz improvisation—as seems to be the consensus among jazz musicians around the world—it seems highly pertinent to know more of what it implies to develop aural facilities apt for jazz improvisation. The lack of knowledge calls for a comprehensive and systematic investigation of the enabling conditions of aural jazz improvisation.

Against this background, the purpose of this article is to flesh out in practical pedagogic, phenomenological, and psychological terms how the aural imitation method works. The general idea is to investigate critically Berliner's assumption that learning jazz by aural imitation is "just like" learning a mother tongue.

1.2. Doing as Musicians Have always Been Doing

This article seeks structural similarities between the propaedeutic process of imitating recordings, on the one hand, and critical traits belonging to the first months of language acquisition, on the other. The work of developmental psychologists Trevarthen, Stern, and Tomasello will be used to review the perspectives on child development. Phenomenologist Merleau-Ponty's work will help articulate a structure of embodied aural behavior. There are some differences between the theorists[24] and ongoing debates regarding specific aspects of their conclusions.[25] However, for the purpose of this article, the theories are compatible and uncontroversial enough. The core agreement evolves in what we, inspired by Tomasello,[26] can call usage-based approaches to human communication. Rather than postulating innate grammars as Chomsky did,[27] our theorists seek general behavioral structures embedded in, and developed by, the use of empirical communicative means.

One innovative approach of this article is to consider the use suggested by the aural method itself. Rather than relying on autobiographic interviews with expert musicians (as most theories on jazz improvisation do today), we will pursue how the aural and propaedeutic practice suggests theoretical perspectives on imitative listening, somewhat analogous to how paintings, for Merleau-Ponty,[28] suggest fundamental aspects of seeing. By analyzing three stages in the method, we will build up a concept of what it implies to develop professional aural skills apt for aural jazz improvisation. Thus, we will also provide a new interdisciplinary bridge between music education, phenomenology, and developmental psychology. The combination of practical and theoretical insights allows for novel analyses of conditioning aspects of jazz improvisation, and, more generally, of overlapping traits in musical and linguistic communication.

The specific version of the aural imitation method we will use stems from the Norwegian Jazz Program at NTNU in Trondheim. Two of the authors, Inderberg and Aksdal, have taught at this institution since 1979 and 1990, respectively. Solli also started as a jazz student in at NTNU before studying philosophy. Inspired by the golden standards set by the oral African American jazz institution, Inderberg and Aksdal have applied and investigated the aural imitation method for more than forty years. This does not mean that the students must imitate and play jazz music originating in the African American tradition, say, swing-based, tonal jazz (although we will soon use Cannonball Adderley's music as our point of departure below). It only means that the students are encouraged to do so—by their own free will—as improvising musicians always have been doing. They are encouraged to imitate recorded music very precisely by ear—meticulously singing, clapping, and stamping along with the audible music, before transmitting what they hear to their musical instruments. Within limitations indicated in the next section, the music can be of almost any style whatsoever. The imitative process itself prepares the student for playing in bands with higher degrees of precision and creative independence.

1.3. Embodied Understanding: Horizontal and Vertical Implications

We will try to demonstrate in some detail how the aural imitation method facilitates a form of integrated musical learning. It is mainly the moving and expressive body that learns to communicate musically—not the analytical intellect. The aural imitation method initiates not a cognitive concept but an intuitive and atheoretical aural grasp of the musical sense. How the student *thinks* about the music is irrelevant compared to how she or he hears and *does* the music through enactive efforts.

To conceptualize this form of embodied understanding, Merleau-Ponty's concept of body schema will function as an overarching term.[29] Body schema describes the human as a functional and expressive whole. That is, it expresses no absolute schisms between mind and body, nor between intrabodily experiences and extrabodily environment. Mind and body are entangled, the "inner" blends into the "outer," and vice versa. Moreover, body schema involves a primordial knowledge—a form of perceptual organization that goes on before things eventually are articulated in the mind in the form of an *I think*. "We discover beneath intelligence and beneath perception a more fundamental function," writes Merleau-Ponty.[30] Prior to explicit cognition, the body's intelligence organizes the perception of spatiality, temporality, and intersubjective relatedness. The body schema gives us a "global, practical, and implicit notion of the relation between our body and things, of our hold on them."[31] This intentionality of the body is compared to the vague and inexplicit sensibility of the so-called sixth sense.[32]

In commentaries and invokings of Merleau-Ponty's theory, it has become customary to separate body schema and body image.[33] Body image includes

immediate consciousness, conceptual constructs of self or environment, and conscious feelings about the body: "sometimes straight forward and sometimes indirect and symbolic."[34] This is summarized by Gallagher: "A body image consists of a system of perceptions, attitudes, and beliefs pertaining to one's own body. In contrast, a body schema is a system of sensory-motor capacities that function without awareness or the necessity of perceptual monitoring."[35]

In our analysis, the aural imitation method facilitates a process of musical integration primarily on the level of body schema. "A movement is learned when the body as understood, that is when it has incorporated it into its 'world,' and to aim at things through it; it is to allow oneself to respond to their call," writes Merleau-Ponty.[36] This sits well with the potential outcome of the imitative learning process: the students incorporate the musical language into their particular body schema, allowing them to respond meaningfully to the calls of the music. The term "movement" is a little misleading, however, as the learning harbors a communicative dimension of hearing via the other[37] in the developed usage of joint musical attention.[38]

With Gallagher, it makes sense to keep this learning process distinct from the more straightforward cognitive aspects of the body image. While experience shows that *sometimes*, the learning process can be noticed and appreciated in the form of "now I get it," it might just as easily happen without any specific notification. Like a child who just picks up a new word without even noting it, the student picks up a new means of communication without any particular experience or conscious monitoring. Still, transformation and emancipation have occurred. The intelligence of the body has adapted the music. Established is a form of tacit knowledge[39]—except *tacit* is perhaps not the best word, insofar as the understanding is auditory, musical, and communicative. The knowledge resounds.

However, while the term *body schema* seems to fit the phenomenon, it would be wrong to assume that the learning process is altogether unconscious. With Berliner[40] and Dewey,[41] one could instead say that the student potentially learns to *think* in the musical language she or he imitates, that is, in the medium-specific musical significations. So construed, one might say that the learning process alters something on the level of body image as well—except, again, the word *image* is not appropriate. (Does it not reflect a certain dominance of the visual, which might be at odds with the thorough aural orientation of oral communities? We leave the question unanswered.) The student has potentially increased his or her auditory understanding of how symbolic behavior unfolds in aural, musical, and intersubjective relatedness. The student has learned to hear and enact a system of perceptions embedded in rhythm, tonality, and *ways* of using the musical language. In other words, the student has instituted in him- or herself what musicians typically refer to as musical *ear*.[42]

Curiously, our semantic clarifications along the way are not mere semantics. It illustrates the fact that the aural imitation method is a hitherto unexplored practice. To avoid unnecessary complexity, however, we will refer to embodied learning primarily in terms of body schema.

This study is no by means the first to call upon Merleau-Ponty to conceptualize embodied musical and communicative learning.[43] What we have not seen before, however, is the radical perspective on aural and imitative learning. Our analysis will suggest novel perspectives on how Trevarthen's and Stern's notion of musicality and multimodality can be read into the intelligence of Merleau-Ponty's body schema via the aural imitation method. We will try to demonstrate how the method of minute auditory replication of music can be said to animate latent communicative resources already present in the body schema since early childhood.

In effect, the following analysis will indicate how Berliner's assumption harbors two sets of implications, one *horizontal* and the other *vertical*. When we talk about structural similarities between the method and first-language acquisition, we unfold, in effect, a horizontal comparison of two radically distinct experiences. For as common sense would agree to, imitating recordings as a relatively mature human being is not the same as being a newborn infant about to learn its first language. Nevertheless, across the differences, Trevarthen, Stern, and others will help us draw out structural similarities between the two modes of imitation. The similarities evolve in the shared struggles for a full match between imitator and the imitated, the prime role of rhythm, and the usage and development of joint attention.

At the same time, the horizontal comparison also implies a vertical relation of ontogenetic formation. According to Stern, Trevarthen, and Merleau-Ponty, infant imitation *preforms* and *overdetermines* whatever goes on later in life—including the aural imitation of recordings. The imitative experiences represent what Merleau-Ponty would call a "change in the structure of consciousness, the institution of a new dimension of experience, and the deployment of an a priori."[44] Infant imitation institutes a general style of behavior effective in the body schema—a style that is not surpassed later in life but determines subsequent behavior in largely general manners.[45] Welsh puts the same point: "Our primordial existence is not just the earliest days, weeks, and months of our life, but an anonymous, subjective presence in our current existence."[46] Hence, against this background, the aural imitation method *draws on* vital resources already latent in the body schema. The method cultivates resources already present in the imitator's general abilities for communicative musicality.

1.4. The Educational Process about to Be Explored

Before we continue the analysis, it is worth reflecting upon the educational process about to be explored. To what extent is this article *exclusively* about

learning the jazz language? What about other musical genres? Besides, what do we mean by the word *jazz* in the first place? Jazz is an essentially heterogeneous term. Attempts to find a shared criterion capturing all the music associated with this term have resulted either in controversies or in concepts that exclude much music that the public would ordinarily call jazz.[47] How are we going to relate to this fact?

We will continue to use the word *jazz* in a pragmatic sense, describing the music originating in the African American tradition in the twentieth century, soon picking up a recording of Cannonball Adderley for exemplification. However, the learning process described around this music is *not* limited to Adderley's jazz idiom or any specific genre. Rather, as far as we can see, learning processes are general. They draw on universal human resources that arguably are in play in all human reciprocal interaction and music-making. Apart from the fact that the authors genuinely like jazz, we choose this music because it brings to the fore these vital and enabling behavioral structures. Moreover, Adderley's music proves the fluent use of a musical language transmitted and unfolded by ear, regulated by a free yet very disciplined reciprocal unfoldment of musical sense among the players, a clear-cut rhythmic logic, and the hegemonic tonal system in Western music: twelve-tone equal temperament and major-minor tonality.[48] Within the criteria suggested in the next section, we could have used a wide range of other musical forms, taken from all corners of the musical world. While it falls outside this article's scope to pursue alternative analyses, we encourage the reader to apply the aural imitation method to other kinds of music and, if needed, to modify the theory thereafter.

But what about the formation of *knowledge*? Suppose we accept that the aural imitation method teaches a musical language; what does this process have to do with the educational process generally associated with higher education?

It is known that the history of Western science and education systems coincides with the history of literacy.[49] Western systematic thought began with the introduction of the alphabet. Writing is a powerful technology rendering possible what Ong calls separation of the "knower" and the "known."[50] When somebody writes down an idea or an argument, it can circulate among readers, independent of the author. The reader does not have to share time and space with the knower but has access to his/her ideas across geographical and temporal distances. Music notation exemplifies the same independence. It allows musicians to play new songs without hearing them first.[51]

However, writing is not just a handy technology for information distribution. It has significantly affected how Western societies evaluate knowledge.[52] The written word typically has a higher status than the spoken word. It is considered more accurate, truer, and even more real than real-time

communication. According to Ong, it is no coincidence that learned Latin once was *the* academic language.[53] The fact that nobody spoke Latin as a native language after some point in history gave the language authority and a blank appearance of "knowledge." Learned Latin was *book language*. It was raised above the "trivia" of the mother tongue.

While Learned Latin no longer is academia's language, values from the same culture are still operative. For instance, a considerable body of pedagogical material approaches jazz as if the music could be learned by reading.[54] Jazz has been "decoded" into specific genres, each with its "underlying" theoretical formula. With Ong, it is tempting to describe what is taught here as Learned Jazz. This is a language of the books, cut loose from the perceptual reality as a resounding and communicative phenomenon. Learned Jazz appeals more to students' intellectual capacities than the fundamental communicative abilities unfolding in the oral domain of everyday communication.

Contemporary music education and research prove a similar heritage. Music students are often encouraged to reflect on their practice as part of their educational processes.[55] In texts or verbal conversations, they are supposed to contextualize and justify their own choices, preferably with the help of theory, to develop a methodological and academic awareness of their playing. In other words, to become "educated," they have to draw out abstract figures describing what the musical process "really" is about. The same tendency characterizes much of the focus on artistic research, wherein the word *research* seems to be synonymous with the meta-reflections on the music.[56] Why is it not enough just to play music in these educational and investigational contexts? Why are aspiring and expert musicians expected to analyze what they are doing in words instead of just building their capacities for aural-musical communication and develop the sonorous-artistic medium? Probably, things are as they are because of the Western literary tradition's values. The rational language of academia is verbalized and sharable, just like a text. Only thus is the learning process recognized as the formation of knowledge proper. (Curiously, evolving here might also be a reason that professional jazz musicians often forget to mention or elaborate on their capacities for aural communication when interviewed about jazz improvisation. Maybe these capacities are simply ignored as banal preconceptions due to the values associated with literacy?)

By contrast, this article illuminates and defends the formation of an oral, aural, and communicative knowledge, which is thoroughly nonscriptural. For hermeneutical background, it is worth noting that the recording industry represents another powerful technology for separating the knower and the known, except that the medium is sonorous rather than visual, unlike the written text. The listener does not have to be present with the player but has direct access to the acoustic (that is, nonwritten) music. The African

American community embraced the possibilities given by this new technology.[57] It fitted right into the culture's general oral orientation, with the emerging jazz community specializing in learning from records.

Moreover, as pointed out by Sidran,[58] the African American oral tradition behind this technology appropriation represents no weakened version of the Western literary tradition but a real and potent alternative, with long roots in West African oral cultures (in pluralistic and heteronomous multiplicity). Here, education implied incorporating expansive systems of rhythm, melody, and harmony.[59] Everyday life, sermons, and rituals involved real-time generation of multistructured symbolic orders of polyphonic communication. The oral domain embodied play and seriousness, spirits of the living and the dead, family, friends, and foes, and the present, past, and future. In other words, unlike the Western idea of Learned Latin, the West African education systems unfolded not in languages dissociated from everyday life but in profound continuity with it. The educational systems encouraged a sense of continuity with life and a sense of participation because the practices *were* themselves participatory.[60] The oral forms of education exemplified what Gadamer would call genuine forms of *Bildung*, that is, the ongoing and nonfixable process of "keeping oneself open to what is other—to other, more universal points of view."[61]

Finally, when jazz crystallizes itself as an art form on the American continent, it draws on the aural knowledge of West African traditions.[62] The oral tradition prevails in jazz as "ear music."[63]And, as far we can tell, one potent way to gain access to some of the oral tradition's universal aspects is by practicing the aural imitation method.

At least, this is what we will try to demonstrate. While we will set aside questions regarding cultural identity, myths, storytelling, and the master–apprentice model often associated with oral cultures, we will try to actualize the oral/aural knowledge in a contemporary educational context. In contrast to what we just called Learned Jazz, the aural imitation method is firmly rooted in everyday life communication, articulated in the horizontal and vertical implications of Berliner's assumption. At the same time, the imitation facilitates an understanding of music untranslatable into words and texts (this text included). Not only do oral traditions foster rhythm and intonation systems that typically are impossible to capture fully in writing,[64] but the aural learning process is also profoundly subjective, embodied, and musical. Just as it is inconceivable to explain to other people exactly what the perception of *red* means to you, it is impossible to explain the embodied feel of a groove[65] or how a melodic phrase sits *just right* relative to the bass and the drums. At stake is a bodily feel and a way of doing things in and with the musical language. Whereas a painter might explore the red color in a painting, the musical understanding built by aural imitation can only be expressed in music. The music is the means, medium, and goal of the

educational process. That said, the aural learning process does not exclude or compete with the competence of reading music (which often is required in contemporary professional contexts). It simply develops the musical skills along another axis, developing the ability to generate music in real time.

The following analysis has no ambition to say everything that needs to be said regarding the aural imitation method and how it can be used in educational contexts, let alone to present a full picture of developmental ontogeny or of what music *is*. But we hope to introduce a comprehensive sketch that builds up a concept of aural facilities in jazz from below, sketching out what happens in the process of launching a potential student into the complex communicative forms of aural jazz improvisation. We also hope to invoke key traits of behavior that arguably render the auditory learning process possible, from early childhood onward, and demonstrate their pertinence for contemporary aesthetic education. Let's see how it goes.

2. Incorporating the Music

In the previous section, we noted how the practice of imitating jazz recordings first emerged in an oral tradition. Experiences from NTNU confirm that oral preservation and communication is the way to deal with the method. That is, also within formal institutions, it should be taught in a dialog between teacher and student and then with a sensitivity for subjective tastes and modes of hearing. Written presentations of the practice, like the ones below, can exemplify certain critical traits. Nevertheless, they will always remain reductive as to what the method really is about, namely, to facilitate a dialogical and open-ended exploration of music and the potential of communicative musicality. The musical exploration lives in live, creative, and oral communication between listening subjects.[66]

Having said that, the process can also be carried out alone. The tradition for autodidactic learning is as old as the aural imitation method itself. Think of the young Charlie Parker submerged in the process of imitating recordings, as portrayed by Russell.[67] The music itself is the ultimate teacher. This point is crucial at NTNU: as much as the teachers admire masters of the jazz tradition, the music forms the center of the aural imitation method, not the masters themselves. The music is prime.[68]

Now, at NTNU, the process typically begins like this: the student gets the task of picking out a piece of recorded music. It can be a whole tune, but it might just be a specific part of it, say, sixty or ninety seconds. For some time ahead, this music will serve as a model and vehicle for aural training. Both the student and the teacher will explore the music in as much detail and as precisely as possible. As pointed out already, the music can be of many different styles. The method works independently of aesthetic preferences. However, experience shows that the musical example should be exemplary.

The more the music exemplifies real, communicative musicality, the more it has to teach the student. What does this mean in practice?

It means, first, that the music must exist in an audible form. From the perspective of aural training, written presentations of the tune (for instance, available in so-called fake books) will have no value. After the music is appropriately learned by ear, it can be written down. But writing should be the last thing, as a teacher at Berklee College emphasizes.[69] In other words, the aural imitation method must be completely dissociated from the general idea of music notation as a means of playing music without hearing it first. The notation is just for mnemonic purposes, in case that is needed (often, it is not).

Second, the student should have a sense of the music being either partly or generally improvised, that is, generated in real-time communication. It can be a solo performance, but if the music is generated by a minimum of three people, the student can get a clearer idea of the aural communicative dialectics going on in the musical language. More fundamentally, the record must exemplify a musical language developed and used by more than one person, notably other humans. Imitating self-made music, fully computer-generated music, or private utterances from one individual will gain little if any results. The same holds for so-called noise music. To the degree that it lacks the human communicative dimension associable with an orally transmitted musical language,[70] noise music is unusable as a model for learning precisely that form of communication.

Third, the music should be stringent, coherent, and articulate. It should exemplify a musical language in a relatively clear-cut way, rather than being vague or indeterminate.

Fourth, and perhaps more controversial, the music should preferably exemplify *heard* music, that is, music generated "purely" in and through aural musical communication. In addition to what we just called Learned Jazz, much jazz improvisation relies on visual, social, conceptual, or ecological protocols (perhaps also as a response to the values of literacy?).[71] Without denying the validity of this music, experience shows that the student will potentially learn other qualitative nuances by imitating music generated and transmitted by ear. Aural music embeds knowledge of how to manipulate, remember, and communicate in auditory and rhythmic means.[72]

Fifth, and this will soon bring us back to Malloch and Trevarthen's concept of communicative musicality, the music should spark a genuine interest, desire, and curiosity in the student. It should exemplify a musical language that the student would want to acquire for him/herself, independently of social norms and expectations. What this means in practice can always be problematized. But one pedagogic rationale of letting the student choose the music is the signal of responsibility. The teacher is there to guide, not to teach in a normative sense. More importantly, the free choice stimulates the

sheer joy of learning. Experience proves that interest, desire, and curiosity generate a playful listening attitude toward the music, which in itself generates more learning. Learning comes indirectly as the result of an autonomous, playful process. To put it in a quasi-formula

Desire ↔ Listening ↔ Play or Playfulness = Learning.

While the importance of play is well known in the pedagogic and philosophical literature,[73] the teachers at NTNU explicitly take the principal role of desire and play from the informal African American jazz tradition. Following Sidran,[74] it seems reasonable to say that, when the aural imitation method first emerged, the inventors probably did not conceive it as a method. The practice was "just" a way that young people followed their drives and curiosity for music. Young people started to imitate their musical heroes because they liked the music. They wanted the expressive powers of their heroes: their sound, melodic and harmonic sense, technical skill, or whatever. This "just for fun" principle is what NTNU tries to preserve in a cultural era otherwise dominated by standardized goals. In contrast to what seems to be the case at Berklee College,[75] there are no specific aesthetic ideals to be reached at NTNU. There is no curriculum, no definitive corpus defined by the institution, that the students have to acquire to become educated. Instead, everything is allowed to hinge on the student's interest, motivation, and curiosity created by the music. Thus, in effect, the aural imitation method practiced at NTNU distinguishes itself not only from ideas handed over by the Western literary tradition but also from the master–apprentice model often in play in oral cultures. In a nondirective and open sense, the institution is there to guide and support the student into a process ultimately driven by the student's own desire and playful attitude.

Interestingly, while this idea seems somewhat radical compared to other contemporary studies on jazz education,[76] it makes sense in light of Malloch and Trevarthen's communicative musicality: humans often imitate *just for fun*.[77] The inborn drive for imitation is carried out not primarily to solve problems but for the mere fun of mastering new challenges, the pleasures of showing off, for the joy of being with others, or just for the fun of acquiring a new mode of expression. From this perspective, the institutionalized desire and play principle seems to work with—not against—the inborn musical drive for imitation, synchronization, and cooperation. The same drive that once propelled the student into a linguistic community is allowed and stimulated to move him/her further into musical discovery.

Now we see why only music generated and articulated by real others would work as an example. Only music used in real communicative contexts sets the talent for communicative musicality genuinely into motion. For, as Galper puts it, "When you are attracted to a particular artist, there is a reason for it. There's a resonance that has been created between you and

that artist. You respond emotionally and psychologically to that music."[78] But we also see why the music example should be stringent and articulate. For just as a reasonably coherent structure of a language is needed for the child to keep the drive into the development of linguistic abilities, so does stringent music generate more audible learning. Alternatively, as Kant and Gadamer would say,[79] only an articulate and exemplary product of another human's communicative abilities can substantially initiate a reschematization of one's cognitive powers; only stringent products of others would tease, animate, and push the will, need, and facility to hear and generate more meaningful relations, outside the habitual loops of private subjectivity.

2.1. Embodiment

Merleau-Ponty reminds us of a crucial fact: when we talk about embodied activities, we always talk about a real human subject that lives and walks the earth along with other living bodies. To be an embodied subject is not to be a subject situated in a body but to be a phenomenal body that is its own body: "We never move our objective body, we move our phenomenal body."[80]

For the sake of further concretization, let us imagine a female student who has picked out Cannonball Adderley's "Autumn Leaves."[81] She does so mainly because she likes Adderley's funky swing, timbre, and tone; Miles Davis's crisp phrasing; and the transparent sound of the band but also because she wants to develop her aural understanding of tonality and harmony. This does not necessarily mean that she will play in Adderley's idiom for the rest of her life, but for the time being, there is resonance, drive, and curiosity.

The music satisfies the conditions outlined above. It exemplifies the use of oral communicative language. The beat and syncopated flow of rhythm illustrate the immensely complex African American rhythmic sensitivity, whose history predates the African diaspora.[82] The tonal language exemplifies the use of the Western tonal system whose fuller account goes back to ancient Greece[83] but also synergies between African tonality and Christian hymns occurring during the periods of slavery[84] and the standardization of the equal temperament emerging in 1917.[85] It also exemplifies the African American collective, aural, polyphone, and spontaneous musical interaction, whose tradition started centuries before American slavery.[86] And, finally, it exemplifies the personal and collective styles or sounds of the musicians.[87] It illustrates the unique and articulate sound emerging when Adderley, Miles Davis, Hank Jones, Sam Jones, and Art Blakey met and attuned to each other within the shared medium of auditory communication.

Thus, Adderley's music makes a perfect example. It concretizes, articulates, and exemplifies the use of a specific musical language unfolding in that precise context—everything in medium tempo and transparent sound.

This concretization, we will come to see, enables a general potential outcome of imitative learning.

The exercises presented below are specific to the imitation of music of Adderley's kind, most notably in its focus on harmonic learning. However, a student could have picked out music from entirely different traditions, say, Indian raga music, Chinese traditional music, or Norwegian folk music, which also follow the oral tradition.[88] If so, the designs of the aural imitation method would have to be adjusted to fit each of the respective musical languages and the cultural background of the student.

Now, with Adderley's music at hand, the student is guided into a process that somewhat artificially can be divided into three stages, as suggested by Inderberg:[89] "embodiment" (Table 1); "tonal and harmonic orders and the use of the instrument" (Table 2); and "subjective (and personal) variation" (Table 3). These stages represent the gradual transition from something as close to one-to-one imitation as possible to more symbolic, flexible, and general musical behavior—or more personal expressive freedom rendered possible by the musical language.

2.1.A. CHALLENGING EXERCISES

Presented in Tables 1–3 are practical instructions. Strictly speaking, the instructions are worthless unless transformed into real use and musical exploration. Now, there is no way of getting around the fact that, for most people, putting the instructions into use is challenging and time-consuming. By itself, this fact calls for the dimension of play, a feeling of fun, curiosity, and intrinsic motivation. To respond adequately to the challenges and get something out of the tasks, the student literally needs the exploratory drive of musicality. Without the joy of musical exploration, she is likely to begin "cheating" at some point, or just give up.

For the sake of description, let us outline typified if not idealized behavioral correlates suggested by the exercises. More precisely, this section aims to invoke pertinent theoretical perspectives suggested by Table 1. (Tables 2 and 3 will be analyzed in Part 2 of the current article.)

2.1.B. AUDIBLE MODE OF ATTENTION

As the title of Table 1 suggests, the first and foundational stage of the imitative process anchors the musical learning in the body schema. The artistic freedom potentially facilitated by the exercises hinges on the security of embodied integration.

The student directs her attention toward the music *as heard*. She listens to the music in a processual sense, that is, as the music flows and emerges from moment to moment. The aural directedness is radical. Not at any point does the exercise encourage an analytical attitude. In sharp contrast to what

Table 1. Phase 1: Embodiment

Step	Description	Tool
Act 1	a. Listen to the entire recording with imitating the rhythmic wholeness in mind. b. Focus on the predominant rhythmic elements. Listen to individual instruments and divide the rhythmic pattern into sub-elements. Use hands, feet, and voice to express these rhythmic sub-elements.	Singing, clapping, foot-tapping
Act 2	Listen to the melodic material in the recording. Play the rhythm of the melody lines with your hands, voice, or feet.	Singing, clapping, foot-tapping
Act 3	Beat/tap the rhythmic foundation from Act 1b while singing the rhythm of the melody.	Singing, clapping, foot-tapping
Act 4	Listen to the recording's melodic material again. Sing motif, theme, and longer melodic phrases. Find the register that suits you. Change octaves as needed. At the same time, try to imitate vibrato, timbre, and other musical characteristics. If there is a text, learn this or some of it.	Singing, clapping, foot-tapping
Act 5	Sing melodic lines while beating/tapping the rhythmic wholeness that you found in Act 1b. Sing the melody lines while beating/tapping the rhythm of other instruments.	Singing, clapping, foot-tapping
Act 6	Focus on the bass instruments in the main parts of the composition. Sing the melodic line of the bass instruments. Sing these bass lines while letting the main melody at this point in the composition sound in your ear.	Singing, clapping, foot-tapping
Act 7	a. Listen to key parts of the harmony unfolding in the recording. Sing the root of each chord. When you listen to the harmonic material, you can hear chords with, e.g., five or three tones or chords with no root. However, you should aim for its simplest representation (seventh chords or triads in root position). b. Sing the chord tones from root up to and including the seventh. Now sing all the chords in longer sequences the way you hear them on the recording: in the right inversion and with any added tones. Find a tempo that suits you; start without a fixed rhythmic pulse. Make audio recordings of yourself.	Singing, clapping, foot-tapping

Table 2. Phase 2: Tonal and harmonic orders and the use of the instrument

Step	Description	Tool
Instr. 1	Repeat the exercises in Act 1a and b. Play the individual rhythmic elements and optionally the rhythmic whole that can be performed naturally on your instrument. Use feet and, if possible, voice to complement your playing. Let the basic rhythm of the recording that you imitated in Act 1b "sound" in your ear. After a few minutes of practicing this, emphasize parts of this rhythm on the instrument.	Instrument, singing, clapping, foot-tapping
Instr. 2	Listen to the core melodic material of the composition/recording. Play the melody rhythm while "listening" to your perceived version of the melody line. Select one or more random tones. Repeat Act 4. Listen to your inner hearing while playing the same melodic material on your instrument in multiple registers. Play with different levels of dynamic, timbre, and tempo.	Instrument, singing, clapping, foot-tapping
Instr. 3 and Instr. 4	Listen to your inner hearing's version of the rhythmic basis of the recording. Play the melody lines from Act 4 simultaneously. Play as if you were singing in the instrument. Play the same melodic material while allowing individual rhythmic elements to be expressed using feet (or voice). Vary dynamic, timbre, and tempo. Remember to make a recording of your playing. Is there a match between your hearing and playing on the recording? Percussionists who do not play a melodic percussion instrument sing the exercises in Instr. 3, Instr. 4, and Instr. 5.	Instrument, singing, clapping, foot-tapping
Instr. 5	The tone that represents the center of the chord (root) is played (see Act 7). One note in each chord throughout the recording should match the root tone in your hearing. Play the four lower notes in all the chords from the root to the seventh. Then arpeggiate the chords up and down in a sequence that you spontaneously decide. Use the full range of the instrument (see Act 7). Make music from this exercise. No exercise should be just technical.	Instrument, singing, clapping, foot-tapping
Instr. 6	Practice your perception of period by, e.g., playing arpeggiated chords (see Act 7b), main melodic lines (see Act 4) or melodic bass lines (see Act 6). Keep the recording's main rhythm going in feet or voice. Make recordings so you can check if you are keeping the periods. Play diverse sub-elements from this table with varying intensity, timbre, register, and dynamics.	Instrument, singing, clapping, foot-tapping

Table 3. Phase 3: Subjective (personal) variation

Step	Description	Tool
Pers. 1	Let the basic rhythmic foundation (Act 1b) sound in your ears and sing, play, and beat new rhythmic ideas/patterns simultaneously. This must occur spontaneously. Vary the length of ideas from short motifs to longer themes. In this exercise, play harmonically and melodically freely. Make recordings of your own playing. What do you like and dislike?	Instrument, singing, clapping, foot-tapping
Pers. 2	Pick some melodic motifs from the recording. Sing this melody line along with the recording so many times that you know it by heart. Play it together with the recording. Play the same thing alone while the rest of the recording sounds in your ear. Sing and then play spontaneous melodic lines. Be sure to convey ideas from your own musical imagination. Record your playing along with the recording as well as your soloing. What do you think of your own playing? What do you want to change? Percussionists: Sing exercises in Pers. 2.	Instrument, singing, clapping, foot-tapping
Pers. 3	Concentrate on the harmonic progression of the recording. Sing the chord sequence you hear on the recording. Break the same chords in arpeggio exercises on your instrument. Let the rhythmic foundation of the recording go into your ear while you spontaneously play new chords. Preferably, piano is used, but singers and melodic instrumentalists sing/play broken chords in inversions of their own choice. Set your harmonic imagination free.	Instrument, singing, clapping, foot-tapping

is often suggested in the pedagogic literature on jazz improvisation,[90] the point of the exercises is *not* to identify rhythmic measures as such (say, 4/4 or 3/4), names of tone relations ("there's a third up; there's a sixth down"), scales, or standardized organizational forms (such as AABA form, or 12-bar blues form). The point is not to catch the conceptual rules unfolding in the music, that is, to draw out, schematize, and categorize the auditory flow into fixed or semi-fixed metastructures that can be conceptualized independent of the musical flow. From the aural perspective, these structures are secondary products, supervening on what goes on in the perceptual domain.[91]

In short, the sole point of the exercises is *just* to listen very carefully to the music as an audible phenomenon, and replicate through enactive effort the audible phenomenon in singing, stamping, clapping, or sound produced in mind (Act 6). While some cognitive reflection is required to perform the

tasks, the exercises themselves encourage no reflective grip as such, only aural attentive awareness. The mode attention does not evolve on the cognitive level of "I think that. . . ." Instead, it tries only to stick imitatively to the music as the student hears it, comprehensively, and with the moving body proper. Alternatively, as Trevarthen would say,[92] the music will be understood through the atheoretical yet communicative potential of the body. Something in the immediate, unrational, unverbalized, conceptless, totally atheoretical potential of consciousness is set in motion.

2.1.C. FULL MATCH—NO APPROXIMATION

The imitative attention called for is meticulously exact. Full match between the music *as heard* and the music *as imitated* is the ideal—at least in a guiding or regulative sense. The student is encouraged to be aware, imitate, and map out, by replication, minute auditory distinctions of phrasing and articulation (Acts 1–3). She tries her best to enact the accents, stress, and timing, down to the exact timespan of tones, the exact timespan *between* tones, or the auditory quality of vibrato and timbre (Act 5).

It is worth noting how the stringent attitude toward the music by itself is a norm developed in the oral jazz tradition. It is no coincidence that Wilf's study from Berklee College reports the same attitude.[93] For as Bill Evans articulates, there's no approximation in jazz: "If you try to approximate something that is very advanced and don't know what you're doing, you can't advance."[94] Wynton Marsalis hits the same point in a way that also captures the exemplary character of the music imitated by our student: "Jazz is not just, 'Well, man, this is what I feel like playing.' It's a very structured thing that comes from tradition and requires a lot of thought and study."[95] The exact character of the music demands an exact correlate of imitative behavior. Otherwise, there will be no substantial contact with the musical sense.

In other words, according to norms developed in the jazz tradition together with the aural imitation method, the student just *has* to be precise to hear the advanced musical organization unfolding in Adderley's specific use of rhythm, tonality, style, and joint musical communication.

2.1.D. RHYTHM IS PRIME

The student approaches the music through its overall rhythmic organization (Act 1). Rhythm is prime, both in regard to the timeline (it comes first) and by being the main structuring aspect. Notably, in Act 1b, we see how the identification of individual instruments is subordinate to the general rhythmic pattern they form together. The rhythmic structure of melodies and chords is prioritized. Melodies are *first* imitated as rhythmic patterns (Acts 2 and 3) before the student "fills in" the tones of the melodies (Act 4).

In effect, the student now listens very carefully to how the Adderley Quintet swings, by trying to map out exactly how everything relates to

everything. Each accent, attack, and length of tone contributes to the formation of the dynamic, collective beat of the band. Everything belongs to or parses into coherent rhythmic wholes. Special stretching of the beat, for example, in Adderley's and Davis's characteristic styles of phrasings emerges *in* the reciprocal dynamic of the collective.

Again, it is worth noting how the rhythmic approach is entirely in line with the evaluations of musicians. Whereas jazz musicians can differ substantially in *how* they swing or groove, the rhythmic organization as such is generally regarded as fundamental, that is, independently of jazz idiom.[96] The logic of rhythm is *the* organizational logic, as drummer Ralph Peterson puts it:

> [I]f you miss a note and the *rhythm* is logical, then the idea comes across . . . whether you hit the note dead center or not. But if you miss time—because music is organized sound in time . . . if you blow the time you're more likely to do irreparable damage to the music.[97]

Roholt's Merleau-Ponty-inspired phenomenology of rhythm reflects the same fundamental point: "In genres such as jazz, hip-hop, pop, and rock, a groove is the glue that holds together a recording or performance, a central element around which musicians coalesce."[98] Roholt brings to the fore several other observations that seem to be relevant to our context, first of all the fact that the coalescing force of a beat is a normative, effective, and emergent phenomenon. The beat is a dynamic and seemingly self-generating gestalt, "bouncing off" as if it created itself. Philosopher Scruton identifies a similar point: "When we hear music, we do not hear sound only; we hear something *in* the sound, something which moves with a force of its own."[99] The beat emerges as an effective and intrinsic push that moves the music according to its organization.

Moreover, Roholt demonstrates how the beat, from the perspective of auditory perception, is a qualitative phenomenon. The normative force *is* only in the ways that the music unfolds, that is, in the qualitative contours of its auditory and temporal flow. Between two seemingly (for an untrained ear) identical series of notes, there can be all the difference in the world, due to how one swings and the other just doesn't. And the whole difference can unfold in exactly how the series progresses, that is, in their styles and manners of unfoldment. As Roholt demonstrates, the swing unfolds in minute expressive details and variations in the musical flow. The functional role of one instrument or part of the instrument fits just right into the auditory context from moment to moment. That said, there is no way to fixate how to swing. Unless one takes an analytical attitude (which is something Roholt refrains from doing, just as our student does), the beat *is* only in the ways that it "bounces off," "pushes," "pulls," "leans forward," is "laid-back," or "in the pocket," and so on. And it can only be

grasped as perceptual wholes, not, as Roholt demonstrates in a critique of Iyer, in terms of milliseconds.[100]

What else characterizes the beat as wholes? In a positive and generative sense, the beat is a *perceptual indeterminate*, Roholt demonstrates.[101] The beat is a normative, objective, and real phenomenon available for perceiving subjects, always open to being perceived differently. Two musicians playing together can hear and respond to the same beat while not responding to the beat in the same manner. Indeed, to perceive, understand, and respond adequately to rhythmic nuances embedded in the variations is a defining mark of the aural skill of professional jazz improvisation.[102] The aural rhythmic facility implies not just a reasonably steady pulse but also the ability to "absorb a large amount of rhythmic variability without being thrown," as musician Ralph Peterson remarks.[103] The beat is like an elliptical perceptual force around which the musicians negotiate.[104]

Hence, against this background, we could say that this is what the student is about to explore by imitating the music: she tries to catch the normative organizational force in Adderley's music. She tries to grasp and replicate the rhythmic logic embedded in the quality of the auditory flow. Roholt's points add on the nonanalytical attitude we indicated a minute ago: in trying to catch the rhythmic sense, the point is less to hear *what* the musicians play than *how* they play it. The precise aural attitude is the ultimate condition to catch the pertinent qualitative nuances.

At the same time, profound ambiguity distinguishes a student's mode of precision. Regardless of how closely the student replicates the music in her behavior, the beat of Adderley's combo will always remain a perceptual indeterminate. The beat will stay an audible and generative X. It will be the glue that holds everything together and propels the music forward without being reducible to anything but precisely this organizational and moving audible force. Better put, while the student digs into more and more qualitative nuances, she will uncover for herself ever new and potent perceptual indeterminacies. There will always be *more* to learn from the rhythmic logic of Adderley's music.

2.1.E. MUSICAL WHOLES, GENERATIVE POTENTIALS, ENACTION

When the student has worked with the overall rhythmic form for some time (Act 1), she brings in features belonging also to melody and harmony (Act 2 ff). Crucially, from now on, she approaches no specific feature of the music in isolation.

The music is not approached as if it consisted of discrete elements or linear sequences coupled, for instance, in the forms of rhythm plus melody plus harmony. Instead, it is always a matter of hearing musical wholes, that is, synergetic and dynamic configurations of rhythm, melody, and harmony. To borrow Evan Thompson's description of dynamic coemergence, the

musical whole arises from its parts, but the parts also arise from the whole: "Part and whole co-emerge and mutually specify each other."[105]

The rhythmic, melodic, or harmonic features can either be imitated together in the direct sense or played out against each other (Act 5). Either way, the music is approached as unfolding wholes. Even when the student turns explicitly to the harmonic system (Act 7), she does so not by isolating individual tones and sequences but by seeking to catch intrasystematic relations. The point of singing bass lines, thirds, fifths, and sevenths in the right voicings is not to identify individual sounds or sequences but to hear relations between sounds and sequences.

Melody and harmony unfold in the logic of rhythm but also imply organizational forces of their own, given the major and minor tonality. How the norms of the tonal language are conceived in our usage-based approach will be the main subject later in this article. For now, however, we note that the tonal language also brings in a dimension of generative potentials embedded in the flow of sound. To borrow a phrase from Scruton, for descriptive purposes, "[W]hen we understand a piece of tonal music, it is because we have grasped the tonal [and rhythmic] order which generates the musical surface."[106] In the tonal organization, something drives the music in specific, nonaccidental directions. Qualitative forces in the music gather form and dissolve, seemingly generating its course of unfolding.

Many theorists think it is a good idea to approach the generative potentials of Western tonality in terms of well-formedness, preference rules, and the like.[107] These descriptions are far too formal and intellectual for our usage-based approach. From the vantage point of imitative and communicative hearing, the generative potentials of music are just perceptual phenomena. The musical orders are hearable, understandable, and communicable—in sonorous communication only. Regulated by the norm of rhythm and tonality, the acoustic music has a perceptual thickness, as Merleau-Ponty would say[108]—a perceptual excess embedded in how the rhythmic and tonal language sketches out its sense.

In calling on Merleau-Ponty's perceptual thickness, we are close to what Schiavio and others outline as sonic environment.[109] Paraphrased coarsely, these theorists demonstrate how music is an enactive phenomenon. Both the ability to listen to music and the ability to produce it in the first place are enactive skills. Hearing music is inseparable from the subject's readiness to act. The music unfolds as a sonic intermediate reality whose organization is revealed and acted out by the skilled perceiver. In joint musical practices such as a jazz band, members participate in the collective sense-making. They co-enact the sense, as Schiavio and Høffding phrase it: "[E]ach musician *brings forth* a domain of meaning, co-creating the sound environment in which he or she is embedded."[110]

To a considerable extent, this perspective sits well with the suggestions of the aural imitation method. By replicating the audible, rhythmic, and tonal orders, the student is about to explore a sonic environment by enacting it. She tries out its sonic possibilities and constraints—how beat, melody, and harmony offer and limit specific types of action. The student can enact the musical sense, or train to do so, because the musical sense exists as a sonic field of possible enaction. Call this the *sonic environment approach*.

However, while the sonic environment approach positively captures some aspects of the imitative process, Merleau-Ponty's concept of perceptual thickness takes us further into the generative dimension of musical exploration. Although talking about music as perceptual thickness also implies the ability to enact the music qua sonic environment, the enacted environment is not just sonic; rather, it is a symbolic communicative aural language. Following Merleau-Ponty,[111] the music streaming out of the student's loudspeaker unfolds audible symbolic matrices. The combination of African American rhythmic sensibility, African American modes of joint aural music-making, Western tonality, and collective and individual styles reveals complex networks of qualitative forces that regulate, distribute, and generate themselves along with stringent and genuinely open systems of equivalences. This, we could say, is what Evans[112] and Marsalis[113] indicate when they describe their music as *advanced* and *very structured things*: everything in the music makes communicative sense. Everything in the music unfolds thick auditory, generative, and stringent perceptual matrices used in live and polyphonic communication.

Now to describe this complex auditory and communicative network as a "mere" sonic environment would be too neutral and passive. We need a thicker concept of the musical organization emerging from and for the knowable aural instance. First, the significance of the music emerges between subjects that always, and by necessity, perceive a little differently. As humans always perceive from their perspective,[114] a productive ambiguity will always permeate aural communication in musical forms. The music can always be heard differently. Second, the symbolic matrices that unfold in the music are not passive environmental features but semi-active, semi-autonomous perceptual realities organized by the plastic norms of the symbolic matrices. The enacted music unfolds rhythmic and harmonic generative potentials, that is, audible invariants that unfold in and through the constant variations of the music. These potentials are not just acted out but enact back, so to speak. Analogous to how verbal language can attain a "life of its own" that shapes and propels the behavior of the perceiver,[115] the musical potentials generate their course of progress. Thus, the music is not a neutral and passive sonic in-between, but a medium wherein *something* is communicated with the possible risk of communicative breakdown.

2.1.F. HOLISTIC AND CROSS-MODAL ENACTION

Training the facility to catch-and-enact the qualitative nuances of the music engages the whole body. Hearing is not just hearing with the ears but with the hands, feet, voice, and *mental* voice (Act 6). The rest of the body is there, too, of course, as the anonymous background of what Merleau-Ponty calls the "body proper." The body is there as a functional whole, with an embodied mind, with a fairly established body schema for possible movements, and with the life and history of an individual.[116]

We note how the forms of enaction stimulated by the exercises are genuinely cross-modal, cross-synchronal, and thoroughly determined by qualitative standards. The student is encouraged not only to sing, clap, stomp, and hear music in mind but to enact the behavioral forms in parallel. More precisely, the multiple enactive forms are carried out not only in simultaneous orders but also with and against each other. See, for instance, Act 3. Here, the student stamps the rhythm of the whole music while singing the specific rhythm of the melody. In Act 6, she sings the bass, while the intramental "inner ear" perceives the theme. In other words, the student maps out the rhythmic, melodic, and tonal features of the music in qualitative and cross-modal behavior. Besides, as we recall that the beat and the tonal generative potentials unfold primarily as qualitative features, we note how the current behavioral forms are determined less by *what* is enacted than *how* everything is enacted.

In effect, the whole is about to be enacted by the whole functional body. While prioritizing the auditory directedness of attention, the auditory access to the music is, by itself, comprehensive or holistic. The auditory domain is extended into all modalities and across modalities. The music as a whole perceptual sense is replicated and taken up, so to speak, by the entire phenomenal body.

The pivotal role of bodily aural understanding is well known among aural musicians. "[T]he intuition and the ears often 'know' more than the intellect," Galper points out.[117] It's the aural body that understands and plays. Fred Hersch states the same: "In order to swing, not just to approximate swing, the rhythm has to come from the body. . . . [I]t's the movement of the body that inspires you to play."[118] Roholt is on the same track. The gluing force of rhythm is only conceivable through the use of the whole body: "feeling a groove *just is* to 'get' it."[119] The *feel* in question is not this or that sentiment but the feel of an affective moving body that potentially *gets it*: "This grasping involves listening (of course) but it also involves a kind of active, practical, non-theoretical knowing. We come to understand grooves *by moving*."[120]

In our context, Roholt seems to hit the core purpose of the multimodal exercises described above, except for the fact that we need to apply his descriptions to the understanding of the whole music (not just rhythm). The cross-modal practices seem to use and stimulate the noncognitive,

motor-intentional understanding. The musical immediacy encouraged by the exercises implies a practical and affective mode of intentionality. By gluing attention to minute details in the music and trying to replicate the complex behavioral forms, the student seeks to incorporate the musical sense into her body schema of possible conduct.

In effect, we are yet again brought back to Trevarthen's concept of communicative musicality. What the musicians and Roholt describe in ways pertinent to the aural imitation method seems to be modifications of the immediate, unrational, unverbalized, conceptless, and atheoretical potential associated with musicality.[121] With this in mind, we turn to the ontogenetic formation.

3. Primary Imitation

The previous section observed how Table 1 suggests structures of an embodied aural attention. This section will first articulate structural and horizontal similarities between the aural imitation method and first-language acquisition. Then, work from developmental psychologists will help us see how vital traits of early childhood can be said to prepare for the learning process about to take place in the student.

Now faced with the advanced structures of Adderley's music, the imperatives of exact imitation, and difficulties of enacting the music cross-modally, our student has deliberately set herself in a position structurally similar to the situation of a helpless child. The student lets herself be overthrown by the other's musical understanding, as Gadamer would say.[122] To catch up on the advanced sense unfolded by the competent language users (that is, the musicians), the student has to grow her communicative abilities—just like a child who has to grow into the complexities of the linguistic community.

To develop the structural similarity further, we invoke some critical characteristics of prelinguistic imitation. What characterizes the imitative attitude of an infant? According to Stern, "Primary consciousness is the yoking together, in a present moment, of the intentional object and the vital background input from the body."[123] In the yoking together, the infant glues itself affectively and entirely to the parent's or the caregiver's expressive behavior, with no clear sense of self vis-à-vis the other and no meta-awareness of what they are doing. As Stern points out, "Infants cannot know *what* they do not know, nor *that* they do not know."[124] The infants live the intersubjective orders bodily and affectively, fully submerged *in* their phenomenal bodies, and the contact with the mother, father, or care provider (henceforth, mother). They are embedded in the shared dynamics emerging around events and things, that is, in the shared perceptual experiences.[125]

The yoking together is meticulously exact—at least as a regulative ideal. The child strives for the full match, and already by the age of two weeks,

it can imitate structures of expressive behavior with high precision, coherence, and sensitivity for minute variances when awake and feeling safe.[126] However, as pointed out by Fuchs and De Jaegher, the imitative interaction between child and mother is also a "messy" process of affective "matches," "mismatches," and "interactive errors."[127] Miscommunication is normal. Typically, infants and mothers match only 30 percent of the time of normal play.[128] The mismatches become a productive force in the imitative interaction. It generates an improvisational process aligned to "repair" the mismatch and bring the situation to rest. According to Fuchs and De Jaegher, "Reparation becomes a key process, as it conveys the experience that a miscommunication ends up in understanding . . . again."[129] Repeated experience of successful repair potentially affects the infant's sense of agency, trust in others, and bonding capacity.

Conceived by Trevarthen,[130] the affective match-and-mismatch process typically unfolds in rhythmic behavior. Rhythm is prime—not only for jazz musicians but in human life as such. Even in their mother's womb, infants learn how life comes to expression in temporal, rhythmic patterns—in forms of vitality.[131] (We return to the forms of vitality in a moment.)

What Trevarthen calls musical proto-conversation sets in a few weeks after birth. The mother sets the tone. The child is not communicative unless appropriate invitations are given by the mother. The mother invites the child by singing simple songs, moving, and dancing gently along with the baby or by breathing in relaxed ways, repeating short, evenly spaced words with a resonant yet relaxed "breathy" and moderately high-pitched voice.[132]

Rhythm does not only come from the mother, however. The baby's mind is not a mere receiver of time. It is a generator of time, Trevarthen stresses.[133] Newborns have superb timing capacity. They modulate their behavior vis-à-vis the mother in what Trevarthen describes as "fine and rapid . . . glides and leaps of pitch or volume of voice, eye-brow flashes, prebeat syllables, suffix morphemes, rhythmic details and embellishments, rapid hand gestures, quick head moves, shifts of gaze."[134] The temporal contours are also embedded in the "fundamental beat of repeating movement, short bursts of expression, repetition of rhythmic groups of movement, exaggerated dynamic expressive 'sentic' forms, and precise modulation of the intensity or force of expression in a moderate to weak range."[135] Newborns can discriminate rhythmic sequences independent of tempo, and they are largely sensitive to tempo variations. They are able to detect split-second deviations from simultaneousness, showing that they, to some extent, prefer contingent variations.[136]

Curiously, Trevarthen documents how the two-way imitational play between child and mother forms into stringent melodic forms of *introduction*, *development*, *climax and resolution*, and *coda*.[137] The proto-conversational play has the form of a narration without semantic content. It is the proto-version

of call-and-response interaction. With Benjamin,[138] we can say that intermediary thirds are formed. Between the partners, there emerge dynamic, perceptual realities that propel behavior in both partners.

We now return to the aural imitation method. In light of the developmental perspectives, it makes sense to say that, when the student, as a mature being, volitionally glues herself to the music, she takes up an attitude that is structurally similar to "non-knowing" of the infant. She uses an aural mode of attention: an open, nondirective, and flexible mode of awareness that by itself likens the infant yoking together with the expressive behavior of the mother. The student imitates the child's intuitive readiness to move rhythmically with others and ways that the infant expresses and receives sympathetic awareness through imitative enactment.[139] In other words, she methodologically practices the nonmethodological attitude of the child's yoking together. Rather than *thinking* the music, she tries to be *in* the music as a perceptual sense. She tries to "yoke" into the auditory and qualitative sense of the music.

Moreover, the student glues herself to the rhythmic forms of the music, just like the infant who glues itself affectively and comprehensively to the mother's voice, gaze, and expressive movements. Where the child imitates the timespan and qualities of the mother's behavior, the student imitates attacks, lengths of the tones, and the timespan between tones, in minute differentiations between simultaneousness and split-second deviations. In other words, just like the infant, the student takes up an attitude of communicative musicality. She tries to attune herself fully to or in the temporal organization of the perceptual object, just like the child who imitates the energetic ebbs and flows of the shared events.

But are we not missing something here? Trevarthen's musical communication is two-way. The two-way performance of rhythm emerges when the mother attunes to the rhythm of the child, by mirroring, inclusive and caring behavior, and the child attunes to this responsive behavior. By contrast, when the student listens to the recordings, there is (obviously) no response from the recorded musicians. In this sense, the attunement is not two-way and dialogical but one-way. How, then, can we compare the two modes of affective attunement? Does it make sense at all?

Yes. The comparison makes sense because, if we think of it again, we see that the encounter between the recorded music and the student *is* dialogical, just as the reading of a text is dialogical in Gadamer's philosophical hermeneutics.[140] Responding to the various calls of the music, the student is in dialogue not with the musicians in a direct sense but in a mediated sense, that is, in dialogue with their means of communication, namely, the music. The supreme quality of the music "forces" the student into an open-ended process of "repair." In the productive process, she has begun seeking the ultimate match. The ultimate match might never come, but she learns while

trying. Digging into more nuances in the complex and ambiguous musical sense, she has to revise how she can catch-and-enact the music.

3.1. Forms of Vitality, Holistic Awareness, and Multimodal Behavior

We have now begun indicating structural similarities between the mode of attention suggested by the aural imitation method and the imitative behavior of infants. It is essential to keep in mind, however, that, by juxtaposing the two modes of imitative behavior, we are not comparing behavioral forms that are external to each other. The fact that every music student was once an infant suggests that the infant imitation is structurally and temporally prior to the more mature form of imitation. The rhythmic interaction described by Trevarthen has already taken place, instituting a fundamental rhythmic understanding in the student's body schema, as Merleau-Ponty would say.[141] Any subsequent rhythmic learning supervenes on the first experiences of the match, mismatch, and rhythm. In other words, when the student at some point in life starts to imitate the rhythmic forms of Adderley's music, she already knows how to perceive rhythmic structures—how to parse reality "into groupings that have some kind of coherence," as Stern would say.[142] She possesses an encultured, implicit, relational competence embedded in her imitative capacity. She can already follow and unfold the perceptual sense of the ever-deferred match with another's communicative expression.

In this subsection and the next, we will expand the picture by illuminating how the student's potential for global attention, multimodal behavior, and atheoretical understanding of the music also originated in primary imitation. To do so, we turn briefly to Stern's concept of forms of vitality. (Forms of vitality are also part of Merleau-Ponty's analysis of body schema;[143] but to avoid unnecessary complexity, we leave Merleau-Ponty's descriptions out in the current context.)

For the infant, qualitative contours of temporal unfoldment (for example, rhythmic forms) are organizing forces not only vis-à-vis the mother but in the perception of life in general: "Infants act as though two events sharing the same temporal structure belong together."[144] Temporal structures have distinguished perceptual sense, defined by how they begin, flow through, and end. The qualities of how this happens might very well group events that, from an adult perspective, can seem different. That is, the infant can group events that share rhythmic patterns in a direct sense (say, the same song used in two distinct contexts). But in a general sense, the events group together due to the modes and manners of shared qualities such as *fleeting*, *wavering*, *accelerating*, *explosive*, *fading*, *pulsating*, *slow*, *easy*, *tense*, *forceful*, or by manifesting *crescendos* or *decrescendos*.[145] These forms can unfold in minuscule timespans—like in the stress put on one word, in the time between inhaling and exhaling, or in the quick raising of an eyebrow.

Minute as they may be, these forms of behavior are, according to Stern, anything but trivial. Potentially, the forms express the modes and manners of life. They are the accents of what Trevarthen calls communicative musicality. In very concrete and potential ways, the stress of the one word can exemplify the rhythm of how communal life is lived in *that* precise context. Long before verbal language sets in, the infant learns to perceive the general affective qualities embedded in the temporal contours of the event.

However, forms of vitality pertain not only to musicality in the general sense but also to the holistic attention invoked by the aural imitation method. According to Stern, the infant perceives vitality through genuinely holistic attention or awareness. The whole body is involved, as when a rush of joy or fear flows through the body or when gentle dancing movements of the mother sets in. Lacking clear distinctions between intrabodily and extrabodily perception, perceiving vitality is for the infant neatly associated with proprioception. Perceiving something out there, for instance, in the tone of voice in the mother, is coextensive with an intuitive self-perception of one's own body. But the proprioceptive awareness does not stop where the physical body stops; it extends into the face and expressions of the mother. In other words, the proprioceptive awareness of vitality forms is profoundly subjective *and* intersubjective at the same time.

Also relevant to the exercises described above is the fact that the child's awareness of vitality is genuinely multimodal. The child experiences the forms directly as global, amodal perceptual qualities. Perhaps better put, the forms *are* the fluidity between various kinds of sense experience. The perceptual grouping of events is not tied to one sense modality at the time but involves the whole body and all modalities at the same time. Properties of duration, beat, and rhythm are readily perceived in all modalities, and the infants are experts in transferring the properties of events across sense modalities.[146] Infants recognize or translate temporal patterns experienced in one sensor domain, say, hearing, into haptic or visual forms, or vice versa. And they immediately "know" how information from one modality corresponds to other sense modalities.

In agreement with Trevarthen's notion of inborn rhythmic abilities, Stern holds that the ability to perform cross-modal translations are neither learned nor constructed. Infants are born with an intuitive and general cross-modal fluency: "No learning is needed initially, and subsequent learning about relations across modalities can be built upon this innate base."[147] That said, the cross-modal fluency is also intrinsically formed through the interplay with the mother. The fluency is a distinguishing aspect of how the child imitates and relates to the mother. The fluency embeds the syncretic sociability, as Merleau-Ponty would say.[148] The intramodal transpositions between the sense modalities embed the modes and manners that the infant lives the affective bond with the mother.

3.2. Re-Formation of Potential Abilities

With Stern, we can now add another axis in the structural similarity sketched out above. The fact that the student imitates the music cross-modally is similar to how children imitate the amodal expressive forms of people around them. The student glues herself to the music by transposing the musical sense into various modes of behavior, just like the child continually transposes behavioral patterns through the fluency of cross-modal behavior.

Now, to see how forms of vitality influence the aural learning process, we must begin with seeing how the forms influence the life of all of us, not just music students. Just like rhythm and communicative musicality is part of every reasonably normal childhood, the Sternian sensitivity toward forms of vitality is part of any typical adult perception. The issue of vitality never goes away. It remains a real human experience:

> We breathe impressions of vitality like we breathe air. We intuitively evaluate their emotions, states of mind, what they are thinking and what they really mean, their authenticity, what they are likely to do next, as well as their health and illness on the basis of the vitality expressed in their almost constant movements.[149]

We all possess the kind of prereflective "sixth sense" similar to the infant. We intuitively hear, see, and feel the timing, stress, and duration embedded in the voice, breath, and movements of others, immediately synchronizing behavior to the information, long before explicit cognition kicks in—if it kicks in at all. Merleau-Ponty-phrased forms of vitality are intrinsic to the ways that our body schemas radiate over their milieus. It is fundamental to the global, affective awareness of functional bodily knowledge.[150] The fluent transpositions between sense modalities are part of spontaneous perceptual life and the know-how of being alive alongside other living human beings. Forms of vitality are part of the system of equivalences that are the ongoing processes of own body proper.

However, peculiar to the student's situation are the ways she can be said to *concretize*, *use*, and *challenge* this "sixth sense" that we all somehow have. The student concretizes awareness by gluing herself to the music; she uses its potential by seeking new perceptual details in the music, and she challenges its perceptual scope by dealing with rather complex rhythms and strenuous cross-modal exercises.

Let us rephrase the point in terms of potential growth of understanding. In gluing attention to Adderley's music and doing her best in enacting the cross-modal behavior suggested by the exercises, the student now has to discover something anew to get it right. The stringent music initiates a determined push [*Anstoß*], as Gadamer would say:[151] it "forces" her into a process where she has to revise her habituated modes and manners of hearing and enacting perceptual sense. She has to use her global awareness

differently and alter her intermodal fluency, discovering other pathways of unfolding sense across the modalities.

Now, to some extent, it is possible to do the exercises described above mechanically, that is, without understanding. For as Roholt makes clear, no method in the world can guarantee that the body just gets it: "Perceiving a groove requires a kind of ability or *facility* for perceiving grooves."[152] That said, experience shows that, if the student is patient enough, enactive imitation enables something to happen. Even if one starts by imitating the music somewhat mechanically, the multimodal exercises make it easier to catch the musical sense. As pointed out above, this catching can even go unnoticed: it can be subject to a joyous discovery, but it can also be something that happens almost in passing.

For the sake of illustration, let us say the student has struggled with Adderley's beat for some time, enacting the music somewhat mechanically and inflexibly, when this transformation happens (or just did happen). Suddenly, she discovers the rhythmic glue in the music. She hears how Sam Jones hits the bass just a split second before Art Blakey hits his ride cymbal, and how Cannonball cuts through and stretches his phrases precisely as he does, due to a common organizing principle. She hears the ambiguous yet normative force around which every attack, length of tone, and the timespan between the tones make sense. In other words, the student grasps the beat as generative potential. She hears the generative force that "bounces off" and "pushes" the music from within. Moreover, at the same moment she hears the beat *in* the music, she suddenly just knows how to enact the rhythmic order herself. More precisely, she discovers how to let the beat "bounce off" and "push" itself, as if the flow of syncopated rhythms created itself from moment to moment. In other words, the transformation within understanding harbors a moment of emancipation: Where the student first was fully dependent on Adderley's recording to imitate the swing, she now swings from herself.

Interpreting the embodied transformation further, we could say, with Stern, that the student discovers a new sensitivity and flexibility in herself or expands her capacity for being sensitive and flexible in spontaneous, vital behavior. She has further developed her proprioceptive, aural awareness. She has a new way of finding behavioral ease and balance in coordinated, synchronized, and musical behavior. Simultaneously, a new unity has formed itself in behavior. She has discovered a new glue in behavior, so to speak: she just *knows* how the many modes of behavior—the stamping, singing, clapping, and dancing—all participate in the same flow of energy. The rhythmic behavior is imbued with the same feel of vitality, as Stern would say.[153] She feels a new gathering force in the cross-modal enaction, how every act is a variation of the same beat.

In other words, the student has incorporated the generative rhythmic potential of the music into her body schema of possible behavior. She has

acquired a new or extended form of cross-modal fluency, as she can now translate the stamping, singing, or clapping into the same behavioral pattern. To borrow a citation from Merleau-Ponty, "The adjustment of motor excitation to [the aural] excitations is accomplished by their common participation in certain musical essences."[154] The whole body participates in the same musical essence. Things are no longer laid out side by side, as they were when she started out but are allowed to envelop each other, as every sense modality has become transformed into the same "bouncing" or "pushing" behavior.

In our context, it makes sense to say that the transformation that has now occurred within the student's understanding is structurally similar to the ways the child "falls into" the successful rhythmic interactions and predictable cycles of behavior emerging between infant and mother. At the same time, the transformation was prepared for in her innate musical potential and her ontogenetic formation. The capability for rhythmic flexibility was already there, in a potential sense. All the student had to do was release some of the potential by channeling the spontaneous flow of vital energies into a new medium of communicative musicality. The potential release came with the aural imitation.

Editorial note: This is the Part 1 of the article. Part 2 will appear in the spring 2022 issue.

Notes

1. Bjørn Alterhaug, "Improvisation on a Triple Theme: Creativity, Jazz Improvisation and Communication," *Studia Musicologica Norwegia* 30, no. 2 (2004): 97–118; Paul Franklin Berliner, *Thinking in Jazz* (Chicago: University of Chicago Press, 1994); Andy Hamilton, *Lee Konitz: Conversations on the Improviser's Art* (Ann Arbor: University of Michigan Press, 2007); Ingrid Monson, *Saying Something: Jazz Improvisation and Interaction* (Chicago: University of Chicago Press, 1996); Ben Ratliff, *The Jazz Ear: Conversations over Music* (New York: Times Books, 2008); David Schroeder, *From the Minds of Jazz Musicians: Conversations with the Creative and Inspired* (New York: Routledge, 2018); Eitan Y. Wilf, *School for Cool: The Academic Jazz Program and the Paradox of Institutionalized Creativity* (Chicago: University of Chicago Press, 2014); George Lipsitz, "Improvised Listening: Opening Statements Listening to the Lambs," in *The Improvisation Studies Reader: Spontaneous Acts*, ed. Ajay Heble and Rebecca Caines (New York: Routledge, 2014), 27–34.
2. Berliner, *Thinking in Jazz*, 92.
3. Simha Arom, *African Polyphony and Polyrhythm: Musical Structure and Methodology* (Cambridge: Cambridge University Press, 1991); Hal Galper, "The Oral Tradition," Hal Galper, http://www.halgalper.com/articles/the-oral-tradition/2012/, accessed May 21, 2021; Kenneth E. Prouty, "Orality, Literacy, and Mediating Musical Experience: Rethinking Oral Tradition in the Learning of Jazz Improvisation," *Popular Music and Society* 29 (2006): 317–34, https://doi.org/10.1080/03007760600670372, accessed May 21, 2021; Ben Sidran, *Black Talk: How the Music of Black America Created a Radical Alternative to the Values of Western Literary Tradition* (New York: Payback Press, 1981); LeRoi Jones, *Blues People: Negro Music in White America* (New York: Harper Perennial, 2002).

4. Ross Russell, *Bird Lives! The High Life and Hard Times of Charlie (Yardbird) Parker* (Boston: Da Capo Press, 1996); Paul Maher and Michael K. Dorr, eds., *Miles on Miles: Interviews and Encounters with Miles Davis* (Chicago: Lawrence Hill Books, 2009); Chris DeVito, ed., *Coltrane on Coltrane: The John Coltrane Interviews* (Chicago: Chicago Review Press, 2010); Eunmi Shim, *Lennie Tristano: His Life in Music* (Ann Arbor: University of Michigan Press, 2007).
5. Alfred Bennet Spellman, *Four Jazz Lives* (Ann Arbor: University of Michigan Press, 2004); Schroeder, *From the Minds of Jazz Musicians*; Jan Garbarek, "The Garbarek Sound," *Irish Times*, June 30, 2001, https://www.irishtimes.com/news/the-garbarek-sound-1.315743, accessed May 21, 2021.
6. Hal Galper, "Able Bodied—An Interview with Hal Galper," *JazzImprov Magazine* 3, no. 3 (2001), Hal Galper, https://www.halgalper.com/interviews-2/jazzimprov-magazine-interview/, accessed May 21, 2021.
7. Berliner, *Thinking in Jazz*, 96.
8. Wilf, *School for Cool*.
9. Elin Angelo, "Music Education as a Dialogue between the Outer and the Inner: A Jazz Pedagogue's Philosophy of Music Education," in *Philosophy of Music Education Challenged: Heideggerian Inspirations*, ed. Øyvind Varkøy and Frederik Pio (New York: Springer, 2015), 169–84.
10. Roland Baggenaes, *Jazz Greats Speak: Interviews with Master Musicians* (Lanham, MD: Scarecrow Press, 2008); Chris Clarke, "An Oral History of Jazz in Britain," *Oral History* 18, no. 1 (1990): 66–72; Barry Velleman, "Speaking of Jazz: Jazz Improvisation through Linguistic Methods," *Music Educators Journal* 65, no. 2 (1978): 28–31; Wayne D. Bowman, "Doctoral Research in Jazz Improvisation Pedagogy: An Overview," *Bulletin of the Council for Research in Music Education/Research in Jazz Education* 96 (1988): 47–76; Guro Gravem Johansen, "Learning from Musicians Better Than Me: The Practice of Copying from Recordings in Jazz Students' Instrumental Practise," in *Aural Perspectives: On Musical Learning and Practice in Higher Music Education*, ed. Inger Elise Reitan (Oslo: Norwegian Academy of Music, 2013), 75–96.
11. Berliner, *Thinking in Jazz.* Also reflected in Monson, *Saying Something*; Ratliff, *The Jazz Ear*; Schroeder, *From the Minds of Jazz Musicians*; Wilf, *School for Cool*.
12. Berliner, *Thinking in Jazz*, 95.
13. Mellonee V. Burnim and Portia K. Maultsby, eds., *African American Music: An Introduction* (New York: Routledge, 2014); Monson, *Saying Something*; Shane White and Graham White, *The Sounds of Slavery: Discovering African American History through Songs, Sermons, and Speech, Volume 1* (Boston: Beacon Press, 2005).
14. Stephen Malloch and Colwyn Trevarthen, "Musicality: Communicating the Vitality and Interests of Life," in *Communicative Musicality: Exploring the Basis of Human Companionship*, ed. Stephen Malloch and Colwyn Trevarthen (New York: Oxford University Press, 2009), 1–11; Stephen Malloch, "Mother and Infants and Communicative Musicality," *Musicae Scientificae Musicae Scientificae* 3, no. 1 (1999–2000), 29–57; Colwyn Trevarthen and Stephen Malloch, "The Dance of Wellbeing: Defining the Musical Therapeutic Effect," *Nordisk Tidsskrift for Musikkterapi* 9, no. 2 (2000): 3–17; Colwyn Trevarthen, "The Musical Art of Infant Conversation: Narrating in the Time of Sympathetic Experience, without Rational Interpretation, before Words," *Musicae Scientiae* 12, no. 2 suppl. (2008): 15–46.
15. Terrence William Deacon, *The Symbolic Species: The Co-Evolution of Language and Brain* (London: Norton, 1997); Steven Mithen, *The Singing Neanderthals: The Origins of Music, Language, Mind and Body* (London: Orion Books, 2011); Oliver Sacks, *Musicophilia: Tales of Music and the Brain* (New York: Pan Macmillan, 2018); Fred Cummins, "The Territory between Speech and Song: A Joint Speech Perspective," *Music Perception* 37, no. 4 (2020): 347–58, https://doi.org/10.1525/mp.2020.37.4.347, accessed May 21, 2021; Jon Roar Bjorkvold, *The Muse Within: Creativity and Communication, Song and Play from Childhood through Maturity* (New York: HarperCollins, 1992).

16. Colwyn Trevarthen, "Musicality and the Intrinsic Motive Pulse: Evidence from Human Psychobiology and Infant Communication," *Musicae Scientiae* 3, no. 1 suppl. (1999): 155–215, 158.
17. Trevarthen, "Musicality and the Intrinsic Motive Pulse," 157.
18. Malloch and Trevarthen, "Musicality," 4.
19. Colwyn Trevarthen, "The Self Born in Intersubjectivity: The Psychology of an Infant Communicating," in *The Perceived Self: Ecological and Interpersonal Sources of Self-Knowledge*, ed. Ulric Neisser (Cambridge: Cambridge University Press, 1993), 121–73, 121.
20. Berliner, *Thinking in Jazz*, 97.
21. See Roger Scruton, *The Aesthetics of Music* (Oxford: Clarendon Press, 1997).
22. The investigated aspects are important enough, but the aural-communicative dimension distinguished for jazz-musicianship is left more or less unvisited. See, for instance, Richard Ashley, "Musical Improvisation," in *Oxford Handbook of Music Psychology*, ed. Susan Hallam, Ian Cross, and Michael Thaut (Oxford: Oxford University Press, 2009), 413–20; David Borgo, *Sync and Swarm; Improvising Music in a Complex Age* (New York: Continuum, 2005); Peter Culicover, "Linguistics, Cognitive Science, and All That Jazz," *The Linguistic Review* 22, nos. 2–4 (2005): 227–48, https://doi.org/10.1515/tlir.2005.22.2-4.227, accessed May 21, 2021; Bill Dobbins, "Jazz and Academia: Street Music in the Ivory Tower," *Bulletin of the Council for Research in Music Education* 96 (Spring 1988): 30–41; Frederick Seddon and Michele Biasutti, "A Comparison of Modes of Communication between Members of a String Quartet and a Jazz Sextet," *Psychology of Music* 37, no. 4 (2009): 395–415, https://doi.org/10.1177/0305735608100375, accessed May 21, 2021; Steve Torrance and Frank Schuman, "The Spur of the Moment: What Jazz Improvisation Tells Cognitive Science," *AI & Society* 34, no. 2 (2019): 251–68, https://doi.org/10.1007/s00146-018-0838-4, accessed May 21, 2021; Benjamin Givan, "Rethinking Interaction in Jazz Improvisation," *Music Theory Online* 22, no. 3 (2016): 1–24; Jeff Pressing, "Improvisation: Methods and Models," in *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition*, ed. John Sloboda (Oxford: Oxford University Press, 1988), 129–78; Philip N. Johnson-Laird, "How Jazz Musicians Improvise," *Music Perception* 19, no. 3 (2002): 425–42; Alfred Pike, "A Phenomenology of Jazz," *Journal of Jazz Studies* 50 (1974): 88–94; Gary Potter, "Analyzing Improvised Jazz," *College Music Symposium* 30, no. 1 (1990): 64–74; Charles Suhor, "Jazz Improvisation and Language Performance: Parallel Competencies," *ETC: A Review of General Semantics* 42, no. 2 (1986): 133–40; Mark J. Steedman, "A Generative Grammar for Jazz Chord Sequences," *Music Perception* 2, no. 1 (1982): 52–77; Bruno Nettl, "Thoughts on Improvisation: A Comparative Approach," *The Musical Quarterly* 60, no. 1 (1974), 1–19; Bruce Ellis Benson, *The Improvisation of Musical Dialogue: A Phenomenology of Music* (Cambridge: Cambridge University Press, 2003); Julie Dawn Smith, "Playing like a Girl: The Queer Laughter of the Feminist Improvising Group," in *The Other Side of Nowhere: Jazz, Improvisation and Communities in Dialogue*, ed. Daniel Fischlin and Ajay Heble (Middletown, CT: Wesleyan University Press, 2004), 279–95; David Borgo, "The Ghost in the Music, or the Perspective of an Improvising Ant," in *The Oxford Handbook of Critical Improvisation Studies*, ed. George E. Lewis and Benjamin Piekut (New York: Oxford University Press, 2014), 91–114.
23. See, for instance, interviews with Keith Jarrett and associates in Mike Dibb, *Keith Jarrett: The Art of Improvisation* (Euro Arts, 2005), DVD. Explored here are topics related to high-skilled artistic accomplishment but not the fundamental aural abilities enabling the artistic practice in the first place. The aural abilities are just taken for granted.
24. Beatrice Beebe, Dorienne Sorter, Judith Rustin, and Steven Knoblauch, "A Comparison of Meltzoff, Trevarthen, and Stern," *Psychoanalytic Dialogues* online, 13, no. 6 (2008): 777–804, https://doi.org/10.1080/10481881309348768, accessed

May 21, 2021; Talia Welsh, *The Child as Natural Phenomenologist: Primal and Primary Experience in Merleau-Ponty's Psychology* (Evanston, IL: Northwestern University Press, 2013).

25. Andrea Schiavio, Dylan van der Schyff, Silke Kruse-Weber, and Renee Timmers, "When the Sound Becomes the Goal: 4e Cognition and Teleomusicality in Early Infancy," *Frontiers in Psychology* 8, no. 1585 (Sept. 25, 2017), https://doi.org/10.3389/fpsyg.2017.01585, accessed May 21, 2021; Shaun Gallagher, *How the Body Shapes the Mind* (Oxford: Clarendon Press, 2005); Maxine Sheets-Johnstone, *The Primacy of Movement: Expanded Second Edition* (Amsterdam: John Benjamins Publishing Company, 2011); Mattias Solli, "Towards an Embodied Hermeneutics: Gadamer, Merleau-Ponty, and Nondirective Meditation" (PhD diss., Norwegian University of Science and Technology, 2017), https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2433436/Mattias%20Solli_PhD.pdf?sequence=1, accessed May 21, 2021; Evan Thompson, *Mind in Life* (Cambridge, MA: The Belknap Press of Harvard University, 2007).
26. Michael Tomasello, *Constructing a Language: A Usage-Based Theory of Language* (Cambridge, MA: Harvard University Press, 2003).
27. Noam Chomsky, *Language and Mind* (New York: Harcourt Brace Jovanovich, 1972); Noam Chomsky, *Reflections on Language* (New York: Pantheon, 1976).
28. Maurice Merleau-Ponty, "Eye and Mind," in *The Merleau-Ponty Reader*, ed. Ted Toadwine and Leonard Lawlor (Evanston, IL: Northwestern University Press, 2007), 352–78.
29. Maurice Merleau-Ponty, *Phenomenology of Perception*, ed. Donald Landes (London: Routledge, 2012).
30. Merleau-Ponty, *Phenomenology*, 137.
31. Maurice Merleau-Ponty, *The Primacy of Perception*, ed. James Edie (Evanston, IL: Northwestern University Press, 1964), 5.
32. Merleau-Ponty, *Phenomenology*, 212.
33. Shaun Gallagher, "Body Image and Body Schema: A Conceptual Clarification," *The Journal of Mind and Behavior* 7, no. 4 (1986): 541–54; Shaun Gallagher, "Body Schema and Intentionality," in *The Body and the Self*, ed. José Luis Bermúdez, Anthony Marcel, and Naomi Eilan (Cambridge: The MIT Press, 1995), 225–44; Gallagher, *How the Body Shapes the Mind*.
34. Gallagher, "Body Image and Body Schema," 546.
35. Gallagher, *How the Body Shapes the Mind*, 24.
36. This citation is from the 1958 translation of Merleau-Ponty's *Phenomenology of Perception* (London: Routledge, 2003), 160–61. All other citations are from the 2012 edition with Donald Landes's new translation.
37. Maurice Merleau-Ponty, *Signs*, trans. Richard McCleary (Evanston, IL: Northwestern University Press, 1987).
38. We disagree with contemporary attempts to reduce improvised music to the auditory result of movement, as suggested by Iyer, who states, "Music is understood as the sound of human bodies in motion." Vijay Iyer, "Improvisation, Action Understanding, and Music Cognition with and without Bodies," in *The Oxford Handbook of Critical Improvisation Studies*, ed. George E. Lewis and Benjamin Piekut (Oxford: Oxford University Press, 2014), 1–13, 3. We also take issue with the empiristic theory of Pressing, "Improvisation."
39. Bengt Molander, "Tacit Knowledge and Silenced Knowledge: Fundamental Problems and Controversies," in *Skill and Education: Reflection and Experience. Artificial Intelligence and Society*, ed. Göranzon B. and Florin M. (London: Springer, 1992), 9–31; Michael Polanyi, *The Tacit Dimension* (Chicago: University of Chicago Press, 2009).
40. Berliner, *Thinking in Jazz*.
41. John Dewey, *Art as Experience* (New York: Perigee, 2005), 47.
42. Berliner, *Thinking in Jazz*; Hamilton, *Lee Konitz*; Monson, *Saying Something*; Ratliff, *The Jazz Ear*; Schroeder, *From the Minds of Jazz Musicians*; Wilf, *School for Cool*.

43. See, for instance, Greg Corness, "The Musical Experience through the Lens of Embodiment," *Leonardo Music Journal* 18, no. 1 (2008): 21–24; Amy Cimini, "Vibrating Colors and Silent Bodies: Music, Sound and Silence in Maurice-Merleau-Ponty's Critique of Dualism," *Contemporary Music Review* 31, nos. 5–6 (2012): 353–70; Kirsten Fink-Jensen, "Attunement and Bodily Dialogues in Music Education," *Philosophy of Music Education Review* 15, no. 1 (2007): 53–68; Cecilia Ferm and Ketil Thorgersen, "Aesthetic Communication in Music Education: Student's Awareness," International Symposium on the Philosophy of Music Education, Ontario, Canada, 2007; Mariusz Kozak, "Listeners' Bodies in Music Analysis: Gestures, Motor Intentionality, and Models," *Music Theory Online* 21, no. 3 (2015), https://mtosmt.org/issues/mto.15.21.3/mto.15.21.3.kozak.html, accessed May 21, 2021; Simon Høffding, *A Phenomenology of Musical Absorption* (Cham: Palgrave, 2018); Andrea Schiavio and Simon Høffding, "Playing Together without Communicating? A Pre-Reflective and Enactive Account of Joint Musical Performance," *Musicae Scientiae* 19, no. 4 (2015): 366–88, https://doi.org/10.1177/1029864915593333, accessed May 21, 2021.
44. Merleau-Ponty, *Phenomenology*, 32.
45. Daniel N. Stern, *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology* (London: Karnac, 1998).
46. Welsh, *The Child as Natural Phenomenologist*, 57.
47. Mark Gridley, Robert Maxham, and Robert Hoff, "Three Approaches to Defining Jazz," *The Musical Quarterly* 73, no. 4 (1989): 513–31.
48. We do *not* claim that the Western tonal language is better or "more universal" than other musical languages, *only* that it has a certain hegemonic status in the Western context.
49. Eric A. Havelock, *Preface to Plato* (Cambridge, MA: The Belknap Press of Harvard University, 1963); Walter Jackson Ong, *Interfaces of the Word: Studies in the Evolution of Consciousness and Culture* (Ithaca, NY: Cornell University Press, 2013); Walter Jackson Ong, *Orality and Literacy: the Technologizing of the Word* (London: Routledge, 2012).
50. Ong, *Interfaces of the Word*.
51. Richard Taruskin, *The Oxford History of Western Music*, vol. 1 (Oxford: Oxford University Press, 2005), Chapter 1.
52. Havelock, *Preface to Plato*; Ong, *Interfaces of the Word*; Ong, *Orality and Literacy*.
53. Ong, *Interfaces of the Word*, 21.
54. For an overview, see Robert Witmer and James Robbins, "A Historical and Critical Survey of Recent Pedagogical Materials for the Teaching and Learning of Jazz," *Bulletin of the Council for Research in Music Education* 96 (1988): 7–29.
55. Anita Prest, "The Importance of Context, Reflection, Interaction, and Consequence in Rural Music Education Practice," *Journal of Research in Rural Education* 28, no. 14 (2013): 1–13; Eva Georgii-Hemming, Karin Johansson, and Nadia Moberg, "Reflection in Higher Music Education: What, Why, Wherefore?" *Music Education Research* 22, no. 3 (2020): 1–12.
56. Mine Dogantan-Dack, ed., *Artistic Practice as Research in Music: Theory, Criticism, Practice* (London: Routledge, 2016).
57. Sidran, *Black Talk*; William Howland Kenney, *Recorded Music in American Life: The Phonograph and Popular Memory, 1890–1945* (Oxford: Oxford University Press on Demand, 2003); Jones, *Blues People*.
58. Sidran, *Black Talk*.
59. Linda Hunter, "Transformation in African Verbal Art: Voice, Speech, Language," *The Journal of African Folklore* 109, no. 432 (1996): 178–92; Doris Green, "African Oral Tradition Literacy," *Journal of Black Studies* 15, no. 4 (1985): 405–25; Harold Scheub, "A Review of African Oral Traditions and Literature," *African Studies Review* 28, no. 2–3 (1985): 1–72.
60. Paraphrasing Ong, *Interfaces of the Word*, 21.

61. Hans-Georg Gadamer, *Truth and Method*, trans. Joel Weinsheimer and Donald G. Marshall (London: Continuum, 2004), 15. "The universal viewpoints to which the cultivated [gebildet] man keeps himself open are not a fixed applicable yardstick, but are present to him only as a viewpoint of possible others" (15–16).
62. Sidran, *Black Talk*. A philosophical consideration of what happens when the oral tradition's music transforms into art is extrapolated in Mattias Solli, "Tradisjon, Individualitet og Spontanitet: Gadamer og jazz," in *Oppløsningen av det Estetiske. Kunstteori og Estetisk Praksis*, ed. Ståle Finke and Mattias Solli (Oslo: Universitetsforlaget, 2021), 151–72.
63. Berliner, *Thinking in Jazz*, 92.
64. For a blatant example of how representatives from the literary tradition find themselves unable to write down the intonations developed in the African tradition, see William Francis Allen, Charles Pickard Ware, and Lucy McKim Garrison, eds., *Slave Songs of the United States* (New York: A. Simpson and Company, 1867), v–vi. This text also demonstrates a racist and imperialist attitude vis-á-vis the African and oral tradition.
65. Tiger Roholt, *Groove: A Phenomenology of Rhythmic Nuance* (New York: Bloomsbury Publishing, 2014).
66. Galper, *The Oral Tradition*.
67. Russell, *Bird Lives*.
68. Here, we differ from Wilf, *School for Cool*, who describes the aural imitation method as a way to identify oneself as much as possible with the master's creative mind. For a critique of similar ways of conceiving hermeneutical processes, see Gadamer, *Truth and Method,* Part II.
69. Wilf, *School for Cool*, 122.
70. David C. Rubin, "Oral Traditions as Collective Memories: Implications for a General Theory of Individual and Collective Memory," in *Memory in Mind and Culture*, ed. P. Boyer and J. V. Wertsch (Cambridge: Cambridge University Press, 2009), 273–87; Ong, *Orality and Literacy*.
71. Borgo, "The Ghost in the Music"; Smith, "Playing like a Girl"; Marshall Soules, "Improvising Character: Jazz, the Actor, and Protocols for Improvisation," in *The Other Side of Nowhere: Jazz, Improvisation, and Communities in Dialogue*, ed. Daniel Fischlin and Ajay Heble (Middletown, CT: Wesley University Press, 2004), 268–97.
72. Rubin, "Oral Traditions."
73. Johan Huizinga, *Homo Ludens: A Study of the Play Element in Culture* (Boston: Beacon Press, 1970); Gadamer, *Truth and Method*; Bjorkvold, *The Muse Within*.
74. Sidran, *Black Talk*.
75. Wilf, *School for Cool*.
76. Kenneth E. Prouty, *Knowing Jazz: Community, Pedagogy, and Canon in the Information Age* (Jackson: University Press of Mississippi, 2013); Prouty, "Orality, Literacy, and Mediating Musical Experience"; Dobbins, "Jazz and Academia."
77. Malloch and Trevarthen, "Musicality."
78. Hal Galper, "What Is Practicing? Masterclass with Hal Galper," filmed Dec. 17, 2013, available on Youtube.com, https://www.youtube.com/watch?v=rPovnp3Dly42013, accessed May 21, 2021.
79. Immanuel Kant, *Critique of Judgement*, trans. Paul Guyer (Oxford: Oxford University Press, 2008); Gadamer, *Truth and Method*.
80. Merleau-Ponty, *Phenomenology*, 108.
81. Julian "Cannonball" Adderly, "Autumn Leaves," from the album *Somethin' Else* (Blue Note, 1958). In accordance with what we said earlier, this article lets practice suggest the theoretical perspectives. That is, rather than relying on autobiographic interviews with actual subjects, we will try to model the structures of behavior emerging around the aural-imitative practice itself. What this means will be clearer in the analysis of Tables 1–3.

82. Arom, *African Polyphony and Polyrhythm*; John Miller Chernoff, *African Rhythm and African Sensibility: Aesthetics and Social Action in African Musical Idioms* (Chicago: University of Chicago Press, 1979); Jones, *Blues People*.
83. James Barbour, *Tuning and Temperament* (New York: Da Capo Press, 1972).
84. Olly Wilson, "The Significance of the Relationship between Afro-American Music and West-African Music," *The Black Perspective in Music* 2, no. 1 (1974): 3–22.
85. Ross W. Duffin, *How Equal Temperament Ruined Harmony (and Why You Should Care)* (New York: W.W. Norton, 2007).
86. Arom, *African Polyphony and Polyrhythm*.
87. Robert R. Foulkner and Howard Becker, *"Do You Know . . ."* (Chicago: University of Chicago Press, 2009); Vijay Iyer, "Exploding the Narrative in Jazz Improvisation," in *Uptown Conversation: The New Jazz Studies*, ed. Robert O'Meally, Brent Edwards Hayes, and Farah Jasmine Griffin (New York: Colombia University Press, 2004), 393–403.
88. Gregory D. Booth, "The North Indian Oral Tradition: Lessons for Music Education," *International Journal of Music Education* 9, no. 1 (1987): 7–9; Philip V. Bohlman, *The Study of Folk Music in the Modern World* (Bloomington: Indiana University Press, 1988); Chris Goertzen, "The Norwegian Folk Revival and the Gammeldans Controversy," *Ethnomusicology* 42, no. 1(1998): 99–127.
89. Tables 1–3 are suggested by John Pål Inderberg, "Er det mulig å gi undervisning i musikalsk improvisasjon? Et forsøk på å beskrive en metode," unpublished handout (1996).
90. Shelly Berg, *Alfred's Essentials of Jazz Theory: Lessons, Ear Training* (Van Nuys, CA: Alfred Music Publishing, 2004).
91. For a relevant analysis of the nonconceptual yet organized nature of the aesthetic phenomenon, see Kant, *Critique of Judgement*, 45–50, where he describes the free play [*Spiel*] of the cognitive powers initiated by aesthetic phenomena. In encountering beauty, the power of imagination [*Einbildungskraft*], understood as the spontaneous organization of plurality of impressions into unity, on the one hand, and the general power of subsuming diversity under a unitary concept, on the other, are brought into a dynamic and reviving harmony. No rule or concept can capture the free play.
92. Trevarthen, "The Self Born in Intersubjectivity," 121.
93. Wilf, *School for Cool*, 19.
94. Louis Cavrell, "The Universal Mind of Bill Evans," TV documentary, 1966, https://www.youtube.com/watch?v=QwXAqIaUahI1966, accessed May 21, 2021.
95. Berliner, *Thinking in Jazz*, 63.
96. See, for instance, Derek Bailey, *Improvisation: Its Nature and Practice in Music* (Ashbourne, UK: Da Capo Press, 1993); Ratliff, *The Jazz Ear*; Schroeder, *From the Minds of Jazz Musicians*.
97. Monson, *Saying Something*, 29.
98. Tiger Roholt, *Groove: A Phenomenology of Rhythmic Nuance* (New York: Bloomsbury Publishing, 2014), 6.
99. Scruton, *The Aesthetics of Music*, 19–20. Admittedly, Scruton's description has a certain formalist ring to it. But formalism is not relevant in the current context. We use the phrase to describe music as thick perceptual sense, as elaborated below.
100. Vijay Iyer, "Embodied Mind, Situated Cognition, and Expressive Micro-Timing in African-American Music," *Music Perception* 19, no. 3 (2002): 387–414.
101. Roholt thus expands on Merleau-Ponty's conception of perceptual indeterminacy; see Merleau-Ponty, *Phenonemology*; Shaun Kelly, "Seeing Things in Merleau-Ponty," in *The Cambridge Companion to Merleau-Ponty*, ed. Taylor Carman and Mark B. N. Hansen (Cambridge: Cambridge University Press, 2005), 74–110.

102. Monson, *Saying Something*; Berliner, *Thinking in Jazz*.
103. Monson, *Saying Something*, 28.
104. Berliner, *Thinking in Jazz*, 151.
105. Thompson, *Mind in Life*, 38.
106. Scruton, *The Aesthetics of Music*, 234. Again, the point is not to evoke formalism but to describe the relatively autonomous force of music qua perceptual sense.
107. Irene Deliege, "Grouping Conditions in Listening to Music: An Approach to Lerdahl & Jackendoff's Grouping Preference Rules," *Music Perception* 4, no. 4 (1987): 325–59; Fred Lerdahl and Ray Jackendoff, *A Generative Theory of Tonal Music* (Cambridge, MA: MIT Press, 1990); John Sloboda, ed., *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition* (New York: Clarendon Press, 1988); Steedman, "A Generative Grammar."
108. Merleau-Ponty, *Phenomenology of Perception*, 192.
109. Joel Krueger, "Affordances and the Musically Extended Mind," *Frontiers in Psychology* 4, no. 1003 (2014), 1–6, https://doi.org/10.3389/fpsyg.2013.01003, accessed May 21, 2021; Micheline Lesaffre et al., "Participatory Sense-Making in Joint Musical Practice," in *The Routledge Companion to Embodied Music Interaction*, ed. Micheline Lesaffre, Pieter-Jan Maes, and Marc Leman (New York: Routledge, 2019), 31–39; Jakub Ryszard Matyja and Andrea Schiavio, "Enactive Music Cognition: Background and Research Themes," *Constructivist Foundations* 8, no. 3 (2013): 351–57; Simon Høffding and Andrea Schiavio, "Exploratory Expertise and the Dual Intentionality of Music-Making," *Phenomenology and the Cognitive Sciences* (2019), https://doi.org/10.1007/s11097-019-09626-5, accessed May 21, 2021; Joel Krueger, "Enacting Musical Experience," *Journal of Consciousness Studies* 16, nos. 2–3 (2009): 98–123; Andrea Schiavio, Dylan van der Schyff, Julian Cespedes-Guevara, and Mark Reybrouck, "Enacting Musical Emotions: Sense-Making, Dynamic Systems, and the Embodied Mind," *Phenomenology and the Cognitive Sciences* 16, no. 5 (2017): 785–809; Mark Reybrouck, "Music as Environment: An Ecological and Biosemiotic Approach," *Behavioral Sciences* 5, no. 1 (2015): 1–26.
110. Schiavio and Høffding, "Playing Together without Communicating," 370. Italics original.
111. Maurice Merleau-Ponty, *Institution and Passivity: Course Notes from the Collège De France (1954–1955)*, trans. Leonard Lawler and Heath Massey (Evanston, IL: Northwestern University Press, 2010), 18–20.
112. Cavrell, *The Universal Mind*.
113. Berliner, *Thinking in Jazz*, 63.
114. Thomas Fuchs and Hanne De Jaegher, "Enactive Intersubjectivity: Participatory Sense-Making and Mutual Incorporation," *Phenomenology and the Cognitive Sciences* 8, no. 4 (2009): 465–86, https://doi.org/10.1007/s11097-009-9136-4, accessed May 21, 2021; Merleau-Ponty, *Phenomenology of Perception*; Charles Keil, "Participatory Discrepancies and the Power of Music," *Cultural Anthropology* 2, no. 3 (1987): 275–83; Roholt, *Groove*.
115. Gadamer, *Truth and Method*; David McNeill, *Gesture and Thought* (Chicago: University of Chicago Press, 2005); Merleau-Ponty, *Signs*.
116. Merleau-Ponty, *Phenomenology*, 100ff.
117. Galper, *The Oral Tradition*.
118. Quoted in Berliner, *Thinking in Jazz*, 152.
119. Roholt, *Groove*, 4. Italics original
120. Roholt, *Groove*, 4. Italics original.
121. Trevarthen, "The Self Born in Intersubjectivity," 121.
122. Gadamer, *Truth and Method*.
123. Stern, *The Interpersonal World of the Infant*, xviii.
124. Stern, *The Interpersonal World of the Infant*, 46. Italics original.
125. Stern, *The Interpersonal World of the Infant*, 128.

126. Colwyn Trevarthen, "First Things First: Infants Make Good Use of the Sympathetic Rhythm of Imitation, without Reason or Language," *Journal of Child Psychotherapy* 31, no. 1 (2005): 91–113, https://doi.org/10.1080/00754170500079651, accessed May 21, 2021; Trevarthen, "Musicality and the Intrinsic Motive Pulse."
127. Fuchs and De Jaegher, "Enactive Intersubjectivity," 479.
128. Edward Tronick and M. Kathrine Weinberg, "Depressed Mothers and Infants: Failure to Form Dyadic States of Consciousness," in *Postpartum Depression and Child Development*, ed. L. Murray and P. J. Cooper (New York: Guilford, 1997), 54–84; Edward Tronick and Jeffrey Cohn, "Infant–Mother Face-to-Face-Interaction: Age and Gender Differences in Coordination and the Occurrences of Mismatches," *Child Development* 60, no. 1 (1989): 85–92.
129. Fuchs and De Jaegher, "Enactive Intersubjectivity," 479.
130. Trevarthen, "First Things First."
131. Maurice Merleau-Ponty, *The Structure of Behavior*, trans. A. L. Fisher (Pittsburgh, PA: Duquesne University Press, 2011); Daniel N. Stern, *Forms of Vitality* (Oxford: Oxford University Press, 2010).
132. Trevarthen, "Musicality and the Intrinsic Motive Pulse," 178.
133. Trevarthen, "First Things First," 92.
134. Trevarthen, "The Self Born in Intersubjectivity," 151.
135. Trevarthen, "The Self Born in Intersubjectivity," 135.
136. Trevarthen, "Musicality and the Intrinsic Motive Pulse," 180.
137. Colwyn Trevarthen, "Play with Infants: The Impulse for Human Storytelling," in *The Routledge International Handbook of Early Childhood Play*, ed. Tina Bruce, Pentti Hakkarainen, and Milda Bredikyte (London: Routledge, 2017), 198–215.
138. Jessica Benjamin, "Beyond Doer and Done: An Intersubjective View of Thirdness," *The Psychoanalytic Quarterly* 73, no. 1 (2004): 5–46.
139. Trevarthen, "First Things First."
140. Gadamer, *Truth and Method*.
141. Merleau-Ponty, *Institution and Passivity*, 18–19.
142. Daniel N. Stern, "Face-to-Face Play: Its Temporal Structure as Predictor of Socioaffective Development," in *Rhythms of Dialogue in Infancy: Coordinated Timing in Development*, ed. Joseph Jaffe et al. (Boston: Wiley/Society for Research in Child Development, 2001), 144–49, 147.
143. David Morris, *Merleau-Ponty's Developmental Ontology* (Evanston, IL: Northwestern University Press, 2018); Merleau-Ponty, *The Structure of Behavior*.
144. Stern, *The Interpersonal World of the Infant*, 85.
145. Stern, *Forms of Vitality*, 7.
146. Stern, *The Interpersonal World of the Infant*, 49.
147. Stern, *The Interpersonal World of the Infant*, 48.
148. Maurice Merleau-Ponty, *Child Psychology and Pedagogy: The Sorbonne Lectures, 1949–1952*, trans. Talia Welsh (Evanston, IL: Northwestern University Press, 2010), 255ff.
149. Stern, *Forms of Vitality*, 3.
150. Merleau-Ponty, *Phenomenology*, 102.
151. Gadamer, *Truth and Method*.
152. Roholt, *Groove*, 41. Italics original.
153. Stern, *Forms of Vitality*, 84.
154. Merleau-Ponty, *The Structure of Behavior*, 121.

Learning Jazz Language by Aural Imitation: A Usage-Based Communicative Jazz Theory (Part 2)

MATTIAS SOLLI, ERLING AKSDAL, JOHN PÅL INDERBERG

Abstract. *How can imitation lead to free musical expression? This article explores the role of auditory imitation in jazz. Even though many renowned jazz musicians have assessed the method of imitating recorded music, no systematic study has hitherto explored how the method prepares for aural jazz improvisation. The article uses Berliner's assumption that learning jazz by aural imitation is "just like" learning a mother tongue. The article studies three potential stages in the method, comparing them to the imitative, rhythmic, multimodal, and protosymbolic behavior of infant perception (building on the works of Stern, Trevarthen, and Merleau-Ponty). The demonstrations of the aural imitation method draw on pedagogic experiences accumulated since 1979 at the Jazz Program at the Norwegian University of Science and Technology. By analyzing structures of behavior suggested by the method, the article indicates key traits that render aural jazz improvisation possible, such as a fundamental sense of rhythm, formation of*

Mattias Solli, PhD, is a Norwegian philosopher. He is a former jazz student and jazz musician, currently working as a researcher in the Department of Philosophy and Religious Studies, Norwegian University of Science and Technology (NTNU). In addition to jazz music, his research interests are phenomenology, hermeneutics, aesthetics, and ethics. Solli is a co-editor (with Ståle Finke) of *Oppløsningen av det estetiske: Kunstteori og estetisk praksis* (2021).

Erling Aksdal is a Norwegian jazz pianist and composer. As a pianist, he has performed and recorded with Chet Baker, Elvin Jones, Warne Marsh, and Lee Konitz, as well as many notable Norwegian musicians. He has served as council member of Association Européenne des Conservatoires (AEC) and was a founding member of AEC's Pop Jazz Platform. Aksdal currently heads the Jazz Program at the Department of Music, NTNU.

John Pål Inderberg is a Norwegian jazz musician and bandleader and a Professor Emeritus of the Department of Music, NTNU. As a saxophonist, he has performed and recorded with Warne Marsh, Lee Konitz, Chet Baker, Bob Brookmeyer, Gil Evans, Chick Corea, and Pat Metheny, as well as with many notable Norwegian jazz musicians. Inderberg taught the saxophone to several generations of Norwegian jazz musicians.

Journal of Aesthetic Education, Vol. 56, No. 1, Spring 2022

symbolic behavior, joint musical attention, and the facility to "hear via the other." In conclusion, we critically address a frequent theoretical model describing musical improvisation as a synthesis of discrete elements or building blocks.

This is Part 2 of the article (for Part 1, see JAE *55.4). It continues with the section enumeration.*

4. Enhancing the Process

The previous sections in Part 1 (*JAE* 55.4) began fleshing out Berliner's assumption, that is, the assumption *that learning jazz by aural imitation is "just like" learning a mother tongue*. The all-important role of rhythm in the aural imitation method suggested the relevance of rhythmic protolinguistic behavior. We indicated how understanding rhythm by aural imitation includes a bodily transformation. The aural imitation method is about to form a general functional ability embedded in an auditory awareness toward general potential in the music and spontaneous behavioral impulses.

Let us introduce the next stage in the aural imitation method to see how the practice suggests more theoretical perspectives (Table 2; table included also in Part 1). Our focus will be on the learning of the Western tonal language. This does not infer that rhythmic learning is considered finished or surpassed, only that we shift attention toward another aspect of the musical language as a whole.

4.1 Embodying Tonal and Harmonic Relations

4.1.A. ENHANCED PROCESS

Everything happening in Table 2 is a further development of what we discussed in Part 1. The exercises bring in no new building blocks, but they intensify the mimetic enactment along axes established in the first phase.

4.1.B. ENHANCED MULTIMODALITY

The exercises initiate an enhanced differentiation between the mental and the bodily production of the music, which further reinforces the multimodal coordination tasks introduced in Table 1 (see Part 1). While rhythmic features are played on the instrument, the melodic material is heard in mind only (Instr. 2), and while the instrument is used to break chords, the keynote is continuously heard in the inner ear (Instr. 5).

In addition, the student is encouraged to record the playing (Instr. 4). Recording oneself functions as a reality check; rather than relying on mere intrasubjective evaluations, the student now brings in external documentation of the process.

Table 2. Phase 2: Tonal and harmonic orders and the use of the instrument

Step	Description	Tool
Instr. 1	Repeat the exercises in Act 1a and b (Acts 1–7 are presented in Table 1). Play the individual rhythmic elements and optionally the rhythmic whole that can be performed naturally on your instrument. Use feet and, if possible, voice to complement your playing.	Instrument, singing, clapping, foot-tapping
	Let the basic rhythm of the recording that you imitated in Act 1b "sound" in your ear. After a few minutes of practicing this, emphasize parts of this rhythm on the instrument.	
Instr. 2	Listen to the core melodic material of the composition/recording. Play the melody rhythm while "listening" to your perceived version of the melody line. Select one or more random tones.	Instrument, singing, clapping, foot-tapping
	Repeat Act 4. Listen to your inner hearing while playing the same melodic material on your instrument in multiple registers. Play with different levels of dynamic, timbre, and tempo.	
Instr. 3 and Instr. 4	Listen to your inner hearing's version of the rhythmic basis of the recording. Play the melody lines from Act 4 simultaneously. Play as if you were singing in the instrument. Play the same melodic material while allowing individual rhythmic elements to be expressed using feet (or voice).	Instrument, singing, clapping, foot-tapping
	Vary dynamic, timbre, and tempo. Remember to make a recording of your playing. Is there a match between your hearing and playing on the recording?	
	Percussionists who do not play a melodic percussion instrument sing the exercises in Instr. 3, Instr. 4, and Instr. 5.	
Instr. 5	The tone that represents the center of the chord (root) is played (see Act 7). One note in each chord throughout the recording should match the root tone in your hearing.	Instrument, singing, clapping, foot-tapping
	Play the four lower notes in all the chords from the root to the seventh. Then arpeggiate the chords up and down in a sequence that you spontaneously decide. Use the full range of the instrument (see Act 7).	
	Make music from this exercise. No exercise should be just technical.	
Instr. 6	Practice your perception of period by, e.g., playing arpeggiated chords (see Act 7b), main melodic lines (see Act 4) or melodic bass lines (see Act 6). Keep the recording's main rhythm going in feet or voice.	Instrument, singing, clapping, foot-tapping
	Make recordings so you can check if you are keeping the periods. Play diverse sub-elements from this table with varying intensity, timbre, register, and dynamics.	

However, on a more profound phenomenological note, listening to self-recording means a further cross-linking in the student's behavior, namely, between a mode of listening *that is and can only be subjective* and a way of listening to music that is *principally intersubjective*. Although only the student can hear her own inner, mental presentation of the music, the recording documents a musical expression available to others. Any discrepancy between the subjective and the intersubjective both broadens the understanding of one's own game and prepares for actual interaction with other people.

4.1.C. USE OF INSTRUMENT IS SECONDARY TO THE HOLISTIC AUDITORY AND MUSICAL ACTIVITY

The student involves the musical instrument gradually, comprehensively, and in ways fully determined by the music. The instrument is brought in via participation in the rhythmic (qualitative and vital) form, before gradually enacting melodic and harmonic structures.

Most importantly, however, imitational work is still holistic. It involves the whole body. The task of coordinating breath, fingers, and limbs to produce sounds on the specific instrument is not a limited task of its own but integrated into the overall embodied enactment. In this sense, producing sounds on the musical instrument is secondary to the general embodied, expressive engagement in the music. The musical instrument is accidental compared to the musical essentials enacted by the whole body.

Let us say, as a thought experiment, that the instrument had been brought in from the start. In that case, the student would be tempted to figure out the music not from the vantage point of the ear but from physical movement. Mechanical movements would easily be primary, whereas the ear would be secondary. Consequently, the modes of perceiving and replicating music would be limited by contingent embodied factors, that is, by not only individual technical skills but also the instrument-specific ways of being in the music. While chord-instrument players like pianists and guitarists would approach the music through the possibilities of putting down chords vertically, players of melodic instruments like saxophone or trumpet would approach the music horizontally. Habits would stay in the way of the general musical sense unfolding in Adderley's music.

Whereas much pedagogic literature on playing improvisational jazz focuses on the instrument and only indirectly on the ability to hear the musical sense,[1] outspoken aural musicians underline the need to think *ear first*. As Galper states, the idea that the musical instrument is the instrument is an illusion: "*You* are the instrument. Everything you are working on is internal to you." The body proper is the instrument. "Your tools are your mind, your body, and your emotions. They are trainable."[2]

Now, in a limited sense, the current usage of the musical instrument can be compared with the famous phenomenological example of a blind person's use of a stick.[3] For the blind man, the world becomes available through

the gradual tip-tapping of the stick. He explores and maps out the environment through physical movements, interaction, and sensor-motor feedback. Analogously, the student now uses her musical instrument to explore the musical world. Moving the fingers over the keys of the saxophone or the keyboard of the piano is skillful probing stretching toward possibilities *out there*. Both stick and musical instruments are prolongations of the sensor-motor capacities of the phenomenal body.

The limitation of the analogy marks the difference to the sonic-environmental approach. For, as we pointed out previously, the environment about to be explored by the student is not "just" a sonic environment but a symbolic, expressive language. The skillful probing is a way to explore a communicative system of equivalences—the symbolic system of the major-minor tonality—through the musical instrument. In this respect, the use of the musical instrument is comparable to the moving of lips, vocal cords, and lungs in speaking with others. The musical instrument is part of the communicative, expressive powers of the body proper in the sense that the moving stick is not.

In practice, the implementation of the musical instrument will require much exploratory effort from the student. There might be new sounds or difficult technical passages that need to be figured out.

While the subsequent sections will expand on the music as a symbolic communicative system, we need to set aside the questions about the use of the instrument, in particular, the tricky issues of the relationship between aural and motor intentionality. Suffice to say that, when a child starts imitating the linguistic sounds of people around, motor intentionality (say, of lip movements or guttural sound production) is secondary compared to the expressive drive to *say something* in relation with others. Similarly, the student uses the instrument just to sound "like that" as much as possible, as Galper puts it,[4] adjusting breath, fingers, lips, and so forth, to the unfolding musical sense.

4.1.D. ENHANCED FOCUS ON TONAL–HARMONIC RELATIONS

The student trains more explicitly to recognize the root of the chord, while at the same time practicing spontaneous voicings of the chords as a whole (Instr. 5). Where the first step (Table 1) established primary auditory contact with keynotes and key relations, the following steps (Table 2) encourage a somewhat freer emulation of the harmonics, exploring the tonal relations latent in the harmonies. In effect, the student begins to juxtapose a vertical and horizontal mode of listening within the harmonics. She strives to hear both how the tones of chords are "piled" vertically and how they can be varied and flexed in voicing while still being the same chord.

As it has been all the way, the point of the exercises is not to plan tonal pathways but to try them out *directly*, with the instant feedback of the increased sensitivity of the aural attention. The student trains to discriminate more

and more precisely within the musical wholes of the major-minor tonality. And in this specific sense, too, the usage of the instrument is subordinate to the aural investigation of the tonality. The student does not use the instrument to hear tonal relations but, contrariwise, to enact the musical relations that she hears.

In this enhanced focus on tonal–harmonic relations, we find the theme that we now need to elaborate on: Where the previous section pursued specifics belonging to rhythm and rhythmic learning, we turn now to the questions about what it implies to learn the Western tonal system, with a particular interest for the generative orders of the tonal system.

4.2 Hearing Generative Tonal Sense

What does it mean to learn the generative syntax of the tonal system in the ways suggested by the exercises described above? To elaborate on an answer, we first need to see what it implies to target the tonal system as a thick audio-perceptual sense (that is, not an intellectual or computational system and not as a system associated with visual notation).[5] Scruton illuminates how the Western tonal system represents a stringent syntax that is both limited and unlimited at the same time. On the side of the limitation, the system consists of *only* twelve tones. Besides, granted that the music is tonal (that is, not nontonal or atonal, as the musical languages fostered in some of the free-jazz traditions), it is organized by inevitable tensions and releases. In Scruton's definition, tonal music requires that the following four conditions are met:[6]

1. The melodic line feels fully "closed" only when it comes to rest on a certain privileged tone (the tonic).
2. The final move on to the tonic has (in standard cases) the character of a "cadence"—a loosening of tension.
3. Octaves are heard as equivalent—so that the effect of closure is duplicated at the octave.
4. Other tones are heard in relation to the tonic—as more or less distant from it, as tending toward or away from it.

A tonal *key*, thus, represents a perceptual norm or a center of aural gravity. All other notes have a force relative to the center. This brings melody and harmony into close and constant relations. All melodic horizontal lines get harmonic implications relative to the key, and all harmonies link to how the melody and the key are heard.[7]

On the other side, the potential *use* of the perceptual system is unlimited, especially when the equally tempered version of the tonal system is used.[8] On standardized contemporary instruments, the Western tonal system is equally transposable into all keys, starting on all twelve tones. This possibility gives the keys a certain perceptual indeterminacy.[9] Each key is defined by a field of possible tones of diatonic and chromatic organization. It unfolds in

relation to the other keys, according to the almost perfect circle of fifths. Each key relates by a fifth to its "neighbor" key, except the key starting on the seventh tone, which relates to its "neighbor" with an imperfect fifth. The melodic order of scales that compose a key generates a synthesis of melodic and harmonic perception. The diatonic scales can be heard as a system whereby the two harmonic affinities—the octave and the fifth—are worked into the very substance of the tonal music, and their intrinsic relation is resolved.[10]

The generative potential of this system cannot be "used up." It can be unfolded in endless musical varieties, a fact that Adderley, Coltrane, Shorter, Bley, Schneider, Bach, Beethoven, Shostakovich, and others demonstrate in their unlimited creative potential of the system. While their works sound distinct and personal, all these artists follow the same perceivable laws for polyphonic musical processes, harmonic relationships, cues, scale relationships, and melodic and chromatic formations.

The potentialities of *this* system are what our student is about to explore. By trying out the various modes of listening within the harmonies, she explores the aural syntax developed through centuries of aural artistic investigations. However, as was the case with rhythm, if she grew up with songs in major-minor tonality, the student does not start from scratch when she starts imitating the record. She already knows the fundamentals. In the womb, the fetus is already surrounded by natural ratios of the harmonic series produced by its mother's voice, literarily affecting the formation of the auditory capacity.[11] Besides, the postnatal rhythmic interactions described above typically have harmonic implications in overtones and melodies.[12] Thus, in a certain sense, the tonal syntax is already part of the ontogenetic and phylogenetic institution of the student's body schema. She spontaneously hears within the aural gravity of the tonic. Even before she begins the methodological ear training, she perceives something *in* the musical organization, as Scruton puts it above. She hears the moving force, how specific forces in the music propel it in certain directions, as when a tune *moves* toward the tonic. In other words, she already has an aural concept of the generative syntax of the Western tonal system. She hears generative potentials of harmony, the intangible yet normative *something* organizing the music from within.

However, it is one thing to possess a more or less vague preunderstanding of the tonal syntax but another to have a clear understanding of the harmonic system as such, that is, *an understanding of the system qua aural system*. The latter form of knowledge is both a goal and a crucial necessity among professional musicians playing tonal jazz.[13] To compose and generate music in real time, as Bill Evans puts it,[14] the musician needs a familiarity with the harmonic syntax as profound and nuanced as the mother tongue. The musician needs to *breathe* harmony.[15]

Hence, as a future aural musician, the student will need an advanced understanding of how to catch and enact tonality in the spur of the moment,

filling in the adequate melodic and harmonic solutions called for. She will need the facility to perceive—with high precision—the melodic and harmonic implications acted out by herself and others (often many others) while simultaneously responding adequately in the same musical language. In other words, she will need a discriminating sensitivity for what goes on in the music and creative capacity for tonal production. In short, she will need to be fluent in the Western tonal language. She will need to have the whole language at her disposal—and the readiness to use it.

4.3 Symbolic Behavior and Joint Musical Attention

If aural fluency within the Western tonal language is the desired outcome of imitational work described above, what, then, are the critical structures of behavior implied in fluency? If we know the contours of an answer to this question, we can better see how the method works toward that goal and how ontogenetic development prepares for the leap into fluency.

In our context, it makes sense to suggest that a full-fledged aural understanding of the tonal generative syntax must have something to do with the awareness and multimodal fluency associated with forms of vitality. Skilled aural musicians reporting of their tonal fluency seem to have incorporated and internalized the order of the tonal system to the degree that the energy and multimodal fluency of the forms of vitality are transposed into the keys of the twelve tones. That is, they perceive perceptual groupings organized not only due to shared temporal forms but also due to their participation in the tonal organization. The body schemas of the musicians transpose not only between sense modalities but also along with the system of equivalences offered by the tonal system. The human life is potentially transposed into the all the keys of the twelve tones (so to say).

What complicates the analysis, however, is the fact that the harmonic system is not "just" forms of vitality but a generative *system* in the sense that forms of vitality, for themselves (whatever this means), are not. The tonal system is a full *language*, a complex *whole*, whose intrasystemic pathways are strictly organized by the twelve tones and the nonarbitrary rules of the major-minor tonality.

Compared to the forms of vitality described above, the tonal system constrains behavior in new ways. Just like verbal language begins to narrow down the expressive potential of a child around nine months of age,[16] the twelve tones and the tonal syntax of major-minor tonality narrow behavior. Not just anything goes: unless the unfolding musical sense calls for it, a major third hammered out on a minor chord just sounds *wrong*. However, just like verbal language opens new possibilities of expression for the child to express itself vis-à-vis the other, the restrictions of tonality *also* enable full new possibilities of expression and new modes of "being with the other," as Stern would say.[17] The aural musician has acquired a complete and

intersubjective medium that makes her capable of expressing *herself* in live communication *through* the aural generative syntax shared with others.

To elaborate on the tonal generative facilities in agreement with Trevarthen's and Stern's theories, we conceive the facility as a specialized form of *symbolic behavior*. Now, in other contexts, the words *symbol* and *symbolic* often become associated with intellectual cognition or verbal signs. For instance, Reybrouck associates symbolic with mental representation, thinking, computation, and conceptualization.[18] By contrast, in our usage-based approach, symbolic behavior equals neither intellectual computation nor linguistic expression as such. Instead, it designates the domain-general, perceptual ability to let *something symbolize something else*—in radically open, transformative, and intersubjective relations:[19]

> Symbols imply the mental ability to grasp something as an invariant under a diversity of aspects and perspectives. Thus symbols imply the ability to grasp something as an *object*, in the phenomenological sense of something that remains invariant through perspectival variation and is graspable for the subject and also available for other subjects.[20]

Assuming we do not overintellectualize the word *mental* (which is something Thompson thoroughly criticizes) and do not protest against the use of the word *object* in musical contexts (music is not an object, but in the phenomenological sense, it is an *intentional object for the subject*), it seems reasonable to say that aural musicians have cultivated symbolic behavior in the form of spontaneous, real-time music-making. They can grasp and vary harmonic invariants embedded *in* the tonal and rhythmic language. Along the complex, intrasystemic pathways of the rhythmic and tonal syntax, they can catch and enact generative potentialities embedded in the musical flow. They can hear how there is something *in* the sound, something that moves with a force of its own, immediately responding within the lawfulness of the tonic atmosphere.[21]

Thus, aural facilities qua symbolic behavior imply the ability to identify and produce indeterminate yet normative perceptual potentialities emerging within the hearable constraints of the twelve tones and the laws of tonality. The musicians can differentiate precisely how the music *leads* according to the suggestions of the octave, the circle of fifths, the diatonic and chromatic scales, and the rest of aural lawfulness that regulates the polyphonic musical processes. They can catch and enact lateral relations opened by the possibility of varied expressions of the same intrasystematic sense.[22]

However, as also indicated by Thompson, cultivating symbolic behavior cannot be reduced to the facility to perceive and manipulate symbols for themselves. It implies an all-important intersubjective dimension. Symbol manipulation is an activity executed *with, against*, and *in accordance with* how other human beings perceive the same symbolic order. As Trevarthen points

out above, musicality is a *communicated talent*—a talent for communicating in live, direct, or virtual company with others.

Recall that jazz musicians generally recognize aural facilities as fundamental. Though there will be empirical variation as to how much the idea of auditory interaction is accomplished, there should be no doubt that the musicians who set the standards of the business fulfill the ideal. The recordings of Adderley, Coltrane, Armstrong, Parker, Mingus, Ellington, and Evans are exemplarily prototyped documents of *music heard*.[23] They are aural experts and teachers in the highest artistic and pedagogic sense, to borrow Archie Shepp's phrasing.[24] Moreover, it should not be controversial to say that these musicians are masters of the collective and polyphone musicianship first emerging in the African context.[25] Their recordings document the spontaneous, auditory, and musical behavior unfolding within the constraints and possibilities of shared musical languages. They document the genuine accomplishment of the human potential for real-time communicating in live and instantaneous communication. (For proof, listen to their recordings.)

Now following Tomasello,[26] it seems reasonable to say that the aural facilities of high-skilled aural musicians imply a form of *joint attention*—or joint *musical* attention, as we will call it. Joint musical attention implies the abilities to

- direct attention toward the same music as heard by others. (This is a banal yet crucial condition for collective music-making);
- hear not only how things are played but also how they could be played, that is, the facility to perceive the rhythmic and tonal generative potentials latent in the music;
- follow the musical attention of the other. This implies the facility to perceive the rhythmic and tonal generative potentials about to be acted out by fellow musicians;
- lead the attention of the other toward self-perceived musical potentialities. This brings in a mutual dialectic in the live company. Both parties lead the attention of the other;
- learn through aural imitation. Imitation is not only a propaedeutic concern but conditions also the activity of pursuing the same musical sense as unfolded by peer musicians.

These criteria are modifications of Tomasello's list defining joint linguistic attention, a fact reflecting the shared background of music and language in human musicality. Moreover, while the criteria indicate a self/other relationship, their implications can be rephrased with a focus on the musical sense unfolding between the listening subjects. From the vantage point of the music, joint musical attention implies the abilities to

- hear how perceptual and musical categories of similar and distinct musical gestalts are formed and dissolved;

- form perceptual and musical categories of how similar and distinct musical gestalts are formed and dissolved;
- hear musical transpositions based on similar functional roles of the musical gestalts;
- form musical transpositions based on similar functional roles of the musical gestalts.

To appreciate these points, we need to see the phenomenological correlation between the aural abilities on the side of the listening *subjects* and the generative potentials in their intentional objects, namely, the music. The music is the intermediary reality, the communicative sense that unfolds, reflects, and embeds the aural horizons of the players. Joint musical attention, then, is the ability to hear the rhythmic and harmonic generative potentials in substantial co-perception with others, as the meaningful possibilities unfold in real-time interaction. It is the ability to catch multiple aural horizons latent in the music, emerging because other people hear the same music differently. For, as we recall that humans *always* perceive more or less differently,[27] the generative potentials can ever be pulled in different directions in joint musical attention. There is no one way to hear the unfolding music, but many—in fact, as many as there are possible ways of listening.

To illustrate, let's say a pianist suddenly hears a latent sharp nine in an unfolding dominant seventh chord, suggested by the flat intonation of the saxophonist. By subtle manipulation of the music, the pianist can now lead the attention of the others toward this potentiality, perhaps by indicating a substitute chord. The others might respond to this initiative directly or by "transposing" what they hear into a rhythmic structure that, for them, serves a similar functional role in the spur of the moment. The crux of joint musical attention evolves in this facility to hear how rhythmic and harmonic gestalts form and dissolve themselves within the context of living, instantaneous, and polyphone communication.[28]

Crucially, this description of joint musical attention will prove essential to see where the student's process potentially is headed. Learning the tonal language and cultivating joint musical attention are but aspects of the same process. It is worth noting how our definition of joint musical attention differs from contemporary concepts of joint musicianship. Seddon, Seddon and Biasutti, and Phillips-Silver and Keller seek to clarify how skilled musicians are capable of rapid mind-reading and understanding the plans of the other.[29] In criticizing their accounts, Schiavio and Høffding suggest an account of joint musical awareness, focusing on how a group of interviewed musicians is thoroughly absorbed in their doing, with neither time nor need to read the mind or emotional states of the peer musicians.[30] However, despite substantial differences in other regards, none of these theorists takes substantially into account how collective musicianship is a *mediated activity*. That is, they do not consider the fact that skilled musicians listen

to each other in and through a thick symbolic communicative medium. The music is just there, so to say, as the neutral sonorous medium between the players.

In our context, by contrast, questions about degrees of conscious awareness of self and others are irrelevant, compared to the fact that skilled musicians listen to each other precisely in and through the shared artistic and communicative medium. The *jointness* in the joint musical attention expresses no direct relation between subjects but is a reflected relation. The musicians listen to the same musical sense, and they unfold it together. They share attention by unfolding the same generative potential from moment to moment.

5. Hearing Symbolic Indications

The previous subsections suggest a series of perspectives on what it implies to learn the Western tonal system by ear and to be able to use the language in live communication with others. Along with these perspectives, the exercises described above are a way to train symbolic behavior. Importantly, rhythmic learning has embedded a symbolic dimension. In cultivating her overall capacity to identify and vary coherent rhythmic unity across the multimodal variation, the student has enacted subtle modes of selfsameness under a diversity of aural perspectives. In Thompson's definition, she trained to grasp something as an invariant under a variety of aspects and perspectives.

However, with the exact replication of the tones and relations between tones (Table 1) and the slightly more spontaneous variation encouraged in Table 2, symbolic learning takes a more specified form, as the student is about to explore the aural selfsameness and equivalences suggested with the tonal atmospheres. She tries out the various modes of listening latent within the keys. She is about to hollow out the multiple, systematic pathway variances implied in the tonal invariances. In other words, the student trains her capacity to catch and enact the generative harmonic potentials. She tries to grasp the invariances unfolding *in* the tonal language, something that moves with a force of its own, propelling the music in various directions.

In Scruton's parlance, the student tries out how the tonal atmosphere of the chords and the sequence harbors latent possibilities of variations within the musical atmosphere of their functioning.[31] She tries out how the octaves can be heard as the "same again" across the tonal differences, how various voicings of the II-V-I cadences create slightly different tensions and releases in the music while still fulfilling the same functional roles, and how the harmonic and melodic potential of the tonic separates into other diatonic and chromatic relations. In other words, she tries to hear how tones and relations between tones belong to the same tonal atmosphere and how the tonal

atmosphere can be strengthened or weakened by tone sequences moving more or less distant to the key, tending toward the center or away from it. She tries to hear unused potentialities within the tonal atmosphere—foreign tones excluded from the principal regions of the key, also creating transitions and potential tensions within the atmosphere. In short, the student explores the tonal language as a perceptual system of equivalences sketching out the generative potentials embedded in the Western tonal language.

The all-important crux, then, is to see how this mode of symbolic learning has an intrinsic dimension of intersubjective and communicative relatedness. In line with the general argument of this article, the student is not replicating and using an abstract, computational generative system. She is about to learn a genuine symbolic language used in real-time communication between real human beings, namely, the Adderley Quintet. By replicating, enacting, and incorporating *their* music, she is about to imitate and incorporate their means of communication in joint musical attention, that is, their aural indications embedded in the harmonic (and rhythmic) generative potentials.

Herein evolves new profundities implied in Berliner's assumption. As we will come to see, this mode of symbolic learning involves structural similarities to a new phase in ontogenetic language development related to the entry into verbal, symbolic language by the age of seven to nine months. By exposing these factors (everything we invoke will eventually turn out relevant), we will finally be ready to conceive the jucier aspects of how the aural imitation method works and how imitating recordings prepares for collective music-making.

5.1 Generalized Interactions and the Indicative Role of Joint Attention

To see what resources Table 1 and Table 2 cultivate into joint musical attention and fluent use of the tonal language, we need to invoke perspectives explaining how the ability of symbolic behavior starts showing itself and how it forms into linguistic capacities. Considering the syntactical dimension of the tonal system, we also need aspects belonging to the learning of linguistic grammar.

In the developmental narrative introduced by Trevarthen[32] and Stern,[33] symbolic behavior begins to show itself soon after birth. That unity in rhythm and forms of vitality emerge across the many sense modalities indicates a certain selfsameness or generality in the constant variance of life is indeed a protoversion of symbolic behavior. Besides, rhythm and forms for vitality, too, have directions. They *lead* somewhere; they indicate.

However, encircling the mode of protosymbolic generality further (in a way that will prove fruitful to capture the generality of aural learning), Stern suggests the term RIG, an acronym for "Representations of Interactions that have been Generalized": "[I]nfants have some abilities to abstract,

average, and represent information preverbally."[34] From two to seven months, together with the formation of the core self, they begin to show capacities to aggregate experience and distill, or abstract out, an average prototype out of perceptual variety. "RIGs are flexible structures that average several actual instances and form a prototype to represent them all."[35]

RIGs have a somewhat ambiguous status when it comes to representing brute reality. On the one hand, the distilled prototype represents *all* the events that made it come into being. On the other hand, the prototype does not represent *any* of the events as such. That is, the RIG does not necessarily correspond to anything in a one-to-one sense: "A RIG is something that has never happened before in exactly that way, yet it takes into account nothing that did not actually happen once."[36] The ambiguity regarding "realness" makes up a flexible and generative moment. Things and events are indexed and reindexed in a fluid and dynamic fashion. Attributes of many different kinds gradually form meaningful networks. Invariants emerge in the constant variance of perceptual life.

Now, put simply, RIGs are protostructures for *words*. When the child is eighteen to twenty-four months and begins to use and understand linguistic symbols, this process is a continuation of the generalization that has been going on for a while. Language fills in the need for more advanced representations and communication of experience across the constant change of perceptual singularity.

Crucially, the ability for joint attention is critical for the process from RIGs to linguistic symbol manipulation. In Trevarthen's framework, joint attention represents a shift from *primary* to *secondary intersubjectivity*, emerging around nine months.[37] At this age, the infant moves from coordination of self and others based mainly on timing, form, and intensity, to the inclusion of objects and more explicit engagement in cooperative exchange of referential gestures. The infant shows increased initiative-taking to the systematic combining of purposes to partner and object. The infant begins to generate meaningful acts in a new sense, such as rudimental demands, refusals, and inquiries, or more awareness of objects. "We are born to generate shifting states of self-awareness, to show them to other persons, and to provoke interest and affectionate responses from them."[38]

In Stern's framework, the process of sharing attention starts around seven to nine months, when the infant develops a new sense of self, which Stern calls the *subjective self*.[39] The infant now begins to show a new awareness of self vis-à-vis others. It "discovers" that there are other minds out there as well as its own.[40] Self and other are no longer only core entities of physical presence, action, affect, and continuity; they also include mental states, such as feelings, motives, and intentions, things that lie beyond the physical happenings in the domain of core-relatedness. The infant shows a new organization of the subjective perspective, defined by a qualitatively new sense of

self vis-à-vis the other. Mental states can be "read," matched, aligned with, or attuned to, in a more articulate sense. The infant shows capacities for sharing a focus of attention, for attributing intentions and motives to others and apprehending them correctly, and for attributing the existence of states of feeling in others and sensing whether they are congruent with one's states or feelings.[41]

The gesture of pointing and the act of following another's line of vision are among the first overt acts that permit inferences about the sharing of attention or the establishment of joint attention.[42] While the infant shows a preliminary form of the ability to follow the gaze directions of others *before* nine months, the ability to share attention suggests a new ability to perceive *pointing* matures around that age. The child still maps out affectively and *very closely* the behavior of the mother but also begins imitating the *goal* of the mother's actions in another sense. "To imitate is not to do what the other does, but to arrive at the same result," as Merleau-Ponty would say.[43] When the mother looks in one direction and the child does the same, the child does not *copy* the movement as such but imitates the aim of the looking gesture, which is to attend to the same as the mother.

Interestingly enough, recent studies accentuate precisely the auditory aspects of this early ability to share attention. Launching the term *teleomusicality*, Schiavio, van der Schyff, Kruse-Weber, and Timmers demonstrate how infants between six and ten months begin to *aim* toward something *in* the sounds of the surrounding.[44] The goal seems to be both to create a nonboring and meaningful environment for its own sake *and* to share this auditory sense with the mother. By trying to imitate the sounds of the mother, infants develop their repertoires of goal-directed actions that will allow them to explore the environment in a meaningful (for example, musical) way. At the same time, infants respond creatively to the sonic situation. By mastering certain actions, they also develop adequate perceptual abilities that seem to motivate further ways of interaction with the world. That is, by understanding the goal of a given sound-related activity performed by another individual (for example, the caregiver or peer), infants could typically begin to play with their own sounds as invitations to more mutual understanding embedded in the auditory phenomena.[45]

Joint attention implies both ability to transcend egocentrism by decentering attention into the interest and purposes of the mother and, the other way around, the ability to steer the attention of the other, thus expanding the "control zone" of the ego.[46] Differently put, together with the ability to *follow* another person's attention to distal objects and events outside immediate interaction comes the reciprocal ability to *direct* the attention of others to distal objects by pointing, showing, and using other nonlinguistic gestures.[47]

The child engages in the process: it validates whether joint attention has been achieved, and if it has failed, the child will initiate more interaction to

gain the joint perspective. Thus, joint attention is not passive reception but active participation *with*, *against*, and *according to* the other's way of perceiving things. The child begins a dialectic mediation of perspective. It becomes a subject in an interpersonal exchange of perspectives. In other words, a new dimension emerges in what we call the intermediary third. The thirdness takes up the distinct perspectives of partners.

In Stern's framework, the leap into the mutuality of *my* and *your* attention indicates the rudimentary formation of a new self, *the verbal self*.[48] The indicative dimension of joint attention is the key to the development of what later will turn into an explicit understanding of verbal symbols. It is also at the core of understanding generative rules and procedures for interactions.[49] Within the framework of affective attunement and mutually created meaning, the child will gradually recognize stable traits in "how we do things." It will begin to identify complex invariants in behavior and to perceive how these invariants contain latent potentialities of novel behavior. It will become aware of how overt behavior is one of several possible manifestations of the same, as Stern puts it.[50] In other words, the child will become aware of how something can symbolize something else beyond the concrete reality of RIGs. It will perceive how each manifestation of something has some degree of substitutability and potential variability latent in how others perceive the same object.

Naturally, it falls outside the scope of this article to pursue further how the child acquires verbal language of words and grammar. Suffice to say that what Tomasello calls early grammaticalization is rendered possible by joint attention.[51] Gradually and then from inside the rhythmic communication full of RIGs and exchange of gaze and expressive gestures, grammatical and generative structures of the mother tongue will emerge for the child as organizing forces embedded in communication. The child will need no meta-awareness of the syntax qua syntax but will eventually pick up general ways to generate utterances. He or she will follow the direction of the gesture—the sense of the movement, as Merleau-Ponty would say: "[B]eginning with the first phonetic oppositions, the child *speaks,* and only afterward will he learn to apply the principle of speech in diverse ways."[52] From the inside of the phonetic dimension of communicative sense, the child will begin to apply general principles embedded in the communication.

5.2 Guided by Music

As we return to the aural imitation method, we recall first how the student and teacher collaborate in the process of exploring the music. By listening together, they genuinely share attention. They mutually guide the other's attention toward nuances in the music, either by saying things like, "Listen to how the saxophone phrases here, compared to here," or raising an eyebrow after a hefty passage or just by singing or playing what they have

in mind. They can indicate or point toward specific nuances in the music, accumulating a growing auditory awareness within the shared musical language.

On a more intriguing and potent level, however, it now makes sense to say that the student also shares attention with the musicians she listens to on the record—not directly, of course, as if they were standing next to her, *but in a mediated sense*. That is, she does not necessarily engage in empathetic attunement with the musicians (as Seddon would suggest[53]), nor does she use "mental imaginary" to plan the productions of her own sound or to predict the upcoming sounds of the players heard on the record.[54] She "just" listens *within* the same musical language as the Adderley combo once used in live communication. She lets herself be guided by musical indications embedded *in* the music as an intermediary third. Gradually, she can become aware of indications unfolding in subtler details of the music, and through the enactive efforts, she can become capable of using similar musical indications. Finally, we see also why the student could have picked a solo performance as a model for imitation. Solo performances unfold a language of indications.

Let's say our student has imitated and incorporated (very precisely) how the bass plays the root in the falling fifths relations of "Autumn Leaves" and how the piano, trumpet, and saxophone voices out the other tones of the chords vertically and horizontally (Acts 6–7, Table 1, Phase 1), before she now begins to vary the voicings according to Instr. 5 (Table 2, Phase 2).

The current point is this: The student does not perform these forms of enaction in a void but *relative* to the music once carried out by Adderley and his combo. She directs her attention toward the same music as they did at the time of recording. But she also trains her ability to hear and accomplish other latent possibilities in the same music. She tries to hear not only how the generative potentials of "Autumn Leaves" once were heard and fulfilled in a recording studio in New York but also how the potentials *could* have been heard and carried out. In other words, she does not *invent* new options for the musical language as much as she tries to fulfill possibilities latent in the musical language. These "dormant" or "quiescent" pathways are already part of the musical organization. All she has to do is bring them forth by hearing them and acting them out.

In our framework, it makes sense to say that the ability to hear distal musical potentialities is rendered possible by the protosymbolic imitation of joint attention. Without the ontogenetic formation of joint attention, it is improbable that the student would be able to follow indications within a complex tonal system. Without the ability to understand the directing of pointing, she would probably not be able to hear unaccomplished, intrasystematic leads unfolding in the musical language. If she had not experienced

the mutual dialectics creating an intermediary third, she would not have had the ability to hear and generate musical potentials in real-time polyphonic synchronization with other subjects.

At the same time, it makes sense to say that the aural imitation method implies a further development of the student's general capacity to follow the attention of others. She is "forced" to expand on her abilities to follow subtle indications embedded in the intersubjective symbolic language. The music pushes her into a general regrouping in her ways of perceiving general sense latent in the concrete auditory unfoldment.

In prolonging our previous observations regarding rhythmic perception and understanding, we note how the auditory attention toward the harmonic sense also harbors a potential proprioceptive dimension. Discovering new pathways in the music is coextensive with the student's discovery of new channels in her behavior. The energy, spontaneity, and self-relation associated with forms of vitality are about to be channeled into the symbolic sense of the tonal language. Multimodal behavior is about to acquire a new medium of expression within the stringent pathways of the tonal language. She has to discover and establish new latent behavior within herself, letting adjustments of motor excitation participate in the harmonic generative potentials of the music.

5.3 Learning the System as a Whole

As we know from the previous sections, it is one thing to hear and enact local or partial latencies belonging to the tonal language and another to have at one's independent disposal the whole tonal language in the ways associated with tonal fluency. How, then, do the exercises of Tables 1 and 2 help the student crack the system as a whole? Alternatively, put in a quasi-Saussurean way, how can she construct a tonal langue ready to be transformed into musical parole at any time, without the support of an external audible source like the record or the instrument?

One key evolves in Stern's RIG concept. Recall how RIG conceptualizes how learning, already from the early and protolinguistic phase of ontogeny, involves a certain generality embedded in the concrete. That is, long before she learned to handle general categorizations of words and linguistic symbols, the student had begun to distill, or abstract out, average prototypes out of perceptual variety. She had started to form flexible structures that averaged several actual instances and formed a prototype that represented all the pertinent situations.

In the aural imitation context, something similar must somehow take place. The student must hear that the partial tonal sequences and generative potentials that she imitates embed a generality—namely, the generality of the *whole* tonal system. Although it is utterly impossible to hear through *all*

possible pathways in the tonal language, it *is* possible to grasp the general functionality of the whole system.

The harmonic falling fifth relation characterizing "Autumn Leaves" gives the student (and us) a hint. By what we said above, this harmonic progression makes use of one of the characteristic features of the Western tonal system: the circle of fifths. Simply put, "Autumn Leaves" makes abbreviated, musical use of precisely this trait of the system.

Herein unfolds the generality that the student somehow must grasp in the audible phenomenon: she needs to hear how the concrete use of generative potentials manifests a systematic *part* of a systematic *whole.* That is, she needs to hear how the tune exemplifies or "symbolizes" the generative syntax of the major-minor tonality, the almost complete circle of fifths, the twelve tones, and the diatonic and chromatic transitions. In other words, she needs to hear how there are practically unlimited aural possibilities embedded in the limited sequences. The individual sequences and the generative potentials embed and exemplify the generative syntax of a system that ultimately transcends the concrete realizations of *that* tune.

While it is easy to overintellectualize the learning process involved in the potential cracking of the whole system, Stern's RIG helps us keep track of the perceptual and embodied mode of learning. The ways that intangible wholeness of the system is present in "Autumn Leaves" is structurally similar to how each manifestation of a RIG represents the average or distillation of several of several experiences. For the student, the system as a perceptual whole is something that never actually can be heard (how could it be?), yet the concrete tune that she is about to imitate takes into account nothing that does not belong to the whole system. In other words, "Autumn Leaves" is what Stern would call a perceptual prototype. It is a distillation of the transgressing, imperceptible whole.

Exposing the RIG point further, we consider briefly a study discussed by Merleau-Ponty as our point of departure. Against the general background that children often learn to discriminate colors relatively late, a group of young children was asked to distinguish objects with a small set of colors.[55] The moment the children first learned to recognize and name two or three colors, the Sternian generality of RIGs kicked in. Suddenly, at some moment, the children were able to identify and discriminate *colors*:

> [W]hat is acquired is not properly speaking the discrimination of these two qualities as such; it is a general power of comparing and distinguishing *colors*: all pairs of colors benefit from the distinction of red and green and differential behavior progress not from one to the other, but by a finer discrimination with regard to all of them.[56]

In the same strike of understanding, the children had learned to *see* more distinctions and to name these distinctions. The color RIG is generalized,

nuanced, and "symbolized," so to say. The combination of grown-up guidance ("look here and try to distinguish this from that") and a limited number of examples helped the children broaden their sensual, qualitative, and symbolic form of perception.

Now, just like the children learning to sort colors, the jazz student has narrowed down the options of variations. The children in Merleau-Ponty's case handle *three* colors, not endlessly many colors, and *this* limitation helps them understand the perceptual generality *exemplified* by the colors. Similarly, the student handles *one* recorded tune organized by the Western tonal system (not many recordings at once), and *this* specific limitation is the vehicle that potentially will make her able to perceive the system as a whole. Moreover, in analogy to Merleau-Ponty's case, this general learning will go hand in hand with an increased awareness for perceptual details. Just like the child learning to *differentiate* between colors, the student will begin to distinguish within the tonic atmosphere in more precise manners. Diatonic organizations harbor chromatic transitions, and V–I sequences harbor a potential II–V7, or whatever.

In contrast to Merleau-Ponty's case, however, the generality potentially grasped by the student is a symbolic network of *sense generating sense*, that is, of a complete, generative syntax. By the stringent constraints discussed above, each new discrimination within the system will potentially lead to more nuances regulated by the same order. Moreover, in contrast to Merleau-Ponty's case, the ultimate guide to this latent generality unfolds in the perceptual phenomenon itself, that is, *in the music*. Adderley's recording harbors its own explanation: it contains indicative forces powerful enough for the student to crack more or less the whole system.

The indicative role of the music was implied when we said that the student joins attention with the recorded musicians *in a mediated sense*. Their use of the tonal language is indicative: *how* the musicians unfold the concrete generative potentials can lead the student's ear in the direction of the general invariant syntax of the tonal language. Analogous to how the student as a child once picked up the invariants of linguistic grammar without necessarily being *told* that "this is grammar," she can pick up the qualitative indications toward generality unfolding in the flow of the music.

Let's say the student has yet to make the final discovery. She manages to swing and hear and generate relevant tonal sequences but has not cracked the system as a whole. This will come in the next section.

6. Subjective (Personal) Variation

Before we say more, we turn to the final stage of the aural imitation method (Table 3, identical to Table 3 presented in Part 1).

Table 3. Phase 3: Subjective (personal) variation

Step	Description	Tool
Pers. 1	Let the basic rhythmic foundation (Act 1b) sound in your ears and sing, play, and beat new rhythmic ideas/patterns simultaneously. This must occur spontaneously. Vary the length of ideas from short motifs to longer themes. In this exercise, play harmonically and melodically freely. Make recordings of your own playing. What do you like and dislike?	Instrument, singing, clapping, foot-tapping
Pers. 2	Pick some melodic motifs from the recording. Sing this melody line along with the recording so many times that you know it by heart. Play it together with the recording. Play the same thing alone while the rest of the recording sounds in your ear. Sing and then play spontaneous melodic lines. Be sure to convey ideas from your own musical imagination. Record your playing along with the recording as well as your soloing. What do you think of your own playing? What do you want to change? Percussionists: Sing exercises in Pers. 2.	Instrument, singing, clapping, foot-tapping
Pers. 3	Concentrate on the harmonic progression of the recording. Sing the chord sequence you hear on the recording. Break the same chords in arpeggio exercises on your instrument. Let the rhythmic foundation of the recording go into your ear while you spontaneously play new chords. Preferably, piano is used, but singers and melodic instrumentalists sing/play broken chords in inversions of their own choice. Set your harmonic imagination free.	Instrument, singing, clapping, foot-tapping

6.1 Enhancing the Process Even Further

Again, everything encouraged in this third phase of the aural imitation method is a further development of the process started with the earlier stages presented in Tables 1 and 2 (partly discussed in Part 1, and partly in the Sections 4–5 above.). We pay extra attention to the following features.

6.1.A. STRENGTHENED MENTAL HEARING

The student enhances the facility to let the music resound in her mind. She pays attention to and explores how the mental or "intrabodily" music sounds relatively independent of the real voice and the instrument. She tries it out in cross-modal behavior, acted out relative to stamping, clapping, singing, and the use of the musical instrument.

6.1.B. STRENGTHENED PERSONAL VARIATION

More explicit than before, the student is now encouraged to act out personal variations within the musical language spontaneously. By doing so,

she stretches the limitations and possibilities of rhythm, melody, and harmony—not by thinking and cognitive evaluation but by impromptu enactment. Rhythmic and melodic motifs from the record function as springboards for embellishments of various lengths.

6.1.C. ENHANCED DYADIC CHARACTER OF ATTENTION

Viewed together, everything we have said here implies a strengthening of the dyadic structure of the auditory attention established above: The mode of attention going "inwards" (toward the self-produced mental music) and "outwards" (toward the self-generated musical variations hearable to others) are ways of releasing and hearing one's spontaneity *relative to* the music that streams out of the loudspeakers. The strengthened contact with the ultimately subjective intramental way of listening goes hand in hand with the intensified contact with the expressive behavior that blends into the intersubjective resounding music.

In other words, for the mode of aural attention encouraged in these exercises, the *personal* ways of enacting the music imply no conflict between the "inner" and the "outer"—just as there is no conflict between the self-generated music and the music streaming from the loudspeakers. The "inner" and the "outer" are not poles of expression but rather dimensions of the same musical sense being explored and expressed. Everything accomplishes the same musical language: Every mode of behavior is played out with, against, and in accordance with each other—and relative to the music that resounds from the record.

7. Leaping into Fluency

To analyze the implications of Table 3, we need not import many new theoretical perspectives but rather unpack how the potential outcome of the exercises accomplishes aspects belonging to an intersubjective relatedness that has been there all the way. In the introduction, we noted how musicality, according to Trevarthen,[57] is a communicated talent—a talent for communicating in live company. Through the critical investigations of Berliner's assumption, we have tried to indicate how intersubjective relatedness is part of every reasonably normal childhood in the forms of protoconversational, rhythmic interaction and the gradual formation of joint attention.

Based on what we have seen, it makes sense to suggest that imitating and incorporating the rhythmic and tonal forms of the music by ear works because the method stimulates, cultivates, and canalizes energies associated with the imitative relatedness of musicality. The imitative method utilizes, animates, and transforms general musical potential into a specialized, aural skill. The fundamental ability to move rhythmically and in self-generated behavior and the ability to hear and generate symbolic orders within the

tonal system were already there, formed and carried out elsewhere in the general ability of joint attention and linguistic facilities.

However, as we might remember, we left our student a little hanging in the previous section. She had begun to hear latent possibilities within the tonal language but had not crossed the symbolic threshold for real. She could hear latent pathways embedded in the tonal syntax but had yet to crack the tonal grammar of the whole. How, then, can the exercises described in Table 3 help the student take the leap into fluency? What, more specifically, would the leap have to do with the intersubjective relatedness which has been there all the way?

One crux of an answer evolves in the intimate relationship between the dyadic character of the aural attention and the indicative aspect of joint attention exposed in the previous sections. Recall how the formation of joint attention implied both a new ability to transcend egocentrism by the decentering of attention into the interest and purposes of the other and a more active ability to steer the other's attention and in that sense expand the "control zone" of the ego. In other words, the child learned to *be manipulated by* the behavior of the other and *to manipulate* the behavior of the other. One general goal of the manipulative activity was to achieve and sustain joint attention, which seems valuable for its own sake for the infant (and the mother).

In effect, striving with the exercises in Table 3, the student now seeks a similar mutuality. For as we just saw, when she varies the music according to her spontaneity, she *relates* to the music through her independent manipulation of it. At the same time, having her ears and body glued to minute details in the music streaming out of the loudspeakers, she picks up how the musical voices relate both to each other *within* the band and to herself as a potential listener. She picks up how subtle manipulations and negotiations go on in the group and how each initiative both singularly and collectively manipulates her spontaneous behavior. In other words, she becomes aware of the aural, dialectical encounter between self and other and how the intermediary music dynamically propels behavior in all parties.

This attended encounter of aural indications and reciprocal manipulations evolved what we could call the key to the tonal fluency and the general understanding of the tonal system.

7.1 Hearing via the Other

Let us say this alteration happens: Suddenly, the student *hears* the general, harmonic glue that holds the tonal organization together. She hears how the same tonal lawfulness organizes every contingent tonal initiative coming from herself and the others and how the generative lawfulness ultimately transgresses the contingent manifestations. She hears how *these* precise phrases and chord progression exemplify the general order of the tonal system *as such*.

Simultaneously, she realizes that her earlier attempts to enact the tonal orders came a little too much *from herself*. She was ego-centered, as Stern would say.[58] She listened to the tonal generative potentialities mainly from her subjective aural perspective. By contrast, after the transformation, she is more *alter-centered*. She relates differently to the whole music. She "takes in" how Adderley and his band relate to each other differently. This is the moment when she "falls into" the tonal fluency. The musical sense "comes to" her—just like verbal language comes to her in a spontaneous conversation with a listening and responding friend. What has happened?

To unpack the structure of the potential transformation, we turn briefly to the philosophical literature on understanding, which reflects the transformative encounter between self and others in ways that now seem pertinent. "It's enough to say that we understand in a *different way if we understand at all*," writes Gadamer.[59] Understanding *differently* means, for Gadamer, a genuinely nondirected openness for how other humans understand: it "involves recognizing that I myself must accept some things that are against me, even though no one forces me to do so."[60]

Merleau-Ponty makes a similar point. To understand another human in a conversation is to *speak via the other*:

> When I speak [and] understand, I experience the presence of others in myself or of myself in others. . . . To the extent that what I say has meaning, I am a different "other" for myself when I am speaking; and to the extent that I understand, I no longer know who is speaking and who is listening.[61]

The *other* discussed by Gadamer and Merleau-Ponty *can* be a real person, but it can also be a fictive, generalized other, located in oneself, so to speak.[62] The point is that the person who *understands* must decenter from a mere private way of conceiving things. Understanding implies a genuine openness to different ways of perceiving.

Gadamer and Merleau-Ponty stress how openness to the other implies the emancipation of what we, with Benjamin,[63] call the intermediary third. Where neither of the parties control the human encounter of perspectives but contribute and attune to the other's ways of understanding, the intermediary is allowed to play itself out, according to its intrinsic norm. The intermediary is accomplished in full-fledged symbolic form. Something is allowed to emerge, something that remains an invariant through an unlimited amount of perspectival variation.

Our previous sections have prepared for Gadamer's and Merleau-Ponty's points. The rhythmic interaction and forms of vitality implied implicit relational knowledge—a sense of direction in synchronized behavior. Joint attention conceptualized the sharing of perspectives, and the concept of joint musical attention exposed the structure of the mutual aural and indicative dialectics among fluent aural musicians. What Gadamer and

Merleau-Ponty currently help us indicate represents the accomplishment of the relational competence. The more mature ability to speak "via the other" is rendered possible by the early development. The formation of the verbal self and the capacity for intersubjective relatedness enables the more mature distribution of aural and musical sense.

At the same time, Gadamer and Merleau-Ponty will now help us encircle how understanding the general syntax of the tonal system and the ability to *hear* and generate tonal sense "via the ear other" are *genuinely aspects of the same transformation*.

In our context, it makes sense to say that, by attending to her spontaneous variations of the music heard in mind, the student is already relating to a fictional or depersonalized other—a way *one* could listen to the music. Simultaneously, she relates to the music produced by real others, which is to say, Adderley's combo. She actively relates to the auditory product emerging when these exact individuals once practiced joint musical attention mediated by the tonal language. So construed, the student is already engaged in a pluralism of distinct aural horizons. She actively relates to multiple highly specific ways that a human ear can enact the generative potentials of the Western tonal system.

Moreover, in the process of trying out personal variation, the allocentric, open, and nondirective mode of aural attention allows everything to blend. It becomes unclear *who* is listening and who is playing, as Merleau-Ponty would say. The student becomes a "different other" for herself, as everything becomes mediated through the tonal language.

Finally, the student hears something general *in* the constant variations. She understands how every concrete tonal organization pivots around or exemplifies something general—something that ultimately remains the same across every contingent tonal variation. This selfsameness is the general syntax of the tonal system. However, the general order is inseparable from the possible aural perspective—the *different* ways of hearing the generative potentials. And this is what the student now hears: how there are *infinite ways of listening embedded in the same tonal organization*. She perceives how the generative potentials are universal pools of possible auditory enaction—pools that cannot be "used up" because there will always be yet another way to listen within the syntax.

In other words, the student has understood the tonal syntax as a symbolic and intersubjective order that enables the current form of tonal communication.

7.2 No Building Blocks but Transformation and Differentiation

In this article, we have tried to describe the formation of an aural, embodied, and communicative knowledge, which is nonscriptural by nature. Nothing in the student's learning process can be fixed conceptually and passed over

in written or semi-written forms. Ultimately, it can only be communicated in music.

To conclude, we can note how our usage-based approach to improvisation differs from a theoretical model that dominates much contemporary research on jazz improvisation and jazz education. Many theorists believe it is a good idea to approach jazz improvisation as if the musical behavior was an activity of synthesizing discrete units. Call this the *building block approach*.

Berliner exemplifies the approach. When he reflects (briefly) on how imitation can prepare for improvisation, he chooses to pick up Nettl's famous metaphor,[64] describing improvisation as a combination and recombination of building blocks: "Many students begin acquiring an expansive collection of improvisational building blocks by extracting those shapes they perceive as discrete components from the larger soloes they have already mastered and practicing them as independent figures."[65] Wilf follows Nettl's and Berliner' path: "[I]mprovisation involves imitation insofar as it is a recombination of previously available building blocks created by other improvisators."[66] Philosopher Benson sees no trouble in doing more or less the same: "For improvisation is a sense of 'putting together.' One takes the basic rhythmic and chord structures of the genre in which one works, and puts them together in different ways."[67] In fact, across the broadest spectrum of theories with interest in improvisational and musical behavior, the same approach is used in various ways.[68]

The building-block approach is perhaps a residue of Hume's empiristic description of human perception,[69] or it may be a conception taken from the Western literary tradition, where the building blocks are like letters in a sentence. Either way, the leap into tonal fluency that we just tried to illuminate would equal the ability to execute rapid combinations of individual tones or sequences of tones. Our student would now be capable of producing a specific set of notes within the acquired framework of implicit and explicit rules. She would also be able to combine various rhythmic structures in different ways, perhaps even into novel combinations.

From the perspective of a perceiving subject, this way of describing the learning process might seem all right. Working with the exercises described above, the student can very well isolate distinct phrases and rhythmic patterns and "put them together" with other isolated elements in focus, thus having the sense of performing a synthesizing activity. However, from the reflected point of the usage-based approach elaborated in this article, the building-block approach is downright wrong. For, as we have seen, not one of the exercises described encourages a synthesizing activity. Not one of the tasks inspires the construction of a musical language by adding rhythm *plus* harmony *plus* tonal sequences or whatever. Instead, constructing the musical language was *always* about hearing musical *wholes* (or matrices) of rhythm, tonality, harmony, and individual style. From the vantage point of

the ear, the generative potentials make sense only as musical wholes or aural matrices.

Besides, in a fundamental philosophical sense supported by this article, *there is no such thing as discrete parts of music* or building blocks. The least fragment of a rhythmic pattern or tonal sequence presupposes and resounds the entire lifeform in which it originated. We have seen how rhythm organizes the first relationship, how the general ability to follow and accomplish tonal sense originates in the ontogeny of joint attention, and how rhythm and tonality have the longer histories in phylogenetic evolution. Each musical whole belongs to a larger whole, implying culture, family, the local language, and everything else that characterizes life with others. Thus, it is more apt to say that constructing a musical language by aural imitation is learning to discriminate within larger wholes, rather than approaching the process as a synthesis of parts into a whole. The differentiation also has an artistic and expressive dimension. To borrow phrasing from Welsh, "The very idea of 'being an individual' can only take place against the relief of a shared social world from which I seek to individuate myself."[70] Just as imitation prepares the child for the emancipative transformation within verbal language, the aural imitation method prepares for emancipative transformations from inside the musical sense.

Our imagined student now packs her horn and goes to meet, play, and improvise with her band. The precise musical language is hers—and it unfolds in the collective.

Notes

Thanks to Lars Sigfred Evensen, Ståle Finke, Bengt Molander, Hans Magnus Solli, Lasse Thoresen, Ingebjørg Seip, Roger Jeffs, Michael Duch, Pradeep A. Dhillon, and members of the Research Group for Aesthetics and Phenomenology at NTNU for highly valuable readings and commentaries on various drafts of the manuscript. Thanks to Eldbjørg Raknes and Vigleik Storaas for important background information on aural jazz musicianship. Double thanks to Njål Ølnes for commenting on an early draft of the manuscript and being a well-informed and inspiring discussion partner in the writing process. Thanks to the Faculty of Humanities at NTNU for financing the project.

1. Bob Mintzer, *12 Contemporary Jazz Etudes: B-Flat Tenor Saxophone* (Van Nuys, CA: Alfred Publishing, 2004); Bob Mintzer, *Playing Jazz Piano* (Van Nuys, CA: Alfred Publishing, 2004).
2. [Hal Galper,] "Hal Galper's Master Class—*The Illusion of an Instrument*," May 2, 2010, https://www.youtube.com/watch?v=y_7DgCrziI82010, accessed Aug. 9, 2021.
3. Maurice Merleau-Ponty, *Phenomenology of Perception*, ed. Donald Landes (London: Routledge, 2012), 144; Alva Noë, *Action in Perception* (Cambridge, MA: The MIT Press, 2006), 1.
4. Hal Galper, "Able Bodied—An Interview with Hal Galper," *JazzImprov Magazine* 3, no. 3 (2001): 3, Hal Galper, https://www.halgalper.com/interviews-2/jazzimprov-magazine-interview/, accessed Aug. 9, 2021.

5. The history of the Western tonal system is neatly interwoven with the Western literary tradition, as demonstrated by Richard Taruskin, *The Oxford History of Western Music*, vol. 1 (Oxford: Oxford University Press, 2005). However, in accordance with what was said in Section 1.4, Part 1, we set aside this perspective, approaching the musical system merely as an audible phenomenon.
6. Roger Scruton, *The Aesthetics of Music* (Oxford: Clarendon Press, 1997) 240.
7. Scruton, *The Aesthetics of Music*, 248.
8. James Barbour, *Tuning and Temperament* (New York: Da Capo Press, 1972); Ross W. Duffin, *How Equal Temperament Ruined Harmony (and Why You Should Care)* (New York: W.W. Norton, 2007).
9. Scruton, *The Aesthetics of Music*, 248ff.
10. Scruton, *The Aesthetics of Music*, 250.
11. Martine Van Puyvelde et al., "Tonal Synchrony in Mother–Infant Interaction Based on Harmonic and Pentatonic Series," *Infant Behavior and Development* 33 (2010): 387–400, https://doi.org/10.1016/j.infbeh.2010.04.003, accessed Aug. 9, 2021; Beatriz Ilari, "Music and Babies: A Review of Research with Implications for Music Educators," *Update: Applications of Research in Music Education* 21 (2002): 17–26; Stanley N. Graven and Joy V. Browne, "Auditory Development in the Fetus and Infant," *Newborn and Infant Nursing Reviews* 8 (2008): 187–93, https://doi.org/10.1053/j.nainr.2008.10.010, accessed Aug. 9, 2021; F. Hicks, "Theatre Nursing: The Power of Music," *Nursing Times* 88 (1992): 72–74.
12. Jon Roar Bjorkvold, *The Muse Within: Creativity and Communication, Song and Play from Childhood through Maturity* (New York: HarperCollins, 1992).
13. Paul Franklin Berliner, *Thinking in Jazz* (Chicago: University of Chicago Press, 1994); David Schroeder, *From the Minds of Jazz Musicians: Conversations with the Creative and Inspired* (New York: Routledge, 2018); Eitan Y. Wilf, *School for Cool: The Academic Jazz Program and the Paradox of Institutionalized Creativity* (Chicago: University of Chicago Press, 2014).
14. Louis Cavrell, "The Universal Mind of Bill Evans," TV documentary, 1966, https://www.youtube.com/watch?v=QwXAqIaUahI1966, accessed Aug. 9, 2021.
15. Berliner, *Thinking in Jazz*, 94.
16. Daniel N. Stern, *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology* (London: Karnac, 1998), 162ff.
17. Stern, *The Interpersonal World*, 162.
18. Mark Reybrouck, "Musical Creativity between Symbolic Modelling and Perceptual Constraints: The Role of Adaptive Behaviour and Epistemic Autonomy," in *Musical Creativity: Multidisciplinary Research in Theory and Practice*, ed. Irène Deliège and Geraint Wiggins (Hove, Sussex, UK: Psychology Press, 2006), 58–76; Reybrouck, "Music as Environment: An Ecological and Biosemiotic Approach," *Behavioral Sciences* 5, no. 1 (2015): 1–26.
19. Terrence William Deacon, *The Symbolic Species: The Co-Evolution of Language and Brain* (London: Norton, 1997); Hans-Georg Gadamer, *Truth and Method*, trans. Joel Weinsheimer and D. G. Marshall (London: Continuum, 2004); Maurice Merleau-Ponty, *The Structure of Behavior*, trans. A. L. Fisher (Pittsburgh, PA: Duquesne University Press, 2011); Daniel N. Stern, *Forms of Vitality* (Oxford: Oxford University Press, 2010); Michael Tomasello, *Constructing a Language: A Usage-Based Theory of Language* (Cambridge, MA: Harvard University Press, 2003); Tomasello, *Origins of Human Communication* (Cambridge, MA: MIT Press, 2010).
20. Evan Thompson, *Mind in Life* (Cambridge, MA: The Belknap Press of Harvard University, 2007), 76; Italics original.
21. While Reybrouck's computational framework is largely distinct from ours, he too describes music in terms of indication; see Mark Reybrouck, "Music as Environment," and Reybrouck, "Music Cognition and Real-Time Listening: Denotation, Cue Abstraction, Route Description and Cognitive Maps," *Musicae Scientiae* 14 (2010): 187–202.
22. Thompson, *Mind in Life*, 76.

23. Hal Galper, "The Oral Tradition," Hal Galper, http://www.halgalper.com/articles/the-oral-tradition/2012/, accessed Aug. 9, 2021.
24. Ben Sidran, *Black Talk: How the Music of Black America Created a Radical Alternative to the Values of Western Literary Tradition* (New York: Payback Press, 1981), xiv.
25. Galper, *The Oral Tradition*; Mellonee V. Burnim and Portia K. Maultsby, eds., *African American Music: An Introduction* (New York: Routledge, 2014); Simha Arom, *African Polyphony and Polyrhythm: Musical Structure and Methodology* (Cambridge: Cambridge University Press, 1991).
26. Tomasello, *Constructing a Language*, 3–4.
27. Merleau-Ponty, *Phenomenology of Perception;* Charles Keil, "Participatory Discrepancies and the Power of Music," *Cultural Anthropology* 2, no. 3 (1987): 275–83; Tiger Roholt, *Groove: A Phenomenology of Rhythmic Nuance* (New York: Bloomsbury Publishing, 2014).
28. For compatible descriptions of polyphonic openness in real-time aural music-making, see Berliner, *Thinking in Jazz*; Ingrid Monson, *Saying Something: Jazz Improvisation and Interaction* (Chicago: University of Chicago Press, 1996); LeRoi Jones, *Blues People: Negro Music in White America* (New York: Harper Perennial, 2002); George Lipsitz, "Improvised Listening: Opening Statements Listening to the Lambs," in *The Improvisation Studies Reader: Spontaneous Acts*, ed. Ajay Heble and Rebecca Caines (New York: Routledge, 2014), 27–34.
29. Frederick Seddon, "Modes of Communication During Jazz Improvisation," *British Journal of Music Education* 22 (2005): 47–61, https://doi.org/10.1017/S0265051704005984, accessed Aug. 9, 2021. Frederick Seddon and Michele Biasutti, "A Comparison of Modes of Communication between Members of a String Quartet and a Jazz Sextet," *Psychology of Music* 37, no. 4 (2009): 395–415, https://doi.org/10.1177/0305735608100375, accessed Aug. 9, 2021. Jessica Phillips-Silver and Peter Keller, "Searching for Roots of Entrainment and Joint Action in Early Musical Interactions," *Frontiers in Human Neuroscience* 2, no. 26 (2012), https://doi.org/10.3389/fnhum.2012.00026, accessed Aug. 9, 2021.
30. Andrea Schiavio and Simon Høffding, "Playing Together without Communicating? A Pre-Reflective and Enactive Account of Joint Musical Performance," *Musicae Scientiae* 19, no. 4 (2015): 366–88, https://doi.org/10.1177/1029864915593333, accessed Aug. 9, 2021.
31. Scruton, *The Aesthetics of Music*.
32. Colwyn Trevarthen, "Musicality and the Intrinsic Motive Pulse: Evidence from Human Psychobiology and Infant Communication," *Musicae Scientiae* 3, no. 1 suppl. (1999): 155–215; Trevarten, "First Things First: Infants Make Good Use of the Sympathetic Rhythm of Imitation, without Reason or Language," *Journal of Child Psychotherapy* 31, no. 1 (2005): 91–113, https://doi.org/10.1080/00754170500079651, accessed Aug. 9, 2021
33. Daniel N. Stern, "Face-to-Face Play: Its Temporal Structure as Predictor of Socio-affective Development," in *Rhythms of Dialogue in Infancy: Coordinated Timing in Development*, ed. Joseph Jaffe et al. (Boston: Wiley/Society for Research in Child Development, 2001), 144–49, 147; Stern, *The First Relationship* (Cambridge, MA: Harvard University Press, 2002); Stern, *The Interpersonal World of the Infant*.
34. Stern, *The Interpersonal World of the Infant*, 97.
35. Stern, *The Interpersonal World of the Infant*, 110.
36. Stern, *The Interpersonal World of the Infant*, 110.
37. Colwyn Trevarthen and Penelope Hubley, "Secondary Intersubjectivity: Confidence, Confiding and Acts of Meaning in the First Year," in *Action, Gesture and Symbol: The Emergence of Language*, ed. A. Lock (London: Academic Press, 1978), 183–229; Penelope Hubley and Colwyn Trevarthen, "Sharing a Task in Infancy," *New Directions for Child and Adolescent Development* 1979 (1979): 57–80; Beebe et al., "A Comparison of Meltzoff, Trevarthen, and Stern."
38. Colwyn Trevarthen, "What Is It Like to Be a Person Who Knows Nothing? Defining the Active Intersubjective Mind of a Newborn Human Being," *Infant & Child*

Development 20 (2010): 119–35, 119, https://doi.org/10.1002/icd.689, accessed Aug. 11, 2021.
39. Stern, *The Interpersonal World of the Infant*, 26–34 and 124–61.
40. Stern, *The Interpersonal World of the Infant*, 27.
41. Stern, *The Interpersonal World of the Infant*, 27
42. Stern, 129.
43. Maurice Merleau-Ponty, *Child Psychology and Pedagogy: The Sorbonne Lectures, 1949–1952*, trans. Talia Welsh (Evanston, IL: Northwestern University Press, 2010), 22.
44. Andrea Schiavio, Dylan van der Schyff, Silke Kruse-Weber, and Renee Timmers, "When the Sound Becomes the Goal: 4e Cognition and Teleomusicality in Early Infancy," *Frontiers in Psychology* 8, no. 1585 (Sept. 25, 2017), https://doi.org/10.3389/fpsyg.2017.01585, accessed Aug. 9, 2021.
45. Schiavio et al., 6.
46. Stern, *The Interpersonal World of the Infant*, 129.
47. Phrase borrowed from Tomasello, *Constructing a Language*, 3.
48. Stern, *The Interpersonal World of the Infant*, 162ff.
49. Stern, *The Interpersonal World of the Infant*, 133.
50. Stern, 161.
51. Tomasello, *Constructing a Language*, 8–17.
52. Maurice Merleau-Ponty, *Signs*, trans. Richard McCleary (Evanston, IL: Northwestern University Press, 1987), 40.
53. Seddon, "Modes of Communication during Jazz Improvisation."
54. Phillips-Silver and Keller, "Searching for Roots."
55. Merleau-Ponty, *The Structure of Behavior*, 96–97. Merleau-Ponty cites Kurt Koffka, *The Growth of the Mind* (New York: Brace, 1925).
56. Merleau-Ponty, *The Structure of Behavior*, 96–97.
57. Trevarthen, "Musicality and the Intrinsic Motive Pulse," 158.
58. Stern, *The Interpersonal World of the Infant*, 124–37.
59. Gadamer, *Truth and Method*, 296. Italics original.
60. Gadamer, *Truth and Method*, 355.
61. Merleau-Ponty, *Signs*, 97.
62. Stein Bråten, "Participant Perception of Others' Acts: Virtual Otherness in Infants and Adults," *Culture and Psychology* 9 (2003): 261–76.
63. Jessica Benjamin, "Beyond Doer and Done: An Intersubjective View of Thirdness," *The Psychoanalytic Quarterly* 73, no. 1 (2004): 5–46. See also Part 1, Section 3, of the current article.
64. Bruno Nettl, "Thoughts on Improvisation: A Comparative Approach," *The Musical Quarterly* 60, no. 1 (1974), 1–19, 13.
65. Berliner, *Thinking in Jazz*, 101.
66. Wilf, *School for Cool*, 134.
67. Bruce Ellis Benson, *The Improvisation of Musical Dialogue: A Phenomenology of Music* (Cambridge: Cambridge University Press, 2003), 136.
68. Aaron L. Berkowitz, *The Improvising Mind: Cognition and Creativity in the Musical Moment* (Oxford: Oxford University Press, 2010); Robert Gjerdingen, *Music in the Galant Style* (Oxford: Oxford University Press, 2007); Alfred Pike, "A Phenomenology of Jazz," *Journal of Jazz Studies* 50 (1974): 88–94; Jeff Pressing, "Improvisation: Methods and Models," in *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition*, ed. John Sloboda (Oxford: Oxford University Press, 1988), 129–78; Reybrouck, "Music as Environment."
69. David Hume, *Enquiries concerning Human Understanding and concerning the Principles of Morals* (Oxford: Clarendon Press, 1975).
70. Talia Welsh, *The Child as Natural Phenomenologist: Primal and Primary Experience in Merleau-Ponty's Psychology* (Evanston, IL: Northwestern University Press, 2013), 49.