

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 *Opplysningen 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.

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, unverb-
alized, 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 interhu-
man relations nor something irrational as such. It is instead a communica-
tive structure of behavior spontaneously unfolding between communicat-
ing human beings. It is *unrational*, *atheoretical*, and *nondirective*. 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 accord-*
dance with others, 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 imita-
tion 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 mu-
sician to develop his/her potential into the specialized skills of au-
ral jazz improvisation?

Berliner leaves these questions hanging, and no systematic study within
phenomenology, psychology, or pedagogy has picked them up. While phi-
losophers have pondered the relationship between music and language
for ages²¹ and while a vast and multidisciplinary literature is produced on
jazz improvisation, the intriguing phenomenon of learning how to impro-
vise 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 theorists²⁴ 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 other³⁷ 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 knowledge³⁹—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 Berliner⁴⁰ 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 *performs* 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 groove⁶⁵ 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 Greece⁸³ but also synergies between African tonality and Christian hymns occurring during the periods of slavery⁸⁴ 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	<p>a. Listen to the entire recording with imitating the rhythmic wholeness in mind.</p> <p>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.</p>	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	<p>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.</p> <p>At the same time, try to imitate vibrato, timbre, and other musical characteristics. If there is a text, learn this or some of it.</p>	Singing, clapping, foot-tapping
Act 5	<p>Sing melodic lines while beating/tapping the rhythmic wholeness that you found in Act 1b.</p> <p>Sing the melody lines while beating/tapping the rhythm of other instruments.</p>	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	<p>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).</p> <p>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.</p>	Singing, clapping, foot-tapping

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

Step	Description	Tool
Instr. 1	<p>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.</p> <p>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.</p>	Instrument, singing, clapping, foot-tapping
Instr. 2	<p>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.</p> <p>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.</p>	Instrument, singing, clapping, foot-tapping
Instr. 3 and Instr. 4	<p>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).</p> <p>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?</p> <p>Percussionists who do not play a melodic percussion instrument sing the exercises in Instr. 3, Instr. 4, and Instr. 5.</p>	Instrument, singing, clapping, foot-tapping
Instr. 5	<p>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.</p> <p>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).</p> <p>Make music from this exercise. No exercise should be just technical.</p>	Instrument, singing, clapping, foot-tapping
Instr. 6	<p>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.</p> <p>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.</p>	Instrument, singing, clapping, foot-tapping

Table 3. Phase 3: Subjective (personal) variation

Step	Description	Tool
Pers. 1	<p>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.</p> <p>In this exercise, play harmonically and melodically freely. Make recordings of your own playing. What do you like and dislike?</p>	Instrument, singing, clapping, foot-tapping
Pers. 2	<p>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.</p> <p>Play the same thing alone while the rest of the recording sounds in your ear. Sing and then play spontaneous melodic lines.</p> <p>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?</p> <p>Percussionists: Sing exercises in Pers. 2.</p>	Instrument, singing, clapping, foot-tapping
Pers. 3	<p>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.</p> <p>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.</p> <p>Set your harmonic imagination free.</p>	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, unverbilized, 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 say¹⁰⁸—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 Evans¹¹² and Marsalis¹¹³ 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, unverbilized, 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 Norwegica* 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 Scientifcae Musicae Scientifcae* 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, *Musophililia: 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 Improve," *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, Dorianne 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. Göransson 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 *Opplysningen 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-à-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: Columbia 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 Gamemeldans 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=QwXAqlaUahI1966>, 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, *Phenomenology*; 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 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.
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.