**Vijay Iyer**

# Microstructures of Feel, Macrostructures of Sound: Embodied Cognition in West African and African-American Musics

**Ph.D. Dissertation, University of California, Berkeley.**
**1998**

Original Table of Contents (with hyperlinks) from the online version




http://archive.cnmat.berkeley.edu/~vijay
http://www.vijay-iyer.com

## List of Audio Examples

A compact disc containing the audio examples cited throughout the text is available at the *Center for New Music and Audio Technologies* (CNMAT), 1750 Arch Street, Berkeley, CA 94720, (510) 643-9990 x 300. However, the text is understandable without this supplementary material. All musical selections are small excerpts from original recordings, unless otherwise indicated. Selections in parentheses were created by the author.

| **Track#** | **chapter** | **page #** | **artist or (example)** | **title** |
|---|---|---|---|---|
| 1. | Ch. 2 | 17 | Art Tatum | "This Could Be Love" |
| 2. | Ch. 3 | 21 | James Brown | "Give it Up or Turn it Loose" |
| 3. | Ch. 4 | 37 | Thelonious Monk | "Misterioso" |
| 4. | | | | "Let's Call This" |
| 5. | | | | "Four in One" |
| 6. | | | Charlie Parker | "52nd Street Theme" |
| 7. | | | Thelonious Monk | "Monk's Point" |
| 8. | | | | "Light Blue" |
| 9. | | | | "Gallop's Gallop" |
| 10. | | | | "Trinkle, Tinkle" |
| 11. | Ch. 4 | 48-49 | John Coltrane | "Spiritual" |
| 12. | Ch. 4 | 50 | Anlo Ewe drumming | kaganu pattern |
| 13. | Ch. 4 | 52 | Cecil Taylor Orchestra | rehearsal excerpt |
| 14. | Ch. 4 | 58 | Eric Dolphy | "Out There" |
| 15. | | | Ornette Coleman | "Check Out Time" |
| 16. | | | Miles Davis | "In a Silent Way" |
| 17. | | | James Brown | "Talkin' Loud & Sayin' Nothing" |

| | | | | |
|---|---|---|---|---|
| 18. | | | De La Soul | "Down Syndrome" |
| 19. | | | De La Soul | "Wonce Again Long Island" |
| 20. | | | Busta Rhymes | "Do My Thing" |
| 21. | Ch. 4 | 63 | Duke Ellington | "Diminuendo & Crescendo in Blue" |
| 22. | | | John Coltrane | "Chasing the Trane" |
| 23. | | | Aruna Sayeeram | "Mahaganapatim" |
| 24. | Ch. 4 | 64 | James Brown | "Get Up (I feel like being a) Sex Machine" |
| 25. | | | John Coltrane | "Transition" |
| 26. | Ch. 4 | 65 | T. Monk feat. Coltrane | "Trinkle, Tinkle" |
| 27. | Ch. 4 | 69 | Ahmad Jamal | "Poinciana" |
| 28. | Ch. 5 | 79 | Los Muñequitos | "Fundamento Dilanga" |
| 29. | | | James Brown | "I Don't Want Nobody to Give Me Nothing" |
| 30. | Ch. 5 | 89-90 | (Anlo-Ewe bell pattern) | |
| 31. | | | (Yoruba bell pattern) | |
| 32. | | | (other bell pattern) | |
| 33. | | | (Yoruba bell pattern, grouped) | |
| 34. | Ch. 5 | 90 | (3:4 composite rhythm) | |
| 35. | Ch. 5 | 92 | (RTP examples) | |
| 36. | Ch. 5 | 100 | Mustapha Tettey Addy | "Oshika" |
| 37. | Ch. 6 | 111 | Los Muñequitos | "Oyelos de Nuevo" |
| 38. | Ch. 6 | 112 | (asynchrony part 1) | |
| 39. | | | (asynchrony part 2) | |
| 40. | Ch. 6 | 113-114 | (streaming part 1) | |
| 41. | | | (streaming part 2) | |
| 42. | Ch. 6 | 115 | (spreading part 1) | |

| | | | | |
|---|---|---|---|---|
| 43. | | | (spreading part 2) | |
| 44. | Ch. 6 | 116 | John Chowning | vibrato & auditory scene analysis demo |
| 45. | Ch. 6 | 118 | (swing part 1) | |
| 46. | | | (swing part 2) | |
| 47. | Ch. 6 | 118-119 | (backbeat part 1) | |
| 48. | | | (backbeat part 2) | |
| 49. | Ch. 6 | 121 | Thelonious Monk | "I'm Confessin'" *(complete)* |
| 50. | Ch. 6 | 122 | excerpt of same | |
| 51. | Ch. 6 | 123 | Ahmad Jamal | "But Not for Me" *(complete)* |
| 52. | Ch. 6 | 124 | excerpt of same | |
| 53. | Ch. 9 | 142 | John Coltrane | "Chasing the Trane" |
| 54. | Ch. 9 | 144 | Busta Rhymes | "Dangerous" |
| 55. | Ch. 9 | 145 | Squarepusher | "Journey to Reedham" |
| 56. | Ch. 9 | 145 | Terminator X | "Vendetta... The Big Getback" |
| 57. | Ch. 9 | 146 | Miya Masaoka | live duo with pianist Vijay Iyer |
| 58. | Ch. 9 | 147 | David Wessel | live duo with vocalist Shafqat Ali Khan |
| 59. | Ch. 9 | 147 | George Lewis | "Voyager Duo II" |
| 60. | Ch. 9 | 148 | Madonna | "Frozen" |
| 61. | Ch. 9 | 148 | James Brown | "Please, Please, Please" |

## Acknowledgements

I followed a circuitous path to this degree, and it would not have been possible without the tremendous support and guidance of the dissertation committee members. Since I met Professor David Wessel, he has been a wonderful friend, colleague, and mentor, as well as a fellow avid music buff. When I left physics in 1994 and was casting about in search of a career, he graciously took me on and introduced me to many exciting new ideas. It was his consistently innovative vision that led to the construction and execution of this interdisciplinary program. As is the case with his direction of the Center for New Music and Audio Technologies (CNMAT), my degree program exemplifies Professor Wessel's extraordinary ability to create novel, nurturing environments for new kinds of research. I am extremely fortunate and grateful to have enjoyed the benefits of his brilliance and generosity.

Professor George Lewis is a similarly visionary and inspiring individual. In addition to his awesome prowess as an improvising trombonist, he has initiated and facilitated a remarkable variety of interdisciplinary discourses on improvised music through his teachings and writings. A pioneer improviser, computer musician, theorist, and composer, not to mention an astute cultural critic, Professor Lewis has influenced my thought profoundly both as an artist and an academic. His clear guidance, advice, and friendship have proven invaluable throughout this process, and I especially thank him for his help with the manuscript itself. I also thank Professor Olly Wilson for instilling in me a true appreciation for the continuity between West African and African-American cultures. This notion became the foundation for the thrust of this work, and for that I am indebted to him. In addition, I am thankful for his input in the initial stages of the application process for this interdisciplinary program, and for his deft leadership of the qualifying exam. Overall he has been a source of wisdom and common sense not just in terms of content, but also with regards to negotiating the bureaucratic hurdles. Professor Ervin Hafter has served as my scientific conscience throughout this endeavor. I learned a great deal by observing his methods of inquiry in the hearing-sciences seminar. In terms of my own work, his commitment to rigor and his high standards kept me on my toes, but his generous spirit and his respect for my ideas saved me from getting raked over the coals. For his input and supportive comments I am extremely grateful. Professor Don Glaser has shown great open-mindedness and warmth over these few years. I am truly honored to have interacted with a scientific mind of his caliber, a Nobel laureate who has been on the cutting edge of a number of disparate research fields. Like the others on this committee, he is an eminent renaissance man with a wide range of interests, and has served as an inspiration for me.

At CNMAT in 1995-96 we had an ongoing discussion group in rhythm. One of the central individuals in this group was the bassist, composer, and computer scientist Jeff Bilmes. Jeff has influenced my work in profound ways, both as an artist and as a scientist. Some of the work here used his MIT masters thesis as a foundation, and my quartet from 1995 to 1997 featured his bass playing as a different sort of foundation. We have had a number of heated but productive arguments over the years about how best to articulate and represent these musical ideas that we hold intuitively to be self-evident. Matt Wright also took part in the rhythm research group and blessed us with a wonderful down-to-earth approach and utmost clarity. I wish also to thank CNMAT and its people for indulging me with space, support, and friendly vibes. The people include Adrian & Eleanor Freed, Sami Khoury, Cyril Drame, Amar Chaudhary, Ron Smith, and Edmund Campion, who was the other half of an interesting series of exploratory two-piano improvisations. I especially thank CNMAT administrators Richard Andrews and Leslie Delehanty for keeping the U.C. bureaucracy at bay and for making sure the checks came in on time. I am extremely grateful to Gibson Guitars, whose financial support via CNMAT made this entire effort possible.

In forming the ideas in this work, I have drawn a great deal from my experiences as a professional musician. Among the most important influences in this part of my life over the last four years is saxophonist Steve Coleman. In our many lengthy discussions, emails, and musical interactions, we exchanged a vast amount of information on a wide range of subjects. He helped direct and focus my thinking on rhythm (particularly on the notion of "feel"), improvisation, and intuition. His input has been as important as that of any of the committee members. I also wish to thank all the musicians I have worked with since 1992, including Liberty Ellman, Elliot Humberto Kavee, Jeff Brock, Brad Hargreaves, Rudresh Mahanthappa, Derrek Phillips, Will Power, Mohammed Bilal, Simone White, DJ Pause, Miya Masaoka, E. W. Wainwright, Robert Rhodes, Eric Crystal, Aaron Stewart, Sean Rickman, Anthony Tidd, Chander Sardjoe, Rahsaan Fredericks, Devin Hoff, J. D. Parran, Kevin Ellington Mingus, Kash Killion, Francis Wong, Harold Lee Yen, D Armous Boone, John Tchicai, Cecil Taylor, and many, many others. All of my musical collaborations have influenced my thinking for this thesis. I was fortunate to have a few colleagues in the field of music perception and cognition who were willing to discuss my work, including Ed Large and Bruno Repp.

Before I began this program, I enjoyed a healthy two years in the physics department. During this time I benefited from the wisdom and guidance of the late Professor Leo Falicov, the warmth and kindness of the department employees Anne Takizawa, Donna Sakima, and Mercy Wang, and important friendships with Lorin Benedict (with whom I enjoyed many extracurricular activities) and Rob Carpick. I also thank Professor Alan Portis for accomodating me as his GSI during my transition year.

I thank my immediate family (Raghu, Sita, Pratima, and Kurt) for being extremely supportive throughout this process. It is probably because of my parents that I did not leave graduate school entirely in 1994; it is also thanks to them that I had the perseverance to try something off the beaten path. My sister helped to remind me of who I was during my California years; I might have lost my moorings without the luxury of laughing and commiserating with her. I thank Kurt for his exemplary groundedness and stability.

And most importantly, I have no idea what would have become of me had it not been for one charmed moment in Sproul Hall when I met Christina Sunita Leslie. The time since then has brought countless moments of pure, unspeakable joy, which managed to ease the lengthy gestation of twin doctorates. For her immeasurable understanding, emotional support, companionship, balance, depth, and ongoing dedication to our life together, I will always feel utterly blessed.

# 1. Introduction

The fundamental claim of this thesis is that music perception and cognition are embodied activities. This means that they depend crucially on the physical constraints and enablings of our sensorimotor apparatus, and also on the sociocultural environment in which our music-listening and -producing capacities come into being. This claim shows a strong similarity to that of John Blacking (1973), who wrote, "Music is a synthesis of cognitive processes which are present in culture and in the human body: the forms it takes, and the effects it has on people, are generated by the social experiences of human bodies in different cultural environments." (Blacking 1973: 89) I shall present some further evidence in its support, by showing how exemplary rhythms of certain kinds of music may relate to such embodied processes. I shall argue that rhythm perception and production involve a complex, whole-body experience, and that much of the structure found in music incorporates an awareness of the embodied, situated role of the participant. The claim that music perception and cognition are embodied activities also means that they are actively constructed by the listener, rather than passively transferred from performer to listener. In particular, the discernment of entities such as pulse and meter from a given piece of music are not perceptual inevitabilities for any human being, but are strongly dependent on the person's culturally contingent listening strategies. In addition, I argue that certain kinds of rhythmic expression in what I shall call groove-based music are directly related not only to the role of the body in music-making, but also to certain cultural aesthetics that privilege this role.

The work in this thesis lies on the outskirts of most contemporary research in rhythm perception and cognition. In particular, it avoids the Pandora's-box searches for beat-finders (Large 1994) and models of expressive tempo variation (Todd 1989, Desain & Honing 1996, Clarke 1988), in favor of a focus on somewhat different elements of rhythm. The aforementioned work has tended to focus on models derived from European classical music, which are valid but do not apply as universally as they might claim. Many of the aspects of rhythm that I discuss here have direct relevance to the performance and perception of popular musics, dance musics, and music that lacks a concept of "score" or "composer." The musics that concern me most directly in the present work are those that have arisen from African cultures, both in their native and diasporic manifestations. To this end I have drawn extensively from the writings and teachings of the Ghanaian master percussionist C. K. Ladzekpo (1995), as well as the theoretical and empirical discourses of many other musicians working in these genres. This work thus represents a nexus of cognitive-scientific ideas with cultural and artistic considerations.

I wish to make clear immediately that I myself am a professional pianist, improviser, and composer, associated with the vast genre known as jazz. (The term "jazz" is an ambiguous and even controversial one, which I accept, for now, as referring to a certain African-American cultural model with hugely varied manifestations.) Many of the ideas in this thesis grew out of my experiences working in jazz, funk, and hip-hop bands over the last decade. I entered this field because I saw it as a way of studying what fascinated me the most about the music of which I am a part – namely, its rhythmic vitality. As I became acquainted with the rather young field of music perception and cognition, I became frustrated by the overwhelming incompatibility between the priorities of the music cognition research program and my own musical experiences. A majority of research on rhythm perception and cognition has either focused on a style of music that exists in a rather rarefied form in the world today, namely European classical music from the Renaissance period to the pre-modern period. If any sense of urgency is detected in these pages, it is most probably because of my desire to surmount this issue, which I see as a problem of scope. I hope the reader will indulge me in such instances.

This thesis is divided into somewhat independent sections. Chapter 3 consists of a summary of the twin theoretical frameworks of embodied and situated cognition. These relatively recent paradigms frame cognition in terms of the body, its physical and temporal situation, and its social and cultural moment. They provide the raw materials for a more expansive view of music cognition that involves all of these dimensions, as elaborated at length in chapter 4. I argue that these ways of framing music cognition complement the abstract, symbolic models set forth in this area of research so far.

The remainder of the thesis applies the framework of embodied musical cognition to a few different specific concerns. Chapter 5 is a discussion of meter, in which I study the crucial role in attention that meter can play, and address the cultural contingency of meter perception. Chapter 6 is a summary of some kinds of expressive timing that appear in groove-based musics, including some simple audio examples. These examples lead into Chapter 7, which touches on different models of the physical activity of rhythmic tapping, before summarizing a paper that was given at the 1997 International Computer Music Conference, co-authored by Jeff Bilmes, Matt Wright, and David Wessel. We set forth a rhythmic representation that allows explicit control of expressive microtiming in the presence of an isochronous pulse. It was software using this representation that facilitated the construction of the examples in Chapter 6. Finally, chapter 8 wraps up this project by exploring some of its implications.

We must begin with some preliminary groundwork. The following chapter, a definition of relevant terms, provides an opportunity to examine some basic problems that we encounter in a multidisciplinary study. In the course of defining the terms that are relevant to our studies of rhythm cognition, I point out some of the assumptions, implications, and limitations inherent in their multivalent usage.

## 2. Defining Terms

Given that this is an interdisciplinary work with a multiperspectival audience, it is necessary to assure ourselves that we are all talking about the same things. Cognitive science, itself an interdisciplinary field, is plagued by the slippage caused by multiple simultaneous meanings of commonly used terms. Philosophers and neuroscientists can barely agree on the meaning of the word "mind," the supposed primary object of study. Similarly, the commonly used terms of music theory and music psychology are clouded by imprecise connotations. Often their scientific meanings, vague as they are, are eclipsed by their colloquial and artistic usages. We must be careful and thorough when using such terms in a scientific context, especially when attempting to describe cross-cultural phenomena. Invariably, these terms have multiple dictionary meanings, multiple meanings implied by common usage, and multiple meanings agreed upon by a community of scholars in the field of music perception and cognition.

In this section I examine a number of terms that arise in the study of rhythm perception, cognition, and production. I will focus on the last body of definitions, and will attempt to remain aware of the slippage among these three kinds of meaning. The nature of an interdisciplinary work makes such semantic mismatches inevitable, as does the nature of an interdisciplinary field such as cognitive science.

**Cognitive science.** The field of cognitive science consists of an interdisciplinary study of the structures of the human mind. These structures include our sensory/perceptual apparatus, such as vision, audition, olfaction; internal mental processes such as language, thinking, reasoning and problem solving; motor control and the organization of skilled behavior such as speech and musical performance; memory; consciousness; attention; and many other aspects of mind. All of these subfields are clearly intertwined. Disciplines included are psychology, biology, neuroscience, philosophy, anthropology, linguistics, sociology, and computer science; more recently, the academic music world has devoted some of its resources to the study of the cognitive science of music.

Often the claim is made that cognition is information processing: "Cognitive science is the study of information processing, and insofar as a discipline studies that, then it is part of cognitive science." (Hardcastle 1996: 8) This claim is true in the most general sense, in that the mind is continually taking in information and dealing with it in some way. In the past, the further claim was made that the mind is merely an example of a formal symbol-manipulating device that executes mental programs, where any such device would do (e.g. von Neumann 1951). As such, it was assumed that one could ignore the physical and biochemical particulars of the brain. Theories of mental processes then consisted of abstract mathematical models that had the correct input-output characteristics and various functional or causal relations. This kind of theorizing suggested a fundamental dichotomy that was assumed between the body and the mind – a problematic dichotomy that dates back to Plato.

Eventually the neurosciences began to make enough advances that it became advantageous to adopt the perspective that an organism's cognition is intimately connected with its strategy for survival within particular ecological environments. Nervous systems are not von Neumann's idealized general-purpose computers. Rather, our neural architecture evolved to fulfill a certain range of specific needs and to facilitate certain activities. For example, the general organization of the brain is best explained if we view the nervous system as a generator of motor output. The cerebellum is connected almost directly to all areas of the brain – sensory transmissions, reticular (arousal/attention) systems, hippocampus (episodic memories), limbic system (emotions, behavior). All areas of our brain seem geared to coping with their functions as they pertain to problems of motor control. The brain is understood in enough detail that one can model them computationally at a fairly low level. (Hardcastle 1996: 6-7)

And furthermore, recent authors (e.g. Shore 1996) have viewed cognition as inextricably linked with the environment that gives rise to it. It has been suggested that "there is reason to suspect that what we call cognition is in fact a complex social phenomenon... 'Cognition' observed in everyday practice is distributed – stretched over, not divided among – mind, body, activity and culturally organized settings (which include other actors)... [Cognitive] 'activity' takes form differently in different situations." (Lave 1988: 1)

Overall, the field of cognitive science consists of research on parallel fronts: to name a few, we have the neuroscientific study of the structure of the brain, the psychological study of mental abstractions, and the socioanthropological study of the shaping of these mental structures by culture.

**Cognition.** In view of the above description, this term operates as a frame for a huge variety of activities and mental processes. It is occasionally distinguished from perception, to denote "higher"-level abstractions. The term also carries the implication that these processes can be described scientifically; i.e. the use of the term "cognition" references the discipline of cognitive science and all of its elaborations.

**Perception.** Perception typically refers to the activity of processing physical input (e.g. pressure waves, photons) into convenient abstractions (e.g. pitch/timbre, color). In the past, the term has been seen as a set of processes distinct from cognition, but in modern terms (in part due to the research agenda of the vast field of cognitive science) seen as subsumed under the umbrella term cognition. But this distinction is often blurred; hence we have a journal called *Music Perception*, and a Society for Music Perception and Cognition, both of which are concerned with the same research agenda. Perception and cognition are often used in ways that attempt to transcend the influence of culture. However, it has been argued more recently that what is commonly called "perception" should be viewed as a *practice* – an open-ended, intentional activity that is accomplished actively by the musical participants, while profoundly influenced by the perceivers social context. (Berger 1997, Bourdieu 1977) Perception should not be seen simply as a raw, sensory inevitability, like a sensation. Nonetheless, there are aspects of perception that are universal, attributable to the human sensory apparatus alone, and can be studied in this way. In the work that follows, I appeal to both sides of this incarnation of the nature/nurture debate.

**Cognitive model.** A cognitive model may comprise a "circle & arrow theory" of how some aspect of cognition is structured (e.g. information processing stages), or a set of equations with the proper input-output specifications and some internal structure that is believed to represent some aspect of cognition. In studying a cognitive model, one considers issues such as predictive power and model uniqueness. In other words, one examines whether the model can foresee any traits of the aspect of cognition it claims to govern, and also whether success of the model logically excludes other possible models with the proper I/O mapping.

**Representation.** A representation is nothing more than some way of organizing, manipulating, and storing information. Because of the overlap between cognitive science and computer science, a cognitive representation is often discussed in the same terms as a computational data structure, or a set of such structures – usually seen as a disembodied, symbolic abstraction, possibly shared by some group of computational or mental processes. The internal organization of a representation, i.e. its data structures and attributes, might be meant to reflect a theory of mind. In such cases, to propose a certain cognitive representation, say for rhythm, means to posit an assertion about cognition; in this sense a representation can be a kind of model. The use of representations is linked traditionally to the information-processing or "cognitivist" view of the mind.

Varela et al. (1991) describe the "cognitivist" view of cognitive science as follows: Intelligent behavior presupposes the ability to represent the world internally; an agent acts by representing relevant features from its environment; the success of the agent's behavior depends on the accuracy of its representations. While this may not seem terribly controversial, the added claim of cognitivism might: namely, that these representations are physically realized in the form of a symbolic code in the brain or in a machine. (Varela et al. 1991: 40)

More recently, Clancey (1997) has broken down this distinction by exploring the difference between a *mechanism* and a descriptive model. He says that "the descriptive modeling literature often equates knowledge, knowledge representations, representations, mental models, knowledge base, concepts..." (Clancey 1997: 50) but claims that instead, "we have a memory for coordinated, interacting processes, not for the descriptions of them per se. These processes correspond to the activation, recategorization, and coordination of perceptual-conceptual-motor sequences and other temporal relations, including rhythm and simultaneity." (*ibid*.: 68-69) The distinction is made between actual information and the description of that information, akin to the distinction between a city and a map of that city. Clancey also points out that descriptions are not the only form of representation involved in cognition, and storage is the wrong metaphor for memory. (*ibid*.: 221) As an alternative example he invokes the *self-organization with memory* of neurobiological systems, where the information lies in the physical organization of interconnected neural pathways, qualitatively different from the mechanism of storage and retrieval systems. (*ibid*: 224) In such cases it is unclear where one would locate the role of a representation, other than as a purely descriptive picture of the system.

**Music.** A definition of music would seem to be necessary, but I will not attempt such a maneuver. However, it is enlightening to discuss problems one might encounter in constructing such a definition.

(1) Many disparate activities may be classified as musical. The two primal sound-generating activities of rhythmic motion of the limbs and melodic outpourings of the voice form a basis for many kinds of music. A whole continuum of instruments in the wind, string, and percussion families derive from, and even refer to, these two fundamental acts. Also, musical activities can appear non-musical in certain contexts, and vice-versa; furthermore, the designation "musical/non-musical" is highly culturally contingent. What passes for non-music (e.g., sine-tone sequences in psychoacoustic experiments) can still be perceived musically, i.e., perceptually organized according to a listener's culturally contingent music-listening strategies. It is not clear where in our perceptual/cognitive systems we would mark the cutoff between sociocultural contingency and psychological fact. Hence, some such experiments may not be as close to measuring cognitive universals of music as some experimenters might believe.

(2) Music possesses different status and roles in different cultures and subcultures. In the west we have many musics associated with various communities: concert music's high-culture spectacle, the colloquial pulsating functionality of an urban hit single, a Hollywood film score's emotional manipulations, the precise environmental design of elevator music. In many such cases it does not make sense to discuss "the musical object" divorced from its context. A full understanding of the perception and cognition of music must include these disparate functions of music. For one rarely attends to elevator music with the attention that one is obliged to give in a concert hall, nor does one often give concert music the level of physical engagement normally expected on a dance floor. It appears that these sociocultural boundaries often serve to delimit not only music's functionality but also our reception,

attention, and understanding. When speaking of music cognition, we should also address musical functionality, and hence we should keep these social factors in mind.

**Grouping.** By *grouping* usually we mean the perceptual or cognitive unification of some series of contiguous events or stimuli, due to proximity and/or similarity (Lerdahl & Jackendoff 1983). It can also mean a unification of non-contiguous events or stimuli due to periodic or non-periodic repetition (Parncutt 1994). In music-perception studies, grouping usually refers to the temporal domain, though synchronous tones can be perceptually grouped also, to form a complex timbre or harmony.

**Rhythm.** I propose that we construe rhythm broadly, as *any* perceived or inferred temporal organization in a series of events. The organization itself need not be cognized thoroughly; it may merely be perceived to exist. The perception of rhythm occurs usually because of some kind of perceptual grouping of events. Rhythm need not evoke a sensation of pulse (e.g. circadian rhythms, conversational speech, most Western contemporary concert music), nor need rhythm be "intended" by the producer of these events (ocean waves, sewing machines); nor need it be constrained to any sensory modality, social context, or timescale. In particular, rhythm can be but need not be a raw sensory phenomenon. Note that this listener-centered definition allows that one person may perceive rhythm where another does not.

**Pulse.** Literally, pulse denotes any periodicity inherent or perceived in any rhythm or combination of rhythms. It also strongly connotes *isochrony* (i.e., a fixed tempo), often connotes some degree of perceptual "salience," and weakly connotes an approximate frequency range between 1.2 and 3.3 Hz (a fuzzy category known as the *tactus* range, also the range of the human heartbeat pulse, human locomotion, and the infant sucking reflex). However, from a scientific perspective, these connotations could be dispensed with entirely in the definition; they seem to derive from the stricter meaning via linguistic and cultural usage of the term itself. *Beat* is roughly equivalent to pulse, except that it is more variable; in some contexts it more strongly connotes the tactus range, whereas in others it functions at any timescale. Beats can also function as an abstract quantity, so that notes may last fractions of a beat. Some authors (Lerdahl & Jackendoff 1983) have insisted that beats are to be seen as points in time, rather than as intervals of time. However, common usage suggests other possible meanings. Hence beats can both mean the discrete time point at which the interval occurs (as in "start this note on the beat") and the continuous interval between such time points (as in "hold this note for three-and-a-half beats").

**Tactus.** The tactus has been long understood to mean the moderate-tempo pulse present in most rhythmic music. Typically when asked to tap a finger or foot to a piece of music, listeners choose a regular time period that is in the approximate range of 300 to 800 milliseconds, averaging a little slower than 2 beats per second (Fraisse 1982). As the music gets faster, a listener is inclined to find progressively slower pulses such that they fit within this range, and vice-versa. The tactus range is also the range of "spontaneous" tempo, that is, of the tempo produced by the typical person asked to tap a steady pulse. This range coincides with a moderate walking pace, a human heartbeat, the rate of jaw movement in chewing, and the infant sucking reflex. It is also a fairly comfortable rate at which to tap a foot or a finger, since it is neither too fast for motor control, nor too slow for accurate, regular timing. Hence the tactus seems to correspond to natural timescales involved with human motion; we might imagine a chipmunk to have a faster tactus.

**Polyrhythm.** This term literally means multiple rhythms appearing simultaneously; it is simply polyphony viewed in its rhythmic dimension. Polyrhythm also frequently connotes multiple cyclically recurring rhythms, but only because the term is used often in conjunction with African musics, in which cyclic rhythms are commonplace. Cyclicality itself is not inherent in polyrhythm. Contrary to how it is often discussed in the literature, the individual rhythms in musical polyrhythm are usually more complex than mere periodic pulses. (I would call the latter construct *polycycle,* or *trivial polyrhythm,* not because it is trivial to reproduce but because it is assembled from trivial individual rhythms; an example would be what we call "three-against-four.")

**Meter.** Most generally, meter is a periodic grouping of a musical time unit. Traditionally in European concert music, meter connotes a hierarchy of weak and strong beats. However, as I shall elaborate in chapter 5, meter can exist without such a hierarchy. Meter denotes a subharmonic (or grouping) of a pulse, and might also imply a higher harmonic (or subdivision) of the same pulse. That is, it can simultaneously group and subdivide pulses into regular units. For example, the time signature 6/8 denotes a cycle of two pulses each divided into three equal subunits. Note that meter is treated as a periodic grouping of pulses – i.e. as a cognitive/perceptual phenomenon, not as an objective reality of the acoustic signal. However, this distinction is often elided, so we might speak of the meter of a piece of music. (See chapter 5 for more on this issue.)

**Expressive timing**. Some theorists have tried to divide music into its *structural* and *expressive* components. This distinction also tends to fall along the same lines as *discrete* versus *continuous* elements. Expressivity in performance is taken to mean that which deviates from regularity; one can be expressive with intonation, with dynamics, with tempo and other kinds of timing. The regularity of a group or unit can be taken to mean, according to Seashore (1938), the norm set by the unit itself; hence the understanding of expression amounts to a kind of statistics. Expressive timing has come to mean the ways in which performers deviate from strict metronomicity.

The separation of the structural and the expressive is of course a problematic distinction, for it (1) suggests that composers cannot be expressive, and (2) presupposes a distinction between the fixed and the regular, between the composition and the performance. At worst, however, this distinction amounts merely to an unfortunate choice of wording, or an impoverished definition. But just as the term "representation" has been appropriated by the cognitive-science community, so the term "expression" has been redefined provisionally by the cognitive-musicology community to refer to a subset of what is typically seen as expressive performance. While I am aware of this issue, I will use the terminology as it is used in the literature on expressive timing. Namely, the term refers to the differences, along the temporal dimension, between a "score" and a "performance," or in the case of improvised groove-based music, between regularity of the underlying meter and fluidity of the performed rhythms.

**Microtiming.** Microtiming, as I have been using it, refers to expressive timing at the sub-tactus level, characterized by high-frequency activity. It is complementary to tempo modulation, which has a low-frequency emphasis. It corresponds to Bilmes's (1993) concept of deviation, but microtiming is more general since it doesn't connote an ideal metric referent (as in deviation *from* something). Microtiming refers to the entire range of sub-tactus, non-notatable rhythmic expression, pertaining both to music and to speech, from which much musical rhythm originates.

**Groove.** I believe that African and African-American dance musics and their descendant genres should be treated in terms of a "groove" – which might be described (but not defined) as an isochronous pulse that is established collectively by an interlocking composite of rhythmic entities. A groove tends to feature a high degree of regularity but also conveys some sense of animation. Groove involves an emphasis on the *process* of music-making, rather than on the syntax (Keil & Feld 1994). The focus is less on coherence and the notes themselves, and more on spontaneity and *how* those notes are played. Groove concerns the animation and decoration of time as it is shared by musicians and audience. This relates to the functional role of African and African-American musics in their communities. It is worthwhile to point out the common observation that both African and African American peoples exhibit a cultural tendency to treat music on a functional basis. That is, music is not merely treated as a work of art for art's sake, but as an activity that is integral to life, partially structuring everyday reality.

A salient feature of groove-based musics seems to be the attentiveness to an additional unifying rhythmic level below the level of the tactus. For example, if the quarter note is the tactus, one may also focus on the sixteenth note to heighten rhythmic precision. It is verified experimentally that discrimination of long temporal intervals is more variable than discrimination of short ones (Weber's law), and the sum of the variances of *n* subdivisions is a factor of *n* smaller than the variance of the total interval. Hence we actually gain accuracy in timing a moderate pulse by subdividing it. According to Fraisse (1982), music listeners typically divide rhythmic intervals into two categories, long and short. These intervals are usually in the ratio of 2:1, indicating that the smaller interval is a subdivision of the larger one; also, the long interval is usually in the tactus range, whereas the short one lies in the subtactus range. Fraisse notes that these two categories have different perceptual implications. During the long intervals, we can be aware of the passage of time, whereas we do not sense temporal extent during the short intervals. However, we can have qualitative awareness of the grouping of numbers of such brief intervals (Fraisse 1956, cited in Clarke 1999), such as their "two-ness or three-ness, [or] accentedness or unaccentedness" (Brower 1993: 25). For this and other reasons, the smallest operative musical subdivision of the tactus has been referred to as a "temporal atom," which was then abbreviated by Bilmes (1993) to "tatum" in homage to the master pianist, Art Tatum **[CD-1]**.

Groove has no correlate in European concert music, and is therefore indescribable by models derived from it. Groove-based musics do not often feature the phrase-final lengthening, *ritardandi, accelerandi*, *rubati,* or other expressive tempo modulations of European classical music; rather, they involve miniscule, subtle microtiming deviations from rigid regularity, while maintaining overall pulse isochrony. This mode of rhythmic expression has a whole tacit grammar unto itself, with its own set of esthetics, techniques, and methods of development. To our knowledge, while much research effort has focused on the investigation of the aforementioned tempo-modulating phenomena (e.g. Longuet-Higgins 1982, Todd 1989, Repp 1990), very little attention has been devoted to expressive timing in the context of an isochronous pulse or groove. Sometimes they are described as "small accelerations and decelerations," (Magill & Pressing 1997), i.e. in terms of a larger construct called tempo, as if to imply the existence of some kind of musical time independent of the musical events that shape it.

Bilmes (1993) developed a model for groove-based expressive timing that features two simultaneous isochronous pulses, one at the foot-tapping *tactus* level (with a period typically between 300 and 800 ms), and another, the temporal atom or *tatum,* at the smallest operative subdivision of that pulse (typically 80 to 150 ms). The onset time of a note occurring on a specific tatum (i.e. a specific sixteenth, twenty-fourth, or other such note) can be transformed by a continuous *deviation* from perfect quantization. Hence rhythmic expression can occur at the tatum level without perturbing the overall tactus or tempo. This representation is described and expanded upon in chapter 7.

**Attention.** Attention can be described or defined in numerous ways. As Jones & Yee (1993: 70) put it, "Ultimately, definitions of attention become theories of attention." It has been described variously as the allocating of info-processing resources to a specific source of information, frequently to the neglect of others; the differential processing of simultaneous sources of information; or simply, the mind's ability to focus and concentrate. It is believed by many that musical meter provides us with an attentive mechanism – a temporal template against which to process information in time, reducing demands on memory. This issue is discussed in chapter 5.

Now, armed with these terms, we will turn our attention to the cognitive science of rhythm perception and cognition.

## 3. New Paradigms: Embodied Mind, Situated Cognition

Music provides an especially interesting laboratory for the study of cognition. Because so much musical behavior is non-linguistic in nature, music tends to challenge dominant linguistic paradigms, which reduce all cognition to rational thought processes such as problem solving, deductive reasoning, and inference. Unfortunately, most research in music perception and cognition has focused on a very narrow band of human musical phenomena, namely the tonal concert music of pre-20th-century Western Europe, as filtered through contemporary European-derived performance practices. Hence we have an abundance of tonal-music-inspired models and representations for perceptual and cognitive phenomena, focusing almost entirely on pitch organization in the large-scale time domain. Some examples are theories of recursive formal hierarchies (Lerdahl & Jackendoff 1983) and of musical meaning from deferred melodic or harmonic expectations (Narmour 1990, Meyer 1956). Lerdahl and Jackendoff (1983) contend that in the way that information-processing stages are organized, musical cognition is fully analogous to linguistic cognition. Such models suppose that the cognition of music consists of the logical parsing of recursive tree structures to reveal greater and greater levels of hierarchical organization.

A side effect of the traditional linguistics-based approach has been the adoption of a major tenet of the Anglo-American "analytic" philosophical tradition, namely that rational language as a logical system has been taken to form the basis of virtually all thought and meaning (Prem 1996). Even in the work of Lakoff (1987, discussed in the following chapter), which purports to frame linguistic understanding and meaning in terms of the human body and its environment, the theory centers around language. Such work would have it that no non-linguistic modes of discourse exist, and that all meaning is attached to language. However, music provides a clear counterexample, with its basis in bodily activity and its strong emotional component.

To be sure, some musical features may have elements in common with language. Bregman (1990) and Handel (1990) suggest independently that similar functions operate in the perception of both speech and music at early levels of processing in order to organize the complex acoustic signal into extended patterns of events, presumably subject to some rules. At a higher level, the expressive manipulation of rhythmic timing in musical performance invites comparison with the use of timing in speech for semantic emphasis. Similarly, the regular grouping of pulses into an intelligible template occurs both in music and in metered verse. But such functions are of a different order from the abstract formulations of linguistic theories. Expressive timing and grouping have less to do with rules governing sentence structure than they do with either low-level auditory processing or the high-level shades of meaning possible in spoken language. Although aspects of musical behavior may have elements in common with linguistic behavior, a large part of musical understanding seems to operate overall quite separately from the realm of rational language and inferential reasoning.

Though often posited as musical universals, many of the results of Lerdahl and Jackendoff (1983) do not carry over effectively to the vast majority of non-Western musics, nor do they account fully for the perception and cognition of Western tonal or atonal music. The inapplicability of these linguistics-derived models to other musics is quite glaring in the cases of West African and African-American musics such as jazz, rumba, funk, and hip-hop. In these cases, certain salient musical features, notably the concept of groove, seem to have no analogue in rational language. Although groove is a highly subjective quality, music that grooves can sustain interest or attention for long stretches of time to an acculturated listener, even if "nothing is happening" on the musical surface. A prime example is James Brown's music **[CD-2]**, which frequently has precious little melodic or harmonic material and is highly repetitive, but would never be described as static. The fact that groove carries enough weight to override other musical factors in certain kinds of musical experience suggests that the traditional linguistics-based viewpoint does not suffice in describing the entirety of music cognition.

A major reason for this mismatch between tonal-music grammars and most music of the world is not (as is commonly thought) differing levels of musical sophistication or complexity, but rather a major cultural disparity in approaches to rhythmic organization and musical form. I claim that an essential component of this disparity is the status of the body and physical movement in the act of making music. The role of the body in various musics of the world becomes clearer when one observes the function that music and dance assume in these cultures, the common cultural/linguistic metaphors associated with musical activity. All of these observations have led us to study the role of the body in cognition in general.

## Cognitivism & Traditional Cognitive Science

Until the late 1980s, generally in cognitive science, and particularly in artificial intelligence, the logic paradigm had prevailed. The *cognitivist* or *objectivist* point of view involved the assumption that thought amounted to the mechanical manipulation of abstract symbols, and that the mind was an abstract machine, manipulating symbols by algorithmic computation in the way a computer does. All meaning arose via correspondences between symbols (words, mental representations) and things in the external world. These symbols formed internal representations of external reality, independent of any limitations of the human body, the human perceptual system, and the human nervous system. The mind was seen as a mirror of nature, and human thought as abstract and disembodied. Human bodies and their environments were thus incidental to the nature of meaningful concepts or reason. The materials (or "hardware," in computer terminology) with which brains think and senses perceive were believed irrelevant to the abstract processes ("software") that these systems conduct. The brain was merely a specific instance of a computing engine, of which all such manifestations would be formally equivalent (von Neumann 1951). Hence, in principle, a brain could be replaced by a computer, and in particular computers can do anything that brains can do; machines that mechanically manipulate symbols that correspond to things in the world were believed capable of meaningful thought and reason. (Lakoff 1987: xii-xiii) The concept of a representation, an abstract data structure for storing and manipulating such symbolic information, is central to such cognitivist theories of mind. Typically, a cognitive representation has taken the form of a "message board" – a disembodied abstraction whose shared contents provide an interface among various independent information-processing units.

These tacit assumptions spawned generations of experiments in artificial intelligence in which symbolic, language-based reasoning was taken to be the cornerstone of intelligent behavior. A classic example of this approach appears in the medical consultation program, MYCIN (Shortliffe 1976). As a so-called "expert system," the program was provided with task-specific data taken from real-life experts, thereby enabling it to achieve substantial competence in a restricted domain of analytical reasoning. However, when running such an expert-system program, the computer merely manipulates abstract symbols according to a body of explicitly formulated rules and guidelines. One could never claim that the program "understood" the ins and outs of human health issues, but merely that it contained a representation of such information and a model for manipulating this information based solely on its symbolic characteristics. Furthermore, its expected input was highly restricted in range, such that grammatical errors or ambiguous terminology could, for example, cause it to diagnose a rusty Chevy as having the measles (Lenat and Feigenbaum 1992: 197; see also Clancey 1997: 29-45). Such text-inspired approaches to knowledge have modeled intelligence as disembodied symbol manipulation.

The above description raises the hotly debated issue of the difference between "real" intelligence and "simulated" intelligence. Presumably, MYCIN could pass some restricted kind of Turing test, in which an impartial expert might judge the computer's input-output behavior to *appear* intelligent. However, I must stress that a microworld representation of intelligence is not human intelligence. For intelligent behavior encompasses not just symbolic manipulation and deductive reasoning, but also interaction with others, attunement to one's surroundings, and awareness of the relationship between oneself and one's world, not to mention intelligent action, creativity, physical coordination, emotion, and countless other behavioral manifestations. One finds that in traditional cognitive science, the emphasis on things "mental" has been at the expense of the physical, and the disproportionate attention paid to "the rational" has often occurred in opposition to the emotional and the intuitive. Yet everyday human intelligence includes all of these dimensions.

**Embodied Cognition**

Thankfully, recent conceptual developments in cognitive science move towards the inclusion of such dimensions. In particular, cognitive scientists have begun to infer connections between the structure of mental processes and physical embodiment. The viewpoint known as *embodied* or *situated* cognition treats cognition as an activity that is structured by the body and its situatedness in its environment – that is, as embodied action. In this view, cognition depends upon experiences based in having a body with sensorimotor capacities; these capacities are embedded in an encompassing biological, psychological, and cultural context. Sensory processes (perception) and motor processes (action), having evolved together, are seen therefore as fundamentally inseparable, mutually informative, and structured so as to ground our conceptual systems. (Varela et al. 1991: 173)

The embodiment hypothesis suggests an alternative basis for cognitive processes. Perception is understood as perceptually guided action. Our eyes move to frame the visual field and to focus on regions of that field; our heads move to facilitate binaural localization. Such behavior is facilitated through elaborate feedback mechanisms among sensory and motor apparatus. Hence, cognitive structures emerge from the recurrent sensorimotor patterns that enable the perceiver to guide his or her actions in the local situation. That is, the emergent, reinforced neural connections between the senses and the motor system form the basis for cognition. The mind's embodiment provides natural biases for inductive models and representations, and thus automatically grounds cognitive processes that might normally be considered disembodied. This view provides a sharp contrast from the standard information-processing viewpoint, in which cognition is seen as a problem of recovering details of the pre-given outer world. (Varela et al. 1991: 173)

In this light, the mind is no longer seen as passively reflective of the outside world, but rather as an active constructor of its own reality. In particular, cognition and bodily activity intertwine to a high degree. In this perspective, the fundamental building blocks of cognitive processes are control schemata for motor patterns that arise from perceptual interaction with the body's environment. The drives for the cognitive system arise from within the system itself, in the form of needs and goals. (Prem 1996)

Neuroscientific evidence corroborates this viewpoint. As was pointed out earlier, we can make more sense of our brains and bodies if we view the nervous system as a system for producing motor output. The cerebellum is connected almost directly to all areas of the brain – sensory transmissions, reticular (arousal/attention) systems, hippocampus (episodic memories), limbic system (emotions, behavior). All areas of our brain seem geared to coping with their functions as they pertain to problems of motor control. (Hardcastle 1996: 7) Such evidence from neuroscience allows for postulating shared mechanisms for low-level control of embodied action and higher-level cognition; motor plans for limb movement could interact with goal-oriented abstract plans. The mind thereby becomes a *distributed* entity, an emergent characteristic of the whole sensory-central-motor neural system, existing in the elaborate network of interconnections that extend throughout the body.

**Situated Cognition**

The above characterization of the embodied mind covers merely half of the picture. If we grant that cognition is structured at least to some degree by bodily experience, then we must understand the body to be immersed in an environment that shapes its experience. Hence the philosophy of embodiment also stresses temporal, physical, and sociocultural situatedness.

It has been shown that the framing of cognition in terms of the body and its environment provides not only limiting but also enabling constraints for cognition. Work in animal behavior has addressed the potential links between sensory and motor systems, as in the classic experiment by Held & Hein (1958). In this study, a group of kittens were raised in the dark and exposed to light only under controlled conditions. A first group of animals was allowed to move around normally, but each of them had to pull around a cart in which rode a member of the second group. The two groups thus shared nearly identical visual experience, but one group experienced the world actively and the other passively. Upon release after a few weeks, the first group of kittens behaved normally, but the second group behaved as if blind, bumping into objects and falling over edges. Hence objects in the world are apprehended not simply by visual extraction of features, but rather by the visual guidance of action. (Varela et al. 1991: 175) In a similar but more humane fashion, it has been observed that infants who can walk have qualitatively different reactions to certain stimuli, such as slopes and falloffs, than infants who can't (Thelen & Smith 1994: 217-220).

There is reason to believe that such programmability of the brain extends far beyond childhood. While it is commonly known that cognitive development proceeds rapidly along with brain growth and cortical myelinization during the crucial first years of life (Passingham 1982: 112ff), it is not often recognized that many networks of the brain retain a susceptibility to reprogramming throughout an individual's life (Laughlin et al. 1992: 41). The remarkable case studies of Ramachandran and Blakeslee (1998) and Sacks (1985) further attest to the adaptability and plasticity of the brain throughout adulthood. Hence we may discern a continuum of neural structures ranging from "hard-wired" evolutionary traits to highly flexible, environmentally adaptive features (Shore 1996:17). The existence of this continuum supports the embodied cognitive paradigm, which encompasses the body and its environment.

In addition to the universals of cognition based upon having a body and its sensorimotor systems, we study the particular social and cultural factors that contribute to the development of mind:

> There is reason to suspect that what we call cognition is in fact a complex social phenomenon... The point is not so much that arrangements of knowledge in the head correspond in a complicated way to the social world outside the head, but that they are socially organized in such a fashion as to be indivisible. 'Cognition' observed in everyday practice is distributed – stretched over, not divided among – mind, body, activity and culturally organized settings. (Lave 1988: 1)

We may rely upon various attributes of our physical, social, and cultural environment to support or augment our mental capacities. Lave (1988) studied the arithmetic of adults of various backgrounds in the grocery store. Lave's results showed that in making purchase decisions, these situated agents employed a flexible real-time arithmetic in order to select better prices per unit weight, continually taking into account the constraints imposed by the layout of the stores, the capacities of their home refrigerators, and the dietary requirements of their family members. Such shopping prowess – skill at situated arithmetic – was rarely reflected in subjects performance on grade-school math problems. Cognition as demonstrated in *practice* was found to be not at all in the same realm as cognition in an abstract, un-situated setting. Similarly, Clark (1997: 213-216) discusses the ambiguity of the boundary between "mind" and "world," along the lines of a distinction between "user" and "tool." When a bird drops a nut from a great height to crack it open, does the ground become a tool? (Clark 1997: 214) Rather, the bird is exploiting an aspect of its environment to extend its physical capabilities; the concept of a tool dissolves.

Along with the embodied/situated paradigm came the gradual recasting of the notion of a representation. In a classic paper entitled "Intelligence without Representation," Brooks (1991) reported research in situated robotics, in which artificial intelligence was "approached in an incremental manner, with strict reliance on interfacing to the real world through perception and action" such that "reliance on representation disappears." (Brooks 1991: 139) In Brooks's view, the intelligent system is decomposed into independent, parallel activity producers, which all interface directly to the environment (rather than to each other) such that their respective perceptions and actions can override one another in the system's resultant observed behavior. Brooks aimed to simulate entire embodied, situated intelligent systems. Abandoning the cause of simulating human intelligence, he considered simulated insect intelligence to be the most feasible first step. In doing so, he cast his work in an evolutionist's light; in his opinion, nature has conducted research on situated agents for billions of years, so he modeled his work after the gradual accrual of complexity in the evolution of species.

While rather extreme in its views, Brooks's work broke down the traditional notion of a representation. In particular, an important distinction is made in the embodied viewpoint between the point of view of the subject and that of an "objective" observer of the subject in its environment. (Prem 1996) The notion of a representation was seen to stem directly from this supposedly objective position of the observer, and hence its actual role in cognition for the situated agent was questioned. Later, Coelho (1995) suggested that we not throw out the baby with the bathwater. While it was granted that purely representational accounts were probably "too limited for modelling the recurrent processing that is the norm in the brain" (Coelho 1995: 311), the validity of the partial role of representations in the mind is proven by our capacity for memory formation and recall. Above all, the notion of representation has been retained as a valuable descriptive tool in the study of cognition, but not as a structural necessity in many cases of cognition itself. Hence caution must be used in relying on representations to define a cognitive model.

In sum, the theory of embodiment encompasses both neuropsychological and socio-environmental views of cognition. Embodied cognition stresses physical, temporal, and functional situatedness, and enforces interaction between the agent's body and its environment. Such a holistic view prevents some inappropriate simplifications and unrealistic assumptions because it enforces dealing with unexpected contingencies, provides specificity, and incorporates energetic and resource considerations. (Mataric 1996) And, quite significantly, the embodied view of cognitive science allows for direct cultural interaction, which is undeniably crucial for both language and music.

## 4. Music Cognition and Embodiment

How might we connect the theoretical framework of embodied cognition with the study of music? First, we should examine the role of the body in music perception, cognition, and production, and attempt to take into account the realities of our perceptual systems. Let us address connections between aspects of musical time – rhythm, timing, meter, phrasing – and the body. People often speak of a musical groove as something that induces motion. In describing his aesthetic criteria for rhythm tracks, a colleague of mine involved in hip-hop music distinguished between a musical excerpt that "makes me bob my head" and one that doesn't (Bilal 1997). Many of us have witnessed motion induced in infants or toddlers via music, but this behavior is not universal, involuntary, or even reliable. This capacity to entrain to a regular aural pulse may be an evolutionary vestige of a previously useful ability that has more recently fallen into disuse. While nobody can account directly for this phenomenon, it clearly involves regular, rhythmic bodily movement as a kind of sympathetic reaction to regular rhythmic sound – that is, as a kind of dance.

Recent neuropsychological studies of music perception have affirmed the cognitive role of body motion in music perception and production. From a meta-analysis of studies of brain-damaged patients with lesions localized in various regions of the brain, it was suggested that the "rhythmic component ... of an auditory image cannot be activated without recruiting neural systems known to be involved in motor activity, especially those involved in the planning of motor sequences" (Carroll-Phelan 1994; see also Peretz 1993). Such neuropsychological data have allowed hypotheses about the induction of a sense of beat or pulse in terms of the so-called sensorimotor loop, which includes the posterior parietal lobe, pre-motor cortex, cerebro-cerebellum, and basal ganglia. In the sensorimotor perspective, a perceived beat is literally an imagined movement; it seems to involve the same neural facilities as motor activity, most notably motor-sequence planning. (Todd 1997) Hence, the act of listening to music involves the same mental processes that generate bodily motion.

One might suppose that musical gestures might be more efficacious in eliciting such sympathetic behavior if they represent aspects of human motion somehow. Such sounds might include the dynamic swells associated with breathing, the steady pulse associated with walking, and the rapid rhythmic figurations associated with speech. Note that each of these three examples occurs at a different timescale; characteristic frequencies of the first regime might fall in the range of .1 to 1 Hz, the second 1 to 3 Hz, the third 3 to 10 Hz. In fact, it is interesting to observe the correspondences in frequency range in these groups of behaviors:

1) breathing, moderate arm gesture, body sway "phrase" 0.1 - 1 Hz

2) heartbeat, sucking/chewing, locomotion, intercourse, head-bob "tactus" 1 - 3 Hz

3) speech/lingual motion, hand gesture, digital motion "tatum" 3 - 10 Hz

Versions of this list were suggested by Fraisse (1982) and Todd (1997). It is a plausible hypothesis that musical activity on these three timescales might exploit these correspondences.

A variety of simple truisms support this view. For example, most wind-instrument phrase lengths are naturally constrained by lung capacity. Tactus-heavy urban dance music often makes sonic references to foot-stomping and to sexually suggestive slapping of skin. Blues guitarists, jazz pianists, and *quinto* players in Afro-Cuban *rumba* are said to "speak" with their hands and fingers. All such instances involve the embodiment of the musical performer and the listening audience.

## Time and Timing

Let us return to the domain of expressive timing. In groove-based contexts, rhythmic expression occurs at an extremely fine timescale – rapid enough to rule out a simple auditory-feedback mechanism for its implementation (Fraisse 1982). This relates to an age-old question in neuroscience, known as the problem of serial order in behavior (Lashley 1951). The question is how to explain our assimilation and production of very fast sequences of events in time, given that human reflexes and neural transmission speeds would seem to be too slow to account for them. Lashley cites the common experience of mistakes in serial order of rapid sequences, such as typing, as evidence of hierarchical organization in this kind of behavior.

There is evidence that temporal, rhythmic, and grouping judgments and productions employ different modes of processing for times under roughly one half-second than they do for longer times (Fraisse 1956: 29-30, cited in Clarke 1999; also Preusser 1972, Michon 1975). These short-time processes are described variously as "pre-cognitive," "sensory," or "immediate" – as a kind of sensation, recognition, or gestalt perception, rather than a kind of analytical or counting process. It has been suggested that this cutoff corresponds to the transition between so-called echoic and short-term memory, as indicated by the timescales involved and by other experiments (Michon 1975, Brower 1993).

Consequently, these different regimes of memory should distinguish musical rhythms above and below this approximate cut-off as qualitatively different phenomena. For pulse-based music, this cut-off lies in the middle of the tactus range, about 300 to 800 milliseconds; rhythmic material below this is perceived categorically as combinations of subdivisions of a main regulating pulse, and durations above it are considered to be on the level of metric grouping of pulses. By this division, as I shall discuss in the next chapter, echoic memory covers the immediate timescale of rhythmic activity, whereas short-term or working memory covers meter and phrases. These different types of memory involve different kinds of processing. We *entrain* to a pulse based on the echoic storage of the previous pulse and some matched internal oscillator periodicity; we *feel* the relationships among strong and weak beats (accentual meter); we *count* times between phrases or bars (metric grouping); and we *recognize* sub-pulse rhythms qualitatively (Brower 1993). An embodied account of rhythm perception and cognition would need to factor in these inherent distinctions of human memory.

The role of different kinds of memory points to the need for different models to explain rhythmic expertise at such a fine scale. A hint comes from bat and owl echolocation, in which neural delay-line architectures serve to give the creatures much higher temporal resolution than neural transmission would seemingly allow (Feldman 1997). One could say that the animals' temporal acuity exists "in" these long neural pathways – in the physical structure of the perceptual apparatus. A working hypothesis, inspired by the existence of such structures, is that precisely timed rhythmic activity involves the entire body in a complex, holistic fashion, combining audio, visual, and somatosensory channels.

According to the embodiment hypothesis, cognitive structures emerge from reinforced inter-modal sensorimotor coupling. In this view, short-time rhythm cognition might include physical sensation, visual entrainment, and sonic reinforcement, unmediated by a symbolic representation. Cognition on the part of musicians – especially on polyphonic, multi-limb instruments such as drums or piano – apparently involves the physical act of making music as a primary ingredient. Consider the components of the sensory-motor image associated with rhythm perception that are rooted in echoic memory: a phrasal/body-sway-oscillator component (respirator-based), a tactus/foot-tap-oscillator component (locomotor-based) (Todd 1994), and a tatum/multiple-finger-tap-oscillator component (speech- or digit-based). According to the embodied-cognition viewpoint, what have been previously called our internal representations may consist of no more than these very sensorimotor couplings.

**Kinesthetics**

Words like *kinesthetic*, *haptic*, and *proprioceptive* refer to the psychology of bodily feedback. They all refer to the sensation of bodily position, presence, or movement resulting from tactile sensation and from vestibular input. We rely on such awareness whenever we engage in any physical activity; it helps us hold objects in our hands, walk upright, lean against walls, guide food into our mouths, and swallow it. In these cases, there is a strong interaction between kinesthetic and visual input. Similarly, in the playing of musical instruments, we must treat sonic and kinesthetic dimensions as interacting parameters; we must bear in mind the spatio-motor mode of musical performance.

All too often, theorists and psychologists have treated musical motion in terms of abstract, time-varying auditory images, while ignoring the motions exerted by the performer. Musical motion is seen as bound up with structural abstractions in pitch space or other sound worlds, involving the play of forms against one another. A typical view is evident in the following quote from noted composer-theorist Roger Sessions. "The gestures which music embodies are, after all, invisible gestures; one may almost define them as consisting of movement in the abstract, movement which exists in time but not in space, movement, in fact, which gives time its meaning and its significance for us." (Sessions 1950: 20, quoted in Shove & Repp 1995: 58) A recent review of the concept of musical motion by Shove and Repp (1995) highlights the often overlooked fact that musical motion is, first and foremost, audible human motion. To amplify this view, Shove and Repp make use of Handel's (1989: 181) three levels of event awareness: the raw psychophysical perception of tones, the perception of abstract qualities of the tones apart from their source, and lastly the apprehension of environmental objects that give rise to the sound event. This last level is aligned with the "ecological level" of perception as suggested by Gibson (1979). At this level, "the listener does not merely hear the sound of a galloping horse or bowing violinist; rather, the listener hears a horse galloping and a violinist bowing." (Shove & Repp 1995: 59) In this ecological framework, the source of perceived musical movement is the human performer, as is abundantly clear to the listener attending to music as a performance event (ibid : 60). We connect the perception of musical motion at the ecological level to human motion. This suggests that musical perception involves an understanding of bodily motion – that is, a kind of empathetic embodied cognition.

For musicians, a major part of musical competence involves the bodily coordination of limbs, digits, and for wind instruments, breathing. Such bodily awareness is most demanding on polyphonic instruments, where multiple sonic streams are generated simultaneously. (In this way, drums and keyboard instruments are the paradigm for body-centered polyphony. The drum set and the organ are the only four-limb instruments; piano should be considered a three-limb instrument including the use of the pedals, which are often coordinated strongly with the sounds generated by the hands.) For musical performers, the difference between musical and human motion collapses; the rhythmic motions of the performer and of the musical object are essentially one and the same (Shove & Repp 1995: 60).

Blacking (1973) raised the issue of kinesthetics in musical performance in comparing two types of kalimba ('thumb piano') music among the Venda community of South Africa. One very physical type, practiced by amateur boys, featured complex melodies that appeared to be secondary artifacts of patterned thumb movements; the regularity of the movements generated the jagged melodic result. The other type, a more popular style practiced by professional musicians, had simpler melodies with small intervals and flowing contours, directed more by an abstract melodic logic than by a spatiomotor one. Baily (1985, 1989) and Baily & Driver (1992) have studied the ergonomic factors that constrain and shape performance and musical structure for various Western and Persian plucked string instruments. They argued that "the spatio-motor mode should be regarded as a legitimate and commonly used mode of musical thought" and that musical creativity may involve "finding new ways to move on the instrument." (Baily & Driver 1992: 59) Especially in instances of American rock guitar, they observed, "musical patterns are remembered and executed not solely as aural patterns but as sequences of movements, and that the music is therefore represented cognitively in terms of movement patterns which have visual, kinaesthetic, tactile, as well as auditory repercussions." (ibid. : 62) From this it was concluded that "the spatial layout of notes and the physical structure of the guitar provides a framework for musical conceptualization, a compositional tool used for the development of musical ideas, an interface to be manipulated and acted upon in certain specific ways." (ibid. : 70)

From my experience with jazz improvisation on the piano, I have found that the kinesthetic/spatiomotor approach and the melodic approach form dual extremes of a continuum. One augments one's aural imagination by exploring the possibilities suggested by the relationship between the body and the instrument, and one judges the result of such experimentation by appealing to one's abstract musical processing capacities and aesthetics. Among pianists who have exploited this relationship in jazz, Thelonious Monk has been the most influential. His compositions and improvisations provide an exemplary nexus of kinesthetics and formalism. Often his pieces contained explicitly pianistic peculiarities, including the repeated use of pendular fourths, fifths, sixths, and sevenths (as in "Misterioso" **[CD-3]** and "Let's Call This" **[CD-4]**), whole-tone runs and patterns ("Four in One" **[CD-5]**, bridge to "52nd Street Theme" **[CD-6]**), major- and minor-second dyads ("Monk's Point" **[CD-7]**, "Light Blue" **[CD-8]**), and rapid figurations and ornamental filigrees ("Gallop's Gallop" **[CD-9]**, "Trinkle, Tinkle" **[CD-10]** - see below). All of these idiosyncrasies fit, so to speak, in the palm of the pianist's hand, while often wreaking havoc for horn players (or, even worse, vocalists). Frequently, Monk's use of such kinesthetically derived material juxtaposed their relative ease of delivery on the piano with their melodic or harmonic ambiguity. He incorporated these elements as fundamental pieces of his improvising style. Also, when transferred to other instruments such as the saxophone, many of his piano-inspired compositions had revolutionary implications for the improvising soloist. Among the most exemplary of Monk's recorded work of this nature are his versions of "Trinkle Tinkle" with saxophonist John Coltrane. (Recordings of the compositions mentioned above can be found on Monk 1986, 1994).

*Opening bars of Thelonious Monk's composition "Trinkle, Tinkle" (author's transcription, piano right hand only)*

Upon examination of the first two measures of this passage, one notices that amidst all its rhythmic complexity, it repeatedly employs consecutive fingerings. Such physical patterns are simpler and apparently more primal for finger coordination than any nonconsecutive pattern. Monk was able to place these simple patterns in unconventional rhythmic and melodic relationships to yield new compositional and improvisational possibilities.

For non-polyphonic instruments (winds, brass, and many bowed and plucked strings) the role of kinesthetics may be different. Playing involves less "split" consciousness among limbs; for the most part, the two hands act together. But in groove-based music, there is an implicit challenge of relating what one is playing to an internally generated pulse – the metronome sense (Waterman 1952) mentioned below in the chapter on meter perception, or the imagined movement (Todd 1997) described above. The legendary trumpeter Doc Cheatham spoke of this relationship: "[Playing]'s like dancing; it's the movement of the body that inspires you to play. You have to pat your foot; you get a different feeling altogether than when you play not patting your foot." (quoted in Berliner 1994: 152) Here he is speaking not of tapping the rhythm he is playing, but tapping the underlying pulse in contrast to what he is playing. Similarly, a colleague who plays bass in numerous dance-oriented salsa bands noted the new dimension of rhythmic awareness that he experienced once he had learned the dance steps associated with the music he was playing. (J. Bilmes, private communication, 1996) All musicians in the group perform these rudimentary salsa dance steps while playing onstage; this elicits a compound rhythmic consciousness along the lines of Cheatham's playing-while-tapping. Evidently, part of what it means to *groove* or to *swing* involves the continual, embodied awareness of the relationship of the pulse to the generated musical material. Different musicians have different opinions about the necessity of physically generating the pulse; for example, in my collaborations with saxophonist Steve Coleman, he has often asked me to stop tapping my foot, apparently because it distracts him. But most musicians, including Coleman, seem to agree about the importance of feeling the pulse in one's body. Hence we can regard the sense of groove as at least partially kinesthetic; it involves relating actions and sounds to the sensation of pulse, which we treat as a virtual movement.

Indeed, it should be seen as no coincidence that one's sense of rhythm is referred to as "feel." A certain kind of awareness is required to be able to tap into this embodied sense of groove. Often musicians believe that this sense can vanish if one attempts to scrutinize it. Musicians often instruct each other not to "think too much" about rhythm, apparently meaning not to analyze it symbolically with numbers or words. Instead, acute rhythmic prowess tends to be a skill that is developed over time, generally in a mindful but undeliberate fashion. Overall, a fair amount of mystique is attached to rhythm perception and performance; there is a relative poverty of terminology or pedagogy associated with these finer points of rhythm.

No treatment of the kinesthetics of musical performance would be complete without at least a mention of dance. Dancing to music is found in all cultures in a vast diversity of manifestations, in secular, religious, and ritual contexts. In many societies music and dance are bound up in one general practice, such that it may be immaterial to suppose one or the other to be the primary activity (Gregory 1997: 127). Arom (1991) describes music in sub-Saharan Africa as a motor activity, almost inseparable from dance, and comments that hearing music often instantly induces body motion among many inhabitants of this region. Many Central and West African vernacular languages have no word for music alone, and few dissociate the concept of rhythm as an abstract component of music. Rhythm is thought of as the stimulus for the bodily movement to which it gives rise, and is given the name of the dance (Gregory 1997: 127). In the Anlo-Ewe culture of southern Ghana, the term that most closely approximates our usage of "music" has been translated by Ladzekpo (1995) as "dance-drumming."

Our understanding of music is enhanced if we interpret our common practices of foot-tapping, head-bobbing, and finger-snapping as a generalized kind of proto-dance, one that arises from the imagined movement associated with beat perception. If we frame groove-based music as *meant* to be danced to, even in these minute ways, then a possible explanation of the elusive sense of groove begins to reveal itself. The physical sensation of groove, either in performance as a musician or in co-performance as a listener, involves both the real bodily movement and the imagined movement supplied by the perception of the isochronous pulse. The former entrains to the latter via both auditory and kinesthetic feedback. Over time, the physical motion is strongly reinforced through repetition; cognitive structure emerges from this reinforced, cross-modal sensorimotor coupling. Hence, groove as performed may stem from this kind of *overlearned* kinesthetic pattern or sensorimotor program.

**Musical Bodies in Culture**

The embodied-cognition viewpoint suggests that a musician's internal representations are intimately tied to his or her connection with the instrument, which forms part of the music-making environment. Musical abstractions certainly exist, but I claim that how an individual musician chooses over time to interact with that instrument gives rise to the majority of the musician's cognitive apparatus. The musician's relationship with the instrument can leave its trace on the music itself – that is, it can be communicated musically. Barthes made such an observation from a listener's point of view:

> I can hear with certainty – the certainty of the body, of thrill – that the harpsichord playing of Wanda Landowska comes from her inner body and not from the petty digital scramble of so many harpsichordists (so much so that it is a different instrument). As for piano music, I know at once which part of the body is playing – if it is the arm, too often, alas, muscled like a dancer's calves, the clutch of the finger-tips (despite the sweeping flourishes of the wrists), or if on the contrary it is the only erotic part of a pianist's body, the pad of the fingers whose 'grain' is so rarely heard... (Barthes 1977: 189)

Barthes believed that he could hear an essential bodily approach in a given musician's music – that he was able to "know the dancer from the dance," as in Yeats's famous line. Such interpretations are only possible given a substantial amount of detailed background knowledge about the specific music and the technique of playing the instrument. One would require an understanding of how the gross bodily traits to which he refers could be encoded in music. The emergent structure communicated therein is informed largely by cultural norms. Again, when we speak of cognition via the body and its interaction with its physical environment, we must also discuss the social and cultural forces that construct the concept of the body. An important culturally based conceptual distinction between European and African musics involves precisely this status of the body – the degree to which the physical situatedness of the music-making or listening body is acknowledged. A journeyman jazz pianist might observe and employ different aspects of the piano from those that a journeyman classical pianist might exploit. Thelonious Monk and Cecil Taylor have treated pianos quite differently from Glenn Gould and Vladimir Horowitz, in part because of differences among these artists' respective cultural sensibilities.

The difference between this African-American dance-music model and the European concert-music model relates to the role of the body in these respective cultures and genres. In a witheringly sharp description of this contrast, McClary (1991) writes,

> In many cultures, music and movement are inseparable activities, and the physical engagement of the musician in performance is desired and expected. By contrast, Western culture – with its puritanical, idealist suspicion of the body – has tried throughout much of its history to mask the fact that actual people usually produce the sounds that constitute music. As far back as Plato, music's mysterious ability to inspire bodily motion has aroused consternation, and a very strong tradition of Western musical thought has been devoted to defining music as the sound itself, to erasing the physicality involved in both the making and the reception of music... (McClary 1991: 136)

(It was precisely this puritanical tendency that inspired Barthes's writing on the grain of the voice (1977), discussed below.) This description recalls the aforementioned tendency of theorists, cognitive psychologists, and Western composers to prevent the reality of embodied performance movement from entering into the abstract concept of musical motion. By contrast, many musics of the world that are not associated with a socially strict high-art tradition, and especially West African and African-American music, feature a body-based approach to music-making. By this I mean that they do not regard the body as an impediment to ideal musical activity, and that instead, many musical concepts develop as an extension of physical activities such as walking or repetitive tasks. The above discussion of Thelonious Monk suggests that his highly experimental musical techniques emerged in an environment where he felt perfectly at ease exploring the relationship between his body and the piano, even allowing his musical ideas to be subject to this relationship.

This distinction between cultural models pertains particularly to the respective approaches to rhythm. In my experiences both as a European-style orchestral violinist and as a keyboardist in jazz and hip-hop/funk contexts, I have noticed a strong cultural disparity in the respective roles of the body in rhythmic activity. As youths in violin sections of school and community orchestras, my peers and I were often discouraged from tapping our feet or swaying rhythmically. Such behavior was made to seem gauche and inappropriate, and moreover it threatened to draw attention away from the conductor's visual pulse. But in many contemporary dance-oriented bands in which I work, we often quite purposefully employ a kind of rhythmic bodily entrainment. This serves not only to allow interpersonal visual-rhythmic interaction to facilitate a collective groove, but also to help each musician to feel the relationship of his part to his own internally generated physical pulse.

Note that in both models, the sense of pulse is continually reinforced; all participants are microadjusting constantly. However, in the conducted case, as is necessary with an unamplified group of such large physical extent, the visual dimension is primary, whereas in the latter, collective case, the groove is maintained chiefly through the sonic dimension and supplemented with visual input. In reviewing preliminary studies of asynchrony among ensemble musicians, Rasch (1988) observed in a survey of recorded classical chamber works that for ensembles of 10 or more persons, on average, conducting is required, and not for 9 or fewer. But while the classic baroque ensemble, jazz big bands and modern salsa bands can have 15 to 20 members, such groups rarely have true conductors; rather, they chug along rhythmically via a collective sense of pulse, deferring to a single designated musician in passages with fermatas or rubati. In the great swing-era jazz bands, the role of a conductor was often merely ornamental, fulfilling no crucial function for the execution of the music aside from the possible cueing of entrances. This suggests that other musical elements contribute to rhythmic precision – notably the guiding role of drums, percussion, piano, and harpsichord, whose sharp attacks present unambiguous cues for collective rhythmic synchrony. But in addition to these percussion instruments, one must not underestimate the role of proprioceptive and visual feedback among the musicians and, in many cases, from the dancers.

## Embodiment and Metamusical Language

Thought has gestalt properties and other overall structures that go beyond mere handling of symbolic building blocks by logical rules. The efficiency of cognitive processing, as in learning and memory, depends on this overall structure of the conceptual system. (Lakoff 1987: xiv-xv)

With these realities in mind, Lakoff and Johnson (1980) set forth the claim that metaphorical structures underlie the way we understand aspects of everyday life. For example, common statements such as "You're wasting my time!" and "How do you spend your days?" suggest an underlying conceptual equation of time with money, resources, and commodity. (Lakoff & Johnson 1980: 7-9) Such elementary metaphorical structures are often reflective of aspects of embodiment, be they bodily or environmental. As an example, note our tendency to treat conceptual abstractions as visual entities. We say "I see" to mean "I understand," as things become "clear" or "transparent." This would appear to stem from a primal sense of understanding that is visual in nature – a privileging of the sense of vision among the modalities. This seems ecologically valid, as vision remains our most relied-upon and most continually varying modality.

Linguistic theories have studied the role of physical and spatial metaphors in language (Lakoff 1987, Lakoff & Johnson 1980). Lakoff (1987) developed the concept of image schemas, which are rough conceptual pictures that we use to organize our understandings of abstract concepts. For example, the "container" schema makes it possible for people to be *in* love, *out of* trouble, and so forth. Lakoff argued that many of these schemas (such as the up-down, center-periphery, container, and movement schemas) are indeed kinesthetic, that is, derived from somatic experiences that "preconceptually structure our experience of functioning in space." (Lakoff 1987: 372) Such theories suggest both bodily and environmental bases for cognition.

We can cast a similar glance at the "metaphors we play by" – the language and schemas we use to conceptualize and talk about music. Surveying some common tropes of metamusical language ought to lead to larger, underlying *ur*-metaphors, which may shed light on music cognition. Some commonly cited examples are aids for visualization of abstraction: pitch as height, timbre as color. (In my experience teaching music to children, I have observed that they tended to have no a priori sense that certain pitches were "higher" than others, even when they were well aware of gradations of pitch. This seemed to suggest that such an abstraction is merely arbitrary convention.) Others may have ecological significance, as the connection between loudness and size; a "big" sound usually comes from a physically large source. Here I point out a few other common musical metaphors.

**Time as space.** It is unsurprising that one of the most pervasive tropes in metamusical language is a sense of spatial dimension and extent. As with time in general human experience, musical time can appear to have forward *momentum* or to stand *still*; time is spatialized in an overall horizontal sense that seems to grow out of our experience walking around in the world. However, rhythmic time also carries a vertical implication, akin to a sense of gravity. As mentioned above, verticality is commonly understood in the realms of pitch and harmony; we have high notes and low notes, stacked voicings and root movement. Less often acknowledged is the way we verticalize time in the presence of a pulse. We play *upbeats* and *downbeats*. Rhythms can be *grounded* or *floating*; time can be *suspended*; a bassist can *walk* a steady pulse; a drummer can play *four-on-the-floor*. This common underlying trope amounts to a verticalization of rhythmic phase, i.e. of "circular" time. This provides a compelling connection between rhythmic pulse and the act of walking, in which feet are raised and lowered in a repetitive manner. These two notions of time demonstrate a grounding in physical embodiment; an extended, rhythmic piece of music can carry a metaphorical suggestion of a walking journey, characterized by regular rhythmic pacing coupled with the gradual visual flow of one's surroundings. Gibson (1975) has argued that this experience of walking through a stationary environment underlies our ecological (and, he argues, illusory) understanding of time as a continuously flowing quantity.

**Music as speech**. Often, music bears metaphorical attributes of speech and conversation. Monson (1996) has given an elaborate treatment of this metaphor in the context of jazz improvisation. One often hears instances of this metaphor in African-American musical pedagogy, where "'to say' or 'to talk' often substitutes for 'to play.'" (Monson 1996: 84) Such usage underscores what musical performance does have in common with speech as an activity or behavior, as well as what music has in common with language as a symbolic system. Among the traits that link musical performance to speech, we see that:

- Musical performance is a *process*, a salient mental and physical activity that takes place in time.
- Musical performance is *interactive*, characterized by dialogue, call-and-response, and collective synchronization.
- Music has *semiotic* dimensions, which enable sonic symbols to refer actively to other parts of the same piece, to other music, or to extramusical and contextual phenomena.
- Aspects of performed musical activity can be *performative*, existing only in the context of performance and meaningful only by virtue of their being performed. (See below for more on performativity.)

Note that these aspects of speech and performed music are not restricted to the domain of semantics; that is, they are not solely concerned with the "intrinsic" meanings of words or notes. Rather, these specific aspects depend upon the act of performance.

**Music as life.** A final way of framing music metaphorically is as life itself. Among many jazz musicians, a most valued characterization is that a certain musician has his or her own, instantly recognizable sound, where "sound" means not only timbre, but also articulation, phrasing, rhythm, melodic vocabulary, and even analytical skills. Generally it came to mean a sort of "personality" or "character" that distinguishes different improvisers. Though it is a complement if someone told you that you "sound like Coleman Hawkins," it is even higher praise to be described as "having your own sound." Trombonist and improviser George Lewis writes,

> "Sound", sensibility, personality and intelligence cannot be separated from an improviser's phenomenal (as distinct from formal) definition of music. Notions of personhood as transmitted via sounds, and sounds become signs for deeper levels of meaning beyond pitches and intervals. (Lewis 1996: 117)

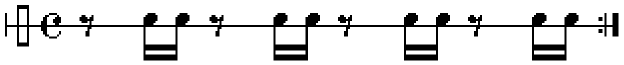
This view supports the widespread interpretation of improvisation as personal narrative (Lewis 1996: 117), as that which gives voice to the meaningful experiences of the individual. Ground-breaking pianist Cecil Taylor wrote of equally ground-breaking saxophonist John Coltrane **[CD-11]**,

> In short, his tone is beautiful because it is functional. In other words, it is always involved in saying something. You can't separate the means that a man uses to say something from what he ultimately says. Technique is not separated from its content in a great artist. (Taylor 1959)

Often, then, an improviser's original playing style is bound up with his or her (possibly idiosyncratic or self-styled) technique. Usually the autodidactic approach plays a large role for improvisers, for whom the creation of music is embodied in one's relationship to one's instrument; hence the inseparability of "sound," or pure musical approach, from a "phenomenal definition of music" – a personal sense of what music is and what it is for.

The notion of personal sound functions as an analytical paradigm, a kind of down-home biographical criticism. An individual's sound, rhythmic feel, and overall musical approach are seen as an indicator of who he or she "is" as a person. Musicians' interactive strategies in music might be seen as an indicator of their interpersonal behavior; their rhythmic placement with respect to the pulse may reflect how "fiery" or "cool" their temperaments run; their melodic inventiveness and harmonic sophistication might parallel their offstage urbanity and wit. Admittedly, such stereotypical characterizations beg to be broken down; rarely does a musician's offstage personality fit such conventional wisdom. Indeed, one could also view "musical personality" as a kind of *mask* that the performer wears onstage, Signifying on his or her offstage identity as well as on performance itself. But in either case, the notion of personal sound, relating musical characteristics to personality traits, reveals much about how music and life can be conceptualized together. It is best seen as a manifestation of a cultural model in which music is a way of life, and vice-versa.

We can also examine this metaphor of music as life in a different light as it appears in West African music. Ladzekpo (1995) counts among the most important aspects of his musical education the understanding of different musical elements in terms of their worldly counterparts. A particular rhythm might be seen as the "artistic animation" of a real-life character. As a concrete example, he has described a rhythmic ostinato that occurs on the off-beats as the "party animal," because of the way it seems to jump off the ground repeatedly and draw attention to itself:



(In the audio example **[CD-12]**, a click track is heard alongside this drum pattern, with the high click on the downbeat of the written measure.) Also, steady rhythmic pulse is seen as "purpose in life," and rhythmic obstacles such as cross-rhythms as "challenges" to that sense of purpose. Mastery of the music's rhythmic complexity is seen as a kind of strength, an ability to keep life balanced. Ladzekpo describes the music in the Anlo-Ewe culture as representative of the "complex fundamental disposition of mankind" (Ladzekpo 1995).

**Metaphors in motion: the music of Cecil Taylor.** An extended example may provide more illumination. The above metaphors of music as life and as speech were made quite concrete in my experience as a performer with pianist Cecil Taylor's "creative orchestra" in 1995. This was his music for large ensemble, which forty very fortunate Bay Area musicians had the opportunity to study and interpret under his guidance. Taylor's approach spoke volumes about improvised music as a collective activity. Early on, when we were repeatedly questioning him about the role of the written material, he said, "This [written material] is the formal content of the piece; what I want is for all the players to bring their individual languages to the interpretation and execution of the piece." Taylor desired that we create a collective embodiment of his material by filtering it through our individual "languages," framing the music as speech, individual sound as personal narrative.

In such a context the emphasis is rarely on being faithful to set compositions in terms of pitch content, melody or harmony. Compositions in an improvising context mean something else entirely – perhaps a jumping-off point, or music-generative methods that all can agree upon to some extent. In our week of daily sessions with Mr. Taylor, the earlier sessions led us to believe that he was a stickler for detail. I recall that we spent the first 3-hour rehearsal on one postage-stamp-sized corner of one of his scores; he would continually repeat and rework the material bit by bit, singing or dancing a certain phrase for us, or asking us to permute the written pitches in a certain way. But towards the end of the week, his requirements grew less stringent, his guidance less direct; he would simply set us in motion and leave the room for a while. I realized that somehow he had taught us his language – his sense of phrasing and repetition, his attention to detail, the way he rigorously reworks and dissects a turn of phrase. Once this had happened, we were free to bring our own ideas to this context – to embody his language. When he returned to the rehearsal room, he would find that we had made something out of his "hieroglyphics." Evidently, Taylor's aesthetic privileges the sound of personalities interacting over conventional concepts of form. Because of the heightened role that group interactivity played, it felt at times as though we had formed a small musical civilization, rather than an orchestra.

Indeed, our group experienced in microcosm the conflict, strife, and tension that a society experiences in macrocosm. Much of this was enacted on a musical level in the performance on October 26, 1995. For example, when some musicians reached the stage, they abandoned their allegiance to the unwritten, brittle orchestral aesthetic that had been developed over the course of rehearsals **[CD-13]**, choosing instead to yield to the temptation to play nonstop with furious intensity. This behavior raised the issue of the distribution of (physical) power – for clearly, a tenor saxophonist can honk and shriek with enough force to drown out a section of six violinists, and a drummer can bury a pianist's efforts with ease. It was found that the louder instrumentalists possessed the privilege to control the intensity level directly, while the softer instrumentalists were forced to defer to such control. Fellow musician Matthew Goodheart (1996) has observed the added role played by the self-serving musical choices made by certain individuals who wanted to "Play With Cecil" and get noticed by the legendary pianist for possible career advancement. Also, in the absence of a more dictatorial leader figure or a hard and fast text to which to adhere, we found ourselves in frequent disagreement as to what was "supposed" to be happening or what to do next. Different factions formed to conduct their own unified small-group activities, allowing for the emergence of pockets of apparent order in the sonic chaos. The resultant performances featured truly sublime flashes of fortuitous beauty and moments of brilliantly focused small-group improvisation, amidst often inscrutable orchestral noise. The metaphor of music-as-life was borne out in our experience of ensemble-as-social-group.

What do these underlying metaphors teach us about music? It becomes clear from the above discussion that especially in the realm of jazz, an understanding of music grows out of one's relationship to one's body, instrument, peers, and broader culture. Such conclusions are also drawn by Berliner (1994). His overall claim is that one acquires the knowledge and skills called for in jazz improvisation chiefly through the combination of immersion in an acculturated community of practitioners and hours and hours of self-directed experimentation on one's instrument – that is, through a confluence of situated and embodied learning.

**Perceptual invariants**. We can view the notion of sound as a carrier of identity from a perceptual standpoint, in the same way that one might describe the recognition of a specific person's speaking voice, using the notion of *invariants*. As Gibson (1979) described it, the perception of an environment that both changes and persists involves extracting invariants of structure from a continually varying bath of stimulation, and noticing the variation relative to these underlying invariants. In a similar vein, Shaw and Pittenger (1978) suggested the possibility of invariants that are functions of time, such as a repetitive motor. They distinguished between *transformational* and *structural* invariants. Transformational invariants are relational aspects of the information that specifies the identity of a particular pattern of change; hence one might hear speaking as a certain kind of use of the human voice, as opposed to singing, for example. By contrast, structural invariants are relational properties specific to the source object undergoing a particular style of change. These properties might include the invariant structural features of this *particular* speaker's vocal tract and other parts of her body that might give rise to the production of vocal sound. It seems that they might also include learned usage of those organs in specific ways, such as regional accent and vocabulary. For these structural invariants are not only confined to body; they also involve memory, history, personal choice – in short, the person's individuality. In African-American musical contexts, especially in the case of musical improvisation, a personal sound contains the musical trace of the musician's body usage (as with saxophone timbre, for example), as well as of his or her conceptual approach (as it is conveyed in the course of improvisation). A variety of musical attributes, ranging from instrumental timbre to improvised musical choices, may be seen as manifestations of underlying invariants in the musician's embodied worldview. Hence we may align the personal sound construct with these perceivable structural invariants.

### Paralinguistics, Performativity, Signifying

Some recent results from psycholinguistics illuminate the role of *paralinguistic* phenomena such as hand gestures in conversational speech. McNeill (1998) examined footage of individuals explaining cartoon scenarios, and found that hand gestures served not just to illustrate but to augment what was described verbally. In this way, language and gesture are seen to be *coexpressive*, meaning that the sonic and visual dimensions sum to form a larger meaning than either one independently would convey. McNeill's aim is to challenge the modular information-processing model of cognition, which in its strictest sense does not treat different processes as interacting. In the modular view, complex cognition is broken down into something resembling subroutines or modules in a computer program. Rather, claims McNeill, visuals and sound connect to produce a whole; they need to be treated as interacting, not modular, so that context can be incorporated.

Similarly one might imagine that visual and other contextual factors in a musical performance co-articulate musical meaning along with the sonic trace. In keeping with some post-structuralist scholarly work, we may call these elements *performatives*. The term "performative" has grown to encompass a wide range of phenomena, but it was first coined by J. L. Austin (1962) in reference to a certain class of speech act, namely a verbal utterance that fulfills a function by virtue of its being spoken. A commonly cited example is a wedding officiant's statement, "I now pronounce you husband and wife." In speaking that declarative sentence, the officiant also executes an action – one that can only be done by enunciating that statement. Hence the utterance accomplishes more than conveying a true or false statement; it is also an act with non-linguistic, real-world consequences. Later, Austin and others (Forguson 1969; see also Parker & Sedgwick 1995) pointed out that "there really is no good reason to distinguish between performative and other sorts of utterances at all. All utterances have their 'performative' role to play in discourse..." (Forguson 1969: 419) Hence it was acknowledged that a variety of additional meaning arises via inflection, stress, and most importantly, situational circumstance. When the Pope says, "Bless you," as an Englishman kneels at his feet, it means something quite different from when a Polish immigrant says, "Bless you," to a sneezing American passerby, even if the intonations, inflections, and intensities are identical. The dimension along which these two cases differ may be called the performative dimension.

In a related essay, Barthes (1977) pointed out that performance of composed music also carries this "extra" dimension. In addition to the meaningful intramusical dynamics, supplemental meaning is generated by the presence of a music-making body, and the sonic traces it leaves behind. Hence the "grain" of the voice, by announcing the vocalist's physical presence, signifies a rupturing of the disembodied, self-contained world of the classical work. The personhood of the performer insinuates its way into (classical) music performance through its roughness, its resistance, its departure from the ideal. The physicality and resistance of the voice point to its producer, the performer, and to the act of it being produced. The grain of a musical performance reminds the listener of the physical sensation of using the voice, or other parts of the body: "The 'grain' is the body in the voice as it sings, the hand as it writes, the limb as it performs." (Barthes 1977: 188)

These variable features of performance give music no small part of its expressive powers. Dunn and Jones (1994) provide a multiplicity of perspectives on embodiment in Western female vocal music. It is seen that the meaning of a vocal utterance is constituted not simply by its semantic content but also by its sonorous content. By focusing on the essential role played by the 'purely sonorous' (i.e. musical, non-verbal) features of the female voice in 'the construction of its non-verbal meanings,' by making explicit the fact that sonorous features must be conceptually linked to the production of vocal sound through a person's body, and by studying the various factors of acculturation that affect the reception of vocal sound, the authors provide a complex account of the status of sound as "performed." Vocal meaning derives from 'an intersubjective acoustic space,' and any attempt to articulate that meaning must necessarily 'reconstruct ... the contexts of ... hearing'." We thereby "recognize the roles played by 1) the person or people producing the sound, 2) the person or people hearing the produced sound, and 3) the acoustical and social contexts in which production and hearing occur. The 'meaning' of any vocal sound, then, must be understood as co-constituted by performative as well as semantic/structural features." (Dunn 1994:2-3) The performative dimension in music and speech, as described above, overlaps many other salient musical/linguistic dimensions (the interactive, the processual, the semiotic) to some degree. In fact it could be said that all socially situated human acts – music, speech, writing, sport, political acts, worship, etc. – contain a performative element, for they all involve non-verbal meanings generated by situated bodies in intersubjective cultural spaces.

How might performativity operate in music that has no score? Indeed one might ask, how might performativity *not* operate in music that only exists when performed, when in motion? To be sure, improvised music must project a great amount of meaning along the performative dimension, since so much of how one listens to it depends on one's understanding of its contextual factors. These aspects might include the role of improvisation as a trope for the present, interactivity as the conveyor of a shared sense of time, and the attention to the role of the body and the specific surroundings in music-making activity. In my experience, I have noticed one of the most common questions asked by novice listeners after a jazz performance to be, "What percentage of that was improvised?" One requires some understanding of the conditions and assumptions that give rise to the sounds one hears; the notes and tones do not explain themselves to an outsider.

In African-American musics, it could be said that a large part of this performative dimension coincides with the dimension of Signifying (Gates 1988). The stylized term is given dozens of meanings that play off of each other; Signifiying can refer to "a way of encoding messages or meanings which involves, in most cases, an element of indirection... [i.e.] an alternative message form [which] may occur in a variety of discourse ... Signifying is troping." (Gates 1988:80-81) The governing idea behind Signifying is verticality – that is, the free play of rhetorical associations to conjure up multiplicities of meaning beyond the literal. In theorizing about African-American literary discourses, Gates identifies the importance of Signifying – i.e. of verticality, of inter-textuality, of history, of multiplicity, of reference to shared knowledge – in the production and communication of meaning.

It seems fair to align the concept of Signifying with the notion of the performative, which also denotes the nonliteral meanings conjured up through nonverbal channels. Hence, just as the very performative activity of Signifying accounts for a significant portion of the generation of meaning in spoken and written language, one might expect a similarly large amount of information to be conveyed via Signifying in the performative aspects of music. For example, much meaning in African-American music is generated through continual referencing, be it explicit or implicit, of a background wealth of cultural information. In jazz, this might happen on a surface level by quoting a well-known melody in the course of an improvised solo, or by paraphrasing a melodic or rhythmic fragment that somebody else just played. Or, it might happen at a subtler, more deeply coded level, through a construct such as timbre (e.g. does Eric Dolphy's alto saxophone sound reference that of Charlie Parker? **[CD-14]**), or in the way that a piece is constructed (do Ornette Coleman's compositions Signify on the melodic gestures of bebop? **[CD-15]**), or most abstractly, in a musician's "sound" or "attitude" (does Miles Davis's sense of space, timing, and melody convey a sense of the blues? **[CD-16]**). In hip-hop it may occur in the choice of musical material, as with the widespread, often blatant use of samples of classic funk and soul tunes **[CD-17, 18]**, or in lyrics, as when MC's (rappers) *represent*, addressing their origins or their home turf

**[CD-19]**. Also, an MC is often characterized by his or her *flow*, a flexible concept (analogous to *sound* in jazz) that can refer in different contexts to rhythmic acuity, lyrical prowess, or general persona **[CD-20]**. The heightened role of such performative parameters in these various African-American musics provides a case for the conceptual role of sociocultural situatedness in their reception, perception, cognition, and production.

In light of this discussion of performativity, it seems that the application of Meyer's (1953) concepts of musical meaning to jazz leaves something to be desired. Meyer theorized that emotion and meaning in music boil down to the deferral of expectations implied by intramusical dynamics. In developing this theory, Meyer adhered to a rather arbitrary distinction between "designative meaning," in which a stimulus may become meaningful by referring to something that is different from itself in kind, and "embodied meaning," in which the reference is to something like the stimulus itself. Roughly, this distinction corresponds to the difference between the ethnomusicological understanding of music as a body of references to certain cultural practices on the one hand, and the objectivist understanding of the "intrinsic" dynamics of the music itself, somehow disembodied from its cultural context. Meyer chooses (and thus believes it possible) to focus on the latter, which, ironically, he labels "embodied meaning." Whatever its applicability or inapplicability to European concert music, I believe that this distinction collapses in much African American music, which is often arguably as much a *definition* of the oral culture that produces it as it is an *outgrowth* of it. For a music that is so conscious of its own origins, one cannot neglect the dimension of meaning made possible by Signifying – namely the possibilities of multiplicities of meaning set forth by the metonymies of oral culture itself. To deny the dimension of history, of the Signifying possibilities inherent in the play of one piece of music against the memory of its predecessors, is to rob the music of a greater part of its meaning.

## Time and Temporal Situatedness

Yet another fundamental consequence of physical embodiment and environmental situatedness is the fact that *things take time*. The concept of time must structure our conception of physically embodied cognition from the start. Smithers (1996) draws a useful distinction between processes that occur "in-time" and those that exist "over-time." The distinction is similar to that between process-oriented activity, such as speech or walking, and product-oriented activity, such as writing a novel or composing a symphony. In-time processes are *embedded* in time; not only does the time taken matter, but in fact it contributes to the overall structure. The speed of a typical walking gait relates to physical attributes like leg mass and size, and shoulder-hip torsional moment; this is why we cannot walk one-tenth or ten times as fast as we do. Similarly, the rate at which we speak exploits the natural timescales of lingual and mandibular motion as well as respiration. Accordingly, we learn (or more likely we are hardwired) to process speech at precisely such a rate. Recorded speech played at slower or faster speeds rapidly becomes unintelligible, even if the pitch is held constant. The perceived flow of conversation, while quite flexible, is sensitive to the slowdown caused by an extra few seconds taken to think of a word or recall a name.

Over-time processes, by contrast, are merely *contained* in time; the fact that they take time is of no fundamental consequence to the result. Most of what we call computation occurs over time. The fact that all machines are considered computationally equivalent regardless of speed suggests that time was not a concern in the theory of computation, and that the temporality of a computational process is theoretically immaterial. In so-called "real-time" systems, typically one exploits the blinding speed of modern microprocessors to allow computation so fast that one doesn't notice how much time is taken. However, this is not what the mind does when immersed in a dynamic, real-time environment; rather, it exploits both the constraints and the allowances of the natural timescales of the body and the brain as a total physical system. In other words, Smithers (1996) claims, *cognition chiefly involves in-time processes*. Furthermore, this claim is not limited simply to cognitive processes that require interpersonal interaction; it pertains to all thought, perception, and action.

In intersubjective activities, such as speech, musical performance, or rehearsal, one remains aware of a sense of mutual embodiment. This sense brings about the presupposition of "shared time" between the listener and the performer. This sense is a crucial aspect of temporality of performance, especially from a communication point of view. The experience of listening to music is qualitatively different from that of reading a book. The experience of music requires a "co-performance" that must occur within a shared temporal domain (Schutz 1964). This sense of co-performance is made literal in musical contexts primarily meant for dance; the participatory act of marking time with rhythmic bodily activity physicalizes the sense of shared time. The meaning derived from such physical participation contrasts with the "contemplative" mode of music listening as practiced in the European concert hall, in which any kind of body motion by an audience member is typically met with negative social feedback.

### The Temporality of Musical Performance

We may consider how various cultural models may affect cognition by framing, though not entirely dictating, notions of time. Shore (1996: 62) gives examples of temporal models that can orchestrate culturally specific time frames. These include incremental, decremental, cyclical, rhythmic, and biographical models, as well as context-framing devices. A jazz performance might contain many of these models simultaneously. It might utilize a cyclic form (such as a song form or "chorus"), in which time is broken into rhythmic segments (beats, subdivisions, meter), and of which chunks may form complete episodes (such as an individual solo, which might last a number of choruses). The entire performance might be governed by a sense of appropriate length, thereby involving an overall incremental model.

Furthermore, the performance situation might be understood as a context-framing device. In his study of music of a certain community in South Africa, ethnomusicologist J. Blacking wrote, "...Venda music is distinguished from nonmusic by the creation of a special world of time. The chief function of music is to involve people in shared experiences within the framework of their cultural experience." (Blacking 1973: 48) There is no doubt that this is true to some degree in all musical performance. We can take this concept further in the case of improvised music. The process of musical improvisation in a jazz context can be seen as one specific way of framing the shared time between performer and audience. The experience of listening to music that is understood to be improvised differs significantly from listening knowingly to composed music. The main source of drama in improvised music is the sheer fact of the shared sense of time: the sense that the improviser is working, creating, generating musical material, in the same time in which we are co-performing as listeners. Part of what we seem to experience as listeners to any music is an awareness of the physicality of the "grain," and a kind of *empathy* for the performer, an understanding of effort required to create music. In improvised music empathy extends beyond the concept of the physical body to an awareness of the performers' coincident physical and mental exertion, of their "in-the-moment" (i.e. in-time) *process* of creative activity and interactivity. Thus improvisation heightens the role of embodiment in musical performance.

Time framed by improvisation is a special kind of time that is flexible in extent, and in fact carries the inherent possibility of endlessness, similar to that pointed out in Shore (1996) in the case of baseball games. Instances like Paul Gonsalves's 27 choruses (over 6 minutes) of blues on Ellington's "Diminuendo and Crescendo in Blue" **[CD-21]** and Coltrane's sixteen-minute take on "Chasin' the Trane" **[CD-22]** – significantly, both live recordings – attest to the power that the improviser wields as framer of time, deciding both the extent and the content of the shared epoch.

## Temporal Situatedness & Musical Form

Accordingly, music that privileges improvisation requires a different concept of musical form from music that is through-composed. In the former case, musical form can be described in terms of temporal situatedness. It is enlightening to consider the concept of form in the classical improvised music of India **[CD-23]**:

> Syntactical forms are virtually unknown in the music of India. Instead we hear long, cyclical, chain structures and a general progression of organic growth that reveals the guidance of quite different formal models and metaphor. The tactics of form go hand-in-hand with the prevailing models of structure: hierarchical and syntactical forms are naturally implemented by such tactics as contrast, parallelism, preparation, rise, transition, and the like; serial forms [as in Indian music], however, tend to be modular, decorative, incremental, progressive, and open-ended. The Indian version of musical structure tends to emphasize variation of the module: by permutation of its elements, by inflation and deflation of patterns, by pattern superimpositions, and by progressive organic development. (Rowell 1988)

Improvised African and African-American music can share many of these traits, particularly in the long-term organization of material. The major role of improvisation in many oral musical traditions, combined with the important function of groove, make possible alternative notions of musical form that do not conform to the recursive hierarchies of tonal-music grammars. A teleological concept of form, in which the meaning of music is taken to be its large-scale structure, may be replaced with an alternative, modular approach, in which the meaning of music is located in the free play of smaller constituent units. Such notions of musical structure appear in many African and African-American musics. Instead of long-range hierarchical form, the focus is on fine-grained rhythmic detail and superpositional rhythmic hierarchy. Thus, large-scale musical form *emerges* from an improvisatory treatment of these short-range musical ingredients – that is, from the in-time manipulation of simple components in a modular conceptual organization.

A prime example is James Brown's frequent practice of "taking it to the bridge" **[CD 24]**. A given piece might consist of two different musical spaces or grooves, the transitions between which are cued musically by the vocalist. Hence each section may be arbitrarily long, since the only thing to delineate it is an improvised cue to the next section. Before the performance of the piece, Brown and his band may not know exactly what will happen when; rather, they know what the raw materials are and how to manipulate them during performed time. As another example, jazz drummer E. W. Wainwright (private communication, 1997) described to me a practice of creating large-scale temporal form out of an open-ended though metrically distinct musical environment, as it was done by John Coltrane's legendary quartet in the early 1960s (cf. the title track to *Transition* **[CD-25]**). In such pieces, the group would be improvising in 4/4 time, using a certain collection of pitches as a loose framework, such as a mode over a D pedal point. Eventually, formal small-section boundaries would emerge by the systematic doubling of the musical period. As was told to Wainwright by Elvin Jones (the quartet's drummer and Wainwright's teacher), the group would initially accent the beginning of every four bars, using intensity as well as rhythmic, melodic, and harmonic parameters. As the piece unfolded, they would expand the period to eight bars, then sixteen, and so on. The larger the period became, the greater heights the intensity and dissonant tension could reach, and the more effective the unified release at the beginning of the next period. As Jones told Wainwright, this practice emerged organically over the course of hundreds of improvised performances, never having been discussed verbally by any band members. These two examples suggest that aspects of musical form can stem from the sense of shared, lived time, and the way variations are carried out while embedded in time.

In addition, in jazz and other musics, intramusical hierarchical organization may very well be decentered in favor of referential, associative, or functional relationships (Honing 1993). Formal emphasis might be more on repetition, on reference to a shared body of knowledge, or on maintaining a relationship to a composite rhythmic pattern, and less on the recursive derivation of a background meaning by grouping sections into chunks. In other words, the emphasis on "the moment" as a consequence of embodiment allows for different kinds of formal derivation.

As an example, consider again the saxophonist John Coltrane. He was known early in his career for playing long, impressive, exploratory solos that projected excitement and forward motion nonetheless, full of blisteringly fast runs, filigrees, and arpeggios, as on Monk's "Trinkle, Tinkle" **[CD-26]**. Coltrane's improvisations were less hierarchically unified than was typical for the idiom, and more serial or sequential; sometimes it was said that his solos lacked direction or went on too long. Many have tried to establish "motific development" in Coltrane's individual improvisations as that which creates structure (Dean 1992, Jost 1981), but it seems to me that this is merely a consequence of a greater formation – Coltrane's "sound," his holistic approach to music, which yields these elements. I do not wish to imply that Coltrane had no mind for "structuring" an individual solo; but these sorts of analyses stem from the critical tools of the listener rather than the improviser. As a musician, I personally believe that the improviser is concerned more with making individual improvisations relate *to each other*, and to his or her conception of personal sound, than he or she might be with obeying some standard of coherence on the scale of the single improvisation.

In this way, the temporally situated view of music cognition facilitates a nonlinear approach to musical narrative. Musical meaning is not conveyed only through formal hierarchies, motific development, contour, and temporal deferral of expectations; it is also *embodied* in improvisatory techniques. Musicians tell their stories, but not merely in the traditional linear narrative sense. An exploded narrative is conveyed through a holistic musical personality or *attitude.* That attitude is conveyed both *musically* through the skillful, individualistic, improvisatory manipulation of expressive parameters in combination, as well as *extramusically* in the sense that these sonic symbols "point" to a certain physical comportment, a certain way-of-being-embodied. In improvised music practices in general, the emphasis tends to be not on the single isolated performance but on the developing body of concepts or expressions as it exists over long periods of one's life. The only invariable guideline for a solo or a group improvisation is to feel in the end as though you have "said something" (Monson 1996). The details of how this is accomplished are as variable in music as they are in speech.

### Embodiment as a Complement to Cognitivism

In concluding this lengthy chapter, it should be pointed out that the claims of embodied and situated cognition operate *alongside* the symbolic cognitive methods that traditional knowledge-based systems might exploit. There is a clichéd distinction between literate and oral cultures that runs, "The literate man stores information through writing; the oral man stores information through physical assimilation: he *becomes* the information." (Sidran 1971: 9) Actually, both kinds of "storage" occur in everyday life in the literate world; information is distributed among various embodied and situated dimensions, including learned sensorimotor patterns, written or memorized symbolic information, and social customs. These delineations correspond to three "vantage points" that we can consider to be part and parcel of cognition: "the development of the individual, the local support conditions leading to the mastery of the symbol system and materials of the domain, and the cultural setting which gives meaning and structure to the entire expression." (Davidson & Torff 1992: 120-121)

As an example, consider a concert violinist's performance of a composed piece from the standard repertoire. We have now discussed many ways of studying the violinist's performance. In working on the piece, she develops a personal, non-transferable interpretation, which emerges as a highly specialized behavior from her hours of physical practice. A performance of this piece would represent both the retrieval of memorized, symbolic information and the enaction of physically assimilated behavior. The piece itself may possess a deeply hierarchical intramusical logic full of formal interplay, and the violinist may highlight these formal elements via certain expressive performance choices. Other performance variations in timing, dynamics, and intonation might stem from nervousness, fatigue, caprice, or the soloist's attempts to be audible over the polyphonic backdrop of a hundred accompanying musicians. One may also study how the performance may be framed socially in a concert hall as elite high-culture spectacle, rich in performative elements, from the "alienating social ritual of the concert itself," to the enhanced social distance between the audience and "the 'artist' in evening dress or tails," to the "listener's poignant speechlessness as he/she faces an onslaught of such refinement, articulation, and technique as almost to constitute a sadomasochistic experience." (Said 1991: 3) All of these dynamics inform our reception of the performed music.

Improvised music provides us with another example. A jazz pianist improvising over a standard tune would certainly require a working understanding of functional harmony, meter, and form, both in general and specifically applied to the song in question. This knowledge would fall mostly in the abstract, symbolic domain. However, a variety of other requirements draw on the situatedness of the pianist's body vis-à-vis the piano. These elements include sensorimotor functions like the placement and control of hands on the keyboard and foot on the pedal, the coordination of the digits, and the harnessing of these activities to an internally entrained pulse. In addition, cultural and learned factors such as the musician's relationship to the instrument, genre, and associated lifestyle may find their way into the improvising process in the form of personal sound, choice of musical material, or adherence to preestablished norms. The artist might signify on established versions of the piece (including his own, as when pianist Ahmad Jamal performs "Poinciana" today, knowingly quoting and modifying his distinctive, extremely popular 1958 version **[CD-27]**). Or, the pianist might highlight his or her version by sheer contextualization (as with the quintessentially modern Thelonious Monk's old-fashioned, buoyant stride-piano version of the then-40-year-old tune popularized by Louis Armstrong, "I'm Confessin'" in 1963 **[CD-49]**). The pianist might deploy expressive timing and accent manipulations that highlight the relationships between the performed rhythm and the pulse, the melody and the chord changes, the melody and the (unheard) lyric, the right hand and the left hand, the left-hand pattern and a bassline from another piece, or any number of other kinds of rhythmic interplay.

The often implied characterization of the social and symbolic as high-level and the embodied as low-level is misleading, for these functions may interact with each other bilaterally. In particular, one should *not* claim that the high-level processes "direct" the low-level, for in some cases it is not clear that there is any such hierarchical organization. Indeed, the tendency to posit such a hierarchy stems from our prejudice of mental processes as more "elevated" than physical ones. Further consideration of the example at hand suggests that the cognitive organization between "body" and "mind" can be *heterarchical*, or non-hierarchically distributed. For instance, in the midst of an improvisation, the temporally situated pianist is always making choices. These choices are informed not simply by which note, phrase, or gesture is "correct," but rather by which activities are executable at the time that a given choice is made. (Similar observations have been made by Sudnow 1978.) That is, a skilled improviser is always attuned to the constraints imposed by the musical moment. This requires an awareness of the palette of musical acts available in general, and particularly of the dynamically evolving subset of this palette that is *physically* possible at any given moment. In this way, for example, the improvising pianist is more likely to choose piano keys that lie under her current hand position than keys that do not. Such weak constraints (which may be overridden, with physical and melodic repercussions) combine holistically with formal directives such as melody and harmony (which may also be overridden). Indeed, improvisation – musical and otherwise – may be understood partially as a dialectic between formal/symbolic and situational/embodied constraints.

Hence the functions of situated or embodied cognition neither replace wholesale, nor obey blindly, but rather *supplement* and *complement* the abstract, symbolic cognitive processes that we usually associate with "thinking." In this chapter, I have attempted to show how the theoretical concepts of embodied and situated cognition can similarly enhance the study of music perception and cognition.

# 5. On the Perception of Meter

In this chapter I shall turn to the concept of meter, and its role in shaping our musical perceptions. The first question to ask is why we have meter at all. I shall stress that meter has some special status among the various time spans of music, and in particular that its function is both established and constrained by our embodied perceptual and cognitive abilities. Later I will discuss some subtleties in the question of how we perceive meter in situated musical contexts, by examining its role in instances of West African music.

### Rhythmic Timescales

The extremely influential techniques of Schenkerian analysis (Schenker 1979) are characterized by concepts of hierarchy, progression, and recursion. Musical structure is understood on three levels: foreground, middle ground, and background. Analysis is seen as a process of isolating, identifying, connecting and integrating these three levels of musical perception, aiming for reduction of musical surface structures to a series of nested levels. The perspective is hierarchical in its layered organization and progressive in the implications between events within levels. The role of recursion is metaphoric, in that a large piece can be seen as a giant cadence, or composition as the large-scale embellishment of an underlying harmonic progression. The *Ursatz*, or underlying background structure, is seen as the point of analysis and the mark of a good composition. The theory presupposes that music's purpose is its background and not its foreground. This assumption seems to bias the very concept of analysis in favor of long-term form; ultimately, the entire theory legitimizes its own goals.

According to Lerdahl and Jackendoff (1983), who have a strong allegiance to the Schenkerian point of view, grouping structure in music is said to be recursive because it is supposed that there is a uniformity from level to level. That is, grouping structure can be elaborated iteratively and indefinitely by the same rules (*ibid*: 16). They suggest that "the elements of metrical structure are essentially the same whether at the level of the smallest note value or at a hypermeasure level (a level larger than the notated measure)." (*ibid*: 20) There are many objections to such a formulation. It must be granted that this is a theory of perception and cognition, not of composition; hence it does not presume that such uniformity exists in the music, but rather that in listening to music we tend to expect such regularities. However, an immediate problem with this kind of analysis is that it presupposes a kind of classic large-scale order or unity; notes hang together coherently as melodies, phrases answer other phrases, sections balance each other. The concept of coherence (often phrased by contemporary composers as "what works") is an old problematic, ascribable to many structuralist and pre-structuralist literary theories. It has been deconstructed by the observation that coherence or artistic unity is projected by some theory of coherence, and not by any intrinsic properties of the work itself (Street 1989). As another critique of the notion of recursive hierarchy in music, Narmour (1990) supposes instead that there can be multidirectional networks and communication between levels of the hierarchy.

In any case, if taken literally, the claim of recursive grouping hierarchies in the temporal domain must be rejected on simple cognitive grounds, because of the qualitative differences among the three kinds of memory, namely the echoic store, short-term or working memory, and long-term memory. As Brower (1993) has pointed out, echoic memory covers the immediate timescale of rhythmic activity ("foreground rhythmic structure"), whereas short-term memory covers meter and phrases ("middle-ground"), and long-term memory covers the timescale of musical sections and movements ("background"). The three different types of memory involve different kinds of processing. We *entrain* to a pulse based on the echoic storage of the previous pulse and some matched internal oscillator periodicity, we *feel* the relationships among strong and weak beats (accentual meter), we *count* times between phrases or bars (metric grouping), and we *associate* features of one section with memories of a previous section that is no longer represented in serial order. So for example, we can experience syncopation at the metric level, and so-called expansion and contraction at a higher level, but these are qualitatively not the same cognitive phenomenon. Hence it appears that the notion of recursion does not apply literally as we switch among these different timescales, because they operate in very different ways from a cognitive perspective.

The above arguments suggest that we tend to organize music in terms of certain ranges of musical times that are commensurate with the timescales of human activity, in keeping with the remarks of the previous chapter on embodiment. Among the pieces of the rhythmic puzzle that we discussed were three qualitatively distinct embodied regimes of temporal perception: the quick detail of speech and digital motion, the steady tactus-level pulse, and the phrasal breath. These distinctions, together with the arguments about grouping and memory discussed above, all enter into the following discussion of meter.

**Pulse Perception**

One of the most crucial aspects of rhythm perception that has been established experimentally is the mind's role in entraining to a periodic input signal. This ability allows us to tap our foot to music. Given a periodic auditory pulse as input, we have a low tolerance for error in the signal; we easily detect any sudden deviation from periodicity above a small threshold. One might suppose that the mind would detect such errors in any of three ways: (i) compare each inter-onset interval (IOI) to each previous IOI, (ii) match the periodic signal to an internally generated periodic oscillator, or (iii) match each IOI to an internal ideal period.

Schulze (1978) carried out the following experiment to test each possibility. Subjects listened to a periodic series of tones with one of three possible kinds of small temporal perturbations, as shown below. From elementary theoretical considerations, it was possible to rank the errors detected by each method for three different experimental conditions. Estimates of the relative discrimination index for each of the methods (i)-(iii) are shown for each kind of perturbation. Each line represents one experimental configuration, in which an isochronous series of tones or clicks, separated by an inter-onset interval τ, is modified by a perturbation Δ. In the experiment, the change occurred after a large number of unchanged intervals, to establish regularity. (Table adapted from Shulze 1978.)

| comparison patterns (successive IOI's) | | | | | | discrimination index | | |
|---|---|---|---|---|---|---|---|---|
| ($\tau$ =300ms; $\Delta$ =10 or 15 ms) | | | | | | (i) | (ii) | (iii) |
| 1. $\tau$ | $\tau$ | $\tau+\Delta$ | $\tau+\Delta$ | $\tau+\Delta$ | $\tau+\Delta$ | lo | hi | hi |
| 2. $\tau$ | $\tau$ | $\tau+\Delta$ | $\tau$ | $\tau$ | $\tau$ | med | med | lo |
| 3. $\tau$ | $\tau$ | $\tau+\Delta$ | $\tau-\Delta$ | $\tau$ | $\tau$ | hi | lo | med |

It turns out that the second alternative (ii) is correct. In other words, when we hear a regular rhythmic pulse, we don't match the intervals one by one, nor do we compare each interval to an internal ideal interval, but rather we *co-perform* this periodic pulse internally. This kind of internal, synchronized pulse generation is known as entrainment. A relatively simple phenomenon, entrainment is the essence of pulse perception. It has turned out to be a rather prodigious computational problem (Desain & Honing 1991).

**Meter**

The next step up from entrainment is a more complex phenomenon known as meter perception. Much music of the world is organized metrically, that is, according to a periodic grouping of pulses; this grouping may or may not be explicit in the musical surface. Most generally, meter is a periodic grouping of a musical time unit; it can connote *but does not strictly imply* a hierarchy of weak and strong beats. Meter denotes a subharmonic (or grouping) of a pulse, and might also suggest a higher harmonic (or subdivision) of the same pulse. That is, it can simultaneously group and subdivide pulses into regular units. For example, the time signature 6/8 denotes a cycle of two pulses, each divided into three equal subunits. The best understanding of meter is as a periodic grouping of perceived pulses – i.e. as a cognitive/perceptual phenomenon, not as an objective reality of the acoustic signal. However, this distinction is often elided, so we might speak of the meter of a piece of music. In doing so we would make some assumptions about the listener's and the musician's musical presuppositions. Studies of rhythm perception of both test patterns and musical performances show that meter itself is an ambiguous if not wholly imaginary property of the audio signal (Parncutt 1994). A given rhythm, whether intentionally musical or not, can evoke a number of contrasting pulse sensations and metric orientations in different listening subjects or in different trials with the same subject. In this regard, as we shall see below, West African and many African-American musics are no exception.

To circumvent this problem, one could create a distinction between the functions of meter in music production (as a generative principle) and in music perception (as an information-processing strategy). Or, more compellingly, one could avoid this distinction by defining meter as an internal periodic template that frames and temporally grounds both perceptions and actions. This latter definition is in keeping with a model proposed by Povel (1984). This model was proposed in response to experimental data involving both perception and imitation of temporal sequences, in which it was found that both musically trained and untrained subjects map temporal sequences onto an interval structure. In the model, the first step in temporal processing of music consists of a segmentation of the sequence into equal intervals (i.e., beats) bordered by events. In a second processing step, intervals smaller than the beat interval are expressed as a subdivision of the beat interval in which they occur. These beats and subdivisions serve to orient perception and action in a musical environment; meter may be seen thus as an *attentive* mechanism, as discussed below.

One small shortcoming of the above model, also admitted in a treatment by Lerdahl & Jackendoff (1983: 97), is that this definition of meter is challenged by asymmetric meters containing odd numbers of beats or an irregular pulse. In many cases (e.g. Eastern European, Greek, and Middle Eastern folk musics, South Indian classical music) one entrains to a series of long and short beats at the tactus level. These different kinds of pulses almost always occur in the duration ratios of 3:2 or 2:1, implying a common small temporal unit at a lower metric level. From these cases, it is apparent that a concept of metrical structure should not enforce regularity of pulse if it claims to apply universally, but instead should allow for a repeating group of irregular tactus-level pulses of either 2 or 3 constituent units.

Similarly, the accentual view of meter – in which a metric template is seen as always reinforced by accentual and microtiming variations – does not appear to carry over to many non-Western musics. One should not regard the global musical preponderance of "syncopation" (off-beat accents) as a vast set of exceptions to the "normal" accentual rules of meter, but rather as convincing counterexamples to such proposed accentual rules. Various signal-processing techniques have been developed to derive a pulse and a meter from musical recordings, but typically such attempts yield notoriously variable results for many such "syncopated" kinds of music such as jazz (e.g. Todd et al. 1998). This is because meter is not necessarily inherent in any audio signal. It is a perceptual and cognitive construction, derived from some perceived periodic patterning of perceived accents (including, paradoxically, accents imposed by the imagined meter itself) but also from some set of assumptions about meter.

Indeed, indicated meters appear in most Western sheet-music scores because of the composer's desire to exploit these assumptions, to allow the implications of a time signature to shape the performer's understanding and subsequent rendering of the piece. Aspects of meter can be conveyed in expressive musical performance. In Western tonal music, a given meter can suggest a certain template for phenomenal accent patterns and sub-tactus expressive timing variations. Meter provides a basic guideline for manipulating loudness, timbre, and timing in keeping with the composer's intention. Additionally, there exists a large vocabulary of diacritical marks to extend the connotations of meter. Due to all of this, we tend to speak of the meter of a piece objectively, because it is simply the time signature written to the left of the first bar of written music. This tendency is carried over erroneously to many ethnographic analyses of nonwritten musics, where the "meter" is simply that perceived by the transcriber. In a study of a foreign music, one should be aware of the cultural baggage attached to a musical concept like meter. There may be an analagous concept of orienting a periodic grouping of pulses, but it need not imply all the other attributes of meter that we recognize in Western music. Later I examine this possibility in the study of West African music.

Sometimes meter is also used to connote larger periodicities (as in "hypermeter"). Lerdahl and Jackendoff (1983) characterize metrical structure as a hierarchical nesting of beats to arbitrary degrees of recursive depth. This formulation covers too broad a range of timescales and presupposes temporal "unities" where they need not exist. In attempting to apply this theory to some rather simple musical examples (Lerdahl & Jackendoff 1983: 250-253), they encounter unexpected difficulties, and find that the supposition of large-scale periodicities beyond the barline level is much less robust and highly error-prone. It appears that meter functions most effectively as relatively small periodic groupings of the main beat, and quite variably at higher levels.

London (1997) has claimed that polymeter does not exist; that is, one cannot perceive two meters at once. Since meter exists only in the minds of the listeners and the performers, two meters thus cannot "exist" at once in the same mind. Instead one can hear ambiguous rhythms in which different metric candidates compete. The result, however, is not the simultaneous perception of two contrary meters but the canceling of the effects of one meter by the other. One can be aware of the possible presence of two different meters, and one can switch actively between perceiving them, as one can between the vase and the two faces in the classic visually ambiguous figure. One can also hear metrically vague music, in which no periodic grouping of pulses suggests itself. Certainly two people may simultaneously hear different meters in the same music. But one person cannot entrain simultaneously to two separate pulse groupings. In the classic figure-ground relationship, only one meter can function as the ground. This claim against polymeter seems to apply only to groupings of pulses in the vicinity of the tactus, however. One can be aware of multiple subdivisions of the tactus (as in some examples of Afro-Cuban rumba, where we perceive both quadruplet and triplet subdivisions in abundance **[CD-28]**). Similarly, one can simultaneously hear different-length groups of measures (as in James Brown's "I Don't Want Nobody to Give Me Nothing" **[CD-29]**, in which we easily apprehend that the vocals are in four-bar phrases and the horns and drums cycle a three-bar unit). Hence the claim against polymeter refers more to the level of *entrainment* than to perception in general - that is, to the act of *imagining the movement associated with the rhythm*, rather than the act of "hearing" the rhythm passively.

Our attentional capacities are sharply delineated by the timescale implied by meter, namely, that of the tactus. It is suggested by many (Gjerdingen 1989, Jones 1986, Jones & Boltz 1989, Jones & Yee 1993) that meter provides us with an *attending* mechanism – a temporal template against which to process information in time. Meter is seen as an *invariant* of the musical environment, computed from recurrent temporal patterns. As an invariant, it gives us a baseline from which to discern musical variation (such as musical rhythm and expressive timing) more efficiently. This view is supported in experiments in which subjects have more difficulty reproducing rhythms that do not have an obvious common time unit than those that do (Jones 1986). Modelling metric oscillations with neural networks is quite straightforward. It seems plausible that a small signal from a modest, neurally based metric counter could be used to modulate the performance of a much larger neural network according to a given meter, yielding a form of metrically modulated attention (Gjerdingen 1989: 78).

In differentiation among various timescales of music, Jones (1993) and Jones & Boltz (1989) has described different music-listening strategies as different kinds of attention, namely *future-oriented* (long-term, memory-based) vs. *analytic* (short-term, perception-based) attending. Actually, this distinction is more relational than absolute; there are intermediate regions between the two extremes. In any case, we can associate analytic attending with the timescale of musical meter.

These distinctions lead one to believe that, in its most useful form, meter measures a narrow range of rhythmic phenomena, limited roughly to what has been called the *psychological present.* Meter functions as an orienting principle among some small number of (often regular) tactus-level beats – between two and eight, roughly. A metric grouping of beats stretches over a duration between approximately 800 ms and 5 s, which lasts longer than the duration of the echoic store (< 1 s) but shorter than the span of working memory (~ 30 s). Apparently, then, we must scale back our definition of meter itself to denote only groupings of the tactus. Hence hypermeter is not a true species of meter, nor is a smaller subdivision of the main beat ("micrometer"). One will find apparent exceptions to this rule in some adagios from the tonal-music literature, where a written measure might stretch over ten seconds, or in some allegros or prestos, where a written measure might last a half-second or less. Typically in such cases, however, the written meter and the perceived metric grouping are far from identical, the latter being both somewhat variable from one individual to the next, and somewhat constrained by our attentional capacities.

In light of these arguments against both polymeter and hypermeter, it becomes clear that the perceptual groupings of the musical tactus into twos, threes, fours, and so on (perhaps up to eights) retain special status among rhythmic perceptual phenomena. This unconscious periodic grouping of pulses as some kind of *gestalt* would saturate at "the magical number seven plus or minus two" (Miller 1956) at which point perception with such ease becomes impossible. At metric lengths beyond this, we require either a *conscious* counting principle to hold our place, or a grouping procedure for "chunking" the beats into larger units. In addition, familiarity is an issue with meter. The sound and feeling of a given meter needs to be experienced, or else it will not function in this unconscious way. This is why, for example, five-beat groupings seem alien to westerners but are quite common in Indian popular music. More remarks on cultural aspects of meter appear in a later section.

### Models of Meter Perception

As mentioned above, meter perception has prompted a number of researchers in music cognition to develop models with varying degrees of success. Large (1994) has developed a compelling model for beat perception, which he has subsequently (1997) applied to meter perception. He modeled the phenomenon of beat perception as a pattern of coordination that arises between an internally generated periodic process (a self-sustaining oscillator) and a periodicity within a complex external rhythm (Large 1994, Large & Kolen 1994). The listener and musical input are treated as a single dynamical system consisting of the external (driving) rhythm and the internal (driven) oscillator; the attractor states of this dynamical system correspond to instances of entrainment. The internal oscillator's dynamics are assumed to adapt to the external periodicity, thereby modelling the robustness of ordinary beat perception to the subtle, systematic timing deviations that occur in human performance.

Large (1997) has also developed the notion that networks of these oscillators can combine hierarchically to respond to multiple periodicities in the driving signal. The model allows hierarchies of entrainment, whereby a fundamental pulse frequency (the beat) and its subharmonics (slower periodicities) coordinate with distinct phase relationships. This is presented as a model for meter perception. While it provides some compelling results that are robust under expressive tempo variation, it appears that this model neglects crucial aspects of our perceptual apparatus that cause us to distinguish between pulse and meter. As discussed above, the notion of a tactus provides a primary pulse range that sits squarely in the echoic-memory regime, whereas the larger timescales associated with meter invoke a different variety of cognition (Brower 1993). In his model, by treating all periodicities as equivalent regardless of length, Large fails to ground his model in the embodied temporal processing abilities of human beings. Hence, while the model works very well with the expressive tempo variations of tonal music, it is not clear to what degree it represents the ways that we perceive and interpret such structures.

Another recent effort to construct a model for rhythm perception (Todd et al. 1998) stems instead from a strongly embodied sensorimotor theory. The authors have developed a model that takes into account the natural biomechanical frequencies of the body and the temporal characteristics of the auditory system, and treats beat induction as an active rather than a passive process. The computational implementation takes audio input and synchronizes a simulated locomotive action. When applied to something as culturally contingent as music, however, the implementation encounters problems getting the phase relationship right for the pulse of the audio signal, not to mention the meter. The model seems to err on the side of overemphasis on the sensory-motor (embodied) component, failing to account for the possibilities of supplementing or even overriding this component with learned input. The only learning discussed is that which tunes the sensory systems to their embodied dynamics and the motor control systems to the temporal-motional properties of the *physical* environment. However, as musical beings, we know that musical perception involves more than physical and raw sensory input. We cannot forget the role of culture in shaping our perception, and this includes rhythm perception, as I shall argue below.

## Rhythm and Meter in West African Music

Studying what others have said about meter in the music of other cultures helps us understand the most basic, possibly cross-cultural functions of meter. In doing so, however, we have to unravel the cultural biases that all too often mark such cross-cultural studies. As a case study in the remainder of this chapter, I shall present my reactions to a wide variety of viewpoints from various ethnomusicologists, psychologists, and theorists about meter, pulse, and rhythmic grouping in West African music. In my opinion, many of these studies are characterized by two striking trends: a misapplication of music-theoretic and music-cognition models of rhythm and meter stemming from studies of Western tonal music on the one hand, and a well-meaning application of misunderstood African cultural and musical principles on the other. Although the models suggested by Lerdahl and Jackendoff (1983) and others go to great pains to keep meter and rhythmic grouping separate, the precedents set in African musicology by Western scholars tend to confuse the two. Furthermore, they tend either to ignore or to oversimplify descriptions provided by musicians and musicologists who come from West African cultures. Many highly influential treatments are fraught with misleading statements about such things as polymeter in African music. For example, Jones (1959) transcribed musical rhythms such that each phrase begins at the downbeat of a new meter, resulting in a jumble of meters of varying lengths and negligible perceptual salience. Below I attempt to address the contradictory viewpoints that works such as Jones's have engendered.

The Ghanaian percussionist C. K. Ladzekpo, a prominent performer and teacher of Anlo-Ewe dance-drumming in northern California, has written an extensive web document on the technique, practice, and cultural functions of this art form. He begins his section on rhythmic techniques thus:

> In a complex interaction of beat schemes of varying rhythmic motions, the human mind normally seeks a focal point. Among the Anlo-Ewe, one of the integral beat schemes is dominant and the rest are perceived in cross rhythmic relationship to it. This dominant beat scheme is considered the main beat because of its strong accents in regular recurrence that pervade and regulate the entire fabric...
>
> In practice, the beat scheme of four units is the most commonly used. At any given tempo, the rhythmic motion of this beat scheme is the most moderate (not too slow or fast) and the most convenient as a focal point...
>
> To better comprehend a main beat, it is structured so that it measures off equal increment of pulsations, the first of which normally bears an accent. These integral fractions or background pulsations are the major ornamental forces that give a main beat its distinct texture, flavor and character...
>
> The recurrent grouping of the main beats normally creates a fixed musical period or measure. While it is possible to create several measure schemes by varied groupings of the main beats, two types of such groupings are the most frequent in the development of Anlo-Ewe dance-drumming. The first most useful measure scheme consists of four main beats with each main beat measuring off three equal pulsations as its distinctive feature. The next most useful measure scheme consists of four main beats with each main beat flavored by measuring off four equal pulsations. These beat schemes are roughly equivalent to 12/8 time and 4/4 time in Western music.
>
> In contrast to the Western measure concept of accenting the first beat of each measure, the Anlo-Ewe concept maintains regular accents on all the main beats. However, there is a tendency to end phrases as well as the entire composition on the accented pulsation of the first main beat implying further movement or flow. (Ladzekpo 1995)

Ladzekpo clearly states that the chief rhythmic organizing principle is an isochronous pulse in an appropriate tactus range, typically grouped into metric units of four pulses, and also subdivided into smaller isochronous subdivisions of three or four. The first of each such group of subdivisions is usually accented, but the first of each group of four pulses is no more accented than the other three. However, the first pulse of a group of four is understood as the end (as opposed to the beginning) of many musical phrases. The first pulse thus has privileged status among the four pulses and corresponds somewhat to the Western notion of a downbeat, although it functions slightly differently. In the original document, all of Ladzekpo's statements are supplementally depicted in Western musical notation to minimize cross-cultural ambiguity.

The strong accents to which he refers should not be taken to mean sonic accents. As noted in Pantaleoni (1972) in reference to the music of this same ethnic group, the gradations in volume accent in West African drumming do not necessarily correspond to those implied by a European time signature such as 12/8 or 4/4. In fact, the volume accents alone may rarely reinforce any unified meter at all, by Western standards. To the "experienced" listener in the dance-drumming environment, however, accents can also be of a phenomenal (sonic, visual, tactile) or purely cognitive (metric) nature. Not only do we have some metric sense that frames the rhythmic activity with conceptual accents; also, we see the dancers' grounded footwork and the musicians' subtle body language, and we feel the sticks beating the drums and the ground shaking. Such paramusical input is operative in any performed music, much as paralinguistic phenomena are in most linguistic communication (McNeil 1998);

these aspects are consequences of situated cognition.

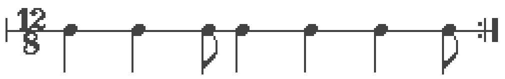
Much West African music is characterized by variable and repetitive rhythms played in conjunction with a basic timeline, such as a fixed pattern played on a bell. The bell part functions as a point of reference, a constantly repeating pattern that the other musicians use to orient themselves metrically. But one should be careful with the connotations of this last statement. Whatever parallel concept of meter exists in these cultures, it does not necessarily imply the same template of accents and stresses that a Western meter does. I am using meter in the sense offered in the previous section – namely, as a perceived periodic grouping of tactus-level pulses. The extra connotations associated with Western meter, such as an accentual template and a basis for expressive tempo modulation, are not included in this general understanding of meter.

(To complicate the issue, the presence of African diasporic musics in the West, and their more and more frequent rendering in Western notation, has smoothed out the difference. For example, a piece of commercially available sheet music can bear next to the tempo marking a suggestion such as "swing," or "slow funk groove," or even "African 6/8 feel," thereby implying a rhythmic treatment that is decidedly non-European, that is indeed derived from African rhythms (Wilson 1974), and that is not implied by the time signature alone. Hence there has been a gradual expansion of the possibilities of written music to include non-Western elements, to the point that one cannot draw a sharp boundary between Western notational concepts and non-Western musical practices.)

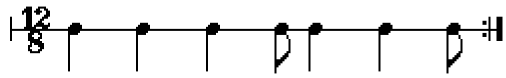
A guiding aesthetic principle in West African musics privileges rhythmic groupings that would challenge or "Signify on" the integrity of the pulse. It must be noted that an important function of the performance of drum music is the psychological balance achieved by internalizing seemingly conflicting rhythmic unities at the physical level (Ladzekpo 1995). Part of the crux of the activity of dance-drumming involves maintaining one's balance (i.e. one's sense of the steady isochronous pulse, metaphorically viewed as "purpose in life") in the midst of a torrent of interlocking cross-rhythms (metaphorically viewed as "obstacles") (Ladzekpo 1995). This does not necessarily imply that the intention of the composer/performer is to induce polymeter, nor does it imply the opposite. Rather, the musical construction is not framed in terms of meter, but in terms of crossed rhythmic groupings; these groupings can be serial or periodic or both, to use Parncutt's (1994) terminology. When two meters appear to the listener, it is because there are periodic groupings of short, serially organized rhythmic fragments or phrases, and their periodicity seems to imply a different meter from the primary one. In the sense described above, the music becomes metrically ambiguous. Most commonly it is some variety of triple meter that seems to appear over some variety of duple meter. But in context it is seen as a challenge, not an alternative, to the prevailing meter; frequently the latter is so obvious subjectively to the practitioner that it is not reinforced sonically very much at all. Waterman (1952) also emphasizes the importance of the subjective beat in African music, the underlying pulse that is not necessarily beaten out literally. He stresses that the appreciation of African music necessitates a "metronome sense." European music emphasizes the main beats, the upbeat and the downbeat. By contrast, African musicians assume implicitly that their audience is imagining these fundamental beats without difficulty; the audible musical material will fashion its rhythms around this mental beat, often using polyrhythms, cross-pulses, offbeat melodic accents, and much indirection and subtlety. Hence perceived groupings in the audio signal need not delineate the meter, and in fact frequently run counter to it.

So, where is meter in West African and related musics? If the Anlo-Ewe music described by Ladzekpo is any indication, meter is there, solidly, unproblematically, in the mind of the practitioner, and equally solidly in the mind of the acculturated listener, even if it is not obviously apparent in some objective rendering of the acoustic signal. One learns the main beat, its subdivisions, and its metric grouping; and then one learns the rhythm of the bell pattern, which simultaneously groups the main beats into larger units of four and subdivides them into smaller units of four or three. This simultaneous grouping and subdivision denoted by this referent rhythm suggests a way of orienting one's attention and physical entrainment to the musical signal; i.e., it connotes a set of assumptions in much the same way that a European time signature connotes a related but distinct set of assumptions. When learning to perform other rhythms in this music, one's attention is continually directed towards the concurrent bell pattern. This is done not because the asymmetric bell supplants the regular pulse, but because the bell gives the most consistent depiction of the pulse and its organization. Sonic accentual reinforcement is unnecessary and (uncoincidentally) is not a guiding aesthetic.

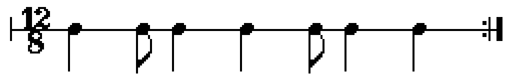
Thus the main beat and its metric grouping are articulated in a rather indirect fashion – not with a continual parade of accentual reinforcement, but with suggestion and complexity. The meter is *encoded* in the rhythm itself. The way in which this is done is unambiguous but highly culturally specific. (In the audio examples that follow, a click track is heard alongside the written patterns, with the high click on the first beat of the written measure.) For example, the standard 12/8 bell pattern of the Anlo-Ewe people is heard phrased as **[CD-30]**



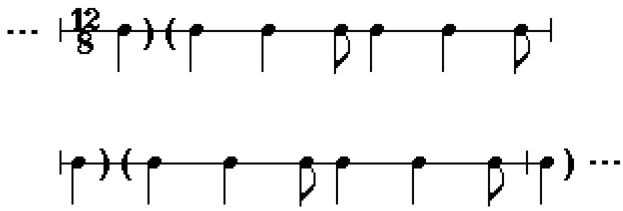

whereas the same bell pattern would be heard with a different starting point by the Yoruba people, as **[CD-31]**




and by yet another ethnic group as **[CD-32]**

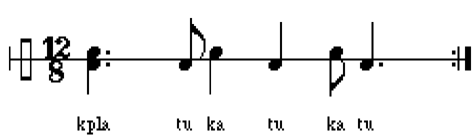



Furthermore, the patterns are not necessarily grouped according to the barlines; for example, the Yoruba pattern is said to be grouped such that the *final* note is the so-called downbeat (Anku 1992), i.e. **[CD-33]**




where the parentheses delineate the grouping over the barline. (In the audio example, the grouping is delineated by a slight accent in intensity on the first note of each group.) Note that respective boundaries implied by grouping and meter do not reinforce each other in this case; this possibility is allowed for in the respective definitions of grouping and meter. This example shows that both the placement of the tactus and the grouping of elements varies from one community to the next with respect to a single rhythmic pattern.

Typically the constituent beats and subdivisions are not literally "counted" by the African practitioners. Hence literally there is no "1." Rhythms are frequently represented instead as linguistically derived mnemonics, an obviously favored technique in an oral culture. An example is the simple rhythm known in the West as "three over four" and represented by Ladzekpo as **[CD-34]**



This example is onomatopoeic, but there are many documented examples of rhythmic mnemonics with semantic meaning as well (Anku 1992).

(There has been a chicken-or-egg debate about the status of these linguistic phrases: are the rhythms derived from language, or are the phrases arbitrary mnemonic devices to represent the rhythms? Ladzekpo (1995) stresses the ubiquity of symbolic and metaphoric encodings in African music, which seems to support the former view. In any case, the fact that linguistic proverbs have been long associated with many rhythms in many African drum musics suggests that the music *carries* these meanings, whatever their etiology.)

The bell patterns mentioned above may be seen as musically encoded time signatures in a self-consistent musical language, implying a different set of basic assumptions than are used in the West. In many West African musical contexts, one learns the proper placement of rhythms not merely in relation to an underlying abstract meter but to the bell pattern itself; no explicit reference need be made to an underlying metric abstraction. An outcome of this approach is the organization of phrases relative to the so-called downbeat, referred to by Ghanaian musicologist W. Anku as the Regulative Time Point (RTP) (Anku 1992). Instead of referring to a specific subdivision of a specific beat in the metric hierarchy, a rhythmic phrase is said to have a specific relationship to the RTP. Improvisations based on this phrase will maintain this relationship, yielding the repeated use of what in Western terms might be called an anacrusis or "pick-up." In the figure below are some examples of rhythmic themes with different RTP relationships, i.e. different phase relationships to the underlying meter. The groupings are specified by the parentheses around each theme. A phrase with RTP *n* corresponds to the RTP placement on the *n*th relevant subdivision from the beginning of the phrase. For example, the bell pattern is designated as RTP 9 because the ninth eighth note of the phrase (beginning after the grouping parenthesis) is the RTP or downbeat, as indicated by the central hash mark.


The examples are taken from Anku (1992), who describes these rhythmic phrases as *structural sets* because of the way they organize subsequent rhythmic material. The RTP corresponds to the downbeat in the indicated time signature. I have mixed representations here to ease understanding; Western meter is used to explain the West African RTP notion. (In the audio example **[CD-35]**, again you hear a metric click track to guide the ear, and again the groupings are delineated by a slight accent in intensity on the first note of each group. Once established, the bell pattern is sustained throughout, while themes 1, 2, and 3 are presented in succession.) Again, these phrases are often linguistically based, and thereby carry the encoded meaning of a proverb. The placement of the RTP in such a phrase may suggest the added element of linguistic emphasis (which in music we might call metrical accent), as in the difference between "I *walked* to the store" and "I walked to the *store*."

Hence, not only does the RTP relationship thoroughly specify when to begin a phrase in the metric cycle implied by the bell; more importantly, this relationship delineates the distribution of emphasis. This is similar but not identical to the accentual template implied in European music by a time signature, and in particular by the downbeat. However, the RTP relationship carries a stronger implication of repetition and cyclicality than is connoted by anacrusis and downbeat. A certain section of drum-ensemble music might be characterized by a certain RTP relationship, maintained each time the metric cycle comes around; the change to a new RTP relationship characterizes the transition to a new section. This treatment amounts to a different way of conceptualizing meter, in which no reference need be made to any mathematical abstraction or its accompanying accent pattern.

The crucial orienting role of the bell pattern and the complex interplay among the interlocking cyclic patterns on like-sounding instruments demands a high degree of skill in both auditory scene analysis (Bregman 1990) and selective attending (Jones 1993). These cognitive facilities are rarely highlighted in speaking of Western tonal music, which tends to be less stratified and less rhythmically dense at the tactus level. However, the centrality of streaming and attending in music perception and cognition should be apparent both generally and in the specific case of drum music. I return to the notion of streaming in the next chapter, where I point out how certain expressive rhythmic manipulation can enhance segregation of similar streams.

## Meter Perception and Culture

Much has been written about African cultural concepts of time and how they might influence the approach to musical time. Ethnotheoretic discussions of how time is represented in the minds of African musicians (Merriam 1982, Stone 1985) have failed to appeal very extensively to African musicians themselves for information. Extramusical cultural investigations, usually in the form of literature searches, have given rise to rather extreme speculation about the "nature" of African time in the mind of "the African." While ostensibly aiming for cultural sensitivity, such studies have often applied oversimplified versions of selected cultural models rather abusively. A so-called "cyclic" mythology of life and death was seen to correspond to a beat cycle (Stone 1985) even though Western music also has beat cycles in the supposed absence of such a conception. Similarly, a cultural lack of ideal incremental measurement of time was seen as an argument against the existence of a steady underlying pulse (Merriam 1982) in a music that is quite unarguably pulse-based. These ethnomusicologists seemed determined to find essential differences between Western and non-Western ways of thinking, and to be able to display such differences in a pat description. In doing so they commit an epistemological error of presuming that such essential differences in worldview can be contained in one language but not in another. In other words, by assuming the role of impartial observer, they implicitly construct their own vantage point as a superior one for discerning such difference (Agawu 1992).

In this way they seemed to recapitulate the controversial stance of Whorf (1956), who asserted (despite a poverty of evidence) that the Hopi language contained such a radically incommensurate conceptualization of time from English that the Hopi themselves must experience a severely different reality from ours. (Whorf's well-meaning but flawed claims were later refuted in Malotki (1983)). The West African musicologist K. Agawu wrote, in direct response to such misinformed efforts in the study of African music, "[W]e have to be careful not to fall prey to simple dichotomies, especially those propped up by a vulgar essentialism." (Agawu 1992: 266) Elsewhere, he cautions:

> What must finally be resisted is the impulse to construct an Africa that is *always* different from the West. No doubt some differences cannot be suppressed, but one may be surprised by the extent to which the need for, and circumstances of, music-making in Africa resemble conditions in other parts of the world, and bespeak a basic human need for artistic expression. (Agawu 1995: 4-5)

And again, on the "project of 'exoticizing'" African music: "[S]uch relentless pursuit of difference must be resisted ... because it is likely to blind us to those areas in which difference actually occurs. For difference is perceivable only against a horizon of non-difference, and until we have constructed such a horizon, our efforts to understand African rhythm will continue to founder." (Agawu 1995: 193)

I do not mean to dispute concepts of mosaic time, nor to downplay the role of cyclical structures, nor to ignore the importance of learning the whole simultaneously with the parts, nor to dispel any other valid insights about African cultural models. Rather, in keeping with Agawu's admonitions, I only mean to be careful about how these ideas are described and employed, especially in regards to musical "time-reckoning." Just as we would not tolerate an assertion that all Western music is linear all the time from start to finish because all Western thought is linear and rational (nor could we really explain what such a claim is supposed to mean), so we cannot make such crude generalizations about the entirety of African music based on a vague and poorly explained cultural tendency.

While many rhetorical assertions are made that African musicians do not measure time or do not have a linear concept of time, nowhere are these claims *proven* in connection with music (nor is it very forcefully proven that all Western musicians *do*). A recent attempt to prove such a claim (Magill & Pressing 1997) is by its own admission (Magill & Pressing 1997: 189; Pressing, private communication, 1998) scientifically inconclusive, although rhetorically quite assertive in its conclusion. The study is predicated upon a rather misleading dichotomy between "an asymmetric timeline-ground model, which represents a computational elaboration of traditional African understanding, and a pulse-ground model, which is based on Western ideas of regular meter" (Magill & Pressing 1997: 189). In other words, it is supposed that the "Western" concept of steady pulse does not exist in "traditional African understanding," and that African music's irregular "timeline" rhythm takes the place of pulse. The quote from Ladzekpo (1995) in the previous section suggests that no such dichotomy exists for him. Curiously, Ladzekpo and his brothers are cited as a reference in Magill & Pressing, as is a humorously long list of other Western and African "experts" (Magill & Pressing 1997: 190, 197), provided as if to demonstrate enough agreement as to render any challenge moot. When I pointed out Ladzekpo's description of the pulse as the cultural equivalent to a sense of "purpose in life," Pressing (private communication, 1998) claimed that Africans have a different way of explaining it to westerners that really does not correspond to their inner thoughts. To this I must respond, why the conspiracy?

In their rigorous statistical analysis of the bimanual tapping of authentic West African rhythms by an authentic Ghanaian percussionist N. K. Owusu, Magill & Pressing (1997) seem to have thrown the experiment off by asking their subject to refrain from playing in his usual way:

> Mr. Owusu would often implement various flams in the course of a trial without conscious awareness. This apparently common stylistic device at first played havoc with the selection criteria, and as a result, it was necessary to remind him repeatedly at the beginning of each trial to avoid flamming... Mr. Owusu also showed a marked tendency toward acceleration over the course of a trial. This is not surprising as it is a common stylistic trait used in Asante drumming music; however, because it risks creating uninterpretable nonstationarity for long runs, he was also regularly reminded to maintain a steady tempo. (Magill & Pressing 1997: 204)

Pressing (private communication, 1998) reported that in fact Owusu would accelerate and decelerate over the course of a measure. I suspect that both the flams and the supposed tempo variation were actually a form of expressive timing against an isochronous pulse, as discussed in chapter 7. Furthermore I can imagine that such rhythmic expression was deeply intertwined in his performing practice, inseparable from the rhythms themselves, to the extent that forbidding him from playing this way created a highly irregular playing situation for him – irregular enough that the results are necessarily inconclusive.

When learning West African music, one is instructed constantly to focus on one's relationship to the bell. But this should not be taken so literally as to ignore the pulse; for one is also continually reminded to remain grounded, which in my estimation is closer to staying with the (most often duple) pulse. This sense of pulse is highly contextual. It is undeniable that a given rhythm can be "pulsed" in a variety of ways; West African musicians are trained in the abilities to shift their metric perspective on a rhythm and to hear and generate cross-rhythms. But it is not necessarily the case that advanced West African musicians hear all possible metric groupings simultaneously to such a degree that any semblance of meter dissolves. There is nearly always one prevailing meter, against which all other metric possibilities provide a sense of contrast. In any case, despite frequent efforts to portray African music as "participatory" (Keil and Feld 1994), there is no question that West African musicians are highly specialized, possessing knowledge and understanding that goes far beyond that of the average West African person (Agawu 1995: 116). Hence it is not clear that the extreme competence of one highly trained individual may generalize to an entire subcontinent. Surely, "inner" musical time remains a mysterious and unexplained dimension of music cognition in general. But that does not imply, as some would have it in African music (Magill et al. 1997; Pressing, private communication, 1998), that there is no commonly agreed upon pulse or time unit among members of a musical ensemble or its listening audience, nor even that there is a deemphasized reliance upon such units.

Occasionally, attempts at understanding of cultural differences can turn out instead quite pejorative. When the same Ghanaian musicologist offered his voice to the fray (Agawu 1986) and suggested a balanced view in which we look at the music itself for partial answers it may contain, he was all but berated for confusing the theoretical discussion with his "indigenous" perspective (Stone 1986). The implication was not simply that every musicologist is contaminated by his or her cultural perspective, but rather that certain contaminants are more suspect than others. What this reveals, and what my present work attempts to address, is the prevalent tendency to treat non-European perspectives on music as extraneous, foreign, or "Other." Unquestionably, one should resist the assumption that because musicians are of a different culture, their cognitive processes are mysterious. It is possible to allow for difference without fetishizing it. Rather, the definitions and governing concepts of musicology and music perception must either expand to apply to the music of the world, or at least have their limitations and problematics acknowledged.

With Western tonal music, usually we defer to the authority of the composer by heeding the time signature indicated in the score rather than imagining our own. We ought to do the equivalent for the music in question here, for which typically no composer is credited. The equivalent authority would be the practitioner, the carrier of knowledge, as we would deem any musician in the ensemble, from the master drummer to the bell player. We could simply ask such a person where the beat is, when the cycle begins, and so forth; or we could ask more indirectly how one would dance or clap to the rhythm. The meter as imagined by the practitioner is apprehended by the listener from the music chiefly due to his or her familiarity with the material – and specifically with its *production,* with its activity as dance-drumming. This does not necessarily imply *training* so much as it does *acculturation.*

When participating in the dance-drumming environment (which I have been fortunate to do in workshop contexts organized by Ladzekpo from 1994 to 1997), I usually cannot help but perceive the pulse, for it is quite readily reinforced by the sum total of the aural, visual, and physical/somatic input. And if my rhythmic footing is lost momentarily, I can always use the timeline, the bell pattern, to tell me which way is "up" – or "where's 1," as it might be worded in conventional Western terms. But this is only possible because I have learned how to parse the bell pattern and accompanying rhythms and body language from a carrier of this tradition; I have been (however minimally) acculturated or attuned to the details of this musical activity.

I have also found myself in similar contexts to which I was not at all acculturated, as in a recent trip to Senegal in collaboration with some North American and British jazz and funk musicians, four Afro-Cuban percussionists, and five Senegalese drummers. We Westerners had great difficulties guessing which way the African musicians' rhythms were intended to be heard, as did they with ours. (As one might expect, this musical language barrier was clearly less extreme between the Senegalese and the Afro-Cubans.) At one point some of us were taken to the Senegalese group's local performance at a crossroads in Dakar. Though some of us had been exposed to other forms of West African music, we had extreme difficulty understanding how to perceive their particular music and dance, which to us was highly ambiguous, metrically and otherwise. We found ourselves watching for the incidental dancing and jumping of small children, our only perceptually "salient" orienting guideline.

What the above anecdote suggests is that in these cultures (and in all cultures, I would argue), meter perception is a *practice* – an open-ended, intentional activity that is accomplished actively by the musical participants, while profoundly influenced by the perceivers' social context. (See Berger 1997 and Bourdieu 1977 for elaborations of this concept.) Meter perception is not simply a raw perceptual inevitability for which mathematical models of pulse salience and gestalt concepts of grouping will easily substitute. On the contrary, one requires crucial background knowledge to perceive the "correct" meter. Indeed, quite frequently an untrained listener from the West will perceive an "incorrect" meter in a piece of West African music by employing a standard minimum-syncopation principle. In the audio example **[CD-36]**, one might be tempted to hear a triple meter, but the high-pitched bell supplies the correct duple pulse.

Hence any model of rhythm perception and cognition must include stages at which incoming rhythms are compared to known rhythms, matched against known meters, and situated among broader expectations about musical events. It also must involve some degree of what may be called *active* perception, by which is meant the assessment of various alternative readings of the musical signal, and the switching among them, all carried out *in time* and continually revised and updated. In short, it must treat perception to some degree as a *practice*. Finally, it must replace the *communicative* model of musical discourse, i.e.

musician -> musical object -> listener

in which the musical signal simply carries information from the sound's creator to its perceiver, with a *mutually constructive* paradigm (Nattiez 1990),

musician -> musical object <- listener

in which both musician and listener are constructive, situated participants in a larger musical *fact* – for example, in an intercultural encounter in which different musical presuppositions hold on either side. And in fact we might consider expanding this paradigm to account for the musician's and listener's respective cultures, and to allow for the ways in which the so-called musical object might also "construct" the musician and listener.

### Summary

The point remains that a great many conflicting ideas have been bandied about regarding African rhythms, from Westerners in both ethnomusicology and cognitive psychology, and also from African musicologists. I would summarize my arguments as follows:

- We need to keep our definitions consistent in order to avoid the confusion that occur so readily with terms such as grouping, rhythm, pulse, and meter, and in particular to avoid confusion between denotative and connotative meanings. As a general construct, meter most often denotes the smallest perceivable periodically repeating groupings of the perceived tactus-level temporal units. Allowance must be made for irregular pulse constructs as they occur in Mediterranean, Middle Eastern, and South Asian musics.
- Meter *need not* denote a hierarchy of strong and weak beats; such a hierarchy is considered a secondary construct. However, the perceived meter is an abstract template that lies upon the musical material *in a certain place*; its particular relationship to the material is crucial. In other words, meter does include a *phase* relationship - that is, a concept of a regulative time point, which in the West we call the downbeat or the "one," but is conceived slightly differently elsewhere. So when we perceive meter, we mark one of the beats as the regulative point, where the metric cycle begins; this placement affects our perception of the rhythms that follow.
- Meter perception also includes among its ingredients the perception of the pulse and its operational subdivisions.
- We need to be aware of the nature-nurture questions entangled in terms like perception and cognition. Is meter perception a "natural" process, or does it involve some acculturation? Does meter just happen, or do we make it happen? If we make it happen, why do we know how to do so? Perception of meter is considered to be at least partially voluntary, in that a skilled musician or listener can hear a given rhythm in any number of meters. Meter perception is also strongly directed by learned schemas, and hence contains a culturally contingent component.
- We need to be clear about distinctions between music theory and music practice, and about how well this distinction transfers across cultures. Does a theory equal a cognitive representation? Can volumes of prose make one understand what it's like to play African music as an African person in an African community?

It should be noted that some meter-like concept of small or large groupings of a pulse exists in a large number of musics of the world, including many European, African, South Asian, Southeast Asian, and East Asian musics. Sometimes this meter is reinforced visually, as in the conducted orchestral music of Europe, in the dancing that is inseparable from much West African drumming, and in the hand signals produced by the musicians and audience in South Indian classical music. Other times it is encoded musically, as in timeline-based and *clave*-based musics of the West African peoples and their diaspora, in the two-and-four hi-hat or backbeat of African-American jazz and popular musics, and in the accentual reinforcement of tonal music. Generally, the cultural elaboration of meter involves multisensory signals that must be *decoded* according to a body of culturally determined guidelines. Such guidelines may conflict with perceptual input, especially if we accept the reality of the intercultural encounter, and even more so if we grant that perception itself is in part a culturally determined practice. The "experienced listener" constructed in Lerdahl and Jackendoff (1983, p. 3) is indeed adept at such idiomatic decoding. Given the variety of types of musical organization, it is unlikely that a single, perceptually based model for beat and meter cognition would apply universally. However, given the cross-cultural existence of *some* concept of meter (as opposed to none), the tendency to organize tactus-level pulses into periodically recurring groups should considered as an important, possibly universal component of pulse-based music.

These points are mere tips of vast conceptual icebergs in musicology, anthropology, and psychology. Above all, this chapter is intended simply to caution against the dangerous cultural slippage that occurs when hypothesizing about both the cognitive universals and the cultural particulars of music.

Like Agawu, I am calling for a restraint of the urge to revel in facile notions of difference.

Agawu's unique vantage point, that of a native West African raised on Northern Ewe music and culture and trained professionally in European music and musicology, allows him to make invaluable criticisms of the typical Western representations of Africa. His stance carries over quite effectively to the study of African-American cultures and musics. As a professional musician working in a number of African-American genres as well as a student of the cognitive science of music, I find myself in a similarly unique position, situated both to debunk some of the widespread problematics in the study of rhythm cognition and to suggest new alternatives. In my work I do not wish to represent African America as a frozen entity that allows easy description and that is free from interaction with mainstream European-American culture. Rather, I endeavor to show how the study of African and African-American musics can highlight previously unexamined aspects of rhythm cognition that are nonetheless globally relevant.

## 6. Microtiming Studies

I have experienced one of the most interesting musical revelations of my life, gradually over the last several years, in studying West African dance-drumming and in playing jazz, hip-hop and funk. The revelation was that the simplest repetitive musical patterns could be imbued with a universe of expression. I have often witnessed the Ghanaian percussionist and teacher C. K. Ladzekpo stopping the music to chide his students for playing their parts with no emotion. One might wonder how much emotion one can convey on a single drum whose pitch range, timbral range, and discrete rhythmic delineations are so narrow, when the only two elements at one's disposal are intensity and timing. Yet I have become convinced that a great deal can be conveyed with just those two elements. Some investigations into how this can happen are set forth in this chapter.

### Rhythmic Expression in African and African-American Musics

Some of the arguments in this chapter draw upon cultural aspects of music listening. Working from the documented historical lineage between West African and African-American cultures, Wilson (1974) has identified a constellation of conceptual tendencies that exist in the musics of that vast diversity of cultures. Among the musical preferences and principles he enumerated were the following:

- rhythmic contrast
- stratification
- antiphony (i.e. "call and response")
- connection between music and physical body movement
- percussivity
- continuity between speech and sound
- heterogeneous sound ideal
- tendency to fill up musical space
- concept of music as meaningful "in motion" – as part of everyday life

These and other concepts can serve as the beginnings of a pan-African musical aesthetic, since so many of these notions appear so often in so many different kinds of West African and African-American music. A great majority of this music falls in the category of groove-based music that I have mentioned, meaning that it features a steady, virtually isochronous pulse that is established collectively by an interlocking composite of rhythmic entities and is intended for or derived from dance. This somewhat inadequate description should not be viewed as a *definition* of the concept of groove; indeed, to some degree, that definition is what we are searching for with this work. One could say that, among other functions, groove *gives rise to the perception of a human, steady pulse* in a musical performance.

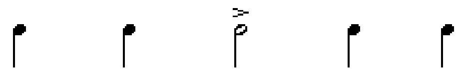
In groove-based music, this steady pulse is the chief structural element, and it may be articulated in a complex, indirect fashion. In groove contexts, musicians display a heightened, seemingly microscopic sensitivity to musical timing (on the order of a few milliseconds). They are able to evoke different kinds of rhythmic qualities, such as apparent accents or emotional mood, by playing notes slightly late or early relative to their theoretical metric location. While numerous studies have dissected the nuances of expressive *ritardandi* and other tempo-modulating rhythmic phenomena (Repp 1990, Todd 1989, Desain & Honing 1996), to our knowledge there have been few careful quantitative studies that focus on expressive timing with respect to an isochronous pulse. In groove-based contexts, even as the tempo remains constant, fine-scale rhythmic delivery becomes just as important a parameter as, say, tone, pitch, or loudness. All these musical quantities combine dynamically and holistically to form what some would call a musician's "feel." Individual players have their own feel, i.e. their own ways of relating to an isochronous pulse. Musical messages can be passed at this level. A musician can pop out of a polyphonic texture by a "deviation" from strict metricality, or a set of such deviations. As I shall attempt to demonstrate below, these kinds of performance variations create an attentional give-and-take to emphasize different moments interactively. This and similar techniques are manipulated with great skill by experienced musicians playing together, as a kind of communication at the "feel" level. We claim that this variety of expressive timing against an isochronous pulse contains important information about the inner structure of groove.

Often when the topic of musical communication is breached, one is tempted to wonder what is being *said* amidst all this communication. This raises the question of what actually constitutes a musical message, or, for that matter, musical meaning in general. Here, I feel, one should draw upon the processual notion of communication, as a collective activity that harmonizes individuals, rather than on the telegraphic model of communication as mere conveyance of literal, verbal meanings. For example, the musical notion of antiphony, or call and response, can function as a kind of communication, and nothing need be "said" at the literal level to make it so (although we need not rule out the possibility of musically encoded symbolic meaning). What definitely *is* happening is that the interactive format, process, and feeling of conversational engagement are enacted by the musicians. In a context like jazz, the presence of this kind of dialogic process can be constant throughout a performance, as *sustained antiphony.* I am arguing that a significant component of such a process occurs along a musical dimension that is non-notatable in Western terms – namely, what I have been calling microtiming.

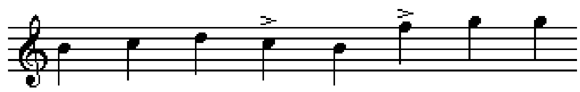
**Previous Microtiming Studies**

Miniscule timing deviations from metronomicity are frequently miscast as "discrepancies" (Keil 1995), "motor noise," or "inaccuracies" (Rasch 1988). But there has been a small thread of research dedicated to the uncovering of structure in these so-called inaccuracies. It turns out that these deviations both convey information about musical structure and provide a window onto internal cognitive representations of music. One of the most compelling examples of this direction in research is provided by Drake & Palmer (1993). They proposed three types of accent structure:

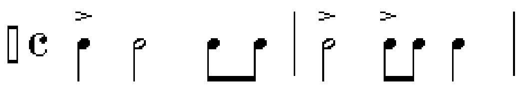
- *rhythmic grouping* accents, in which a longer note tends to end a rhythmic group and is perceived as accented (a.k.a. the "gap" principle), as indicated by the accent mark:



- *melodic contour* accents (turns or leaps in pitch contour), as in:



- *metric* accents (so-called strong beats), such as:



In studies of timing of numerous skilled classical pianists, they found systematic deviations from strict regularity that correlated with these accent structures. Their qualitative findings are summarized below (table adapted from Drake & Palmer 1993).

In these results, it is clear that small performance variations in timing, intensity, and duration enhance aspects of musical structure. Drake and Palmer concluded that these performance variations facilitate listeners' segmentation of musical sequences, since the accent structures serve to break up a musical sequence into smaller, more tractable chunks. Furthermore, Drake and Palmer found that the expressive effects that stemmed from rhythmic grouping tend to dominate melodic-accent or metric-accent effects; the

| *Type of accent structure* | **Intensity variation** | **Inter-onset interval** | **Overlap/ Articulation** |
|---|---|---|---|
| **Rhythmic grouping accent** | Last event is louder<br><br>Notes crescendo throughout group | Penultimate event is elongated (i.e., last event is delayed) | Penultimate event is more staccato<br><br>Articulations are proportional to note durations |
| **Melodic accent** | Event on a turn is louder | Event before or on a melodic turn or leap is elongated (i.e., melodically accented event is delayed) | Event before or on melodic leap is more staccato |
| **Metric accent** | Stronger beats are louder | Last beat of metric cycle is elongated (i.e., downbeat is delayed) | Stronger beats are less staccato |

former tended to override the latter two when the music yielded conflicting accent interpretations. That the metric accents would tend to be overridden is unsurprising, since one would expect expressive timing to break up the regularity of repetitive metric accents. But the fact that rhythmic grouping effects dominate melodic-accent effects suggests a more general primacy of rhythm over melody, in both production and perception. Also worth noting is that these expressive variations allow *graduated change* – flexible, continuous variability. A moderate amount of expression is the norm, and performances that are low or high in expressivity will stand out as extreme.

In his studies of microtiming variation in small chamber groups, Rasch (1988) conducted statistical analyses of inter-musician differences in note onsets from recorded ensemble performances. Although he averaged out all musical structure, including metric accents and tempo variation, he found that generally, in a string trio, the violin's lead voice tends to lead by 5 to 10 milliseconds, the cello tends to follow, and the viola's middle voice tends to lag by another 5 to 10 milliseconds. It is unclear how accurate these findings are, however, because the standard deviation for each of the instruments was around 35 milliseconds. But at any rate, Rasch's hypothesis that there might be systematic variation in ensemble performance is a valuable one.

The above studies focused on European classical-music performance, which would not fall into the realm of groove-based music because of its reliance on tempo variation for expressive purposes. Indeed, the above results indicate that beats are frequently lengthened or shortened by the performer. Also, as discussed in the previous chapter, the treatment of metric organization as implying a series of weak and strong beats does not apply particularly well to West African or African-American musics; in these contexts there is no such thing as a metric accent, in terms of performance variation. Although the above studies are valuable, their stylistic scope does not coincide with ours.

It would be instructive to conduct a similar microtiming analysis for a percussion ensemble, particularly in instances of groove-based music which is much less forgiving in the realm of tempo variation and *rubati* than a string trio might be. Bilmes (1993) conducted a timing analysis of a recorded performance of Los Muñequitos de Matanzas, an Afro-Cuban *rumba* group **[CD-37]**. In a performance averaging 110 beats per minute (such that what would be a notated sixteenth note lasts around 135 milliseconds), both the *quinto* and the *segundo* (lead and middle conga drum, respectively) tend to play about 30 milliseconds ahead, or "on top." On the other hand, the *tumbao* (low conga drum) had a much broader distribution, nearly as often late as early. It should be noted that here the precise moment of the beat was *not* determined by the norm set by these three instruments themselves, as it was in the case of the string trio. Rather, the beat was established by a reference instrument, in this case a *clave* or a *guagua*. Hence it was possible for all three instruments to be ahead of the nominal beat, which was not the case for the string trio. In Bilmes's work, the *average* inter-drum asynchrony was not calculated; indeed, such a measure would ignore any relationship between timing and musical structure. But a frequency analysis of the microtiming variations revealed systematic structure. For example, the repetitive *segundo* part displayed a strong peak corresponding to its frequency of repetition, showing that the microtiming variations were not at all random.

Given this apparent systematicity of fine-scale rhythmic expression in groove contexts, we can take cues from the results of Drake & Palmer (1993) and Rasch (1988) discussed above, as well as from our expanded view of cognition involving the theory of embodiment, to make guesses about the function of such rhythmic expression. Thus we hypothesize that microtiming variations in groove music play any of the following roles:

- highlight structural aspects of the musical material,
- reflect specific temporal constraints imposed by physical embodiment, and/or
- fulfill some aesthetic or communicative function.

I shall now address all of these possibilities via a few examples.

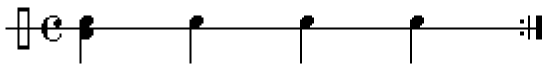
## Examples of Expressive Microtiming

### Asynchrony

The asynchronous unison attacks described above support a scientifically meaningful explanation. Rasch (1988) reports an earlier study of asynchrony: a 1977 experiment in which he investigated the effect of onset difference times on the perception of quasi-simultaneous tones. The threshold for perceiving the upper of two quasi-synchronous tones could be decreased drastically [from between zero and -20 dB to about -60 dB] by introducing an onset difference time of, say, 30 ms. In the latter case, the threshold was largely independent of factors other than temporal ones... In supra-threshold conditions asynchronization contributes to the apparent transparency of the compound sound multi-tone stimuli. (Rasch 1988: 80)

So a function of these quasi-synchronous attacks, or flams, could be to aid in the perception of the timbral constituents of the unison attack.

The accompanying audio example demonstrates this tactic **[CD-38, 39]**.

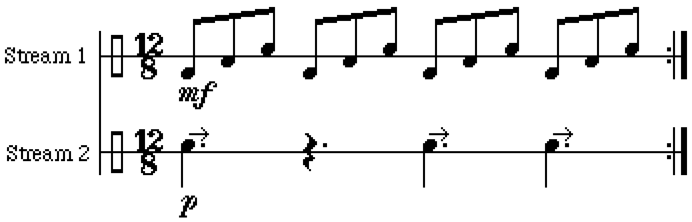


In the first version, the double stroke on the first beat is essentially synchronous; in the second version, a delay of 30 milliseconds is introduced between the two notes, resulting in a small flam. This serves to enhance the perception of the two timbral constituents.

In some musical situations in which blending is preferred, this kind of multitimbral asynchrony may be undesirable, but it is often a valued musical trait in groove-based music. On Wilson's list of African and African-American aesthetic concepts is the notion of a heterogeneous sound ideal, a tendency to value the presence of a variety of contrasting timbres. Another important cultural characteristic (not mentioned earlier) is a *collectivist* ideal, in which music is construed as a communal activity among groups of people. The rhythmic asynchronies described above aid in the perception of a multiplicity of timbres, as well as, in the ecological view of music perception (Gibson 1979, Shove & Repp 1995), the multiplicity of human bodies behind those timbres. That is, rhythmic asynchrony contributes both to the heterogeneous sound ideal and to the sense of collective participation. To be sure, exact synchrony is impossible with groups of people anyway; but this principle applies even when different sounds are played by one individual, as on the modern drumset. So here is an instance in which a kind of subtle rhythmic expression fulfills both a perceptual and an aesthetic function.

**Streaming**

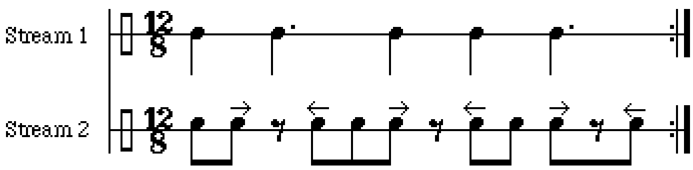
It is well established that auditory stream segregation is a function of both pitch and timbre (Bregman 1990). From our work, it appears that microtiming can also contribute to streaming. This claim builds on the role of asynchrony in facilitating the perception of multiple tones. The audio examples **[CD-40, 41]** consist of a steady stream of triplets on tom-toms, together with a series of sparse tom-tom strokes at a lower volume. The musical material is shown below.




In the first audio example, the unison strokes are as simultaneous as is allowed by the MIDI protocol (i.e., within a couple of milliseconds of each other). In the second example, the second stream is delayed by 30 milliseconds with respect to the first, as indicated by the arrows in the figure, but kept at its same low volume. In the former case, the different timbres fuse into one stream, whereas in the latter case, the second stream is clearly audible as a separate entity. This example shows clearly how such miniscule timing variations can contribute to streaming. This technique is especially important in a context where the aesthetic tendency is to "fill up the musical space" (Wilson 1974). Timing variations can allow an instrument that is sonically buried to draw attention to itself in the auditory scene. Thus the presence of multiple instruments of similar timbres, as in a West African drum ensemble or a large jazz ensemble, need not be viewed as enforcing the subordination of individual identity. Individual musicians can improvise at this microrhythmic level to create an attentional give-and-take. This streaming effect also serves an aesthetic function, in that it enhances the perception of different rhythmic groups as separate animate entities with distinct "personalities" as Ladzekpo stresses (1995).

**Spreading**

It was not until the advent of automated machinery that human ears were ever treated to inhuman rhythmic precision. The fact is that sonic trace of temporal constraints imposed by the body are often perceived as aesthetically pleasing, while inhuman rhythmic regularity often is not. These audio examples **[CD-42, 43]** consist of two versions of the "same" rhythm, shown below.




The first rendition is executed as close to the theoretical ideals the computer allows - that is, rigidly, mathematically accurate. The second features timing inflections designed to imitate an aspect of human performance. The difference is not simply the injection of random temporal slop. Rather, it involves the spreading apart of consecutive attacks played by the same hypothetical limb or digit. An individual effector such as a limb, hand, or digit has a time constant associated with its motion; the nerves and muscles have a brief

The rhythmic expression added in the above example is systematic; the first of each group of three taps is about 30 milliseconds early, and the last is about 30 milliseconds late, as indicated by the arrows in the figure. In addition to enhancing perceived separation, this example depicts the encoding of bodily movement in musical material. Nearly all listeners are familiar with the kind of motion suggested by these synthetic tapped rhythms, but that motion is strongly implied only by the second, "imperfect" version. Again, this description recalls the embodied, ecological view of musical perception, in which the listener perceives the *source* of the sounds, rather than the sounds themselves. In a music that embraces physical body motion (Wilson 1974) and that is contiguous with everyday experience, this sonic trace of the body is a valued aesthetic.

## Coding for Invariance

The above three examples demonstrate the notion of invariance. At the most basic level, expressive microtiming represents a departure from regularity, so it is likely to be noticed in relief against the more regular background. Gibson (1975) claimed that our perceptual systems are attuned to variants and invariants in the environment; they code for change. As an example, consider the way that vibrato or a trill can facilitate auditory scene analysis by drawing our attention to a particular instrument in an otherwise blended orchestral texture **[CD-44]**. The microvariation of a single pitch is enough to make that voice pop out in the auditory scene.

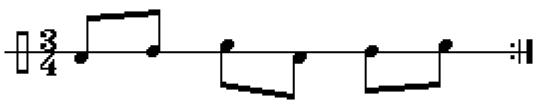
We can make a similar generalization with rhythm. That which is regular, or invariant, in an isochronous-pulse context is the norm set by the regularity of pulsation, along with its salient multiples and subdivisions; that which is irregular comprises the variable rhythmic material along with its continuous expressive variation. Microrhythmic expression signals a departure from the implied norm, hence marking a particular sound or group of sounds as worthy of attention or analysis by our perceptual systems. This argument contributes to an ecological view of rhythm perception, in which we are attuned to variations in an otherwise regular environment.

## Swing

A kind of rhythmic expression that seems to be indigenous to African-American culture is that found in jazz of the first half of this century. Known as swing, this kind of structure can be thought of as modified duple subdivisions of the main pulse, or as modified triplet subdivisions, or both concurrently. As duple subdivisions, they divide the interval of a pulse into two unequal portions, of which the first is slightly longer. They are occasionally rendered in triplet notation as a quarter note followed by an eighth note, but this exaggerates the typical swing ratio, which is usually in the gray area between duple and triple and is strongly tempo-dependent (typically lower for fast tempi and higher for slow ones). An individual musician has a particular range of preferred ratios and particular ways of manipulating them, which together form crucial dimensions of that individual's sound, rhythmic feel, and musical personality.

In a related experiment on rhythm, Fraisse (1982) has studied the ability of musically trained and untrained subjects to reproduce rhythmic patterns of varying degrees of complexity. "Arrhythmic" sequences with arbitrary relationships between time intervals caused the most difficulty. In more regular rhythmic cases, subjects tended to simplify the ratios between intervals, almost always settling on exactly two classes of time interval: long (400-800 milliseconds) and short (200-400 ms). People tend to understand rhythms to feature two and only two interval lengths, roughly in the ratio of 2:1. This drive towards rhythmic simplicity recalls some of the classical perceptual laws, namely the principle of economy in organization (Fraisse 1982). Usually as performed or as "preferred," the ratio is lower – in fact closer to swing, about 1.75:1, about 57%

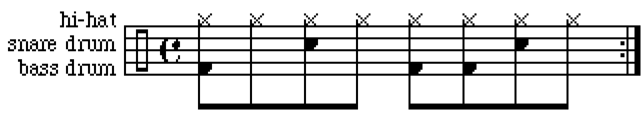
However, it is not apparent why the interval would be divided unequally in the first place. It would seem even simpler and more economical if there were no such difference in duration between the first and second of two consecutive swung notes. But the point is that this difference facilitates the perception of higher-level rhythmic structure. An immediate consequence of the swing feel is that it suggests the next level of hierarchical organization. In conventional terms, the swung eighth-note pairs are perceptually grouped into the larger regular interval, that is, the quarter note. If all subdivisions were performed with exactly the same duration, it would be more difficult to perceive the main beat. The lengthening of the first of two swung notes in a pair amounts to a durational accentuation of the beat. (Often in practice, the second note of the swung pair is given a slight accent in intensity, as if to compensate for its shorter duration.) Hence swing enhances the perception of the main pulse, as the examples **[CD-45, 46]** demonstrate:



The first version plays all eighth notes exactly equivalently and is therefore metrically indistinct, whereas the second version introduces a slight swing, which immediately marks the pulse.

**In the Pocket: Backbeat Delay**

The notion of a backbeat is indigenous to the modern drum kit, an instrument pioneered by African-Americans in this century. It consists of a strongly accented snare drum stroke or handclap on beats two and four of a four-beat metric cycle, where the beat is typically a moderate tactus rate **[CD-47]**.




The backbeat appears to have arisen in the middle of this century, as the popular swing rhythm yielded to the even more popular, more bombastic rock and roll rhythms of artists such as Little Richard and Chuck Berry.

In his musical interpretation of Stuckey's (1987) study of the culture of enslaved Africans and its influence on modern African-American culture, Floyd (1995) discusses the the important African diasporic cultural ritual known as the *ring shout* as a distinctive space in which, among other things, music and dance were fused. This activity "helped preserve ... what we have come to know as the characterizing and foundational elements of African-American music," including "constant repetition of rhythmic and melodic figures and phrases," "hand clapping, foot patting, and approximations thereof," and "the metronomic pulse that underlies all music." (Floyd 1995: 6) As a cultural model, the ring shout serves for Stuckey as a hermeneutical point of departure in the study of African-American art forms. It provides an alternative lens through which to view these later practices, a lens that is grounded on African, rather than European, concepts and aesthetics. (See Rosenbaum 1998 for more documenttation of the ring shout.)

The backbeat that is so prevalent in postwar African-American popular music seems to reference the role of the body in the ring shout – the bass drum (struck with a foot pedal in the modern drumset) and snare drum (struck manually with a stick) replacing the stomp and clap, respectively. In fact, a real or synthetic handclap sound is often superimposed on the backbeat's snare drum sound in popular urban dance music. The hard-edged repetitiveness of the backbeat embodies the cyclic, earthy atmosphere of the ring-shout ritual. Though sometimes dismissed as dull and monotonous, the backbeat taps into the hypnotic, functional role of repetition in such rituals, in which steady, moderate tempo, rhythmic ostinati, and physical body motion (stomping and clapping) were combined in a collective setting to create a shared multisensory experience. It seems plausible that the earliest musical activities of humankind possessed many of these qualities. The backbeat is best understood as a contemporary, popular remnant of what is probably some very ancient human musical behavior, filtered through a sophisticated, stylized African ritual and through centuries of African-American musical development.

The curious point about the backbeat in practice is that when performed, it displays a microscopic lopsidedness. If we consider the downbeat to be exactly when the bass drum is struck, then the snare drum is very often played ever so slightly *later* than the midpoint between two consecutive pulses **[CD-48]**. Often musicians are aware of this to some degree, and they have a term for it: the drummer is said to play "in the pocket." While perhaps unaware of the exact temporal details of this effect, a skilled musician or listener in this genre hears this kind of expressive microdelay as "relaxed" or "laid back" as opposed to "stiff" or "on top." This effect is much subtler than the salient rhythmic categorization of the long and short durations of swing. It is a miniscule adjustment at the level of the tactus, rather than the substantial fractional shift of rhythmic subdivisions in swing.

What function does this delay structure have? Perhaps the delay functions as a kind of accent, since it involves the postponement of an expected consequent (Meyer 1956). It seems plausible that the optimum snare-drum offset that we call the "pocket" is that precise rhythmic position that maximizes the accentual effect of a delay without upsetting the ongoing sense of pulse. This involves the balance of two opposing forces: the force of regularity that resists delay, and the backbeat accentuation that demands delay.

Note that the concept of a backbeat, and the slight delay associated with it, does not pertain if a single voice is used for both the downbeat and the backbeat. (As an example, the urban dance-music genre known as "house" features an isochronous bass drum on all four beats, with the snare-drum backbeat occasionally dropping out.) The effect seems tied to the difference between the two sounds, and perhaps also to the actual sounds themselves and the imagined bodily activity that gives rise to it. In a related study, Fraisse (1982) reports,

> In speaking of synchronization, it is necessary to specify what is synchronized with what. In effect, if one measures the temporal separation between a tap of the forefinger and the sound, one finds that the tap slightly anticipates the sound by about 30 msec. The subject does not perceive this error systematically... Moreover, this error is greater if the sound is synchronized with the foot. The difference between hand and foot permits us to think that the subject's criterion for synchronization is the coincidence of the auditory and of the tactile-kinesthetic information at the cortical level. For this coincidence to be as precise as possible, the movement of tapping should slightly precede the sound in order to make allowance for the length of the transmission of peripheral information. This length is all the greater when the distance is longer. (Fraisse 1982)

This delay architecture amounts to the subject's hand coming *after* the foot for perceived synchronization, since the anticipatory "error" is greater for the foot. This seems to predict that a regularly alternating stomp-clap pattern would contain a microscopic asymmetry similar to that found in the modern backbeat. Given that the bass drum both references and is played by the foot, and similarly the snare drum both points to and involves the hand, it is possible that this resultant delay structure was transferred to the drumset. Though these arguments are quite speculative, it is plausible that there is an important relationship between the backbeat and the body, informed by the African-American cultural model of the ring shout.

## Rhythmic Expression: Two Musical Examples

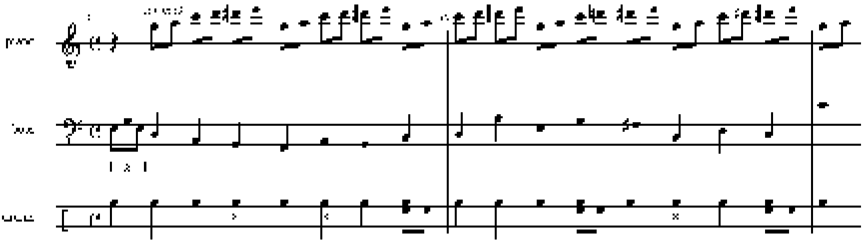
**Thelonious Monk plays "I'm Confessin'." [CD-49]** One of the most fascinating skills displayed by Monk and many other pianists of the genre is a high degree of independence between the two hands, to the degree that one hand can appear to perform rhythms that are ambiguously if at all related to those performed by the other. Often, as in stride piano, this takes the form of a steady pulse or repetitive bass rhythm in the left hand (the "ground"), and upper-register, rhythmically free melodies in the right hand (the "figure"). A classic example is Monk's 1963 solo recording of "I'm Confessin' (That I Love You)" (Monk 1998). In this piece, after carrying on in this expressive stride fashion for some time, the last two bars of the first chorus give rise to an improvised melodic fragment that rhythmically seems to stretch and tumble into the next bar **[CD-50]**.

In this excerpt, the melodic structure in the right hand temporarily overrides and upsets the underlying rhythmic structure, only to be righted again. We can interpret Monk's unquestionably gripping display here as the rhythmic equivalent of a struggle, one that threatens the norm of established pulse regularity set by what has come before. It seems to offer an example of a case in which such regularity is sacrificed briefly to allow for a case of extreme rhythmic expression. But note that the sense of pulse is never lost; Monk leaves out a couple of quarter-note chords in the left hand, but otherwise provides strong and accurate pulse reinforcement in the stride style. The rhythmic underpinning of the left hand compensates for the apparent deviation from regularity.

When I played the recording of this piece for a roomful of cognitive science undergraduates, most of whom presumably had no familiarity with jazz, this excerpt elicited a burst of spontaneous laughter. Something about Monk's delivery is communicative enough to transcend what one might expect of the traditional confines of genre. Nearly upsetting the regular pulse, Monk takes a chance and chooses to follow through on a melodic idea that momentarily takes him rhythmically far afield.

The question of whether Monk "intended" to play this in exactly this way is a pejorative one, akin to rectifying the role of "mistakes" in jazz (as in Walser 1995). From the perspective of an improviser, the notion of a mistake is supplanted by the concept of displaying one's interaction with the structure suggested by the sonic environment. It is never clear what is "supposed" to happen in improvised music, so it makes no sense to talk about mistakes. This improvisation-friendly framework allows for the possibility of musical exploration and experimentation, including impromptu rhythmic variation of the sort described here, without invoking a notion of mistakes.

**Ahmad Jamal plays "But Not for Me." [CD-51]** A wonderfully extemporaneous, playful spirit is captured masterfully in pianist Ahmad Jamal's 1952 trio version of the standard tune "But Not for Me." In this piece, Jamal manipulates his relationship to the pulse actively and voluntarily through the skillful use of microtiming variation. Nearly every single phrase in Jamal's rendition contains some interesting microrhythmic manipulations, but here I will focus on one fragment, namely the end of the first chorus into the beginning of the second chorus. In measure 31, Jamal initiates a repeating three-beat figure in the four-beat metric context. This additive rhythmic technique is a common one in African-American music, and Jamal carries it out to a humorous extreme, letting the blues-inflected figure cycle twelve full times (nine measures). The first four measures of this passage are displayed below. I have adhered to the convention of representing swung rhythms with regular eighth notes, but it should be understood that there is much more to this passage than meets the eye. In particular, Jamal plays this figure extremely behind the beat, so much so as to enhance the humorous effect of the repeating melodic figure by casting it in starker relief against the more ordinary rhythmic background **[CD-52]**.



In these four measures, the quarter note averages 469 milliseconds (128 beats per minute). The note events in the piano that are displayed as occurring *on* the beat tend to begin actually around 40% of a beat *later* than the drummer's rimshots, which are indicated with x's above. This places him more than a triplet behind the beat. Furthermore, Jamal's second eighth note in each swung pair tends to occur about 85% of the way through the beat. This means that the swing ratio here is effectively inverted; the *first* eighth note in a delayed pair lasts about 45% of a beat (less than half), and the *second* lasts about 55% (more than half). It would appear that the perception of swing arises due to complex variations in timing, intensity, or articulation; in this case, it is not merely a matter of achieving the "correct" microrhythmic ratio.

How does Jamal pull off this apparent rhythmic violation of an inverted swing? The answer seems to lie in his 40% *phase shift* relative to the beat established by the accompanying instruments. If, while maintaining this phase relationship, he were to adhere to the usual swing ratio of around 57%, then the second note in a swung pair would be close enough to the onset of the next beat (only a few percent early) that it would be heard as on-the-beat. By employing a relative anticipation of the second eighth note in each pair, Jamal avoids this problem, instead sounding squarely "between" the beats. The 40% delay also affords him enough rhythmic ambiguity so that the inverted swing does not sound jarring. Also, Jamal enhances the sense of swing by accenting the second of each pair (a common technique, as mentioned earlier). So here is a case in which one kind of rhythmic expression interacts with another; the usual long-short relationship of swing is altered in order to accommodate the "laid-back" quality of the melodic figure.

What is accomplished by playing in this laid-back, behind-the-beat fashion? One might expect the same simple perceptual effects (such as enhancing stream segregation) if he instead played *ahead* of the beat, for example. Playing behind the beat is definitely a cultural aesthetic in African-American music, especially jazz. In an unpublished study, Bilmes (1996) found that a West African drummer played equally as often ahead of as behind the beat, whereas one might observe casually that skilled jazz improvisers tend to play much more often behind than ahead. From the ecological point of view, playing behind the beat might be normally associated with a physical or mental state of relaxation, or might suggest a causal relationship in which the musical material is a *reaction* to the pulse. Such hypotheses would demand further investigation.

In this chapter I have discussed some aspects of rhythmic expression that are quite distinct from the common body of European classical musical performance techniques typically discussed. Instead of (or in addition to) expressive concepts like *rubato*, *ritardando*, and *accelerando*, we have seen deliberately asynchronous unisons, subtle separation of rapid consecutive notes, asymmetric subdivisions of a pulse, and microscopic delays. As further illustration, we have seen extremely deft manipulation of fine-scale rhythmic material in examples from the jazz idiom. I have chosen to focus on African and African-American musics because they often feature these concepts in isolation from the possible interference of tempo variation, and because they tend to involve percussive timbres which facilitate precise micro-rhythmic analysis. I have argued that African and African-American musics incorporate aesthetics that value these kinds of micro-rhythmic expression. However, I believe that these techniques are found to varying degrees in all music, including the European classical genres. In the next chapter, I present a representation for rhythmic structure that allows for the explicit manipulation of expressive microtiming of the variety discussed above.

# 7. Describing Rhythmic Behavior, Representing Rhythmic Structure

In this chapter I discuss a class of models for rhythmic tapping that have developed in the literature over the last few decades. I discuss what I see as a shortcoming of these models, from which I suggest a supplementary view that allows for different mental processes at different timescales. In this light, I then describe work initiated by Bilmes (1993) and continued by a group consisting of Jeff Bilmes, Matt Wright, David Wessel, and myself (Iyer et al. 1997). This work culminated in the design and implementation of a novel representation for rhythmic structure, incorporating considerations specific to groove-based music such as those discussed in the previous chapter.

### Rhythmic Tapping

The well-known studies of Wing & Kristofferson (1973), Verborg & Hambuch (1984), Jagacinski et al. (1988), and Verborg & Wing (1996) are concerned with the unraveling of cognitive command structure in the timing of motor activity. Wing & Kristofferson (1973) conducted experiments in which subjects tapped a finger periodically at a moderate, steady rate. Subjects tapping was initially matched with periodic tones, and then the tones ceased and the tapping was unpaced. The experimenters studied how accurately the original period was upheld in unpaced tapping. From their data they developed a model in which it was assumed that there were two independent sources of variability in this unpaced phase: 1) variation in timing of centrally-generated, feed-forward, periodic commands, and 2) variation in implementation (mechanical noise in effector, nerve response delays, and so forth). The variability in tapping was taken to be the sum of these two random variables. By doing elementary statistics on the interval values, one could determine which of these two variables gave rise to a given change in the overall variability. For example, it was found that the typical negative correlation between successive intervals ("negative covariance at lag-1") is associated entirely with variance in implementation. In other words, the natural "swing" inherent in a person's tapped steady pulse is due to variability not in the centrally-generated command originating from the brain, but in the motor processes associated with the body.

Pressing et al. (1996) worked with experienced musicians to find that systematic variations in so-called polyrhythmic tapping, measurable through variance-covariance analysis of tapped intervals, imply a hierarchical cognitive structure in rhythm production. Specifically, there is a difference in the microtiming inflections of a musician tapping 4:3 (four equally spaced taps over the length of three tapped beats) from one tapping 3:4, even though theoretically they are the "same" rhythm. This systematic variation fits with a hierarchical model involving a central clock, a separate process of referential timing, and motor delays.

The general cognitive model proposed by the above authors for the production of polyrhythms is as follows:

1. A central clock-like command process directly triggers the motor delay process of the "ground" rhythmic stream, played by one hand.
2. Via an intervening subpulse (subdivision) process, the immediately preceding ground-stream clock element cues the motor delay in the "figure" stream.

It is possible that the results showing the difference between 3:4 and 4:3 are a recasting of the work of Drake and Palmer (1993, discussed in the previous chapter) at a lower cognitive level. If there is indeed some amount of microtiming that is related to meter, then it could show up in the difference between the so-called figure and ground stream. Pressing et al. make no reference to meter, but it is clear that the ground stream functions metrically in this context. The central command clock functions as a pulse; the inter-pulse intervals are subdivided for accurate timing of the figure stream. However, Magill and Pressing (1997) have attempted to posit an asymmetric central clock, based on a view of West African music as treating its so-called timeline patterns as the "ground" rather than a figure. This work was discussed and critiqued earlier, in the chapter on meter.

**A Subtactus Clock**

A problem I see with all of the references cited above is that nobody has ever proposed a model for polyrhythmic tapping in which the central clock pulse is the common subdivision of the generated rhythms. Instead, they propose internal timekeepers that are equal to the length of one of the tactus-level pulses. In the above model, subdivision of the ground pulse only occurs when necessary; it does not continue in the absence of subdividing material. This is odd, considering that the oft-cited Povel (1981) has shown that people tend to perceive rhythms in terms of a common underlying unit, and have the least difficulty perceiving rhythms that conform to such a structure.

To be sure, the notion of a second, faster cognitive clock seems to go against some theories of rhythm perception. According to Brower, we process their structure of rapid (sub-tactus) rhythms in a more qualitative fashion. Instead of measuring individual durations against the background of an internally generated metric grid, the listener recognizes sub-pulse rhythms by their qualities of "evenness or unevenness, twoness or threeness, accentedness or unaccentedness, and so on." (Brower 1993: 25) She cites Preusser (1972), who found experimental evidence for a difference between integrated, immediate, passive processing of rapid rhythmic gestalts, and intellectualized, cognitive, active processing of slower rhythms. She says that performance variations (e.g., 2:1 becomes 1.75:1, as described in Fraisse 1982) are evidence that the fast clock is not the most accurate for music. Brower's claims about the unreliability of fast clocks appear to stem from generalizations about unskilled rhythmic behavior.

However, my own experience as a performer has suggested that what we call a groove is easier to reach if the members of the ensemble collectively focus on the ongoing small subdivisions of the beat, whether or not they are filled by musical events. Similarly, C.K. Ladzekpo teaches us to feel the abstract, ongoing rhythmic subdivisions as a constant reservoir of rhythmic intensity, a rapid, energetic, incessant movement, which he calls the "yell." (Ladzekpo 1995) The awareness of this continual activity helps the performer animate a simple rhythmic pattern. It seems quite similar to the "imagined movement" that we have discussed in the case of pulse perception, but now at the faster rate of the pulse subdivision. By maintaining a sense of these abstract subpulses, and indeed by imagining them not silent but *loud*, the performer not only enhances rhythmic precision but also derives cues for appropriate durations and intensities for individual notes.

Contrary to Brower's claims, Ladzekpo's teachings suggest that one can learn to internalize this fast subpulse such that one need not rely on its physical reinforcement. This relates to a model proposed by Ivry (1998) which proposes a bank of timers in our neural apparatus, each related to the time constants of our various limbs, digits, and other effectors as they perform various tasks. If we are able to generalize from locomotor-type activities to an abstract concept of musical pulse, it is equally possible that, *with practice*, we can learn to internalize a faster clock. The latter would be related to the temporal structure of digital, manual, and lingual motion, which, as I mentioned in chapter 3, occur at a timescale that is substantially faster than the tactus/locomotor rate.

If all this is granted, it seems as though one could posit two simultaneous central clocks, one of whose frequencies is a multiple of the other. Again, *with practice*, an individual would learn to yoke together the faster clock associated with the physical activity of music-making, such as rapid finger motion, with a slower tactus-level pulse. I emphasize the notion of practicing this behavior because, as demonstrated in the previous chapter, groove-based musical activity involves highly skilled and precise temporal acuity, far from the simplicity of a typical tapping experiment. Also, significantly, groove-based musical activity is quite corporeal in nature; it is not just an abstract form of knowledge, but also a concrete skill requiring physical dexterity.

### A Tripartite Model

Bilmes (1993) has developed a tripartite model for expressive timing in performance of groove-based music. In addition to the salient moderate-tempo pulse or *tactus*, another important pulse cycle is defined at the finest temporal resolution relevant to a given piece of music. It is called the temporal atom or *tatum* (in homage to the great African-American improvising pianist, Art Tatum), the smallest cognitively meaningful subdivision of the main beat. Multiple tatum rates may be active simultaneously, particularly in ensemble performance. In Western notation, tatums may correspond typically to sixteenth-notes or triplets, though they may vary over the course of a performance. As noted above, groove-based music is characterized in part by focused attentiveness to events at this fine level. The tactus and the tatum provide at least two distinct clocks for rhythmic synchronization and communication among musicians.

In Bilmes's scheme, a performance displays musical phenomena that may be represented on three timescales. First, the musical referent or "score" corresponds to the most basic representation of the performed music as it would be notated in Western terms, using quantized rhythmic values (tatums) that subdivide the main pulse. All note-events are represented at this level. Secondly, at relatively large timescale, inter-onset intervals are stretched and compressed through tempo variation. This variation may be represented as a tempo curve – a function of musical time vs. score time. However, particularly in percussive music, there is no real musical continuum separate from the note-events; score time is quantized in units of tatums. In fact, the tempo curve operates on tatums, modifying their durations such that their sequential sum corresponds to the integral of the tempo curve. In this way, the tatum may be regarded as a sampling rate.

Thirdly, the tatum-relative temporal *deviations* capture many of the expressive microtiming variations discussed in the previous chapter. Deviations quantify the microscopic delays or anticipations of note-events relative to the theoretical tatum onsets. In other words, they represent the microscopic values by which note onsets differ from rigid quantization, *over a metronomic background*. Deviations take on continuous values from -0.5 to +0.5 tatums, so that all possible rhythmic placements are allowed. In the case of multiple simultaneous tatum rates, this may allow for a redundant representation, in that a given note-event may be described in a number of different ways. For example, swung notes may be rendered as deviated sextuplets or as differently deviated sixteenth notes. We include this ambiguity purposefully, because such ambiguities occur frequently and naturally in the types of music under study.

In his 1993 work, Bilmes used signal-processing techniques to extract each of these three quantities from a musical performance. The work demonstrated that analysis of deviations can shed some light on musicians' internal representations of the rhythmic content of their performances, particularly in ensemble contexts that feature fixed and variable rhythmic groups.

### Representation & Implementation

More recently, we have elaborated upon the above model to develop a powerful representation for musical rhythm. Implemented in MAX, a graphical, object-oriented music-programming environment (Puckette 1991), the representation includes features such as pitch, accent, rhythmic deviations, tempo variation, note durations (which are found to carry important rhythmic information and are therefore treated independently), and probabilistic processes. It may be used in conjunction with MIDI instruments or other synthesizers or sound modules.

To facilitate the use of the representation, we have developed various Editors and Players, for creating and musically enacting the data structures, respectively. One of our main goals has been interactive music performance, so our Players have been designed with real-time control in mind. A Player agent steps through the data structures, scheduling and playing the note-events. Multiple data structures are handled with ease thanks to MAX's parallel architecture. Players may also improvise by selecting from banks of rhythmic data or by creating new structures in real time. The Editors consist of graphical user interfaces for creating and modifying data structures in the representation.

The basic unit of our representation is the Cell, a data structure containing a Duration and any number of Note_layers. A given Note_layer contains either a discrete, regular Tatum_grid whose elements contain Notes, or a linked list of Notes occurring at fractional points of the cell duration. The presence of multiple Note_layers allows a rich variety of rhythmic possibilities at the most fundamental level, including multiple tatum rates, "a-tatum" rhythms, hierarchy, and stratification. A Note is a vector containing data about the note type, loudness/velocity, note duration, and microrhythmic deviation (in the discrete case). Thus, expressive microtiming against a metronomic background stands on equal footing with other continually modulating musical parameters.

Among the various tools for manipulating the data structures, we provide a way to exaggerate or de-emphasize rhythmic features by the use of non-linear (power-law) compression and expansion. This technique applies most effectively to deviation and accent data, where subtle expressive features may be either softened or enhanced via a continuous controller.

**Modularity**

Note that the overall design privileges hierarchy at the intra-cellular level, and emphasizes "heterarchy" or modularity at the multi-cellular level. This prioritization favors a modular approach to musical organization. As was pointed out in the chapter on music and embodiment, the modular concept of musical form has special relevance to African and African-American musics. For example, rhythmic textures often arise from the superposition of various cyclic musical patterns. A prime instance of this trait occurs in Afro-Cuban *rumba*, which features fixed cyclic clave and wood-block ostinati, relatively stable repetitive low- and mid-range conga drum patterns, and a variable, heavily improvised *quinto* (high conga drum) part, all combining to form an extremely rich emergent texture. This modular approach may also occur at a higher hierarchical level. Musical pieces may have a number of different repeated sections or "spaces" that cycle for arbitrary lengths of time; the transitions among these spaces are often cued in an improvisatory fashion, quite possibly without a preordained large-scale temporal structure or a strictly linear notion of overall musical time. As mentioned earlier, the music of James Brown provides many examples of this type. The linguistics-derived structural notion of large-scale recursive depth may be replaced or supplemented, for the musician or the listener, by a concept of large-scale organizational breadth. We characterize these methods of musical organization as modular; large musical structures are assembled from small, fully formed constituent units. This mosaic concept functions as an important aesthetic guideline in African and African-American musics, appearing in many different manifestations in the cultures of the continent and their diaspora.

In our implementation, Cells may be combined either in series or in parallel into larger Cells, or they may be cycled indefinitely. The system employs a novel method for handling large numbers of complex rhythmic structures, using features of the MAX collection object, a versatile and open-ended data structure. This method facilitates many of the modular combination techniques mentioned above, such as the stratification of different-length rhythmic cycles, creation of composite beat schemes, repetition, rhythmic progression, and most importantly, improvisatory manipulation of these structures. Thus, a user may create a number of different Cells and select rapidly from among them in real time, superimposing, serializing, or otherwise manipulating the constituent units.

## Applications

The richness of control over many meaningful musical quantities distinguishes our representation from those in more common usage, such as music notation programs, drum machines, while the focus on musical modularity gives the program a different emphasis from traditional sequencers. In addition, as mentioned above, the representation supports creative applications in improvised performance with electronic instruments. We use our implementation to collect rhythmic data from musicians for analysis, and to develop hypotheses and models of rhythm cognition. We have also begun to employ probabilistic processes to construct a useful preliminary representation of rhythmic improvisation, which has been exploited in performance.

## 8. Implications for Music Cognition, Musicology, and Computer Music

My overall thesis has been that music perception and cognition are embodied activities, depending crucially on the tangible features of our sensorimotor apparatus, and also on the sociocultural environment in which music perception, cognition, and production are situated. I have presented some specific evidence in its support, by showing how certain rhythms of African-American music may relate to such embodied processes. I have claimed that musical perception and cognition are active constructions, rather than passive experiences, of the listener. In particular, the perception of pulse and meter are not perceptual inevitabilities, but are strongly dependent on the person's culturally contingent listening strategies. I have also argued that much of what we experience listening to performed music relates to a ecological recognition of, and even an empathy for, the bodily motion of which the musical sounds are a result. These sonic traces of bodily motion can be appreciated as such, and even aesthetically privileged in certain cultures, while neglected or suppressed in others.

The music-cognition community has been somewhat slow to acknowledge fully the role of culture in shaping our ways of perceiving music. Consider the search for the universals of human music cognition. In a recent, rather controversial and quite flawed lecture on possible evolutionary explanations for the existence of music (prefaced by the disclaimer, "I don't know anything about music, *but*..."), evolutionary biologist Steven Pinker made one compelling claim: if we wish to study the basics of music cognition, we should appeal to the musical experiences of the masses (such as hip-hop, the Eurovision song contest, disco, etc.) rather than the art music of high culture. (Pinker 1997) I concur with this claim, but challenge him in his later assumptions that the perception and cognition of music is foremost a solitary, pensive act. His model listener was an idealized, radically de-situated one, with headphones on and eyes closed c in fact an invocation of the classic Western autonomous listener described in many European music-theoretic texts and dating back to Plato, as McClary (1991) has pointed out (see below). When Pinker was asked whether he had any ideas about how group psychology might have affected the evolution of music, he responded that if any such effects did exist, they were profoundly secondary. While it is clear that music listening involves an individual's cognitive systems, one could argue that language comprehension does as well; but nobody would claim that language – particularly spoken language – exists for solitary, pensive activity. Language serves as a means of interpersonal communication, not just of factual information but also of emotions, interrelationships, and imaginings. One could say the same of music. In fact, the connections between music and language are quite far-reaching, as I discuss further below.

In the same lecture, Pinker argued that music could be seen as a "pleasure technology" – a concentrated dose of auditory patterns that happen to give pleasure for other evolutionary reasons. For example, sensory systems grant a sense of pleasure when they receive optimal input – clear, analyzable signals – probably because such signals were favored evolutionarily. Hence one might hypothesize that simply organized musical material would be more popular than pieces with much surface complexity. But this is not the case for many popular musics; salsa and Afro-Cuban *rumba* **[CD-37]** would be judged as quite complex by cultural outsiders. It seems as though a simplicity criterion might be one of many competing criteria in the perception of music as pleasurable or not. One can also think of other functions of music that are not clearly derived from pleasure. For example, consider a musical mnemonic such as the "Alphabet Song," which nearly all young, English-speaking children learn. It aids in the memorization of a large number of symbols by associating it with a familiar melody (that of "Twinkle, Twinkle Little Star") and parsing it into six chunks of decreasing size. Dance music provides another example; not all dance is merely pleasure-related (as in many religious, ritual, and narrative dance forms), but many social situations involving dance seems to reinforce collective unity, which itself could enhance pleasure.

Only someone who knew nothing about music could miss the ugly implications of Pinker's suggestions to study instances of "quasi-music," among which in condescending fashion he included rap music alongside train whistles and wood chopping. Blacking cautions against this "evolutionary" treatment of the development of musical styles (Blacking 1973: 55-56). There is such an absence of description or understanding of the music of so many cultures that it is impossible to judge any music to be primitive. Often a musical style appears simplistic by one culture because it is judged by false criteria. We see this in our own culture: by other musical standards, hip-hop music is musically derivative and dull, because (for example) it doesn't have enough chord changes or melodies. I need not point out that hip-hop culture has its own highly developed, distinctive and elaborate esthetics and standards, nor that it requires highly developed improvisational skills in a variety of domains: vocal, instrumental (turntables), and dance. What should be stressed is that Pinker's ethnocentric comment and his arguments in general betray his ignorance of these dimensions. This instance is emblematic of some of the larger problems in this field, namely the tendency to generalize from a poverty of data, to fetishize certain varieties of musical complexity, and to remain blind to non-European parameters of musical expression.

In any case, questions of complexity defer, in hip-hop as well as in almost all other cases of music and language, to questions of function and utility. Indeed, in Pinker's own book, *The Language Instinct* (1994), he points out that all human languages have pretty much the same degree of complexity; they seem to arise fully formed, regardless of a culture's technological level. Pinker roundly rejects the notorious Sapir-Whorf "relativism" hypothesis (Whorf 1956), which claimed that language and culture shape one another to the degree that certain cognitive abilities, like color classification, are enhanced or stunted by cultural and environmental factors. Instead, Pinker argues, humans are born with a baseline of hardwired cognitive capacities, among which ranks language. In this way, Pinker derives his ideas from Chomsky (see, for instance, Chomsky 1975), and in particular Chomsky's findings of universals or "super-rules" of human language. As an example, he cites data that show that children are able to create the requisite complexity of a full-fledged language, like a creole or American Sign Language, even if their parents speak a pidgin or a shabby version of ASL. Perhaps we could hypothesize that music contains a similar baseline of complexity and a similar set of super-rules, which is distributed among rhythmic, melodic, and other components. But in his talk, Pinker claimed that music showed extreme variation in complexity across cultures, with tonal music representing some sort of pinnacle (Pinker 1997). Blacking cautions against attaching too much significance to musical complexity:

> The issue of musical complexity is irrelevant in any consideration of universal musical competence. First, within a single musical system greater surface complexity may be like an extension of vocabulary, which does not alter the basic principles of a grammar and is meaningless apart from them. Second, in comparing different systems we cannot assume that *surface* complexity is either musically or cognitively more complex. In any case, the mind of man is infinitely more complex than anything produced by particular men or cultures. (Blacking 1973: 34-35)

Furthermore, the research reviewed by Dowling (1988) shows that very young children produce spontaneous songs that incorporate but by no means mimic elements of adult productions. Just as in language development, song development seems to go through a set of ordered, rule-governed stages, somewhat independently of external input. However, as with language, the raw materials for these cognitive processes do not appear in a vacuum; the child requires some basic stimulation to exercise these capacities. It appears as though there does exist some hardwired baseline of musical understanding; however, this basic cognitive ability may atrophy if not nurtured sufficiently in early years, as is the case with language.

Would variations in surface complexity signal variations in fundamental structure? Or might there be other factors? For example, extreme surface complexity may serve as a kind of exclusivity, in the same way that extremely jargonistic language might delineate a certain small professional community. Might increased complexity amount to a kind of augmented vocabulary, rather than a complexified grammar? Perhaps a general musical grammar would include interpersonal factors as well as individual ones; most probably it would include the conceptual scaffolding for embodiment, dance, and collective rhythmic synchronization as well as rules governing melodic and rhythmic structure. If we are to follow this evidence, we may consider the possibility that universal musical competence exist, and consider how it might manifest.

Ultimately, however, we must question the utility of a concept of musical universals. While it has been documented by Brown (1991) (via thorough and painstaking analysis of as many documented ethnographies as possible) that every culture known to man has music and dance – that they themselves are human universals – we should be careful with the limits of the assertion that the same principles underlie all musics of the world. Whatever the role of musical universals, the *particulars* seem to matter just as much. For example, cross-cultural studies suggest that listeners experience great difficulty in intuiting the emotional content of unfamiliar music from another culture (Gregory & Varney 1996). Furthermore, even people of "the same" culture may fail to decode a given piece's emotional content in the same way. Often musical portrayals of exuberance and rage can have similar surface characteristics; for example, these similarities have led to much-contested interpretations of saxophonist John Coltrane's music as alternatively angry or joyous **[CD-53]**. Similarly, musical depictions of sexuality and of violence can be mistaken for one another. (Wessel 1998) Hence, one of music's most unarguable strengths, namely its capacity for emotional expression, appears to be the result of cultural associations rather than purely intramusical dynamics. Just as different cultures have different words for joy or sorrow, they may just as well use different sonic gestures to connote these emotions. The cultural factors that give rise to musical activity provide the richness that distinguishes one music from another, and they do so in a manner that is productive, not limiting.

I will now turn briefly to the implications of this thesis in the realm of computer music. The early days of computer music saw pieces that focused on manipulation of timbral parameters in direct reference to work in music perception and cognition. One might suppose that this work suggests perceptual issues that can be addressed through music itself. The most relevant issues along these lines relate to the body, and to its status in contemporary music.

McClary writes, "The advent of recording has been a Platonic dream come true, for with a disk one can have the pleasure of the sound without the troubling reminder of the bodies producing it. And electronic composition makes it possible to eliminate the last trace of the nonidealist element." (McClary 1991: 136) The implicit prejudices about computer music – that it sounds inhuman, digitized, random, and so forth – are addressed explicitly by the present work. As is the case with programs that improvise "convincing" musical output, programs that generate human-sounding rhythms focus our attention on the role of the same human performer that they might seem to replace. From the few psychological and cultural considerations discussed herein, we could construct a handful of heuristics about rhythmic expression in a groove context; these heuristics can be applied to relevant musical material in an intelligent way. To summarize what was set forth in chapter 7:

- Bear in mind the important aesthetic of microrhythmic asynchrony, as well as its functional role in contributing to rhythmic auditory streaming,
- Watch out for physical impossibilities. Our aesthetics of musical gesture derive from our experiences with bodily gesture. Extremely rapid rhythmic figurations, while not impossible in human performance, nearly always bear the microrhythmic stamp of embodiment. For example, fast figures played in one limb or finger are slightly spread apart in time. We are used to hearing the sound of physical work involved in making music, and this sound includes a specific temporal microstructure.
- We can give a kind of inner life to seemingly monotonous rhythms such as a backbeat or swing, through subtle yet consistent deviations from metronomicity, especially if those deviations somehow contain the sonic trace of physical embodiment.

A simple understanding of such minor adjustments could help musicians working with computers to create music that is rhythmically vital and rich in texture.

Popular music of recent decades has grown quite aware of these possibilities, as its use of technology has catered to its fickle and ever-changing audience. You only need to tune in to any urban radio station to hear that rather convincing electronic tracks have replaced the drummer. Observe the trajectory from the quantized, otherworldly sounds of the Roland TR-808, an analog-synthesis drum machine popular in the early 1980s, to the plasticity of sampled recordings of real drumming manipulated by contemporary software tools. This path suggests the narrative of a popular aesthetic, born of widely available technology, whose participants attempt to make inexpensive rhythmic accompaniment that sound as funky and fresh as their human counterparts, and the counter-narrative of the role of technology in shaping those aesthetics. The uncanny, inhuman sounds of the TR-808 are now enjoying a resurgence thanks to a retro craze, and thanks to the influence of history and memory on popular taste. A contemporary song using the sound of the TR-808 **[CD-54]** implicitly signifies on the past. Every pop tune or "ghetto classic" exists in a universe of signifying associations with other such songs.

Performance variation, musical expression, microtiming – they all suggest the presence of a human body making music. Humans necessarily exhibit some deviation from rigid quantization. Hence, the absence of these deviations implies the absence of a musical body. But this absence can be as musically meaningful as its presence; the strategic use of "robotic" rhythms can suggest a disembodied, techno-fetishistic, futuristic ideal (as in contemporary electronica), or it can embody a Signifying riff on technology, history, and memory (as in contemporary hip-hop referencing the sounds of its beginnings **[CD-54]**).

Often, popular computer music plays in the gray area between bodily presence and electronic impossibility. Again, an example from electronica displays this playful ambiguity **[CD-55]**. A sampled "beat" i.e., a brief recording of a human drummer is sliced into small temporal units. These units are played back in rearranged orders, sped up or slowed down, multiply triggered, and otherwise manipulated electronically. Because the original sampled recording bears the micro-rhythmic traces of embodiment, the result sounds something like a human drummer improvising with often amusing flourishes and ample metric ambiguity. Momentarily regular, almost human-sounding pseudo-drumming devolves into inhumanly rapid sequences of rhythmic attacks, fast enough to resemble digital noise. Such electronic manipulation of familiar musical sounds serves to problematize the listener's ecologically sound image of a human drummer.

Another prime example of the play of embodiment in contemporary popular music is the hip-hop DJ, who treats the turntables as a kind of percussion meta-instrument **[CD-56]**. Using strategically chosen segments of a vinyl record, the DJ moves the record back and forth with one hand, while creating amplitude envelopes with a fader on a mixer in the other hand. The sound generated is of two general types: one is a percussive scratch derived from rapid motion of the record, and the other is a recognizable, meaningful fragment of recorded music or sound. The latter stroke type often hides the sophisticated, impeccably timed physical gestures involved in their creation, as these gestures are unrelated to the sonic material. The scratch sound, however, bears a direct sonic resemblance to the physical motion involved. There is an interesting continuum between these two general types. A fragment of recorded sound can be manipulated percussively in a manner that temporarily overrides its referential content, causing it to refer instead to the physical materiality of the vinyl-record medium, and more importantly to the embodiment, dexterity and skill of its manipulator.

This play with the ambiguity of embodiment also appears in some more experimental realms of computer music. Improvising kotoist Miya Masaoka **[CD-57]** augments the physical capacities of her wooden, stringed classical instrument with electronic sensors that drive a bank of synthesizers and samplers. The sensors track her physical gestures as well as the pitch material from each string. Her creative mapping of this data to the electronic sound sources results in a sort of electronic-acoustic hybrid instrument, which she calls the "Koto-Monster." As an improviser, she makes use of this expanded palette with an organic seamlessness that blurs the boundary between the acoustic (physical, embodied) and the electronic (artificial, disembodied) realms. In a similar vein, Laetitia Sonami has developed a sleek lady's glove into a sensitive gestural controller that tracks dozens of dimensions of manual movement. In her performances, she transforms primal gestures of the hand into sonic elements that seem to bear the trace of these gestures. Her sound material often consists of non-melodic, sampled sounds that also reference their own physical sources (speaking voices, animal sounds, wind, etc.). The result is as fascinating to watch as to hear, as one discerns an emergent connection between the hand motions and the disembodied sound material, and one becomes aware that a versatile instrument is being played with great skill.

Also in the realm of sensitive, expressive controllers, David Wessel **[CD-58]** has developed a productive framework for improvising with the Buchla Thunder, a novel electronic instrument with two dimensions of continuous control (position and pressure) for each fingertip. In his setup, each pressure controller acts as a volume fader that brings up a computer-driven rhythmic process, which is always in motion. The position control of each finger is set to manipulate various rhythmic parameters of the associated process, such as density, timing, or timbre. The ten fingers can create richly variable musical gestures that act as larger constructs on top of an implied rhythmic undercurrent. The musical material thus generated has the quality of hand gestures, but the musical totality is not the simple result of these gestures. Rather, it is the novel interaction between the hand motions and the computer rhythm engine that gives rise to the hybrid musical texture.

Working in a slightly different paradigm, improvising trombonist George Lewis **[CD-59]** has built a computer program that improvises polyphonically, with other musicians or without them. Quite expansive in timbral scope, it gives the sense of an improvising orchestra that produces focused, flowing music. Its output relates to its human colleague's sonic input just enough to convince us that it is listening, without sounding imitative. When unstimulated by musical input, it "takes a solo," seemingly unfazed. Aesthetically it fits quite squarely into the world of improvised music of the last three or four decades. It draws its inspiration from the collective improvisations of artists like the Art Ensemble of Chicago, Anthony Braxton, and Muhal Richard Abrams, all of whom (along with Lewis himself) are associated with the Chicago-based African-American musical collective known as the Association for the Advancement of Creative Musicians.

Indeed, the program makes musical choices with enough depth and wisdom that one easily forgets that it is a purely disembodied computer program. In listening to this piece of artificial intelligence, we begin to perceive what we call a *sound* c the sonic traces of a creative personality. Lewis's work addresses these issues of embodiment, by creating a distinct sense of embodied artistry out of a laptop computer and some synthesizer modules.

We can imagine a multitude of further possibilities of exploration of a meaningful continuum between these two poles of absence and presence. For her recent electronica album, pop diva Madonna says that she wanted to explore the possibility of giving that music's characteristically inhuman sound "a soul." (Rule 1998) Her solution was heartfelt (if exceedingly banal) lyrics delivered by an instantly recognizable, celebrity voice **[CD-60]**. But now that we have begun to analyze the actual sonic trace of the human body in the microrhythmic content of instrumental music, perhaps we can further problematize the longstanding cliché that electronic music fails to provide a sense of soul. For what is soul in music, if not a powerfully embodied human presence? **[CD-61]**

# Bibliography


Agawu, V. K. 1986. "'Gi Dunu','Nyekpadudo,' and the study of West African Rhythm." *Ethnomusicology* **30**(1): 64-83.

______. 1992. "Representing African Music." *Critical Inquiry* **18**(2): 245-266.

______. 1995. *African Rhythm: A Northern Ewe Perspective.* Cambridge: Cambridge University Press.

Anku, W. 1992. *Structural Set Analysis of African Music.* Legon, Ghana: Soundstage Production.

Arom, S. 1991. *African Polyphony and Polyrhythm.* Cambridge: Cambridge University Press.

Austin, J. L. 1962. *How to Do Things with Words.* Cambridge, Mass.: Harvard University Press.

Baily, J. 1985. "Music Structure and Human Movement." In Howell, Cross, & West, eds., *Musical Structure and Cognition* (London: Academic Press), 237-258.

_____. 1989. "Principles of rhythmic improvisation for the Afghan *rûbab*." *International Council for Traditional Music UK Chapter Bulletin* **22**: 3-16.

Baily, J. & Driver, P. 1992. "Spatio-Motor Thinking in Playing Folk Blues Guitar." *The World of Music* **34**(3): 58-71.

Barthes, R. 1977. "The Grain of the Voice." In *Image, Music, Text.* S. Heath, ed./trans. (New York: Hill & Wang), 179-189.

Berger, H. M. 1997. The practice of perception: Multi-functionality and time in the musical experiences of a heavy metal drummer. *Ethnomusicology* **41**(3): 464-488.

Berliner, P. 1994. *Thinking in Jazz: The Infinite Art of Improvisation.* Chicago: University of Chicago Press.

Bilal, M. 1997. Private communication.

Bilmes, J. A. 1993. *Timing is of the Essence: Perceptual and Computational Techniques for Representing, Learning, and Reproducing Expressive Timing in Percussive Rhythm.* Masters thesis, Massachusetts Institute of Technology.

______. 1996. Private communication.

Blacking, J. 1973. *How Musical is Man?* Seattle: University of Washington Press.

Bourdieu, P. 1977. *Outline of a Theory of Practice,* translated by R. Nice. Cambridge: Cambridge University Press.

Bregman, A. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, Mass.: MIT Press.

Brooks, R. 1991. "Intelligence without Representation." *Artificial Intelligence* **47**: 139-159.

Brower, C. 1993. "Memory and the Perception of Rhythm." *Music Theory Spectrum* **15**: 19-35.

Brown, D. E. 1991. *Human Universals.* New York: McGraw-Hill.

Carroll-Phelan, B., & P. J. Hampson. 1996. "Multiple Components of the Perception of Musical Sequences: A Cognitive Neuroscience Analysis and Some Implications for Auditory Imagery." *Music Perception* **13**(4): 517-561.

Clancey, W. 1997. *Situated Cognition: On Human Knowledge and Computer Representation.* New York: Cambridge University Press.

Clark, A. 1997. *Being There: Putting Brain, Body, and World Together Again.* Cambridge, Mass.: Bradford/MIT Press.

Clarke, E. 1988. "Timing in the performance of Erik Satie's 'Vexations'." *Acta Psychologica* **50**: 1-19.

______. 1999. "Rhythm and Timing in Music." In Deutsch, D., ed., *The Psychology of Music* (2nd edition) (New York: Academic Press), 473-500.

Clynes, M. and Walker, J. 1982. "Neurobiologic Functions of Rhythm, Time and Pulse in Music." In Clynes, M., ed., *Music, Mind, and Brain* (New York: Plenum Press), 171-216.

Coelho, O. 1995. "Are Representations still necessary for understanding Cognition?" In Niklasson, L. & Boden, M., eds., *Current Trends in Connectionism* (Hillsdale, NJ: Lawrence Erlbaum Associates), 311-320.

Condon, W. S. 1982. "Cultural Microrhythms." In Davis, M., ed., *Interaction Rhythms* (New York: Human Sciences Press), 77-102.

Davidson, L. & Torff, B. 1992. "Situated Cognition in Music." *The World of Music* **34**(3): 120-139.

Dean, R. 1992. *New Structures in Jazz and Improvised Music since 1960.* Philadelphia: Open University Press.

Desain, P. and Honing, H. 1996. "Physical Motion as a Metaphor for Timing in Music: The Final Ritard." *Proceedings of the International Computer Music Conference*, 458-460.

Dowling, W. J. 1988. "Tonal structure and children's early learning of music." In Sloboda, J., ed., *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition* (New York: Clarendon Press), 113-128.

Drake, C. & Palmer, C. 1993. "Accent Structures in music performance." *Music Perception* **10**(3): 343-378.

Dunn, L. & Jones, N. 1994. *Embodied Voices: Representing Female Vocality in Western Culture.* Cambridge, U.K.: Cambridge University Press.

Feldman, J. 1997. Class lecture at U.C. Berkeley, September 10.

Floyd, S. 1995. *The Power of Black Music.* New York: Oxford University Press.

Forguson, L.W. 1969. "In Pursuit of Performatives." In Fann, K.T., ed., *Symposium on J. L. Austin* (New York: Humanities Press), 413-419.

Fraisse, P. 1956. *Les Structures Rhythmiques.* Louvain, Paris: Publications Universitaires de Louvain.

______. 1982. "Rhythm and Tempo." In D. Deutsch, ed., *The Psychology of Music* (New York: Academic Press), 149-180.

Gates, H.L. 1988. *The Signifying Monkey: A Theory of African-American Literary Criticism.* New York: Oxford University Press.

Gibson, J. J. 1966. *The Senses Considered as Perceptual Systems.* Boston: Houghton Mifflin.

______. 1975. "Events are perceivable but time is not." In Fraser, J.T. & N. Lawrence, eds., *The Study of Time II* (New York: Springer-Verlag), 295-301.

______. 1979. *The Ecological Approach to Visual Perception.* Boston: Houghton Mifflin.

Gjerdingen, R. 1989. "Meter as a Mode of Attending: A Network Simulation of Attentional Rhythmicity in Music." *Intégral* **3**: 67-91.

Goodheart, M. 1996. "Freedom and Individuality in the Music of Cecil Taylor." Masters thesis, Mills College, Oakland, California.

Gregory, A. H. 1997. "The roles of music in society: the ethnomusicological perspective." In Hargreaves, D. J. & A. C. North, eds., *The Social Psychology of Music* (New York: Oxford University Press), 123-140.

Handel, S. 1990. *Listening.* Cambridge, Mass.: MIT Press.

Hardcastle, V. 1996. *How to Build a Theory in Cognitive Science.* Albany, NY: SUNY Press.

Held, R. & Hein, A. 1958. "Adaptation of disarranged hand-eye coordination contingent upon re-afferent stimulation." *Perceptual-Motor Skills* **8**:87-90.

Honing, H. 1993. "Issues in the representation of time and structure in music." Contemporary Music Review **9**: 221-239.

Hughes, D. W. 1991. "Grammars of Non-Western Music: A Selective Survey." In Howell, P. et al., eds., *Representing Musical Structure* (New York: Academic Press), 327-362.

Ivry, R. 1998. Lecture delivered at University of California, Berkeley, Sept. 11.

Iyer, V., Bilmes, J., Wright, M., & Wessel, D. 1997. "A Novel Representation for Rhythmic Structure." *Proceedings of the 1997 International Computer Music Conference* (San Francisco: International Computer Music Association), 97-100.

Jagacinski, R., Marshburn, E., Klapp, S., and Jones, M. 1988. "Tests of parallel versus integrated structure in polyrhythmic tapping." *Journal of Motor Behavior* **20**(4): 416-442.

Jones, A. M. 1959. *Studies in African Music.* London: Oxford University Press.

Jones, M. R. 1986. "Attentional Rhythmicity in Human Perception." In Evans, J. & Clynes, M., eds., *Rhythm in Psychological, Linguistic, and Musical Processes* (Springfield, Ill.: Charles Thomas Publishers), 83-99.

Jones, M. R. & Boltz, M. 1989. "Dynamic Attending and Responses to Time." *Psychological Review* **96**(3): 459-491.

Jones, M. R. & Yee, W. 1993. "Attending to auditory events: the role of temporal organization." In S. MacAdams & E. Bigand, eds., *Thinking in Sound: The Cognitive Psychology of Human Audition* (Oxford: Oxford University Press), 69-112.

Jost, E. 1981. *Free Jazz.* New York: Da Capo Press.

Keil, C. & Feld, S. 1994. *Music Grooves.* Chicago: University of Chicago Press.

Ladzekpo, C. K. 1995. Foundation Course in African Dance Drumming. http://cnmat.CNMAT.Berkeley.EDU/~ladzekpo/Foundation.html

Lakoff, G. 1987. *Women, Fire, and Dangerous Things.* Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. 1980. *Metaphors We Live By.* Chicago: University of Chicago Press.

Large, E. 1994. *Dynamic Representation of Musical Structure.* Unpublished Ph.D. dissertation. The Ohio State University, Columbus, Ohio.

______. 1997. "Modeling the Perception of Meter in Music Performance." Talk presented at the annual meeting of the Society for Music Perception and Cognition, MIT, Cambridge, Mass.

Large, E. & Kolen, J. 1994. "Resonance and the Perception of Musical Meter." *Connection Science* **6**(1): 177-208.

Lashley, K. S. 1951. "The Problem of Serial Order in Behavior." In L. A. Jeffries, ed., *Cerebral Mechanisms in Behavior: The Hixon Symposium* (New York: John Wiley & Sons), 112-146.

Laughlin, C., McManus, J., D Aquili, E. 1990. *Brain, Symbol, & Experience: Towards a Neurophenomenology of Human Consciousness.* New York: Columbia University Press.

Lave, J. 1988. *Cognition in Practice: Mind, Mathematics, and Culture in Everyday Life.* New York: Cambridge University Press.

Lenat, D. & Feigenbaum, E. 1992. "On the thresholds of knowledge." In Kirsh, D., ed., *Foundations of Artificial Intelligence* (Cambridge, Mass.: MIT Press), 47.

Lerdahl, F. and Jackendoff, R. *A Generative Theory of Tonal Music*. Cambridge: MIT Press, 1983.

Lewis, G. 1995. Unpublished essay.

_____. 1996. "Improvised Music since 1950: Afrological and Eurological forms." *Black Music Research Journal* **16**(1): 91-119.

London, J. 1997. "Can you hear two meters at once? Metric ambiguity, metric vagueness, and music theory." Talk given at the annual meeting of the Society for Music Perception and Cognition, Cambridge, Mass.

Longuet-Higgins, H. C. and Lee, C. S. 1982. "The Perception of Musical Rhythms." *Perception* **11**: 115-128.

______. 1984. "The Rhythmic Interpretation of Monophonic Music." *Music Perception* **1**(4): 424-441.

Magill, J. M. & J. L. Pressing. 1997. "Asymmetric cognitive clock structures in West African Rhythms." *Music Perception* **15**(2): 189-222.

Malotki, E. 1983. *Hopi Time: A Linguistic Analysis of Temporal Concepts in the Hopi Language.* Berlin: Mouton.

Mataric, M. (ed.) 1996. *Embodied Cognition & Action: Papers from the 1996 AAAI Fall Symposium.* Menlo Park, CA: AAAI Press.

McNeill, D. 1998. *Hand and Mind.* Chicago: University of Chicago Press.

McClary, S.. 1991. *Feminine Endings: Music, Gender, and Sexuality.* Minneapolis, MN: University of Minnesota Press.

McClary, S. & Walser, R. 1994. "Theorizing the Body in African-American Music." *Black Music Research Journal* **14**(1): 75-84.

Merriam, A. 1982. "African musical rhythm and concepts of time-reckoning." In *African Music in Perspective* (New York: Garland Publishers), 443-461.

Meyer, L. 1956. *Emotion and Meaning in Music.* Chicago: University of Chicago Press.

Michon, J. A. 1975. "Time Experience and Memory Processes." In J. T. Fraser and N. Lawrence, ed., *The Study of Time II* (New York: Springer-Verlag), 302-313.

Miller, G. 1956. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* **63**: 81-97

Minsky, M. 1985. *The Society of Mind.* New York: Simon and Schuster.

Monson, I. 1996. *Saying Something: Jazz Improvisation and Interaction.* Chicago: University of Chicago Press.

Narmour, E. 1990. *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model.* Chicago: University of Chicago Press.

Nattiez, J.-J. 1990. *Music and Discourse.* Princeton, N.J.: Princeton University Press.

Pantaleoni, H. 1972. "Three principles of timing in Anlo Dance Drumming." *African Music* **5**(2): 50-63.

Parker, A. and Sedgwick, E., eds. 1995. *Performativity and Performance.* New York: Routledge.

Parncutt, R. 1994. "A perceptual model of pulse salience and metrical accent in musical rhythm." Music Perception **11**(4): 409-464.

Passingham, R. 1982. *The Human Primate.* Oxford: Freeman.

Peretz, I. 1993. "Auditory agnosia: a functional analysis." In S. McAdams & E. Bigand, eds., *Thinking in Sound: The Cognitive Psychology of Human Audition* (Oxford: Clarendon Press), 199-230.

Pinker, S. 1994. *The Language Instinct.* New York: William Morrow & Co.

______. 1997. "The Evolution of Music." Plenary session given at the annual meeting of the Society for Music Perception and Cognition, Massachusetts Institute of Technology, August 1.

Povel, D.-J. 1984. "A theoretical framework for rhythm perception." *Psychological Research* **45**: 315-337.

Preusser, D. 1972. "The effect of structure and rate on the recognition and description of auditory temporal patterns." *Perception & Psychophysics* **11**(3): 233-240.

Puckette, M. 1991. "Combining Event and Signal Processing in the MAX Graphical Programming Environment." *Computer Music Journal* **15**(3): 58-67.

Ramachandran, V. S. & Blakeslee, S. 1998. *Phantoms in the Brain: Probing the Mysteries of the Human Mind.* New York: William Morrow & Co.

Rasch, R. A. 1988. "Timing and synchronization in ensemble performance." In Sloboda, J. A., ed., *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition* (Oxford: Clarendon Press), 70-90.

Repp, B. H. 1990. "Patterns of Expressive Timing In Performances of a Beethoven Minuet by Nineteen Famous Pianists." *Journal of the Acoustical Society of America* **88**: 622-641.

Rosenbaum, A. 1998. *Shout Because You re Free: The African American Ring Shout tradition in coastal Georgia.* Athens, Georgia: University of Georgia Press.

Rowell, L. 1988. "The idea of music in India and the Ancient West." *Acta Philosophica Fennica* **43**, 323-342.

Rule, G. 1998. "William Orbit: The Methods and Machinery behind Madonna's *Ray of Light.*" *Keyboard* **24**(7): 30-38.

Sacks, O. 1985. *The Man who Mistook His Wife for a Hat and Other Clinical Tales.* New York: Summit Books.

Said, E. 1991. *Musical Elaborations.* New York: Columbia University Press.

Schenker, H. 1979. *Free Composition* (E. Oster, trans.). New York: Longman. (Original work published 1935.)

Schulze, H.-H. 1978. "The detectability of local and global displacements in regular rhythmic patterns." *Psychological Research* **40**: 173-181.

Schutz, A. 1964. "Making Music Together." In *Collected Papers II: Studies in Social Theory* (The Hague: Martinus Nijhoff), 159-178.

Searle, J. 1980. "Minds, brains, and programs." *Behavioral and Brain Sciences* **3**: 417-457.

______. 1996. Unpublished lecture at University of California, Berkeley.

Sessions, R. 1950. *The Musical Experience of Composer, Performer, Listener.* Princeton: Princeton University Press.

Shaw, R. & Pittenger, J. 1978. "On perceiving change." In Pick, H. & Saltzman, E., eds., *Modes of Perceiving and Processing Information* (Hillsdale, NJ: Erlbaum), 187-204.

Shore, B. 1996. *Culture in Mind: Cognition, Culture, and the Problem of Meaning.* New York: Oxford University Press.

Shove, P. & Repp, B. 1995. "Musical motion and performance: theoretical and empirical perspectives." In Rink, J., ed., *The Practice of Performance* (Cambridge, U.K.: Cambridge University Press), 55-83.

Shortliffe, E. 1976. *Computer Based Medical Consultations: MYCIN.* Elsevier.

Smithers, T. 1996. "On What Embodiment Might Have to do with Cognition." In M. Mataric, ed., *Embodied Cognition and Action: Papers from the 1996 AAAI Fall Symposium (Technical Report FS-96-02)* (Menlo Park, CA: AAAI Press), 113-116.

Southern, E. 1983. *The Music of Black Americans.* New York: W. W. Norton & Co.

Stone, R. 1985. "In search of time in African music." *Music Theory Spectrum* **7**.

Stone, R. 1986. "Commentary: The value of local ideas in understanding West African rhythm." *Ethnomusicology* **30**(1): 84-88.

Street, A. 1989. "Superior Myths, Dogmatic Allegories: The Resistance to Musical Unity." *Music Analysis* **8**: 77-123.

Stuckey, S. 1987. *Slave Culture: Nationalist Theory and the Foundations of Black America.* New York: Oxford University Press.

Sudnow, D. 1978. *Ways of the Hand.* Cambridge, Mass.: Harvard University Press.

Swinnen, S.P. and Walter, C.B., "Constraints in coordinating limb movements," in Colley, A.M. & Beech, J.R., ed., *Cognition and Action in Skilled Behavior* , Elsevier Science Publishers B.V. (North-Holland), 1988.

Taylor, C. 1959. "John Coltrane." In *Jazz Review*, January, p. 34.

Thelen, E., & Smith, L. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action.* Cambridge, Mass.: MIT Press.

Todd, N. 1989. "Towards a Cognitive Theory of Expression: The Performance and Perception of Rubato", *CMR* **4**: 405–416.

_____. 1994. "The Auditory 'Primal Sketch': A Multiscale Model of Rhythmic Grouping." *Journal of New Music Research* **23**:25-70.

_____. 1997. Message sent to the auditory-l listserve, November 22.

Todd, N., Lee, C., & O Boyle, D. 1998. "A Sensory-Motor Theory of Rhythm, Time Perception and Beat Induction." *Journal of New Music Research* (in press).

Varela, F., Thompson, E., & Rosch, E. 1991. *The Embodied Mind: Cognitive Science and Human Experience.* Cambridge, Mass.: MIT Press.

Verborg, D., and Hambuch, R. 1978. "On the temporal control of rhythmic performance." In Requin, J., ed., *Attention and Performance VII* (Hillsdale, NJ: Lawrence Erlbaum Associates), 535-555.

von Neumann, J. 1951. "The General and Logical Theory of Automata." In Jeffress, L., ed., *Cerebral Mechanisms in Behavior: The Hixon Symposium* (New York: John Wiley & Sons), 1-41.

Walser, R. 1995. "'Out of Notes': Signification, Interpretation, and the Problem of Miles Davis." In Gabbard, K., ed., *Jazz Among the Discourses* (Durham: Duke University Press), 165-188.

Waterman, R. A. 1952. "African influence on the music of the Americas." In *Acculturation in the Americas* (ed. S. Tax). University of Chicago Press, Chicago. Reprinted in Shelemay, K. K. (ed.) 1990. *The Garland library of readings in ethnomusicology. Vol. 3: Music as culture.* New York: Garland.

Wessel, D. 1998. Private communication.

Whorf, B. 1956. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf.* J. Carroll, ed. Cambridge, Mass.: MIT Press.

Wilson, O. 1974. "The Significance of the Relationship between Afro-American Music and West-African Music." *The Black Perspective in Music* **2**: 3-22.

______. 1981. "The Association of Movement and Music as a Manifestation of a Black Conceptual Approach to Music Making." In Heartz, D. & Wade, B., eds., *International Musicological Society Report of the Twelfth Congress* (Basel: Bärenreiter Kassel), 98-105.

Windsor, W. & Clarke, E. 1997. "Expressive Timing and Dynamics in Real and Artificial Musical Performances: Using an Algorithm as an Analytical Tool." *Music Perception* **15**(2): 127-152.

Wing, A. & Kristofferson, A. 1973. "The Timing of Interresponse Intervals." *Perception and Psychophysics* **14**: 5-12.

## Discography


Addy, Mustapha Tettey. 1998. *Secret Rhythms* (a.k.a. *Drummer By Nature*). Jork, Germany: WeltWunder Records.

Brown, James. 1991. *Star Time.* Compact disc compilation of original releases 1956-1984. New York: PolyGram Records.

Busta Rhymes. 1996. *The Coming.* New York: Elektra Entertainment Group.

______. 1997. *When Disaster Strikes.* New York: Elektra Entertainment Group.

Coleman, Ornette. 1990. *Love Call.* Compact disc reissue of original 1968 recording. Hollywood, CA: Blue Note/Capitol Records.

Coltrane, John. 1993. *Transition.* Compact disc reissue of original 1970 release, recorded 1965. New York: GRP Records.

______. 1998. *The Complete 1961 Village Vanguard Recordings.* Compact disc reissue of original recordings. New York: GRP Records.

Davis, Miles. 1998. *Panthalassa: The Music of Miles Davis 1969-1974.* Compact disc issue of remixed original recordings. New York: Columbia Records.

De La Soul. 1996. *Stakes is High.* New York: Tommy Boy Music.

Dolphy, Eric. 1982. *Out There.* Compact disc reissue of original 1960 recording. Berkeley, CA: New Jazz Records.

Ellington, Duke. 1990. *At Newport.* Compact disc reissue of original 1956 recording. New York: Columbia Records.

Jamal, Ahmad. 1980. *What's New.* Compact disc compilation of original 1952 & 1958 releases. Woodland Hills, CA: Telstar Records.

Lewis, George. 1993. *Voyager.* Tokyo: Avant Records/Disk Union.

Madonna. 1998. *Ray of Light.* New York: Maverick/Warner Brothers Records.

Monk, Thelonious. 1986.*The Complete Riverside Recordings.* Compact disc compilation of recordings 1955-1961. Berkeley: Riverside/Fantasy Records.

______. 1994. *The Complete Blue Note Recordings.* Compact disc compilation of recordings 1947-1958. Hollywood: Blue Note/Capitol Records.

______. 1998. *Monk Alone: The Complete Columbia Solo Studio Recordings: 1962-1968.* Compact disc compilation. New York: Columbia Records/Sony Music.

______. 1998. *Live at the It Club: Complete.* Compact disc reissue of original 1965 release, recorded 1964. New York: Columbia Records.

Los Muñequitos de Matanzas. 1990. *Cantar Maravilloso.* London: GlobeStyle Records.

Parker, Charlie. 1988. *One Night at Birdland.* Compact disc reissue of original 1950 recording. New York: Columbia Records.

Rakim. 1997. *The 18th Letter: The Book of Life.* New York: Universal Records.

Sayeeram, Aruna. 1995. *The Lyrical Tradition of Carnatic Music 1.* Paris: Makar.

Squarepusher. 1997. *Big Loada.* Sheffield, U.K.: Warp Records.

Tatum, Art. 1991. *The Complete Pablo Solo Recordings.* Compact disc reissue of original 1953-55 recordings. Los Angeles: Pablo Records.

Taylor, Cecil. 1995. Cassette recording of rehearsal with the Cecil Taylor Creative Orchestra, from the author's private collection.

Terminator X. 1991. *Terminator X & The Valley of the Jeep Beats.* New York: Rush Associated Labels / Columbia Records