Book Report on:

Rhythm in Psychological, Linguistic, and Musical Processes, edited by James R. Evans and Manfred Clynes. (Springfield, Illinois: Charles C. Thomas Press, 1986)

A Special Report for the Cognitive Eurhythmics Community

by Eric Barnhill

The book *Rhythm in Psychological, Linguistic and Musical Processes*, published nearly 20 years ago, is a treasure trove of scientific studies connecting rhythm with many dimensions of human functioning. The subject of rhythm, and its role in the human organism, continues to sit outside the paradigm of mainstream therapy, treatment and medicine, and so these fascinating studies, through rigorous and reviewed, have had only peripheral impact on the treatment of people in clinical, therapeutic or educational contexts.

These ideas, however, intersect greatly with what we do in Cognitive Eurhythmics class. I therefore thought that a review of parts of the book would be of great value not only to pass on information parents and others will likely find interesting and valuable, but also to help clarify and explain some of the goals and techniques of Cognitive Eurhythmics to parents and others who are interested.

The heart of this volume is Chapter 2, "Attentional Rhythmicity in Human Perception," which presents a new, rhythmic model for the way the mind uses attention. As its author correctly notes, attention is one of the fundamental acts of consciousness - how the mind directs its energies - and traditional studies of attention have been tremendously impoverished. These theories use overly simplistic models divorced from reality, and attempt to draw conclusions independent of any speculation of mechanism. That is to say, attention is studied without any reference to how it works, what it might be useful for, or why we evolved to have a faculty of attention in the first place.

The author of Chapter 2, Mari Riess Jones, first reflects on the shortcomings of contemporary theories of attention. The most common current model of attention is called a "resource capacity" model. In this model, different amounts of attention are parceled out in different directions, some more than others. Attention can become more focused in one area, to the neglect of others, or more diffuse, and seems to always be happening on a minimal level even in what are called "unattended channels", for example when one hears one's name out of the otherwise undifferentiated noise of a crowd.

While Jones admires the way the model accounts for selectivity, she counters that "none of these approaches directly addresses the nature of attentional energy. Nor do they attempt to specify determinants of energy capacity limitations in various tasks." Further, attention is considered some sort of self-contained force, ignoring its relationship to "the complex issue[s] of skill acquisition and perceptual learning." Despite the known central importance of attention to cognitive performance, therapy, and education, people are assumed merely to have more or less of "it" without much explanation of what it is, or its

impact on other aspects of cognitive performance. Part of the reason that models of attention ignore the relationship to learning and performance, she asserts, is that "much of the evidence used to support these models is based on tasks that lack a coherent structure to which people may gradually become attuned," but were rather "random and contentless tasks".

Jones also explains "schematic" models of attention, which she praises for correctly observing that "attending is anticipatory," and that our attention tends to proceed along a prior schema of expectations for what we are going to see. She credits the researchers involved for seeing this, probably because the attentional tasks they gave were substantial and subjects showed an improvement in attention over time. Still, despite their recognition that we structure our attention in some sort of way, the researchers did not come any closer to the nature of that structure except to observe that attention "got better with practice".

Jones then presents her "Rhythmic Model of Attending", which asserts that "people and other animals target attending over predetermined time intervals toward events in space and time in a rhythmical fashion". The missing link that brought her to her theory was her realization that conventional theories of attention carried an implicit premise that "absolute time, or sequence rate, and not relative time, or sequence rhythm, are assumed to determine attentional resources." Rather than think of attention as some linear force, she believes the nature of attention is revealed if we think of "time patterning of attentional energy within an organism: attending may either lag behind, be in phase, or run ahead of reference events in the world." Such a theory has "less emphasis on capacity limits", which do not convincingly explain the extreme variations in capacity that we demonstrate, and emphasizes "temporal relationships that exist within an organism and its environmental surrounds", which accounts for among other things, the widely confirmed fact that "information pickup is facilitated...by temporal predictability," such as mnemonic devices and rhymes. "In the rhythmical interpretation," Jones writes, "[attentional] energy is explicitly tied to natural rhythmicities of living things." This aspect of attention was misunderstood in previous experiments, because the stimuli attended to were contentless and did not account for "the dynamic structure of stimulation."

Jones accompanies her theory with a robust explanation for how this attention really works. Stimuli can be attended to in multiple ways, and we must choose the strategy that allows us to selectively attend to what we want. "Often a temporal patterning or periodicity within the stimulus co-varies with its to-be-attended affordance. Temporal patterns associated with speech phonemes, for example, differ from others associated with sentential meaning...Alternatively, to catch the sarcasm of a speaker's utterance...attending is to longer and differently patterned time-based cadences in the speech stream." This example lists three different ways we may attend to the same words: as sound, as words with meaning, and for a sarcastic tone underneath the surface meaning. How are we able to select which to attend to? "Perceived needs may call forth within the organism particular 'calibrated' shapes of energy over time and these attentional shapes selectively guide attending to 'lock-into' matching affordances

embedded within an ostensibly ambiguous stimulus pattern." Rather than a fixed pool of attention that is directed around the environment, the brain creates "goal-oriented attentional schemes", shapes of attention that rhythmically unfold over time enabling us to track the attributes of a stimulus that we want to follow in a "relativistic temporal interplay of scheme and environment." Attention thus always carries a predictive element, as the attention's shape must predict future peaks and troughs of attention.

Observing a predictive element in attention would align it with what we know about perception and motor action. Rather than a stimulus-response model for taking perceptions into action, we now know from movement science research that no movement is initiated without the brain making a predictive guess as to its outcome and setting a frame of reference with which to judge and correct error. It is no surprise that attention functions with a similar mechanism, setting a projective rhythmic pace that allows it to track whatever element in all the noise and confusion that it wants to follow. As attention is educated, it will gain in skill and subtlety at tracking the affordances in the environment relevant for the person, and timing whatever response is judged appropriate; hence, attention is not surprisingly at the service of perception, motor learning, and coordination of action.

Jones applies the rhythmic model of attention to language, and in so doing opens up the door connecting music and the musical nuance embedded in language that is expanded upon in later chapters. "There are distinctive patterns of loudness over time revealed in human speech and music which convey, through intonation, intended (and sometimes unintended) meanings. Here meaning between speaker and listener is communicated at a tacit level through the attentional act of inner mimicry, a mimicry induced by entrainment of attentional energies of the listener with intonations of the speaker....[I]ntonation patterns in grammatical contexts effectively guide listeners to pick up the relevant new material which follows." As we will see below, intonation in fact allows us to process information in a far more powerful and sophisticated manner by allowing us to set a predictive structure to the input that enables us to track a much longer and more complex thought without taxing short-term memory.

To expand on this model in a way Jones didn't, a spoken sentence has two dimensions: the semantic, dealing with the meaning of the sentence right now, expressed in the word, and the predictive, creating a shape for our attention to follow, expressed in the music. Such a hypothesis connects to speech research that has determined that we only hear about 70% of a given sentence, making sense out of it mostly by inference, and work with children that shows rhythmic cues as an essential aspect of language comprehension even through age 10. In a different direction, Jones' research also synchronizes completely with the theories of the great musicologist Leonard B. Meyer, who felt that the fundamental act of music was to stir our emotions by creating expectations and then artfully violating them.

Of the chapters I found important in this book, the Rhythmic Theory of Attending is the most powerful and fruitful. The other three that I found useful - dealing with memory,

speech, and reading - can be looked at as elaborations of offshoot issues raised by this theory of attention.

In the case of memory, Payne and Holzman examine the common knowledge that information digested in a rhythmic framework, such as a rhyming or mnemonic device, is better understood and digested in Chapter 3, "Rhythm as a Factor in Memory". Rhythmic arrangement of information turns out to be an essential part of how we store and retrieve it, and viewed with a discerning eye, leaves its stamp on all that we claim to recall accurately. "Subjects have been found to impose a rhythmic organization on material they attempt to remember even when that material is unpatterned." As an example, "free recall of brief English phrases appears to be organized in terms of stress rhythm, i.e. accent, or syllables."

I can add that this is confirmed by more recent research by Sloboda on immediate recall of melodies, in which "the most fundamental feature that is preserved...is the metrical structure...coded into two bar phrases." Melodic and harmonic representations are much less accurate, through they are usually substituted by metrical and harmonic equivalents.

For the most part, the essay by Payne and Holzman uses examples that hold little interest, using pure tones with fixed durations. These examples, like so many in the psychology of music, inevitably throw the baby out with the bathwater by being what Jones, above, termed random and contentless. They do mention a very striking bit of research, though, which indicates that "auditory sequential memory deficit is a major factor in at least some retarded readers' performance. This accords with Jones...who showed that retarded readers were unable to hold...a rhythmic pattern in memory."

In Chapter 2, Jones enumerates one incarnation of the Rhythmic Theory of Attention by describing an "attentional act of inner mimicry" between a speaker and a listener, " an inner activity of "entrainment of attentional energies of the listener with intonations of the speaker." If Jones captures the inner essence of this behavior, William S. Condon puts the flesh on it in his excellent Chapter 4, "Communication: Rhythm and Structure":

"A speaker's body is observed to move in organizations of change which are precisely synchronized with the articulatory structure of his own speech across multiple levels. This is a unified, rhythmic, and hierarchic organization of great precision which has been called 'self-synchrony'. Further, and surprisingly, the body of a listener moves in organizations of change which are precisely synchronized with the articulatory structure of a speaker's speech, and often with inanimate sounds as well. This has been called interactional synchrony or entrainment. It appears to be a universal characteristic of normal listener behavior and has been observed in many different cultures. It is also a basic characteristic of infant behavior and has been observed as early as twenty minutes after birth."

Condon then describes disruptions in this entrainment and self-synchrony as a fundamental trait of a host of prominent disorders:

"Sound-film microanalysis of various disorders such as autism, dyslexia, hyperactivity, Huntington's disease, Parkinson's disease, schizophrenia, and stuttering has revealed 'jumps' and 'jerks' in the body following sounds. This leads to the hypothesis of an abnormal *multiple entrainment* to sound." "Multiple" meaning the entrainments of various parts of the body are not synchronized and therefore independent. Though this seems like a diverse array of conditions to say the least, it is very striking that they all share the characteristic of multiple entrainment. Equally striking is that the abnormal entrainments always break into exactly four parts, four separate entrainments, firing at different rates.

One might be tempted to ask what body parts we use when listening to, say, Hamlet's soliloquy, but through years of studying slow-motion films, Condon concluded that "the body parts are not the unit of behavior. The organization of the relationships of change of the movements of the body parts (and speech) constitute the units of behavior." These relationships occur not just through time but among many layers moving simultaneously at different rates. This leads Condon to conclude that "the movements, gestures, and speech that one sees and hears when a person speaks can be interpreted as wave forms which are hierarchically organized."

I can add that this account of entrained movement connects to a hypothesis from Berthoz about how the brain views and enacts movement. Rather than see movements as a concatenation of muscle functions, which as roboticists could tell you is impossibly complex, the brain invokes phase shifts that automatically contract a series of muscles in a pattern. These phase shifts, and the way they are manipulated to account for external conditions, are best accounted for using matrix-based mathematical functions such as eigenvectors, tensors and Lyapunov equations. The brain itself, then, invokes movement in the body through complex waveforms, and modulates that movement through further waveform filters. This synchronizes perfectly with a Rhythmic Theory of Attending in which attention is directed in rhythmic cycles, as well as a theory of perceptual and motor learning in which these phase shift patterns are adopted by the motor centers of the brain based on information absorbed from moving in the world.

Turning back to the brain, Condon comes to the same conclusion about the relationship between entrainment and brain activity that Jones does: "If the hierarchic organization of the behavioral waves is synchronous with, or a reflection of, brain waves...it may suggest that the brain waves are operating together synchronously and hierarchically like the behavioral waves." In a final suggestion reminiscent of hypnosis, he adds, "it would also suggest that the brain wave organization of a listener may entrain with the structure of the incoming speech of a speaker." This entrainment with speech is so intuitive that in tests, given a soundless film of a speaker's body language and a copy of what is said, a judge can identify perfectly what words were said when.

Chapter 5, "Aspects of Rhythmic Structure in Speech Perception", is a shorter essay that contains two fascinating points. Using as his launching point early attempts to make speech-recognition software for computers, James G. Martin describes how his research

on how speech is produced, parsed, and comprehended led him to conclude that overarching rhythmic structures are essential for parsing continuous speech.

Computers have great difficulty assessing continuous speech, and to this day the successful programs for speech do not effectively comprehend naturally spoken speech, and only overcome major defects in understanding speech with raw processing power.

According to Martin, the biggest failure in the methodology of speech recognition software is the assumption that speech can be understood as a series of syllables concatenated together. On the contrary, he says, through a series of experiments he deduced that speech has "internal wholistic organization; the sequential elements are organized in terms of relative timing and not concatenation." New sounds are not simply appended to the previous ones; rather the syllables are all timed relative to each other within the range of "the breath group, a phrase-like unit, as the domain of the pattern." This is an exact parallel to Jones' claim that prior theorists of attention wrongfully saw "absolute time or sequence" instead of "relative time or rhythm" as the unit within which attention was functioning.

A good elucidation of this comes from the next chapter: "Concatenated accounts of rhythmic activity presume that patterning unfolds in a successive or 'left-to-right' series and limit its predictive power to adjacent elements. Control is peripheral as opposed to central. For Martin, the locus of relative timing is the *accented element* [emphasis mine]...Speech production is made based on a priori decisions made first to determine accented syllables which have high-information yields."

A huge amount of speech information, then, is encoded in the breath group as an "intact utterance". This unit has "an internal representation of the production constraints and temporal redundancy in the signal. Such constraints and temporal redundancy allow future pattern elements to be predictable in real time; early elements in a pattern generate moment-to-moment expectancies and hence an efficient perceptual device may be expected to have an anticipatory 'feedforward' component." After presenting some experimental data, he concludes, "Evidence that the rhythmic and segmental aspects of continuous speech are not perceived independently suggests that a future model of speech perception by humans should incorporate at some early stage a rhythmic expectancy component." Indeed, this is done in the chapter on reading comprehension, in which prosody is credited for being able to expand our cognitive powers by allowing us to tend to longer and more complex stretches of information, without taxing short-term memory, by using the clues of intonation to set up predictions about the structure of the content that is on its way.

Having established this global phenomenon of rhythmic expectancy, Martin describes one way this manifests itself on the local, or segmental, level, in an manner that particularly thwarts speech recognition software: the phenomenon of "anticipatory coarticulation." Martin's study of the physical production of phonemes shows that every sound pronounced is impacted by the sounds that come before and after. For example, "lip protrusion occurs two segments or earlier before the rounded vowel of 'sloop', but not

before the unrounded vowel of 'sleep'". The actual method of making the "s" in both cases was different, depending on the vowel that comes next, or as he puts it more generally, "every part of the continuous stimulus may be transformed by the context preceding and following it." Though our own predictive sense is trained enough that we adjust to these coarticulations without thought, it puts a huge amount of variety in how the same intended sound is actually pronounced, making it require huge amounts of processing power for a computer to guess from all the different ways a consonant or vowel might be pronounced in its various contexts.

Chapter 6, "Rhythm as a Factor in Mediated and Nonmediated Processes in Reading", the final chapter this review will consider, is closely related to its predecessor, moving from rhythmic aspects of speech perception to rhythmic processes used in reading; both are focused on how we process language in its various modalities.

The author, Madlyn L. Hanes, first defines her use throughout the paper of the terms "mediated" and "nonmediated": "Mediated refers to strategic processing necessary when difficulty arises...nonmediated, on the other hand, refers to fluent processing whereby comprehension is accumulating readily...Mediated and nonmediated states of processing are interdependent and interchange periodically in skilled reading."

"The principle linguistic correlate of rhythm is prosody," Hanes writes, which will probably make sense to the reader from, among other things, rhythmic analysis of poetry. "Prosody and its constituent features – stress, pitch, and juncture – accompany running speech, generating an intonation contour which serves to bound speech production into meaningful units. The boundary effect created by the prosodic features yields another linguistic correlate of rhythm which is critical to language reception, including reading...the ability to parse sentences into meaningful units be it clause, phrase, or sentence gestalts." This assertion is parallel to all that has come before, including the Rhythmic Theory of Attending and its claims about attentional energy, as well as the chapters on language: rhythm serves to help us understand language by giving us larger-scale units to help us put the meaning of sentences together.

This observation, articulated in so many ways in this book, turns out to bear directly on current ideas on parsing in reading comprehension: "parsing enables the processing of information, probably within as large a unit as possible, without losing meaning or taxing the short-term memory." As a fundamental reading strategy, "speech recoding, a type of inner speech, relies on prosody to echo parsed segments of the sentence when short-term memory is taxed." She elaborates: "correction strategies...use a morphosyntactic analysis...and have the effect of reconstructing and sustaining the accompanying prosodic contour. "

To put it simply, the prosody, or music, of the words is what allows us to organize speech in our minds into units large enough to make meaningful units comprehensible without taxing the short term memory, that is to say, it is an absolutely critical feature for higher-level language processing and necessary for fluent and speedy reading. "An information processing model of rhythm," she writes, "is hierarchical in structure rather than

concatenated," and therefore "the property of fluency in automatic reading is akin to the sense of timing in fluid, *nonlanguage motor* functioning (which is rhythm)." [Emphasis mine]

As students of the mind have expressed since at least William James, surely the brain uses, for the most part, the same faculty for both understanding speech and generating it, for to have a totally separate apparatus for each would be incredibly inefficient and inconsistency-prone. Similarly, if fluent reading shares properties and behaviors with fluid motor functioning, it's likely they share common areas of the brain to do their work, underlining more than ever the connections between our capacity for rhythmic movement and our higher cortical functions.

Difficulties in reading, Hanes asserts, prove the rule ever more strongly: "The absence of prosody from written text has been regarded as the root of reading difficulty in cases where the reader is unable to compensate for the lack of prosodic information through parsing skill, rendering the task of reading unwieldy and unnecessarily cumbersome. Parsing assigns the stretch or unit of processing, establishing syntactic constraints from which the able reader derives the semantic constraints and identifies (approximates) detail related to words and word-parts."

So how does this fit into the exercises in the Cognitive Eurhythmics curriculum? These studies, along with other research I cited, demonstrate that the first-level activity of our brains is the making of rhythmic predictions. The nervous system is constantly pelted with sensations; but to get from sensation to perception, as Berthoz puts it, the brain has to make a prediction, and as we understand from this book, a rhythmic prediction, which it then corrects and adjusts based on feedback from the environment. The mind isolates phenomena in the world through an act of synchronous attention, setting up a projection about the timing it will need to track the particular aspect of the world that it wants to follow, and correcting for error on the fly. It acts through setting up predictions about movement, enacting patterns of phase-shifts that enable the body's countless complexities to be accounted for by a few waveform variables, choosing a frame of reference from which to evaluate the movement, correcting for error via further waveform filters, and shifting the frame as needed to be evaluating the action from the most advantageous standpoint.

Though our scientists and philosophers are all heavily biased to see human beings as creatures of language and explicit reasoning, the fact is that the great majority of the problems the brain is solving at every moment are the ones characterized above, and this leaves an indelible stamp on all of our "higher functions" if we have an eye trained to see it. Recall of phrases and music reveals that our memory organizes sounds into rhythmic patterns even if there was no rhythm there to begin with. When we speak or read, we undergird our words with a complex use of musical nuance that enables our brains to process and organize a large and complex variety of material through the establishment of projections about the structure of that material, and redundancies in coarticulation. Our entire "higher functions" of language, reasoning, and memory all stand upon the framework of rhythmic prediction and projection.

Having established this mountain of concepts and theories, we are now in a position to provide a perfectly simple definition of the Cognitive Eurhythmics curriculum. Cognitive Eurhythmics requires a student to transfer analogous rhythmic activity from one modality to another - piano music, percussion, individual movement, group movement, voice, and language - in order to train, educate, tune up, synchronize and repair the rhythmically predictive apparatus that operates as the base of all human functioning.

Some implications need to be drawn out to avoid giving an impoverished picture. In particular, many individual functions need to be tuned up before this analogous transfer can be handled well by the student. The brain must possess a strong and subtle self-image of the body, a confidence in its movement, and a curiosity about using it to explore new things, or the many activities involving the body will not transfer through optimally. Emotionally, the use of subtleties in language and nuance requires both a feel for emotional expression, an ability to express it, and the courage and social confidence to express it. Psychically, there must be some capacity to analyze one's own behavior in discrete terms, such as remembering what one has just done, following a logical progression, being able to count the syllables in a word, etc., in order to hold onto the concepts at hand long enough to solve the puzzle of transferring them into a different modality. At the same time, Cognitive Eurhythmics exposes very clearly any weak points in the above attributes to the child, in the safe, fun atmosphere of creative games and physical challenges. When a child "gets it" in one modality, it's easy for the child to see how he or she is not getting the very same idea in another, and awareness often goes a long way to fixing the problem itself. For this reason I consistently describe Cognitive Eurhythmics as training the faculties of attention, coordination, socialization, and self-expression. What is harder to describe is that these attributes are trained to create a fundamental capacity of rhythmic prediction that is the universal substratum of all human functioning, and that if this base functioning can be optimized, it is hard to say whether any of the child's current limitations will be left standing, and it is not too much to suppose there will be nothing but freedom as far as the eye can see.