Response to Fernando Benadon

Matthew W. Butterfield

[1] What a pleasure it is to have one's work read with such care by as astute and thoughtful an observer as Fernando Benadon. In his commentary on my essay “The Power of Anacrusis,” Benadon has offered some provocative criticisms of my analysis of Herbie Hancock’s “Chameleon” (Benadon 2007). I would like to address these criticisms here.

[2] Benadon first questions the process by which I derived idealized timings of beats and beat subdivisions in the synthesized bass solo at the outset of “Chameleon.” At issue, specifically, is the timing of the syncopated bass note on the “uh” of beat one in each bar: how does one determine whether it arrives early or late? I took the actual and projected length of each measure and arrived at timing figures by a simple process of division (see my Table 1). Benadon proposes instead to derive such timings from more “local” criteria “to underscore the importance of ecological validity in the testing of timing-related hypotheses” [par. 6]. His figures corroborate my own, but on what he believes to be more solid ecological footing.

[3] I agree on the importance of ecological validity in deriving idealized timing figures for each beat and its subdivisions. I question, however, whether Benadon’s proposed metrics have more ecological validity than my use of the measure (or projected measure) for this purpose. An ecological approach to perception should model a subject’s perceptual engagement with his or her environment. It should take account of the structural attributes of the environment (with a special focus on invariant features), the perceptual and cognitive faculties of the subject, and his or her historical and cultural situatedness—all of which help the subject to understand what it is that is going on in a particular situation. If we are to derive ecologically valid idealized timing figures for “Chameleon,” then, we must find the metric that best models the perceptual strategies of the listener in engaging with the structural features of the rhythmic pattern at hand.

[4] Finding the measure unsuitable “for assessing the microtemporal placement of values as small as the sixteenth note” [par. 4], Benadon proposes two alternative methods based on more local criteria (see his example 2 and Table 1). The first employs the formula y–(x/2) to derive the ideal timing for the “uh” of beat one. In strictly mathematical terms, this makes sense, but do listeners actually make judgments about the duration of y (and the timing of the “uh”) on the basis of their experience of x? It is possible to judge y as half of x, of course, but only if one first experiences x as a salient duration that is
available for comparison. This can happen only if \( x \), upon its completion, realizes what Hasty calls “projective potential” (1997, 84–91). Thus in Example 1a, if we perceive projective potential \( P \), then \( x \) is available to serve as a measure for \( y \), and it should be relatively easy to anticipate the timing of the \( A_2 \) despite the implied hemiola, especially if the \( B_2 \) arrives at just the right moment to confirm the realization of projected potential \( P \). However, there will be a projective potential \( P \) and a definite duration \( x \) available for comparison only if we experience this passage in \( \frac{3}{8} \), as shown in example 1b. In \( \frac{4}{4} \), by contrast, a projective potential \( P \) can come only at the cost of the quarter-note tactus, as shown in example 1c. In this case, for there to be an \( x \), there must be a \( P \), and if there is a \( P \), there can be no projection \( Q-Q' \), and consequently no \( R \) for the duration of a quarter note.\(^2\) Instead, if \( P-P' \) is realized, \( R \) will be unhearable, and the \( B_2 \) on the “and” of beat two will sound more like a new beginning—like a new downbeat—than a syncopation anticipating a beat three. There would be nothing to indicate simple quadruple meter, and metric confusion would arise for the listener by the end of the bar.

\[ \text{[5]} \] How, then, do listeners experience meter in this passage, and what durations are relevant for their feeling of timing? Example 1d shows a more detailed analysis of this passage than I provided as example 14 in “The Power of Anacrusis.” Low-level projections in the anacrustic group preceding beat one generate eighth-note (\( P-P' \)) and quarter-note (\( Q-Q' \)) periodicities from the opening \( G_2 \) and \( A_2 \) respectively. \( B_2 \) emerges unambiguously as beat one—as a new beginning—because it marks the conclusion of a melodic process, it receives a dynamic accent, the interonset interval (IOI) preceding it is slightly elongated with respect to its predecessors, and the IOI following it is categorically longer still resulting in an agogic accent. These factors set the \( B_2 \) up to be heard as a new “dominant beginning” that makes all earlier beginnings and the projections stemming from them past and inactive for the becoming of new durations. This is not to say that the eighth-note (\( P-P' \)) and quarter-note (\( Q-Q' \)) projections in the anacrustic group are ineffectual and meaningless, however. Rather, projective potential \( R \) inherits from its predecessors both the eighth note and the quarter note as meaningful durations.\(^3\) Projective potential \( R \) is then a definite potential—i.e., its expected duration is stipulated by prior events—and its realization in the absence of a sounded beat two depends on our ability to perceive the quarter note as the tactus even at the outset of the performance. In other words, if there were no \( Q-Q' \), there could be no \( R \), and we would be unable to experience the \( A_2 \) as the “uh” of beat one—i.e., as a syncopation anticipating a second quarter-note beat.

\[ \text{[6]} \] I believe that this interpretation best models listener perception of this passage. Crucially, no salient duration \( x \) is available here—or in any other iteration of this passage—to serve as a measure in the calculation of the duration of \( y \). Consequently, though derived from more local criteria than the measure, Benadon’s formula \( y-(x/2) \) appears to me to lack ecological validity since it does not correspond in a coherent and meaningful way to the perceptual strategies of a listener engaged with this passage.

\[ \text{[7]} \] Benadon’s second proposed metric \( y-z \) strikes me as equally problematic. Here, the timing of the “uh” is gauged in relation to the “and” of beat two. This will give an accurate measure only if we can be sure that the “and” of beat two arrives “on time,” as it were—i.e., only if we can depend on the “and” of beat two as a fixed point of reference in each bar. My timing analysis does not usually bear this out, however. Instead, as my Tables 1 and 2 show, I find considerable fluctuation in the timing of notes played there (ranging from –37 ms to +23 ms in the first 12 bars alone), tending towards increasing delay after the entrance of the drums in m. 5. Moreover, I must question whether or not either a performer or listener would typically perceive the “uh” as the midpoint in a division between beat one and the “and” of two. Trying to hear it this way, I find, nullifies the effect of the syncopation anticipating beat three—it is simply contrary to the expressive effects of the pattern.

\[ \text{[8]} \] To derive ideal timing figures for each beat and beat division, we need a more or less fixed point of reference in each bar. Benadon’s \( y-z \) formula offers the “and” of two for this purpose. By contrast, in using the actual and projected length of the measure to derive my timing figures for mm. 1–4 of “Chameleon,” I have privileged beat one for two important reasons. First, empirical evidence suggests that listeners’ perceptual acuity is maximized around peaks of attentional energy—in other words, listeners find beats more perceptually salient and clearly defined than offbeats. In their study of the perception of interleaved melodies, for example, Dowling, Lung, and Herrbold (1987) found that listeners were more sensitive to melodic alterations on the beat than off. Similarly, Large and Palmer found that “internal coupling among oscillators [i.e., the coincidence of two or more pulse layers] improves tracking, particularly at temporally variable metrical levels. One reason for this improvement is that reduced variability at one metrical level allowed oscillations at that level to stabilize the tracking at
more variable metrical levels through coupling among oscillators” (2002, 31). Because beat one marks the coincidence of the most cyclic periodicities in each measure—it involves the internal coupling among the most oscillators—it emerges as the highest peak of attentional energy. What this means is that beat one is the most stable beat in each measure and the most logical point of reference for deriving ideal timing figures of other beats and beat divisions, especially in the absence of sounded beats two and three, as in “Chameleon.”

[9] Second, and probably as a result of its perceptual salience, both performers and listeners take beat one as the basis for metric projection—in other words, for their feeling of the becoming of specific durations. Low-level projections can emerge elsewhere in the bar, but they cannot override a dominant beginning on beat one without entailing metric confusion, as discussed in example 1c above. Projections arising from beat one are especially important in dance music, for dancers need a stable point of reference in each bar to make accurate judgments about the timing of their movements. One could argue that dancers orient themselves around backbeats in African-American groove-based musics, but there can be no perception of a backbeat without a prior perception of a beat one.

[10] For these reasons, I can only conclude that the measure and projected measure constitute the most ecologically valid metrics from which to derive idealized timing figures for the entire bar in the opening measures of “Chameleon.” Benadon’s proposed metrics, though indeed more local than the measure, involve metrically unstable or perceptually implausible durations to be used as a basis for the judgment of duration; neither of his proposed formulas seems to me to model accurately the listener’s engagement with this rhythm pattern. We could perhaps use the quarter-note duration of beat four in each bar as a more local means of gauging the quarter-note positions of beats two and three in the ensuing bar, or even simply average the durations of the three anacrustic eighth notes to determine subsequent eighth-note positions—as I did for bar 1 since there was no measure yet available for reference—but as Benadon rightly observes, Hancock tends to play each of these pick-ups at the end of each bar early. That leaves us with the measure itself as the least arbitrary, most stable, and, I believe, the most ecologically valid duration to serve as a metric for each bar.

[11] Benadon further questions whether the backbeat delays I identify are actually heard as being late. He offers five audio examples that compare the Hancock original with four alternate versions in which he has “corrected” the timing of the bass line’s “uh” in accordance with the four metrics offered in our respective methods. Benadon suspects that he “would not be able to guess in a blindfold test which of the above five versions of the ‘Chameleon’ bass line is the original, much less say which one has ‘a more relaxed quality that just feels more at ease’” [par. 9]. I share his skepticism—the discrepancies are indeed quite small, after all—though I do hear distinctly more punch in his audio example 1e, which involves the most radical “correction” of the original’s timing.

[12] Nevertheless, Benadon raises the important question of what the discrepancy threshold must be to generate tangible expressive effects. He is certainly right to assert that our ability “to measure minute imperfections (of rhythm, of intonation, of timbre) resulting from human production does not mean that they are automatically expressive in nature” [par. 15]. At issue is exactly how much discrepancy is required to constitute an expressive deviation. Benadon finds that “deviations of 20–40 ms seem large enough to be expressive, whereas deviations of about 5 ms are not only impossible to detect, but also devilishly tricky to pinpoint with confidence” [par. 14]. He does nevertheless seem to accept a minimum threshold of about 10 ms, and I am inclined to agree with him—especially with respect to the beat four backbeats of “Chameleon,” which I did admittedly over-interpret—though with some qualifications.

[13] To begin with, the higher discrepancy range of 20–40 ms might best apply to melody parts. Soloists have wide latitude in varying their timing—in fact, “Signifyin(g)” on the “time-line” is an important expressive resource for improvisers. Keil’s example (quoted by Benadon, par. 11) of a soloist whose “phrasing is consistently behind the pulse and then for one dramatic instant squarely on top of it” is very much to the point (Keil 1966, 346). No surprise, then, that in his own study of expressive microrhythm Benadon observes rather substantial behind-the-beat delays in soloists—in “I Hear a Rhapsody,” for example, Coltrane generally sits 50–80 ms back of the beat (Benadon 2006, 77). Likewise, Friberg and Sundström (2002) find mean downbeat delays among a variety of well known jazz soloists clustering in the 50–80 ms range at tempos between about 120–200 bpm (see their Figure 6).

[14] Rhythm section players have considerably less latitude, however. In groove-based music, tempo is not supposed to
fluctuate—at least not perceptibly. The instruments charged with sustaining the groove are expected to keep time as precisely as possible without rushing or dragging, and without allowing the discrepancy between them to vary too widely. They need to provide a stable background rhythm against which expressive timing in melody parts can operate. In the context of an ostinato groove pattern, if a timing deviation is large enough to rise to the level of conscious awareness, it has probably left the range of expressive nuance and will likely be interpreted by most listeners as a timing error. In other words, if the “uh” in the “Chameleon” bass line actually sounds either early or late, it probably sounds wrong.

[15] Thus whereas conscious perception of expressive timing is acceptable and often even desirable in a melody part, it is generally aesthetically undesirable in a groove pattern. It should not surprise us, then, to find that the deviations do tend to be smaller in the rhythm section than in melody lines. My timing figures for “Chameleon” bear this out—bass and drum onsets are seldom very far apart and neither deviates substantially from the ideal beat locations I have defined. Similarly, Friberg and Sundström 2002 shows an average discrepancy of less than 20 ms between the bass and ride cymbal onsets across a range of tempos on recordings of bassists Ron Carter, Robert Leslie Hurst III, and Gary Peacock; this contrasts markedly with the downbeat delays of soloists in their study, as discussed above. Rhythm section timing is just tighter—it has to be! Consequently, it is more difficult to register consciously, whether as an explicit timing discrepancy or as a more vague quality of feeling.

[16] There is another reason for this difficulty, as well. Gibsonian ecological theory assigns special importance to the invariant properties of our environment. We look for invariance to define an unchanging background against which change can be understood and acted upon. Thus in formulating his theory of meter based on Gibsonian ecological principles, London situates invariance in a more or less isochronic pulse, and identifies meter as a behavioral response to it: “In many contexts, we synchronize our attentional energies to the rhythms of the world around us. This synchronization is achieved by latching onto temporal invariants, that is, similar events that occur at regular intervals. Meter is a specifically musical instance of a more general perceptual facility of temporal attunement or entrainment” (2004, 25). Such entrainment of meter is not entirely passive, however:

   In musical contexts, metric attending involves both the discovery of temporal invariants in the music and the projection of temporal invariants onto the music. We actively seek and generate temporal structure through our attending behaviors. The way we attend to the present is strongly affected by our immediate past; once we have established a pattern of temporal attending we tend to maintain it in the face of surprises, noncongruent events, or even contradictory invariants. (25)

In other words, not only do we look for invariance, but we also seek to impose it on even a modestly fluctuating background for the purposes of enabling coherent action and exploration. This is the basis for categorical perception, in fact: we tend spontaneously to interpret the durations of successive events in the simplest temporal ratios congruent with the meter, even when their actual durations entail a more complex relationship (Clarke 1987). London’s “Many Meters Hypothesis” makes a similar claim: with experience, we come to interpret the metrical timing patterns characteristic of particular styles (e.g., swing) or performers (e.g., the Count Basie Orchestra) not as expressive deviations at all, but as unique meters imbued with a more general expressive quality or feel (2004, 153–156). In other words, we have a natural perceptual inclination to normalize timing deviations within an ostinato groove pattern, and this makes it quite difficult to perceive them as directly expressive, especially since they tend to be smaller than deviations played in the melody part(s).

[17] If timing deviations in a groove pattern are meaningful, then, conscious perception cannot be the decisive factor. As Benadon observes, however, this leaves us in the more tenuous position of defining qualities of feeling that supposedly emerge from these extremely subtle timing patterns. Discerning the effects in terms of feeling is challenging enough, especially since they are often subliminal. Finding language to describe them presents another difficulty, and achieving intersubjective corroboration may well be impossible—after all, our feelings are quite subjective! If we believe that they are real and meaningful—and I think Benadon would agree that they are—how do we bring the often subliminal effects of participatory discrepancies in a groove pattern to the surface of our conscious experience of the quality of feeling it generates?

[18] First, it takes repetition. Benadon’s assessment of his own audio examples suggests that he might be looking for a more
robust and immediate expressive payoff than PDs can offer in a typical groove pattern. Because they are so small, however, the effects of expressive timing are more of an emergent quality of a groove; i.e., they stem from repetition. Repetition allows us to test and confirm our subjective responses to the pattern. It gives us time to engage with a particular timing pattern, to learn our way around it, to begin to feel its most subtle nuances—time to live with it. We might compare this to the process of choosing a color to paint a room. You might settle initially on red, for example, but there are countless shades of red. How do you choose from among them? You have to study each one and allow some time for its affective qualities to emerge. You become sensitive to the subtle nuances of feeling each shade of color invokes only by repeatedly renewing your perception, by holding the color in your mind for a while. A momentary glance simply does not generate the same affective response. Perhaps, then, Benadon’s audio examples—and the first four measures of “Chameleon”—are too brief to generate much discernable feeling at the subsyntactical level? Perhaps the expressive variations would become more salient if the examples were extended further.

Second, sensitivity to the effects of expressive microtiming requires practice. Rhythm-section PDs are largely background phenomena. Their effects are extremely subtle and not nearly as tangible as events that happen on the syntactical level. Indeed, this is why I proposed that syntactical effects always override the subsyntactical. (This position is considerably at odds with Keil, incidentally, who privileges the power of participatory discrepancies and leaves little room for syntactical contributions to engendered feeling.) Learning to engage expressive timing as a conscious component of groove perception is a little like learning to taste fine wine. To novices, the only meaningful distinction is between red and white. With practice and experience, however, one learns to discern subtle differences in quality and taste and to develop a language with which to describe them. One learns to distinguish a chardonnay, for example, from a sauvignon blanc, and then to differentiate California chardonnays from French or Australian, etc. In making this analogy, I do not mean to claim more experience than or superior perceptual faculties to Benadon or anyone else. But experienced professional rhythm section players, who probably have the most sharply attuned acuity of perception for timing, do find these subtle discrepancies in timing quite meaningful, and some are quite gifted at translating the effects into language. Bassist Rufus Reid, for example, observes that “[t]here’s an edge I feel when I’m playing walking bass lines on top of the beat. It’s like if you are walking into the wind, you feel a certain resistance when your body is straight, but you feel a greater resistance if you lean into the wind” (Berliner 1994, 351).

If PDs are to constitute a meaningful aspect of our engagement with groove patterns, then, we must overcome our natural proclivity to normalize timing deviations, we must be prepared to accept small deviations as potentially consequential, and we must be willing to engage them through repetition and practice. My current research involves testing a variety of listening strategies against listening explicitly for timing. I find that when I ask listeners to tune in to a particular beat, it is next to impossible to make any claim about the timing unless the discrepancy is considerably exaggerated. In the case of a repeated pattern, however, and when listening for different qualities of feeling, it does become more possible and performance does seem to improve with practice, though I have not yet found consistently reproducible results.

Finally, Benadon finds that the eighth-note pickups played by the bass at the end of each bar sound “as metronomic as is humanly possible,” though they are played consistently early by about 20 ms, and that, “according to [my] line of thought and interpretation of magnitudes, that these pickups are less relaxed and more nervous” [par. 16]. The terms I employed to define magnitudes of discrepancy were borrowed from Michael Stewart (1987) via Prögler (1995), and were intended as guidelines for the expressive effects of backbeats, and not as rules delimiting explicit thresholds across all tempos. I had not actually thought of them in terms of pickup notes, where there might be a different range of expressive effects. Repp’s study, which I discussed in “The Power of Anacrusis,” does indicate that listener’s find it more difficult to identify a “duration increment” (i.e., a delay) than a “duration decrement” (i.e., an early onset) “in positions that typically exhibit expressive lengthening in performance,” and vice versa (1998, 131). It is possible, then, that Benadon hears these pickups as deadpan in their timing because an early onset is what their syntactical structure invokes. I do hear energy emerging from those pickups, however—a definite drive toward the ensuing downbeat. To some extent, this is surely tied to syntactical function—anacrusis generates energy, after all. But I do think it comes in part from the on-top timing of these notes. This is perhaps why Hancock’s recording of “Chameleon” is so much livelier than my synthetic rendering of the groove, in which the eighth-note pickups are indeed “as metronomic as is humanly possible.”
In conclusion, I would like to thank Fernando Benadon for his thoughtful and provocative commentary on my work, and I would also like to thank the editors of *MTO* for allowing me space to respond. Benadon’s own writings are essential reading for anyone interested in the topic of expressive microtiming, and his criticisms have raised important questions and enriched and clarified my own thinking on this topic. I find his call for the ecological validity of empirical data on timing an important methodological constraint, though I disagree on what specific metric best guarantees such validity. I do agree that the miniscule delays in Harvey Mason’s snare hits on beat 4 are probably too small to be consequential, but I think we should be prepared to accept a smaller threshold for the expressive effects of PDs in the rhythm section than in melody parts. We should look for these effects to emerge most palpably with repetition rather than in more limited contexts and to become more consciously available and expressively meaningful with practice.

Matthew W. Butterfield  
Franklin & Marshall College  
Department of Music  
P.O. Box 3003  
Lancaster, PA 17604  
matthew.butterfield@fandm.edu

**Works Cited**


### Footnotes

1. For a lucid summary of ecological theory and its applicability to music perception, see Clarke 2005, especially 17–47.  
   Return to text

2. Similarly, see the discussion of example 9.5 in Hasty 1997, 108–112.  
   Return to text

3. With respect to a given projection’s “inheritance,” see the discussion of metrical particularity in Hasty 1997, 148–167.  
   Return to text

4. Incidentally, Collier and Collier (2002) also derived their timing figures for the stop-time sections of Louis Armstrong’s “Potato Head Blues” and “Cornet Chop Suey” from the measure: “Because we felt that the measure was a fundamental reference unit, each time was computed as time elapsed since the beginning of its measure” (468).  
   Return to text

5. See, for example, Samuel Floyd (1995)’s analysis of Jelly Roll Morton’s “Black Bottom Stomp” (123).  
   Return to text

6. Prögler shows a wider variation in timing between bass and drums, but the more limited sample size of his study and the rather odd process by which he generated his figures call his data into question (1995, 32–46).  
   Return to text

### Copyright Statement

*Copyright © 2007 by the Society for Music Theory. All rights reserved.*

[1] Copyrights for individual items published in *Music Theory Online (MTO)* are held by their authors. Items appearing in MTO may be saved and stored in electronic or paper form, and may be shared among individuals for purposes of scholarly research or discussion, but may not be republished in any form, electronic or print, without prior, written permission from the author(s), and advance notification of the editors of MTO.

[2] Any redistributed form of items published in MTO must include the following information in a form appropriate to the medium in which the items are to appear:

This item appeared in *Music Theory Online* in [VOLUME #, ISSUE #] on [DAY/MONTH/YEAR]. It was authored by [FULL NAME, EMAIL ADDRESS], with whose written permission it is reprinted here.

[3] Libraries may archive issues of MTO in electronic or paper form for public access so long as each issue is stored in its entirety, and no access fee is charged. Exceptions to these requirements must be approved in writing by the editors of MTO, who will act in accordance with the decisions of the Society for Music Theory.

This document and all portions thereof are protected by U.S. and international copyright laws. Material contained herein may be copied and/or distributed for research purposes only.

Prepared by Brent Yorgason, Managing Editor and Stefanie Acevedo, Editorial Assistant