



## *Tactus* in Performance: Constraints and Possibilities

Peter Martens

NOTE: The examples for the (text-only) PDF version of this item are available online at:

<http://www.mtosmt.org/issues/mto.12.18.1/mto.12.18.1.martens.php>

KEYWORDS: meter, beat-finding, tactus, performance, computational modeling

ABSTRACT: A student string quartet was coached by the author to perform a set of seven musical excerpts twice, keeping the same tempo in each performance but feeling and expressing a different main beat (tactus) in each performance. Two empirical studies were conducted to determine the degree to which the quartet's intentions were communicated to study participants. In the initial study, participants viewed the full A/V performances and were asked to tap their dominant hand along with the main beat of the music. In a second study, the participants completed the same tapping task in response to either audio-only or video-only versions of the same performances. Finally, the audio and video of these performances were analyzed separately using the meter-finding computer model of Janata & Tomic (2008).

Overall, the quartet's intention significantly influenced participants' choice of tactus under the A/V and video-only conditions, but not under the audio-only conditions. Thus visual information is key to tactus communication even in an ostensibly sonic art form. In individual excerpts, however, aspects of metric structure appeared to constrain tactus choice, and objective visual and aural cues uncovered by computer analysis were not always matched by participant responses. Together, the results shed light on the extent to which this type of communication depends on the combination of tempo, a performance's aural and visual components, and musical structure.

*Received February 2012*

---

[1] **Example 1** reproduces the opening measures of a movement for string quartet by Haydn. If the reader were to auralize this excerpt in the neighborhood of the indicated tempo, I can make a few predictions about that reader's temporal orientation within the music. First, the reader will have become a *listener*, which is to say that the process of imagining the music activates similar brain activity as would processing actual acoustic events that correspond to the score (Hubbard 2010). Second, the listener is very likely to be organizing the excerpt's pitches and rhythms in relation to some consistent pulse that is represented in the notation, felt as a counting beat or tactus (Jones and Boltz 1989).<sup>(1)</sup> In this excerpt, the most comfortable pulses with respect to tempo are the quarter- and half-note pulses at the tempo indicated, 144 and 72 BPM, respectively. Researchers have explored the constraints that tempo places on pulse perception (Parncutt 1994; Drake and

Bertrand 2001); in the context of fully musical material, however, recent research has shown that tempo alone does not dictate a specific tactus within the fairly large window of 30–300 BPM (McKinney and Moelants 2006; Martens 2011). These later studies posit several factors that can direct a listener’s attention to one or another pulse within that window: the presence of subdividing pulses, genre-specific cues such as a backbeat pattern, years of musical training, and the specific instrument on which that training focused are just a few. If tactus-seeking listeners are busy attending to this noisy and opinionated crowd—tempo, rhythmic features, metric structure, and individual preferences based on prior experience—can the performer be heard above the din?

[2] The purpose of the exploratory research presented here is to investigate precisely this question: Can performers influence or direct listeners’ attention toward a specific tactus? The notion that tactus is communicated during performance is implicit in its very definition, and is encoded in our pedagogy and theories of beats and meter,<sup>(2)</sup> yet we have few detailed principles for performers. Will performers succeed simply based on a conscious decision, adeptly executed? Are often-neglected visual components of a viewed-and-heard performance important, even crucial, for this type of communication? In what follows, I will present results from two empirical studies and a computer model that speak to these issues, and I will set up each of these approaches immediately below in order to weave together their results later on.

### Experiment 1

[3] Seven excerpts of common practice music were arranged for an undergraduate string quartet at Texas Tech University School of Music, the titles, composers, and scores of which are given in [Appendix A](#). The pieces themselves, and a performance tempo for each excerpt, were chosen to allow for two pulses in the music to be plausible tactus choices. I will refer to the two tactus options generically as the “slower” and “faster” tactus throughout this article; across the seven excerpts, the slower tactus ranged from 42–76 BPM, the faster from 102–52 BPM. For example, the faster and slower tactus options in Example 1 were the two pulses already mentioned, at 144 and 72 BPM. Of the seven excerpts used in the study, four were in simple duple or quadruple meters and three were in simple triple meters; in the triple meter excerpts the triple ratio existed between the two tactus options.

[4] The quartet met for two rehearsals prior to a recording session in which they performed all of the excerpts. During the rehearsals I dictated a performance tempo for each piece and coached the quartet to express *either* the slower or faster tactus in each excerpt as their main beat, without discussing the other option at all. I had chosen a mixture of the slower and faster tactus options across the seven excerpts, and within each meter type. These performances were videorecorded. Four weeks after this initial recording session I asked the quartet to meet again, this time coaching them to express the other tactus in each excerpt, but at the same overall performance tempo. The hiatus was meant to allow their memory of the excerpts to fade, especially memories of specific strategies that they had employed to express the first tactus. These performances were also recorded using the same performance space and recording equipment. As one might imagine, the most difficult aspect of this second session was to match the performance tempo of each excerpt to that of the first recording, given that the quartet was feeling a different main beat. There was indeed some tempo variation between sessions, and two recordings were digitally manipulated so that the faster and slower tactus versions of each excerpt differed in tempo by no more than 5% of the faster tactus rate, as expressed in BPM. After this procedure, the two versions of each excerpt were judged by the author and two independent expert musicians to be indistinguishable in terms of tempo.<sup>(3)</sup>

[5] These videorecordings were viewed by 44 trained musicians, all of whom were undergraduate or graduate music students, or music faculty members at Texas Tech University School of Music. Directions to these individuals were simple: “tap along with what you feel is the main beat of the music with your dominant hand.” Each participant viewed and heard all seven excerpts only once, comprised of a mixture of slower and faster tactus versions across the seven excerpts, and a mixture of recordings from both recording sessions, presented in random order. During the experimental sessions, excerpt recordings were separated by 10-sec. distractor stimuli, which consisted of animated Lego video clips, free of regular visual periodicities, and accompanied by non-periodic music and speech. Here is one example of a recording followed by a distractor stimulus.

[6] Participants’ tapped responses were assessed using videorecordings of their dominant hand/arm/torso, and they were judged to have chosen a tactus successfully after executing eight consecutive taps in sync with one of the pulses in the target

excerpt. Slight discrepancies in tapping isochrony or phase were accepted as not detrimental to the goals of the study. The recording given as **Example 2** is one version of the excerpt shown in Example 1. I encourage the reader to listen again to Example 2 and tap along with the recording; then click on the “version” link below to see whether or not the quartet’s intention was communicated successfully.

[version](#)

## Experiment 2

[7] Experiment 2 was a followup study, using the same recordings as Experiment 1, but with the audio and video of the stimuli separated. 43 trained musicians, none of whom had participated in Experiment 1, responded to these recordings in the same fashion. These participants, however, responded to all 14 original stimuli, *hearing* one version of each excerpt and separately *viewing* the other version. Thus each participant did respond to both the slow and fast version of each excerpt, but in opposite modalities. An informal post-test survey confirmed that no participant had recognized the link between the slow video and fast audio of a given excerpt, or the fast video and slow audio, and thus no order effects would be expected. For example, having chosen a tactus in the faster tactus audio of excerpt #1, a participant would not have been predisposed toward that same tactus when viewing the slower tactus video.

## Computer model

[8] The audio and video of all 14 stimuli were processed by a reson filter-based computer model for meter identification demonstrated in [Tomic and Janata 2008](#). The model outputs shown in **Example 3** are visual representations of energy at periodicities that exist in these signals, processed in real time; the brighter the color, the greater the energy. Example 3a shows the model’s output or “energy profile” for an audio waveform containing a nominally isochronous pulse at 35 BPM. The horizontal lines represent energetic periodicities, with real time progressing along the x-axis from left to right and Hz along the y-axis, translated to BPM where useful. Notice that the strongest energy band in Figure 3a appears at 35 BPM rate itself, with other energy bands located at integer multiples of that pulse. Thus the model finds energy at pulse divisions that may not be audible in music, somewhat akin to the upper harmonics of a complex tone. In a fully metrical stimulus these beat divisions would be reinforced or damped according to the actual rhythms present. Finally, note that *groupings* of the actual pulse are not assumed.

[9] These design elements are evident in Examples 3b and c, which are the model’s assessment of meter in the audio signals of both performances of *Haydn’s op. 33, no. 3, IV*. At the most basic level, one can see differences that, while not stark, are at least consistent with the quartet’s intentions; note the increased energy at 36 BPM (a notated two-bar span) in the slower tactus audio. One might wonder that there is no noticeable energy band corresponding to the eighth note between 280 and 290 BPM in either profile. This is a result of a gain function built into the model, which amplifies frequencies within a central tempo window, having a peak amplification at 134 BPM. This design feature is meant to mimic listeners’ cognitive constraints relative to tempo (one can see that there is virtually no energy above 5 Hz/300 BPM in the profiles), and thus the model judges the eighth-note pulse in the excerpt from Haydn’s op. 33, no. 3 (see Example 1) as non-salient (see [Martens 2011](#), [London 2004](#), and [Parncutt 1994](#)).

[10] The Tomic and Janata computer model was developed to analyze audio input; the video of the string quartet performances was converted for its use as well, via the steps summarized below and illustrated visually in **Example 4**. Using ImageJ software, the recordings were first stripped of their audio; the video was split into individual frames at the standard cinematic rate of 30 frames per second (Example 4a); the frames were rendered black and white, and had their central portions blacked out to eliminate flickering pixels caused by the reflection of fluorescent lighting on the polished wood floor (Example 4b). Performances of the study excerpts resulted in between 500 and 800 frames each, and pairs of adjacent frames were analyzed for frame-to-frame pixel changes that expressed the minute changes in performers’ bodies, instruments, and bow positions. Examples 4b and c show a single pair of adjacent frames; the number of pixels that changed from one frame to the next was computed within ImageJ (shown visually in Example 4d). This process essentially converted physical movement into a string of integers; these strings of integers were then fed into the Tomic and Janata model. <sup>(4)</sup> Instead of the

model extracting periodicities based on amplitude peaks as it did while processing the audio waveforms, however, the model was slightly reconfigured to locate minima in the video integer strings. These minima capture moments of collective ictus in the quartet's movements, the moments of zero acceleration at which gestures, a down-bow stroke for instance, end and are about to be reversed, as when a bounced ball hits the floor.

[11] Returning to Examples 3d and e, one can see the model's assessment of energetic periodicities in the physical movements of the performers. Again, there are differences evident between slower and faster tactus versions, with more energy at the faster tactus rate of 140 BPM when it was intended. What these images convey is the sense that the quartet's two performances were objectively different, and were correlated with their intent. What they do not convey with any certainty, however, is that these differences were communicated successfully to listeners. This distinction will become clear below in the responses to individual excerpts.

### Overall Results

[12] **Figure 1** summarizes listener responses to the quartet's intended tactus. Figure 1a shows that listeners who successfully synchronized with the excerpts did choose the intended tactus significantly more often than the other tactus option ( $p < .001$ ). (5) (6) Figure 1b breaks down the data by condition, however, and presents a more complicated picture. Comparing the two blue bars, one can see that listeners tapped the slower tactus more often when it was intended than when it was not (48.2% vs. 29.7%), and likewise for the faster tactus (66.1% vs. 47%). Across all seven excerpts both of these comparisons are significant differences ( $p = .002$ ), but the skeptical reader may instead focus on the within-intent results. Most notably, when the quartet was intending the slow tactus, almost as many participants chose the fast tactus (48.2 vs. 47%). So while the quartet's intentions had *some* effect, is it worth talking about, or pursuing from a pedagogical standpoint? Do we see evidence here of global constraints on performers' intentions?

[13] We might easily assume that a bias toward the faster tactus was driving the overall results, and a straightforward explanation for the quartet's success in communicating the faster tactus—and its relative lack of success communicating the slower tactus—is that the faster tactus rates were simply more comfortable with respect to tempo. The argument would be that listeners simply found the faster tactus rates between 102 and 152 BPM more comfortable than the slower rates in the 42–76 BPM range, despite the quartet's intentions. This explanation would indeed buttress some long-standing assumptions that the most comfortable or most salient tempos exist between 100–120 BPM (see [London 2004](#) for summary and musical application), but as mentioned above, recent research has cast some doubt on the dictatorial role of tempo in choosing a specific tactus.

[14] **Figures 2a and b** show the same results as Figure 1b, but now divided into duple- and triple-meter excerpt groups. In the responses to the triple-meter excerpts (Figure 2a), the slower tactus was in fact chosen more often than the faster tactus when it was intended ( $p < .001$ ). The faster tactus was likewise chosen significantly more often when it was intended ( $p = .048$ ). This was not the case with the duple meter excerpts (Figure 2b), in response to which only 37.5% of participants agreed with the quartet's slower tactus intention. This was true even though the slower tactus rates in the triple meter excerpts were a good deal slower than the slower tactus rates in the duple excerpts; mean tactus rates are shown in the figure's legend. Since pulses in the 40s and low 50s BPM are less comfortable than pulses in the 60s or low 70s BPM, these results contradict the notion of a faster tactus bias, and indeed against any explanation that rests solely on tempo. To complete this picture, note that the mean faster tactus rates were very similar between duple and triple excerpts, 137 BPM and 132 BPM, respectively. Thus it was not the case that the faster tactus rates in the triple meter excerpts were uncomfortably fast relative to those in the duple excerpts, and thus drove participants to the slower tactus rates. Simply put, tempo cannot explain this contrast.

### Individual Excerpts

[15] What can explain participant responses more satisfactorily is a combination of structural features composed into the music and performance factors that the quartet adds. In what follows I will present the human responses to three individual excerpts with which the quartet garnered varying degrees of success in communicating their intentions in their full A/V performances. Results from the audio- and video-only conditions, together with the energy profiles, will aid in teasing out

performance details and in distinguishing between performance factors and factors that seem to be outside of the quartet's control. A summary of responses to all seven excerpts in all conditions is given at the end of [Appendix A](#).

### *Haydn Op. 50, No. 2, IV*

[16] The quartet's two versions of measures 1–16 of [Haydn op. 50, no. 2, IV](#) were met with similar responses. The top left chart in [Figure 3](#) shows participants' tactus choices in response to the full A/V version. There is a non-significant effect of intention evident in the 19% vs. 4.6% of participants who chose the slow tactus; even if this difference were statistically significant, however, it seems counterintuitive to label as a success less than one-fifth of the audience choosing the intended tactus.

[17] My first assumption when compiling these results was that the quartet, and I as their coach, had simply failed to execute the assigned task, that of communicating the slower tactus—here the half note at 55 BPM. This assumption was strengthened by the responses to the audio-only condition, from which the slower tactus is all but absent. I had to revise this opinion, however, after seeing what the computer model found in the audio of these recordings. The greatest amount of energy in the audio signal of both versions corresponds to a division of the faster tactus, the notated eighth note pulse at 220 BPM, and thus it is not surprising that the vast majority of participants chose the quarter note pulse at 110 BPM as tactus. I have previously found that, for many listeners, a “subdivision benefit” for tactus exists; a pulse will be perceived more easily and more strongly when its subdivision is also present ([Martens 2011](#)). A subdivision benefit seemed evident in this excerpt, where the vast majority of listeners chose as tactus a pulse that barely exists in the audio signal in and of itself.

[18] We might expect the energy profile for the slower tactus audio to be much the same, given the participants' fondness for the faster tactus. But, as seen in the figure, there is clearly more energy at the slower tactus rate in the quartet's slower tactus version than in the faster tactus version. Thus in the slower tactus version, the quartet was indeed creating greater dynamic stress at the half-note level, stresses that were absent in the faster tactus version. The quartet's efforts are also evident in the visual domain. Indeed, results from the video-only condition do show that significantly more viewers chose the slow tactus when it was intended than when it was not. In the faster tactus video, not only was the slower tactus utterly unattractive, but the *division* of the faster tactus attracted 30% of responses.<sup>(7)</sup> The video energy profiles for the two versions are similar, but the energy at the 220 BPM rate in the faster tactus version is noticeably greater and more consistent.

[19] What probably drove participants to the faster quarter note tactus when viewing the full A/V performance, then, even in the slower tactus intent, was its prominent division in the audio. A subdivision benefit seems to explain best why increased energy at the slower tactus level did not matter to any great extent. The consistent and prominent eighth note rhythms in this excerpt act here as a structural constraint on tactus, making it difficult for the quartet, despite its efforts, to communicate the half note as tactus. These results also suggest that what little increase in attention was given to the slower tactus in responses to the A/V condition was due to the quartet's gestural cues.

### *Caldara, “Sebben, crudele”*

[20] [Figure 4](#) gives the same complex of human responses and energy profiles for a triple-meter excerpt, an arrangement of [Caldara's “Sebben, crudele.”](#) In the score one can see that a quarter-note pulse dominates in the outer voices and is consistently divided in the inner voices. Based on the Haydn example just discussed, we might expect the quartet to encounter difficulty in communicating the slower tactus here as well, the dotted half note pulse at 46 BPM. In reality, however, 65% of responses did match the intention of the full A/V slower tactus version, while over 63% of responses matched faster tactus intention (both comparisons significant, as shown).

[21] What factors account for the quartet's success with this slower tactus? On the structural side of things, the eighth-note division of the faster tactus, so melodically prominent in Haydn's op. 50, no. 2, was in this excerpt relegated to a sewing-machine accompanimental pattern given to the inner voices, and not so prominent as to draw participants invariably to the faster tactus quarter note; the harmonic rhythm often matches the slower tactus as well. These composed-in features allow

the dotted half note to come across as tactus despite its slow tempo. On the performative side, the videorecordings tell the story perhaps better than the energy profiles. In the faster tactus version, notice the quarter-note pulse in the performers' bodies, evident even in the first violinist's anacrusic gesture.<sup>(8)</sup> In the slower tactus version, there is more pronounced movement at the dotted half pulse, including a clear dotted-half preparatory cue this time led by the cellist (who, after several miscues and botched beginnings of this excerpt during the recording session, was taking matters into his own hands).

[22] Although the audio energy profiles show some possible distinction between versions, they are quite noisy. The clearest difference between the two versions is in the video energy profiles, where we see strong consistent energy at the slower tactus rate when it was intended. This visual difference lines up with the overall results, but this basis for the quartet's overall success is not at all apparent in the audio-only results, and the video-only results were actually skewed opposite the quartet's intention. So with this excerpt we see a combination of gesture and structure-as-sound being effective in communicating the slower tactus in the A/V condition, whereas neither of these factors was effective separately.

### *Haydn, Op. 20, No. 5, II*

[23] Thus far I have posited that structure often acts as a large-scale constraint on performance in terms of the salience of different pulses, and that visual information in particular is only effective in the absence of strong audio cues. But this relationship can be complicated. One final example will show that gesture can actually work against the quartet's intentions, and that structure does not always hold sway in the absence of visual cues (see **Figure 5**). A third Haydn excerpt, this time from his *op. 20, no. 5 quartet, mvt. II*, is also in triple meter, and in the score one can see that the faster tactus (the quarter note) is rarely divided. The subdivision benefit hypothesis would predict that this lack of division decreases the quarter note's salience, and we can see evidence of this structural feature in the responses to the full A/V performance; the faster tactus at 146 BPM garnered not quite 43% of responses when it was intended.

[24] Yet with the video removed, one can see that the faster tactus *was* clearly communicated when intended (the audio-only condition). There is indeed more audio energy at the faster tactus rate when it was intended, and thus it seems that any avoidance of the faster tactus due to a lack of subdivision benefit is video-dependent. But the exact effect of video is difficult to determine; the responses to the video-only condition include no statistically significant comparisons, and these video energy profiles are among the very messiest from the seven excerpts. The video of the faster tactus version does contain a strong periodicity at the two-bar level (lacking, perhaps surprisingly, in the slower tactus version). By contrast, the slower tactus video does contain energy at the 146 BPM rate that is absent in the faster tactus video (but present in the audio). Further, other strong pulses related to the meter do appear in the video of both intent conditions (e.g. the dotted-quarter pulse in the faster tactus condition), but these were often weaker than, or difficult to distinguish from, other periodicities that had no direct relationship to the meter. The only exception seems to be the slower tactus version's half note hemiola, which attracted 30% of that version's viewers. Again, the combination of audio and video seems required to communicate the slower tactus, but a clear explanation is more difficult than in the Caldara excerpt.

[25] Let me summarize the factors that I see operating in these results as follows:

- Tempo is a broad factor in choosing tactus, but not a specific one; there is little evidence of a broad effect in these results since I had controlled for it via the stimulus selection and mandatory performance tempos. To an average performing group that is constantly choosing new repertoire and deciding on overall performance rates, however, tempo will frequently be a factor in communicating tactus.
- The performers' own limitations are a factor; the quartet's members were relatively young and inexperienced, and while the energy profiles show that they were attempting to follow my directives, an advanced level of technical and conceptual mastery is likely required to achieve reliably the sort of communication I have been discussing.
- The predisposition of audience members is a factor; we can never expect a completely unanimous response from an audience in the context of a formal experiment, much less in a concert hall. Communicating an intention to even a simple majority is a worthy achievement, and thus the significance tests have been used throughout the data analysis above in order to give meaning to what often looks like mediocre success on the part of the quartet.

- There is evidence that structure, especially in the form of a subdivision benefit, acts as at least a binary “allow-disallow” function for tactus choices.
- Gesture has an important role in communicating tactus, even with the complex set of constraints outlined in the previous four points.
- Much of what performers can do with musical time falls into the broad category of accent, but includes bodily and instrumental gestures that create visual accent.

[26] By way of conclusion, the overall responses to audio-only and video-only presentations of the excerpts are shown in **Figure 6**. In the audio-only results, no comparisons between intentions (slower tactus intended vs. faster tactus intended) are significant (Figure 6a). We can say, then, that the quartet’s intended tactus did not come across in the audio of their performances. By contrast, *every* between-intention comparison in the video-only results is significant (Figure 6b) and in broad accordance with the intention (i.e. the hemiolic responses are alternatives to the slower tactus, while the faster tactus divisions reflect increased energy at faster pulse rates).

[27] The results that I have focused on here suggest that, within the constraints of tempo and structure, (1) the intentions and resulting decisions of performers do play a role in communicating a musical beat, and (2) visual information is key to this communication—even in this ostensibly sonic art form. Although technology allows a significant percentage of our daily intake of music to be strictly aural, the increased ability to communicate via the visual domain may underlie the continued popularity of music videos, live music performances, and physical movement in response to music.<sup>(9)</sup> The challenge, or even charge, to performers and those with the responsibility of teaching performers, then, is to understand the structural and perceptual constraints involved in the communication of musical time, and to manipulate and work within those constraints in a conscious way. It is only to our and our students’ advantage to study and know individual musical structures in a detailed way, to understand their perceptual implications, and to use this knowledge to our advantage while making music.

---

**Peter Martens**  
**Texas Tech University**  
**School of Music**  
**Box 42033**  
**Lubbock, TX 79409**  
[peter.martens@ttu.edu](mailto:peter.martens@ttu.edu)

---

## Works Cited

- Clendinning, Jane Piper, Elizabeth West Marvin, and Joel Phillips. 2012. *The Musician’s Guide to Fundamentals*. New York: W.W. Norton & Co.
- Drake, Carolyn, and Daisy Bertrand. 2001. “The quest for universals in temporal processing in music.” In *The Biological Foundations of Music*, ed. R. Zatorre and I. Peretz. New York: New York Academy of Sciences.
- Honing, Henkjan. 2007. “Is expressive timing relational invariant under tempo transformation?” *Psychology of Music* 35, no. 2: 276–85.
- Houle, George. 1987. *Meter in Music, 1600–1800*. Bloomington, IN: Indiana University Press.
- Hubbard, Timothy L. 2010. “Auditory imagery: Empirical findings.” *Psychological Bulletin* 136, no. 2: 302–29.
- Huron, David. 2006. *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: The MIT Press.
- Jones, Mari Riess, and Marilyn Boltz. 1989. “Dynamic attending and responses to time.” *Psychological Review* 96, no. 3: 459–91.

- Lerdahl, Fred, and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge, Mass.: MIT Press.
- London, Justin. 2004. *Hearing in Time*. Oxford, UK: Oxford University Press.
- Martens, Peter. 2011. "The ambiguous tactus: Tempo, subdivision benefit, and three listener strategies." *Music Perception* 28, no. 5: 433–48.
- Levitin, Daniel J., and Vinod Menon. 2003. "Musical structure is processed in 'language' areas of the brain: a possible role for Brodmann Area 47 in temporal coherence." *NeuroImage* 20, no. 4: 2142–52.
- McKinney, Martin, and Dirk Moelants. 2006. "Ambiguity in tempo perception: What draws listeners to different metrical levels?" *Music Perception* 24, no. 2: 155–65.
- Parncutt, Richard. 1994. "A perceptual model of pulse salience and metrical accent in musical rhythms." *Music Perception* 11, no. 4: 409–64.
- Tomic, Stefan, and Petr Janata. 2008. "Beyond the beat: Modeling metric structure in music and performance." *Journal of the Acoustical Society of America* 124, no. 6: 4024–41.
- 

## Footnotes

1. Despite its frequent historical use as the metric level that *groups* the beat level, I will use the term *tactus* throughout to label the primary beat unit. This usage follows from the divorce of *Takt* from *tactus* in the late 17th century (see Houle 1987, 28–34) and corresponds with modern usage (e.g., Lerdahl and Jackendoff 1983, 21; and London 2004, 17; and Huron 2006, 176).  
[Return to text](#)
2. See, for example, the succinct definition of Lerdahl and Jackendoff 1983, 21, and the procedures "to determine the meter of a composition by ear" in Clendenning, Marvin, and Phillips 2012, 45–46.  
[Return to text](#)
3. Honing 2007 found that tempo manipulation negatively influenced listeners' judgments of naturalness in musical excerpts, suggesting that expressive timing does not scale with tempo. His altered excerpts, however, were changed by  $\pm 20\%$  of the original BPM, and those in the present were changed by much less, roughly 3% and 5.5% of the original faster tactus rate. No altered naturalness effect or concomitant effect on tactus choice would be expected from these relatively small tempo variations.  
[Return to text](#)
4. While this method of measuring physical movement may be crude compared to current motion capture technology, the overarching desire was to produce performances that were as ecological as possible for the players and viewers. One possible refinement would be to measure the movement of each player separately, especially if combined with responses that included eye-tracking data.  
[Return to text](#)
5. In this and other charts in this paper, tactus rates other than the faster and slower options are omitted unless germane to the immediate discussion. In such cases the percentages shown on the chart will not sum to 100. These "other" rates were always either divisions of the faster tactus or hemiolic pulses (e.g. a half note pulse chosen as tactus in response to a simple triple excerpt).  
[Return to text](#)
6. The p-values reported here and in Figures 2 and 6 derive from two-tailed t-tests for dependent samples; those shown in Figures 3–5 derive from a difference in proportions test for independent samples (z-test).



[Return to text](#)

7. The missing 20% of responses in this chart occurred at seemingly unlikely rate of 165 BPM, corresponding to what would be notated as tripleted quarter notes. This tactus rate is also faster than the faster tactus rate of 110 BPM, and thus these responses further point to strong visual periodicities at the fast end of the tempo range in this performance.

[Return to text](#)

8. This faster tactus version was one of the two digitally altered performances, uncorrected for pitch since participants heard only one or the other version; thus the pitch levels of the two versions differ slightly.

[Return to text](#)

9. The link between rhythm/meter and vision may have a neurological basis; [Levitin and Menon 2003](#) showed that a specific region of the brain active in attending to temporal aspects of musical structure is also active in processing American Sign Language.

[Return to text](#)

---

## Copyright Statement

**Copyright © 2012 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 Michael McClimon, Editorial Assistant*