Listener-Performer Synchronicity in Recorded Performances of Chopin's Mazurkas

MITCHELL OHRINER [1]
Shenandoah University

ABSTRACT: Studies of duration in music performance since the 1980s have focused on the role of duration in the communication of musical structure, in particular grouping and metric structure. This article posits a broader view of duration in performance; that performers choose durations to facilitate or inhibit listener-performer synchronicity. Listeners who maintain synchronicity are continually rewarded for making accurate predictions of when events will happen, and thus the loss of synchronicity might hazard feelings of frustration or isolation. Performers, by choosing durations that are difficult to predict, can therefore enlist these feelings in dynamic narratives of synchronicity that augment trajectories in the score. The article explores two such narratives, found in two idiosyncratic renditions of passages of Chopin's mazurkas and echoed by seventeen participants' attempts to synchronize with them. By ascribing significance not only to the choices of performers and composers but also to how individual listeners attend to those choices, the article aims to widen the circle of agencies ascribed with the ability to affect the meaning of musical works.

Submitted 2013 September 6; accepted 2014 April 3.

KEYWORDS: *entrainment, sensorimotor synchronization, expressive timing, mazurkas, Chopin*

I. INTRODUCTION: *TEMPO RUBATO* AS A SUBSTRATE FOR SYNCHRONICITY

> It is still possible to speak about musical materials—if not 'works themselves'—in relation to matters of value, authenticity, meaning and effect. To do so, however, requires us to identify not what the work, as a bounded object, means, or does in itself, but rather, how it comes to be identified by others who refer to or attend to (and this includes non-discursive, corporeal forms of attention) its various properties so as to construct its symbolic, emotive or corporeal force. Such a strategy ensures that interpretation of music is not used as a resource for, but rather a topic of, investigation (DeNora 2000, p. 30).

IN the past thirty years the measurement of duration in performance has burgeoned as an area of interest across the scholarly music disciplines. Earlier studies (e.g., Bengtsson & Gabrielsson, 1983) suggested that in some repertoires of solo performance the durational variance of "more or less equal" events, to use the characterization of Jonathan Kramer (1988, p. 7), is greater than was previously assumed. In the 1980s and 1990s, Eric Clarke (1993), Neil Todd (1985, 1989), and Bruno Repp (1992) explained these durations as attempts on the part of performers to express aspects of musical structure, primarily grouping structure and meter. Hence this phenomenon is often called "expressive timing."[2] The aspects of structure thought to be conveyed through expressive timing are specific to levels of time-span organization (Lerdahl & Jackendoff, 1983, pp. 124–145): phrases on the order of eight measures tend to feature initial accelerations and terminal decelerations when the durations of measures are observed while beats within measures are variably timed in response to meter, changes in melodic contour, and local ornamentation.[3] More recently, music theorists have drawn on the expressive timing literature to document the practices of idiosyncratic performers rather than features of "works themselves" (e.g., Clarke, 1998; Dodson, 2008 & 2009; Benadon, 2009; Ohriner, 2011 & 2012; Gingras, Asselin & McAdams, 2013).

Another function of durational variability in performance—one more commonly encountered in current music psychology—relates to the role of timing variation in the maintenance of synchronicity between a performer and an attending listener. Repp (2005) terms this process of maintaining synchronicity "sensorimotor synchronization" [hereafter SMS].[4] SMS pertains the phenomenon of entrainment, which Martin Clayton (2012, p.49) defines as the interaction of multiple "independent rhythmical systems" (see also Phillips-Silver, Aktipis & Bryant, 2010; Overy & Molnar-Szakacs, 2009). These rhythmical systems may be different zones of an individual's body, different individuals performing in a group, or different groups of individuals interacting together (Lucas, Clayton & Leante, 2011, p. 51). In this article, I focus on the entrainment and synchronization between a recorded performance and a listener engaging in what Marie Riess Jones and Marilyn Boltz (1989) call analytic attending, that is, the low-level attending to events on smaller time scales that are difficult to predict. Sensorimotor synchronization is usually studied by having a participant tap along to a stimulus—most often a periodic (i.e., metronomic) stimulus—and monitoring the asynchronies between the perceived stimulus and the produced taps.[5] Although studies of SMS in "expressive" musical contexts are limited, Repp (2002) has shown listeners draw on common timing strategies in order to plan their attending to performed music and hence their ability to attend to performances of tonal music shaped through conventional means (e.g., with phrase-ending decelerations) is greater than their ability to attend to performances that defy these conventions (e.g., a timing pattern "phase shifted" within the meter or a random timing pattern).[6]

If correspondences between musical material and performed timing enable listeners to attend to performed music, then presumably performers might choose durations that catalyze synchronicity or, alternatively, elicit asynchrony. In this sense duration in performance is expressive not only of aspects of musical structure, but also of the dynamic relationship between a listener and a performer. It is this state of listener-performer synchronicity (and its absence) that I detail in this article. After describing how listeners are thought to attend to expressively timed music in the second section and what it might mean to lose synchronicity, the remainder of the article describes an experiment in which participants attempted to tap along to two passages from Frederic Chopin's mazurkas, a genre known for extensive tempo fluctuation. The third section describes a novel method for aligning many series of produced taps to a series of performed onsets where the number of taps and onsets may differ. The fourth section focuses on participants' entrainment to two idiosyncratic renditions of two passages in the mazurkas that significantly defy tonal and temporal expectations. By highlighting asynchronies between specific performers and specific listeners, I hope to demonstrate one way in which musical meaning can arise at the intersection of a work, a rendition, and a particular listening experience.

II. VALENCE AND MECHANICS OF EXPRESSIVE ASYNCHRONICITY

Synchronicity, or more specifically the ability to synchronize with a variably timed signal, confers social, cognitive, and emotional benefits (Overy & Molnar-Szakacs, 2009). Many social interactions, including some that extend survival benefits, are enhanced by our ability to coordinate action with others (Wiltermuth & Heath, 2009). Because making music (actively or passively) requires such coordination, musical experiences tend to increase empathetic behavior (Kirscher & Tomasello, 2009), which in turn increases self-assessments of interpersonal relatedness and well being (McPherson, Davidson, & Faulkner, 2012, p. 87). Conversely, an inability to synchronize with a musical performance might undermine this sense of cohesion. As David Huron (2006) writes:

> When an event happens at an expected moment in time, the *prediction response* [original emphasis] is positively valenced. This provides a positive reward for the heuristic used in the prediction and reinforces the use of such a heuristic in making future predictions (p. 184).

This positive valence originates in the effective deployment of attentional and muscular resources. These resources are marshaled periodically in the attempt to coordinate one's action with an actor in the environment (Large & Jones, 1999). An error in the prediction of the timing of the external stimulus may result in unnecessary and potentially taxing maintenance of this state of arousal (or, worse, a failure of the action due to poor preparation). Synchronization minimizes such errors by efficiently summoning attentional resources only when necessary (Large, 2008, p. 216). The emotional benefit of synchronization can thus be understood as an incentive to make accurate temporal predictions, while an emotional penalty may ensue for poor predictions.

Importantly, the emotional implications of asynchrony need not be negative. Asynchrony can surprise and delight as easily as it can frustrate. In collaborative music making (and much passive listening), participants are trying to attend to others and thus periods of asynchronicity hazard feelings of frustration and isolation, but there is also a mode of listening (and perhaps performing) in which synchronicity is not an aim and asynchronicity is not emotionally detrimental. In this sense, asynchronies between performers and listeners are akin to the "participatory discrepancies" described by Keil (1987) and have a similarly wide field of meaning depending on context and style. To cite one possible example, listeners accustomed to the complex phase relationships of a jazz rhythm section may be unperturbed when their own attempts at synchronicity are similarly out-of-phase. The emotional consequences of failing to accurately attend to a performance thus relate to one's ambitions for synchronicity and tolerances for asynchrony.

The ability to entrain to a periodic stimulus is well documented. But how are listeners able to successfully predict event onsets when beat durations vary extensively, as in a repertoire like that of Chopin? And conversely, what are the possible causes of poor prediction? For Justin London (2004), temporal prediction arises from a listener's familiarity with the metrical timing pattern particular to the genre of the performance, an idea he terms the "Many Meters Hypothesis" (pp. 142–160). Under this hypothesis, listeners might have different strategies for predicting onsets in a waltz and a minuet—both triple meters—perhaps by more extensively lengthening the second beat of the waltz. The notion that meter affects timing patterns has a long history of empirical support, including Bengtsson and Gabrielsson's demonstration of a unique timing pattern for the Viennese waltz (1977, p. 34) and Sloboda's demonstration of the role of duration in conveying potentially ambiguous meters (1985).[7]

Bruno Repp (2002, p. 252) presents a different explanation of temporal prediction, one distinct from generic conventions or specific musical content. Imagine four beat onsets (O_1, O_2, O_3, and O_4) delimiting three inter-beat intervals (IBI_1, IBI_2, and IBI_3, see Figure 1). Next, imagine a listener trying to tap to these beats, producing four taps (T_1, T_2, T_3, and T_4), which themselves have three inter-tap intervals (ITI_1, ITI_2, and ITI_3). Repp maintains that listeners will assume that $IBI_2=IBI_1$, and thus listeners will place T_3 at O_2+IBI_1.[8] But if IBI_2 is less than IBI_1, T_3 will be late. Going forward, T_4 will be placed at T_3+IBI_2. The primary prediction that arises from this model is that the value of ITI_n will be less like the value of IBI_n and more like the value of IBI_{n-1}. Repp, following the earlier work of Vorberg and Wing (1996) terms this the "lag-1 cross-correlation" of attending. This idea of lag-1 correlation receives support in Repp's study, but the notion may not address synchronicity in the more ecologically rich context of complete musical performances, wherein listeners can draw on their awareness of the relation between musical content and duration. If one hears a lengthening and relates the lengthening to, say, a melodic apex or metric position, it does not follow that one would map that duration onto a subsequent event that is not a melodic apex or is in a different metric position.

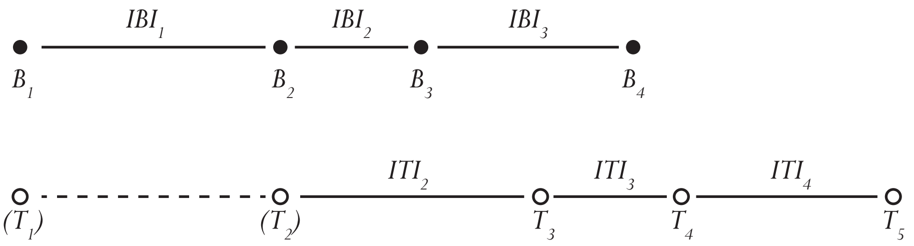

Fig. 1. Repp's model of "lag-1 cross-correlation" of attending (2002, p. 252). $ITI_2 = IBI_1$, $ITI_3 = IBI_2$, etc.

Simon Dixon, Werner Goebl, and Emilios Cambouropoulos (2006) offer another explanation for listener-performer asynchrony. As they demonstrate, timing patterns are better predicted with repeated hearings. As listeners more accurately synchronize with specific performances, ITIs correlate more highly with a smoothed average of IBIs rather than the IBIs themselves. Thus Dixon *et al.* suggest that in tapping to performed music listeners are not trying to synchronize with the performer, but rather to produce as steady a tapping sequence as possible given the perturbations of the performance.[9]

I will evaluate the performance of these models with the data generated in the present article below, but however asynchronies come about, certain perceptual limits should be considered in their interpretation. If an asynchrony is imperceptible to a listener, it should not be assigned the negative valence Huron describes. Most studies of the limits of rhythmic perception focus on the degree to which unequal durations may be perceived as equal (e.g., Friberg & Sundberg, 1995; Madison & Merker, 2002) or the degree to which non-aligned onsets are perceived as aligned (e.g., Vernon, 1937; Rasch, 1979; Palmer, 1996). In monitoring alignment between an attending listener and a performer, the relevant question is the extent to which a listener can perceive a latency between an action in the tactile modality (the tapping) and an expected percept in the auditory modality (the onset). Musicians are accustomed to latencies associated with their instruments: for example, the latency between striking a key and the hammer striking the string(s) of a piano ranges from 30–100ms, depending on velocity (Askenfelt & Jansson, 1990). Presumably, they could excuse similar asynchronies in tapping. Through a very clever experimental setup involving Max Mathew's Radio Baton (a drum mallet whose location in three-dimensional space can be tracked through radio signals), Daniel Levitin *et al.* (2000) suggest a tolerance for inter-modal (i.e., audio-visual) latency between −25 and 42 milliseconds with a fair amount of interpersonal difference (p. 328).[10] While detecting latency between a sound deriving from an action one has produced and the sound derived from the predicted action of another is not the same thing, the study of Levitin *et al.* (2000) most closely resembles the question of perceived asynchrony raised in tapping along to performed music. In the plots that follow in subsequent sections of this article, I include a grey box delimiting the just-noticeable latency threshold of Levitin *et al.* (2000).

As the perceptual threshold posited in Levitin, MacLean, Mathews, Chu, and Jensen (2000) suggests, taps that fall before as opposed to after an event should not be interpreted as equal with respect to their emotional valence. Repp (1998) has found that listeners frequently underestimate decelerations (p. 283). In the common case of a phrase-final deceleration, many participants will, like the performers, decelerate their rate of tapping. But because they underestimate the amount of deceleration, taps will still be early. Yet if synchronizing to a performance means predicting the relationship between aspects of musical structure (e.g., phrase endings) and performed duration, then the fact that both participants and performers are decelerating would suggest that they are in concordance as to what that relationship should be. Therefore, asynchronies arising from underestimations of gradual decelerations have a different affective character than those arising from a difference in the basic construal of the material-timing relationship.

Late taps should be distinguished from early taps for another reason as well. The motor program of a tap relates not only to the predicted onset of the next beat but also the one after that, in so far as the velocity of the rebound from the tap is planned to be conducive to the following tap. If a tap is early, the error is not apparent until the rebound is already underway and can be addressed by slowing the rebound or pausing at its apex. If a tap is late, the error is apparent before the tap is even executed and an increase in energy is necessary to hasten the rebound for the following tap, which, without correction, will probably also be late. And to the extent that an unnecessary expense of energy has an affective consequence (Huron 2006, p. 10), it is reasonable to assume that tapping late has a different and more severe affective consequence than tapping early.

The models proposed by Repp (1998) and Dixon, Goebl and Cambouropoulos (2006) are valuable for generalizing how listeners attend to music with variable beat durations. They are less adept at predicting how listeners will engage with specific musical content or a specific rendition of that content. The experiment described in the next two sections was designed to document this type of specificity by relating the asynchronies of a listener attending to a performance to the structural and expressive-durational details of that performance. By documenting this kind of listener-performer synchronicity in the context of a specific composed work, I hope to demonstrate one way in which music attains DeNora's (2000) "symbolic, emotive, or corporeal force" by those who attend to it.

III. METHODOLOGY

Experimental Protocol

During a one-hour session, seventeen undergraduate conservatory students granted informed consent, filled out a demographic questionnaire, and tapped to three renditions of three excerpts from Chopin's mazurkas, two of which are discussed below.[11] The participants were specifically instructed to "tap to each of the three beats in each measure of each excerpt." They were not presented with a musical score and no participant professed prior familiarity with any of the excerpts. The order of renditions within the excerpts was counterbalanced through a Latin square procedure.[12] Each excerpt ranged in length from 20–45s, and participants tapped on a MIDI keyboard running through the Logic digital audio workstation and saved their taps as MIDI files.[13] They heard each excerpt nine times in a row, separated by three seconds, and were instructed that all the excerpts they would hear would be mazurkas, all would be in a 3/4 meter, and that they should tap to each of the three beats in each measure.[14]

Selection of Pieces and Renditions

In selecting stimuli for the study of listener-performer synchronicity I had four goals: (1) to select pieces that normally exhibit high levels of beat-to-beat tempo variability, (2) to focus on moments of those pieces that feature the denial of tonal or temporal expectations, (3) to identify idiosyncratic renditions of those spans, and (4) to complement those renditions with two others that are more normative for the sake of comparison.[15]

Two passages from Chopin's mazurkas were selected, the opening sixteen measures of the Mazurka in G Major, Op. 50, No. 1, and a passage from the middle of the Mazurka in C-sharp Minor, Op. 63, No. 3. The particular tonal expectations and stereotypical temporal features of performances of these pieces are discussed in Section IV below. As a genre, the mazurkas were chosen because they are widely recognized for their profound tempo variability in performance. Although notated in 3/4, many performers of the mazurkas lengthen the second beat extensively, often at the expense of a shorter beat three. Indeed, Chopin's own lengthening of beat two was reportedly so extensive that several contemporaneous listeners mistook the meter for 2/4 or 4/4, a reaction that baffled the composer (Eigeldinger & Shohet, 1986, pp. 72–3).[16] Besides tempo variation, Grosche, Müller, and Sapp (2010) highlight other challenges presented to those attending to mazurkas, including highly ornamented melodies, soft repetitions of harmonies on consecutive weak beats, and beats lacking events. Before describing the ways in which participants attended to the two chosen mazurka excerpts, I will present a method for documenting participant entrainment that informs the later analyses.

Measuring and Visualizing Listener-Performer Asynchrony

VISUALIZING ASYNCHRONY OF A SINGLE PARTICIPANT

When participants tap to unfamiliar, expressively timed music, there can be stark differences between the number of taps and beat onsets. **Fig. 2**a–e demonstrates the method employed here for documenting listener-performer asynchrony even if the number of taps does not equal the number of beats. **Fig. 2**a shows the expressive-timing profile of four measures of Roberto Poli's rendition of Op. 50, No. 1 by plotting each onset horizontally where it occurs in the course of the rendition (one inch equals 1.14 seconds; this passage is discussed in greater detail below). **Fig. 2**b shows Participant Q's taps to this music in each of the nine trials, though in no individual trial is the number of taps the same as the number of beat onsets. Midway through Trial 3 the duration between taps doubles, suggesting a tap is missing (see "?"). At the same time, the column of taps before m. 9_2 (read: "measure nine, beat two") strongly suggests that the participant repeats an action because of his or her understanding of the temporal structure of the rendition.[17] Because of the difference in cardinality between trials and the occasional gaps in recorded taps, it is unwise to assume any particular relationship between the temporal structure of the rendition and that of the taps produced in response to it. Rather, I seek clusters of taps like the one before m. 9_2 in Figure 2b, allowing that (1) participants will not place a tap in each cluster in each trial and (2) participants may produce extra taps that do not align well with any cluster.

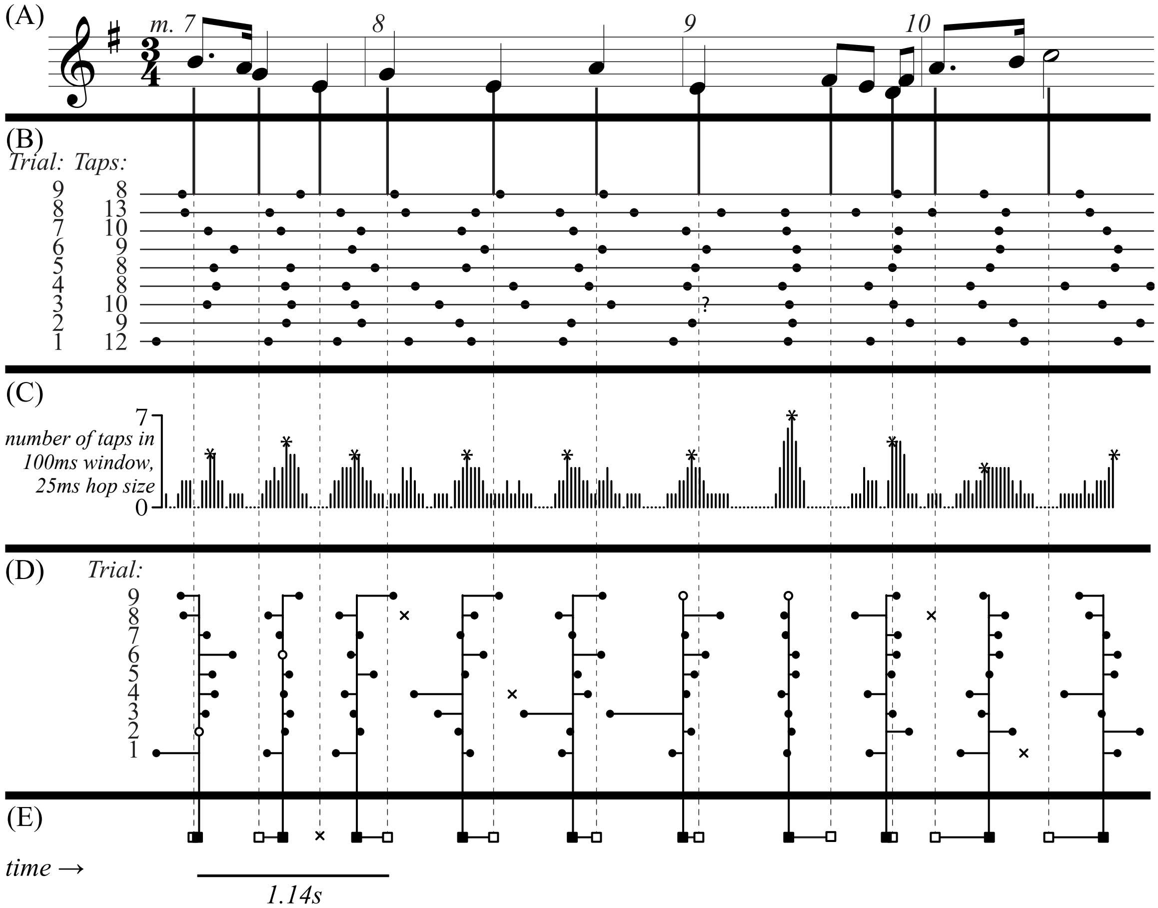

Fig. 2. Taps produced by Participant Q while listening to Roberto Poli's performance of Op. 50, No. 1 nine times. **(A)** Poli's expressive profile, shown by altering the notation such that one horizontal inch equals 1.14 seconds. **(B)** The taps in the nine trials, aligned with the onsets of events in (A), with number of taps produced in each trial indexed at left. **(C)** A demonstration of the clustering method. The y-axis shows a count of how many taps occur within a moving window. Inflection points representing the centers of clusters with more than three taps are marked by an "*." **(D)** Taps from (B) are aligned with the marked clusters of (C). "Extra taps" marked "x." Clusters omitted from a trial marked "o." **(E)** Alignment between clusters of taps (solid lines leading to closed squares) and beat onsets (dashed lines leading to open squares).

Broadly speaking, the clustering method proceeds in two steps: first, I try to uncover a temporal structure of the taps themselves, without reference to when or whether events occur in the music. Then, if such a structure exists, I try to relate it to the temporal structure of the music. By "the temporal structure of taps themselves," I mean clusters of taps that are narrow in duration and are represented in many trials like the one before m. 9_2. To find a participant's clusters of taps, I begin by counting how many taps occur within a window of 100ms, overlapping by 25ms (see Figure 2c); I then run those values through a moving average of three windows. Next, I take the higher inflection points of the values as centers of clusters, discarding those not represented in more than one-third of the total number of trials.[18] These inflection points are marked with an "*" in Figure 2c.[19]

Once a series of clusters has been identified for each participant, I set about aligning the taps of each trial with the series of clusters. I do so with the technique of Dynamic Time Warping (DTW), first described by Sakoe and Chiba (1978) in work on speech recognition, using implementations in the R statistical computing environment written by Giorgino (2009) and Tormene *et al.* (2008). The technique takes two vectors possibly of different lengths (in this case the vector of clusters and, separately, each trial of taps) and creates an alignment that minimizes the average distance between them by doubling some values in both vectors. In this case, "distance" is the average asynchrony between aligned clusters and taps.

Often, in order to minimize the distance between the taps and clusters, the Dynamic Time Warping algorithm will align multiple taps with a single cluster or vice versa; the algorithm may even align one tap with two clusters, one of which is also aligned with another tap. While minimizing distance, this aspect of DTW posits a poor interpretation of synchronicity: if there are two taps near a single cluster, only one should be interpreted as an attempt to tap "with" the cluster; the other is extra or errant. And if there are two clusters near a single tap, the tap should not be aligned with both. Rather, one of those clusters was not tapped to at all in the trial. Therefore, if one tap aligns with multiple clusters, all clusters but the one with least asynchrony are deemed "missing" in that trial and the asynchrony between those clusters and that trial is recorded as NA. On the other hand, if multiple taps align with a single cluster, all taps but the one with the least asynchrony are deemed "extra" and stored separately. As a final step, I go back through the clusters and asynchronies and move the cluster centers so that the mean asynchrony of taps aligned with a cluster is 0.[20] Figure 2d shows the result of this process, with taps aligned to the clusters identified in Figure 2c and an "x" placed at each tap that was not aligned to a cluster. The symbol "o" marks a cluster center in a trial with no tap aligning with that cluster.

Once the taps have been grouped into clusters and the asynchronies between taps and clusters have been recorded, the second step of the alignment proceeds by aligning the clusters of taps to the onsets of beats in the rendition. This step, visualized in Figure 2e, uses the same altered DTW algorithm as the process that produced Figure 2d, and records the asynchronies between clusters of taps and beat onsets, noting that some beat onsets do not have associated clusters of taps (one such occurrence is noted with an "x" in Figure 2e), or that a cluster of taps has no associated beat onset and is considered "extra."[21] (Participant Q produced no extra tap clusters in this passage). One common source for extra clusters is a beat not articulated in the musical surface. But extra clusters of taps also arise when participants are so early for an event that they tap again in an attempt to regain synchronicity. Where the distance between the open and closed squares is small, Participant Q and Poli are robustly entrained; when it is large, they are, in a sense, not making music together.

VISUALIZING ASYNCHRONIES OF MULTIPLE PARTICIPANTS

While Figure 2 demonstrates the varying extent of Participant Q's synchronicity to Poli's rendition of Op. 50, No. 1, it is not up to the task of documenting how many participants entrain to a certain rendition.
Fig. 3 shows the varying entrainment of all twelve participants who tapped to Poli's rendition. Specifically, the figure shows asynchronies between clusters of taps and beat onsets among the participants, identified by letter. Recall that in Figure 2b, a very tight cluster of taps across eight trials before m. 9_2 produced the greatest distance between open and closed squares in Figure 2e. Correspondingly, the value of m. 9_2 for Participant Q in Figure 3 is large and negative. At the right edge of
Fig. 3, six participants have an "x" above their cluster asynchrony value, or below if the value is negative. These signify that those participants produced an extra cluster of taps not aligned with any onset *before* m. 11_1. This is to be expected, since there is no event marking the onset of m. 10_3. The other extra tap cluster in the figure, produced by Participant G before m. 9_2, indicates a severe lack of entrainment at this moment. Because the downbeat is so long, occupying nearly half the measure, this participant, in every trial, tapped to what he or she thought was beat two, realized the error, tapped again, and was still early! I will use illustrations like
Fig. 3 to explore the entrainment of participants at moments of denied tonal and temporal expectations in the analytical discussion to follow.

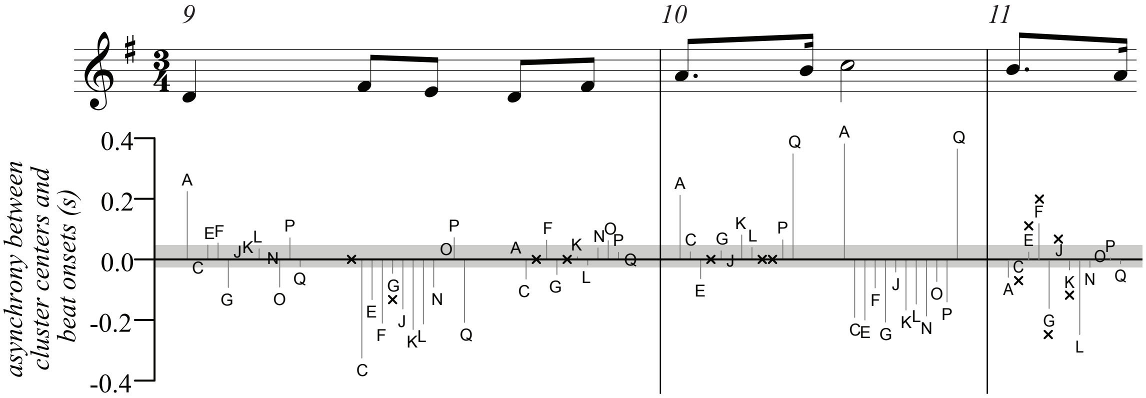

Fig. 3. Entrainment of twelve participants to Poli's rendition of Op. 50, No. 1, mm. 9–11$_1$, arbitrarily lettered from A–Q.[22] Asynchronies between clusters of taps and beat onsets in each participant are identified by letter. Missing clusters of taps marked with an "x" at 0. Extra clusters of taps that occur *prior* to a beat are marked with an "x" above asynchronies (or below, for negative asynchronies). For example, many participants produced extra taps prior to the downbeat of m. 11 because there is no event articulating m. 10$_3$. Gray box near 0 delimits perceptual threshold of -25–42ms described in Levitin *et al.* (2000). Notation is not proportional: each beat occupies the same horizontal space on the page regardless of its timing in performance.

IV. TWO STUDIES OF LISTENER-PERFORMER SYNCHRONICITY

Listener-Performer Synchronicity in a Turn away from the Tonic: Op. 50, No. 1, mm. 1–16

STRUCTURAL FEATURES OF THE G-MAJOR MAZURKA

The first excerpt chosen, from the beginning of the Mazurka in G major, Op. 50, No. 1, is shown in Figure 4. The excerpt is sixteen measures long, grouped as four phrases of four measures each. The first three of these phrases begin with the same melodic material, an ascent outlining the dominant triad. The fourth phrase draws on a portion of the original melody, namely the dotted-eighth/sixteenth figure first seen in m. 2. Cadential goals of the four phrases differ: the first establishes the key with a cadence on the tonic in m. 4. The second ends on a half cadence in m. 8. Both the half cadence and the repetition of mm. 1–3 as mm. 9–11 create a strong tonal expectation for another authentic cadence in m. 12. Yet the move towards the tonic is disrupted through a variety of means. First, an F-natural is introduced in the alto part, altering what would be a tonic chord into a seventh chord (or, including the E in the melody, a thirteenth chord) pointing towards C major, a harmony that never emerges. Second, although the two melodic notes of m. 4 (E5 and D5) recur, the D is now dissonant above the bass G-sharp, rather than consonant above G as it was in m. 4. This G-sharp points towards A minor, a key stated briefly at the beginning of the final phrase en route to the final cadence on the tonic. Temporal expectations are also evaded on the downbeat of m. 12. Unlike every previous measure, there is no articulation in the bass register on the downbeat. Instead, the lowest note in the texture is the alto's grace-note F-natural. Yet without a bass note against which to measure its placement, the metric identity of this note—whether it is before the downbeat or concurrent with it—is difficult to discern. These evasions pertain to DeNora's symbolic force in the music, but how do particular durational decisions and listeners' reactions to those decisions create emotive or corporeal force?
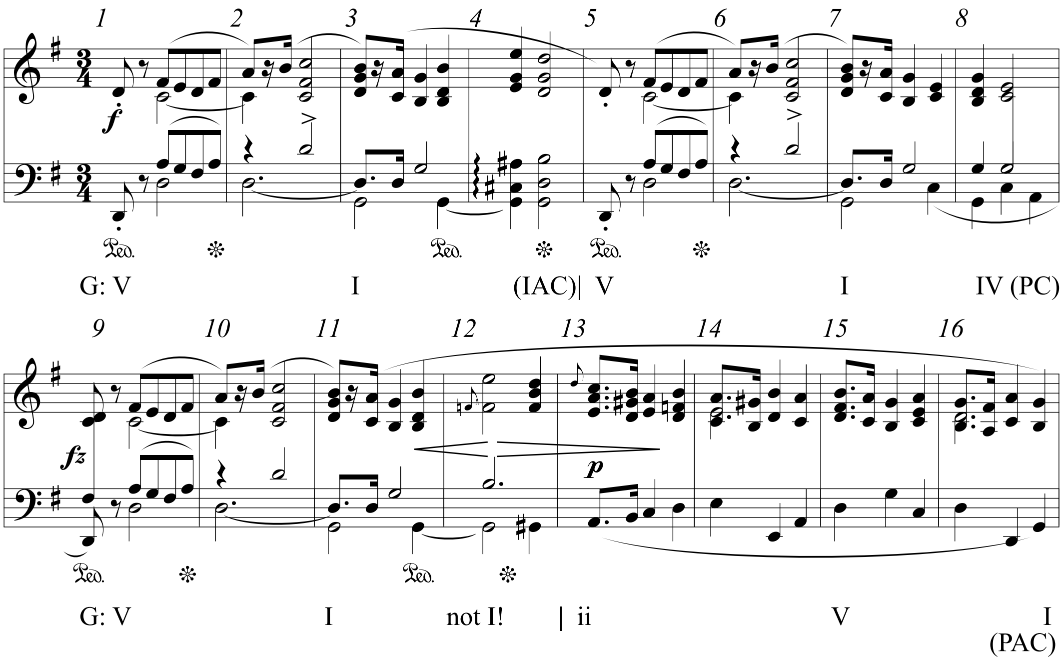
Fig. 4. Frederic Chopin, Mazurka in G Major, Op. 50, No. 1, mm. 1–16.[23]

TEMPORAL FEATURES OF ROBERTO POLI'S 2008 RENDITION

Roberto Poli's 2008 performance of the mazurka was chosen as the idiosyncratic rendition because of the severe challenges it places on attending.

Fig. 5 shows two contours of the durations of beats and measures in performances of the opening of the mazurka. The open dots and dashed lines represent Poli's rendition; the closed dots and solid lines represent an average of twelve other performances (see Recordings Cited below). The y-axis of the plot is logarithmic so that both the durations of beats and the durations of measures can be plotted. Poli's performance is notable in two respects: in the first three phrases, he alternates severe accelerations in odd-numbered measures with slow and steady even-numbered measures. As a corollary, whereas the averaged performance decelerates consistently through each phrase at the level of the four-measure phrase, Poli's durations at that level are much more jagged. Second, he comes to a near stand-still at m. 11_3. That beat, as well as the first two beats of m. 12 (which are articulated as a single half note), last as long as some entire previous measures.

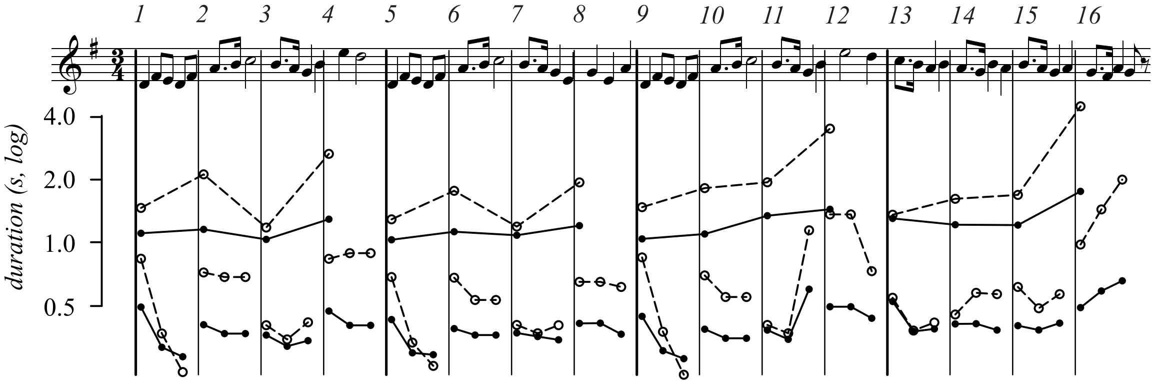

Fig. 5. Durational contour of Roberto Poli's rendition of Chopin, Op. 50, No. 1, mm. 1–16. Dashed lines and open points indicate Poli's durations. Filled dots and closed points indicate durations of 12 other

renditions averaged together. Notice the logarithmic y-axis: upper pair of contours shows durations of measures, segmented every four measures. Lower pair of contours shows durations of beats, segmented every measure. Notation is not proportional.

Figure 6 statistically demonstrates Poli's singular approach and explains my motivation for selecting his performance, as well as the two other comparison performances, as stimuli. The y-axis of the plot shows the standard deviation of beat durations in each performance; this can be regarded as a measure of *tempo rubato*.[24] The x-axis shows median tempo. Poli plays the passage considerably slower than the others and with much more *rubato*. If participants approach attending to Poli's rendition through behaviors and tendencies they've learned listening to other music, it is to be expected that they will have a hard time synchronizing with Poli's performance because it is so distinct from the others. To gauge what part of this difficulty was due to Poli's idiosyncrasies as opposed to the general difficulties of synchronizing with variably timed music, participants also tapped along with renditions of Noriko Ogawa and Lev Oborin, two performers with more normative tempos and amounts of *rubato.*

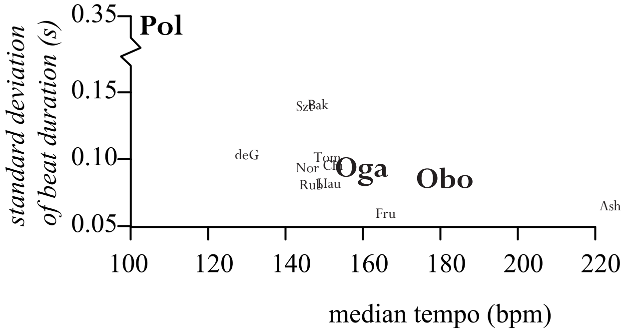

Fig. 6. Median tempo of renditions plotted against the standard deviation of beat durations, a measurement of "global" *tempo rubato*. Performers to which participants tapped are in bold.[25]

PARTICIPANTS' ENTRAINMENT TO POLI'S RENDITION

As discussed above, there are two idiosyncratic features of Poli's rendition: (1) the drastic acceleration of mm. 1, 5, and 9 compared with the slower presentation of mm. 2, 6, and 10 and (2) the near cessation of musical motion around m. 12. Both of these idiosyncrasies make their mark on listeners' abilities to entrain to his rendition. Figure 7 shows participants' entrainment to mm. 1–4. While asynchronicity is low at m. 1_3, few listeners realize the extent to which Poli intends to accelerate moving into m. 2: thus, four of the participants (Participants A, K, L, and P) are slightly late to the downbeat, four others are substantially late (Participants C, E, G, and Q), and two (Participants F and O) do not tap at all. Those participants who tapped very late at the downbeat are still behind at m. 2_2, even though the downbeat is held substantially longer than the last beat of the first measure. This lateness is in conflict with the notion of a lag-1 correlation as described in Repp (2002). In contrast, the third and fourth measures elicit a much more assured entrainment from many participants; there are fewer missed beats and several participants are within the perceptual window of synchronicity from m. 3_2 through m. 4_1. I will return to this pairing of two measures of relatively high asynchronicity with two measures of relatively low asynchronicity in a moment.

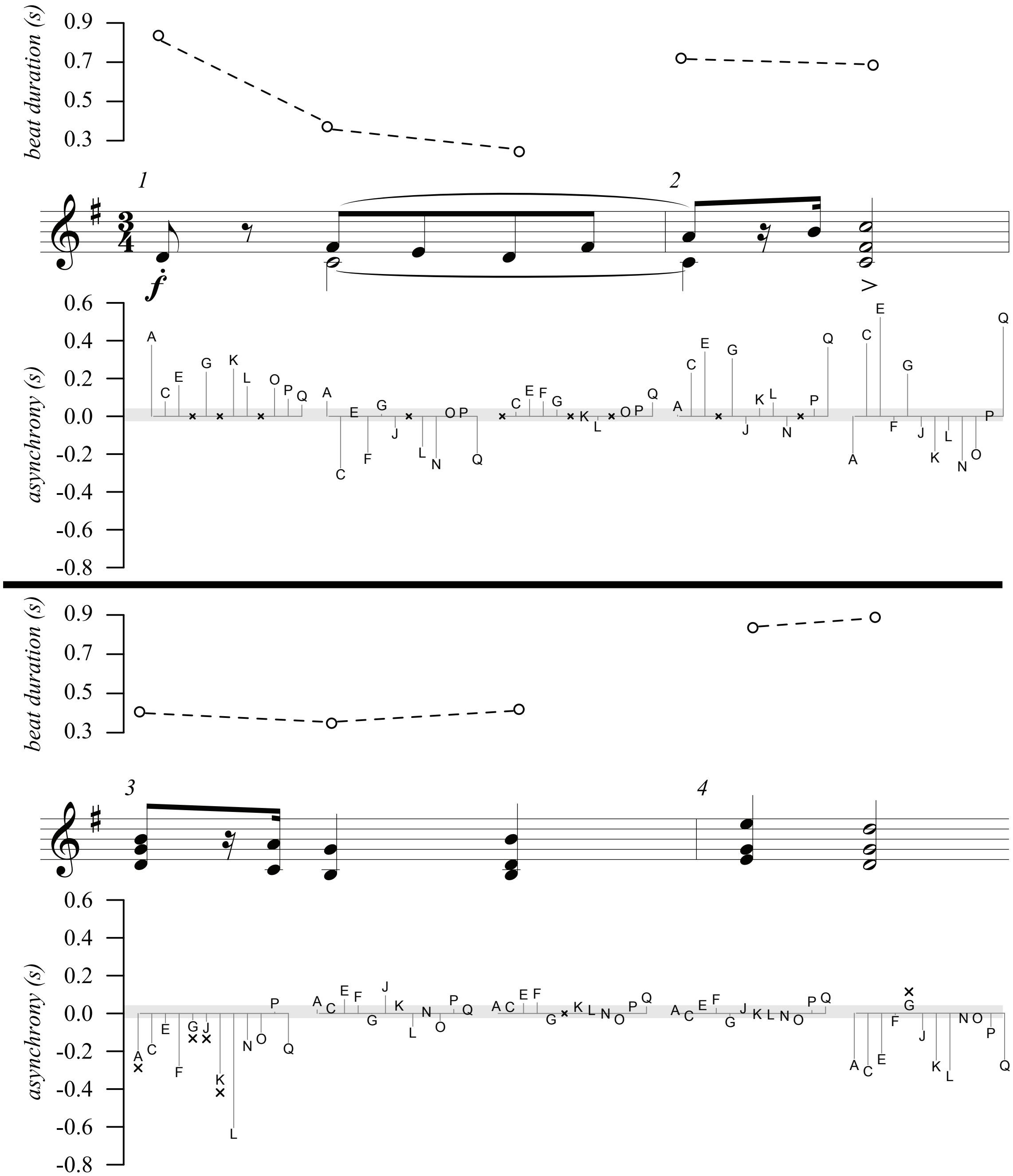

Fig. 7. Entrainment of twelve participants to Poli, Op. 50, No. 1, mm. 1–4. See caption of **Fig. 3** for details. Durational contour is also included.

Figure 8 shows entrainment in mm. 11 and 12, the other idiosyncratic passage of Poli's rendition. (Recall that Poli plays mm. 11_3–12_2 extremely slowly, holding each of those beats as long as the entirety of mm. 3, 5, or 7). Indeed, the duration of Poli's F-natural grace note at the beginning of m. 12 is longer than the previous quarter note duration.[26] While the asynchronies of m. 12_1 are not so severe, the log of extra tap clusters reveals the dissociation listeners experience at this moment. Only one participant correctly places one extra tap cluster for the silent m. 12_2. Others tap as many as four times waiting for Poli to continue. If, as Huron (2006) indicates, failure to predict music's unfolding hazards negative emotional consequences, Poli's m. 12 would seem to mark an apex of stress for attending listeners. How does this break in entrainment affect listeners' perception of the larger passage? Figure 9 segments the sixteen measures into two-measure groups. Most of these groupings are marked in the score by the half notes of mm. 2, 4, 6, 10, and 12. For each two-measure group, which is comprised of six beats, I've taken the mean of the squared error. By "error," I mean the difference between a cluster and its associated beat; lower error bars reflect more accurate tapping. The groups are colored by their motivic and melodic connections (e.g., mm. 5–6 and 9–10 are repetitions of mm. 1–2).

The varying extent of synchronicity within the two-measure groups presents a narrative of dynamic attending to Poli's rendition. An initial contrast is elicited between a group difficult to track (mm. 1–2) and one easier to track (mm. 3–4), as discussed above. As the difficult group recurs in mm. 5–6 and 9–10, participants become more adept at staying synchronized, mitigating the contrast in the difficulty of entrainment between the beginnings and ends of four-measure phrases. The treatment of m. 12 reverses the progress listeners have made in synchronizing as Poli virtually uncouples himself from his audience. Not only is the mean-squared error of mm. 11–12 higher than all the previous groups except the first, but as **Fig. 8** shows, some participants produce several extra tap clusters as well. Yet just as this divorced state is recognized (perhaps by tapping several times and not hearing any expected event), Poli presents a final phrase that, except for the last two beats of phrase-final deceleration, is very easily attended to.[27]

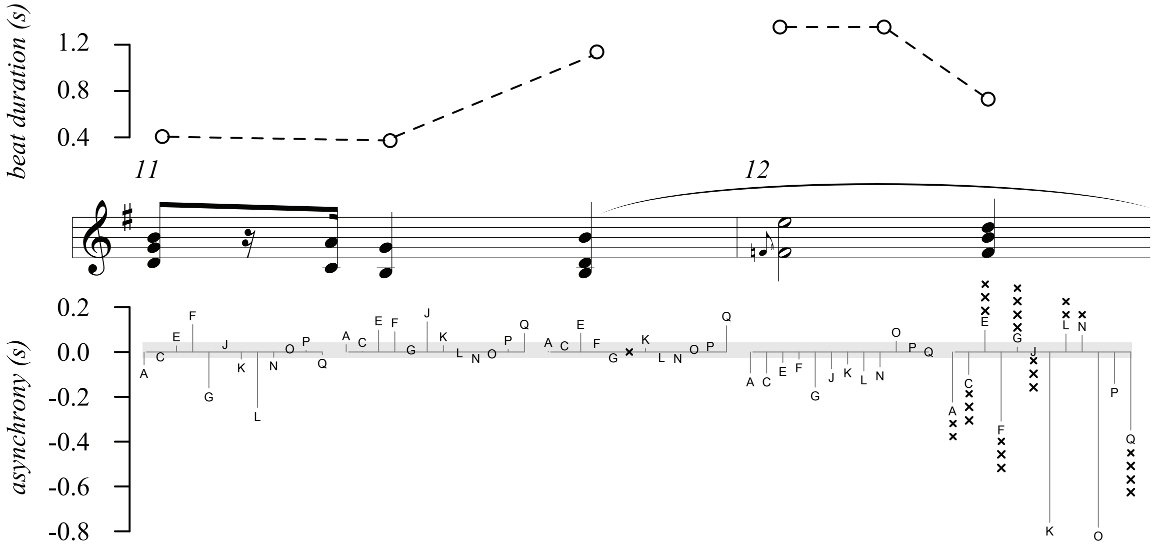

Fig. 8. Entrainment of twelve participants to Poli, mm. 11–12. See caption of **Fig. 3** for details.

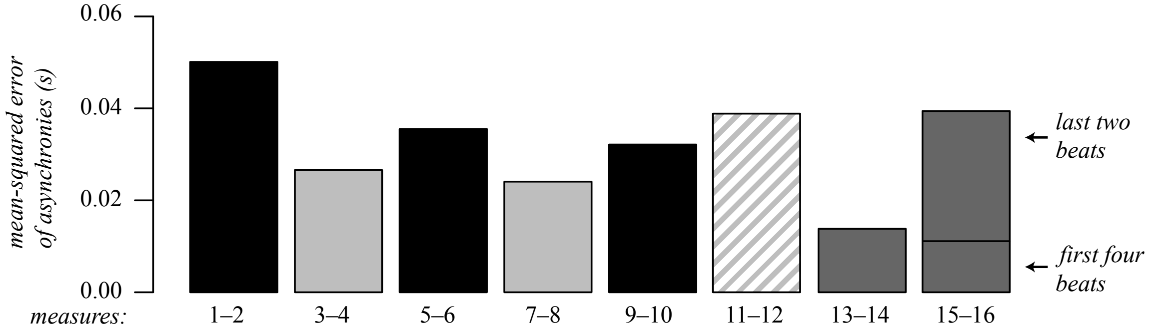

Fig. 9. Mean-squared errors of asynchronies in Poli's rendition of Op. 50, No. 1, segmented every two measures. Two-measure groups are shaded to highlight motivic and melodic similarities.

In a sense, the fluctuating entrainment Poli elicits tracks structural features of the score. A reading of tonal tension, following a model like that of Lerdahl (2004) or Farbood (2006), would reveal m. 12 as the site of greatest tonal distance from G major. A reading of entrainment reveals m. 12 as the site of greatest distance between Poli and his listeners.[28] And yet the experience of tension in Poli's rendition may well relate to how well it is entrained. Perhaps the experience of a listener like Participant N—who is within the perceptual boundaries for most of mm. 9–11, produces the correct number of taps in m. 12, and is only slightly behind at m. 12_3—does not conform to the narrative I outlined a moment ago. So, is the emotive and corporeal experience of tension in m. 12 and resolution approaching m. 16 the result of Chopin's compositional choices, Poli's durational decisions, or the listener's ability for and commitment to entrainment? While I cannot suggest the relative weight of these various factors, I would argue that each has a contribution to make. The way we understand a piece like the G-major mazurka seems to depend in part on not only performer- but also listener-specific factors.

Listener-Performer Synchronicity in an Altered Return: Op. 63, No. 3, mm. 33–57

STRUCTURAL FEATURES OF THE C-SHARP MINOR MAZURKA

The second excerpt, from the Mazurka in C-sharp Minor, Op. 63, No. 3, features another moment that evades tonal and temporal expectations. The first sixteen measures of the excerpt (four phrases of four measures each) constitute the B section of the mazurka. They are highly repetitive melodically and rhythmically—nearly every measure has the same rhythm and the first three phrases end in authentic cadences (the first and second in D-flat major, the relative major of the global tonic, and the third in G-flat major, the subdominant of the tonic of the section). The opening theme returns in the seventeenth measure of the excerpt (m. 49 of the piece, hereafter referred to as A').[29] But in a fashion typical of Chopin, this return is deliberately obscured.[30] The last four measures of the B section do not prepare a return to the tonic key and original theme. Rather, the harmony steadily ratchets through a sequence from G-flat through A-flat to A-natural at the beginning of m. 49. This terminal A-natural eventually resolves to C-sharp minor in m. 52 through a standard nineteenth-century harmonic device that reinterprets a minor seventh (A-natural and G-natural, a dominant function in D major) as an augmented sixth (A-natural and, implicitly, F-double-sharp, pointing to G# as the dominant of C# minor). This harmonic surprise at m. 49_1 also precludes the returning melody from starting on G#4, as it did in m. 1.

Like m. 12 of Op. 50, No. 1, the non-alignment of tonal and thematic return at m. 49 evades several tonal and temporal expectations: the expectation for an event in the soprano register on the downbeat, the expectation for the apparent dominant-seventh chord on A to resolve to D major, and the expectation for the return of the melody to begin with G-sharp. One of the tasks in performing this piece is negotiating these evaded expectations, and thereby influencing how they are experienced by a listener. If a participant is able to immediately entrain to the passage beginning at m. 49, presumably because the performer is choosing a durational patterning reminiscent of the way in which the beginning of the piece is played, then the tonal disorientation of mm. 46–49 may be ameliorated emotionally. On the other hand, if a performer presents durations in and after m. 49 that inhibit synchronicity, the reacquisition of C-sharp as a tonic, as well as the demarcation of the new formal section, may be delayed or obscured.

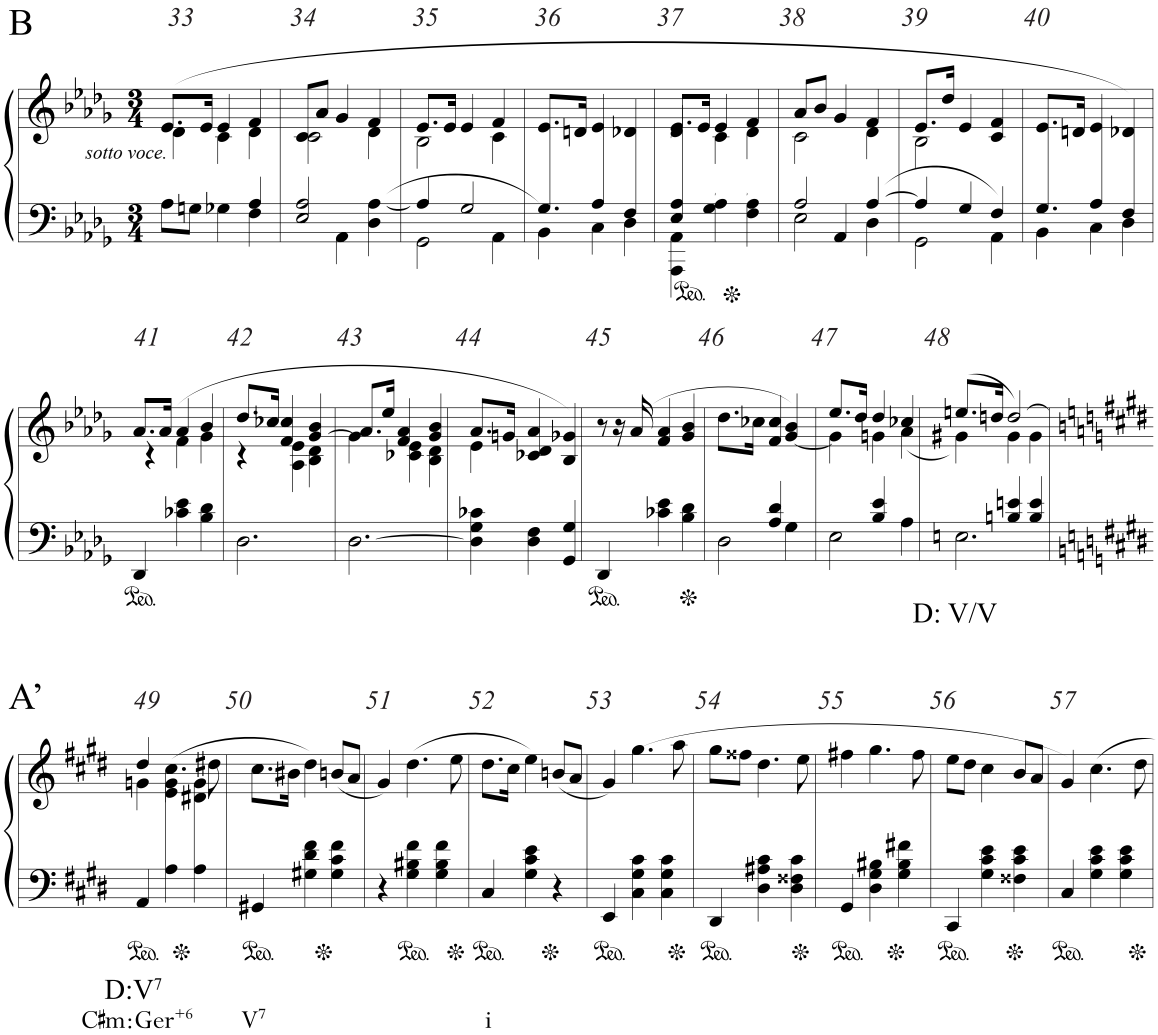

Fig. 10. Chopin, Mazurka in C-sharp Minor, Op. 63, No. 3, mm. 22–49.

TEMPORAL FEATURES OF STANISLAV BUNIN'S RENDITION

As with the Op. 50, No. 1 example, I sought a performance of the passage in question—the transition between B and A'—that was unique within a collection of many renditions. Figure 11 shows the durational decisions of Stanislav Bunin in comparison to the averaged durations of 56 other renditions in a manner similar to Figure 5. As Figure 11 shows, his approach to the B section consists of substantial acceleration-deceleration patterns at the level of the two-measure groups. Unlike other performers, these two measure-groups are fashioned into a larger acceleration-deceleration gesture that unfolds over the entire sixteen measures (as opposed to the average performance's four-measure groupings). As a result of this highly patterned presentation, at both the level of the beat and the level of the two-measure group, participants are highly entrained to Bunin in the B section. From mm. 41$_1$–46$_3$, not a single beat elicits average (absolute value) asynchronies outside the just-noticeable difference threshold.

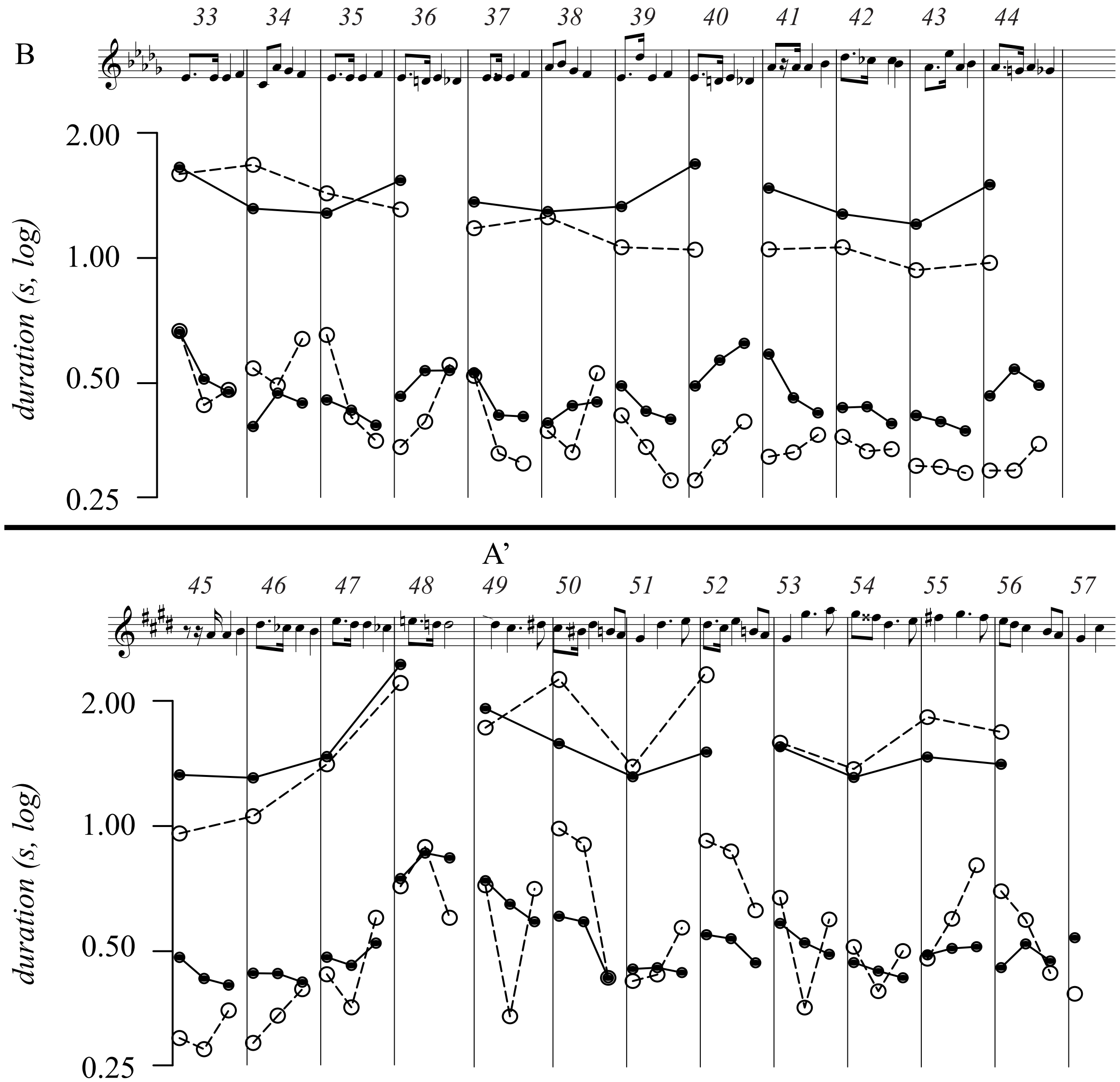

Fig. 11. Duration in Stanislav Bunin's rendition of Op. 63, No. 3, mm. 33–57. Analogous to **Fig. 5**, dashed lines refer to Bunin and solid lines refer to an average of fifty-six other renditions.[31] Higher lines refer to durations of measures and lower lines refer to durations of beats.

But his divergence from more conventional performance styles is most evident in the return of the main theme. Most performers decelerate significantly throughout mm. 45–48 in anticipation of the first measure's return, but then regain a faster tempo soon after the downbeat of m. 49; the remainder of A' usually eschews drastic changes in tempo, in preference of moderate decelerations at the grouping boundary of m. 53. Bunin, on the other hand, presents beat durations throughout mm. 48–49 that in contrast to the triple meter, oscillate between slow and fast. Once the main theme has returned, he decelerates substantially in mm. 50 and 52, returning to a slowness that other performers leave behind in m. 48. Through these returns to a slower tempo, Bunin links the melodic ascent of mm. 50 and 52 with the earlier and considerably tenser rising sequence of mm. 46–48, thereby resisting the release of tension offered by the return to the main theme. Bunin also forgoes the gestures of closure that other performers employ to highlight the cadence of m. 56. Whereas the averaged contour decelerates beginning in m. 55 and continues to slow as the melody descends to C#5 in m. 57, Bunin places his own apex of duration at the beginning of that ascent in mm. 55 and accelerates through m. 57.

PARTICIPANTS' ATTENDING TO BUNIN'S RENDITION

These choices have a profound effect on participants' abilities to entrain to Bunin's rendition, even after hearing it nine times in quick succession (see Figure 12). The oscillation of durations in mm. 48–49 begets oscillating asynchronies in a pattern reminiscent of the lag-1 correlation described in Repp (2002): when a short beat follows a long one, participants tap late to the subsequent onset (e.g., mm. 49_1 and 49_3). The reverse is also true, explaining the earliness of mm. 48_3, 49_2, and 50_1. Once the new phrase is underway, the halting presentation of the first two beats of mm. 50 and 52 elicit anticipatory taps throughout those measures. (Although several participants, e.g., A, C, I, J, and Q, are late for mm. 50_2 and 52_2, these participants also produce an extra tap cluster before the beat. It is understood that they tapped early and then tried to correct the error by tapping again).

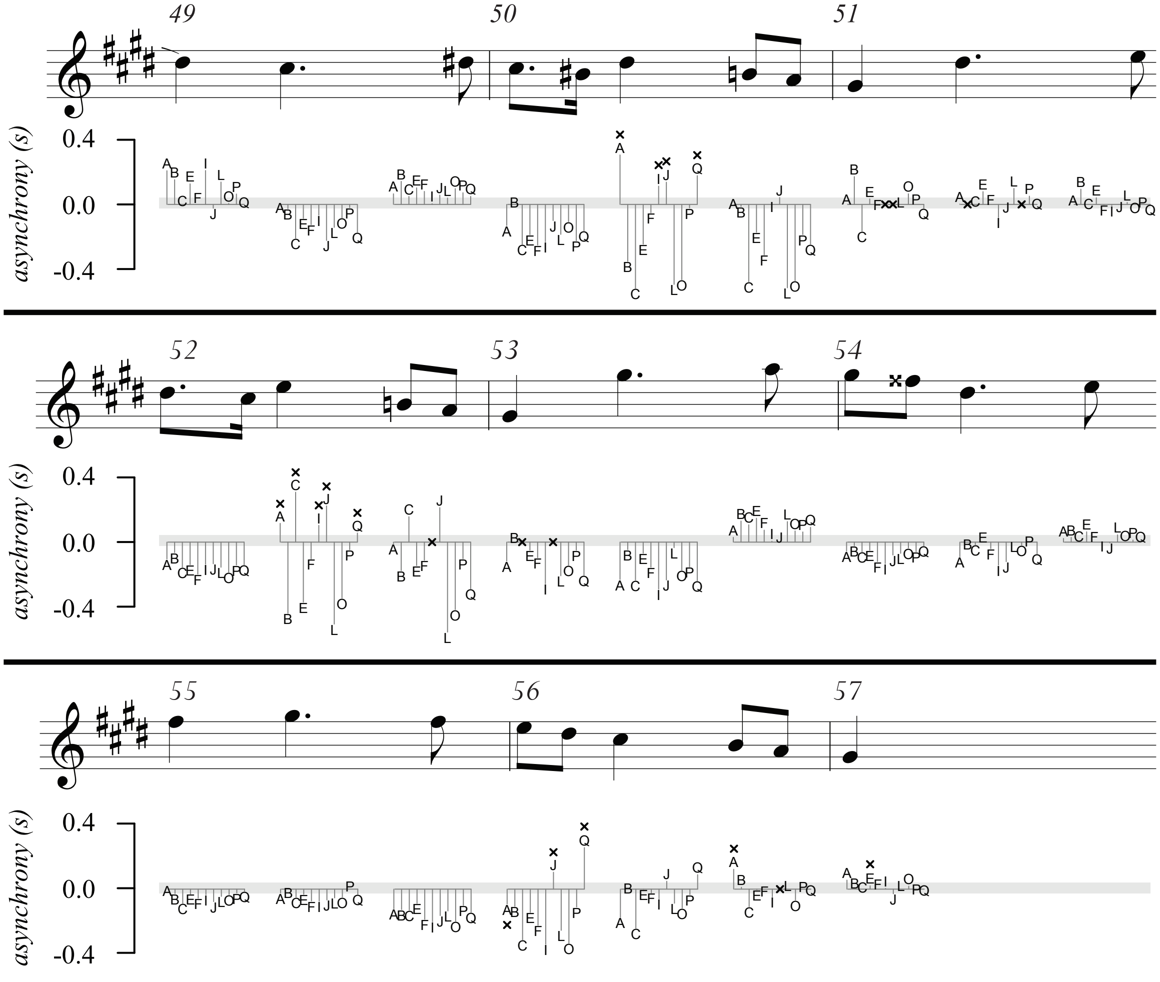

Fig. 12. Entrainment of eleven participants to Bunin's rendition of Op. 63, No. 3, mm. 47_1–57_1. See **Fig. 3** for details.

In the following phrase, asynchronies remain high as virtually every participant is early on almost every beat from m. 53 through m. 56. Not until the downbeat of m. 57 do most participants align with Bunin. This evasiveness can be contrasted with the participants' performance in tapping to the rendition of Frederic Chiu. **Fig. 13** shows mean-squared errors of asynchronies for Bunin on the left and Chiu on the right in mm. 47–57, and includes a line marking the perceptual boundary.[32] On the left, only at m. 57 do participants, on average, produce imperceptible asynchronies. On the right, however, Chiu's rendition enables them to accomplish this immediately on the downbeat of the new section in m. 49, and again at least once in every measure through to m. 56.

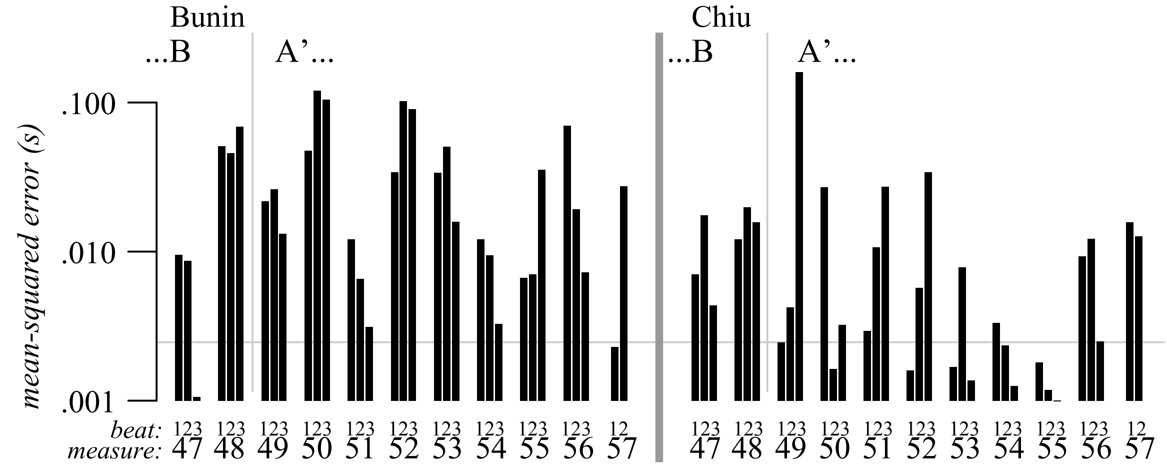

Fig. 13. Mean-squared errors of asynchrony in attending to Bunin (left) and Frederic Chiu (right). Horizontal line at .0018, the square of .042s, the larger, positive perceptual limit discussed above.

The relative ease with which Chiu affords entrainment at m. 49 and after, in contrast to Bunin's evasiveness, affects a listener's experience of the piece on multiple levels. At the local level, listeners may experience the kind of isolation and anxiety that results from an inability to predict the actions of agents in the environment, as described by Huron (2006). At a larger level, this isolation may be compounded by its placement within the form. Insofar as similar durational treatment can highlight thematic and motivic connections, a listener unable to track Bunin's return of the main theme may have a harder time recognizing mm. 49–52 as a return. The hazard of failing to recognize the recurrence of the main theme presents the possibility for two crucially different perspectives on the piece. For a listener who attends well the recurrence of A at m. 49—a listener like Participant C attending to Frederic Chiu, who was within the perceptual boundary for eight of the twelve beats of mm. 49–52—the ABA' ternary form of the mazurka may be apparent at m. 49. For a listener who cannot attend well to the recurrence of A at m. 49—and this would include all the participants attending to Bunin in this study—the recognition of m. 49 as A' may be delayed and its relatedness to A may be undermined. This obfuscation of form may be the composer's wish, and Chopin certainly went to some lengths to disguise m. 49 as the recurrence of the main theme. But subverting the awareness of the thematic return, I would argue, requires both some idiosyncratic performance decisions as well as a diminished ability on the part of listeners to attend to those decisions. In this way, listener-performer synchronicity affects basic questions of musical meaning as much as do the choices of composers and performers.

CONCLUSION: DURATION, TIMING, AND MUSICAL MEANING

As I interpret it, the current study supports the view that listener-performer synchronicity is influenced not only by previous durations but also by listeners' expectations of the relationships between musical structures and duration. This increases the evidence that the relationship between events and taps is more sophisticated than proposed in simple tapping models (e.g. Repp, 2002; and Dixon *et al.*, 2006). Clearly, there are instances where listeners place taps to create inter-tap intervals much shorter than previous inter-beat intervals, in contrast to Repp's model of lag-1 correlation. And the data collected in this experiment does not conform to Dixon *et al.*'s model of perceptual smoothing. Granted, the stimuli presented to these participants contained considerably greater beat-to-beat durational differences than those presented to the participants of the other studies. Thus, lag-1 correlation and perceptual smoothing may still be sufficient models for music with more modest variation in tempo.

Further, I hope the processing method presented in Section III above can augment analyses of emotional trajectories of performances, as those trajectories are shaped not only by the symbolic meaning of the pitches and rhythms chosen by the performer, but also the dynamic experience of a listener trying to predict their unfolding. Nevertheless, a more fully fleshed-out theory of listener-performer synchronicity in durationally variable contexts will likely require additional studies in the mold of those presented here. In particular, future work might clarify the relation between synchronicity and preference, between listeners' expectation and tapping behavior, the impact of the tapping modality on synchronicity, and composers' contribution to listener-performer synchronicity.

I assume that the failure to predict when a performer will place an anticipated event is negatively valenced, especially when a listener is actively trying to make such predictions. This negative valence is supported by Huron (2006) and, as he relates, unexpected event onsets also degrade the accuracy of pitch perception (Jones, Moynihan, MacKenzie, & Puente, 2002). Yet it is not certain that negative emotions attended the participants who tapped four times waiting for Poli to continue past the downbeat of m. 12. In future studies, participants could be asked to express their preference for each performance: perhaps their preference would correlate with the performers they attend to most easily. On the other hand, there may be other factors contributing to preference (e.g., tempo), and preference for an easily attended rendition may be a cause and not an effect.

In this study I've assumed that asynchronies result from violations of temporal expectations and, further, that differences in asynchronies reflect different expectations. Yet those expectations are not represented in any formal way. Another approach might ask participants to perform the piece (or, for non-pianists, perform just the melody line) prior to tapping to numerous recorded renditions. Subsequently, asynchronies in tapping could be compared to differences in timing profiles. Such a method might also elucidate what portion of asynchronies result from motor noise and other factors of production rather than inaccurate temporal prediction. Similarly, while tapping is a convenient way to express temporal predictions, the tapping modality may have its own peculiarities and blind spots. The sort of spontaneous movement explored in Toiviainen, Luck & Thompson (2010) presents far greater challenges in data collection and processing than those encountered here, but may also indicate how listeners change their predications as events unfold in richer detail than, say, the presences of a number of extra taps.

Like earlier studies of performance timing, the range of musical styles encountered in the tapping literature is regrettably small. As in so many others (e.g., Repp, 2002; Dixon, Goebl, & Cambouropoulos, 2006, etc.), the scope of this study is limited to a single genre of a single composer. And yet other appropriate repertoires, especially those outside of eighteenth- and nineteenth-century European notated music, raise more fundamental questions. Much of the music we hear (e.g., much commercially available popular music) is not "expressively performed" in the sense used here because beat durations are normalized through mechanical means. At the other end of the spectrum lies what Martin Clayton (1996) calls "free rhythm," music without a clear sense of pulse (e.g., improvised cadenzas, the *taqasim* of the Levant, various genres of unaccompanied chant, etc.). It may be that a more thorough theory of listener-performer synchronicity will require as a prerequisite a more thorough theory of the perception and production of musical meter and rhythm.

Such a theory is worth pursuing, and the tapping paradigm employed here may be a powerful tool in future work. The first wave of studies of expressive timing cited in the introduction are contemporaneous with the post-structuralist turn in musicology in the 1980s and 1990s (e.g., Kerman, 1980; McClary, 1991; Whittall, 1991; Kramer, 1992). These two bodies of literature contrast in their methods, modes of discourse, and theoretical underpinnings. Yet they have a shared skepticism of the ability of an autonomous musical work to determine its own meaning. This skepticism has led to a flourishing of studies of performance (and performance timing): the notion that performers and composers cooperatively create musical meaning is no longer controversial. By documenting how different listeners attend to differently timed performance, and further by speculating on how the different experiences of attending might alter listeners' perception of musical works, I hope that this study can further expand the group of agents vested with the ability to affect the meanings of works to include listeners as well.

NOTES

[1] The author welcomes correspondence at mohriner@gmail.com. This research was supported by a Faculty Development Grant from Shenandoah University and approved by that institution's Review Board. The author gratefully acknowledges the assistance of Mark Chan in administering the experiment, Yushen Han and Michael Magro for assistance in data preparation, John Walker (2004, 2008) for writing the software that converted MIDI data to comma-separated data, Keith Salley and Gretchen Horlacher for their comments on earlier drafts, an anonymous reviewer for the journal *Music Analysis*, and Renee Timmers for her comments as the editor of *Empirical Musicology Review*.

[2] A review of the expressive timing literature can be found in Gabrielsson (1999, pp. 523–56; 2003, pp. 225–36). The phrase "expressive timing" itself is somewhat controversial, because as Renee Timmers and Henkjan Honing (2002) show, many writers mean quite different things by the first word of the phrase. In this article, I use the term "expressive timing" to mean "unequal and non-repeating beat durations." I do not address in any systematic way whether these durations are unequal because of music-structural features or performers' predilections. Aside from its frequent lack of specificity, Nicholas Cook (1999) has criticized the term "expressive timing" for relegating performers' "expression"—"traditionally seen as the core of performers' individuality (p. 240)"—to the conveyance of musical structure. The first generation of writers on expressive timing, Clarke (2006), Repp (2001), and Todd (1999), have all nuanced their structuralist view of timing in performance in subsequent work.

[3] See Bengtsson and Gabrielsson (1977) on the interaction of timing and meter, Desain and Honing (1994) on that of timing and ornamentation, and Friberg, Bresin, and Sundberg (2006) on that of timing and melodic contour. For a method of distinguishing the contributions of these various factors, see Windsor *et al.* (2006).

[4] Much of the SMS literature concerns synchronization with periodic (i.e., metronomic) signals and addresses questions of preferred rates and rate limits in tapping (e.g., Fraisse, 1982; Repp, 2003; Mikaye, Onishi, & Pöppel, 2004; Van Noorden & Moelants, 1999; Todd, Cousins, & Lee, 2007), abilities for anti-phase tapping (Semjen, 2000; Keller & Repp, 2005) and tapping multiples or divisions of a given rate (Drake et. al, 2000). Another segment of the literature focuses on tapping to complex but repeated cycles of durations (e.g., Patel, 2005) or periodic cycles with isolated perturbations in period duration or phrase (e.g., Large et. al, 2002; Repp, 2011). Participants in the study described below, in contrast to most of the existing literature, are tapping to complex, non-repeating cycles of duration. And my concern is not for how accurate their tapping is on average, but rather for how specific moments of asynchronicity might alter perceptions of the piece in whole or in part.

[5] One issue in studying SMS is distinguishing the perception of a beat with the production of a gesture that expresses the beat. The tapping method has a long history because the difference between perception and production is thought to be small. Still, there may be some lag between the two, and changing the quality of gestures participants are asked to produce (say, arm motions or full-body dancing, etc.) may affect the apparent success of temporal prediction (Leman et. al, 2007; Toiviainen, Luck, & Thompson, 2010).

[6] The ability to synchronize with a conventional timing profile, while impressive, is not as robust as the ability to synchronize with a periodic stimulus.

[7] While the literature supports a metric component to timing variation, there is also considerable variation in the proportions of the measure devoted to each beat across the many measures of a performance. Insofar as London's Many Meters Hypothesis advocates a consistent timing pattern across all measures of a particular genre, it is similar to Manfred Clynes's earlier notion of the "Composer's Pulse" (1983). While these two ideas are distinguished by the explanations their authors put forward for differences in various patterns—the conventions of genre in the case of London and the composer's identity in the case of Clynes—the counter-evidence presented against the "Composer's Pulse" should cast aspersions on the Many Meters Hypothesis as well (Thompson, 1989; Repp, 1990). There may well be more variation in the proportion of measures accorded to each beat *within genres* than *between genres*.

[8] This model is not unlike Christopher Hasty's concept of durational projection (1997, pp. 84–95): as one event ends, its duration is projected forwards to predict the ending of a second event (and equivalently the beginning of third event). Justin London (1999) has criticized Hasty's concept of projection for its apparent denial of remembered events other than the most recent, calling it "a first-order Markov process (p. 265)."

[9] Drake, Penel, & Bigand (2000) also explore SMS with expressively timed music, though their focus is on how such attending relates to mechanically timed performances or performances with intensity accents but without timing variation.

[10] More formally, Levitin et al. define their detection threshold as "the asynchrony beyond which fewer than 75% of responses were incorrectly judged as simultaneous."

[11] The participants are undergraduate music majors at the Shenandoah Conservatory. Their age varies from 18–25, their years of training on their primary instrument varies from 3–18, and their self-evaluated familiarity with Chopin's music varies from listening recreationally for "more than hour per week" to "less than one hour per month." However, none of these factors correlate significantly with the sum-of-squares of asynchronies they produce. One explanation for this is that I did not assess their familiarity with Chopin's style adequately. Another is that the ability to attend to renditions of Chopin's music is one that untrained listeners acquire through cultural exposure. For a discussion of the effects of expert training on aspects of music cognition, see Bigand and Poulin-Charronnat (2006).

[12] Two of the excerpts, mm. 1–16 of the Mazurka in G Major, Op. 50, No. 1, and mm. 33–57 of the Mazurka in C-sharp Minor, Op. 63, No. 3, are discussed extensively below. The third excerpt, mm. 1–16 of Op. 63, No. 3, was included essentially as a training task so that participants might recognize the recurrence of the main theme in m. 49 of that piece.

[13] This study shares with many other studies of entrainment an exclusive focus on the aural modality. Many writers have advocated for the importance of visual cues in synchronization (e.g., Will, 2011; Clayton, 2007; Will, Utter, Clayton, & Leante, 2005; Guttman, Gilroy, & Blake, 2005), and such cues would probably change how participants tap to the excerpts. In particular, one performance of Op. 50, no.1 discussed below features a very long pause that participants found very hard to predict, and, on the original video, the performer gives visual cues that might aid in the maintenance of synchronization. The decision to pursue only the auditory modality was made (1) to broaden the number of recordings available for study and (2) in recognition that the auditory-only modality is encountered, not infrequently, in ecologically valid contexts (e.g., private listening to audio recordings through headphones, obscured sight-lines in live performance, concert attendance by the visually impaired, etc.). While I recognize that including visual cues would alter results, I would also argue that just as performers are able to place events in time so as to surprise listeners, they may also use visual cues to inhibit temporal prediction.

[14] There was some user error in saving MIDI files in Logic, which resulted in the taps being zero-adjusted to the first tap rather than the start of the stimulus audio. These trials have been excluded from the data and explain why the discussions below do not list seventeen participants.

[15] By "tonal expectation," I refer to the predictions listeners of tonal music constantly make as to the nature of the following chord. These expectations are thought to arise from statistical learning and are documented through corpus studies of harmony (e.g., Tymocsko, 2010; Quinn & Mavromatis, 2011; White, 2013; Ohriner, 2013). For example, because a major-minor seventh chord built on the fifth degree of the scale so frequently moves to a major or minor chord on the first degree of the scale, a listener hearing such a chord will expect that transition. In selecting stimuli for the experiment, I chose moments that evade such expectations.

[16] Not all the mazurkas feature this lengthening of beat two, but many feature significantly unequal beats of some kind. In the first piece selected, the Mazurka in G major, Op. 50, No. 1, a dataset of the onsets of beats in thirteen performers shows a metrical durational pattern of <.31,.29,.38>, a very long third beat and a very short second.

[17] Incidentally, the cluster before m. 9_2 illustrates a case of lag-1 correlation, falling where m. 9_2 would have occurred had m. 9_1 been the same duration as m. 8_3.

[18] If the count of taps within 100ms windows repeats values, the repetitions are removed so that inflection points are unique. This tends to place the centers of clusters artificially earlier, an issue that is corrected at the end of the process (see below).

[19] Some clusters are diffuse enough to generate two relevant inflection points in a plot like that of Figure 2c, so I remove detected clusters that are less than 125ms apart. This value, 8Hz, is roughly as fast as participants could possibly tap (Repp, 2003; Keele & Hawkins, 1982), and thus clusters closer together must be an artifact of the clustering method. This method is presented as a tool to examine participant entrainment, not an end unto itself. Once the code was written, I plotted the results and made some manual adjustments. These included converting a number of nearby "extra taps" into a cluster and deleting one of two clusters that was close together, The cluster after m. 10_1 in Figure 2 was one such adjustment. In the plot, it appears that taps never cluster higher than the threshold of 3. Nevertheless, there is evidently a cluster of taps. Many of these manual adjustments could be avoided by altering the width of the 100ms window or altering the hop size of 25ms, but these adjustment create their own artifactual tap clusters.

[20] The method of the determining the clusters necessitates this final step. Recall that I removed repeated values in the counting of taps within 100ms windows. If nine taps, each in a different trial, occurred 10ms apart, the onset of the cluster would align with the ninth one and all the others would appear early, when a better interpretation would align the cluster with the fifth one.

[21] Judging asynchrony from the cluster closest to the onset and deeming the others "extra" may seem arbitrary. I could just as easily always take the first aligned cluster, or the last. What is important is not the quality of the error (its asynchrony) but its quantity (the fact that multiple clusters of taps were produced for a single beat onset).

[22] Although seventeen participants sat for the experiment, five failed to complete the proper protocol for saving their taps. Essentially, their files were saved with the first tap indexed from 0, rather than the beginning of the audio indexed from 0. As a result there is an unrecoverable latency in their data. In retrospect, a clearer protocol (or specialized experimental software) could have prevented this data loss.

[23] By IAC, PC, and PAC, I mean, respectively, an imperfect authentic cadence (ending on a tonic chord without scale-degree 1 in the melody), a plagal cadence (a cadence emphasizing the subdominant chord), and a perfect authentic cadence (ending on a tonic chord with scale-degree 1 in the melody). The roman-numeral analysis accounts only for the most salient chords. Other analysts might code the harmony differently, but I believe most would concur with the surprising nature of the downbeat of m. 12.

[24] There are other ways to formalize a definition of *tempo rubato*. Rather than computing standard deviation of beat durations, one might compute the average change in tempo from beat to beat. The latter measurement is a more "local" measure of *rubato*, whereas standard deviation, which ignores the ordering of beats, is a more "global" measure. In either case, Poli is an outlier, and the ordering of the other performers is not much changed.

[25] Timing data for the thirteen performances of Op. 50, No. 1 were collected manually by the author. See note [26] below for more details.

[26] I make a practice of annotating both the bass and treble events when I can detect chordal "spread," and I do so by marking events on the spectrograph of the wave form in the software package Sonic Visualiser (Cook & Leech-Wilkinson, 2010). Throughout this article, the onset of a beat is understood as the mean of the bass and treble onsets when they are perceptibly separate. Poli's m. 12_1 is the sole exception to this practice. The duration of the grace note in Poli's rendition is 580ms, more than five times the duration of any other spread chord. Thus I take the onset of the grace note as the onset of the beat. If I took the average of the alto F-natural and the soprano E, the asynchronies of participants, already negative in almost every case, would be another 290ms greater.

[27] Recall Repp's finding (2002) that phrase-final decelerations are normally underestimated, leading to the higher mean-squared errors of the final two beats of the phrase.

[28] In fact, mm. 1–2 elicit greater asynchronies than mm. 11–12, though I believe the asynchronies in mm. 1–2 pertain to participants' process of "tuning in" to the tempo of the performance, not errors in prediction caused by evasions of temporal expectations as in mm. 11–12.

[29] Recall that participants tapped to three renditions of the first sixteen measures of the piece as a training exercise so that they might recognize the recurring theme.

[30] William Rothstein (1989, pp. 214–282) has noted a paradox in many of Chopin's pieces: although he often writes in genres with extremely regular phrase lengths (e.g., mazurkas, waltzes, polonaises, etc.), he also goes to great lengths to hide these regular phrase boundaries.

[31] Timing data for fifty-four of the fifty-seven renditions of Op. 63, No. 3 were annotated as part of the Mazurka Project (Center for the History and Analysis of Recorded Music, 2009). For a detailed discussion of the annotation method, see "Reverse Conducting Example," http://www.mazurka.org.uk/info/revcond/ (accessed 29 August 2014).

[32] Squaring the asynchronies, part of the mean-squared error computation, eliminates their sign. Yet recall that the perceptual threshold of asynchrony differed by sign: positive asynchronies as small as 25ms are perceivable, while negative asynchronies need to be nearly twice as large (42ms) to pass the threshold. The gray line denoting the perceptual limit in Figure 13 marks the square of 42ms, the larger of the two limits (of different sign) measured in Levitin *et al.* (2000).

REFERENCES

Askenfelt, A., & Jansson, E. V. (1990). From touch to string vibrations. I: Timing in the grand piano action. *The Journal of the Acoustical Society of America*, *88*, 52.

Benadon, F. (2009). Time warps in early jazz. *Music Theory Spectrum*, *31*(1), 1–25.

Bengtsson, I., & Gabrielsson, A. (1977). *Rhythm Research in Uppsala*. Musical Academy.

Bengtsson, I., & Gabrielsson, A. (1983). Analysis and synthesis of musical rhythm. *Studies of Music Performance*, *39*, 27–60.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition*, *100*(1), 100–130.

Center for the History and Analysis of Recorded Music. (2009). *The Mazurka Project*. http://www.mazurka.org.uk/.

Clarke, E. F. (1993). Generativity, mimesis and the human body in music performance. *Contemporary Music Review*, *9*(1–2), 207–219.

Clarke, E. F. (1998). The semiotics of expression in musical performance. *Contemporary Music Review, 17*(2), 87–102.

Clarke, E. F. (2006). Making and hearing meaning in performance. *Nordic Journal of Aesthetics*, *18*, 24–48.

Clayton, M. R. (1996). Free rhythm: Ethnomusicology and the study of music without metre. *Bulletin-School of Oriental and African Studies*, *59*, 323–332.

Clayton, M. R. (2007). Observing entrainment in Indian music performance: Video-based observation analysis of tanpura playing and beat marking. *Musicae Scientiae*, *11*(1), 27–60.

Clayton, M. R. (2012). What is entrainment? Definition and applications in musical research. *Empirical Musicology Review*, *7*(1–2), 49–56.

Clynes, M. (1983). Expressive microstructure in music, linked to living qualities. In J. Sundberg (Ed.), *Studies of Music Performance No. 39* (pp. 76–181). Stockholm: Royal Swedish Academy of Music.

Cook, N. (1999). Analysing performance and performing analysis. In N. Cook and M. Everist (Eds.), *Rethinking music* (pp. 239–61). Oxford: Oxford University Press.

Cook, N. (2009). Squaring the circle: Phrase arching in recordings of Chopin's Mazurkas. *Musica Humana*, *1*(1), 5–28.

Cook, N., & Leech-Wilkinson, D. (2010). *A musicologist's guide to Sonic Visualiser*. Centre for the History and Analysis of Recorded Music, accessed August 29th, 2014. http://www.charm.rhul.ac.uk/analysing/p9_1.html.

DeNora, T. (2000). *Music in everyday life*. Cambridge: Cambridge University Press.

Desain, P., & Honing, H. (1994). Does expressive timing in music performance scale proportionally with tempo? *Psychological Research*, *56*(4), 285–292.

Dixon, S., Goebl, W., & Cambouropoulos, E. (2006). Perceptual smoothness of tempo in expressively performed music. *Music Perception*, *23*(3), 195–214.

Dodson, A. (2008). Performance, grouping and Schenkerian alternative readings in some passages from Beethoven's 'Lebewohl'Sonata. *Music Analysis*, *27*(1), 107–134.

Dodson, A. (2009). Metrical dissonance and directed motion in Paderewski's recordings of Chopin's Mazurkas. *Journal of Music Theory*, *53*(1), 57–94.

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception*, *18*, 1–24.

Eigeldinger, J. J., & Shohet, N. (1986). *Chopin: pianist and teacher as seen by his pupils*. Cambridge: Cambridge University Press.

Farbood, M. M. (2006). *A quantitative, parametric model of musical tension*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, USA.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). New York: Academic Press.

Friberg, A., Bresin, R., & Sundberg, J. (2006). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology*, *2*(2–3), 145–161.

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *The Journal of the Acoustical Society of America*, *98*, 2524.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of music* (pp. 501–602). San Diego: Academic Press.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software*, *31*(7), 1–24.

Gingras B, Asselin P-Y, & McAdams S. (2013). Individuality in harpsichord performance: disentangling performer- and piece-specific influences on interpretive choices. *Frontiers in Psychology*, *4*, doi: 10.3389/fpsyg.2013.00895. http://journal.frontiersin.org/Journal/10.3389/fpsyg.2013.00895/abstract

Grosche, P., Müller, M., & Sapp, C. (2010). What makes beat tracking difficult? A case study on Chopin mazurkas. In: J.S. Downie & R.C. Veltkamp (Eds.), *Eleventh International Society for Music Information Retrieval Conference* (pp. 649–654). Utrecht, Netherlands: Universiteit Utrecht.

Guttman, S. E., Gilroy, L. A., & Blake, R. (2005). Hearing what the eyes see: Auditory encoding of visual temporal sequences. *Psychological Science*, *16*, 228-235.

Hasty, C. F. (1997). *Meter as rhythm*. New York: Oxford University Press.

Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, Mass.: MIT Press.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*(3), 459–491.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus driven attending in dynamic arrays. *Psychological Science*, *13*(5), 313–319.

Keele, S. W., & Hawkins, H. L. (1982). Explorations of individual differences relevant to high level skill. *Journal of Motor Behavior*, *14*, 3–23.

Keil, C. (1987). Participatory discrepancies and the power of music. *Cultural Anthropology*, *2*(3), 275–283.

Keller, P. E., & Repp, B. H. (2005). Staying offbeat: sensorimotor syncopation with structured and unstructured auditory sequences. *Psychological Research*, *69*, 292–309.

Kerman, J. (1980). How we got into analysis, and how to get out*.* *Critical Inquiry*, *7*(2), 311–331.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102*(3), 299–314.

Kramer, J. D. (1988). *The time of music: New meanings, new temporalities, new listening strategies.* New York: Schirmer Books.

Kramer, L. (1992). The musicology of the future. *Repercussions, 1*(1), 5–18.

Leman, M., Desmet, F., Styns, F., Van Noorden, L., & Moelants, D. (2009). Sharing musical expression through embodied listening: a case study based on Chinese guqin music. *Music Perception*, *26*(3), 263–278.

Large, E. W. (2000). On synchronizing movements to music. *Human Movement Science*, *19*, 526–566.

Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *Psychology of time* (pp. 189–232.). London: Emerald Publishing Group.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159.

Lerdahl, F. (2004). *Tonal pitch space*. New York: Oxford University Press.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, Mass.: MIT Press.

Levitin, D. J., MacLean, K., Mathews, M., Chu, L., & Jensen, E. (2000, May). The perception of cross-modal simultaneity (or "the Greenwich Observatory Problem" revisited). In *AIP Conference Proceedings*, 517, 323–326.

London, J. (1999). Hasty's dichotomy. *Music Theory Spectrum*, *21*(2), 260–274.

London, J. (2004). *Hearing in time: Psychological aspects of musical meter*. New York: Oxford University Press.

Lucas, G., Clayton, M., & Leante, L. (2011). Inter-group entrainment in Afro-Brazilian Congado ritual. *Empirical Musicology Review*, *6*(2), 75–102.

Madison, G., & Merker, B. (2002). On the limits of anisochrony in pulse attribution. *Psychological Research*, *66*(3), 201–207.

McClary, S. (1991). *Feminine endings: Music, gender, and sexuality*. Minneapolis: University of Minnesota Press.

McPherson, G. E., Davidson, J. W., & Faulkner, R. (2012). *Music in our lives: Rethinking musical ability, development, and identity.* New York: Oxford University Press.

Miyake, Y., Onishi, Y., & Pöppel, E. (2004). Two types of anticipation in synchronization tapping. *Acta Neurobiologiae Experimentalis*, *64*, 415–426.

Ohriner, M. (2012). Grouping hierarchy and trajectories of pacing in performances of Chopin's mazurkas. *Music Theory Online*, *18* (1), accessed September 19th, 2014. http://www.mtosmt.org/issues/mto.12.18.1/mto.12.18.1.ohriner.php

Ohriner, M. (2011). What can we learn from idiosyncratic performances? Exploring outliers in corpuses of Chopin renditions. In A. Williamon, D. Edwards, & L. Bartel (Eds.), *Proceedings of the International Symposium on Performance Science 2011* (pp. 635–640). Utrecht, The Netherlands: European Association of Conservatories.

Ohriner, M. (2013). Effects of temporal position on harmonic succession in the Bach chorale corpus. In J. Burgoyne, J. Wild, & J. Yust (Eds.), *Mathematics and computation in music* (pp. 167–176). Berlin: Springer Berlin Heidelberg.

Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception*, *25*(5), 489–504.

Palmer, C. (1996). Anatomy of a performance: Sources of musical expression. *Music Perception*, *13*(3), 433–453.

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B.H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, *163*, 226–238.

Phillips-Silver, J., Aktipis, C. A., & Bryant, G. A. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Perception*, *28*(1), 3–14.

Quinn, I., & Mavromatis, P. (2011). Voice-leading prototypes and harmonic function in two chorale corpora. In C. Agon, M. Andreatta, G. Assayag, E. Amiot, J. Bresson, & J. Madereau (Eds.), *Mathematics and computation in music* (pp. 230–240). Berlin: Springer Berlin Heidelberg.

Rasch, R. A. (1979). Synchronization in performed ensemble music. *Acta Acustica united with Acustica*, *43*(2), 121–131.

Repp, B. H. (1990). Composers' Pulses: Science or Art? *Music Perception*, *7*(1), 423–434.

Repp, B. H. (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's ''Träumerei''. *The Journal of the Acoustical Society of America*, *92*, 2546–2568.

Repp, B. H. (1998). The detectability of local deviations from a typical expressive timing pattern. *Music Perception*, *15*(3), 265–289.

Repp, B. H. (2001). Expressive timing in the mind's ear. In R.I. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 185–200). Lisse, The Netherlands: Swets & Zeitlinger.

Repp, B. H. (2002). The embodiment of musical structure: Effects of musical context on sensorimotor synchronization with complex timing patterns. In: W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and performance XIX* (pp. 245–265). New York: Oxford University Press.

Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior*, *35*(4), 355–370.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*(6), 969–992.

Repp, B. H. (2011). Perception-production relationships and phase correction in synchronization with two-interval rhythms. *Psychological Research*, *75*, 227–242.

Rothstein, W. N. (1989). *Phrase rhythm in tonal music*. New York: Schirmer Books.

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, *26*(1), 43–49.

Sapp, C. S. (2008). Hybrid numeric/rank similarity metrics for musical performance analysis. In: *Proceedings of 9th International Conference on Music Information Retrieval (ISMIR 2008)* (pp. 501–506), accessed September 19th, 2014. http://ismir2008.ismir.net/papers/ISMIR2008_240.pdf.

Semjen, A. (2000). Tapping off the beat: A note on the role of preparation in stabilising anti-phase sensory-motor coordination. *Cognitive Processing*, *1*, 89–98.

Sloboda, J. A. (1985). Expressive skill in two pianists: Metrical communication in real and simulated performances. *Canadian Journal of Psychology*, *39*(2), 273–293.

Thompson, W. F. (1989). Composer-specific aspects of musical performance: An evaluation of Clynes's theory of pulse for performances of Mozart and Beethoven. *Music Perception*, *7*(1), 15–42.

Timmers, R., & Honing, H. (2002). On music performance, theories, measurement, and diversity. In M. A. Belardinelli (Ed.), *Cognitive Processing*, *1–2*, 1–19.

Todd, N. P. M. (1985). A model of expressive timing in tonal music. *Music Perception*, *3*(1), 33–57.

Todd, N. P. M. (1989). Towards a cognitive theory of expression: The performance and perception of rubato. *Contemporary Music Review*, *4*(1), 405–416.

Todd, N. P. M. (1999). Motion in music: A neurobiological perspective. *Music Perception, 17*(1), 115–126.

Todd, N. P. M. (2007). The contribution of anthropometric factors to individual differences in the perception of rhythm. *Empirical Musicology Review*, *2*(1), 1–13.

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception*, *28*(1), 59–70.

Tymocsko, D. (2010). *Local harmonic grammar in Western classical music*. Unpublished manuscript, available at http://dmitri.mycpanel.princeton.edu/mozart.pdf.

Tormene, P., Giorgino, T., Quaglini, S., & Stefanelli, M. (2009). Matching incomplete time series with dynamic time warping: an algorithm and an application to post-stroke rehabilitation. *Artificial Intelligence in Medicine*, *45*(1), 11–34.

van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, *28*, 43–66.

Vernon, L. N. (1937). *Synchronization of chords in artistic piano music*. Unpublished doctoral dissertation, University of Iowa, USA.

Vorberg, D., & Wing, A. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of perception and action, Volume 2: Motor skills* (pp. 181–262). London: Academic Press.

Walker, J. (2004, revised 2008). MIDI2CSV [software]. Available from http://www.fourmilab.ch/webtools/midicsv.

White, C. (2014). Changing styles, changing corpora, changing tonal models. *Music Perception, 31*(3), 244–253.

Whittall, A. (1991). Analysis as performance. *Atti Del XIV Congresso Della Societa Internazionale Di Musicologica* , *1*, 654–60.

Will, U. (2011). Coupling factors, visual rhythms, and synchronization ratios. *Empirical Musicology Review*, *6*(3), 180–185.

Will, U., Utter, H., Clayton, M. & Leante, L. (2005). Influence of auditory and visual information on the tapping responses of vocal alap. Paper presented at the *Entrainment Network Conference*, The Ohio State University, Columbus. 5 May, 2005.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, *20*(1), 1–5.

Windsor, W. L., Desain, P., Penel, A., & Borkent, M. (2006). A structurally guided method for the decomposition of expression in music performance. *The Journal of the Acoustical Society of America*, *119*, 1182–1193.

DISCOGRAPHY

Op. 50, No. 1

TAPPING STIMULI

Oborin, L. (1951/2009). *Lev Oborin*. London: APR.

Ogawa, N. (2008). London. Live performance, no longer available on youtube.com.

Poli, R. (2008). Weston, MA. http://www.youtube.com/watch?v=k-DAcnFGpKo.

OTHER RENDITIONS IN DATA SETS

Ashkenazy, V. (1983); Bakst, R. (1960); Chiu, F. (1999); de Groot, C. (2007); Frugoni, O. (1959); Hauser, A. (2004); Norton, E. (2010); Rubinstein, A. (1938); Sztompka, H. (1959); Tomita, Y. (2005).

Op. 63, No. 3

TAPPING STIMULI, MEASURES 1–16 (TRAINING)

Biret, I. (1990). *Complete Mazurkas*. Munich: Naxos.

Morozova, I. (2008). http://www.youtube.com/watch?v=e8BQc5wwQGU

Rubinstein, A. (1965). *Mazurka, Op. 63, No. 3*. Hannover: Deutsche Grammaphon.

TAPPING STIMULI, MEASURES 33–57

Bunin, S. (1987). *Bunin: Chopin and Debussy.* London: Olympia.

Chiu, F. (1999). *Complete Mazurkas*. Los Angeles: Harmonia Mundi USA.

Sztompka, H. (1959). Krakow: Polskie Nagrania Muza.

OTHER RENDITIONS IN DATA SETS

See http://www.mazurka.org.uk/info/discography/#63-3