Article

MUSICÆ SCIENTIÆ

Musicae Scientiae
1–20

DOI: 10.1177/1029864916685928
journals.sagepub.com/home/msx
SAGE

Musicians body sway embodies musical structure and expression: A recurrence-based approach

Alexander P. Demos
University of Connecticut, USA

Roger Chaffin
University of Connecticut, USA

Topher Logan
University of Connecticut, USA

Abstract
Musicians' sway during performance seems to be related to musical structure. However, it has yet to be shown that examples of the relationship are not simply due to chance. Progress has been impeded by three problems: the assumption that musical structure is constant across performances; the complexity of the movements; and the inability of traditional statistical tests to accurately model the multilevel temporal hierarchies involved. We solved these problems in a study of the side-to-side postural sway of two trombonists as they each recorded two performances of each of two solo pieces in each of three different performance styles (normal, expressive, non-expressive). The musicians reported their phrasing immediately after each performance by marking copies of the score. We measured the rate and stability (mean line) of recurrence (self-similarity) and assessed the effect of serial position within a phrase, using mixed linear models to model the nesting of phrases within pieces, within performances, across expressive styles and musicians. Recurrence and stability of recurrence changed systematically across the course of a phrase, producing sinusoidal-like and arch-shaped phrasing contours that differed with the performance style and length of phrase. As long suspected, musicians' expressive movements reflect musical structure.

Keywords
music and movement, music performance, music phrasing, postural sway, quantification analysis, recurrence

Audiences and researchers alike share the intuition that the swaying of musicians as they play reflects the music they are playing (Davidson, 1994, 2005, 2006, 2007, 2009, 2012; Ginsborg, 2009; Leman & Godoy, 2010; MacRitchie, Buck, & Bailey, 2013; Nusseck & Wanderley, 2009;

Corresponding author:
Alexander P. Demos, Department of Psychology, 1007 West Harrison Street (M/C 285), Chicago, IL 60607-7137, USA.
Email: ademos@uic.edu

Palmer, Koopmans, Carter, Loehr, & Wanderley, 2009). Musicians move more than strictly necessary simply to produce the notes, making it useful to distinguish the *sound-producing* movements directly responsible for producing musical sound from *ancillary* movements less directly responsible for sound production (Jensenius, Wanderley, Godøy, & Leman, 2010; Wanderley, 2002). One function of ancillary movements may be to convey information about performers' expressive intentions. Viewers shown video recordings or point-light displays of musicians' movements are able to accurately identify the emotions the musicians were asked to express, and how expressively they intended to play (Dahl & Friberg, 2007; Davidson & Broughton, 2016; Nusseck, & Wanderley, 2009). Musicians may also use ancillary movement to help guide their performance such as when a singer's arm movements delineate the rise and fall of musical tension across the course of a musical phrase (Leman, 2008; Jensenius et al., 2010; Pierce, 2007, p. 1).

To study performers' movements, researchers have adapted tools and techniques from the study of language-based gestures. Movements of interest, such as pointing and nodding, are identified by inspection of a recorded communication and their meaning is determined from the context (McNeill, 2006). Musicians' gestures are often aligned with rhythmic patterns (Wanderley, Vines, Middleton, McKay, & Hatch, 2005) and cluster at musically important locations (Davidson, 2007; Teixeira, Yehia, & Loureiro, 2015). Beyond this, the gestural approach has proved inadequate (Demos, Chaffin, & Kant, 2014; Leman, 2008). The relationship between context and gesture seems less consistent for music than for speech. Each musician has a different repertoire of gestures which they use in different musical contexts, even within the same performance (Davidson, 2012). Davidson (2007, p. 398) concludes that "perhaps ... it is the *quality* of the movements and not the specific movements themselves" (emphasis added) that matters. Similarly, Nusseck and Wanderley (2009, p. 335) conclude that the experience of the audience "is less dependent on the players' particular body motion behaviors than it is on the players' overall relative motion characteristics".

We took a different approach, using concepts and mathematical tools developed for describing and analyzing the behavior of non-linear dynamical systems (Demos et al., 2014; Latash, 2008). Rather than identifying movements of interest and then examining their musical contexts, we started with a musical context and asked whether it was systematically related to movement. Instead of segmenting continuous movements into discrete gestures, we avoided assumptions about what size or type of movement might be meaningful, allowing data from the performance of an entire piece to answer that question. Instead of attaching sensors to particular locations on the performer or instrument, we measured overall body motion using a force plate. Rather than assuming that the musical structure is fixed and affected the musicians' phrasing in the same way in every performance, we asked performers to report their understanding of the phrasing for each performance (Cook, 2013, pp. 182–208). Instead of using traditional inferential statistics, such as ANOVA or multiple regression, we used statistical procedures developed for multi-level longitudinal/short-time series data (Singer & Willett, 2003) combined with measures developed for describing the behavior of non-linear dynamical systems (Abarbanel, 1995; Marwan, 2008). We explain the reason for each of these choices below (also see Demos & Chaffin, in press).

Trombone performance

The first decisions to be made in studying musicians' movements are the related choices of instrument, movement, and measurement device. Researchers have studied singers (Davidson, 2001, 2006) and instrumentalists on clarinet (Caramiaux, Wanderley & Bevilacqua, 2012; Palmer et al., 2009; Teixeira et al., 2015; Wanderley et al., 2005), piano (Clarke & Davidson, 1998; Davidson, 1994, 2002, 2007; MacRitchie et al., 2013), and violin (Davidson, 1993;

Demos, Frank, & Chaffin, 2012). We measured the postural sway of two trombonists as they each played the same two pieces, which they had prepared for performance. In the present study, we examine the relationship between the trombonists' sway and the music, looking for effects of musical phrasing.

One reason for our choice of the trombone in the present study is that it provides a convenient way of separating ancillary from sound producing movement, at least partially. For trombonists, postural sway on the anterior-posterior (AP) axis is more affected by the sound producing movements of the trombone slide, while sway on the medio-lateral (ML) axis is more ancillary. Although AP and ML sway are both part of the same circular motion of the body, they can function relatively independently or not, depending on the task (Balasubramaniam, Riley, & Turvey, 2000). For the trombone, the two directions of sway are relatively independent, probably because of the need to compensate for movements of the trombone slide affects AP more than ML sway (Demos et al., 2014). We measured postural sway in both directions but report only ML sway because we were mainly interested in the relationship between musical phrasing and ancillary movement.

Most studies of musicians' movements have used motion capture, measuring the movement of reflective markers or sensors placed on various parts of the performer or instrument. Some researchers have focused on a single location of particular interest, such as the clarinet bell (Palmer et al., 2009; Teixeira et al., 2015; Wanderley et al., 2005). Other researchers employ data reduction techniques (e.g., principle components analysis) to merge the movement of multiple sensors in three spatial dimensions into a composite, one-dimensional measure of overall movement (MacRitchie et al., 2013; Teixeira, Loureiro, Wanderley, & Yehia, 2014). We measured postural sway because it reflects all of the postural adjustments required to maintain balance whenever any part of the body moves, thus providing a physical, rather than statistical, composite measure of overall movement (Balasubramaniam et al., 2000; Latash, Scholz, & Schöner, 2007; Mochizuki, Duarte, Amadio, Zatsiorsky & Latash, 2006).

Measuring movement

Another decision for researchers is what metric to extract from the raw measurements to obtain useful information about movement. The movements of interest are the wide variety of movements displayed by performers from short head nods and wiggles to long sweeping movements of the arms that appear to trace the contour of the music (Davidson, 2012). All these are captured by measurements of postural sway (Mochizuki et al., 2006). These complex movements are not necessarily captured by traditional time-series methods, such as auto- and instantaneous correlation, because the same movement may be repeated irregularly, at different locations or time-scales. For example, a performer might swoop only at the start of each phrase or theme, or move in different patterns from one phrase to the next. Traditional analysis methods can detect simple regularities but may be less successful for irregular or complex patterns (Marwan, 2008).

Complex patterns that evolve over time may be better captured by recurrence, a measure of self-similarity across multiple time-scales (Marwan, 2008). We used recurrence quantification analysis (RQA), to examine recurrence across entire performances, one performance at a time. In contrast, Teixeira at al. (2014) used instantaneous correlation to measure the similarity of selected gestures across performances, two at a time. RQA provides a plot of the behavior of a system in phase-space, an abstract mathematical representation of the functioning of the system (Abarbanel, 1995, p. 21). The phase-space of any (non-linear) complex system can be reconstructed from measurement of the system on a single dimension because each dimension

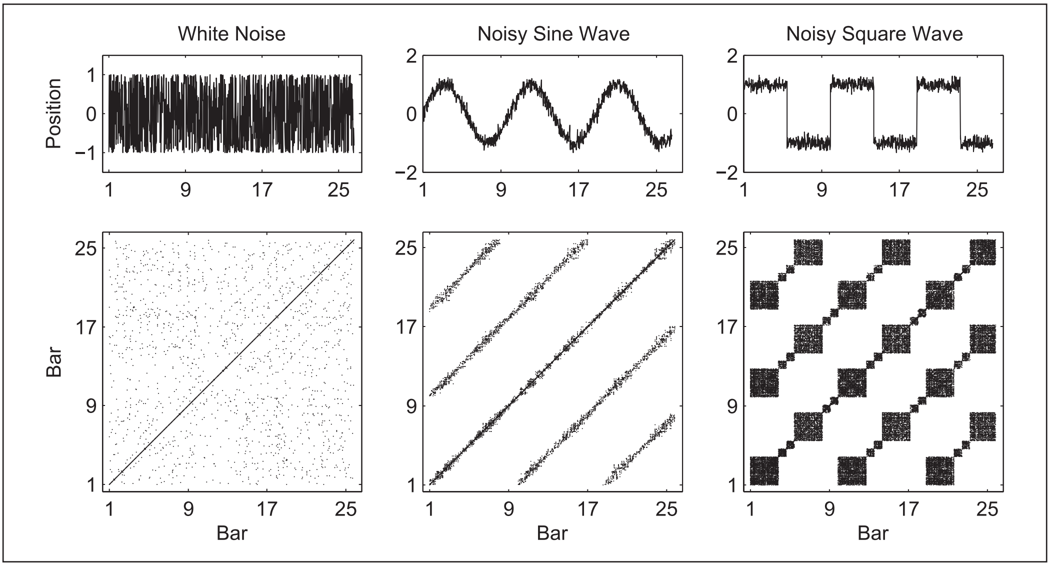

Figure 1. Recurrence quantitation plots (bottom row) illustrating three simple types of movement (white noise/random, sine-wave, and square-wave) that might occur during a musical performance, along with the position data from which they were generated (top row).

contains information about the other dimensions (Takens, 1981). Thus, instead of identifying movement patterns by visual inspection, we simply asked whether recurrence changed reliably across the course of a musical phrase.

Figure 1 shows examples of three RQA plots (bottom row) illustrating three simple types of movement that might occur during a musical performance, along with the position data from which they were generated (top row). The middle column shows a noisy sinewave that might be generated by a performer swaying back and forth. The right-hand column shows a noisy square wave that might be generated by a performer shifting weight back-and-forth from one leg to the other. The left-hand column shows random movement (white noise), providing a baseline against which to compare the more orderly data in the other two columns.

The RQA plots (bottom row) represent similarity in position after phase-space reconstruction. Each dot represents a point of recurrence where position at one point in time overlapped with position at another point in time (in phase-space). Both axes represent elapsed musical time from the start of the performance (in bars). The main diagonal represents a time-lag of zero and so the solid line along the diagonal is the tautological consequence of movements overlapping perfectly with themselves at lag–0. Off-diagonal lines indicate similarity at time lags that increase with distance from the diagonal. The figures are symmetrical, so lines above and below the diagonal are redundant.

Off-diagonal lines are the most informative because they show the temporal structure of the data. Note the difference between the even scattering of data points in the left panel, like a TV with no signal, compared with the bold patterns in the other two panels. Pattern means structure; even scattering means random movement with no structure. Off-diagonal lines indicate the repetition of similar movements at different time lags. In the center and right-hand panels, the spacing between the off-diagonal lines indicates that movements repeated every nine or so bars. If these were real data, we could look at the score to see if they coincide with the repetition of musical elements. In the right-hand panel, we see recurrence at two time-scales. In addition to diagonal lines every nine or so bars, black squares indicate episodes of self-similarity on a shorter time-scale, lasting approximately three bars, such as when a performer shifts weight from one foot to the other.

RQA plots provide a wealth of information that can be quantified by a variety of different metrics that capture different aspects of the patterns (Marwan, Romano, Thiel, & Kurths, 2007). We report the two metrics that provided the clearest relationship to phrasing in our data: rate and stability. The *rate of recurrence* (aka *recurrence*) measures the density of recurrent data points as the proportion of recurrent to non-recurrent data points (0–100%). Recurrence indicates repetitiveness and increases when a musician repeats the same notes, sustains a note, or remains stationary. *Stability of recurrence* (aka *mean line*) is measured by the mean length of the off-diagonal lines, and indicates how long sequences of recurrent movements persist. In Figure 1, stability is lower for the random movements in the left-hand panel than for the more orderly movements in the middle and left-hand panels.

After RQA, the reliability of any apparent relationship between patterns of movement and musical structure must be assessed. This requires that the RQA plots be reduced to a one-dimensional time series for statistical analysis. We collapsed the RQA plots by calculating the rate and stability of recurrence within windows whose size was determined by the musical beats. This facilitated the alignment of RQA metrics (recurrence and stability) with the musical score. The final step was to look for systematic relationships between each metric and the musical phrases using inferential statistics. In summary, the multi-stage process is as follows: 1) phase-space reconstruction, 2) recurrence quantification analysis (RQA), 3) extraction of windowed time series, and 4) statistical analysis by mixed modeling.

Analysis of continuous data

Regardless of the measure chosen, traditional inferential statistics are not well suited to the analysis of music performance data. Music performance typically involves multi-level temporal hierarchies that violate the assumptions of traditional tests based on general linear models (GLM), such as ANOVA and multiple regression. GLM assumes that variance is homogeneous across conditions and that observations are independent, assumptions not usually met by music performance. Variance may not be constant because phrases, sections, and pieces contain different musical material with consequent differences in variability. Observations are not independent because each beat is related to the next and so on up the temporal hierarchy (beats within bars, within phrases, within sections, within pieces). In addition, GLM is restricted to a single level of temporal grouping, limiting examination of the multi-level temporal hierarchies present in music to one level at a time (e.g., Mishra, 2010).

The gestural approach deals with the first two of these problems by arbitrarily segmenting continuous movements into discrete, isolated units, which probably goes some way towards meeting the requirements of GLM. Mostly, researchers have avoided use of inferential statistics, relying instead on description, both qualitative and quantitative, of compelling examples (Davidson, 2009). For example, Teixeira et al. (2014) empirically identified similarity in movements using instantaneous correlation of musicians' "overall motion" profiles created though data reduction methods (i.e., principle components analysis [PCA] of 3-dimensional position data from reflective markers). The authors used inferential statistics (*t*-tests) to show that the duration and velocity of these gestures were lower than normal when playing with a metronome, consistent with previous studies of the relationship between movement and performance style (Davidson & Broughton, 2016). The "strong evidence of the musical significance of the musician's physical movements" was purely descriptive. The authors assert that the gestures described were at locations "where there is more room for expressiveness, due to the greater rhythmic and melodic variation and also to the conclusion of the musical phrase" (Teixeira et al., 2014, p. 9). However, despite the sophisticated statistics, there was no assessment of whether the correspondence between movement and music was due to chance.

In the only study to report a traditional inferential test of the relationship between movement and musical structure, MacRitchie et al. (2013) used ANOVA to evaluate effects of phrasing on the overall movement profiles (PCA of motion capture of the body) of nine pianists performing two Chopin preludes. The motion profiles had wavelike shapes that appeared from visual inspection to correspond with the phrasing reported by the pianists. To show that the correspondence was not due to chance, the researchers measured the distance from the wave maximum in each phrase to the following phrase boundary and compared across performers and phrases (sequentially numbered) in a two-way ANOVA. (The data for the second half of one piece were excluded because the pianists did not agree on the phrasing.) A significant effect for pianists indicated the presence of consistent differences between pianists in where maxima were located in a phrase, earlier for some pianists, later for others. The absence of a comparable effect for phrases suggested that maxima were consistently located at the same point in each phrase, across the different phrases. The authors conclude that "the motion profiles contained repeated periodic patterns for all performers conforming to the underlying phrasing structure for both pieces" (p. 102). Additional support for this conclusion came from lagged autocorrelation functions with peaks at intervals that appeared, from visual inspection, to correspond with phrase boundaries. Despite the ingenious methodology, this visual inspection provides the best evidence of a relationship between movement and phrasing. The null effect of phrasing in the ANOVA demonstrates an absence of difference, not the presence of a reliable relationship, even if the use of ANOVA was warranted, which is doubtful.

Palmer et al. (2009) used functional ANOVA, a non-traditional inferential test, to compare the height of the clarinet bell across the course of three phrases as eight clarinetists played the first eight measures of Mozart's Clarinet Concerto in each of three expressive styles: normal, exaggerated, and inexpressive. Functional ANOVA is a form of functional data analysis that compares different times series by using b-splines to smooth the data (Ramsay, 2006). The clarinet bell rose across the course of a phrase and the analyses showed that the change increased with the expressiveness of the performance. The authors conclude that movement of the clarinet bell was systematically related to phrasing. However, as evidence of a general relationship between movement and phrasing, the data are weak. First, the musical material is limited – three phrases, two of which ended with a beat of rest. Second, inspection of the graph presented suggests that the clarinet bell rose in the first two phrases, but not in the third. Third, the analysis requires the dubious assumption that phrasing remained constant across musicians and conditions. This was not the case in MacRitchie et al.'s (2013) study and there is reason to think that some variation in phrasing is a normal feature of most performance (Cook, 2013, pp. 182–208). In our study, rather than assuming that phrasing would remain constant across performances, we asked the musicians to report the phrasing that they had used immediately after each performance.

Fortunately, recent developments in (short/longitudinal) time-series analysis provide a solution to most of these problems of statistical inference. Mixed-effects models provide a generalized form of multiple regression analysis capable of taking into account the correlated nature of performance data as well as the complex crossing and nesting of temporal hierarchies that routinely occur in music performance as well as the unbalanced number of observations per phrase, section, or piece typical of music (Pinheiro & Bates, 2000; Singer & Willett, 2003). Mixed models allow examination of grouping effects at multiple levels of a musical hierarchy, the interleaving of temporally nested and crossed factors of the sort that occurs when a musician gives multiple performances of the same piece, and differences between musicians and performances in how the music is divided up. This makes it possible to accurately partition variance in the data between various sources (musicians, musical pieces, performances, and levels of musical structure), the better to pinpoint effects of interest, such as phrasing. Mixed models allow testing of "growth curve" serial position effects, such as U-shaped or S-shaped changes across the length of a phrase (Mirman, 2014). Finally, mixed models allow the researcher to treat each independent variable as fixed, random or both, making it possible to increase sensitivity by statistically holding constant variables of less interest (e.g., different lengths of phrase or musical locations). We used mixed-effects models with dependent variables from RQA analysis.

Our study

In summary, we measured the side-to-side (ML) sway of two trombonists as they each performed two solo pieces, twice, in each of three *performance styles* (*normal*, *expressive*, and *non-expressive*), for a total of 24 performances, and used mixed models to look for effects of *phrasing* and performance style. We examined the linear, quadratic (U-shaped) and cubic (S-shaped) components of serial position in a phrase, which we refer to as *components of phrasing*. If movement is unrelated to phrasing, all three components will be non-significant, indicating that the phrasing contour is flat. Effects of any component would indicate that movement and phrasing were systematically related across the various phrases, performances, musicians, and pieces. If we find effects of serial position for rate of recurrence, it will mean that the repetitiveness of movement changes systematically across a phrase. Similarly, if we find effects for stability of recurrence it will mean that the duration of episodes of repetitiveness changes systematically across a phrase. We were also interested to see whether effects of phrasing on movement would depend on performance style.

Method

Musicians and music

The musicians were two male professional tenor trombone players, each with over 25 years of experience, both of whom perform regularly and teach on multiple brass instruments. Both musicians were familiar with the two pieces selected for the study, having taught them to students.

We selected two standards of the trombone literature by Marco Bordogni (1789–1856), transcribed by Joannes Rochut (Rochut, 1928), that differed in musical structure but were of similar length and difficulty, with similar distributions of intervals (Cronbach's $\alpha = .932$). The *more structured* piece, Rochut No. 4, follows a standard ABA form, with a nested question-and-answer structure within each section, and contains 154 beats and 238 notes in F major with a 3/4 meter. The *less structured* piece, Rochut No. 13, is structured as a short fantasy or impromptu, without the larger scale harmonic/melodic structure of a rounded binary or ternary format. It is comprised of four, short thematic sections, and contains 170 beats and 245 notes in E-flat major with a 3/8 meter.

Apparatus

Body movements. We measured postural sway as changes in center of pressure (COP) measurements using a Wii Nintendo Balance Board (Nintendo, Kyoto, Japan). COP reflects movements of head and arms, in addition to the trunk. The Wii Balance Board provides reliable, low cost measurements of COP at 34 Hz (Clark et al., 2010). While lower than in most other studies (e.g., Wanderley et al., 2005), the sampling frequency was more than adequate to capture the frequency of the great majority of the postural sway which was between .25 and 2 Hz. This means we over sampled the data at least 8 times. The Wii Balance Board was connected via Bluetooth to a Dell Inspiron E1505 computer with Matlab 2011b. Matlab interfaced with the Wii Balance Board using WiiLab Toolbox (Ahmed, 2012). Data were collected using the Matlab Psychophysics Toolbox version 3.0 (Brainard, 1997; Kleiner et al., 2007). Data were linearly interpolated to 34 Hz to correct for timing variances and low-pass filtered (Butterworth filter) at 16 Hz. COP was measured in centimeters for medio-lateral sway (ML), i.e., left-to-right. The alignment between the sound recordings and movements were established with the Psychophysics tool box. The location of each note in each performance was located initially by listening, and then more precisely by finding the local minima in the acoustic wave of each performance to determine onset and offset times. This process was repeated twice to increase the reliability of note location. The note locations were used to establish the locations of the beats.

Sound. The performances were recorded simultaneously with the postural data using a Shure microphone and external USB sound mixer (M-Audio) connected to Matlab running the Psychophysics Toolbox. The microphone was placed on a stand approximately 4 feet above the ground, 4 feet from the performer, and 1 foot left of center. The locations of the microphone and the balance board remained constant across performances.

Procedure

We told the musicians that their body movements would be recorded and asked them to stand on the Wii Board without moving their feet, but to otherwise move their bodies in any way they needed. We asked them to prepare the two pieces so that they would be able to perform them fluently from the score when they came to the lab.

The musicians came to the lab separately. Each musician came twice, on two different days, playing only one of the two pieces during each visit. They played the piece selected for the visit six times, twice in each of three *expressive styles* (normal, expressive, and non-expressive), with the two *performances* in each style blocked, i.e., back-to-back. For the normal style, they were asked to play in a way that they considered natural. For the expressive style, they were asked to play with exaggerated expression. Both performers understood this as a direction to exaggerate both dynamics and tempo. For the non-expressive style, they were asked to play with minimal variation in tempo and dynamics, “like a MIDI performance”. Normal performances were always first, to allow the performers to establish a baseline for the other, atypical performance styles. The order of expressive and non-expressive style performances was counterbalanced across pieces and performers.

The musicians were told that they would be asked to report their phrasing by marking it on a copy of the score at the end of each performance. A clean copy of the score was used for each report. The musicians were told that we were measuring their movements, but they were not told specifically that their movements would be related to their reports of phrasing. The analysis of musical structure was provided by the third author, who was one of the musicians and has an MA degree in trombone performance. Two independent music researchers with advanced degrees in music were also asked to mark the locations of the boundaries of the musical structure. For Rochut 4, their markings locations agreed with the musical structure reported by the musician, 100% and 91.67%, collapsing over the upper and lower level boundaries, respectively for each rater. For Rochut 13, the location of musical boundaries reported by the third author agreed with those of the two independent judges for 100% of the top level boundaries and 76.92% and 55.55% of the middle and lower level boundaries respectively. Given the impromptu-like nature of Rochut 13, it is not surprising that musicians differ with respect to the lower level boundaries. The four main melodic sections are so short that further fractionation into sub-phrases is somewhat arbitrary.

Analysis

Step 1: Phase-space reconstruction. Parameters were determined successively for each step of phase-space reconstruction (parameters: time-lag and embedding dimension) followed by RQA (parameter: radius size) (see Abarbanel (1995) and Shockley and Riley (2015) for RQA parameter selection). For phase-space reconstruction, we selected the median lag (42 lags/1.23 seconds) across the 24 performances, using the average mutual information index to identify the first local minimum in each performance, i.e., where the time series was least self-similar. This method makes the extracted components nearly orthogonal. Next, we used false neighbor's analysis to determine the number of embedding dimensions, unfolding the time series into successively higher dimensions until the data points did not overlap spuriously. Finally, we performed the phase-space reconstruction using the median (and mode) number of dimensions ($N = 4$) across the 24 performances.

Step 2: Recurrence quantification analysis (RQA). RQA requires the investigator to select a radius window that determines the distance between data points counted as "near", i.e., recurrent. We used the maximum norm radius and we empirically determined a radius size that ensured that the majority of the performances were near 10% recurrent (Marwan et. al., 2007).

Step 3: Extract time series. We reduced the RQA data for each performance to a one-dimensional time series at the level of musical beats, using Soundforge 9.0 to measure the elapsed time from the beginning of the performance to the local minimum in the acoustic wave corresponding to the onset of each note that fell on the beat. Finally, we computed the rate and stability (mean line) of recurrence of ML sway for each beat.

Step 4: Mixed models. The phrasing reported by the musicians for each performance was coded continuously for each phrase as the phrase length, measured by the number of beats in the phrase expressed as a Z-score, and for each beat as serial position from the start of the phrase, expressed as the percentage of the phrase completed. We refer to phrase length and serial position together as "phrasing" and to effects of serial position on the dependent measures as "phrasing contours".

We used mixed models to examine the effects of phrasing on each dependent measure (Singer & Willett, 2003), using the LME4 package in R (Bates, Maechler, Bolker, & Walker, 2015) following the procedures of Barr, Levy, Scheppers, and Tily (2013). Each model was constructed in an identical manner using the same random effects: serial position (linear quadratic, and cubic), phrase length, performance style, musician, and performance (1st or 2nd). Serial position in a phrase was entered into the models using orthogonal polynomials (Mirman, Dixon, & Magnuson, 2008). In order to interpret the effects for the polynomials, we generated graphs from the models.

We performed separate analyses for each metric of postural sway (recurrence and stability). For each analysis, we examined two models, using a forward modeling procedure. The first model examined the effects of serial position, performance style, and phrase length treating them as fixed effects (predictors). The second model added the interactions of phrase length with the other predictors. The random effects in both models were the same as the fixed effects together with musician, piece, and trial, which were included only as random effects. Model fit was assessed by deviance testing.

We included phrase length, rather than musical piece, as a predictor because preliminary analyses showed that phrase length accounted for differences between the two pieces. We examined the effects of normal and expressive performances styles by comparing each style, separately, to a baseline provided by the non-expressive performances. Values above and below the non-expressive baseline were indicated by positive and negative effects respectively. In modeling stability (mean line), we partialed out recurrence (using linear regression), in order to be able to draw conclusions about stability that were independent of amount of recurrence. (Note: the mean of the mean line values was added back into the stability values shown in Table 3 and Figure 5.)

Results

Phrasing

Figure 2 shows an excerpt of each of the pieces with the phrase markings (shown as slurs) for each musician the first time they performed each piece in the normal style. As can be seen in this excerpt, Musician 1 used shorter phrases than Musician 2.

Table 1 shows the number of phrases into which the musicians divided the musical score in each performance. The most striking feature of the data is the wide range of variation, both within and between musicians and between pieces. Phrasing was mostly consistent for the two performances of the same piece in the same style by the same musician but varied widely across musicians and performance styles. The only reliable difference was between musicians for the more structured piece, for which Musician 1 reported more phrases than Musician 2, $t(10) = 3.79$, $p < .05$.

Phrasing was based on the musical structure. Table 2 shows that most phrases started at a structural boundary and that the two musicians differed in their use of the musical structure. Musician 2 (who provided the report of the musical structure) started phrases at musical boundaries more often than Musician 1 for both pieces, $t(10) = 2.43$, $p < .05$ and $t(10) = 2.75$, $p < .05$, for the more and less structured pieces respectively.

We examined the percentage of boundaries marked as the start of a phrase at each level of musical structure. For the more structured piece, the musicians placed phrase boundaries that coincided with an average of 96.67%, 95.83%, and 63.33% of the top, middle, and bottom level boundaries, respectively across all performances and styles. For the less structured piece, the musicians placed phrase boundaries that coincided with an average of 100% and 60.83% of top and bottom levels boundaries, respectively. In summary, both musicians based their phrasing on the musical structure, but differed in their use of lower-level musical boundaries. Anticipating the possibility of such variation, we asked them to report their phrasing after each performance and used their reports in examining the relationship between phrasing and movement.

Postural sway

Recurrence quantification of postural sway

Figure 3 shows the postural sway of one of the musicians for three performances of the same piece, one in each expressive style. The figures illustrate characteristic patterns present in many

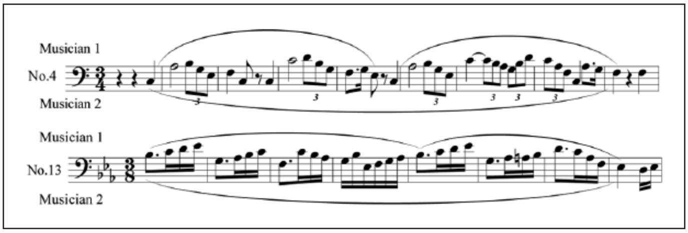

Figure 2. Excerpts from Rochut No. 4 and No. 13 with phrase markings for each musician indicated as slurs for the first performance in the normal style.

Table 1. Number of phrases reported by two musicians in two normal, expressive, and non-expressive performances of two pieces.

Piece	Musician	Performance style & Performance					
		Normal		Expressive		Non-expressive	
		1st	*2nd*	*1st*	*2nd*	*1st*	*2nd*
More-structured	1	12	13	17	16	23	24
	2	8	8	14	5	8	8
Less-structured	1	12	12	19	18	9	9
	2	12	12	9	9	9	9

Table 2. Percentage of phrases coinciding with a boundary in the musical form by two musicians in two normal, expressive, and non-expressive performances of two pieces.

Piece	Musician	Performance style & Performance					
		Normal		Expressive		Non-expressive	
		1st	*2nd*	*1st*	*2nd*	*1st*	*2nd*
More-structured	1	100.00	92.31	70.59	75.00	47.83	50.00
	2	100.00	100.00	78.57	100.00	100.00	100.00
Less-structured	1	83.33	83.33	55.56	58.82	88.89	88.89
	2	100.00	100.00	100.00	77.78	100.00	100.00

of the 24 performances. The top row shows position across the entire performance. Recurrence plots for these data are shown for the entire piece in the middle row and for the first nine bars in the bottom row. Each column shows a different expressive style.

The most important feature of Figure 3 is the presence of patterns in the recurrence plots. The irregular diagonal lines for the normal performance (center) signify short periods of smooth oscillation, as in the center RQA plot in Figure 1. The irregular horizontal lines for the expressive performance (right) signify short periods, at varying time intervals, during which the same position recurred. For both styles, there are also square box-shapes along the diagonal, as in the right-hand plot in Figure 1, indicating that the performer maintained the same position for periods of up to one bar. Of the three performance styles, the non-expressive performance most resembles the white noise in the left-hand plot in Figure 1. The non-expressive plot

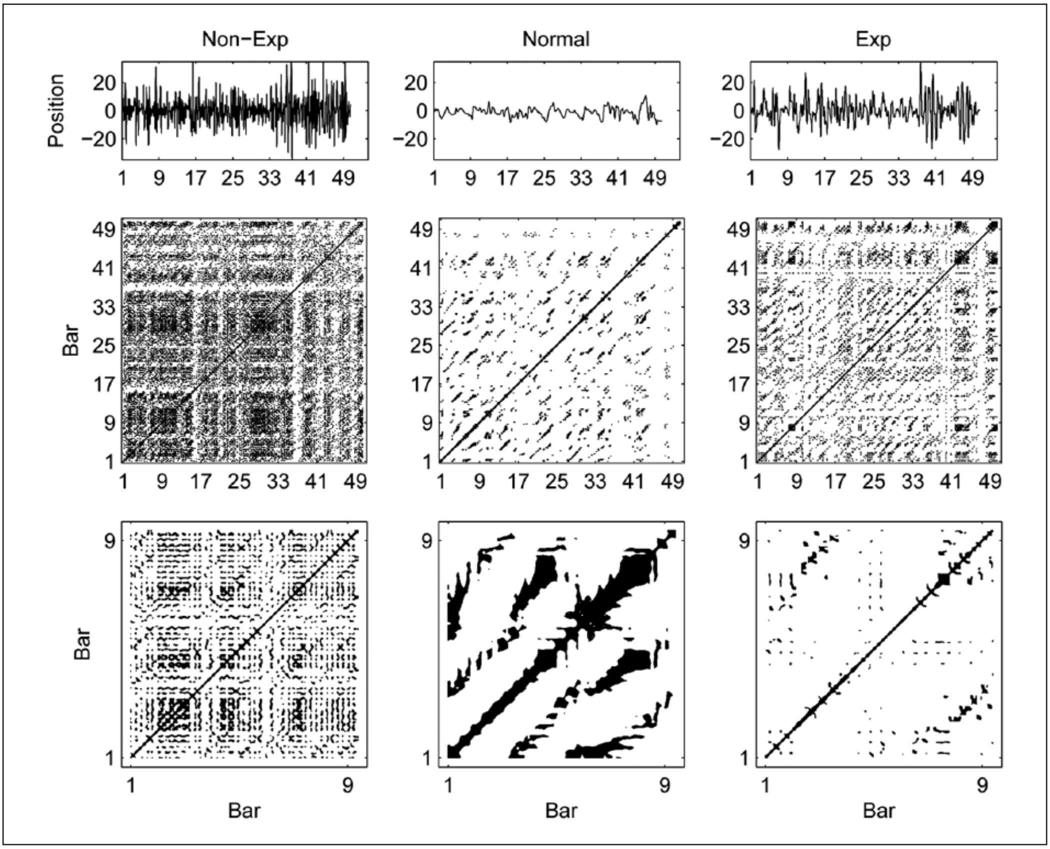

Figure 3. Postural sway of the same musician performing the same piece in three expressive styles for the more structured piece, showing position (top row), the full recurrence quantification of the whole performance (middle row), and a magnified view of the first 9 bars extracted from the whole performance (bottom row).

also appears darker than the normal and expressive plots, indicating the presence of more recurrent data points. The source of these differences is evident in the top row of Figure 3, where the plots of position show that movements in the non-expressive performance were shorter and more sudden, while those in the normal performance were longer and more continuous, with the expressive performance midway between.

The recurrence quantification plots in Figure 3 were then windowed into beats and means extracted within each window (without any overlap), separately for recurrence rate and stability. Figure 4 shows the results of this windowing for the three performances shown in Figure 3. These unidimensional vectors were then analyzed using mixed-effects modeling.

Modeling recurrence and stability

Table 3 shows the results of the mixed-effects modeling of each metric of postural sway (recurrence and stability). The first model examined the effects of serial position, performance style, and phrase length, and the second model added the interactions with phrase length. The

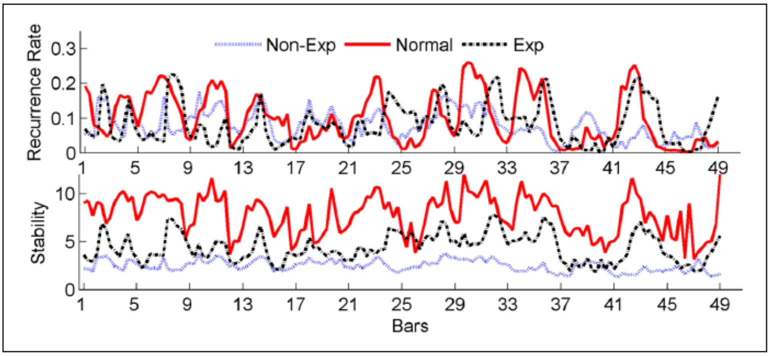

Figure 4. The recurrence rate and stability extracted from the data featured in Figure 3 of the same musician performing the more structured piece in three expressive styles.

addition of the interactions with phrase length in Model 2 significantly improved the fit of the model for both rate and stability due to the presence of three-way interactions indicating that the effects of phrasing differed with the performance style and length of phrase. Therefore, we limit our description of the results to Model 2 for each metric, describing effects in order of their appearance in Table 3. Model 2 includes only simple effects, not main effects, because performance style was involved in every effect and was dummy coded. In both models, the effects of performance style were evaluated by comparing normal and expressive with non-expressive performances. As a reminder, we have italicized "*Non-Expressive*" when labeling this simple effect in Table 3.

To aid understanding of Model 2, Figure 5 displays the model fitted data generated by the model (fixed and random effects). The figure is based on data from all the phrases, in both pieces, for all performances, by both musicians. Serial position effects (and their interactions) from Table 3 are represented by the *phrase contours* shown in the figure, which show how movement changed across the course of a phrase. The two-way interactions between serial position (continuous) and style (categorical) are best seen in the middle panels of Figure 5 (all phrases). The three-way interactions are shown in the left and right panels which were created by dichotomizing phrase length, a continuous measure. Short phrases (13 or fewer beats [80% of phrases]) are shown in the left-hand panel of Figure 5, long phrases (14 or more beats) in the right-hand panel.

Performance style. Sway was significantly more recurrent and less stable in non-expressive performances than in normal and expressive performances. This is indicated by the significant effects for both performance styles which were negative for recurrence (-5.11 and -6.46 respectively) and positive for stability (0.92 and 1.39 respectively). We have already noted these effects in the RQA plots in Figure 3, in which the plot for the non-expressive performance appears darker and the data points more evenly scattered, indicating that the non-expressive performance was more recurrent and less stable. The significant effects of performance style in Table 3 indicate that this was not an isolated case; similar differences occurred throughout the data, more frequently than expected by chance. Thus, there was less self-similarity in normal and expressive performances than non-expressive performances, but when self-similarity did occur it was more stable (lasted longer) in normal and expressive than in non-expressive performances.

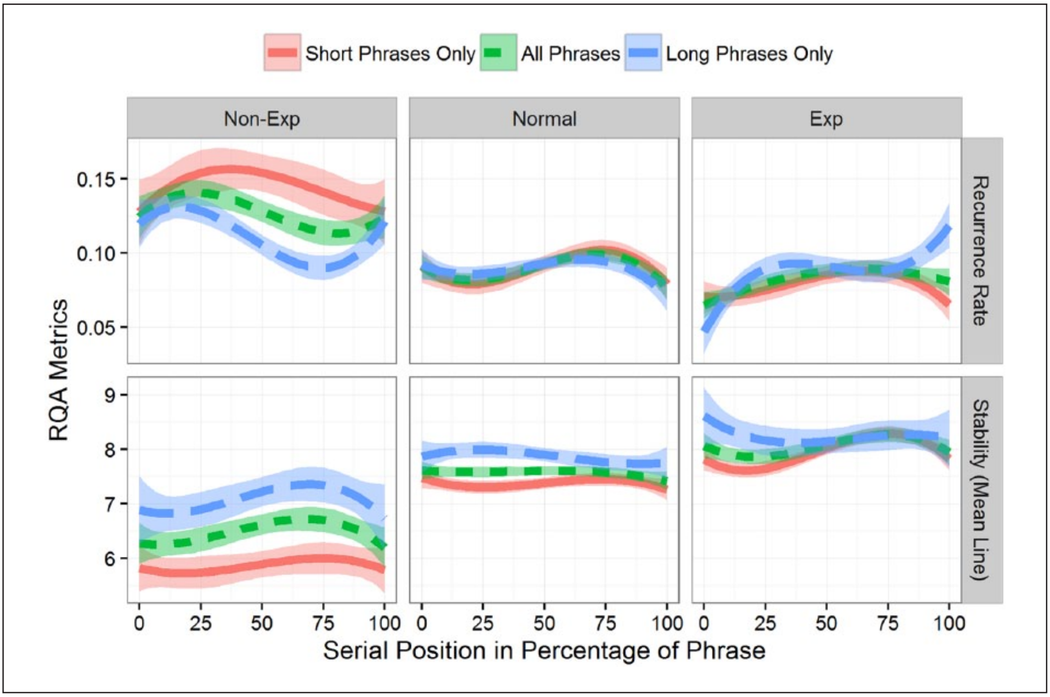

Figure 5. Rate and stability of recurrence of postural sway over serial position: phrasing contours for short, long, and for all (short + long) phrases, separately for non-expressive (Non-Exp), normal, and expressive (Exp) performances.

Phrase length by style. Stability increased with length of phrase and this effect was significantly larger in non-expressive than in normal and expressive performances. This is indicated by the significant positive effect of phrase length for non-expressive performances (1.01) and the significant negative interactions of phrase length for the normal and expressive performances (-1.14 and -1.38 respectively). For rate of recurrence, the interaction of phrase length and style was not significant. Thus, the length of periods of self-similar movement increased with phrase length in non-expressive performances, but not in normal and expressive performances.

Serial position by style. Phrasing contours were significantly curved for both recurrence and stability; the contours differed with the performance style for recurrence but not for stability. The differences between performance styles are best seen in the green short-dashed lines in Figure 5, representing all phrases. For recurrence, the contours were S-shaped for non-expressive and normal performances and arched for expressive performances. This is reflected in Model 2 by significant cubic effects for non-expressive and normal performances and a significant quadratic effect for expressive performances. The cubic effects for non-expressive and normal performances were in opposite directions, reflecting differences in the location of the peaks. There were also significant linear effects, positive for normal and expressive performances and negative for non-expressive performances, indicating that recurrence increased across most of the phrase for the former and decreased for the latter. Thus, the amount of self-similarity in the trombonists' postural sway changed across the course of a phrase and the pattern of change differed with the style of the performance.

Table 3. Forward fitted mixed-effects models of serial position, performance style, and phrase length, separately for recurrence and stability.

Fixed effects	Recurrence Model 1		Recurrence Model 2		Stability Model 1		Stability Model 2	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Style								
Intercept value *Non-Expressive (Non-Exp)*	34.31***	(2.21)	33.60***	(2.53)	6.21***	(0.54)	6.63***	(0.56)
Normal (Norm)	−5.73***	(1.41)	−5.11**	(1.96)	1.35***	(0.17)	0.92***	(0.23)
Expressive (Exp)	−7.51***	(1.58)	−6.46**	(2.21)	1.92***	(0.19)	1.39***	(0.25)
Phrase Length (PL) by Style								
PL: *Non-Exp*	0.64	(1.06)	−1.15	(3.43)	–0.10	(0.07)	1.01*	(0.40)
PL: Normal			1.66	(3.54)			−1.14**	(0.41)
PL: Exp			2.82	(3.82)			−1.38**	(0.45)
Serial Position (SP) by Style								
SP: *Non-Exp*	−43.42	(25.37)	−70.48*	(29.81)	4.54	(2.61)	3.07	(3.11)
SP^2: *Non-Exp*	−13.61	(22.73)	6.32	(27.99)	–2.49	(2.51)	−6.07	(3.10)
SP^3: *Non-Exp*	43.84*	(19.06)	72.89**	(22.45)	–5.75**	(2.20)	−6.97**	(2.68)
SP: Norm	81.48***	(23.54)	97.26***	(26.25)	–6.13	(3.20)	−4.60	(3.53)
SP^2: Norm	−26.41	(23.53)	−44.96	(26.40)	2.57	(3.21)	6.00	(3.55)
SP^3: Norm	−121.85***	(23.10)	−145.4***	(25.75)	4.17	(3.14)	4.87	(3.50)
SP: Exp	76.81**	(24.15)	121.49***	(28.09)	0.20	(3.23)	−1.57	(3.67)
SP^2: Exp	−44.68	(23.87)	−66.99*	(28.28)	1.08	(3.20)	5.01	(3.68)
SP^3: Exp	−47.96*	(23.12)	−41.76	(26.41)	–1.43	(3.11)	−0.38	(3.51)
Serial Position by Style by Phrase Length								
SP: *Non-Exp*: PL			−101.00	(52.30)			−1.56	(6.29)
SP^2: *Non-Exp*: PL			72.94	(50.21)			−14.03*	(6.11)
SP^3: *Non-Exp*: PL			101.70*	(43.58)			−4.88	(5.50)
SP: Norm: PL			61.64	(48.40)			−1.58	(6.20)
SP^2: Norm: PL			−93.93*	(47.26)			16.31**	(6.08)
SP^3: Norm: PL			−84.76	(43.78)			8.27	(5.78)
SP: Exp: PL			156.92**	(56.05)			−13.93*	(6.97)
SP^2: Exp: PL			−88.62	(54.63)			16.26*	(6.81)
SP^3: Exp: PL			18.04	(48.47)			9.39	(6.17)
Goodness of fit								
Deviance	26758.88		26716.78		12396.52		12348.62	
AIC	26830.87		26810.78		12468.53		12442.62	
BIC	27053.06		27100.86		12690.71		12732.69	
X^2 (*df*)	– (36)		42.09*** (47)		– (36)		47.91*** (47)	

* $p < .05$, ** $p < .01$, *** $p < .001$.

For stability, the phrasing contour was S-shaped for non-expressive performances, while the smaller fluctuations for normal and expressive performances did not differ from chance. This is indicated in Model 2 by the significant cubic effect for non-expressive performances and the absence of significant effects for normal and expressive performances. Thus, the duration of episodes of self-similarity peaked in the second half of a phrase in non-expressive performances, and remained constant across a phrase for performances in other styles.

Serial position by style by phrase length. Phrasing contours also varied with length of phrase, creating three-way interactions represented by the red solid and blue long-dashed lines in Figure 5. Recurrence and stability were affected in different ways. For recurrence, phrasing contours were more S-shaped in long phrases for non-expressive performances and more U-shaped in the first half of the phrase for normal performances. For expressive performances, the contour sloped upward over the course of long phrases and downward over the course of short phrases. These differences between the three styles are reflected in the direction and size of each of the components of phrasing and, to a lesser extent, in which component was significant. Thus, the amount of self-similarity changed across a phrase in complex ways that depended on the performance style and length of phrase. Phrasing contours were more pronounced in long phrases for non-expressive and expressive performances. For normal performances, the effects of phrase length were in the opposite direction and smaller.

For stability, the contours again varied with style and phrase length, but were more similar than for rate of recurrence. For non-expressive and expressive performances, there were arch-shaped contours that peaked in the second half of the phrase. These arches were more pronounced in long phrases for non-expressive performances and in short phrases for expressive performances. In normal performances, the contours were flatter and peaked in the first half of long phrases and in the second half of short phrases. Again, the differences are reflected in the direction and size of each of the components of phrasing and, to a lesser extent, in which component was significant. Thus, as for rate of recurrence, phrasing contours for stability of recurrence depended on the performance style and length of phrase and were more pronounced in long phrases and more pronounced for non-expressive and expressive performances than for normal performances.

Discussion

The musicians' sway changed systematically across the course of a phrase in complex but orderly ways, confirming the widely-held view that musicians' movements reflect the music that they are playing (Davidson & Broughton, 2016; Ginsborg, 2009; MacRitchie et al, 2013; Nusseck & Wanderley, 2009; Palmer et al., 2009). Also, playing with more or less expression affected the way the musicians moved, as in previous studies (Davidson, 1994; Teixeira et al., 2014). Unlike previous studies, we avoided making assumptions about the size and types of movement of interest. Instead, we asked whether phrasing had consistent effects on patterns of movement across entire performances. Phrasing did have consistent effects. However, the two musicians used different phrasing (for one piece), and effects of phrasing changed with performance style and length of phrase. This complexity explains the difficulty that researchers have had identifying consistent relationships between movement and music (Davidson, 2007; Wanderley et al., 2005).

Our most important finding is that the relationship of phrasing and movement was statistically reliable. Our methods differed in three ways from previous studies that identified relationships between movement and phrasing but did not evaluate the possibility that they were due to chance (Davidson, 2009; MacRitchie et al., 2013; Palmer et al., 2009; Teixeira et al. 2014; Wanderley, 2002; Wanderley et al., 2005). First, we examined recurrence. Second, we asked the musicians to report their phrasing for each performance and used their reports to examine their movements. Third, we used mixed models to assess the reliability of the recurrence metrics. We discuss each point in turn.

First, recurrence quantification analysis (RQA) provided an effective method of quantifying the patterning of body movement during performance. Instead of asking whether musicians moved up or down, or more or less, across the course of a phrase, we asked whether they made the kinds of complex movement gestures identified in previous research (Davidson, 2002, 2007, 2012; Wanderley et al., 2005). Recurrence is a more abstract metric than raw position or velocity of body sway, and provides a more global measure of the patterning of movement. Further studies are required to better understand how recurrence can be used to examine the relationship between music and movement. We used only two of the many possible metrics of recurrence; others may be useful for other instruments, pieces, and performance contexts (Marwan et al., 2007). We measured postural sway, as have others (see Shockley & Riley, 2015); other types of movement and other body parts can also be analyzed this way.

Second, asking the musicians to report phrasing after each performance was important because each musician used different phrasing, suggesting that the musical structure and the phrasing the musician can adopt can vary for each performance (Cook, 2013, pp. 182–208). While there was high agreement on musical structure between the performer and independent raters for one of the pieces, agreement for the other was lower (between 56% and 77%). Our anticipation of this kind of ambiguity is the reason that we relied on each performer's account of their own phrasing. If we had not done so, the extent of the differences between the two musicians makes it unlikely that we would have found a systematic relationship between movement and phrasing. While overall structure is important for a musician's understanding and interpretation of a piece, the differences in phrasing reported by the musicians for each performance style suggest that musical expression affects phrasing and that changing the expression can alter the choice of phrasing.

Third, mixed models allowed us to respect the performers' reporting of the varied number of phrases and phrase lengths in the two pieces and to test simultaneously for linear, quadratic and cubic components of phrasing on the recurrence metrics. This allowed us to assess the reliability of effects across entire performances, while respecting the temporal structure of the data and avoiding the need to single out particular passages or to treat phrases in isolation from their neighbors.

The effects of performance style on the musicians' movements were broadly consistent with previous reports that musicians' movements change with their style of playing (Davidson, 1994; Teixeira et al., 2014; Wanderley et al., 2005). Movement was less organized in non-expressive performances than in normal and expressive performances. Paradoxically, movements in the non-expressive performances were more strongly related to phrasing. Phrasing contours were more exaggerated and effects of phrase length larger in the non-expressive than in normal and expressive performances. One possible resolution of the paradox is suggested by the musicians' report that playing non-expressively was difficult. We speculate that playing in an atypical, unpracticed style disrupts familiar movement patterns, resulting in less organized movements.

In summary, the trombonists' swaying movements delineated the musical phrases they were playing. It is intuitively obvious to even a casual observer that movement and music are intimately related. However, it has been difficult to provide evidence of this relationship that meets standards of scientific rigor. The complex interactions that we found help to explain why. Until recently, psychology relied primarily on statistical procedures not well suited to this kind of complexity or to the continuous measurements needed for music performance. By using a combination of techniques, we were able to show that the belief that movement and music are closely related is well founded. The conclusions of our science now match those of our senses.

Acknowledgements

Alexander P. Demos is now at the Department of Psychology, University of Illinois at Chicago. We thank our trombonists for their time and musical advice; Kerry Marsh, Whitney Tabor, and Vivek Kant for theoretical advice; Rita Aiello and Frances Spidle for musical advice; Till Frank, Jay Dixon, Pyeong Whan Cho, and Stephanie Del Tufo for advice about the analyses; and Mary Crawford for suggestions about the manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

Abarbanel, H. D. I. (1995). *Analysis of observed chaotic data*. New York, NY: Springer.

Ahmed, A. (2012). WiiLab [Matlab Toolbox]. Retrieved from http://www.colorado.edu/neuromechanics/research/wii-balance-board-project

Balasubramaniam, R., Riley, M. A., & Turvey, M. T. (2000). Specificity of postural sway to the demands of a precision task. *Gait & Posture*, *11*(1), 12–24.

Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Maechler, M., Bolker, B. M., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Caramiaux, B., Wanderley, M. M., & Bevilacqua, F. (2012). Segmenting and parsing instrumentalists' gestures. *Journal of New Music Research*, *41*(1), 13–29.

Cook, N. (2013). *Beyond the score: Music as performance*. New York, NY: Oxford University Press.

Clark, R. A., Bryant, A. L., Pua, Y., McCrory, P., Bennell, K., & Hunt, M. (2010). Validity and reliability of the Nintendo Wii Balance Board for assessment of standing balance. *Gait & Posture*, *31*(3), 307–310.

Clarke, E. F., & Davidson, J. W. (1998). The body in performance. In W. Thomas (Ed.), *Composition-performance-reception: Studies in the creative process in music* (pp. 74–92). Aldershot, UK: Ashgate Press.

Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, *24*(5), 433–454.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, *21*(2), 103–113.

Davidson, J. W. (1994). What type of information is conveyed in the body movements of solo musician performers? *Journal of Human Movement Studies*, *6*, 279–301.

Davidson, J. W. (2001). The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox. *Musicae Scientiae*, *5*(2), 235–256.

Davidson, J. W. (2002). Understanding the expressive performance movements of a solo pianist. *Musikpsychologie*, *16*, 7–29.

Davidson, J. W. (2005). Body communication in musical performance? In D. Miell, R. MacDonald, & D. Hargreaves (Eds.), *Musical communication* (pp. 215–238). New York, NY: Oxford University Press.

Davidson, J. W. (2006). "She's the one": Multiple functions of body movement in a stage performance by Robbie Williams. In A. Gritten & E. King (Eds.), *Music and gesture* (pp. 208–226). Burlington, VT: Ashgate.

Davidson, J. W. (2007). Qualitative insights into the use of expressive body movement in solo piano performance: A case study approach. *Psychology of Music*, *35*(3), 381–401.

Davidson, J. W. (2009). Movement and collaboration in music performance. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 364–376). Oxford, UK: Oxford University Press.

Davidson, J. W. (2012). Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies. *Psychology of Music*, *40*(5), 595–633.

Davidson, J. W., & Broughton, M. C. (2016). Bodily mediated coordination, collaboration, and communication in music and performance. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 615–631). Oxford, UK: Oxford University Press.

Demos, A. P., & Chaffin, R. (in press). Removing obstacles to the analysis of movement in musical performance: Recurrence, mixed models, and surrogates. In M. Lesaffre, P.-J. Maes, & M. Leman (Eds.), *The Routledge companion to embodied music interaction*. New York, NY: Routledge.

Demos, A. P., Chaffin, R., & Kant, V. (2014). Toward a dynamical theory of body movement in musical performance. *Frontiers in Psychology*, *5*(477). doi:10.3389/fpsyg.2014.00477

Demos, A. P., Frank, T., & Chaffin, R. (2012). Understanding movement during performance: A recurrence quantization approach. In A. Williamon, D. Edwards, & L. Bartel (Eds.), *Proceedings of the International Symposium on Performance Science 2011* (pp. 259–264). Utrecht, the Netherlands: European Association of Conservatoires (AEC).

Ginsborg, J. (2009). Beating time: The role of kinaesthetic learning in the development of mental representations for music. In A. Mornell (Ed.) *Art in motion* (pp. 121–142). Vienna, Austria: Peter Lang.

Latash, M. L. (2008). *Synergy*. New York, NY: Oxford University Press.

Latash, M. L., Scholz, J. P., & Schöner, G. (2007). Toward a new theory of motor synergies. *Motor Control*, *11*(3), 276–308.

Leman, M. (2008). *Embodied music cognition and mediation technology*. Cambridge, MA: MIT Press.

Jensenius, A. R., Wanderley, M., Godøy, R. I, & Leman, M. (2010). Musical gestures: Concepts and methods in research. In R. I. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 12–35). New York, NY: Routledge.

Leman, M., & Godøy, R. I. (2010). Why study musical gestures. In R. I. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 3–11). New York, NY: Routledge.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(14), 1–16.

MacRitchie, J., Buck, B., & Bailey, N. J. (2013). Inferring musical structure through bodily gestures. *Musicae Scientiae*, *17*(1), 86–108.

Marwan, N. (2008). A historical review of recurrence plots. *The European Physical Journal Special Topics*, *164*(1), 3–12.

Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, *438*(5–6), 237–329.

McNeill, D. (2006). Gesture: A psycholinguistic approach. In E. Brown & A. Anderson (Eds.), *The encyclopedia of language and linguistics* (pp. 58–66). Amsterdam, the Netherlands: Elsevier.

Mochizuki, L., Duarte, M., Amadio, A. C., Zatsiorsky, V. M., & Latash, M. L. (2006). Changes in postural sway and its fractions in conditions of postural instability. *Journal of Applied Biomechanics*, *22*(1), 51–60.

Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, FL: CRC Press.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, *59*(4), 475–494.

Mishra, J. (2010). Effects of structure and serial position on memory errors in musical performance. *Psychology of Music*, *38*(4), 447–461.

Nusseck, M., & Wanderley, M. M. (2009). Music and motion – How music-related ancillary body movements contribute to the experience of music. *Music Perception*, *26*(4), 335–354.

Palmer, C., Koopmans, E., Carter, C., Loehr, J. D., & Wanderley, M. (2009). Synchronization of motion and timing in clarinet performance. In A. Williamon, S. Pretty, & R. Buck (Eds.), *Proceedings of the Second International Symposium on Performance Science* (pp. 159–164). Utrecht, The Netherlands: European Association of Conservatoires (AEC).

Pierce, A. (2010). *Deepening musical performance through movement: The theory and practice of embodied interpretation*. Bloomington, IN: Indiana University Press.

Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-PLUS*. Berlin, Germany: Springer.

Ramsay, J. O. (2006). *Functional data analysis*. New York, NY: John Wiley & Sons.

Rochut, J. (1928). Melodious etudes for trombone: Selected from the vocalises of Marco Bordogni). New York, NY: Carl Fischer Music.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Methods for studying change and event occurrence*. New York, NY: Oxford University Press.

Shockley, K., & Riley, M. A. (2015). Interpersonal couplings in human interactions. In C. L. Webber, Jr., & N. Marwan (Eds.), *Recurrence quantification analysis: Theory and best practices* (pp. 399–421). Berlin, Germany: Springer.

Takens, F. (1981). Detecting strange attractors in turbulence. In D. A. Rand & L.-S. Young (Eds), *Dynamical systems and turbulence: Lecture notes in mathematics* (Vol. 898, pp. 366–381). Berlin, Germany: Springer.

Teixeira, E. C., Loureiro, M. A., Wanderley, M. M., & Yehia, H. C. (2014). Motion analysis of clarinet performers. *Journal of New Music Research*, *44*(2), 97–111.

Teixeira, E. C., Yehia, H. C., & Loureiro, M. A. (2015). Relating movement recurrence and expressive timing patterns in music performances. *The Journal of the Acoustical Society of America*, *138*(3), 121–216.

Wanderley, M. M. (2002). Quantitative analysis of non-obvious performer gestures. In I. Wachsmuth & T. Sowa (Eds.), *Gesture and sign language in human-computer interaction: International Gesture Workshop* (pp. 241–253). Berlin, Germany: Springer.

Wanderley, M. M., Vines, B. W., Middleton, N., McKay, C., & Hatch, W. (2005). The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research*, *34*(1), 97–113.