Sensitivity to Meter in Auditory Feedback During Music Performance.

Article *in* Psychomusicology: Music, Mind, and Brain · January 2017

DOI: 10.1037/pmu0000166

2 authors:

Peter Pfordresher
University at Buffalo, The State University of …
63 PUBLICATIONS **1,089** CITATIONS

Anastasiya Kobrina
University at Buffalo, The State University of …
5 PUBLICATIONS **0** CITATIONS

Psychomusicology: Music, Mind, and Brain
2017, Vol. 27, No. 1, 54–62

0275-3987/17/$12.00 http://dx.doi.org/10.1037/pmu0000166

Sensitivity to Meter in Auditory Feedback During Music Performance

Peter Q. Pfordresher and Anastasiya Kobrina
University at Buffalo, State University of New York

A long-standing question in research on the role of auditory feedback during production (including music performance) concerns the role of feedback content, which in music refers to pitch categories. Whereas Howell and colleagues claimed that producers are only sensitive to rhythmic synchronization between actions and sound (Howell, Powell, & Khan, 1983), serial shifts of pitch content (e.g., hearing the previous planned pitch at each keypress) disrupt piano performance (Pfordresher, 2005). However, some results suggest that the basis of disruption from altered pitches may relate to higher-order rhythms conveyed by the pattern of melodic accents in auditory feedback that communicate meter (cf. Jones, 1987). Thus, we tested whether participants would be disrupted while hearing a feedback sequence with a conflicting meter. In 2 experiments, participants performed either binary or ternary meter melodies while hearing auditory feedback that could be altered with respect to its metrical organization (thus forming a different pitch sequence), and/or its sequential alignment with the planned melody. Auditory feedback with conflicting metrical accents disrupted performance as did serial shifts of the planned melody. Thus, the disruptive effect of altered pitch may reflect rhythmic organization, albeit in a different sense than was suggested by Howell and colleagues.

Keywords: music performance, auditory feedback, meter, melodic contour, perception and action

Performance of a complex sequential task, such as speech or music, occurs simultaneously with its perceived outcome, known as auditory feedback. The importance of auditory feedback to the fluency of performance is demonstrated by the disruptive effects of altered auditory feedback (AAF). For instance, it has been know for decades that sequence production (in music and speech) can be disrupted by delayed auditory feedback (DAF), a slight asynchrony between actions and resulting sounds (for reviews see Finney, 1999; Howell, 2004; Pfordresher, 2006; Yates, 1963). Fluency in performance clearly relies on temporal coordination between planned actions and concurrent sounds. It is less clear whether performance relies on the *content* of auditory feedback, which in a musical context refers to the pitches associated with individual notes.

For some time, it appeared as though performance may not rely on feedback content. For instance, Howell and Archer (1984) found that disruption from DAF was unchanged when speech feedback was converted to an amplitude-modulated square-wave tone. Finney (1997) further demonstrated that pianists' performances of Bach inventions did not suffer when pitch content was altered in a random-like way. These observations are consistent with Howell's claim that DAF effects result from displacement of auditory rhythms, relative to the rhythmic coordination of actions (Howell, 2001, 2004; Howell, Powell, & Khan, 1983; Howell & Sackin, 2002).

However, Pfordresher (2003) introduced an AAF manipulation of pitch content that can disrupt production, referred to as a *serial shift*. In contrast to other manipulations of content, serial shifts involve presenting pitches from a *planned melody* (i.e., a sequence of actions that are associated with a sequence of anticipated pitch events) at an alternate serial position. For instance, a lag-1 serial shift causes the feedback pitch heard at each keypress to match what the performer produced at the previous keypress. As such the *feedback melody* is serially shifted relative to the planned melody. Serial shifts lead to increased pitch errors in performance, while sparing produced timing (Pfordresher, 2003), can disrupt performance of both pianists and nonpianists (Pfordresher, 2005), can lead to disruption of music performance via keyboard or singing (Pfordresher & Mantell, 2012), and can also disrupt the accuracy of speech production (Pruitt & Pfordresher, 2015). It is important to note that serial shifts are distinct from traditional DAF, even though serial shifts may refer to previous events. For instance, hearing serially shifted pitches from future positions also disrupts performance (Pfordresher & Palmer, 2006).

Although the effect of serial shifts does demonstrate sensitivity to pitch content, this effect alone does not rule out the possibility that their disruptive effect is to some extent based on rhythmic relationships between perception and action. It has often been observed that the patterning of pitch content can convey rhythmic information through melodic accents (Cooper & Meyer, 1960; Ellis & Jones, 2009; Jones, 1987), including breaks or pivots in the melodic contour (for examples, see Handel, 1989, p. 398), and the use of parallelism in broader pitch patterns (Acevedo, Temperley,

---

Peter Q. Pfordresher and Anastasiya Kobrina, Department of Psychology, University at Buffalo, State University of New York.

This research was sponsored by NSF grants BCS-0642592 and BCS-1256964. This work represents a portion of a master's thesis by the Anastasiya Kobrina.

Correspondence concerning this article should be addressed to Peter Q. Pfordresher, Department of Psychology, University at Buffalo, State University of New York, 362 Park Hall, Buffalo, State University of New York, NY 14260. E-mail: pqp@buffalo.edu

& Pfordresher, 2014; Deutsch & Feroe, 1981; Steedman, 1977; Temperley & Bartlette, 2002).

We adopted these features to create stimuli for the present study. Notation for the four melodies we used are shown in Figure 1 along with their meter, represented as a grid. Metrical accents are associated with positions that have higher columns of X's above them. Melodies were created to match either a binary (4/4) or a ternary (3/4) metrical structure. The first binary melody (Figure 1A) used breaks in the melodic contour to signal metrical boundaries, whereas the second binary melody includes a recurring pattern structure based on an alternating melodic contour that is inverted in the second measure. The first ternary melody (Figure 1B) used contour pivot accents to signal measure boundaries, whereas the second ternary melody used contour breaks.

The primary question guiding the present research concerns whether hearing a feedback melody with a conflicting meter (e.g., planning a binary-meter melody while hearing a ternary-meter melody in auditory feedback, and vise versa) has comparable effects on music performance to the aforementioned serial shift. Some support for this prediction comes from an earlier study by Pfordresher (2008). In that study, the sequence of auditory feedback events either matched the planned melody (i.e., the melody that a participant learned and planned to produce), or could be a melody with distinct pitches but the same melodic contour (pattern of rising and falling in pitch). Either feedback sequence could be presented in sequential alignment with produced actions, such that the melodic contour (if not the pitches) matched, or could be serially shifted, so that the melodic contour conflicted with the pattern of actions. Pfordresher found disruption from serial shifts, even when the feedback melody differed in its pitch contents. Disruption thus may occur at least in part because melodic contour accents in the feedback melody conflict with the pattern of planned finger movements. In other words, participants would hear pitches associated with strong metrical accents when planning events associated with weak metrical accents, and vice versa. In this context, the serial shift manipulation may be considered as a kind of phase-shift of the feedback meter, relative to the planned melody. If so, then it is possible that any feedback melody with a salient meter that conflicts with the planned melody may disrupt production, even if differing in both pitch content and in the melodic contour pattern.

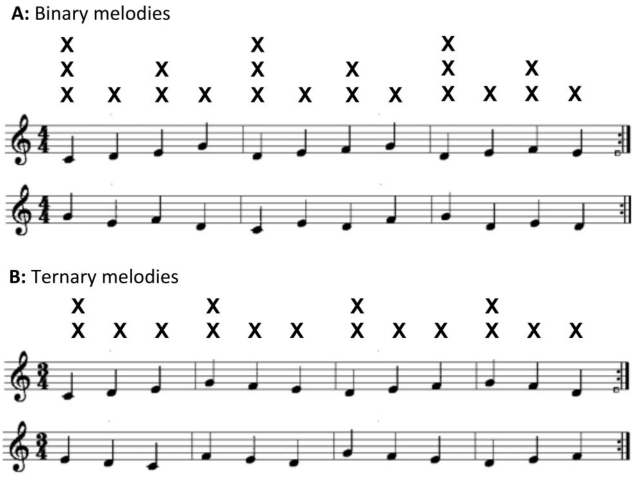

Figure 1. Stimuli used in both experiments in music notation. Metrical grids are superimposed over notation for melodies within each meter type.

As such, in the present experiments, performers heard feedback melodies that either matched the planned melody or formed a different melody with a conflicting meter. Metrical conflict was created by pairing a binary meter in the planned melody with a ternary meter in the feedback melody (Experiment 1), or vice versa (Experiment 2). This amounts to a disruption of meter with respect to its period, rather than phase. Asymmetries in period increase the complexity of frequency ratios and are very difficult to coordinate in motor tasks, as in bimanual coordination (Treffner & Turvey, 1993). If the serial shift effect is metrically based, rather than pitch-based, there should be equivalent or even greater disruption from hearing feedback melodies with a conflicting meter.

Experiment 1

Experiment 1 aimed to replicate and extend results from Pfordresher (2003, 2005, 2008). As in previous studies, participants could hear either the planned melody or a lag-1 serial shift of the planned melody (AAF) during performances. In addition, we introduced new AAF conditions in which the feedback melody conformed to a conflicting meter, which could be presented with its starting point in phase with the produced melody, or could be serially shifted by a lag of 1. There were thus four feedback conditions that resulted from crossing the factors meter (planned, conflicting) and serial shift (none, lag-1), which leads to one normal feedback condition and three AAF conditions. Figure 2 shows an example of the four conditions, for performance of one of the binary-meter melodies. In the normal feedback condition, auditory feedback would match the performed (planned) melody, whereas the other notated melodies represent feedback pitch events that the participant would hear in the other conditions.

In Experiment 1, the planned melody had a binary meter. We thus tested the hypothesis that hearing a feedback melody with a ternary meter will be equally as disruptive as hearing a lag-1 serial shift of the planned melody during performance. Furthermore, effects of serially shifting the alternate meter feedback condition should be negligible, because phase relationships between metrical accents in these conditions is always variable due to phase wrapping of melodic accents across planned and feedback melodies (cf. where measure lines align across the planned melody and either of the altered meter conditions in Figure 2).

Method

Participants. Thirty students were recruited from an introductory psychology course and were given course credit for their participation. Two participants declined to report demographic information. Of the remaining participants, 16 were female and 12 were male. The average age was 19.8 years (range = 18–23). All participants reported being right-handed. Eleven participants (37% of the sample) reported having at least 5 years of experience on a musical instrument and were considered musicians; of these five (17%) were considered keyboardists, with at least 5 years of piano experience (and one organist). Overall, participants reported an

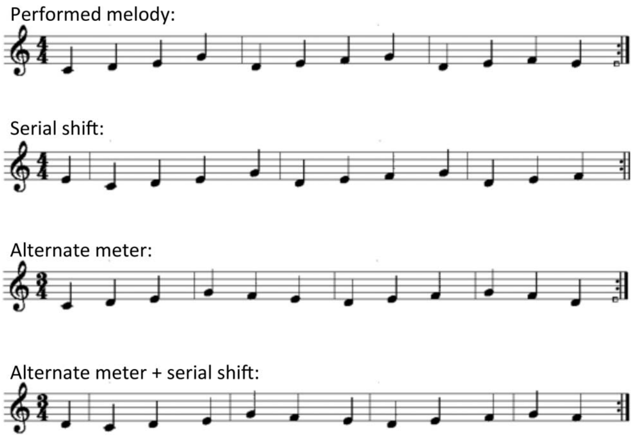

Figure 2. Notated examples for feedback conditions in a trial where the participant performs the melody in the top staff. One repetition of the melody is shown, which would be repeated five times in a trial. Music notation is used for the convenience of the reader, but was never used in the actual experiment (see *Methods* for further details). Note that the performed melody (top staff) also functions as the normal auditory feedback (control) condition.

average of 3 years of experience playing musical instruments (range = 0−16 years). Four participants reported having absolute pitch.[1]

Upon inspection of the data from both experiments, it was apparent that some participants made at least one error on many trials, which creates a ceiling effect for our measure of disruption (see the Data analysis section) and thus obscures any effects of AAF. Based on this problem, as well as the fact that participants ought to do well at reproducing the melody with normal auditory feedback (the condition they practiced), we limited the sample to those participants who reproduced the baseline condition with no errors on at least 75% of trials. In Experiment 1, this led to discarding 10 participants (final sample size = 20). These participants were comparable with the entire sample with respect to age and gender, but not surprisingly the final sample included a higher proportion of musicians than the original sample, with nine individuals having at least 5 years on any instrument (47%) and five pianists (26%).

Stimulus materials. Four 12-note melodies in the key of C major were composed for this set of experiments (see Figure 1). The melodies varied in contour (ascending, increasing from low to high; or descending, decreasing from high to low) and meter (binary or ternary). Melodies were composed so that their melodic contours and implied harmonies should communicate the notated meter. In Experiment 1, participants learned and played one of two binary melodies. This constituted the planned melody for a participant. Feedback melodies could match this planned melody, could be a serially shifted planned melody, could be a different melody with a conflicting (ternary) meter, or could be a serial shift of a different melody with a conflicting meter.

Two ternary melodies were composed to function as feedback melodies with a conflicting meter. Each ternary melody was composed to match the initial melodic contour of the binary melody. The alteration of the feedback melody's meter always led to the participant hearing the ternary melody that began with the matched contour. For instance, the first binary melody (Figure 1A, top staff) begins with upward pitch motion and was matched to the first ternary melody (Figure 1B, top staff), which begins in the same manner. A participant, who learned and performed the first binary melody, would hear the first ternary melody when AAF involved an alteration of the feedback melody's meter. It is important to note that the matching of binary and ternary melodies was restricted to this initial contour pattern.

An independent group of musically trained listeners (lab personnel who were not familiar with this experiment) validated the difference in intended meter in a rating task. Binary meters were rated as sounding significantly more binary than ternary meters, $t(15) = 4.23$, $p < .001$, on a scale from 1 (*very sure binary*) to 7 (*very sure ternary*). However, we should note that the mean rating for ternary ($M = 4.13$) suggested that these melodies were heard as fairly ambiguous, in contrast to binary melodies for which ratings were more solidly within category ($M = 2.38$). This difference likely reflects the relative infrequency of ternary meters and/or the fact that ternary meters are conceptually "compound" in that they contain binary subdivisions of the beat (Lerdahl & Jackendoff, 1983). In any case, the critical point here is that binary melodies were rated as perceptually distinct from ternary melodies with respect to meter.

A special notation was used for teaching the melodies (see Pfordresher, 2005, for examples). In this notation, the first five keys of C major on the keyboard were assigned numbers one through five that were arranged in a row above the corresponding piano keys, with arrows pointing to each key with one corresponding to C and so on. The notation had a picture of the right hand on which the fingers were assigned numbers with the thumb being one and so on. Participants had to use the highlighted finger number with the key on the keyboard.

Conditions and design. The current study used a 2 (meter type: binary vs. ternary) × 2 (serial shift: lag-0 vs. lag-1) within-subjects design. Meter type refers to the meter associated with the feedback melody, whereas serial shift refers to the relationships between melodies with respect to their starting phase. Every participant experienced four auditory feedback conditions based on crossing of these variables. Note that the binary-meter melody was always same as the learned melody, thus the control condition occurred when participant heard the lag-0 binary-meter melody (the other conditions involved AAF). Each of the feedback conditions repeated 10 times during the experiment, yielding 40 experimental trials per participant.[2] Proportion of trials in error and mean interonset intervals were measured for each condition.

Apparatus. Subjects used an M-AUDIO Keystation 49e unweighted piano keyboard. The software program FTAP (Finney, 2001) was used to manipulate auditory feedback, acquire MIDI data, and control a Roland SC 55 mk II midi tone generator.[3]

---

[1] Although absolute pitch possessors generally performed well in the task, they also experienced disruption from AAF and so were retained in the sample.

[2] In both experiments, each experimental condition was repeated 16 times for a participant. However, due to a coding error, data from the first six of these trials were overwritten and so we only consider the final 10 in our analyses.

[3] Recent studies have concerned the merits of this kind of system. For details of the discussion, see Schutz and van Vugt (2016) and Finney (2016).

Participants received auditory feedback and metronome pulses over Sony MDR-7506 professional headphones at a comfortable listening level that could be adjusted upon request.

Manipulations of feedback pitch were generated by using the "fixed pitch" setting on FTAP. This setting causes each keypress to select the next pitch from a list, which cycles around each time a sequence is repeated. The fixed pitch setting allows the presentation of feedback pitches to form a melody that is unrelated to the planned melody (Pfordresher, 2008), and has been used previously to produce feedback pitches from future locations (Pfordresher & Palmer, 2006). In lag-0 feedback conditions, the first keypress triggered the first pitch of the feedback sequence, and then every subsequent keypress triggered the next pitch in the sequence. In lag-1 feedback conditions, the first keypress triggered the last pitch of the feedback sequence, the second keypress produced the first pitch of the feedback sequence, and so on.

Procedure. Participants were informed that they would be learning and performing a simple melody during the experiment, after which they signed the consent form. Participants were assigned to one of the two melody conditions: binary ascending or binary descending. The music notation was explained, and subjects were instructed to memorize the melody. Nonmusicians and musicians who were not pianists were given additional instruction on proper hand and finger posture during piano performance. To assess whether or not participants memorized the melody and understood the task, they were required to accurately perform the melody three consecutive times from memory without any errors. Following the learning phase, the notation was taken away for the rest of the experiment.

For the practice and experimental trials, participants were instructed to start playing the melody after listening to four pulses of the metronome separated by 500-ms interonset intervals. The first 18 participants (12 from the final sample) simply continued after the metronome stopped and attempted to maintain the same rate. After some initial concerns about how well participants could do this, we instructed later participants (8 from the final sample) to synchronize with the metronome, which was maintained for 12 key presses. Later analyses that separated these groups of participants did not reveal any differences in the effects of auditory feedback; thus, we include both groups of participants in the results reported below. All participants were instructed to continue playing the repetitions of the memorized melody until the auditory feedback stopped, which occurred after the 61st keypress (i.e., after five repetitions of the melody, if no notes were added or deleted).

After each trial, participants heard a MIDI tone that cued them to rate their subjective experience of difficulty on that trial. They used white keys on the piano keyboard (29 in all) to make their rating, with maximal difficulty being anchored to the rightmost key, and minimal difficulty to the leftmost key. No auditory feedback was generated from this response.

Participants were asked to treat the experiment as if it was a real-life performance and to continue performing the melody even if they made a mistake. The practice trial used a feedback melody with a conflicting meter for the purpose of giving an example of AAF. The experiment was divided into two halves separated by a short break in order to avoid fatigue. Participants filled out demographics questionnaires during the break. Upon completion of the final trial participants were thanked for their time and debriefed. The entire procedure (for both experiments) lasted ~1 hour.

Data analysis. Errors in performed melodies were detected with software that matches performed pitches with those of an ideal performance (Large, 1993; Palmer & van de Sande, 1993, 1995). From the original set of errors, we dropped all backup errors, defined as a string of repeated insertions, because these errors may not be independent of each other. We also removed any errors that were repeated throughout a trial, which may reflect problems of learning or retrieval that are independent of disruption from AAF. Finally, we removed the first and last repetitions of the melody because these segments often included errors reflecting difficulty initiating the sequence, or errors near the end that reflect aftereffects of earlier errors (e.g., deletions at the end that occur because participants insert an additional event near the middle).

After this preprocessing, the main measure of disruption was the proportion of trials with one or more errors, referred to as the *proportion of trials in error*, which has been used in previous studies that incorporated the fixed pitch manipulation from FTAP (Pfordresher, 2008; Pfordresher & Palmer, 2006). The reason for using a proportion of trials with errors instead of error rate is because errors that involve additions or deletions alter the serial relationship between actions and auditory feedback for subsequent events. The measure we use, which just focuses on the first produced error, circumvents this problem.

Difficulty rating responses were converted to reflect the percentile rank of the key used for the response. We did this by dividing the pitch height of the key, relative to all lower white keys (adjusting for the asymmetry of white-key spacing on the keyboard), dividing that number by the total number of keys, and multiplying that proportion by 100. A similar procedure was used for ratings of self-agency in Couchman, Beasley, and Pfordresher (2012).

Results and Discussion

We analyzed the proportion of trials in error using a 2 (meter type: binary vs. ternary) × 2 (serial shift: lag-0 vs. lag-1) within-subject analysis of variance (ANOVA).[4] Figure 3A shows boxplots relating to each treatment condition. The ANOVA revealed a significant main effect of serial shift, $F(1, 19) = 15.97$, $p < .001$, $\eta_p^2 = .46$, and a significant Serial Shift × Meter Interaction, $F(1, 19) = 5.75$, $p < .05$, $\eta_p^2 = .23$, but no main effect of meter ($p = .11$, $\eta_p^2 = .13$). We analyzed the interaction using two sequential Dunnett's tests. The first test analyzed contrasts between each of the altered feedback conditions with the control condition (binary meter and no serial shift). Each of these contrasts was significant, indicating that every alteration disrupted production. The second test discarded the control condition and analyzed whether the serial shift condition used in previous research (i.e., a shift of the planned melody) differed from either of the other altered feedback conditions, which included the altered melody with a ternary meter. Neither of these contrasts were significant. These contrasts reflect

---

[4] Distributions of both dependent variables differed significantly from normal, and so we also analyzed the data using nonparametric tests: Friedman's test on all four treatment conditions, followed by Wilcoxon signed ranks tests on pairs of means. The results of these tests aligned with parametric tests reported here. We report parametric tests because the discussion is simpler, and the procedures allow for more straightforward corrections of family-wise Type I error.

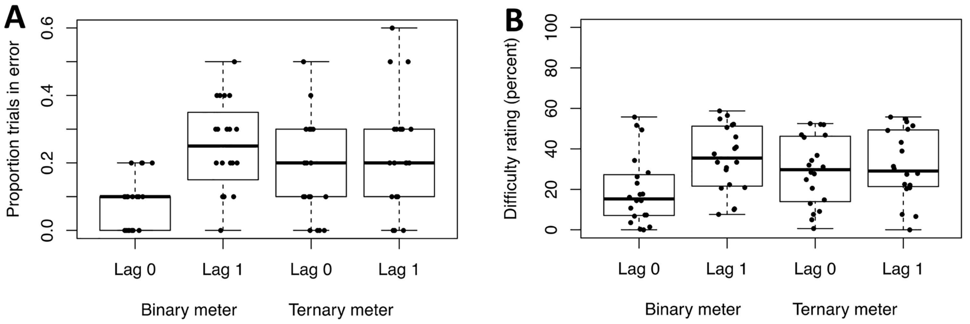

Figure 3. Boxplots of results from Experiment 1, showing the subset of participants who had at least 75% error-free trials in the control condition (lag-0, binary meter). For each condition, rectangles highlight the interquartile range, thick horizontal lines represent the median, and whiskers span from the minimum to the maximum value. Means across trials for individual participants are shown as dots, with random jitter added to *x*-axis values to avoid overlap.

the statistical interaction in the ANOVA, because they show that the serial shift effect is qualified by the meter of the feedback melody. Thus, in keeping with the hypothesis described earlier, the effect of altering meter led to disruption of similar magnitude to serially shifting pitches from the planned melody.

Figure 3B shows the corresponding analysis of difficulty ratings. Results were comparable with error proportions, and the measures were significantly correlated across participants and conditions, $r(78) = .45$, $p < .001$. As with error proportions, the ANOVA yielded a main effect of serial shift, $F(1, 19) = 26.64$, $p < .001$, $\eta_p^2 = .61$, and a significant Serial Shift × Meter Interaction, $F(1, 19) = 16.00$, $p < .001$, $\eta_p^2 = .46$, but no main effect of meter ($p = .16$, $\eta_p^2 = .10$). However, post hoc analyses suggested that serial shifts of the planned melody may have had a stronger effect than other AAF conditions with altered meters. The only significant contrast with the control (binary meter, unshifted) condition was between that condition and the serially shifted binary meter condition (i.e., the serially shifted planned melody). The other two contrasts with the control condition were nonsignificant. At the same time, the second Dunnett's test yielded no differences between the serially shifted binary meter condition and either altered (ternary) meter condition. Thus, with respect to difficulty ratings, the ternary-meter conditions seem to yield intermediate levels of disruption (cf. Pfordresher, 2005).

As described in *Participants*, we discarded data from 1/3 of the sample in order to evaluate results only from those participants who were able to learn the sequence reasonably well. Of course, when participants are dropped from a sample, it is reasonable to question how their data compare with the rest. As such, we evaluated whether each individual's proportion of trials in error for the four feedback conditions matched the data in Figure 3 ordinally. The error proportions for each feedback condition within an individual were correlated with a set of coefficients reflecting the effect found for means from the sample of 20 included participants. The control condition was coded as −3 and the remaining three conditions (which were statistically indistinguishable from each other) were each coded as +1. For each participant (including those not included in the Figure 3 data), a nonzero positive correlation between these coefficients and their error proportions was coded as a match to the effect seen with the used sample, and all other correlations were coded as a nonmatch. Of the 30 participants in the original sample, only two participants (both discarded) failed to match the pattern. Thus, the results shown in Figure 3 were reliable across all participants, including those who were discarded, at least at an ordinal level.

Another question concerns individual differences in musical training. Although we had participants learn simple melodies on notation that is easily understood by nonmusicians, it is possible that responses to AAF could differ across groups. We addressed this by running two additional ANOVAs that included a grouping variable based on categorization of participants as a musician (nine participants) or a keyboardist (five participants). In neither case did the grouping variable interact with any effects related to auditory feedback.

Experiment 2

Experiment 2 builds on Experiment 1 by changing the meter of the planned melody. In this experiment, participants were asked to learn a ternary melody and to perform it while receiving one of four types of auditory feedback. These included normal feedback (now with a lag-0 ternary meter), a lag-1 serial shift of the planned melody, a feedback melody with a binary (conflicting) meter, and a lag-1 serial shift of the binary-meter melody. Our starting prediction was that the results of Experiment 2 would mirror those from Experiment 1, with all altered conditions being disruptive. This prediction was made tentatively, however, given that ternary meters are less common than binary meters in Western music and thus this metrical structure may be less salient overall than that of binary meters.

Method

Participants. Thirty University at Buffalo undergraduate students were recruited from an introductory psychology course and were given course credit for their participation in Experiment 2. One participant declined to report demographic information. Of the remaining participants, 16 were female and 13 were male. Their average age was 19.7 years (range = 18–42). Twenty-eight participants reported being right-handed, and one reported being left-handed. Eleven participants (37% of the sample) reported having at least 5 years of experience on a musical instrument and were considered musicians; of these two (7%) were considered keyboardists, having at least 5 years of piano experience. One participant reported having absolute pitch.

As in Experiment 1, we limited analyses of mean data to those participants who were able to produce the sequence error free in the control condition on at least 75% of trials. Apparently, participants in Experiment 2 had a much harder time achieving this criterion, and 17 participants were discarded (final sample size = 13). The final sample included six musicians (46% of the sample) and one of the two pianists (somewhat surprisingly, the absolute pitch possessor was discarded).

Materials, conditions, procedure, and analyses. The same materials were used as in Experiment 1, with the exception that participants first learned one of the two ternary melodies, which functioned as the planned melody. Likewise, the same design was used with the exception that the control condition in Experiment 2 was performance of a ternary-metered melody with no serial shift. We also analyzed data in the same way as Experiment 1, treating feedback conditions as appropriate to Experiment 2.

Results and Discussion

Boxplots in Figure 4A show distributions of error proportions across conditions. The ANOVA (using the same design as in Experiment 1) yielded only a main effect of serial shift, $F(1, 12) = 8.95$, $p < .05$, $\eta_p^2 = .43$. The main effect of meter was nonsignificant ($p = .12$, $\eta_p^2 = .19$) as was the Serial Shift × Meter Interaction ($p = .31$, $\eta_p^2 = .09$). It is important to note that, as in Experiment 1, participants were disrupted by the presence of an accompanying melody with a meter that conflicts with that of the planned melody, even when the accompanying melody comprises a distinct sequence of pitches from the planned melody. Thus, metrical conflict on its own is sufficient to disrupt production, which suggests that the serial shift effect observed in many studies may have to do with the temporal organization implied by pitch patterns, and is not a pitch-specific effect entirely. However, unlike Experiment 1, alterations of meter only caused significant disruption when the starting point of the melody was shifted back one position relative to the starting point of the planned melody. Note that the dramatic difference in effect sizes across experiments for this interaction ($\eta_p^2 = .23$ in Experiment 1 vs. .09 in Experiment 2) suggests that the nonsignificant interaction here does not reflect the smaller sample size in Experiment 2.

Figure 4B shows a comparable analysis of difficulty ratings. As with error proportions, the ANOVA yielded only a main effect of serial shift $F(1, 19) = 5.06$, $p < .05$, $\eta_p^2 = .30$, but no main effect of meter ($p = .16$, $\eta_p^2 = .16$) and no interaction ($p = .69$, $\eta_p^2 = .01$). Thus, as in Experiment 1, difficulty ratings converged with error proportions. Like Experiment 1, difficulty ratings were positively correlated with error proportions across participants and conditions. This correlation did not reach significance in Experiment 2, however, $r(38) = .12$, $p = .46$.

In Experiment 2, we dropped over half of the sample in order to identify participants who successfully learned the stimulus melody. As such, we evaluated whether the pattern observed in Figure 4A holds for individuals in the entire original sample. We employed a similar contrast analysis to Experiment 1, using coefficients that reflect the results found in Experiment 2 for the used sample. Both nonshifted feedback conditions (control and binary melody) were coded as -1 and both serially shifted conditions were coded as $+1$. In Experiment 2, only 15 participants' data matched the pattern found in Figure 4A. Figure 5 shows the relationship between the categorization of the participants' contrasts and the error proportions in the control condition, which were used to include or exclude individuals. As can be seen, the vast majority of participants included (see dots for individuals with y-values lower than .25) had contrasts that matched the pattern in Figure 4A, and participants who did not learn the sequence well tended to yield error proportions that did not match the pattern shown in Figure 4A. Thus, in contrast to Experiment 1, the metrical effect of auditory feedback on performance depended on how well participants were able to learn the melody.

As in Experiment 1, we addressed individual differences in musical training in a follow-up ANOVA. Because there was only one pianist in the final sample, we only ran an ANOVA with a grouping variable that coded for any kind of musical training (6 participants). As in Experiment 1, this factor did not interact with any effect related to auditory feedback.

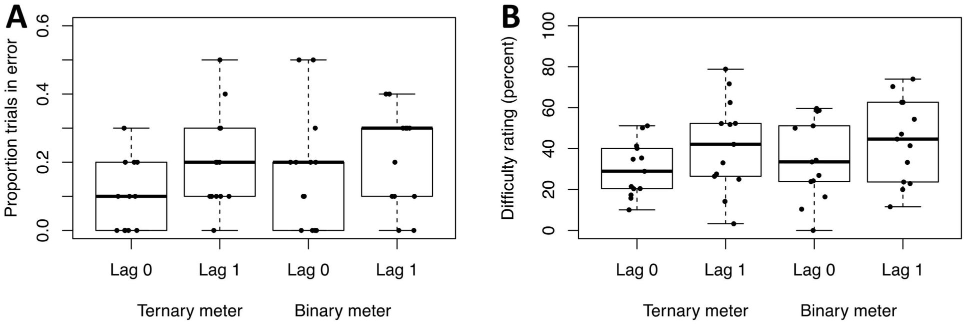

Figure 4. Boxplots of results from Experiment 2, showing the subset of participants who had at least 75% error-free trials in the control condition (lag-0, ternary meter). For each condition, rectangles highlight the interquartile range, thick horizontal lines represent the median, and whiskers span from the minimum to the maximum value. Means across trials for individual participants are shown as dots, with random jitter added to *x*-axis values to avoid overlap.

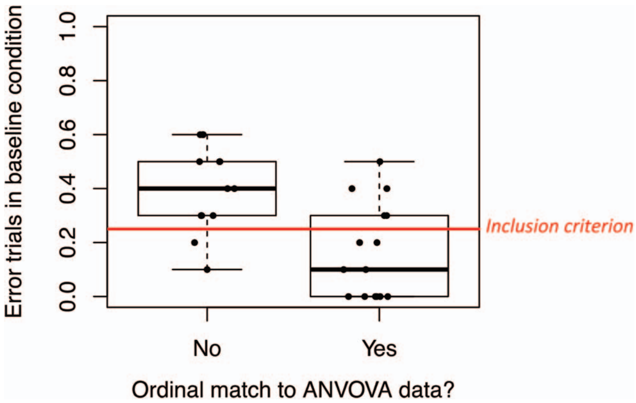

Figure 5. Boxplots of Experiment 2 data based on categorization of individual participants. The full sample is shown. The *x*-axis category reflects whether an individuals' pattern of errors across all four conditions matches the pattern found for the mean across participants. The *y*-axis displays the proportion of trials in error for individual participants within the control condition (lag-0, ternary meter), which was used as a criterion for inclusion in the final sample (see horizontal reference line). For each *x*-axis category, rectangles highlight the interquartile range, thick horizontal lines represent the median, and whiskers span from the minimum to the maximum value. Means for individual participants are shown as dots, with random jitter added to *x*-axis values to avoid overlap. See the online article for the color version of this figure.

General Discussion

The two experiments reported here demonstrate that disruptive effects of AAF involving pitch may partly arise from the temporal organization of pitch patterns, specifically the implied meter. In both experiments, disruption from altered pitch was observed when participants heard a different feedback melody than the one they had planned, but that had a meter that conflicted with the metrical organization of the planned melody. In Experiment 1, where participants planned and produced a binary-meter melody, this was found for feedback melodies with a distinct (ternary) meter, independent of how that meter was phased with the sequence of actions. In Experiment 2, in which participants planned and produced a ternary-meter melody, this was found for melodies with a distinct (binary) meter, but only when the feedback melody was out of phase (serially shifted by a lag of 1) relative to the planned melody.

This is a significant finding because it suggests that sensitivity to the pitch content of auditory feedback may be, at least in part, temporal. As Jones (1987) observed, the pitch contour of a melodic sequence creates accents based on changes in direction, leaps in interval size, and implied harmony. These accents generate a sense of higher-order time structure in the listener that leads to the perceived meter (cf. Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983). Although the disruptive effects of serial shifts must have to do with pitch content, and are distinct in many ways from the effects of asynchronous feedback (for a review see Pfordresher, 2006), the basis of this disruption may be based partly on the fact that pitch can convey temporal structure.

The results of the present study complement those of Pfordresher (2008). In that study, performance was disrupted when participants heard a serially shifted feedback melody that differed from the planned melody but shared the same melody contour. As in the present study, altered pitches disrupted production based on temporal structure implied by melodic contour. However, the present study is the first to report effects of altering the feedback melody's meter. In other words, whereas the manipulations of Pfordresher (2008) can be attributed to phasing of pitch structure, the present data speak to the importance of perception and action aligning with respect to the overarching period (binary vs. ternary).

It is important to note that many AAF manipulations of pitch cause no disruption whatsoever. Thus, it is wrong to conclude from the present data that any kind of alteration to pitch may disrupt performance. Previous studies have shown that pitch alterations that lead to a random-like unpatented sequence that does not repeat cause no disruption (Finney, 1997; Pfordresher, 2005). Likewise, Pfordresher (2005) found that hearing a repeated pattern based on pitches differing from the planned melody (presented in phase with the planned melody) failed to disrupt production. Thus, the disruptive effects of meter in the present data likely reflect the fact that the pitches do form a repeating pattern but one that conflicts metrically with the planned melody. If anything, the fact that alterations of feedback meter in the present study also resulted in changes to the feedback melody should have reduced the disruptive effect of this AAF condition, given previously published results (for more discussion see Pfordresher, 2006).

That being said, other data suggest that the effects of serial shifts are not simply due to temporal patterning of pitches. Whereas Pfordresher (2008) found disruption of production from a serially shifted melody that differed from the planned melody yet shared its contour, there was an important qualification. The planned melody was tonal, and the disruptive effect of shifts vanished when the feedback melody was atonal. Tonality is a salient musical feature that relates to the sense that pitches belong to a specific "key" within which some pitches belong (are stable) and others do not (are unstable). Tonality is typically considered to be based on statistical properties across the entire sequence that are independent of a melody's temporal organization (Krumhansl, 1990). A more recent paper further supports the importance of tonality in relating auditory feedback to the planned melody. Jebb and Pfordresher (2016) replicated the results of Pfordresher (2008) and also showed that when participants produced and planned an atonal melody, a contour matched variation that was itself tonal could disrupt production when serially shifted.

An unexpected limitation of the present study arose in the apparent difficulty of the sequences, particularly ternary sequences of Experiment 2. As a result, more participants were discarded than would be desirable. To a great degree, these attrition rates reflect the kind of dependent variable that we needed to use. Because we focused just on whether any error was present in a trial (for reasons described in the *Methods*), some participants could have made just a single error regularly across trials. In this context, it is worth noting that the mean error rates within trials averaged across participants and trials, were in ranges observed in other experiments, though still higher in Experiment 2 than in Experiment 1 (Experiment 1 $M = 3.8\%$, $SD = 4.9\%$; Experiment 2 $M = 7.3\%$, $SD = 5.4\%$).

Thus, the high attrition rate was in part a byproduct of necessary measure constraints. But an important question remains: Why were errors so much more prevalent in Experiment 2 than in Experiment 1? It seems unlikely that differences in musical training led to these discrepancies. Although Experiment 1 included more keyboardists, both experiments had a similar proportion of musically trained participants (about half the sample in both cases) and, more the point, there was no evidence that musical training influenced responses to AAF. A more likely possibility, we think, is the inherent complexity of binary versus ternary meters. Almost all previous studies on AAF effects of piano used binary-meter stimuli, and the one study that incorporated ternary meters included only trained pianists (Pfordresher, 2003). Binary meters are more prevalent in Western music and may be preferred. For instance, Nozaradan, Peretz, Missal, and Mouraux (2011) discovered that when participants performed the ternary meter imagery task, additional involuntary binary meter activation occurred in the EEG signal. There is a general preference for binary metrical relationships (Essens & Povel, 1985; Fraisse, 1982) that is even found in infants (Bergeson & Trehub, 2006). Infants and adults alike can extract different metrical structures from simple melodies (Bergeson & Trehub, 2006; Hannon & Johnson, 2005; Hannon & Trehub, 2005). Conversely, infants are more sensitive to subtle pitch changes in binary melodies rather than ternary melodies (Bergeson & Trehub, 2006). The ability to perceive binary meter may reflect basic predispositions, whereas ternary meter perception and representation may require experience with music or stimulation from rhythmic body movement (Phillips-Silver & Trainor, 2005). These previous findings are likewise reflected in listener ratings of our stimuli, which suggest that ternary meters were more ambiguous metrically than binary meters, despite being composed using the same principles.

Another puzzle from Experiment 2 has to do with differences across conditions in which participants heard an alternate melody with a conflicting (binary) meter. In Experiment 1 the phasing of the alternate melody did not influence performance. By contrast, in Experiment 2 there was only a main effect of serial shift, suggesting that the nonshifted alternate (binary) melody did not disrupt performance. This is puzzling if one only considers metrical organization. When two sequences have different meters, the phase of one sequence varies with respect to the other (referred to as phase wrapping), and it should not matter whether the feedback melody is “serially shifted.” Here again we have evidence that meter is not the only reason why serial shifts are disruptive, even though meter may play a major role. Another important factor concerns the starting point of each sequence. Pfordresher and Kulpa (2011) examined the time series over which disruption from serial shifts occurred. Their data revealed a substantial increase in errors at the starting point of a repeated melody. Performers thus may be particularly sensitive to feedback relationships when a sequence is initiated, perhaps because of the demands associated with retrieving chunked information from memory (cf., Chaffin, Logan, & Begosh, 2009).

In conclusion, the present data argue for an important role of metrical organization in how performers respond to the pitch content of auditory feedback. Metrical representation is a powerful component of musical structure that is present in all listeners, although it may be more refined among musicians (Palmer & Krumhansl, 1990). In performance, meter plays an important role in the retrieval of musical sequences (Mathias, Palmer, Pfordresher, & Anderson, 2011; Palmer & Pfordresher, 2003). We here show that the use of the pitch content in auditory feedback may be based in part on the way in which pitch patterns convey meter (Jones, 1987).

References

Acevedo, S., Temperley, D., & Pfordresher, P. Q. (2014). Effects of metrical encoding on melody recognition. *Music Perception, 31,* 372–386. http://dx.doi.org/10.1525/mp.2014.31.4.372

Bergeson, T. R., & Trehub, S. E. (2006). Infants perception of rhythmic patterns. *Music Perception, 23,* 345–360. http://dx.doi.org/10.1525/mp.2006.23.4.345

Chaffin, R., Logan, T. R., & Begosh, K. (2009). Performing from memory. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 352–363). Oxford, United Kingdom: Oxford University Press.

Cooper, G. W., & Meyer, L. B. (1960). *The rhythmic structure of music*. Chicago, IL: University of Chicago Press.

Couchman, J. J., Beasley, R., & Pfordresher, P. Q. (2012). The experience of agency in sequence production with altered auditory feedback. *Consciousness and Cognition, 21,* 186–203. http://dx.doi.org/10.1016/j.concog.2011.10.007

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review, 88,* 503–522. http://dx.doi.org/10.1037/0033-295X.88.6.503

Ellis, R. J., & Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 264–280. http://dx.doi.org/10.1037/a0013482

Essens, P. J., & Povel, D. J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception and Psychophysics, 37,* 1–7. http://dx.doi.org/10.3758/BF03207132

Finney, S. A. (1997). Auditory feedback and musical keyboard performance. *Music Perception, 15,* 153–174. http://dx.doi.org/10.2307/40285747

Finney, S. A. (1999). *Disruptive effects of delayed auditory feedback on motor sequencing* (Unpublished doctoral dissertation). Brown University, Providence, RI.

Finney, S. A. (2001). FTAP: A Linux-based program for tapping and music experiments. *Behavior Research Methods, Instruments, and Computers, 33,* 65–72. http://dx.doi.org/10.3758/BF03195348

Finney, S. A. (2016). *In Defense of Linux, USB, and MIDI systems for sensorimotor experiments: A response to Schultz and van Vugt (2015)*. Unpublished manuscript. Retrieved from http://www.sfinney.com/images/pdfs/sf/finney2016a.pdf

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

Handel, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: MIT Press.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical struct.ure learning. *Cognitive Psychology, 50,* 354–377. http://dx.doi.org/10.1016/j.cogpsych.2004.09.003

Hannon, E. E., & Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 12639–12643. http://dx.doi.org/10.1073/pnas.0504254102

Howell, P. (2001). A model of timing interference to speech control in normal and altered listening conditions applied to the treatment of stuttering. In B. Maassen, W. Hulsijn, R. Kent, H. F. M. Peters, & P. H. M. M. van-Lieshout (Eds.), *Speech motor control in normal and disordered speech* (pp. 291–294). Nijmegen, the Netherlands: Uttgeverij Vantilt.

Howell, P. (2004). Assessment of some contemporary theories of stuttering that apply to spontaneous speech. *Contemporary Issues in Communication Science and Disorders, 31,* 122–139.

Howell, P., & Archer, A. (1984). Susceptibility to the effects of delayed auditory feedback. *Perception & Psychophysics, 36,* 296–302. http://dx.doi.org/10.3758/BF03206371

Howell, P., Powell, D. J., & Khan, I. (1983). Amplitude contour of the delayed signal and interference in delayed auditory feedback tasks. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 772–784. http://dx.doi.org/10.1037/0096-1523.9.5.772

Howell, P., & Sackin, S. (2002). Timing interference to speech in altered listening conditions. *The Journal of the Acoustical Society of America, 111,* 2842–2852. http://dx.doi.org/10.1121/1.1474444

Jebb, A. T., & Pfordresher, P. Q. (2016). Exploring perception-action relations in music production: The asymmetric effect of tonal class. *Journal of Experimental Psychology: Human Perception and Performance, 42,* 658–670. http://dx.doi.org/10.1037/xhp0000172

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics, 41,* 621–634. http://dx.doi.org/10.3758/BF03210494

Krumhansl, C. L. (1990). *The cognitive foundations of musical pitch*. New York, NY: Oxford University Press.

Large, E. W. (1993). Dynamic programming for the analysis of serial behaviors. *Behavior Research Methods, Instruments, and Computers, 25,* 238–241. http://dx.doi.org/10.3758/BF03204504

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

Mathias, B., Anderson, M. F., Palmer, C., & Pfordresher, P. Q. (2011). Effects of meter and serial position on memory retrieval during music performance. In A. Williamon, D. Edwards & L. Bartel (Eds.), *Proceedings of the International Symposium on Performance Science* (pp. 405–410). Utrecht, the Netherlands: Association Européenne des Conservatories.

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience, 31,* 10234–10240. http://dx.doi.org/10.1523/JNEUROSCI.0411-11.2011

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 728–741. http://dx.doi.org/10.1037/0096-1523.16.4.728

Palmer, C., & Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review, 110,* 683–712. http://dx.doi.org/10.1037/0033-295X.110.4.683

Palmer, C., & van de Sande, C. (1993). Units of knowledge in music performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 457–470. http://dx.doi.org/10.1037/0278-7393.19.2.457

Palmer, C., & van de Sande, C. (1995). Range of planning in music performance. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 947–962. http://dx.doi.org/10.1037/0096-1523.21.5.947

Pfordresher, P. Q. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 949–964. http://dx.doi.org/10.1037/0096-1523.29.5.949

Pfordresher, P. Q. (2005). Auditory feedback in music performance: The role of melodic structure and musical skill. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 1331–1345. http://dx.doi.org/10.1037/0096-1523.31.6.1331

Pfordresher, P. Q. (2006). Coordination of perception and action in music performance. *Advances in Cognitive Psychology, 2,* 183–198. http://dx.doi.org/10.2478/v10053-008-0054-8

Pfordresher, P. Q. (2008). Auditory feedback in music performance: The role of transition-based similarity. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 708–725. http://dx.doi.org/10.1037/0096-1523.34.3.708

Pfordresher, P. Q., & Kulpa, J. D. (2011). The dynamics of disruption from altered auditory feedback: Further evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 949–967. http://dx.doi.org/10.1037/a0021435

Pfordresher, P. Q., & Mantell, J. T. (2012). Effects of altered auditory feedback across effector systems: Production of melodies by keyboard and singing. *Acta Psychologica, 139,* 166–177. http://dx.doi.org/10.1016/j.actpsy.2011.10.009

Pfordresher, P. Q., & Palmer, C. (2006). Effects of hearing the past, present, or future during music performance. *Perception and Psychophysics, 68,* 362–376. http://dx.doi.org/10.3758/BF03193683

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science, 308,* 1430. http://dx.doi.org/10.1126/science.1110922

Pruitt, T. A., & Pfordresher, P. Q. (2015). The role of auditory feedback in speech and song. *Journal of Experimental Psychology: Human Perception and Performance, 41,* 152–166. http://dx.doi.org/10.1037/a0038285

Schutz, B. G., & van Vugt, F. T. (2016). Tap Arduino: An Arduino microcontroller for low-latency auditory feedback in sensorimotor synchronization experiments. *Behavior Research Methods, 48,* 1591–1607.

Steedman, M. J. (1977). The perception of musical rhythm and metre. *Perception, 6,* 555–569. http://dx.doi.org/10.1068/p060555

Temperley, D., & Bartlette, C. (2002). Parallelism as a factor in metrical analysis. *Music Perception, 20,* 117–149. http://dx.doi.org/10.1525/mp.2002.20.2.117

Treffner, P. J., & Turvey, M. T. (1993). Resonance constraints on rhythmic movement. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1221–1237. http://dx.doi.org/10.1037/0096-1523.19.6.1221

Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin, 60,* 213–232. http://dx.doi.org/10.1037/h0044155

Received May 26, 2016
Revision received September 26, 2016
Accepted December 10, 2016 ■