Melodic Accent as an Emergent Property of Tonal Motion

PAOLO AMMIRANTE
Macquarie University

WILLIAM FORDE THOMPSON
Macquarie University

ABSTRACT: In a previous continuation tapping study (Ammirante, Thompson, & Russo, in press), each tap triggered a discrete tone in a sequence randomly varying in pitch height and contour. Although participants were instructed to ignore the tones, pitch distance and pitch contour influenced intertap interval (ITI) and tap velocity (TV). The current study replicated these findings with original melodies. Results were interpreted as an effect of apparent tonal motion, with deviation in ITI and TV mirroring implied tonal acceleration. Due to overlapping perceptual and motor representations, participants may have failed to disambiguate acceleration implied by tonal motion from the acceleration of their finger trajectory. Dissociative effects of pitch distance on ITI and pitch contour on TV implied that pitch distance influences the initial finger extension while pitch contour influences later finger flexion. Acceleration in ITI and TV were also both correlated with melodic accent strength values from perceptual data (Thomassen, 1982), suggesting that perception and production of melodic accent emerge from shared action associations.

IT has been established that the perception and production of timing in tonal sequences is systematically influenced by the pattern of changing pitches in those sequences (e.g., Boltz, 1998; Drake, Dowling, & Palmer, 1991; Drake & Palmer; 1993; Repp, 1995, 1998, 1999; Tekman, 2001). One account, the *musical experience hypothesis* (Repp, 1992, 1995), proposes that this effect reflects familiarity with the conventions of expressive musical performance. In support of this hypothesis, Repp (1992, 1995, 1998) found that small temporal perturbations from strict timing are more difficult to detect if they are consistent with conventional timing deviations in expressive performance. However, Repp (1992, 1995) also found that perceptual bias was not correlated with listeners' musical experience, which the hypothesis does not predict. If perceptual bias was acquired through familiarity with conventional expressive performance, then it should have been stronger in experienced listeners. Another account, the *perceptual salience hypothesis*, proposes that certain melodic events, particularly contour-changing (CC) tones, stand out more than others (In the melodic pattern C4-E4-D4, where an ascending major third is followed by a descending whole tone, D4 is referred to as the *contour changing tone*.) In a musical context, CC tones may function as *melodic accents*, coinciding with points of stress and prolongation, and serving to delineate larger structural units such as phrases within a musical flow (Lerdahl & Jackendoff, 1983). This association may, in turn, bias perception and production. For example, Drake, Palmer, & Dowling (1991) found that children and adult pianists had more performance errors when CC tones were out of phase with points of rhythmic stress. However, null findings for the perceptual salience of melodic accents have also been reported (Drake, 1993; Penel & Drake, 1998). More generally, it is unclear why a contour change should stand out more than, for example, a change in pitch height, which a contour-preserving (CP) melodic pattern entails.

An alternative hypothesis proposes that transitions between discrete tones varying in pitch height may be perceived to have motion-like properties (see Shove & Repp, 1995 for a review). Boltz (1998) had listeners compare the tempi of pairs of isochronous melodies. Despite instructions to ignore pitch, comparison melodies containing more CC tones were judged to be slower than the target; conversely, comparison melodies containing fewer CC tones were judged to be faster. These data imply that global tempo judgment reflected the accumulation of local perturbations in perceived tempo arising from contour changes. Although this finding is not inconsistent with the perceptual salience hypothesis, Boltz (1998)

alternatively proposed that perceptual bias may have been an effect of *apparent tonal motion*, reflecting listeners' implicit awareness of terrestrial locomotion. When navigating turns, a moving body equipped with a vestibular system must slow down in order to maintain balance, but tends to speed up when moving unidirectionally. Comparable effects of lawful biological motion have been found for continuous visual stimuli. For example, Viviani & Stucchi (1992) found that when the path of an ellipse traced by a dot moving across a screen slowed down at turns and sped up at the straighter parts, it was judged to be of more uniform velocity than a path of actual uniform velocity.

Apparent tonal motion may also account for interactions between timing and pitch distance. (Cohen, Hansel, & Sylvester, 1954; Yoblick & Salvendy, 1970; Shigeno, 1986, 1993; Crowder & Neath, 1994; Mackenzie, 2007; Henry & McAuley, 2009). For example, in a synchronised tapping task to a continuous pacing signal oscillating sinusoidally in pitch, McAnally (2002) found that the tendency for taps to precede each oscillation increased with the pitch distance to peak amplitude. Thus, just as greater spatial extent traversed by a moving body within a fixed time interval means faster motion, greater pitch distance traversed at a steady rate may have implied faster tonal motion. This may have influenced finger movement velocity as reflected in an increased tendency towards early taps.

A previous study (Ammirante, Thompson, & Russo, in press) aimed to replicate this last finding for pitch distance as well as to investigate pitch contour/timing interactions in a continuation tapping task with discrete tones. In continuation tapping, participants first synchronize their taps with a pacing signal after which the signal is removed and the participant continues to tap at the same tempo until the end of the trial. Typically only the continuation taps are subjected to analysis. In this study, each continuation tap triggered a tone. Pitch distance between successive tones varied systematically between trials. Pitch contour also varied between CC and CP tones within trials. Although participants were instructed to ignore them, tones had systematic effects on the timing of continuation tapping. First, consistent with the implications of the McAnally (2002) study, tempo was faster with greater pitch distance. Second, consistent with Boltz' (1998) perceptual findings, the intertap interval (ITI) initiated by CP tones was faster than for CC tones. Moreover, tap velocity (TV: the velocity of the tap trajectory as measured by MIDI) was faster following CP tones than following CC tones.

Taken together, these data supported the apparent tonal motion hypothesis. Tempo was globally faster with accelerated motion implied by increased pitch distance. ITI and TV mirrored acceleration implied by the pitch contour of the just-heard tone. The data did not support the musical experience hypothesis in that years of musical training did not predict the magnitude of the effects. They also did not support the perceptual salience hypothesis in that there was no increase in variability in ITI in the pitched conditions relative to a monotone control condition (Finney, 1997). ITI variability in continuation tapping increases when selective attention is taxed (Sergent, Hellige, & Cherry, 1993). Thus, if CC tones were more salient, cognitive costs should have been associated with their unpredictable recurrence. Indeed, the fact that TV was affected by pitch contour within ~500 msec of tone onset with no increase in variability suggests modulated responses may have been automatic.[1]

Ammirante, Thompson, & Russo (in press) proposed that automatic modulated responses may have been due to overlapping perceptual and motor representations. In this case, participants may have failed to disambiguate acceleration implied by apparent tonal motion from the acceleration of their finger trajectory. This interpretation is compatible with *Common Coding theory,* which proposes that perceptual and motor representations of action are shared (Hommel et al., 2001). Neurological evidence for common coding theory has been demonstrated in monkeys and humans (Rizzolati & Craighero, 2004), where brain activity occurs in the same regions in response to both the perception of an action and performance of the same action. Behavioural evidence supports the prediction from Common Coding theory that "perception and action will modulate each other reciprocally whenever they are similar" (Prinz, Aschersleben, & Koch, 2009: p. 40; James 1890). For example, in the auditory domain, reciprocal modulation has been reported for single and paired tones. Stimulus-response compatibility between the pitch height of a tone and the spatial position of response keys has been shown to modulate action (Rusconi et al. 2006; Lidji et al. 2007) and perception (Repp & Knoblich, 2007).

The notion that the motion-like properties of tones may have automatic modulating effects is not new. For example, Truslit proposed that musical melodies "elicit in the listener *a corresponding motion experience*" (1993 [1938]: p. 50, italics added). He also suggested that the engagement of the vestibular system underlies this experience — a view supported by electrophysiological evidence (Todd, 1999; Trainor et al., 2009). Truslit further proposed that the experience is unconscious and largely inhibited in the listener, "its only outward manifestation [being] subtle tensions of the muscles" (Truslit 1993 [1938]: p.

51), but expressed in music performance through accentuation such as a lengthening of temporal interval (temporal accent) and an increase in intensity (dynamic accent). Thus, the *de facto* temporal and dynamic accents that spontaneously emerged in non-musicians in the previous study, although smaller in magnitude, may have been rooted in the same "corresponding motion experience" communicated by skilled musicians in expressive music performance (Leman, 2009).[2] This shared experience might account for the persistence in musically naïve listeners of a perceptual bias correlated with expressively accentuated performance (Repp, 1992, 1995, 1998). It might also explain why, consistent with non-musicians' responses to random tonal sequences in the previous study, unintentional systematic temporal and dynamic deviations persist in musicians' performances in strict timing from notated melodies (Drake & Palmer, 1993; Penel & Drake, 1998; Repp, 1999). For example, when performing an isochronous melody, CP tones are played with greater velocity than CC tones (Drake & Palmer, 1993, exp.1). However, pitch contour and pitch distance are not always associated with unintentional systematic deviations in performance timing (Penel & Drake, 1998; Repp, 1999). Moreover, the deviations reported by Drake and Palmer may have been related to factors such as familiarity with the melody, visual cues in the notated score, physical constraints, and metrical structure.

The study by Ammirante, Thompson, and Russo isolated tonal motion effects of pitch contour and pitch distance from these confounding factors. However, the sequences in that study bore little resemblance to musical melodies. The current study aimed to replicate their findings with more musical stimuli. Pitch height between successive triggered tones was varied to form original melodies. Each melody was composed from either a major or minor scale and presented with either a sine or piano timbre. There were no specific predictions with respect to modality or timbre, but it was of interest to determine whether previously observed effects should generalize to more complex musical sounds. Melodies involved either repeating tones or stepwise changes, which introduced local changes in pitch distance of between 0 and 2 semitones within sequences. The inclusion of repeating tones expanded the set of three-tone melodic "cells" from 4 in the previous study to 8 (see Figure 6 below), which could then be applied as a moving window to each successive tone in a melody.

It was predicted that ITI and TV should mirror acceleration implied by the just-heard tone. Thus, for pitch contour, positive acceleration implied by a CP tone should initiate faster ITIs and be followed by faster TV. Conversely, CC tones imply negative acceleration and should initiate slower ITIs and be followed by slower TV. For pitch distance, increases should be perceived as a positive acceleration, initiating faster ITIs and faster TV, while decreases should be perceived as a negative acceleration and initiate slower ITIs and slower TV. The influences of pitch contour and pitch distance were assumed to be independent. However, interactions were also investigated by introducing an incongruence in implied acceleration between pitch contour and pitch distance once per melody and under a single circumstance: where a CP cell was a semitone following a tone, as in the sequence C-D-Eb. In this case, pitch contour implies positive acceleration whereas pitch distance implies negative acceleration.

Finally, an algorithm of melodic accent strength derived from perceptual data (Thomassen, 1982) was tested as a predictor of ITI and TV. It is discussed in detail below. Significant perception/production correlations might further support the hypothesis that perception and production of melodic accent emerge from shared action associations (Leman, 2009).

METHOD

Participants

12 undergraduates (9 females) aged 18-32 (M = 22; SD = 4.26) from Ryerson University participated. 6 had participated in a previous tapping experiment. Participants received 10 dollars in compensation. Years of private music lessons ranged from 0 to 7 (M = 2.21; SD = 2.41).

Apparatus

Tonal stimuli were assembled into a sound font and loaded into a software-based tone generator, which was interfaced with a Fatar VMK-161 MIDI keyboard. Stimulus presentation and data collection were done with FTAP (Finney, 2001). Stimuli were presented through headphones at 74 dB SPL. Left-handed participants (*n* = 2) tapped with their index finger on the lowest key of the keyboard (C3); right-handed

participants (*n* = 10) tapped on the highest key (C8). Output intensity was set to a fixed value to eliminate variations in feedback tone intensity between taps.

Stimuli and Procedure

Original isochronous melodies are displayed in Figure 1. Each melody contained the first five tones from either a C major scale (i.e., C D E F G) or a C minor scale (i.e., C D Eb F G). Two C major models (melodies 1 and 5) were each subjected to three different transformations — 1) retrograde, 2) transposition up a perfect fifth and inversion (which entails a change in tonality to C minor), and 3) retrograde of 2) — yielding a total of 8 melodies. Successive tones were 0, 1 or 2 semitones in pitch distance from previous. Melodies were composed so that a melody in the left column and its retrograde in the right column contained at least one occurrence of the 8 cells of interest (see Figure 6). The single incongruence within each melody was varied in melodic position and pitch direction. Motivic patterns were avoided in order to discourage modulated responses based on higher-order perceptual grouping.
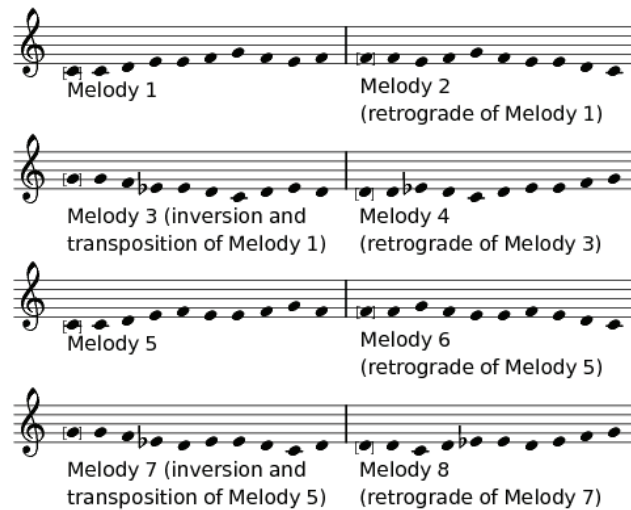
Fig. 1. The 8 original melodies used as feedback sequences. Tones in brackets were repeated 5 times preceding the melody. Responses to the final 8 tones were analysed.

The melodies served as feedback sequences in a continuation tapping task (Stevens, 1886). For each trial, participants first synchronised their taps to a pacing signal, which consisted of 15 ms square waves clicks presented at a fixed inter-onset interval (IOI) of 500 ms. Participants could begin synchronizing with the pacing signal at their discretion. At this point, taps produced no auditory feedback through the headphones. The sound of the pacing signal stopped after 7 taps, and participants now heard tones coinciding with each of their taps. Continuation phase tones were either sine or piano tones matched in intensity and onset/decay times and 200 ms in duration. Continuation phase sequences consisted of 15 tones: 5 repetitions of the first tone of a melody followed by the melody itself. Each successive tap triggered the next tone in the melody. The 5 repeating tones preceding the melody allowed for acceleration that often occurs at the transition between synchronization and continuation phases (Flach, 2005). These taps were not analysed. A trial ended with the 15th continuation phase tap, which repeated the final tone in the melody.

Participants were instructed to maintain the tempo of the pacing signal until the end of the trial and to ignore any changes in the pitch of the feedback tones. They were additionally instructed give an equal weight to each tap and to maintain contact between their fingertip and the key. Thus, deviations were presumed to be unintentional and could not be offset by a longer movement trajectory. To maintain task vigilance, after each trial participants were presented with the standard deviation of the continuation phase ITIs and instructed to try to achieve as low a score as possible.

5 practice trials in which the experimenter was present were followed by 4 blocks of 40 trials each. Trials were blocked by timbre. Participants completed 2 blocks for each timbre. Within a block, 5 repetitions of each of the 8 melodies were completed. Order of presentation for both the trials within a

block and the blocks themselves was randomised. In total, each melody was repeated 20 times. With breaks offered between blocks, the task took approximately 1 hour.

Data Analysis

For each of the 8 melodies, the final 8 continuation taps were subjected to analysis. "Note on" events were registered during the downward displacement by a sensor at a fixed position near the bottom of the key. ITIs were defined as the time difference between MIDI "note on" events. ITIs exceeding 800 ms were likely due to unregistered taps. These trials (1.7% of 1920) were discarded. TV for each "note on" event was a value on a linear scale from 1 to 127.

For each of the 64 tones (8 melodies x 8 tones per melody), a normalized measure of local deviation in ITI was obtained by calculating the percent deviation between successive ITIs according to the formula $100 * ((ITI_{n-1} / ITI_n) - 1)$. The same procedure was done for TV values except that, for ease of reading, the sign of the deviation was reversed. Thus, positive values indicate positive acceleration in both ITI and TV and negative values indicate negative acceleration.

Regression Model

Separate multiple regression analyses were performed on the grand average ITI and TV percent deviation values. There were five factors: Pitch Contour, Pitch Distance Change, Incongruency, Sequence Position, and Pitch Height.

For Pitch Contour, CP tones were coded as +1, CC tones were coded as -1, and repeating tones were coded as 0. Pitch Distance Change coded the occurrence of changes in the pitch distance (in semitones) between successive tones. Null changes between successive tones were coded as 0, increases of 1 and 2 semitones were coded as +1 and +2, respectively, and decreases were coded as -1 and -2. Thus, in the cell D-D-Eb, the Eb is coded as +1 because of the increase in pitch distance from 0 semitones between D-D to 1 semitone between D-Eb; in the cell C-D-E, the E is coded as 0 because there is a pitch distance of 2 semitones between both C-D and D-E, etc. An Incongruency value of 1 marked the final tone in a CP cell containing pitch distances of two semitones followed by one semitone (e.g., C-D-Eb); all other tones were coded as 0. Sequence Position was coded from 1 to 8 for each successive tone in a melody and was included to control for linear tempo drift over the course of a trial. Pitch Height (in semitones) was numbered 0 to 7 and controlled for the influence of the absolute pitch height of individual tones. Codes for a sample melody are shown in Figure 2.
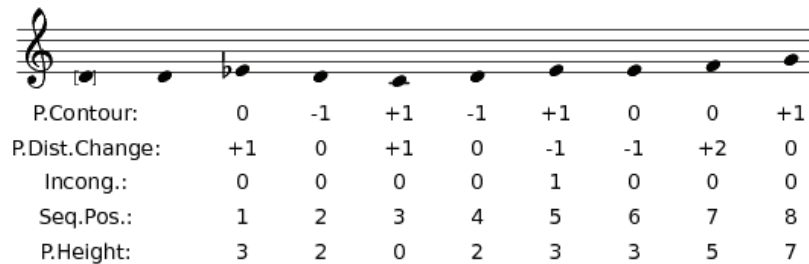

P.Contour:	0	-1	+1	-1	+1	0	0	+1
P.Dist.Change:	+1	0	+1	0	-1	-1	+2	0
Incong.:	0	0	0	0	1	0	0	0
Seq.Pos.:	1	2	3	4	5	6	7	8
P.Height:	3	2	0	2	3	3	5	7

Fig. 2. Regression codes for Melody 4.

RESULTS

Preliminary Analysis

Preliminary analysis of the raw ITI and TV values averaged for each participant revealed no global effects of timbre (piano vs. sine) or tonality (minor vs. major) on either mean ITI or mean TV (all Fs < 1). For the remaining analyses, values were pooled across timbres.

To test for learning effects, slope values from linear regressions of ITI and TV percent deviation against trial repetition number (20 - *n*, where *n* = the number of discarded trials) were calculated for each participant and for each of the 64 tones. It was reasoned that if modulate responses were based on expectancy of upcoming tones, percent deviation values should linearly change over the course of the experiment with emerging familiarity, producing either positive or negative slope values. However, across participants, none of the 64 mean slope values significantly deviated from zero ($p > .05$). In addition, for each participant and for each of the 64 tones, the standard deviation of the 20 - *n* responses was calculated. Separate repeated measures ANOVAs on standard deviation for ITI and TV with melody (8) and tone (8) as within subjects factors each produced null findings (all Fs < 1). Thus, neither pitch contour nor pitch distance affected response variability (Finney, 1997). Taken together, these data suggest that participants followed the instruction of ignoring the tones.

Next, grand average ITI and TV percent deviation values were obtained for each of the 64 tones by first averaging by participant across timbres and repetitions, and then overall.

AUTOCORRELATIONS FOR LAGS 1-4

Dependency between successive taps is typically observed as a negative lag 1 autocorrelation usually between 0 and -0.5 for both ITI (Wing & Kristofferson, 1973) and TV (Balasubramanian, Wing, & Daffertshofer, 2004; Billon & Semjen, 2006; Billon Semjen, & Stelmach, 1996; Wing, 1979). In other words, faster ITI and TV tend to be followed by slower ITI and TV, and vice versa. This may be due to random peripheral delays in execution triggered by a central timekeeper in an open loop system or a closed loop corrective response to the previous tap (Wing, 1977). Positive dependencies at lags greater than 1 might indicate the imposition of a periodic structure on the sequences, which might introduce artifactual deviations in ITI and TV unrelated to pitch contour or pitch distance. Thus, before entering the grand average ITI and TV values into regression analyses, autocorrelations at Lags 1 to 4 were calculated for each dependent variable and for each melody. ITI and TV autocorrelations averaged across the 8 melodies are displayed in Figure 3. As expected, Lag 1 values were significantly negative for both ITI, $t(7) = 3.92$, $p < .01$, and TV, $t(7) = 5.12$, $p < .001$, but did not differ significantly from zero at Lags 2 to 4 ($p > .05$). These data confirmed the presence of the negative lag 1 dependency typically observed and that participants did not impose a periodic structure on their taps at lags greater than 1.[3]
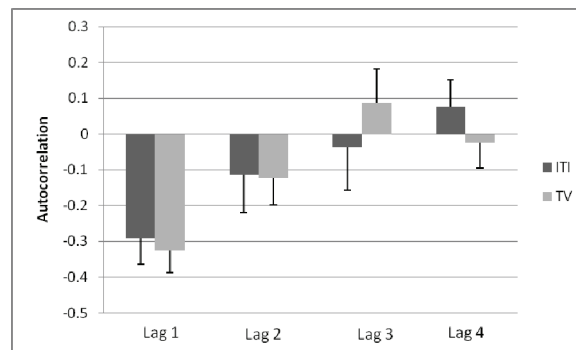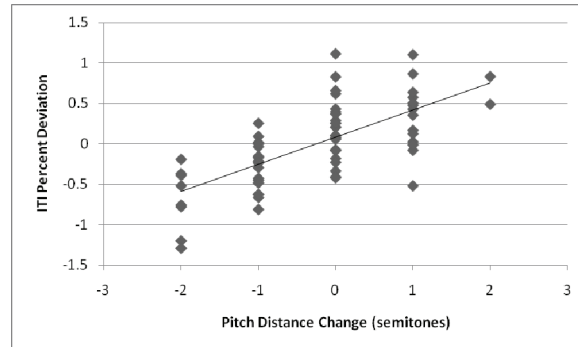

Fig.3. Autocorrelations at lags of 1 to 4 for ITI and TV

ITI Percent Deviation

Grand average ITI percent deviation values were entered into a multiple regression with 5 predictors: Pitch Contour, Pitch Distance, Incongruence, Sequence Position, and Pitch Height. The model was significant, $F(5,58) = 14.89$, $p < .0001$, explaining 56% of the variance. Just 2 predictors were significant: Pitch Distance, $t(63) = 8.04$, $p < .0001$, and Congruency, $t(63) = -3.02$, $p < .01$, explaining 47% and 6% of the variance, respectively.

The correlation between ITI percent deviation and Pitch Distance Change was $r = .69$, $p < .0001$. These data indicate that, as expected, ITI was faster when initiated by an increase in pitch distance and

slower when initiated by a decrease in pitch distance (see Figure 4). With Incongruence partialled out, the correlation was slightly higher, $r = .73$, suggesting that incongruence in tonal motion implied by pitch contour and pitch distance only modestly interfered with the priming effect of the latter on ITI.

Fig. 4. Correlation between grand average ITI percent deviation values for the 64 melody tones and Pitch Distance Change (in semitones) between successive tones.
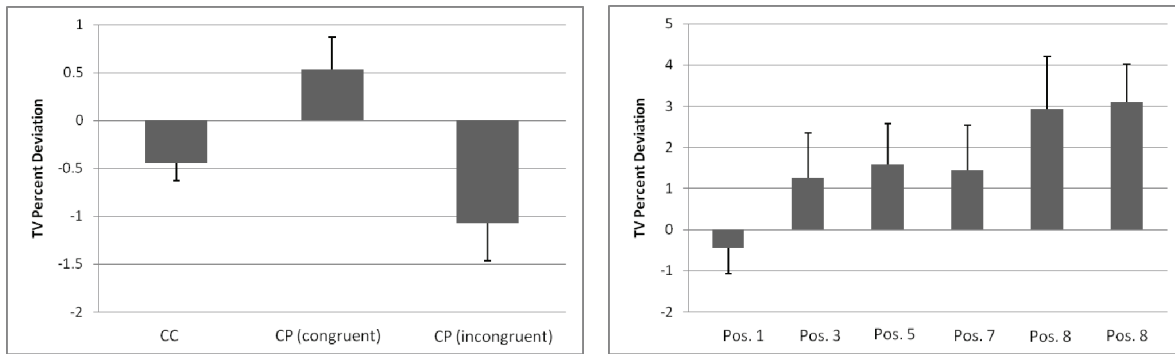

TV Percent Deviation

Grand average TV percent deviation values were entered into a multiple regression with the same 5 predictors. Although less successful than for ITI, the model was significant, $F(5,58) = 4.48$, $p < .01$. Three predictors were statistically significant — Pitch Contour, Incongruency, and Pitch Height — accounting for 28% of the variance.

The zero-order correlation for Pitch Contour was not significant, $r = .17$, $p < .17$, but emerged with Incongruence partialled out, $r = .34$, $p < .01$. This indicated that, as expected, where CC tones implied negative acceleration, the following TV was slower, and where CP tones implied positive acceleration, the following TV was faster. The negative correlation between TV and Incongruence, $r = -.30$, $p < .02$, indicated that the 8 incongruent tones elicited slower TV than the other 56 tones.

The effect of Incongruence on Pitch Contour was investigated further. Recall that incongruences only occurred in a single circumstance once in each melody: when a CP semitone change followed a two semitone change, as in the cell C-D-Eb. In these cases, pitch distance implied negative acceleration while pitch contour implied positive acceleration. Mean TV values for CC, CP (congruent), and CP (incongruent) tones are displayed in Figure 5 (left panel). These data indicate that congruent CP tones affected TV differently than incongruent CP tones. As expected, for congruent CP tones (either a semitone followed by a tone, such as F-E-D, or two successive tones, such as C-D-E), TV was faster (mean = .61%). However, TV for the 8 incongruent CP tones was slower in every case, with a mean of -1.26% that significantly deviated from zero, $t(7) = 4.21$, $p < .01$.

The unexpected negative correlation between TV and Pitch Height, $r = -.26$, $p < .04$, indicated that TV was faster following lower tones and slower following higher tones. Partial correlations were nearly identical, indicating that this association was independent of the other predictors. One possibility is that the presence of a tonal center (the lowest C tone) had a gravitational influence, with faster TV mirroring the acceleration properties of falling objects. As shown in Figure 5 (right panel), this is supported by the fact that, of the 6 occurrences of the C tone, faster TV scaled with its melodic position. TV was only slower following a C tone when it occurred in the first position in melody 8 (following repeated D tones at the transition between synchronization and continuation). On the other hand, where C occurred in the last position in melodies 2 and 6, TV was fastest, with percent deviations of 3.11% and 2.93%, respectively. These data suggest that over the course of a melody the status of C as a tonal center emerged with context. This finding will be discussed further in the Discussion section.

Fig. 5. (left panel) Mean TV percent deviation for CC, CP (congruent), and CP (incongruent) cells; (right panel) TV percent deviation in the 6 occurrences of the C note within melodies as a function of its melodic position.

In sum, these data supported the hypothesis that deviation in ITI and TV should mirror implied tonal acceleration in original melodies. However, whereas in the previous experiment (Ammirante, Thompson, & Russo, in press), a robust effect of pitch contour on both ITI and TV emerged with pitch distance held constant within trials, the introduction of changes in pitch distance within trials in the current study produced more complex interactions. Pitch distance had strong modulating effects on ITI, whereas pitch contour affected TV.

Recall that TV measured flexion velocity at a fixed point near the end of its trajectory. The dissociative effects of pitch distance on ITI and pitch contour on TV might suggest that after the onset of a tone, pitch distance influenced the initial extension trajectory, while pitch contour influenced the flexion trajectory that followed.[4] Doumas and Wing (2007) showed that in continuation tapping the extension-flexion trajectory comprising an ITI is asymmetrical. Extension velocity is only about half as fast as flexion velocity and the extension trajectory takes up a greater proportion of an ITI than the flexion trajectory. Thus, in the current study, an accelerated extension trajectory modulated by pitch distance should still have a strong impact on ITI even with little or no effect on flexion. In contrast, a modulating effect of pitch contour too late to influence extension but soon enough to modulate flexion should have had only a modest impact on ITI. The incongruence data further support this explanation. If it takes both early and late influences for a perceived incongruence between acceleration implied by pitch contour and pitch distance to emerge, then this should have primarily affected flexion. Thus, Incongruence had only a modest intervening effect on ITI, but substantially impacted TV as indicated by the fact that the significant TV/Pitch Contour partial correlation only emerged after removing Incongruence.

Thomassen's (1982) Model

A secondary goal in the current study was to assess previous perceptual ratings for melodic accent strength as predictors of deviation in ITI and TV. Thomassen (1982) presented listeners with pairs of stimuli each containing a single three-tone cell embedded within a longer monotone sequence. Manipulation of the pitch height of the second and third tones in the cell generated 9 possible cells (see Figure 6). Increased intensity of every fourth repeated tone established a metrical context before and after the cell. Participants indicated whether the cell in the second sequence better continued meter than the cell in the first sequence. Melodic accent strength values between 0 and 1 were derived for each cell type. Accent values were the proportion of positive responses. As shown in Figure 6, two sets of accent values were derived: one for cells in which the middle tone continued meter and one for which the final tone continued meter. When applied as a moving window to the tones in a melody, the algorithm predicts accent strength for a given tone within a melody to be the product of these two values.

The algorithm was applied to the 64 tones comprising the 8 melodies. The correlation between ITI and these values was $r = .35$, $p < .01$, indicating that faster ITIs were associated with stronger accent strength values. The partial correlation between the algorithm values and TV was significant with Pitch

Height and Incongruence removed, $r = .27$, $p < .04$. This indicates that faster TV was associated with stronger accent strength values.
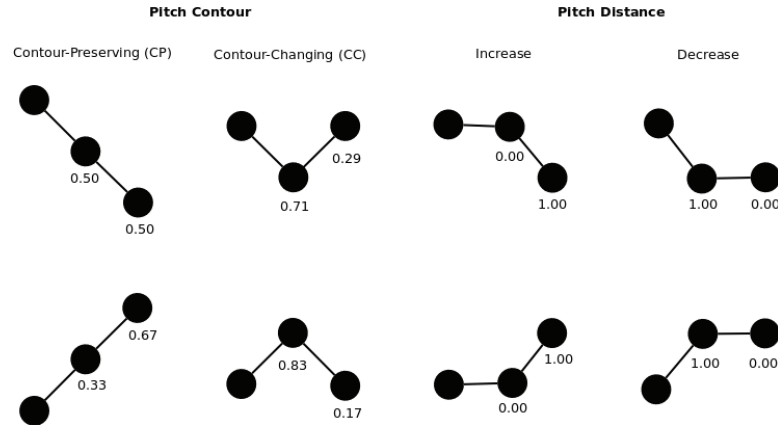
Fig. 6. Middle and final tone melodic accent ratings from Thomassen (1982) for the 8 cells contained in the stimuli for the current study (after Huron & Royal, 1996). When applied as a moving window to the tones in a melody, the algorithm predicts accent strength for a given tone within a melody to be the product of these two values.

DISCUSSION

The current study supported the hypothesis that in a continuation tapping task in which successive taps each triggered a discrete tone, apparent tonal motion implied by changes in pitch contour and pitch distance between tones should systematically modulate ITI and TV. These data extend a previous finding for random tonal sequences (Ammirante, Thompson, & Russo, in press) to original musical melodies. For pitch contour, CC tones elicited slower TV and CP tones elicited faster TV. These data are consistent with the proposal that interactions between timing and pitch contour may reflect implicit awareness of terrestrial locomotion: a moving body equipped with a vestibular system will slow down when navigating a turn but speed up when its trajectory is unidirectional (Boltz, 1998). For pitch distance, a decrease elicited a slower ITI and an increase elicited a faster ITI. These data are also consistent with implicit awareness of moving bodies; acceleration positively covaries with spatial extent traversed within a fixed time interval. There was no evidence of learning effects over 20 repetitions of each of the 8 melodies; nor was variability in ITI and TV affected by either pitch distance or pitch contour. Taken together, these data suggest that participants followed the instruction of ignoring the tones and were not distracted by them.

Based on Truslit's suggestion that the motion-like properties of musical melodies may "elicit in the listener a corresponding motion experience" (1993 [1938]: p. 50), it is proposed that systematic deviations in ITI and TV in the current study arose from participants' failure to disambiguate acceleration implied by apparent tonal motion from the acceleration of their finger trajectory. Common Coding theory predicts that, due to shared neural resources, perception may automatically modulate action where overlap exists between perceptual and motor representations (Prinz, Aschersleben, & Koch, 2009: p. 40; James 1890). Studies of the influence of task-irrelevant stimulus features on action support this prediction (Andres, et al. 2008; Glover et al., 2004; Liepelt et al., 2008; Simon, 1990).

Some unexpected dissociative effects of pitch distance and pitch contour on ITI and TV were observed. ITI was modulated by pitch distance whereas TV was modulated by pitch contour. Dissociations may have been due to pitch distance influencing the initial extension trajectory of a tap and pitch contour influencing the flexion trajectory that follows. In the absence of a direct measure of extension velocity, this interpretation is speculative. On the other hand, structural dissociations between cortical centers active during finger extension and flexion are supported by neuroimaging (Yue et al., 2000; Stenekes et al., 2006). In addition, a division of function is thought to occur in simple movements such as grasping, with initial opening of the grip aperture (extension) guided by an object's spatial properties and movement velocity calibrated during the subsequent flexion phase (Castiello, 2005; Stenekes et al., 2006). Thus, one possibility is that, rather than implying acceleration, the direction and magnitude of a change in pitch distance may

directly modulate calibration of spatial extent of movement during the extension phase. In this case, an accelerated extension trajectory might be a by-product of a motor representation of a longer movement.

There was also an unexpected finding with respect to the absolute pitch height of tones. Whereas using equidistant tones within trials in the previous experiment (Ammirante, Thompson, & Russo, in press), there was no influence of absolute pitch height, in the current study using C major and C minor scales, faster taps followed lower tones. One possibility is that the presence of a tonal center (C tone) exerted a gravitational influence, with TV mirroring the positive acceleration of falling bodies.

Truslit (1993 [1938]) argued for melody as a kind of motion simulation technology, intriguingly proposing that engagement of the vestibular system underlies this experience in the listener. Thus, the apparent translational (pitch distance) and gravitational (pitch contour, pitch height) forces proposed to account for deviations in ITI and TV correspond to those to which the hair cells in each of the two otolithic organs of the vestibular system — the utricle and saccule, respectively — are maximally responsive. Periodicities in the frequency ranges of musical tempo and tone have been shown to elicit vestibular responses (Todd, 1999). In support for the proposal that accent perception and production emerge from action associations originating in the engagement of the otolithic organs, Trainor et al. (2009) recently demonstrated that the perception of metrical accent may be biased by direct vestibular stimulation. It remains to be seen whether a vestibular response may be elicited by changes in pitch height between tones.

Implications for Melodic Accent

Finally, significant positive correlations were found between both ITI and TV deviation values and the accent strength values of the 8 melodies generated by an algorithm derived from perceptual data (Thomassen, 1982). Although often cited as supporting the perceptual salience of CC tones, inspection of Thomassen's data reveals a more complicated picture. In particular, as shown in the second column of Figure 6, middle tones that *preceded* a change of pitch direction received higher ratings of .71 and .83 than the ratings of .29 and .17 for the final tones that actually *initiated* a change of direction. Thus, at issue is whether it is the middle or the final tone of the cell that should be classified as CC. While the final tone as CC has been adopted here, proponents of the perceptual salience hypothesis have argued for the middle tone as CC (e.g., Pfordresher, 2003). This is presumably based on the assumption that accent strength is evaluated after the occurrence of the actual change in pitch direction and then attributed to the tone that preceded it. However, the data from the current and previous experiments demonstrate that the production of a melodic accent may precede a CC tone without succeeding context because responses are measured prior to its availability. In other words, a succeeding CC tone may only be triggered by another tap and, as such, should have no bearing on the previous response.

An account of the association between Thomassen's findings and the current findings is offered that demands no succeeding context. It seems plausible that participants in Thomassen's study may have gauged how well a tone continued the meter by overtly or covertly entraining an effector (e.g., finger or foot) to the preceding metrical context and continuing the metrical pattern during the presentation of the melodic cell. In this case, melodic accent ratings may have been a function of overlap between a motor representation of an accented movement trajectory and a perceptual representation of apparent tonal motion between the cell tones. It has been demonstrated that when participants deliberately produce an accent every nth tap in continuation tapping, both the TV of the accented tap and the ITI that the accented tap completes are faster relative to the following unaccented tap (Billon & Semjen, 1995; Billon, Semjen, & Stelmach, 1996). Thus, if melodic accent strength ratings were a function of representational overlap, then they should be higher for a melodic pattern that implies first positive and then negative acceleration.

This is illustrated in Figure 7 (left panel). At the middle tone in the cell, positive acceleration implied by the increase in pitch distance from the monotone metrical context to the initial change in pitch height overlaps with a motor representation of an accented tap (i.e., faster ITI and/or TV). At the final tone in the cell, negative acceleration implied by the change in pitch direction overlaps with a motor representation of an unaccented tap (i.e., slower ITI and/or TV). As such, the middle tone receives a strong rating of .83. On the other hand, in figure 7 (right panel), where the final tone in the cell is CP, the middle tone rating is weaker (.33). This may be due to the non-overlap between positive acceleration implied by the CP tone and an unaccented (i.e., slower) tap.
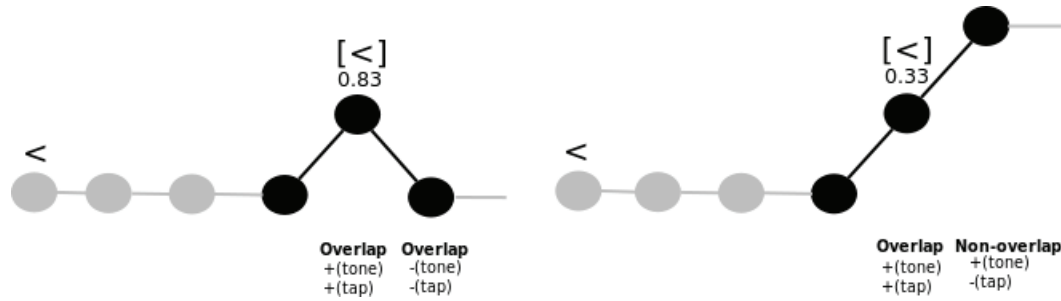
Fig. 7. An interpretation of Thomassen's findings based on action associations. The grey dots indicate the monotone metrical context preceding and succeeding the melodic cell (black dots). The unbracketed accent symbol indicates a dynamic accent. The bracketed accent symbol indicates an overtly or covertly produced accented movement. In the left panel, direction of acceleration for perceptual (tone) and motor (tap) representations overlaps at both the middle and final tones of the melodic cell, prompting a strong accent strength rating for the middle tone (.83). In the right panel, direction of acceleration for perceptual and motor representations does not overlap at the final tone, prompting a weaker rating for the middle tone (.33).

A similar case can be made for the relatively strong accent strength ratings for CP final tones (see Figure 6, column 1). Positive acceleration implied by the CP tone overlaps with an accented tap, and negative acceleration implied by the decrease in pitch distance at the onset of the succeeding monotone metrical context overlaps with an unaccented tap. As such, the final tone receives higher ratings. By contrast, the weaker accent strength ratings for the CC final tones (Figure 6, column 2) may be due to non-overlap between their implied negative acceleration and an accented tap.[5] In sum, what may be captured by Thomassen's algorithm, which predicts accent strength to be the product of a given tone's middle and final tone ratings, is the extent to which, like accented and then unaccented taps, positive and then negative acceleration are implied by the apparent motion of successive tones.

Concluding Remarks

Leman (2009: pp. 84-85) proposed that if 1) perception of musical sound is action-based and 2) motor and perceptual representations of action overlap, then it should be possible to reconstruct this perception through the investigation of action sequences. The findings from the current study based on predictions largely derived from perceptual data (e.g., Boltz, 1998) offer some support to this proposal. Further, the paradigm used here, in which prepared MIDI sequences are triggered by the participant via a simple motor task, offers a controlled means for such investigation in the non-musician population.

NOTES

[1] In a follow-up study, the effect of pitch contour was robust following a pacing signal of 350 ms IOI.

[2] Although intensity was experimentally controlled, under normal conditions faster taps would have also produced dynamic accents.

[3] An alternative regression model accounted for negative lag 1 dependency. In this model, ITI and TV were predicted to be directly influenced by the just-heard tone (i.e., at a lag of 0) and the indirectly influenced by the previously-heard tone (i.e., at a lag of 1) via a negative dependency with the previous tap. This model produced nearly identical results to those reported.

[4] An additional regression analysis investigated effects of the same 5 factors on deviations in *dwell time* between taps (i.e., the temporal interval between flexion offset [MIDI "note on" message], which initiates a tone, and extension onset [MIDI "note off" message]). Null findings indicated that acceleration in ITI could not be attributed to either early or late tap initiation. These data partially support the proposal that the pitch distance/ITI correlation was largely due to differences in acceleration during the extension trajectory. On

the other hand, it does not rule out the contributions of 1) dwell time between extension and flexion where the finger reaches its highest amplitude and/or 2) flexion acceleration at some other point in the trajectory as additional contributors to accelerated ITI. This is currently being investigated using motion capture.

[5] It is also noted that for downward CP tones (top left in Figure 6), the rating for both middle and final values is .50. This finding may also be rooted in overlapping action associations, reflecting congruence at both middle and final tones between an either covertly or overtly produced flexion trajectory and a downward change in pitch direction.

REFERENCES

Ammirante, P., Thompson, W. F., & Russo, F. A. (in press). Ideomotor effects of pitch in continuation tapping. *Quarterly Journal of Experimental Psychology*.

Andres, M., Ostry, D. J., Nicol, F., & Paus, T. (2008). Time course of number magnitude interference during grasping. *Cortex*, Vol. 44, No. 4, pp. 414-419.

Balasubramaniam, R., Wing, A. M., & Daffertshofer, A. (2004). Keeping with the beat: Movement trajectories contribute to movement timing. *Experimental Brain Research*, Vol. 159, pp. 129-134.

Billon, M., & Semjen, A. (1995). The timing of accent production in synchronization and continuation tasks performed by musicians and nonmusicians. *Psychological Research*, Vol. 58, No. 3, pp. 206-217.

Billon, M., Semjen, A., & Stelmach, G. E. (1996). The timing effects of accent production in periodic finger-tapping sequences. *Journal of Motor Behavior*, Vol. 28, No. 3, pp. 198-210.

Boltz, M. G. (1998). Tempo discrimination of musical patterns: Effects due to pitch and rhythmic structure. *Perception & Psychophysics*, Vol. 60, No. 8, pp. 1357-1373.

Castiello, U. (2005). The neuroscience of grasping. *Nature Reviews, Neuroscience*, Vol. 6, pp. 726-736.

Cohen, J., Hansel, C. E. M., & Sylvester, J. D. (1954). Interdependence of temporal and auditory judgments. *Nature*, Vol. 174, pp. 642-644.

Crowder, R. G., & Neath, I. (1994). The influence of pitch and time perception in short melodies. *Music Perception*, Vol. 12, No. 4, pp. 379-386.

Doumas, M., & Wing, A. (2007). Timing and trajectory in rhythm production. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 33, No. 2, pp. 442-455.

Drake, C. (1993). Perceptual and performed accents in musical sequences. *Bulletin of the Psychonomic Society*, Vol. 31, No. 2, pp. 107-110.

Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception*, Vol. 10, No. 3, pp. 343-378.

Drake, C., Dowling, J. W., & Palmer, C. (1991). Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Perception*, Vol. 8, No. 3, pp. 315-334.

Finney, S. A. (1997). Auditory feedback and musical keyboard performance. *Music Perception*, Vol. 15, No. 2, pp. 153-174.

Finney, S. A. (2001). FTAP: A Linux-based program for tapping and music experiments. *Behaviour Research Methods, Instruments & Computers*, Vol. 33, No. 1, pp. 65-72.

Flach, R. (2005). The transition from synchronization to continuation tapping. *Human Movement Science*, Vol. 24, No. 4, pp. 465-483.

Glover, S., Rosenbaum, D. A., Graham, J., & Dixon, P. (2004). Grasping the meaning of words. *Experimental Brain Research*, Vol. 154, No. 1, pp. 103-108.

Henry, M. J., & McCauley, J. D. (2009). Evaluation of an imputed pitch velocity model of the auditory kappa effect. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 35, No. 2, pp. 551-564.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioural and Brain Sciences*, Vol. 24, pp. 849-878.

Huron, D., & Royal, M. (1996). What is melodic accent? Converging evidence from musical practice. *Music Perception*, Vol. 13, No. 4, pp. 489-516.

James, W. (1890). *The principles of psychology* (Vol. 2). New York: Holt.

Leman, M. (2009). *Embodied music cognition and mediation technology*. Cambridge, MA: MIT Press.

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

Lidji, P., Kolinsky, R., Lochy, A., & Morais, J. (2007). Spatial associations for musical stimuli: A piano in the head? *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 33, No. 5, pp. 1189-1207.

Liepelt, R., Von Cramon, D. Y., & Brass, M. (2008). What is matched in direct matching? Intention attribution modulates motor priming. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 34, No. 3, pp. 578-591.

MacKenzie, N. (2007). The kappa effect in pitch/time context. Doctoral dissertation. Ohio State University, 2007. *Dissertation Abstracts International*, Vol. 68, 132 pages.

McAnally, K. (2002). Timing of finger tapping to frequency modulated acoustic stimuli. *Acta Psychologica*, Vol. 109, No. 33, pp. 331-338.

Penel, A., & Drake, C. (1998). Sources of timing variations in music performance: A psychological segmentation model. *Psychological Research*, Vol. 61, No. 1, pp. 12-32.

Pfordresher, P. Q. (2003). The role of melodic and rhythmic accents in musical structure. *Music Perception*, Vol. 20, No. 4, pp. 431-464.

Prinz, W., Aschersleben, G., & Koch, I. (2009). Cognition and action. In: E. Morsella, J.A. Bargh & P.M. Gollwitzer (Eds.), *Oxford handbook of human action.* New York: Oxford University Press, pp. 35-71.

Repp, B. (1992). Probing the cognitive representation of musical time: Structural constraints on the perception of timing perturbations. *Cognition,* Vol. 44, pp. 241-281.

Repp, B. (1995). Detectability of duration and intensity increments in melody tones: A partial connection between music perception and performance. *Perception & Psychophysics*, Vol. 57, No. 8, pp. 1217-1232.

Repp, B. (1998). Musical motion in perception and performance. In: D. A. Rosenbaum & C. E. Collyer (eds.), *Timing of behavior: Neural, Psychological, and Computational Perspectives*. Cambridge, MA: MIT Press, pp. 125-144.

Repp, B. (1999). Control of expressive and metronomic timing in pianists. *Journal of Motor Behaviour*, Vol. 31, No 2, pp. 145-164.

Repp, B., & Knoblich, G. (2007). Action can affect auditory perception. *Psychological Science*, Vol. 18, No. 1, pp. 6-7.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, Vol. 27, pp. 169-192.

Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition*, Vol. 99, No. 2, pp. 113-129.

Sergent, V., Hellige, J. B., & Cherry, B. (1993). Effects of responding hand and concurrent verbal processing on time-keeping and motor-implementation processes. *Brain and Cognition*, Vol. 23, pp. 243–262.

Shigeno, S. (1986). The auditory tau and kappa effects for speech and nonspeech stimuli. *Perception & Psychophysics*, Vol. 40, No. 1, pp. 9-19.

Shigeno, S. (1993). The interdependence of pitch and temporal judgments by absolute pitch processors. *Perception & Psychophysics*, Vol. 54, pp. 682-692.

Shove, P., & Repp. B. (1995). Music motion and performance. Theoretical and empirical perspectives. In: J. Rink (Ed.), *The practice of performance: Studies in musical interpretation*, Cambridge: Cambridge University Press, pp. 55-83.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In: R. W. Proctor & T. G. Reeve (Eds.), *Stimulus-response compatibility: An integrated perspective*, Amsterdam: North Holland, pp. 31-86.

Stenekes, M. W., Hoogduin, J. M., Mulder, Th. Geertzen, J H. B., Leenders, K. L., Nicolai, J.-P. A., & de Jong, B. M. (2006). Functional dominance of finger flexion over extension, expressed in left parietal activation. *NeuroImage*, Vol. 32, No. 2, pp. 676-683.

Stevens, L. T. (1886). On the time sense. *Mind*, Vol. 11, No. 43, pp. 393-404.

Tekman, H. G. (2001). Accenting and detection of timing variations in tone sequences: Different kinds of accents have different effects. *Perception & Psychophysics*, Vol. 63, No. 3, pp. 514-523.

Thomassen, J. M. (1982). Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America*, Vol. 71, No. 6, pp. 1596-1605.

Todd, N. P. M. (1999). Motion and music: A neurobiological perspective. *Music Perception*, Vol. 17, No. 1, pp. 115-126.

Trainor, L., Gao, X., Lei, J., Lehtovaarra, K., & Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. Cortex, Vol. 45, No. 1, pp. 35-43.

Truslit, A. (1993 [1938]). Music as motion: A synopsis of Alexander Truslit's *Gestaltung und Bewegung in der Musik* (B. Repp, Trans. and Ed.). *Psychology of Music*, Vol. 21, pp. 48-72.

Viviani P., & Stucchi, N. (1992). Biological movements look uniform: Evidence of motor-perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 18, No. 3, pp. 603-623.

Wing, A. (1977). Perturbations of timing feedback delay and the timing of movement. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 3, No. 2, pp. 175-186.

Wing, A. (1979). Timing of movement phases of a repeated response. *Journal of Motor Behaviour*, 12, 113-124.

Wing, A., & Kristofferson, A. (1973). Response delays and the timing of discrete motor responses. *Perception and Psychophysics*, Vol. 14, No. 1, pp. 5-12.

Yoblick, D., & Salvendy, G. (1970). Influence of frequency on the estimation of time for auditory, visual, and tactile modalities: The kappa effect. *Journal of Experimental Psychology*, Vol. 86, No. 2, pp. 157-164.

Yue, G. H., Liua, J. Z., Siemionowa, V., Ranganathana, V. K., Ng, T. C. & Sahgal, V. (2000). Brain activation during human finger extension and flexion movements. *Brain Research*, Vol. 856, No. 1-2, pp. 291-300.