Journal of Experimental Psychology:
Human Perception and Performance
2009, Vol. 35, No. 2, 551–564

0096-1523/09/$12.00 DOI: 10.1037/0096-1523.35.2.551

Evaluation of an Imputed Pitch Velocity Model of the Auditory Kappa Effect

Molly J. Henry and J. Devin McAuley
Bowling Green State University

Three experiments evaluated an imputed pitch velocity model of the auditory kappa effect. Listeners heard 3-tone sequences and judged the timing of the middle (target) tone relative to the timing of the 1st and 3rd (bounding) tones. Experiment 1 held pitch constant but varied the time (T) interval between bounding tones (T = 728, 1,000, or 1,600 ms) in order to establish baseline performance levels for the 3 values of T. Experiments 2 and 3 combined the values of T tested in Experiment 1 with a pitch manipulation in order to create fast (8 semitones/728 ms), medium (8 semitones/1,000 ms), and slow (8 semitones/1,600 ms) velocity conditions. Consistent with an auditory motion hypothesis, distortions in perceived timing were larger for fast than for slow velocity conditions for both ascending sequences (Experiment 2) and descending sequences (Experiment 3). Overall, results supported the proposed imputed pitch velocity model of the auditory kappa effect.

Keywords: kappa effect, auditory motion, imputed velocity, perceptual interdependence, timing

Lawful movement trajectories of objects in our everyday environment afford predictions about when a moving object will arrive where. The kappa effect, initially demonstrated for visual stimuli, is the phenomenon whereby deviations from expected stimulus spacing, based on an implied space–time trajectory, tend to produce systematic distortions in perceived stimulus timing. For example, altering the spacing of two adjacent transient visual stimuli in a three-element sequence so that they are spaced closer together or farther apart than expected tends to shorten or lengthen, respectively, the perceived time interval between the two stimulus onsets (Cohen, Hansel, & Sylvester, 1953, 1955; Huang & Jones, 1982; B. Jones & Huang, 1982; Matsuda & Matsuda, 1979, 1981; Price-Williams, 1954; Sarrazin, Giraudo, Pailhous, & Bootsma, 2004). In this article, we report three experiments that investigated an auditory version of the kappa effect in which participants judged the timing of individual elements (tones) in auditory displays (tone sequences) where we manipulated the pitch spacing of tones rather than their separation in physical space.

The Visual Kappa Effect: Dependence of Perceived Time on Space

The canonical task used to investigate the visual kappa effect involves the sequential presentation of three transient visual stimuli followed by a judgment about the timing of the middle (target) stimulus, with the critical instruction to ignore stimulus spacing (Abbe, 1936; Abe, 1935; Cohen et al., 1953, 1955; Huang & Jones, 1982). A diagram of the canonical task is shown in Figure 1. For the canonical task, the spatial and temporal positions of the first and third (bounding) sequence elements are fixed, while the spacing and timing of the middle (target) element vary relative to the spatial and temporal bisection points of the bounding elements. Observers typically judge whether the target stimulus is closer in time to either the first or third bounding stimulus by making a binary *short–long*/*long–short* decision or by explicitly adjusting the timing of the target stimulus so that it subjectively bisects the temporal interval marked by the bounding elements. Under these conditions, observers tend to make systematic timing errors based on to-be-ignored stimulus spacing. In general, when the target stimulus is spaced closer to the first stimulus than the third, observers tend to perceive that the target stimulus is closer in time to the first stimulus than the third, regardless of the target's objective temporal position. Similarly, when the target stimulus is spaced closer to the third stimulus than the first, observers tend to perceive that the target stimulus is closer in time to the third stimulus than the first (Cohen et al., 1953, 1955; Huang & Jones, 1982). Thus, the visual kappa effect is characterized by systematic distortions in perceived timing that are in the direction of the to-be-ignored change in stimulus spacing. Early seminal work on the visual kappa effect reported directional distortions in timing on the order of 12% (Abbe, 1936; Abe, 1935; Cohen et al., 1953).

The dominant explanation of the visual kappa effect is based on an *imputed velocity hypothesis* (Cohen et al., 1955; B. Jones & Huang, 1982; Price-Williams, 1954). This hypothesis proposes that properties of moving objects in the environment are applied by the perceiver to stimulus displays made up of successive transient sequence elements, even though the displays are not explicitly

Molly J. Henry and J. Devin McAuley, Department of Psychology, Bowling Green State University.

Portions of this research were presented at the 48th annual meeting of the Psychonomic Society on November 15–18, 2007, in Long Beach, California. We gratefully acknowledge numerous comments on this work by the members of the Rhythm, Attention, Perception Lab at Bowling Green State University. We are also grateful to Marta Zaleha, Nick Steuver, and Danielle Champney for their help with data collection.

Correspondence concerning this article should be addressed to Molly J. Henry, Department of Psychology, Bowling Green State University, Bowling Green, OH 43403. E-mail: mjhenry@bgnet.bgsu.edu

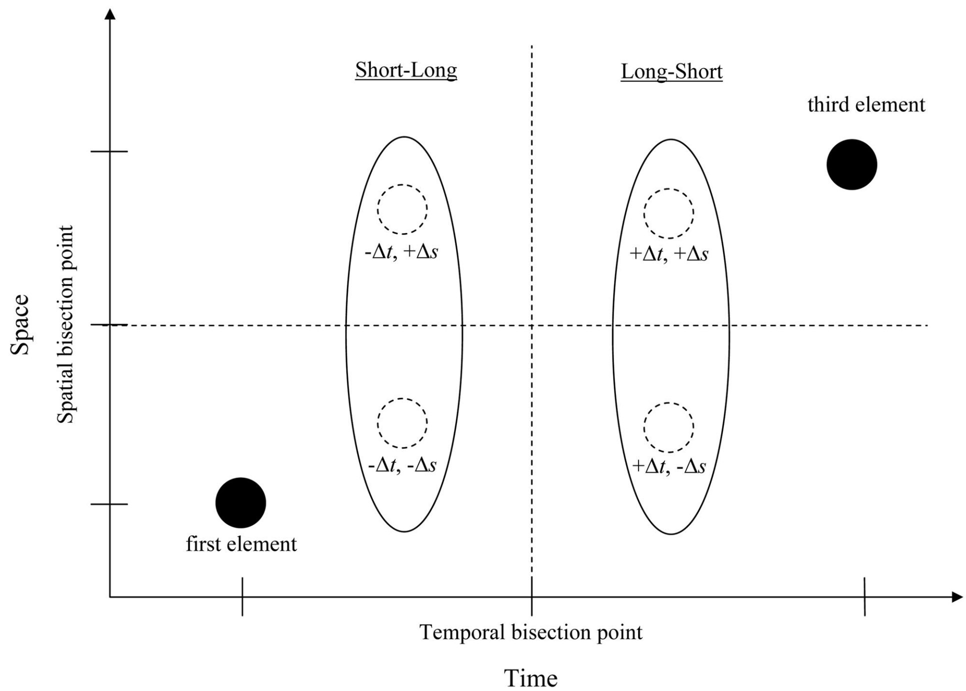

Figure 1. The canonical kappa task. Participants experience three element sequences; the first and third bounding elements (closed circles) have fixed positions in space and time, while the spatial and temporal position of the target element (open circles) is varied from trial to trial relative to the spatial and temporal bisection point of the bounding elements. Participants judge the temporal pattern of the two time intervals delineated by each three-element sequence by responding *short–long* or *long–short*.

moving per se (Cohen et al., 1955; Price-Williams, 1954). However, because perceivers have a wealth of experience with moving objects in the environment that show coherent transformations in space and time constrained by physical laws, they tend to naturally apply these laws to artificial stimulus displays (Shepard, 1984). The implication for transient (and somewhat unnatural) stimulus sequences that are regularly spaced and unfolding in time at a regular rate is that perceivers will impute constant velocity to the display (Cohen et al., 1955; M. R. Jones, Maser, & Kidd, 1978). In the context of the canonical task, observers are hypothesized to generate expectancies about the *where* and *when* of a target sequence element by interpolating the space–time trajectory implied by the fixed spatiotemporal positions of the initial and final bounding elements. When the middle (target) sequence element then violates the constant velocity assumption along the spatial dimension, perceptual adjustments must be made along the temporal dimension in order to maintain the sense of a constant velocity across the display. From this perspective, the critical independent variable with respect to the kappa effect is velocity ($\Delta s/\Delta t$), rather than absolute temporal or spatial separation of the sequence elements.

To provide a more formal quantitative account of visual kappa findings, B. Jones and Huang (1982; see also Anderson, 1974) developed an imputed velocity model, in which judgments about the timing of the target stimulus were considered to be a weighted combination of the objective duration of the first sequence time interval and an expectation for the amount of time needed to traverse the corresponding spatial interval at constant velocity. This model can be formalized as follows:

$$R = f[wt + (1 - w)E(t)] \tag{1}$$

In Equation 1, we will use t to refer to the objective duration of the first time interval in the pair of intervals delineated by the three-element sequences; $E(t)$ then represents the expected duration of the same time interval based on the imputed velocity of the sequence. The expected duration, $E(t)$, is determined by dividing the distance marked by the first spatial interval by imputed velocity, $E(t) = \Delta s/V$; this equation follows from simple algebraic manipulation of the equation for velocity, $V = \Delta s/\Delta t$. The function f maps perceived target timing into a response probability; for the version of the canonical task implemented in the current studies, f was a sigmoid function that mapped the weighted combination of t and $E(t)$ to a probability of a *long–short* response. The weight parameter, $w \in [0,1]$, specifies the relative contribution of t and $E(t)$ to participants' judgments about target timing. When $w = 1$, the objective timing of the target is the sole contributor to perceived duration, and there is no kappa effect. However, when $w < 1$, the perceived timing of the target stimulus is based partially on the expected target timing derived from imputed velocity, and a kappa effect is observed. Smaller estimated values of w correspond to larger kappa effects. In sum, the imputed velocity model predicts that the perceived timing of the target in the canonical kappa task is based on a combination of the objective timing and the expected timing of the target stimulus; the expected timing of the target is derived from imputed velocity and objective stimulus spacing. Model fits to empirical data provide an estimate of the parameter w, which quantifies the relative contribution of imputed velocity to perceived target timing.

Applying the imputed velocity model, Huang and Jones (1982) demonstrated that, at fast velocities, estimated values of w were much smaller than at relatively slow velocities, indicating a greater reliance on expected target timing, $E(t)$, that is, a larger kappa effect. Over the range of velocities examined, estimated values of w varied from 0.10 to 0.90. Huang and Jones, however, were unable to reliably demonstrate kappa effects for all observers in the fastest velocity conditions examined (4° visual angle/90 ms, 4° visual angle/170 ms, Experiment 1). Collyer (1977) was similarly unable to obtain consistent kappa effects at very fast velocities, suggesting the existence of an upper velocity limit to the visual kappa effect. Conversely, at the other end of the velocity scale, Cohen et al. (1953) reported that the magnitude of the kappa effect diminishes, indicative of a lower (slow) velocity limit to the visual kappa effect. In terms of the imputed velocity model, the lower velocity limit to the visual kappa effect corresponds to an estimated w value near 1.0.

In sum, the visual kappa effect is the observation that perceived stimulus timing may be systematically affected by stimulus spacing, critically, even when the spatial dimension is to be ignored. The dominant explanation for the visual kappa effect is based on an imputed velocity hypothesis that posits that perceivers tend to impute constant velocity to displays that exhibit motionlike properties; moreover, perceivers are assumed to develop expectancies for the future space–time locations of sequence elements making up these displays. Violation of these expectancies along the to-be-ignored spatial dimension yields perceptual distortions along the to-be-judged temporal dimension. Overall, the visual kappa effect is found within only a limited range of velocities; increasing sequence velocity within this range increases the magnitude of the visual kappa effect (Cohen et al., 1953; Huang & Jones, 1982; B. Jones & Huang, 1982). Of interest in the current research is whether velocity in the context of an auditory pitch manipulation similarly modulates the magnitude of the auditory kappa effect, to which we now turn.

The Auditory Kappa Effect: Dependence of Perceived Time on Pitch

There have been only a small number of studies that have considered the possibility of an auditory kappa effect (Cohen, Hansel, & Sylvester, 1954; Crowder & Neath, 1994; MacKenzie, 2007; Shigeno, 1986, 1993; Yoblick & Salvendy, 1970). Most of these have manipulated the timing and pitch of sequence elements rather than their separation in physical space (see, however, Sarrazin, Giraudo, & Pittenger, 2007). The motivation for conceptualizing pitch as a spatial dimension comes from a number of sources.

First, in music, notes are typically thought of as having a pitch that varies in tone height; that is, musical notes can be "high" or "low." This is visually apparent by the way that many cultures represent musical notes on a stave. Second, an analogy can be made at the neural level between the topographic organization of the response of the retina to light and the tonotopic organization of the response of the basilar membrane to sound (Crowder & Neath, 1994). Third, there is increasing evidence that mental representations for musical pitch may overlap to some degree with mental representations of physical space. For example, in one recent study, Rusconi, Kwan, Giordano, Umiltà, and Butterworth (2006) found that during performance of a pitch discrimination task, listeners responded more quickly and accurately when the vertical arrangement of response buttons was congruent with the relative pitch height of the correct response. In another recent study, Douglas and Bilkey (2007) found that individuals with normal hearing tend to show interference when performing a pitch discrimination task simultaneously with a mental rotation task, but that tone-deaf individuals do not tend to show this pattern of interference. Further strengthening the connection between pitch and space, tone-deaf individuals tended to show marked impairments on a mental rotation task relative to controls. In sum, there is increasing converging evidence that pitch embodies properties of a spatial dimension.

In the canonical version of the auditory kappa task involving tone sequences, the timing and pitch distance between the first and third elements of a three-tone sequence are held constant, establishing a constant pitch velocity; the timing and pitch of the middle (target) tone vary from trial to trial relative to the pitch–time bisection point of the bounding tones. Participants judge the timing of the target tone, ignoring pitch. If the first two tones are closer together or farther apart in pitch space than expected, then the target tone tends to be perceived as being closer in time to the first or third tone, respectively (Cohen et al., 1954; Crowder & Neath, 1994; MacKenzie, 2007; Shigeno, 1986, 1993).

The same principles used to develop an imputed velocity model of the visual kappa effect can also be applied to the auditory kappa effect (Hass & Hass, 1984). Specifically, pitch velocity can be measured as change in pitch per change in unit time, $V_p = \Delta p/\Delta t$. Thus, in the context of the canonical auditory kappa task, the perceived timing of the target tone is assumed to be based on a weighted combination of the objective timing of the target, t, and the expected timing, $E(t)$, derived from imputed pitch velocity and the distance of the target tone in pitch space from the initial bounding tone, $E(t) = \Delta p/V_p$. As was illustrated in Equation 1 for vision, the weight parameter, w, estimates the relative contributions of t and $E(t)$ to perceived target timing, with smaller values of w associated with larger kappa effects.

In general, studies of the auditory kappa effect have reported weaker effects than those observed in visual kappa studies, with the magnitude of directional errors in temporal judgments on the order of 4%. Indeed, the size of the auditory kappa effect reported by Cohen et al. (1954) was very small in one experiment and failed to reach statistical significance in a second. The failure of Cohen and colleagues to demonstrate a reliable auditory kappa effect may be because velocity varied between two very different values across trials; to vary velocity, they held the temporal separation of the bounding tones constant at 1,500 ms, while the separation of the bounding tones in pitch space was 1000 Hz in one condition and 168.5 Hz in the other. The relevance of this aspect of the experimental design is that, in vision, the kappa effect can be abolished by holding the temporal separation of the bounding tones constant and allowing the physical distance delimited by the sequence elements to vary between two very different values across trials, suggesting that under such conditions, observers are unable to develop a stable velocity referent (Matsuda & Matsuda, 1981). Analogously, it is possible that the two velocities tested by Cohen et al. spanned too large a range for listeners to develop a stable representation of pitch velocity and thus to show a reliable auditory kappa effect. To eliminate this possibility in the present study, we held pitch velocity constant across trials.

The variable velocity explanation of the comparatively weaker auditory kappa effect does not seem likely, however, to completely account for the difference between modalities. Recent studies that have held pitch velocity constant across trials have still tended to find auditory kappa effects of smaller magnitude than those typically observed in vision (Shigeno, 1986). An alternative explanation for comparatively weaker kappa effects in the auditory domain is that the temporal separation of the bounding tones in previous studies may have been too large for listeners to readily impute a constant pitch velocity. Some support for this view comes from an auditory motion hypothesis proposed by Jones and colleagues (M. R. Jones, 1976; M. R. Jones et al., 1978; M. R. Jones & Yee, 1993; MacKenzie, 2007; MacKenzie & Jones, 2005).

The *auditory motion hypothesis* proposes that there is a limited range of pitch velocities that give rise to the perception of motionlike relations in auditory sequences. Within this range, listeners extrapolate the implied pitch–time trajectory of the auditory sequence and use this information to make predictions about the future arrival time of a stimulus at a particular location in pitch space (M. R. Jones, 1976; M. R. Jones et al., 1978; M. R. Jones & Yee, 1993). M. R. Jones et al. (1978) proposed that the case of constant velocity—similar to the imputed velocity explanation of the visual kappa effect—was the simplest with respect to auditory motion and, moreover, that if a pitch–time event deviated from the implied trajectory, listeners would mentally simplify the pitch–time relation to conform to constant velocity. According to the auditory motion hypothesis, the auditory kappa effect should be strongest for pitch velocities that fall within the range that induces attention to motionlike properties of tone sequences. As pitch velocity approaches the upper limit of this range, the induced feeling of motion should become stronger and the magnitude of the kappa effect should increase. Conversely, as pitch velocity approaches the lower limit of this range, the magnitude of the kappa effect should decrease, and the effect should eventually disappear. In the present research, we consider the possibility that previous studies failed to observe a robust auditory kappa effect because the fixed time interval between the bounding tones, which varied between 1,400 ms and 2,800 ms across studies, yielded velocity conditions that were on the lower (slow) end of the range expected to induce a sense of auditory motion. Thus, for the tested range of velocities, the auditory kappa effect would be expected to be weak.

Overview

The aims of this research were twofold. First, we were interested in evaluating the proposed imputed pitch velocity model of the auditory kappa effect. Second, we were interested in testing the auditory motion hypothesis. Three experiments were conducted. In all experiments, participants listened to three-tone sequences and indicated whether they heard a *short–long* or *long–short* pattern. Experiment 1 was a baseline study that examined participants' ability to detect deviations in the timing of the middle (target) tone for constant pitch conditions. The time interval (T) between the onset of the first and third (bounding) tones varied between subjects and took on one of three values: T = 728 ms; 1,000 ms; or 1,600 ms. In Experiments 2 and 3, we investigated the canonical version of the auditory kappa task for the same T values as in Experiment 1. The specific values of T were selected in order to extend previous research on the auditory kappa effect to a wider range of velocities, which we hypothesized would lead to a stronger sense of auditory motion and thus larger kappa effects than previously observed. In Experiment 2, sequences ascended in pitch, whereas in Experiment 3, sequences descended in pitch. The pitch distance between bounding elements was fixed at 8 semitones (ST), yielding a pitch bisection point of 4 ST for the middle (target) tone and three between-subjects velocity conditions: fast (8 ST/728 ms), medium (8 ST/1,000 ms), and slow (8 ST/1,600 ms); expressed in semitones per second, the fast, medium, and slow velocity conditions were 11 ST/s, 8 ST/s and 5 ST/s, respectively. Both pitch and timing of the middle (target) tone varied from trial to trial, but a critical point was that participants were instructed to ignore pitch when making *short–long* and *long–short* judgments. On the basis of the auditory motion hypothesis, increasing pitch velocity was expected to increase the magnitude of the auditory kappa effect for both ascending and descending tone sequences.

Experiment 1: Baseline

Method

Design. Experiment 1 implemented a 3 × 8 mixed factorial design. The time interval (T) separating the onsets of the first and third (bounding) tones took on values of T = 728 ms, 1,000 ms, or 1,600 ms. These three values of T were crossed with eight time levels (± 4%, 8%, 12%, 16%), where time level referred to the position of the middle tone in each three-tone sequence relative to the temporal bisection point (T/2) of the first and third tones (see Figure 2A). The value of T varied between subjects, while time level varied within subjects.

Participants. Twenty-seven Bowling Green State University undergraduates (n = 16 women) between the ages of 18 and 25 years participated in return for course credit. All participants self-reported normal hearing and had a range of formal musical training (M = 2.74 years, SD = 2.78). Participants were randomly assigned to one of the three T values (n = 9 in all conditions). One participant in the T = 728 ms condition was excluded from the final sample due to inattention to the task. Final participant numbers for T = 728 ms, 1,000 ms, and 1,600 ms were therefore n = 8, n = 9, and n = 9, respectively.

Stimuli and equipment. All stimuli were 100-ms tones with a 329.63-Hz fundamental frequency (corresponding in musical terms to the note E4). Stimulus generation and response collection were controlled by an IBM PC-compatible computer running the MIDILAB software package, with a resolution of ≈1 ms (Todd, Boltz, & Jones, 1989). Tone sequences were presented at a comfortable listening level through Beyerdynamic DT770 headphones controlled by a Rane HC6 headphone console and attached to a Yamaha TX81Z tone generator. Responses were made by pressing one of two horizontally aligned buttons on a MIDILAB response box: the left button for *short–long* responses and the right button for *long–short* responses.

Procedure. Participants were instructed that for each three-tone sequence, the middle tone would "move around" in time, thus altering the pattern of durations delimited by the tone sequence. Participants were asked to respond *short–long* when they thought

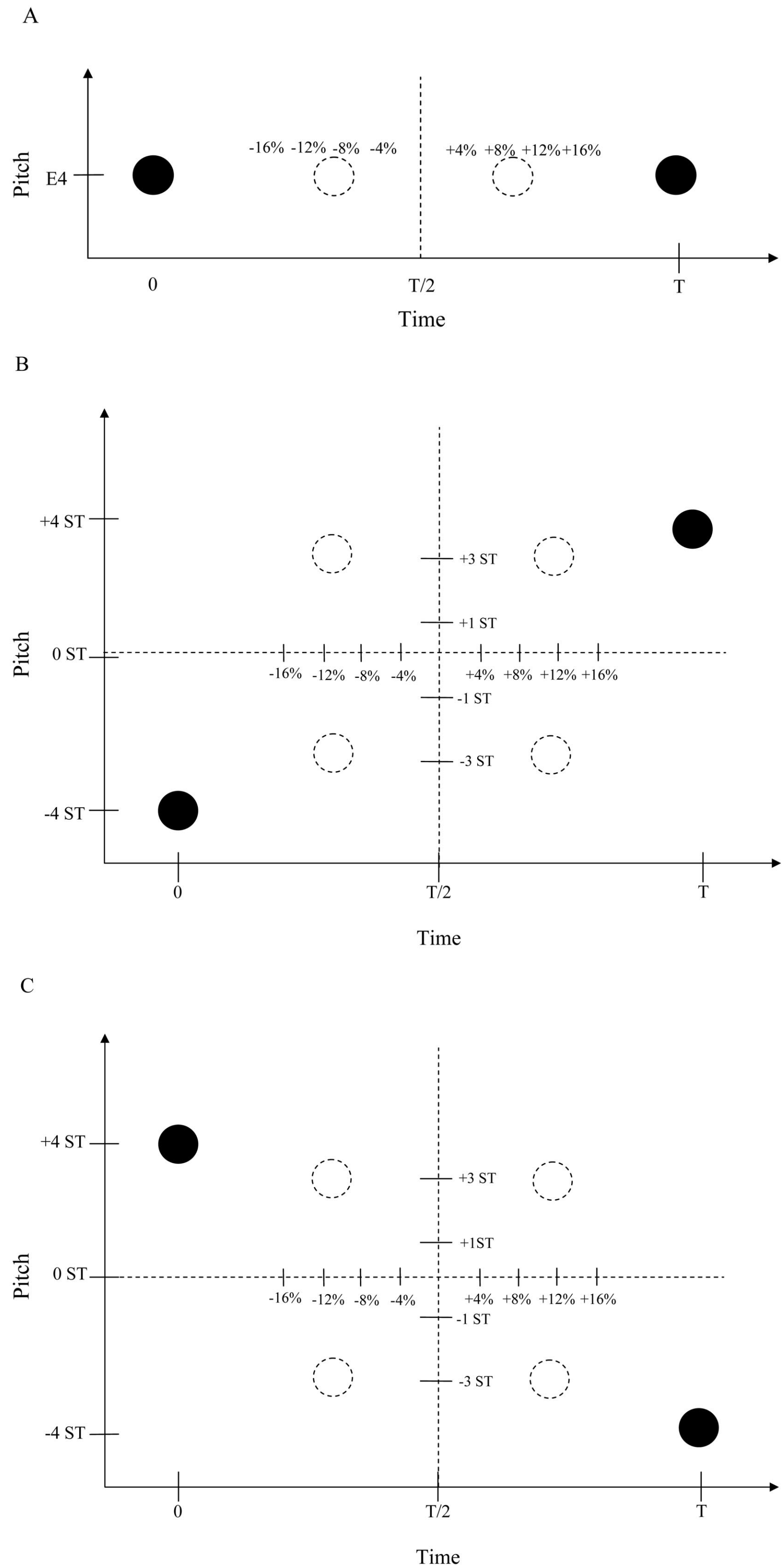

Figure 2. (A) Task diagram for Experiment 1 (baseline). (B) Task diagram for Experiment 2 (ascending sequences). (C) Task diagram for Experiment 3 (descending sequences). The position of the target tone (open circles) in time was varied from trial to trial and took on values that were ± 4%, 8%, 12%, or 16% relative to the temporal bisection point of the bounding tones (T/2). In Experiment 1, pitch was held constant at E4. In Experiments 2 and 3, the pitch separation of the bounding tones was 8 semitones (ST), yielding fast (8 ST/728 ms), medium (8 ST/1,000 ms), and slow (8 ST/1,600 ms) velocity conditions; the pitch of the target tone varied from trial to trial and took on values of –3, –1, 0, +1, or + 3 ST relative to the pitch bisection point of the bounding tones.

the target tone was closer in time to the first tone than the third tone and to respond *long–short* when they thought the target tone was closer in time to the third tone than the first tone. Thus, for each three-tone sequence, participants were asked to make either a *short–long* or a *long–short* response with respect to the temporal pattern created by the pair of time intervals delimited by the sequence.

Participants completed a short training block with feedback followed by four test blocks without feedback. Each test block consisted of 40 trials. Within a test block, the temporal position of the middle (target) tone (time level) varied from trial to trial; participants heard each of the eight time levels five times within a block. Across the four test blocks, participants made 20 responses for each time level. The entire experiment lasted approximately 30 min.

Data analysis. Proportions of *long–short* responses were determined for each participant for each of the eight time levels, averaged over the four test blocks. Just noticeable difference (JND) and point of subjective equality (PSE) for the resulting psychometric curves were then estimated for each participant by using the *z*-transform method prescribed by Macmillan and Creelman (1991; pp. 219–220). On the basis of this analysis, JND is a measure of half the distance between the 25th and 75th percentiles of the cumulative response function, whereas PSE is a measure of the median, corresponding to the temporal position of the target tone judged *long–short* 50% of the time. We defined a directional constant error (CE) score by subtracting PSE from T/2. According to this metric, negative CE indicates a rightward shift in PSE and a tendency to underestimate the objective timing of the target stimulus, whereas positive CE indicates a leftward shift in PSE and a tendency to overestimate the objective timing of the target stimulus. Both CE and JND are reported as percentage (relative) deviations.

Results and Discussion

As expected, there were no reliable distortions in timing across the three values of T; a one-way between-subjects analysis of variance (ANOVA) on CE scores revealed no effect of T, $F(2, 23) = 0.47$, $MSE = 5.59$, $p = .63$, $\eta^2 = .04$; mean relative CEs ($\pm$ *SE*) for T = 728 ms, 1,000 ms, and 1,600 ms were 0.52% ± 0.84, 1.40% ± 0.79, and 0.42% ± 0.79, respectively. In addition, separate *t* tests showed that CEs for the three conditions were not significantly different from zero: T = 728 ms: $t(7) = 0.48$, $p = .64$; T = 1,000 ms: $t(8) = 1.74$, $p = .12$; T = 1,600 ms: $t(8) = 0.86$, $p = .41$. Similarly, with respect to discrimination thresholds, a one-way between-subjects ANOVA on JND scores revealed no effect of T, $F(2, 23) = 1.14$, $MSE = 9.15$, $p = .34$, $\eta^2 = .09$; mean relative JNDs ($\pm$ *SE*) for T = 728 ms, 1,000 ms, and 1,600 ms were 8.91% ± 1.07, 7.91% ± 1.01, and 6.70% ± 1.01, respectively.

The results of Experiment 1 demonstrate that when pitch is held constant, there are no systematic distortions in perceived target timing, as evidenced by CE scores close to zero; moreover, discrimination thresholds for *short–long/long–short* judgments are similar across the three conditions. Overall, discrimination thresholds for Experiment 1 are within the range of those reported in previous research. For values of T between 400 ms and 1,600 ms, Hirsh, Monahan, Grant, and Singh (1990) found discrimination thresholds on the order of 5%–8% for a same–different task in which listeners judged deviations from isochrony in three-tone sequences. For a T value of 1,400 ms, MacKenzie (2007) reported thresholds that were slightly higher (on the order of 12%) for a *short–long/long–short* timing task involving three-tone sequences where the constant pitch of sequence elements varied from trial to trial. In Experiment 2, the same three fixed values of T were examined but were combined with variations in pitch in order to generate three equally spaced pitch velocity conditions.

Experiment 2: Ascending Sequences

In Experiment 2, all sequences ascended in pitch, and the pitch and timing of the middle (target) tone were manipulated relative to the expected pitch–time location based on implied pitch velocity of the sequence. As in Experiment 1, listeners judged the temporal position of the target tone relative to isochrony by responding *short–long* or *long–short*. One critical point was that listeners were instructed to ignore pitch when making judgments about target timing. We predicted the presence of an auditory kappa effect; that is, we expected that listeners' judgments about target timing would be distorted in the direction of the deviation of target tone pitch from that expected based on constant velocity. Moreover, based on the auditory motion hypothesis, increasing pitch velocity was expected to increase the magnitude of the auditory kappa effect, analogous to the effect of velocity on the magnitude of the visual kappa effect (Cohen et al., 1953; Huang & Jones, 1982).

Method

Design. Experiment 2 implemented a 3 × 5 × 8 mixed factorial design. Three pitch velocities (8 ST/728 ms, 8 ST/1,000 ms, 8 ST/1,600 ms) were crossed with five target-tone pitch levels and eight time levels (see Figure 2B). Values of target-tone pitch were –3, –1, 0, +1, or + 3 ST relative to the expected 4 ST pitch level of the target tone based on a constant velocity. The eight time levels for the target tone were the same as in Experiment 1, namely ± 4%, 8%, 12%, and 16% relative to T/2 in each three-tone sequence. Pitch velocity was a between-subjects factor, while pitch and time level of the target tone were within-subjects factors.

Participants. Forty-eight individuals ($n = 31$ women) between the ages of 18 and 32 from Bowling Green State University and the wider Bowling Green community participated in return for course credit or a cash payment of $5. All participants had self-reported normal hearing and a range of formal musical training ($M = 3.79$ years, $SD = 4.42$). Participants were randomly assigned to either the fast (8 ST/728 ms; $n = 16$), medium (8 ST/1,000 ms; $n = 15$), or slow (8 ST/1,600 ms; $n = 17$) velocity condition. Data from 2 participants in the fast velocity condition, 1 participant in the medium velocity condition, and 1 participant in the slow velocity condition were excluded due to inattention to the task or failure to follow instructions. Final participant numbers for the fast, medium, and slow velocity conditions were $n = 14$, $n = 14$, and $n = 16$, respectively.

Stimuli and equipment. Stimuli were three-tone sequences that ascended in pitch and had an 8 ST-pitch separation between the first and third (bounding) tones (see Figure 2B). The fundamental frequencies of the first and third tones were 329.63 Hz (corresponding to E4) and 523.25 Hz (corresponding to C5), respectively. As in Experiment 1, the time interval between the onsets of the first and third tones was fixed at one of three values (T = 728 ms, 1,000 ms, 1,600 ms). This produced three pitch velocity conditions: 8 ST/728 ms (fast), 8 ST/1,000 ms (medium), and 8 ST/1,600 ms (slow). The pitch of the middle (target) tone took on values of –3, –1, 0, +1, and + 3 ST relative to the expected 4 ST pitch bisection point. These relative pitch levels corresponded to values of F4 (349.23 Hz), G4 (392 Hz), A-flat4 (415.3 Hz), A4 (440 Hz), and B4 (493.88 Hz). As in Experiment 1, the onset of the target tone occurred ± 4%, 8%, 12%, or 16% relative to T/2.

As in Experiment 1, stimulus generation and response collection were controlled by an IBM PC-compatible computer running the MIDILAB software package, with a time resolution of ≈1 ms (Todd et al., 1989). Tone sequences were presented to participants at a comfortable listening level through Grado SR-80 headphones attached to a Yamaha PSR-70 MIDI keyboard set to a piano voice.

Procedure. The procedure for Experiment 2 was similar to that for Experiment 1. A short training session preceded testing; participants were asked to judge the timing of the target tone in three-tone sequences, responding *short–long* or *long–short.* Following training, participants were informed that each sequence would ascend in pitch. Again, participants were instructed to ignore changes in pitch and judge only the timing of the target tone. Participants completed two test sessions on separate days. Within each session, participants completed five test blocks with 40 trials per block. Within a test block, participants responded once to each of the 40 combinations of pitch and time level of the target tone. Overall, there were 200 trials per session, and the entire experiment lasted approximately 60 min. In total, 10 observations were obtained per participant for each of the 40 Pitch × Time Level combinations.

Results

In general, large distortions in timing were observed, consistent with a robust auditory kappa effect. A 3 (Velocity) × 5 (Pitch Level) ANOVA on CE scores revealed a main effect of Pitch Level, $F(4, 164) = 11.24$, $MSE = 1{,}131.09$, $p < .001$, $\eta^2 = .22$; judgments about the temporal position of the target tone were systematically affected by the target's pitch. When the relative pitch level of the target tone was negative, that is, when the pitch distance delimited by first tone and target was smaller than expected based on constant imputed pitch velocity, then CEs were generally negative, indicating that listeners were more likely to respond *short–long* than *long–short*. Conversely, when the relative pitch level of the target tone was positive, that is, when pitch distance was larger than expected based on constant imputed velocity, then CEs were generally positive, indicating that listeners were more likely to respond *long–short* than *short–long.* It is important to note that, in addition to the main effect of Pitch Level, there was a significant Pitch Level × Velocity interaction, $F(8, 164) = 4.49$, $MSE = 1{,}131.09$, $p < .001$, $\eta^2 = .18$, indicating that the magnitude of the auditory kappa effect was modulated by pitch velocity. Consistent with the auditory motion hypothesis, the auditory kappa effect was strongest in the fast velocity condition and weakest in the slow velocity condition (see Figure 3). There was also a significant main effect of Velocity, $F(2, 41) = 4.50$, $MSE = 462.88$, $p < .05$, $\eta^2 = .17$. In general, listeners exhibited a bias toward responding *long–short* to a greater extent in the fast velocity condition than in either the medium or slow conditions ($p < .05$, Tukey's HSD). Average CE scores for the fast (8 ST/728 ms), medium (8 ST/1,000 ms), and slow (8 ST/1,600 ms) pitch velocities were 9.01% ± 2.57, 1.11% ± 2.57, and –0.87% ± 2.41, respectively.

Next, we evaluated the proposed imputed pitch velocity model of the auditory kappa effect and used the estimates of w to quantify the extent to which listeners' judgments were based on imputed velocity in each of the three velocity conditions. The model was used to predict proportion of *long–short* responses at each of the eight time levels (± 4%, 8%, 12%, and 16%) for each of the five pitch levels (–3, –1, 0, +1, and + 3 ST) using a single value of w

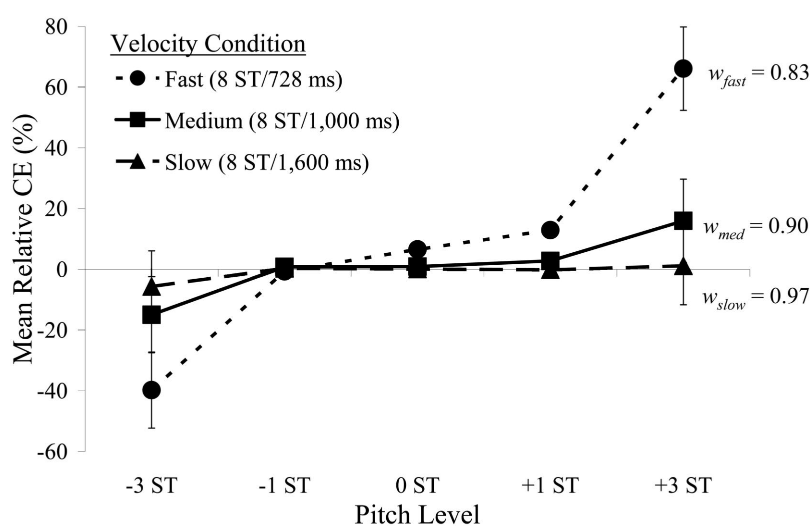

Figure 3. Experiment 2: Mean relative CE (with standard error bars) as a function of pitch level for ascending sequences with fast, medium, and slow velocities. Best fit values of w from the imputed velocity model are reported for each condition. ST = semitone; w = weight parameter; med = medium; CE = constant error.

for each participant. Values of w were adjusted until we obtained the best fitting value that minimized the root mean square error of approximation (RMSEA) between predicted and actual response proportions. Additional details of the method used to obtain the model fits are given in the Appendix. In general, the imputed velocity model provided good quantitative fits to participant data (mean RMSEA = .12). Obtained estimates of w were subjected to a one-way between-subjects ANOVA in order to examine the effect of velocity on the magnitude of the kappa effect. The ANOVA on w revealed a main effect of Velocity, $F(2, 41) = 11.11$, $MSE = 0.01$, $p < .001$, $\eta^2 = .35$. Consistent with the hypothesis that increasing velocity increases the magnitude of the auditory kappa effect, the largest w value obtained for the slow velocity condition (8 ST/1,600 ms), and the smallest w value obtained for the fast velocity condition (8 ST/728 ms). Mean w values for the fast, medium, and slow pitch velocity conditions were 0.83 ± 0.02, 0.90 ± 0.02, and 0.97 ± 0.02, respectively. Pairwise comparisons using Tukey's HSD indicated reliable differences between the fast and slow velocity conditions ($p < .001$) and between the medium and slow velocity conditions ($p = .05$) but only a marginally significant difference between the fast and medium velocity conditions ($p = .08$).

Finally, Table 1 summarizes relative JNDs for each of the five pitch levels in each velocity condition. A 3 (Velocity) × 5 (Pitch Level) mixed measures ANOVA on JNDs revealed a main effect of Pitch, $F(4, 164) = 3.22$, $MSE = 234.86$, $p < .05$, $\eta^2 = .07$, as well as a significant Pitch × Velocity interaction, $F(8, 164) = 2.89$, $MSE = 234.86$, $p < .01$, $\eta^2 = .12$. Overall, thresholds were larger in Experiment 2 than in Experiment 1, and thresholds tended to be largest for the most extreme pitch deviations. The Pitch × Velocity interaction illustrates that the effect of pitch deviations on discrimination thresholds was most pronounced in the fast velocity condition.

Table 1
Estimated Relative Just Noticeable Differences (± SEM) Reported as a Percentage for Five Pitch Levels of the Target Tone for Three Velocity Conditions

Experiment and target tone pitch level	Fast (8 ST/728 ms)	Medium (8 ST/1,000 ms)	Slow (8 ST/1,600 ms)
Experiment 2			
–3 ST	22.30 (6.18)	12.74 (6.18)	18.30 (5.78)
–1 ST	9.52 (2.68)	14.98 (2.68)	8.31 (2.51)
0 ST	11.63 (1.82)	12.81 (1.82)	9.39 (1.71)
+1 ST	13.67 (1.61)	11.00 (1.61)	10.06 (1.51)
+3 ST	37.01 (6.25)	12.39 (6.25)	9.66 (5.85)
Experiment 3			
–3 ST	16.73 (6.41)	10.30 (6.00)	14.75 (7.31)
–1 ST	13.35 (2.79)	11.17 (2.60)	10.66 (3.18)
0 ST	11.17 (1.89)	11.17 (1.76)	8.84 (2.16)
+1 ST	13.58 (1.67)	9.74 (1.56)	9.45 (1.91)
+3 ST	27.31 (6.49)	10.20 (6.04)	13.50 (7.39)

Note. Tone sequences ascended in pitch in Experiment 2, whereas they descended in pitch in Experiment 3. ST = semitone.

Discussion

In sum, Experiment 2 demonstrated a robust auditory kappa effect and a mediating role for pitch velocity. Consistent with the auditory motion hypothesis, the magnitude of the auditory kappa effect increased as a function of increasing pitch velocity. This finding is analogous to the reported effect of velocity on the magnitude of the visual kappa effect (Cohen et al., 1953; Huang & Jones, 1982), but this is the first study in the auditory domain to show a velocity effect based on the timing of the first and third sequence elements. With respect to discrimination thresholds, relative JNDs for Experiment 2 were larger than those reported in Experiment 1, and the largest discrimination thresholds were found for the most extreme pitch deviations (i.e., –3 ST, +3 ST), especially in the fast velocity condition. Increased thresholds with variations in pitch have been reported previously. For example, Hirsh et al. (1990) found larger discrimination thresholds when listeners judged the duration of an interval delimited by tones that differed in pitch relative to intervals delimited by tones of the same pitch.

One question that emerges from this work concerns whether the pitch direction of the tone sequence affects the magnitude of the kappa effect. In the visual domain, the magnitude of the visual kappa effect does not tend to vary as a function of direction of presentation in the horizontal plane (left–right vs. right–left), but when stimulus sequences are vertically aligned, the magnitude of the kappa effect tends to be larger when sequences unfold from top to bottom than when they unfold from bottom to top. This latter finding has been interpreted as evidence for observers' application of environmental invariants (i.e., gravity) to sequences exhibiting motionlike trajectories through space and time (Cohen et al., 1955).

If listeners similarly apply an auditory gravity principle to tone sequences, then as sequences ascend in pitch, they should decelerate in the mind of the listener and induce a *short–long* perception for adjacent time intervals that are objectively equal. Conversely, tone sequences that descend in pitch would be expected to accelerate in the mind of the listener and induce a *long–short* perception for adjacent sequence intervals that are objectively equal. Hubbard (1995) observed, in a study of auditory representational momentum that is consistent with an auditory gravity hypothesis, that the magnitude of forward displacement along an implied trajectory in memory for the final pitch of a descending tone sequence was greater than the magnitude of forward displacement in memory for the final pitch of an ascending tone sequence, specifically at the fastest tested pitch velocity. Experiment 3 examined the auditory gravity hypothesis in the context of the canonical auditory kappa task.

Experiment 3: Descending Sequences

Experiment 3 tested an auditory gravity hypothesis using tone sequences that descended rather than ascended in pitch; the remaining aspects of the design were the same as in Experiment 2. In the context of the canonical auditory kappa task, we expected that the application of an auditory gravity principle would (a) induce stronger kappa effects for descending tone sequences than for ascending tone sequences and (b) encourage a more pronounced *long–short* bias.

Method

Design. The design of Experiment 3 was identical to that of Experiment 2 with the exception that tone sequences descended rather than ascended in pitch.

Participants. Forty-five Bowling Green State University undergraduates (n = 24 women) between the ages of 18 and 25 years participated in return for course credit. All participants self-reported normal hearing and had a range of formal musical training (M = 3.82, SD = 3.59). Participants were randomly assigned to either the fast (8 ST/728 ms; n = 15), medium (8 ST/1,000 ms; n = 16), or slow (8 ST/1,600 ms; n = 14) velocity condition. Data for 2 participants in the fast, 1 participant in the medium, and 4 participants in the slow velocity condition were excluded due to inattention to the task or failure to follow instructions. Final participant numbers for the fast, medium, and slow velocity conditions were n = 13, n = 15, and n = 10, respectively.

Equipment. The equipment used was the same as in Experiment 1.

Stimuli. Stimuli were the same as in Experiment 2, except that the order of the tones was reversed so that the sequences descended rather than ascended in pitch (see Figure 2C).

Procedure. Participants completed 10 test blocks with 40 trials per block. Within a test block, participants heard each combination of pitch level and time level once. As in Experiment 2, each participant provided 10 observations for each of the 40 conditions. The entire experiment lasted approximately 60 min.

Results

As in Experiment 2, large distortions in perceived timing were observed, consistent with a robust auditory kappa effect. A 3 (Velocity) × 5 (Pitch Level) mixed-measures ANOVA on CE revealed a main effect of Pitch Level, $F(4, 140) = 8.47$, MSE = 775.40, $p < .001$, $\eta^2 = .20$. Judgments about the temporal position of the target tone were systematically affected by the position of the target tone in pitch space. Figure 4 shows that when the relative pitch level of the target tone was negative, listeners were more likely to respond *short–long* than *long–short*, and CEs were generally negative. Conversely, when the relative pitch level of the target tone was positive, listeners were more likely to respond *long–short* than *short–long*, and CEs were generally positive.

Importantly, as in Experiment 2, there was a significant Pitch × Velocity interaction, $F(8, 140) = 2.11$, MSE = 775.40, $p < .05$, $\eta^2 = .11$, indicating that the magnitude of the auditory kappa effect was modulated by pitch velocity. Specifically, the auditory kappa effect was strongest in the fast velocity condition. In contrast to the results of Experiment 2, however, the magnitudes of the kappa effects for the medium (8 ST/1,000 ms) and slow (8 ST/1,600 ms) pitch velocity conditions were similar. Application of the imputed pitch velocity model, which we report below, confirmed this. The main effect of Velocity did not reach significance, $F(2, 35) = 1.08$, MSE = 993.64, $p = .35$, $\eta^2 = .08$.

As in Experiment 2, the imputed pitch velocity model was used to predict the proportion of *long–short* responses at each of the eight time levels (± 4%, 8%, 12%, and 16%) for each of the five pitch levels (–3, –1, 0, +1, and +3 ST) using a single value of w for each participant. Values of w were adjusted until we obtained the best fitting value that minimized the RMSEA between the predicted and actual response proportions. Also as in Experiment 2, the imputed velocity model provided reasonable fits to participant data (mean RMSEA = .12) Recall that a smaller value of w is associated with a larger kappa effect. Mean w values for the fast, medium, and slow pitch velocity conditions were 0.85 ± 0.02, 0.91 ± 0.02, and 0.91 ± 0.03, respectively. A one-way between-subjects ANOVA on w revealed a marginally significant main effect of Velocity, $F(2, 35) = 2.67$, MSE = 0.01, $p = .08$, $\eta^2 = .08$.

To address the prediction derived from the auditory gravity hypothesis that the magnitude of the auditory kappa effect should be potentially larger for sequences that descend in pitch than for sequences that ascend in pitch, we compared the results of Exper-

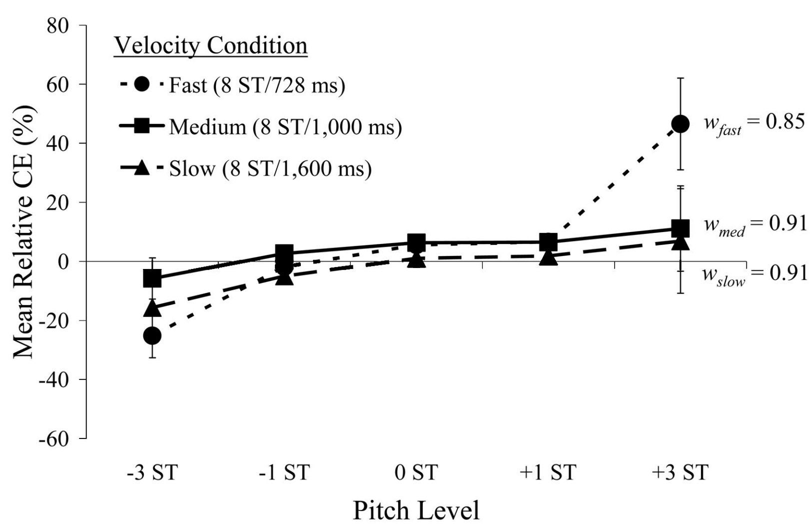

Figure 4. Experiment 3: Mean relative CE (with standard error bars) as a function of pitch level for the fast, medium, and slow velocity descending sequences. Best fit values of w from the imputed velocity model are reported for each condition. ST = semitone; w = weight parameter; med = medium; CE = constant error.

iments 2 and 3 by subjecting w estimates from both experiments to a 3 (Velocity) × 2 (Direction: ascending vs. descending) between-subjects ANOVA. The ANOVA revealed, inconsistent with an auditory gravity hypothesis, a main effect of Velocity, $F(2, 76) = 10.79$, $MSE = 0.01$, $p < .001$, $\eta^2 = .22$, but critically, no effect of Direction, $F(1, 76) = 0.20$, $MSE = 0.01$, $p = .66$, $\eta^2 = .003$, and no significant interaction between Velocity and Direction, $F(2, 81) = 1.87$, $MSE = 0.01$, $p = .16$, $\eta^2 = .05$. Despite the failure of the interaction to reach significance, there was a trend for w to be smaller for descending sequences relative to ascending sequences for the slow pitch velocity condition (ws = .91 and .97, respectively). An independent samples t test confirmed that this difference was significant, $t(24) = 2.66$, $p < .05$, Cohen's $d = 0.98$.

Finally, Table 1 summarizes relative JNDs for each of the five pitch levels in each velocity condition. A 3 (Velocity) × 5 (Pitch Level) mixed measures ANOVA revealed that discrimination thresholds did not differ significantly as a function of Pitch Level, $F(4, 35) = 1.96$, $MSE = 138.54$, $p = .10$, $\eta^2 = .05$, or Velocity, $F(2, 35) = 1.98$, $MSE = 337.90$, $p = .15$, $\eta^2 = .10$. The Pitch × Velocity interaction also did not reach significance, $F(8, 35) = 1.25$, $MSE = 138.54$, $p = .27$, $\eta^2 = .07$.

Discussion

In sum, the results of Experiment 3 demonstrate that for descending sequence, as for ascending sequences, the magnitude of the auditory kappa effect is modulated by pitch velocity. Consistent with the auditory motion hypothesis, a larger kappa effect was observed in the fast velocity condition than in the slow velocity condition. As in Experiment 2, the proposed imputed pitch velocity model provided generally good fits to participants' response proportions; the trend was for w to decrease as a function of pitch velocity. Inconsistent with the auditory gravity hypothesis, there was little evidence for differences between the ascending and descending sequence conditions. However, there was a tendency for the magnitude of the kappa effect to be larger for slow descending sequences than for slow ascending sequences. We will return to this point in the General Discussion.

General Discussion

The general aims of this research were twofold. First, we were interested in evaluating an imputed pitch velocity model of the auditory kappa effect. Second, we were interested in testing predictions of an auditory motion hypothesis concerning the effect of velocity on the magnitude of the auditory kappa effect. Three experiments were conducted to address these aims. Experiment 1 was a baseline experiment in which participants made *short–long/long–short* judgments about the timing of a middle (target) tone in monotone three-element sequences. Two main findings obtained. First, with pitch constant, constant error scores were close to zero for time intervals (T) of 728 ms, 1,000 ms, and 1,600 ms. Second, discrimination thresholds were similar for the three conditions.

Experiments 2 and 3 combined the values of T tested in Experiment 1 with a pitch manipulation in order to create fast, medium, and slow velocity conditions. Both experiments examined the canonical auditory kappa task and held the pitch distance between the first and third (bounding) tones constant at 8 semitones (ST). This yielded three pitch velocity conditions: fast (8 ST/728 ms), medium (8 ST/1,000 ms), and slow (8 ST/1,600 ms). With respect to the first aim, the proposed imputed pitch velocity model provided reasonable quantitative fits to listeners' time judgment responses in the three velocity conditions for both ascending tone sequences (Experiment 2) and descending tone sequences (Experiment 3). The importance of this finding is that it provides initial quantitative support for an imputed pitch velocity model of the auditory kappa effect; moreover, estimated values of the weight parameter (w) provide a method for quantifying the magnitude of the auditory kappa effect in a way that is directly comparable to previous visual kappa findings.

With respect to the second aim, robust auditory kappa effects were found for both the ascending and descending sequence conditions. Consistent with the auditory motion hypothesis, the largest kappa effects were demonstrated at the fastest pitch velocity (8 ST/728 ms), while the smallest kappa effects were found at the slowest pitch velocity (8 ST/1,600 ms). In general, observed values of w indicated that expected target timing based on imputed velocity, $E(t)$, was weighted more heavily in the computation of perceived time with increasing pitch velocity for both ascending and descending sequences.

According to the auditory motion hypothesis, motionlike relations in auditory sequences with nonzero pitch velocity encourage dynamic attending and extrapolation along an implied pitch–time trajectory, thus allowing for the generation of expectancies for future locations of an auditory event in pitch and time. From this perspective, the auditory kappa effect is a consequence of a perceptual simplification resulting from the deviation of a pitch–time event from constant pitch velocity imputed to a tone sequence exhibiting lawful pitch–time transformations. Moreover, as pitch velocities approach the upper limit of the range within which sequences induce a feeling of auditory motion, expectancies generated based on extrapolation of the trajectory are strengthened (M. R. Jones, 1976; M. R. Jones & Yee, 1993), and therefore a larger kappa effect is observable for pitch velocities at the upper (fast) end of this range. Thus, it seems likely that previous studies failed to observe a more robust auditory kappa effect because the bounding tones were separated by time intervals between 1,400 ms and 2,800 ms (Cohen et al., 1954; MacKenzie, 2007; Shigeno, 1986), which would have resulted in velocity conditions that were on the lower (slow) end of the range expected to induce a sense of auditory motion. Therefore, one additional contribution of the current research is the extension of tested velocities into the range within which the auditory motion hypothesis predicts the presence of robust auditory kappa effects.

Experiment 3 extended the results of Experiment 2 by considering the possibility that, in line with an auditory gravity hypothesis, descending tone sequences would show an augmented kappa effect compared with ascending tone sequences, as has been demonstrated for visual stimulus displays by Cohen et al. (1955). In general, the auditory gravity hypothesis was only weakly supported, as the results of Experiment 3 did not reveal a main effect of tone sequence direction (ascending vs. descending) on the magnitude of the kappa effect. However, there was a trend for the magnitude of the effect in the slow velocity condition to be larger for descending sequences than for ascending sequences. One possible reason for this is that the slow pitch velocity lies close to the proposed limit of the auditory motion boundary (M. R. Jones, 1976), and as such, an effect of auditory gravity may have been more readily observable than in the medium and fast velocity conditions. Specifically, if in the mind of the listener descending tone sequences were subject to acceleration, then slow descending tone sequences may have been heard as taking on more motionlike qualities relative to their ascending counterparts, thus inducing an augmented auditory kappa effect.

One question that emerges from the present set of studies is whether, within the range of velocities that induce a feeling of auditory motion, pitch and duration contribute equally to perceived velocity. In a magnitude estimation study varying the physical spacing of sounds, Strybel, Span, and Witty (1998) reported that temporal separation contributes more strongly to perceived velocity than does absolute spatial separation. Thus there is reason to suspect that in the context of the auditory kappa effect elicited through manipulations to pitch space, the time interval between the bounding tones may contribute more strongly to perceived pitch velocity and a sense of auditory motion than do equivalent changes in the pitch interval between the bounding tones. This suggests that changes in pitch velocity based on changes in pitch (Δp) should produce a weaker modulatory effect of velocity on the magnitude of the auditory kappa effect than observed in the present study for equivalent changes in pitch velocity based on changes in time (Δt).

Alternative Perspectives on the Auditory Kappa Effect

There are at least two alternative viewpoints on the auditory kappa effect that require additional consideration in the context of the present findings. First, Huang and Jones (1982; see also B. Jones & Huang, 1982) proposed that, in the visual domain, uncertainty due to difficulty of temporal judgments has the potential to force the perceiver to rely on the to-be-ignored spatial dimension for cues to the temporal position of the target. Thus, when the judgment about the target's temporal position is difficult, it may be biased in the direction of the spatial (pitch) manipulation. If temporal judgments become more difficult at shorter time intervals, then the cues from spatial position will be more heavily weighted in the determination of the appropriate response. The results of Experiment 1, however, indicated that the difficulty of auditory temporal judgments did not vary significantly as a function of temporal separation within the range of rates tested in the current study; however, there was a trend toward slightly larger discrimination thresholds in the $T = 728$ ms condition compared with the two other conditions. This hypothesis, however, cannot account for effects of pitch velocity on the kappa effect that result from varying the pitch separation of the bounding tones while holding T constant (see MacKenzie, 2007, for a demonstration along these lines).

A second viewpoint that requires additional consideration is a response competition hypothesis (Hazeltine, Poldrack, & Gabrieli, 2000). From this perspective, the systematic distortions in timing that characterize the kappa effect may arise from simultaneous activation of "competing" responses associated with incongruent pitch–time characteristics of a stimulus. In the context of the canonical auditory kappa task, a target tone closer in time to the first tone but closer in pitch to the third tone, for example, constitutes an incongruent pairing of stimulus characteristics (see Figure 1; $-\Delta t$ and $+\Delta s$, respectively). Thus, when an incongruent pitch–time pairing activates competing responses, judgments about the value of the to-be-judged stimulus dimension should be biased in the direction of the value of the to-be-ignored stimulus dimension, producing a kappa effect. Although beyond the scope of the present study, a response competition account of the kappa effect makes at least two testable predictions for future research. First, there should be lengthened response times for correct responses to incongruent pitch–time pairs relative to correct responses for congruent pairs. Second, the probability of making an incorrect response to an incongruent stimulus pairing should increase following trials that activated the response associated with the current value of the to-be-ignored stimulus dimension.

Broader Implications

Overall, the results of the current study contribute to a broader theoretical debate concerning the relative independence of pitch and rhythm (temporal) processing. On one side of the debate, pitch and rhythm are viewed as independent dimensions (e.g., Palmer & Krumhansl, 1987), whereas the alternative perspective is that pitch and rhythm processing are fundamentally interdependent (e.g., M. R. Jones, 1976; M. R. Jones & Boltz, 1989). The present set of findings related to auditory kappa support the latter view.

Converging evidence for the interdependence of pitch and rhythm (temporal) processing comes from a number of sources (Boltz, 1998; Dowling, Lung, & Herrbold, 1987; Ellis & Jones, in press; M. R. Jones, Moynihan Johnston, & Puente, 2006; M. R. Jones, Moynihan, MacKenzie, & Puente, 2002; Lebrun-Guillaud & Tillman, 2007). Most relevant in the context of the present auditory kappa findings are the auditory tau effect and auditory representational momentum. The auditory tau effect, the converse of the kappa effect, is the phenomenon whereby the perceived pitch of a target tone embedded in a three-tone sequence is systematically affected by the temporal location of the target. As in the canonical kappa task, the canonical task used to study the auditory tau effect involves the sequential presentation of a three-tone sequence that may be ascending or descending in pitch. Listeners make judgments about the pitch of the middle (target) tone, ignoring target timing. When the timing of the target tone deviates from that expected based on the pitch velocity implied by the fixed pitch–time positions of the bounding tones, perception of the to-be-judged target pitch is biased in the direction of the to-be-ignored change in target timing (Christensen & Huang, 1979; Cohen et al., 1954; Shigeno, 1986, 1993). Thus, a target tone closer in time to the first bounding stimulus than the third tends to be perceived as being closer in pitch to the first stimulus than the third, and vice versa.

Auditory representational momentum for tone sequences has also been interpreted as evidence that listeners anticipate future locations of auditory events through extrapolation along the implied pitch–time trajectory (Freyd, Kelly, & DeKay, 1990; Johnston & Jones, 2006). This view is supported by several findings that fit within the framework of the auditory motion hypothesis. First, the amount of forward displacement along the pitch–time trajectory is generally positively correlated with the duration of the retention interval preceding the response. That is, for longer retention intervals, the remembered endpoint is farther along the trajectory implying motion than for relatively short retention intervals (Freyd & Johnson, 1987). Second, Johnston and Jones (2006; see also Freyd & Finke, 1984) demonstrated that when irregularities are introduced into the trajectories of auditory patterns, representational momentum is abolished, indicating that forward displacement is dependent on expectations derived from regularities in pitch–time relationships; deviations of pitch–time events from the established trajectory inhibit the ability of the perceiver to capitalize on motionlike relations in order to make predictions. Additional research is needed to clarify the nature of the relationship between the phenomena of auditory kappa, tau, and representational momentum in the context of the auditory motion hypothesis.

On the other side of the debate, strong proponents of the perceptual independence view have proposed a strict functional and anatomical modularity in pitch and rhythm processing (Peretz, 2006; Peretz & Coltheart, 2003). Supporting this view are reports of pitch/rhythm dissociations in patients with brain damage (DiPietro, Leeman, & Schnider, 2004; Griffiths et al., 1997; Johnsrude, Penhune, & Zatorre, 2000; Tramo, Shah, & Braida, 2002; Wilson, Pressing, & Wales, 2002; Zatorre, 1988). Patients with acquired amusia are often left with impaired pitch perception but intact rhythm perception/production, as well as unimpaired general auditory faculties (Peretz & Coltheart, 2003). Amusia acquired through brain damage has been associated with lesions to the right auditory cortex (Griffiths et al., 1997; Johnsrude et al., 2000; Tramo et al., 2002; Zatorre, 1988). On the other hand, cases of impaired rhythm perception/production abilities but intact pitch perception are less clear cut and have been reported to occur with damage to either the left or right hemispheres (DiPietro et al., 2004; Mavlov, 1980; Wilson et al., 2002). In related work on congenital amusia, there is an increasing body of evidence that tone-deaf individuals show impaired pitch processing but unimpaired rhythm processing (Peretz, 2006; Peretz & Coltheart, 2003). This double dissociation is typically interpreted as supporting modularity of pitch and rhythm (temporal) processing.

One weakness of a strict independence perspective is the recent finding that when rhythmic sequences take on dynamic pitch–time trajectories, rhythm perception performance by individuals with congenital amusia deteriorates (Foxton, Nandy, & Griffiths, 2006). From this standpoint, it is unclear why the introduction of a pitch manipulation would negatively affect rhythm perception if pitch and time characteristics of musical stimuli are independently processed by functionally isolated modules. A more general criticism of the strict independence view, however, comes from neural network simulations that suggest the use of caution when interpreting double dissociations (Juola & Plunkett, 2000). In the context of an investigation of language processing, Juola and Plunkett (2000) demonstrated that double dissociations can occur with simulated lesions of a neural network model when none are present. This finding implies that the presence of double dissociations may not be the best basis to infer functional and/or anatomical modularity of pitch and rhythm processing.

Conclusions

In conclusion, there are two main contributions of this research. First, we provide support for the proposed imputed pitch velocity model of the auditory kappa effect. Second, we provide additional support for an auditory motion hypothesis (M. R. Jones, 1976; M. R. Jones & Yee, 1993; MacKenzie, 2007; MacKenzie & Jones, 2005). In the present study, the auditory motion hypothesis was tested by examining the possibility that implied pitch velocity ($\Delta p/\Delta t$), manipulated by holding pitch space constant and adjusting the time interval (T) between the first and third tones, modulates the strength of the auditory kappa effect. Participants judged the timing of a target tone embedded in ascending and descending three-tone sequences while ignoring changes in target tone pitch. Consistent with the auditory motion hypothesis, the magnitude of the auditory kappa effect was larger for tone sequences that implied a faster pitch velocity (8 ST/728 ms) than for sequences that implied a slower pitch velocity (8 ST/1,600 ms). Moreover, the magnitude of the auditory kappa effect obtained in the present study was much larger than has been observed previously. An auditory gravity hypothesis received only limited support. More generally, this research provides further support for the perspective that the processing of pitch and temporal characteristics of auditory patterns are fundamentally interdependent.

References

Abbe, M. (1936). The spatial effect upon the perception of time. *Journal of Experimental Psychology, 3,* 1–52.

Abe, S. (1935). Experimental study on the correlation between time and space. *Tohoku Psychologica Folia, 3,* 53–68.

Anderson, N. H. (1974). Algebraic models in perception. In E. Carterette & M. P. Friedman (Eds.), *Handbook of perception: Vol. 2. Psychophysical judgment and measurement* (pp. 258–259). New York: Academic Press.

Boltz, M. G. (1998). Tempo discrimination of musical patterns: Effects due to pitch and rhythmic structure. *Perception & Psychophysics, 60,* 1357–1373.

Christensen, I. P., & Huang, Y. L. (1979). The auditory tau effect and memory for pitch. *Perception & Psychophysics, 26,* 489–494.

Cohen, J., Hansel, C. E. M., & Sylvester, J. D. (1953, November 14). A new phenomenon in time judgment. *Nature, 172,* 901.

Cohen, J., Hansel, C. E. M., & Sylvester, J. D. (1954, October 2). Interdependence of temporal and auditory judgments. *Nature, 174,* 642–644.

Cohen, J., Hansel, C. E. M., & Sylvester, J. D. (1955). Interdependence in judgments of space, time and movement. *Acta Psychologica, 11,* 108–114.

Collyer, C. E. (1977). Discrimination of spatial and temporal intervals defined by three light flashes: Effects of spacing on temporal judgments and of timing on spatial judgments. *Perception & Psychophysics, 2,* 357–364.

Crowder, R. G., & Neath, I. (1994). The influence of pitch on time perception in short melodies. *Music Perception, 12,* 379–386.

DiPietro, M. L. M., Leemann, B., & Schnider, A. (2004). Receptive amusia: Temporal auditory processing deficit in a professional musician following a left temporo-parietal lesion. *Neuropsychologia, 42,* 868–877.

Douglas, K. M., & Bilkey, D. K. (2007). Amusia is associated with deficits in spatial processing. *Nature Neuroscience, 10,* 915–921.

Dowling, W. J., Lung, K. M.-T., & Herrbold, S. (1987). Aiming attention in pitch and time in the perception of interleaved melodies. *Perception & Psychophysics, 41,* 642–656.

Ellis, R. J., & Jones, M. R. (in press). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance.*

Foxton, J. M., Nandy, R. K., & Griffiths, T. D. (2006). Rhythm deficits in 'tone deafness.' *Brain and Cognition, 62,* 24–29.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 10,* 126–132.

Freyd, J. J., & Johnson, J. Q. (1987). Probing the time course of representational momentum. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13,* 126–132.

Freyd, J. J., Kelly, M. H., & DeKay, M. L. (1990). Representational momentum in memory for pitch. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 16,* 1107–1117.

Griffiths, T. D., Rees, A., Witton, C., Cross, P. M., Shakir, R. A., & Green, G. G. R. (1997). Spatial and temporal processing deficits following right hemisphere infarction: A psychophysical study. *Brain, 120,* 785–794.

Hass, E. J., & Hass, K. C. (1984). Alternative to the wavelength interpretation of the auditory kappa effect. *Perceptual and Motor Skills, 58,* 675–678.

Hazeltine, E., Poldrack, R., & Gabrieli, J. D. E. (2000). Neural activation during response competition. *Journal of Cognitive Neuroscience, 12*, 118–129.

Hirsh, I. J., Monahan, C. B., Grant, K. W., & Singh, P. G. (1990). Studies in auditory timing I: Simple patterns. *Perception & Psychophysics, 47,* 215–226.

Huang, Y. L., & Jones, B. (1982). On the interdependence of temporal and spatial judgments. *Perception & Psychophysics, 32,* 7–14.

Hubbard, T. L. (1995). Auditory representational momentum: Surface form, direction, and velocity effects. *American Journal of Psychology, 108*, 255–274.

Johnsrude, I., Penhune, V., & Zatorre, R. (2000). Functional specificity in the right auditory cortex for perceiving pitch direction. *Brain, 123,* 155–163.

Johnston, H. M., & Jones, M. R. (2006). Higher-order pattern structure influences auditory representational momentum. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 2–17.

Jones, B., & Huang, Y. L. (1982). Space-time dependencies in psychophysical judgment of extent and duration: Algebraic models of the tau and kappa effects. *Psychological Bulletin, 91,* 128–142.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83,* 323–335.

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

Jones, M. R., Maser, D. J., & Kidd, G. R. (1978). Rate and structure in memory for auditory patterns. *Memory & Cognition, 6*, 246–258.

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

Jones, M. R., Moynihan Johnston, H., & Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology, 53,* 59–96.

Jones, M. R., & Yee, W. (1993). Attending to auditory events: The role of temporal organization. In S. McAdams & E. Bigand (Eds.) *Thinking in sound: The cognitive psychology of human audition* (pp. 69–112). New York: Clarendon Press/Oxford University Press.

Juola, P., & Plunkett, K. (2000). Why double dissociations don't mean much. In G. Cohen, R. A. Johnston, & K. Plunkett (Eds.) *Exploring cognition: Damaged brains and neural networks: Readings in cognitive neuropsychology and connectionist modeling* (pp. 319–327). New York: Psychology Press.

Lebrun-Guillaud, G. & Tillman, B. (2007). Influence of a tone's tonal function on temporal change detection. *Perception & Psychophysics, 69*, 1450–1459.

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

MacKenzie, N., & Jones, M. R. (2005, November). *The auditory kappa effect revisited.* Paper presented at the 46th annual meeting of the Psychonomic Society, Toronto, Ontario, Canada.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

Matsuda, F., & Matsuda, M. (1979). Effects of spatial separation as a cue of time estimation in children and adults. *Japanese Psychological Research, 21,* 132–138.

Matsuda, F., & Matsuda, M. (1981). The anti-kappa effect in successively presented stimuli: A developmental study. *Japanese Psychological Research, 23,* 9–17.

Mavlov, L. (1980). Amusia due to rhythm agnosia in a musician with left hemisphere damage: A non-auditory supramodal defect. *Cortex, 16,* 331–338.

Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology, 13,* 116–126.

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition, 100,* 1–32.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6,* 688–691.

Price-Williams, D. R. (1954, February 20). The kappa effect. *Nature, 173,* 363–364.

Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition, 99,* 113–129.

Sarrazin, J.-C., Giraudo, M.-D., Pailhous, J., & Bootsma, R. J. (2004). Dynamics of balancing space and time in memory: Tau and kappa effects revisited. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 411–430.

Sarrazin, J. C., Giraudo, M. D., & Pittenger, J. B. (2007). Tau and Kappa effects in physical space: The case of audition. *Psychological Research, 71*, 201–218.

Shepard, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review, 91*, 417–447.

Shigeno, S. (1986). The auditory tau and kappa effects for speech and nonspeech stimuli. *Perception & Psychophysics, 40,* 9–19.

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

Strybel, T. Z., Span, S. A., & Witty, A. M. (1998). The effect of timing and spatial separation on the velocity of auditory apparent motion. *Perception & Psychophysics, 60,* 1441–1451.

Todd, R., Boltz, M. G., & Jones, M. R. (1989). The MIDILAB auditory research system. *Psychomusicology, 8,* 17–28.

Tramo, M., Shah, G. D., & Braida, L. D. (2002). Functional role of auditory cortex in frequency processing and pitch perception. *Journal of Neurophysiology, 87,* 122–139.

Wilson, S. J., Pressing, J. L., & Wales, R. J. (2002). Modeling rhythmic function in a musician post-stroke. *Neuropsychologia, 40*, 1494–1505.

Yoblick, D. A., & Salvendy, G. (1970). Influence of frequency on the estimation of time for auditory, visual, and tactile modalities: The kappa effect. *Journal of Experimental Psychology, 86,* 157–164.

Zatorre, R. J. (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America, 84,* 566–572.

(*Appendix follows*)

Appendix

Method Used to Obtain the Model Fits

The imputed pitch velocity model was fit to individual participant data from Experiment 2 (ascending sequences) and Experiment 3 (descending sequences). Predicted proportions of *long–short* responses, or $P(LS)$, were given by

$$P(LS) = f[wt + (1 - w)E(t)],$$

where t was the objective duration of the first time interval in the pair of time intervals delineated by the three-tone sequence, $E(t)$ was the expected duration of t based on the imputed velocity of the sequence, w was the weight parameter described in the text, and f was a sigmoid function, $f = 1/\{1 + e\wedge [-\gamma(t^* + \theta)]\}$, with gain parameter ($\gamma$) and bias parameter ($\theta$). The sigmoid function f was used to map perceived target timing (t^*), where $t^* = wt + (1 - w)E(t)$, into a *long–short* response probability. Using a single value of w, γ, and θ for each participant, we generated *long–short* response proportions for each of the eight time levels (± 4%, 8%, 12%, and 16%) at each of the five target pitch levels (−3, −1, 0, +1, and +3 ST). Model fits were obtained by minimizing the root mean square error of approximation (RMSEA) between the model's generated response proportions and the actual response proportions, allowing the parameters w, γ, and θ to vary. The values for $E(t)$ were derived from imputed velocity (V_i = 8 ST/T) and the pitch level (P) in semitones of the target tone as follows: $E(t) = P/V_i$. Overall, the model provided reasonable fits to all velocity conditions in both Experiment 2 and Experiment 3. Average RMSEA values for Experiments 2 and 3 were .121 and .122, respectively.

Received July 24, 2007
Revision received March 25, 2008
Accepted March 28, 2008 ■