Velocity perception for sounds moving in frequency space

Molly J. Henry · J. Devin McAuley

Abstract In three experiments, we considered the relative contribution of frequency change (Δf) and time change (Δt) to perceived velocity ($\Delta f/\Delta t$) for sounds that moved either continuously in frequency space (Experiment 1) or in discrete steps (Experiments 2 and 3). In all the experiments, participants estimated "how *quickly* stimuli changed in pitch" on a scale ranging from 0 (*not changing at all*) to 100 (*changing very quickly*). Objective frequency velocity was specified in terms of semitones per second (ST/s), with ascending and descending stimuli presented on each trial at one of seven velocities (2, 4, 6, 8, 10, 12, and 14 ST/s). Separate contributions of frequency change (Δf) and time change (Δt) to perceived velocity were assessed by holding total Δt constant and varying Δf or vice versa. For tone glides that moved continuously in frequency space, both Δf and Δt cues contributed approximately equally to perceived velocity. For tone sequences, in contrast, perceived velocity was based almost entirely on Δt, with surprisingly little contribution from Δf. Experiment 3 considered separate judgments about Δf and Δt in order to rule out the possibility that the results of Experiment 2 were due to the inability to judge frequency change in tone sequences.

Keywords Auditory motion · Velocity perception · Pitch perception · Time perception

M. J. Henry
Department of Psychology, Bowling Green State University,
Bowling Green, OH 43403, USA

M. J. Henry (✉) · J. D. McAuley
Department of Psychology, Michigan State University,
East Lansing, MI, USA
e-mail: henrymol@msu.edu

J. D. McAuley
e-mail: dmcauley@msu.edu

In the environment, the ability to estimate the velocity of a moving object has clear survival value for an animal, since it provides critical information about *where* an object will be *when*. Many animals, including humans, perceive velocity using both their eyes and ears. Auditory information about velocity is particularly advantageous for an animal when it is unable to see the moving object, such as in the dark, in dense foliage, or when the animal is not facing the object. Under such conditions, sound cues to object velocity provide critical information about time of arrival that can, for example, help an animal avoid a collision or escape an attacking predator.

In the auditory domain, another type of motion concerns time-varying changes in frequency. We will refer to this type of motion as *motion in frequency space* to distinguish it from the motion of objects in physical space. Motion in frequency space conveys important information to listeners in both speech and musical contexts (Dilley & McAuley, 2008; Frick, 1985; Jones, 1976; Jones & Yee, 1993; Krumhansl, 1991, 2000; Werner & Keller, 1994). In speech communication, patterns of rises and falls in fundamental frequency (an element of speech prosody) help determine word boundaries, distinguish between questions and statements, disambiguate semantic content, and place emphasis, as well as convey information about the emotional state of the speaker. In music, patterns of pitch change over time convey information about tonality and contribute to perceived accent structure.

More broadly, in the general context of an auditory scene, motion in frequency space also enables listeners to predict the future time course of a stimulus—that is, "what" will happen "when." This has been shown to be quite important in auditory scene analysis for guiding auditory attention to the right place in frequency at the right time (Crum & Hafter, 2008; Jones, Johnston, & Puente, 2006). Notably, in this regard, changes in frequency (pitch) can be continuous in nature (e.g., rises and falls in fundamental

frequency in speech) or may occur in discrete steps (e.g., the notes of a melody played on the piano). This article considers parallels between velocity perception for sounds moving in frequency space and for sounds/objects moving in physical space, with an emphasis on an empirical assessment of perceived velocity in frequency space for continuous and discrete stimuli.

The case for motion in frequency space

The motivation for considering parallels between motion in frequency space and motion in physical space comes from a number of sources. First, 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.[1] Notably, retinotopic and tonotopic organizations are preserved in visual and auditory cortical areas, respectively (Lund, 1988; Merzenich, Colwell, & Anderson, 1982). Second, in stimulus–response compatibility (SRC) pitch discrimination tasks, listeners are faster and more accurate when the mapping between relative pitch height and the corresponding response is congruent (high pitch associated with high button) than when the mapping is incongruent (Douglas & Bilkey, 2007; Melara & O'brien, 1987; Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2006). Douglas and Bilkey additionally demonstrated that tone-deaf listeners, who show deficits in fine-grained pitch discrimination and pitch direction discrimination, are also markedly impaired on a mental rotation task, relative to normal-hearing listeners.

Third, parallels can be made between frequency space and physical space in attentional-cuing paradigms. Directing visual attention to a location in physical space improves reaction time in detecting a visual target at the cued location (Posner, 1980; Posner, Snyder, & Davidson, 1980); similarly, directing auditory attention to a particular frequency

location improves detectability of an auditory signal at or near the cued frequency location (Hafter, Schlauch, & Tang, 1993; Howard, O'Toole, Parasuraman, & Bennett, 1984; Howard, O'Toole, & Rice, 1986). Finally, two spatiotemporal illusions—the tau and kappa effects—demonstrated for physical space in the visual (Bill & Teft, 1969; Cohen, Hansel, & Sylvester, 1953; Jones & Huang, 1982) and auditory (Grondin & Plourde, 2007; Sarrazin, Giraudo, & Pittenger, 2007) modalities have also been reported for motion in frequency space (Cohen, Hansel, & Sylvester, 1954; Crowder & Neath, 1994; Henry & McAuley, 2009; Henry, McAuley, & Zaleha, 2009). Tau effects refer to systematic errors in perceived stimulus spacing that are attributable to deviations from expected stimulus timing, whereas kappa effects refer to systematic errors in perceived stimulus timing that are attributable to deviations from expected stimulus spacing. Notably, tau and kappa effects in both physical space and frequency space can be similarly accounted for by an imputed velocity model (Henry & McAuley, 2009; Henry et al., 2009; Jones & Huang, 1982) that assumes that observers pick up on the motion implied by discrete sequences and use the associated trajectory information to make predictions about the locations and times of future events.

In sum, there is some evidence that frequency embodies properties of a spatial dimension, supporting the general hypothesis that listeners' perception of motion in frequency space shares similarities with perceived motion in physical space. The present study further explored this issue by comparing listeners' ability to estimate the velocities of tone glides that moved continuously in frequency space and of tone sequences that moved in discrete steps. The motivation for making a distinction between velocity perception for sounds moving continuously in frequency space and sounds moving in discrete steps comes from previous work on velocity perception for sounds moving in physical space, to which we turn next.

The case for a distinction between continuous and discrete motion

Velocity, v, is given by $\Delta s / \Delta t$, where Δs refers to the distance traveled and Δt refers to the duration of the movement (or time traveled). For motion in physical space in both vision and audition, a key issue in perception research has historically concerned the direct versus inferred nature of perceived velocity. The data on this point are very mixed, and the issue is generally unresolved (Ahissar, Ahissar, Bergman, & Vaadia, 1992; Altman, Syka, & Shmigidina, 1971; Boring, 1942; Goldstein, 1957; Grantham & Wightman, 1978; Maunsell & Van Essen, 1983; Orban, Kennedy, & Maes, 1981; Smeets &

---

[1] Strybel and Menges (1998) argued that auditory physical space, rather than frequency space, is the more appropriate auditory analogue to visual space and cautioned against describing phenomena such as auditory streaming and perception of tone glides as analogues to visual apparent motion (cf. Antis & Saida, 1985; Bregman, 1990). In general, they reject analogies between visual space and frequency that compare the retinotopic organization of the visual system and the tonotopic organization of the auditory system, on the grounds that neurons at higher levels of the auditory system can be tuned to spatial location and that the use of similar language in the two domains does not necessarily imply similar experience. We agree with elements of this line of argumentation but view the issue of parallels between visual space and frequency space as an issue that is best addressed empirically. The present set of experiments were performed to that end. We refer the reader to Strybel and Menges (1998) for an expanded treatment of their critique of parallels between visual space and frequency space.

Brenner, 1995; Stumpf, Toronchuk, & Cynader, 1992). One question that emerges from the view that velocity is inferred is whether Δs and Δt contribute equally to the perceived velocity of a moving source and, moreover, whether the relative contribution of Δs and Δt to perceived velocity depends on whether the source is moving continuously or in a series of discrete steps (Lappin, Bell, Harm, & Kottas, 1975). Algom and Cohen-Raz (1984, 1987) have drawn a similar distinction between the velocity of a continuously moving object and what they refer to as "cognitive velocity," which must be calculated by an observer using separate estimates of Δs and Δt; a special case of the latter would be the inference of velocity implied by the regular temporal and spatial separation of a sequence of discrete elements.

In the auditory domain, several studies have provided evidence that aspects of velocity perception for continuous and discrete stimuli differ. Perrott and colleagues (Perrott, Buck, Waugh, & Strybel, 1979; Waugh, Strybel, & Perrott, 1979) examined the function relating perceived velocity to objective velocity for sounds moving continuously in physical space. Perceived velocity was linearly related to source velocity, and this relationship closely mimicked the relation of perceived velocity to objective velocity in vision (Waugh et al., 1979). In contrast, for auditory apparent motion stimuli composed of a series of discrete sound events, Strybel, Span, and Witty (1998) found that Δs and Δt cues contributed differently to perceived velocity. In Strybel et al.'s (1998) study, the discrete sound events were noise bursts originating from loudspeakers at three locations; spacing (loudspeaker separation; Δs) and timing (noise burst interonset-intervals and noise burst duration; Δt) of individual sequence elements were varied. Velocity estimates for the apparent motion of the sound sequences were affected by the interonset timing of noise burst onsets and their duration (Δt). However, the spatial separation of the loudspeakers (Δs) had surprisingly little effect on velocity estimates. Moreover, listeners' velocity estimates were better predicted by the total duration of the sequences than by the ratio of total distance to total duration ($\Delta s/\Delta t$). Thus, in contrast with continuous stimuli, for discrete sequences of sounds, total duration is the primary contributor to perceived velocity (Strybel et al., 1998). The general aim of the present study was to extend this line of work by Strybel and colleagues to the study of velocity perception in frequency space.

Overview of the Present Study

Notably, no studies that we are aware of have directly examined the nature of the relationship between objective velocity ($\Delta f/\Delta t$) and perceived velocity for sounds moving

in frequency space. Moreover, very little is known about the relative contribution of frequency change (Δf) and time change (Δt) cues to perceived frequency velocity and whether this is the same for sounds moving continuously in frequency space versus sounds moving in discrete steps, where velocity is only implied. To address these questions, we conducted a series of three experiments in which participants estimated "how *quickly* the sounds they heard changed in pitch" on a scale ranging from 0 (*not changing at all*) to 100 (*changing very quickly*) for stimuli that ascended or descended in frequency for seven different values of $\Delta f/\Delta t$.

In Experiment 1, we considered tone glides that moved continuously in frequency space, whereas in Experiments 2 and 3, we considered tone sequences that moved in frequency space in discrete steps. Velocity information ($\Delta f/\Delta t$) was potentially conveyed to participants in two ways. In a frequency change condition, velocity was manipulated by varying the total frequency change (Δf) of the auditory stimulus while holding total time change (Δt) constant. Conversely, in the time change condition, velocity was manipulated by varying the total time change (Δt) and holding frequency change (Δf) constant; Δf was varied in semitone (ST) units, where a ST corresponds to a musical half-step. The rationale for varying Δf in ST units is based on the view that perceived pitch is best described by an ST scale (Burns, 1999). Use of STs is also consistent with previous research that has considered auditory motion in frequency space (e.g., Henry & McAuley, 2009; Henry et al., 2009; Johnston & Jones, 2006; Jones & Yee, 1993). These manipulations yielded seven velocity levels in both frequency change and time change conditions: 2, 4, 6, 8, 10, 12, and 14 ST/s.

Studies of visual velocity perception have shown that both spatial (Δs) and temporal (Δt) change information contribute to perceived velocity for continuously moving objects (Algom & Cohen-Raz, 1984, 1987), and Waugh and colleagues demonstrated a remarkable agreement between the visual and auditory velocity functions in physical space (Waugh et al., 1979). Therefore, we reasoned that if judgments about the perceived velocity of sounds moving in frequency space parallel those for continuous events moving in physical space, both frequency change (Δf) and time change (Δt) should contribute approximately equally to the perceived velocity of continuous tone glides. Conversely, for tone sequences that move in frequency space in discrete steps, we expected that velocity judgments should be based primarily on time change (Δt) cues, with little to no contribution from frequency change (Δf) cues, as has been previously found for auditory apparent motion stimuli (Strybel et al., 1998). In Experiment 3, we ruled out the possibility that a greater reliance on time change cues in tone sequences

(Experiment 2) is due to a poor ability to judge frequency change, by having participants make estimates of total duration and total frequency change, in addition to velocity. This also permitted us to assess a number of models of velocity perception, two of which were similar to the models previously examined by Strybel et al. (1998).

Experiment 1

Method

Participants and design

Eleven Bowling Green State University undergraduates ($n = 9$, female) participated in return for course credit. Participants varied in their formal musical training (0–12 years; $M = 4.6$ years, $SD = 4.0$ years), and all self-reported normal hearing. The design of the experiment was a 2 (type of change: frequency change, time change) × 2 (direction: ascending, descending) × 3 (starting frequency: 262, 329, or 415 Hz) × 7 (Velocity: 2, 4, 6, 8, 10, 12, or 14 ST/s) within-subjects factorial. For each stimulus, participants estimated the perceived frequency velocity of the stimulus on a scale ranging from 0 (*not changing at all*) to 100 (*changing very quickly*).

Stimuli and equipment

Stimuli were sine-tone glides, ramped over the first five and last five pitch periods[2] to eliminate acoustic artifacts. Tone glides ascended or descended in frequency from one of three starting frequencies (262, 329, or 415 Hz) at one of seven velocities (2, 4, 6, 8, 10, 12, or 14 ST/s; see Fig. 1A). In the frequency change condition, velocity was varied by holding total glide duration (Δt) constant and varying total frequency change (Δf). For this condition, Δt was fixed at 1,500 ms, and Δf between the starting and ending frequencies of the glide took on one of seven values ($\Delta f = 3, 6, 9, 12, 15, 18,$ or 21 ST) to create seven velocity conditions (2, 4, 6, 8, 10, 12, or 14 ST/s). In the time change condition, velocity was varied by holding Δf constant and varying Δt. For this condition, Δf between the starting and ending frequencies of the glide was fixed at 12 ST, and Δt took on one of seven values ($\Delta t = 857$,

---

[2] Here, *period* refers to the duration of each cycle (peak to peak) of the pure tone glide, which has a reciprocal relationship to the frequency in hertz. The amplitude of tone glides was ramped linearly from a value of 0 to maximum amplitude over the first and last five cycles of the sound. Thus, the absolute duration of the onset and offset ramps depended on the starting and ending frequencies, respectively, of the glide.
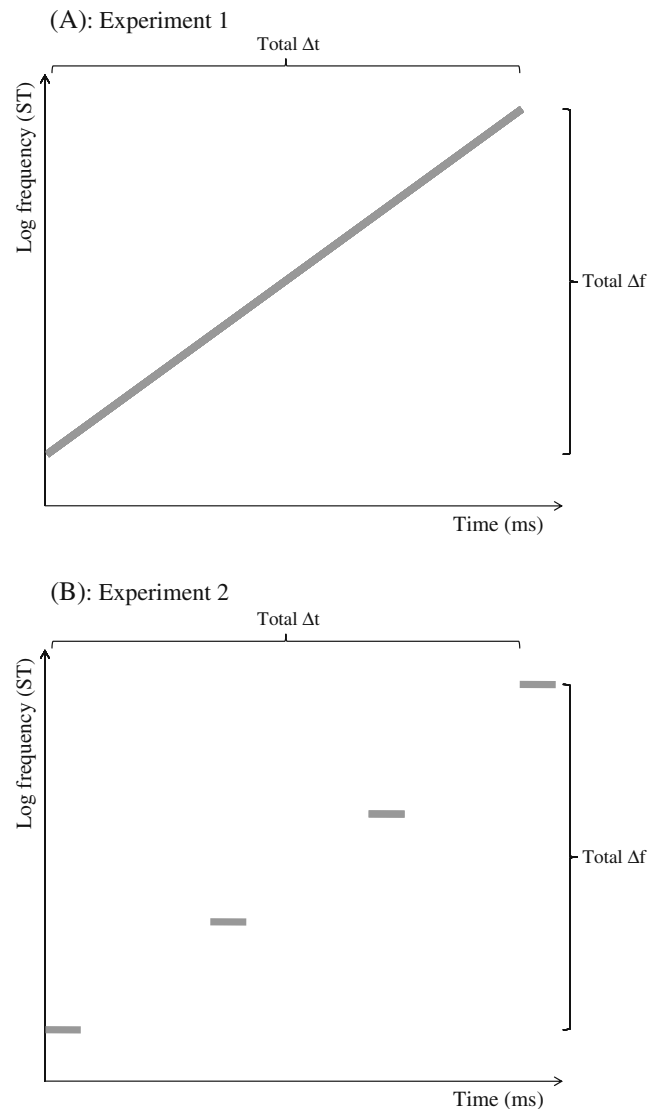

Fig. 1 A In Experiment 1, stimuli were tone glides that ascended or descended in frequency at one of seven velocities, where velocity was specified in units of semitones per second (ST/s). Total Δt and total Δf were defined as the duration and frequency distance, respectively, from the onset to the offset of the glide. Participants rated "how quickly stimuli changed in pitch" on a scale ranging from 0 (*not changing at all*) to 100 (*changing very quickly*). B In Experiment 2, stimuli were isochronous four-tone sequences that ascended or descended in frequency at one of seven velocities matched to Experiment 1. Total Δt was defined as the duration from the onset of the first tone to the onset of the final tone; total Δf was defined as the frequency distance between the first and final sequence tones

1,000, 1,200, 1,500, 2,000, 3,000, or 6,000 ms) to create the same seven matched velocity conditions.

All sounds were generated using Praat software (Boersma & Weenik, 2005). Stimulus presentation and response collection were controlled by E-Prime, Version 1.2 (Psychology Software Tools, Pittsburgh, PA) running on a Dell PC computer. Sounds were presented at a comfortable listening level (~70 dB) over Sennheiser HD 280 Pro headphones (Old

Lyme, CT). Estimates were made by clicking with the computer mouse along a horizontally aligned scale. Estimates of 0 (*not changing at all*) were made by clicking the left endpoint of the scale, whereas estimates of 100 (*changing very quickly*) were made by clicking the right endpoint of the scale.

Procedure

Participants completed a short familiarization block, followed by two experimental blocks. On all the trials, participants were asked to estimate "how quickly each stimulus changed in pitch" on a scale ranging from 0 (*not changing at all*) to 100 (*changing very quickly*). For the familiarization block, participants heard single tones with constant frequency (262, 329 , or 415 Hz) presented for one of three values of Δt (857, 1,500, or 6,000 ms). Participants were instructed to give these stimuli ratings of 0 (*not changing at all*). They were then presented with an example of a continuous tone glide that ascended in frequency at a rate of 8 ST/s (the median velocity) and were asked to give the sample glide a rating greater than 0. Following the familiarization phase, participants completed two experimental blocks, with a short break halfway through. Within each block, participants responded once to each type of change (frequency vs. time) × direction (ascending vs. descending) × starting frequency (262 , 329 , or 415 Hz) × velocity (2, 4, 6, 8, 10, 12, or 14 ST/s) combination and once to each of the 21 constant-frequency tones, which served as catch trials; constant-frequency tones were presented at all three starting frequencies and at all seven values of Δt. Overall, the two experimental blocks consisted of 210 trials. The experiment lasted approximately 45 min.

Results and discussion

An inspection of velocity estimates for zero-velocity (catch) trials showed that one participant gave velocity estimates that were much greater than zero ($M = 32.45$); moreover, this participant appeared to be responding randomly for all trial types, and so the data were excluded from subsequent analyses. The mean velocity estimate for zero-velocity (catch) trials without this participant was 1.41 (±1.41), which did not differ significantly from zero, $t(9) = 1.00$, $p = .35$.

Figure 2A shows mean estimates of frequency velocity as a function of $\Delta f/\Delta t$ (in ST/sec units) for the time change and frequency change conditions; regression lines are superimposed.[3] Slopes and intercepts of regression lines

for each participant were determined for each combination of type of change, direction, and starting frequency. For this analysis, differences in the slope across conditions reveal potential differences in the strength of the relationship between glide velocity and perceived velocity, whereas differences in the intercept across conditions reveal potential shifts in overall magnitude of velocity estimates. We expected that both frequency change (Δf) and time change (Δt) should contribute approximately equally to perceived velocity; in the context of the present analysis, this amounted to a prediction of no difference between slopes or intercepts for the frequency change and time change conditions. Moreover, slopes for both the frequency change and time change conditions should differ significantly from 0. In line with this hypotheses, the observed slopes for the time change and frequency change conditions were very similar (time change, $M = 3.02 \pm 0.47$, $R^2 = .86$; frequency change, $M = 3.53 \pm 0.72$, $R^2 = .97$). Separate 2 (type of change) × 2 (direction) × 3 (starting frequency) repeated measures ANOVAs on slopes and intercepts revealed no main effects or interactions (all $ps \geq .1$). Moreover, single-sample t tests confirmed that for both the frequency change and time change conditions, slopes were significantly different from 0, $t(9) = 6.82$, $p < .001$, and $t(9) = 4.48$, $p < .001$, respectively. Parallel tests for intercepts revealed that intercepts for the frequency change and time change conditions also differed significantly from 0, $t(9) = 5.41$, $p < .001$, and $t(9) = 2.64$, $p < .05$, respectively.

Finally, since participants in the experiment varied in their number of years of formal musical training, we examined the possibility that musical training might be related to the extent to which an individual used frequency change and time change cues to judge velocity. A correlation analysis revealed a marginally significant correlation between years of musical training and slope for the time change condition, Spearman's $\rho = .60$, $p = .09$, but no other correlations were significant, all $ps \geq .23$.

In sum, the results reported here for continuous sounds (tone glides) moving in frequency space are consistent with those in previous studies showing that for continuously moving visual stimuli, spatial separation (Δs) and temporal separation (Δt) contribute jointly to perceived velocity (Algom & Cohen-Raz, 1984, 1987); thus, the present findings support the view that motion in frequency space shares perceptual characteristics with motion in physical space. In a second experiment, we used the same design as in Experiment 1, except that we presented participants with a discrete sequence of tones on each trial, rather than with a single continuous tone glide. The velocity conditions were the same. Our expectation was that contributions of frequency change and time change cues to velocity estimates for tone sequences would be consistent with the findings of Strybel et al. (1998) for auditory apparent motion in physical

---

[3] For continuous tone glides, participants' velocity estimates increased linearly with glide velocity ($R^2 = .93$). This was true for the time change condition ($R^2 = .97$) and the frequency change condition ($R^2 = .86$). However, the logarithmic relationship between perceived velocity and objective velocity provided a better fit to the data for the frequency change condition ($R^2 = .97$) than did the linear relationship ($R^2 = .86$).

Fig. 2 Velocity estimates are shown as a function of $\Delta f/\Delta t$, with regression lines superimposed. A Experiment 1 results: Velocity estimates increased as a function of $\Delta f/\Delta t$ for the time change condition and the frequency change condition. B Experiment 2 results: Velocity estimates increased as a function of $\Delta f/\Delta t$ for the time change condition, but not for the frequency change condition

(A): Experiment 1
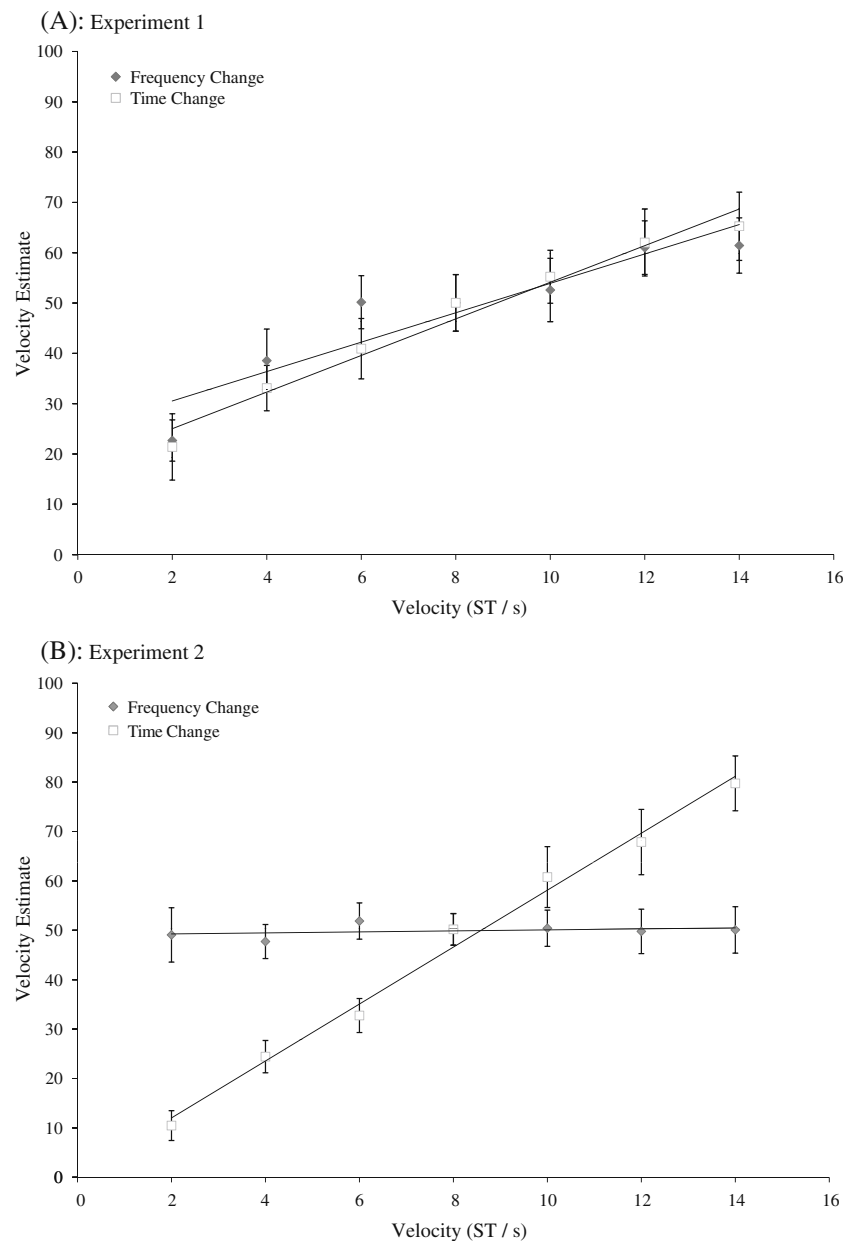
(B): Experiment 2

space. If so, then, unlike Experiment 1, time change (Δt) should contribute more strongly to perceived frequency velocity than should frequency change (Δf).

Experiment 2

Method

Participants and design

Nine Bowling Green State University undergraduates ($n = 8$, female) participated in return for course credit; none had participated in Experiment 1. Participants varied in their formal musical training (0–12 years; $M = 5.8$ years, $SD = 4.2$ years), and all self-reported normal hearing. The design and task for Experiment 2 were identical to those in Experiment 1. The only difference was that participants on each trial heard a tone sequence, rather than a tone glide.

Stimuli and equipment

The stimuli were four-tone isochronous sequences that ascended or descended in frequency from one of three starting frequencies (262, 329, or 415 Hz) at one of seven implied velocities (2, 4, 6, 8, 10, 12, or 14 ST/s; see Fig. 1B). Discrete sequence elements were 100-ms sine

tones that were ramped over the first and last 5 ms to eliminate acoustic artifacts. For the frequency change condition, the stimulus onset asynchrony (SOA) between the first tone and the last tone (Δt) was fixed at 1,500 ms, and the frequency separation between the first and last tones took on one of seven values (Δf = 3, 6, 9, 12, 15, 18, or 21 ST) to create the seven implied velocity conditions ($\Delta f/\Delta t$ = 2, 4, 6, 8, 10, 12, or 14 ST/s); with Δt fixed at 1,500 ms, the SOA between successive tones in each sequence was 500 ms, with equal frequency spacing of successive tones determined by Δf, which varied with velocity. For the time change condition, the frequency spacing of the first and last tones (Δf) was fixed at 12 ST, and the SOA between the first and last tone onsets took on one of seven values (Δt = 857, 1,000, 1,200, 1,500, 2,000, 3,000, or 6,000 ms) to create the same seven velocity conditions; with Δf fixed at 12 ST, the frequency spacing of successive tones was 4 ST, with the SOAs between successive tones equal and determined by Δt. Stimulus generation, presentation, and response collection methods were identical to those in Experiment 1.

Procedure

To mirror the procedure of Experiment 1, participants first heard monotone sequences presented at three frequencies (262, 329, or 415 Hz) and at three values of Δt (857, 1,500, or 6,000 ms); participants were instructed to assign these sequences ratings of 0 (*not changing at all in pitch*). Next, participants heard an example of a sequence with an implied velocity of 8 ST/s (the median velocity) and were asked to give the sequence a rating greater than 0. Following the familiarization block, participants completed two experimental blocks, during which the analogous 21 monotone sequences (rather than constant-frequency tones) served as catch trials. The general composition of experimental blocks was the same as that in Experiment 1.

Results and discussion

Velocity estimates for zero-velocity (catch) trials were slightly, but reliably, larger than zero($M = 2.7 \pm 1.1$), $t(8) = 2.48$, $p < .05$. Nonetheless, velocity estimates for catch trials showed that all the participants were following instructions and assigning velocity estimates that were still close to 0 for zero-velocity sequences.

Figure 2B shows mean estimates of frequency velocity as a function of $\Delta f/\Delta t$ (in ST/s units) for the time change and frequency change conditions. Regression lines are superimposed. As in Experiment 1, slopes and intercepts of the regression lines were determined for each participant for each combination of type of change, direction, and

starting frequency. Here, we expected that velocity judgments should be based primarily on time change (Δt) cues, with little to no contribution from frequency change (Δf) cues; thus, we anticipated that (1) slopes and intercepts for the frequency change and time change conditions would differ significantly, (2) slope for the time change condition would be significantly greater than 0, and (3) slope for the frequency change condition would not be reliably different from 0. Separate 2 (type of change) × 2 (direction) × 3 (starting frequency) repeated measures ANOVAs were conducted on slope and intercept estimates. With respect to the slope measure, only the main effect of type of change reached significance, $F(1, 8) = 61.81$, $MSE = 14.01$, $p < .001$, $\eta_p^2 = .89$. Consistent with expectations, slopes were generally much larger for the time change condition ($M = 5.76 \pm 0.60$), $R^2 = .99$, than for the frequency change condition ($M = 0.10 \pm 0.23$), $R^2 = .12$, $t(8) = 7.86$, $p < .001$. Moreover, although slopes for the time change condition were significantly larger than 0, $t(8) = 9.68$, $p < .001$, slopes for the frequency change condition were not significantly different from zero, $t(8) = 0.43$, $p = .67$. No other main effects or interactions reached significance, all $ps \geq .31$.

Similar to the ANOVA on estimated slopes, the ANOVA on intercepts revealed only a main effect of type of change, $F(1, 8) = 64.51$, $MSE = 987.74$, $p < .001$, $\eta_p^2 = .89$. The intercept for the time change condition ($M = 0.49 \pm 3.49$) was much smaller than the intercept for the frequency change condition ($M = 49.07 \pm 4.42$), $t(8) = 8.03$, $p < .001$. Moreover, the intercept for the time change condition was not significantly different from 0, $t(8) = 0.14$, $p = .89$, indicating that the best-fit line for the time change condition passed through the origin (i.e., $\Delta t = 0$ corresponded to an estimate of 0). The intercept for the frequency change condition was significantly different from 0, $t(8) = 11.10$, $p < .001$. No other main effects or interactions reached significance, all $ps \geq .20$.

Finally, as in Experiment 1, we examined correlations between years of musical training and slopes and intercepts for the frequency change and time change conditions and found that no correlations reached statistical significance, $ps \geq .30$.

In sum, for tone sequences that moved in frequency space in discrete steps, time change information contributed to perceived velocity, but in contrast to Experiment 1, frequency change information did not. Most striking was the finding that all values of Δf produced approximately the same estimate of velocity. These findings can be contrasted with the results of Experiment 1, where, for continuous tone glides, time and frequency information contributed approximately equally to perceived velocity. The results of Experiment 2 were consistent with the findings of Strybel et al. (1998), in that duration (Δt) contributed to the

perceived velocity of discrete sequences, whereas spatial (frequency, Δf) information did not. One possibility for the discrepancy between the results of Experiments 1 and 2 is that participants had much more difficulty perceiving frequency change in tone sequences (where velocity is only implied) than in tone glides, especially with long SOAs between sequence elements. To address this possibility, we conducted a third experiment in which participants estimated velocity, total time change (stimulus duration), and total frequency change in separate blocks for a subset of sequences from Experiment 2. Moreover, we used this additional information to evaluate four models of velocity perception, two of which were similar to the models examined by Strybel et al. (1998) for sounds moving in physical space. Here, the four models regressed velocity estimates on (1) estimates of total time change ($\Delta t'$), (2) estimates of total frequency change ($\Delta f'$), (3) the ratio of estimates of total frequency change to total time change ($\Delta f'/\Delta t'$), and (4) a linear combination of time change and frequency change estimates.

Experiment 3

Method

Participants and design

Fifteen Michigan State University undergraduates ($n = 11$, female) participated in return for course credit; none had participated in the previous experiments. Participants varied in their formal musical training (0–13 years; $M = 2.7$ years, $SD = 3.5$ years). Participants self-reported normal hearing, except for 1 participant, who reported partial hearing loss in one ear; however, the stimuli were presented binaurally, and an inspection of the data revealed no qualitative differences between responses by this participant and the rest of the sample, so their data were included in the analyses. The design of Experiment 3 was a 3 (task: velocity, total time change, total frequency change) × 2 (type of change: frequency change, time change) × 2 (starting frequency: 262 or 415 Hz) × 7 (velocity: 2, 4, 6, 8, 10, 12, or 14 ST/s) within-subjects factorial.

Stimuli and equipment

The stimuli were the same as those in Experiment 2, with the following exceptions. Since no differences in velocity estimates were found in Experiment 2 as a function of direction or starting frequency, we used only ascending sequences and eliminated the 329-Hz starting frequency for the present experiment in order to reduce the number of trials. Stimulus presentation and response collection were

controlled by E-Prime Version 2.0.8.22 (Psychology Software Tools;) but were otherwise identical to those in Experiments 1 and 2.

Procedure

Participants estimated the velocity ($\Delta f/\Delta t$), total frequency change (Δf), and total time change (Δt) of each sequence in three separate blocks. Note that estimates of the total time change correspond to estimates of the total sequence duration. All estimates were made using a 0—100 scale. For velocity estimates, the lower and upper anchors of the scale were *not changing at all* and *changing very quickly* in pitch (identical to Experiments 1 and 2); for time change (duration) estimates, the lower and upper endpoints of the scale were *very short* and *very long*. Finally, for frequency change estimates, the lower and upper anchors were *no pitch change* and *very large pitch change*. Unlike in Experiment 2, participants in Experiment 3 were not given a familiarization block where they heard only the monotone sequences (for which they were instructed to assign the sequences ratings of 0). Task order was counterbalanced across participants according to a Latin square. For each task, participants made judgments about each sequence twice, completing a total of 84 trials in a single block for each task. In total, the three experimental blocks consisted of 252 trials. The entire experiment took approximately 1 hr to complete.

Results and discussion

The results are summarized in Figs. 3, 4, and 5 for the velocity, time change, and frequency change tasks, respectively. Data for 1 participant were excluded from the final analysis, due to what appeared to be random responding for all three tasks.

Responses to monotone sequences for all three tasks

As in Experiment 2, we first considered responses to the monotone sequences, which served as zero-velocity and no-frequency-change (catch) trials for the velocity and frequency change tasks, respectively. For the velocity task, velocity estimates were higher in Experiment 3 than in Experiment 2 ($M = 28.4 \pm 7.4$) and notably significantly different from 0, $t(13) = 3.87$, $p < .01$. Moreover, the average velocity estimate for monotone sequences was similar to (and not significantly different from) the mean time change (duration) estimate for monotone sequences ($M = 32.6 \pm 1.6$), $t(14) = -0.53$, $p = .60$, supporting the general conclusion from Experiment 2 that for discrete sequences, participants use time change, but not frequency change, cues to make velocity estimates. For the frequency

Fig. 3 Experiment 3 results, velocity task: Velocity estimates increased as a function of $\Delta f / \Delta t$ for the time change condition and, to a lesser extent, for the frequency change condition. The contribution of Δt was somewhat smaller for the 263-Hz starting frequency (A) than for the 415-Hz starting frequency (B)

(A): 263-Hz Starting Frequency

(B): 415-Hz Starting Frequency
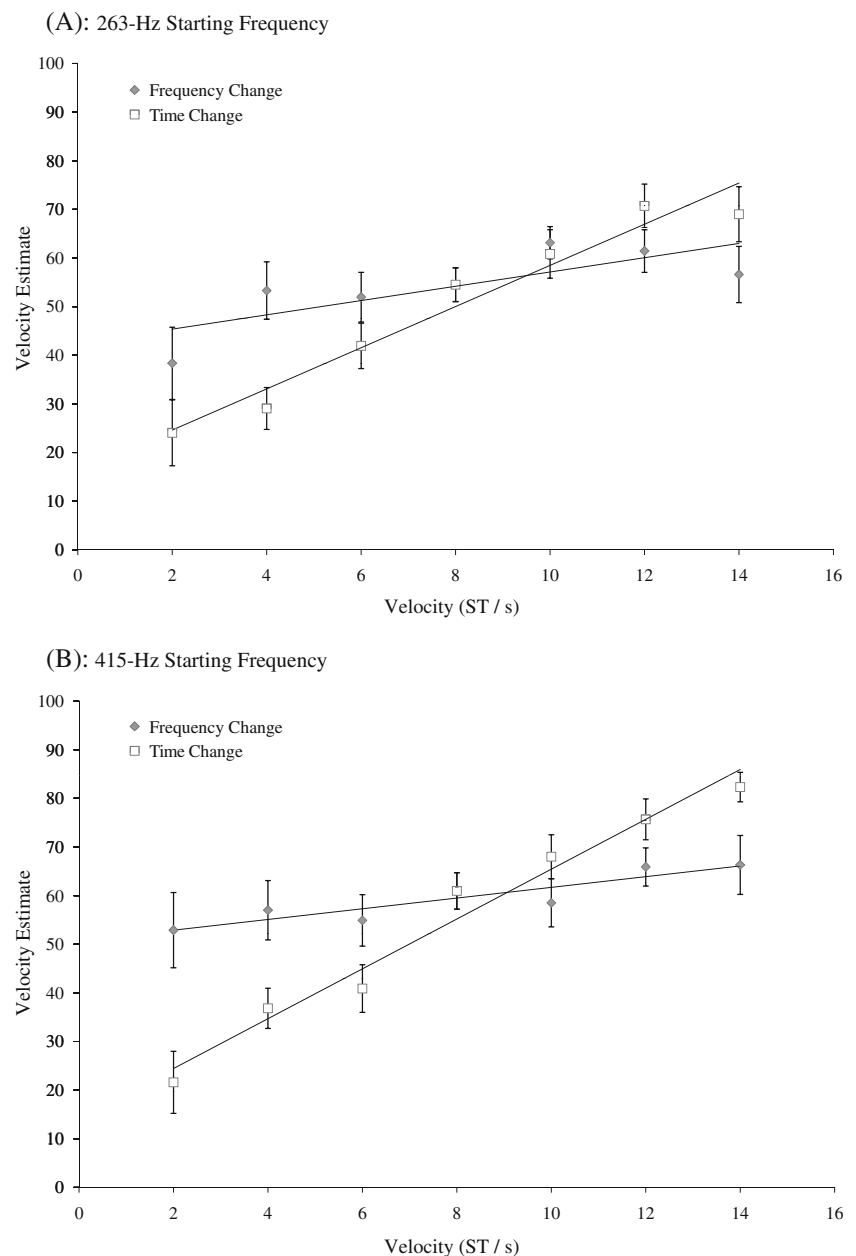
change task, in contrast, the mean frequency change estimate for monotone sequences was 5.7 (±1.9), which was significantly lower than estimates for the velocity task, $t(14) = 3.67$, $p < .01$, and the duration task, $t(14) = 7.72$, $p < .001$.

Velocity task

Figure 3 shows mean velocity estimates as a function of $\Delta f / \Delta t$ (in ST/s units) in the frequency change and the time change conditions, with regression lines superimposed; plots are shown separately for the 262-Hz starting frequency (Fig. 3A) and the 415-Hz starting frequency (Fig. 3B). The main finding from Experiment 2 was replicated here; that is, time change (Δt) contributed much more strongly to

velocity estimates than did frequency change (Δf), regardless of starting frequency. This conclusion was supported statistically by separate 2 (type of change) × 2 (starting frequency) repeated measures ANOVAs on the resulting slope and intercept estimates. The ANOVA on slope revealed a main effect of type of change, $F(1, 13) = 15.60$, $MSE = 10.34$, $p < .01$, $\eta_p^2 = .55$, no main effect of starting frequency, $F(1, 13) = 0.36$, $MSE = 2.66$, $p = .56$, $\eta_p^2 = .03$, and a type of change × starting frequency interaction, $F(1, 13) = 8.45$, $MSE = 0.66$, $p < .05$, $\eta_p^2 = .39$. Consistent with Experiment 2, average slope was much greater for the time change condition ($M = 4.68 \pm 0.46$) than for the frequency change condition ($M = 1.28 \pm 0.52$), indicating a larger contribution of time change (Δt) than of

Fig. 4 Experiment 3 results, time change task: Time change (duration) estimates increased as a function of $\Delta f/\Delta t$ for the time change condition, but not for the frequency change condition. Results are plotted separately for the 263-Hz starting frequency (A) and the 415-Hz starting frequency (B)
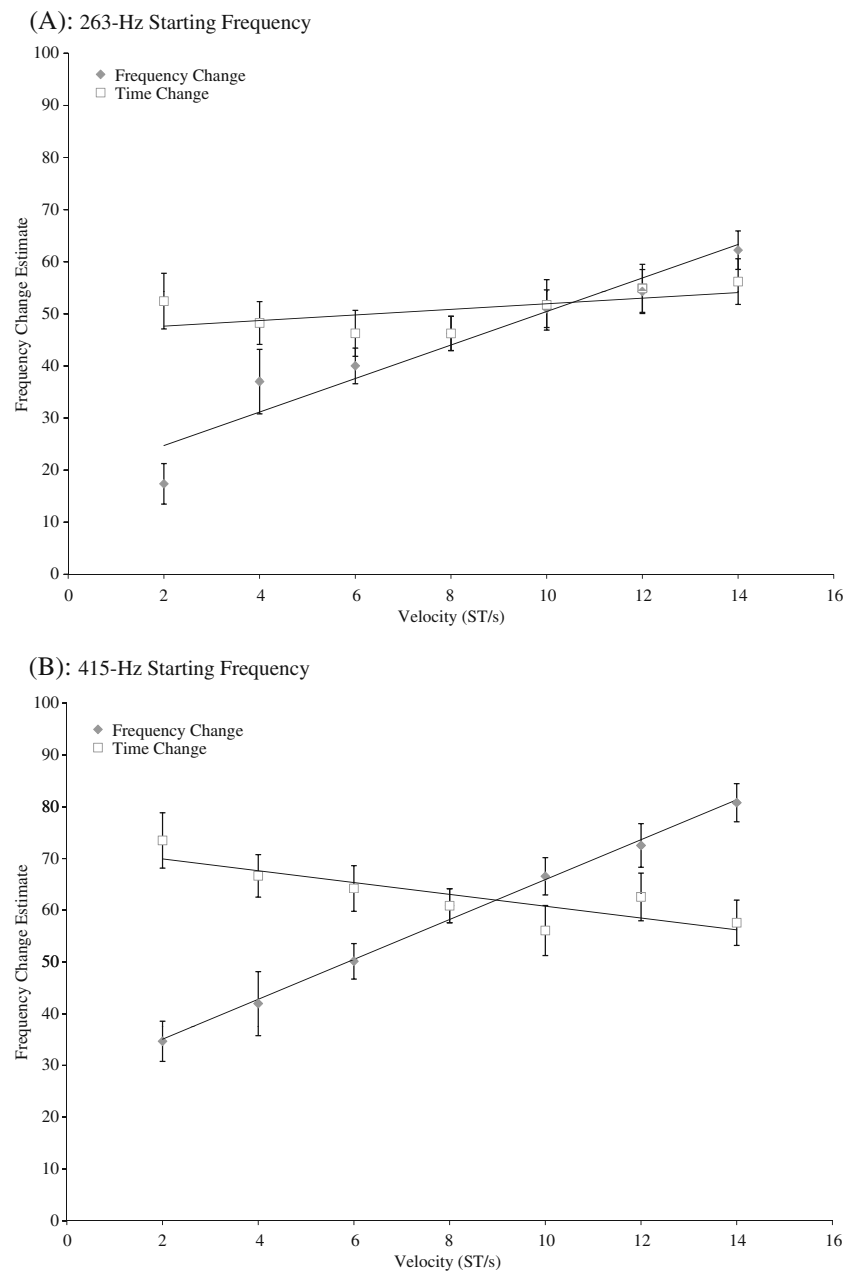

(A): 263-Hz Starting Frequency

(B): 415-Hz Starting Frequency
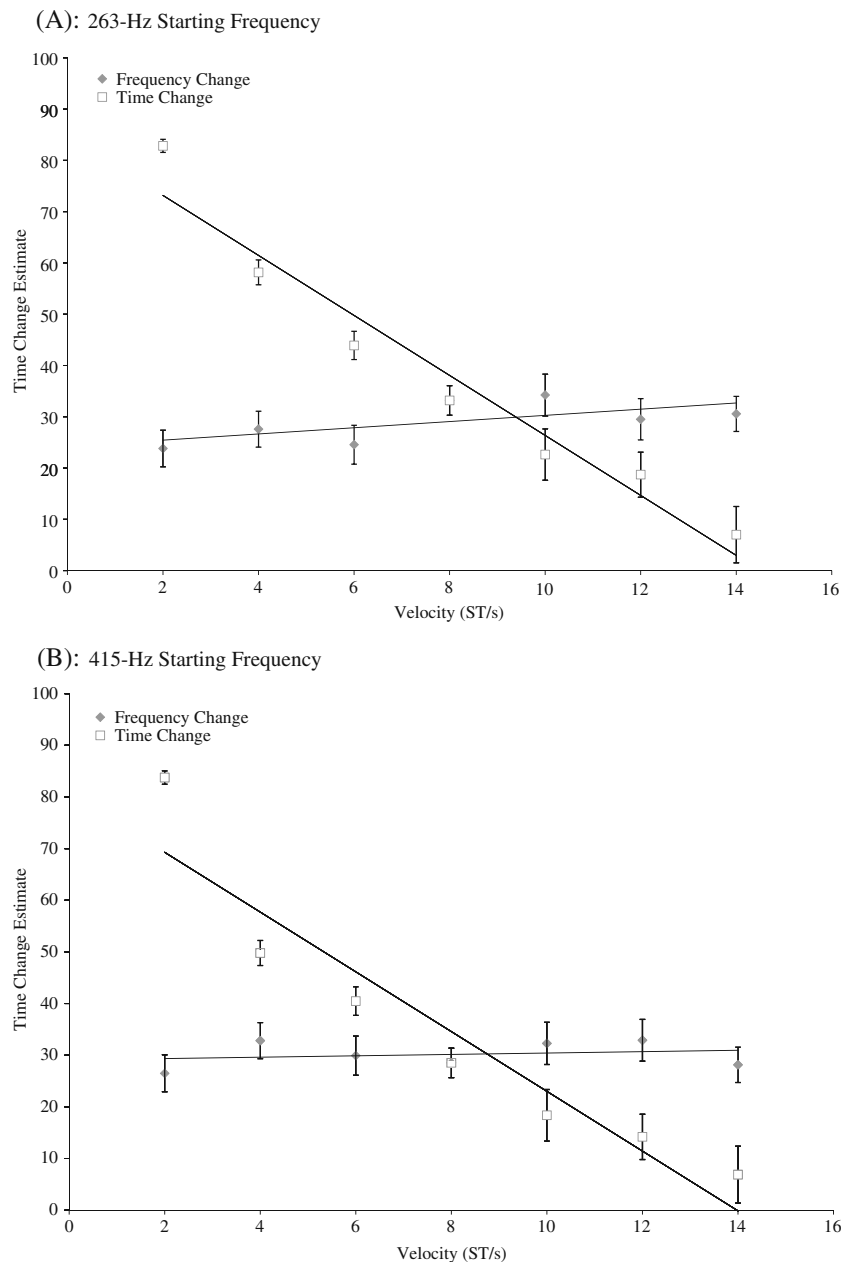
frequency change (Δf) to perceived velocity. To probe the nature of the type of change × starting frequency interaction, paired-samples t tests contrasted the slopes for the frequency change and time change conditions at each starting frequency. The slopes for the frequency change condition were similar for both starting frequencies, $t(13) = 0.63$, $p = .54$. However, the slopes for the time change condition differed significantly between starting frequencies, $t(13) = 2.49$, $p < .05$; the slope was slightly higher for the 415-Hz starting frequency ($M = 5.12 \pm 0.43$) than for the 262-Hz starting frequency ($M = 4.23 \pm 0.64$).

To further probe the nature of the starting frequency effect on slopes in the time change condition, we conducted

a 2 × 3 mixed measures ANOVA with starting frequency as a within-subjects factor and task order (velocity task first, time change task first, frequency change task first) as a between-subjects factor. Although neither the main effect of task order nor the starting frequency × task order interaction reached significance ($ps \geq .15$), this analysis was likely underpowered, and thus it is worth noting that there was some tendency for slopes to depend on both starting frequency and task order. Average slopes for the three task orders are shown in Table 2. Slopes were steeper for the 415-Hz starting frequency when the velocity task or the frequency change task was completed first, whereas slopes were similar when the duration task

Fig. 5 Experiment 3 results, frequency change task: Frequency change estimates increased as a function of $\Delta f/\Delta t$ for the frequency change condition, but not for the time change condition. Results are plotted separately for the 263-Hz starting frequency (A) and the 415-Hz starting frequency (B)

(A): 263-Hz Starting Frequency

(B): 415-Hz Starting Frequency

was completed first. We return to this point in the General Discussion section.

For the ANOVA on intercepts, the same general pattern held. There was a main effect of type of change, $F(1, 13) = 13.04$, $MSE = 209.09$, $p < .01$, $\eta_p^2 = .50$, no main effect of starting frequency, $F(1, 13) = 0.67$, $MSE = 209.09$, $p = .43$, $\eta_p^2 = .05$, and a marginally significant type of change × starting frequency interaction, $F(1, 13) = 3.89$, $MSE = 93.80$, $p = .07$, $\eta_p^2 = .23$. The intercept for the frequency change condition ($M = 46.57 \pm 6.14$) was significantly higher than the intercept for the time change condition ($M = 15.17 \pm 5.10$). Paired-samples t tests performed to probe the nature of the interaction revealed that the intercepts did not differ significantly for the frequency change or time change condition at either starting frequency, $ps \geq .19$.

Time change task

Figure 4 shows mean duration estimates as a function of $\Delta f/\Delta t$ for the frequency change and time change conditions; plots are shown separately for the 262-Hz starting frequency (Fig. 4A) and the 415-Hz starting frequency (Fig. 4B). As would be expected, time change (Δt), but not frequency change (Δf), contributed to duration estimates. Separate ANOVAs on slope and intercept measures

revealed only a main effect of type of change (slope, $F(1, 13) = 106.56$, $MSE = 723.60$, $p < .001$, $\eta_p^2 = .89$; intercept, $F(1, 13) = 116.28$, $MSE = 220.30$, $p < .001$, $\eta_p^2 = .90$). As was expected, the slope for the time change condition ($M = -5.82 \pm 0.52$) was steeper than that for the frequency change condition ($M = 0.37 \pm 0.24$), which was also not significantly different from 0 ($p = .15$). Average intercept estimates for the time change and frequency change conditions were 82.85 (± 5.64) and 26.66 (± 3.38), respectively.

Frequency change task

Figure 5 shows mean frequency change estimates as a function of $\Delta f/\Delta t$ for the frequency change and time change conditions; plots are shown separately for the 262-Hz starting frequency (Fig. 5A) and the 415-Hz starting frequency (Fig. 5B). As would be expected when frequency change is estimated, frequency change (Δf), but not total time change (Δt), contributed to frequency change estimates. This was confirmed by separate ANOVAs on slopes and intercepts, which revealed main effects of type of change for both dependent measures (slope, $F(1, 13) = 51.39$, $MSE = 1.78$, $p < .001$, $\eta_p^2 = .80$; intercept, $F(1, 13) = 62.48$, $MSE = 299.50$, $p < .001$, $\eta_p^2 = .83$). However, patterns of frequency change estimates differed slightly for the 263-Hz (Fig. 5A) and 415-Hz (Fig. 5B) starting frequencies, as evidenced by significant type of change × starting frequency interactions (slope, $F(1, 13) = 6.69$, $MSE = 1.25$, $p < .05$, $\eta_p^2 = .34$; intercept, $F(1, 13) = 5.82$, $MSE = 165.03$, $p < .05$, $\eta_p^2 = .31$). The main effect of starting frequency was significant for the intercept analysis, $F(1, 13) = 39.62$, $MSE = 106.47$, $p < .001$, $\eta_p^2 = .75$, but not for the slope analysis, $F(1, 13) = 2.40$, $MSE = 1.57$, $p = .15$, $\eta_p^2 = .16$.

As was expected, the slope ($M = 3.54 \pm 0.36$) was much steeper for the frequency change condition than for the time change condition ($M = -0.30 \pm 0.29$), which was not significantly different from 0, $t(13) = -1.05$, $p = .31$. The intercept for the frequency change condition ($M = 22.83 \pm 4.64$) was also smaller than that for the time change condition ($M = 59.39 \pm 4.53$), consistent with a positive relationship between frequency change and $\Delta f/\Delta t$ for the frequency change condition. Paired-samples t tests conducted to examine the interactions revealed that the slopes for the frequency change condition were similar for both starting frequencies, $t(13) = -1.25$, $p = .23$. However, the slopes for the time change condition differed significantly between starting frequencies, $t(13) = 2.78$, $p < .05$; the slope was steeper and negative for the 415-Hz starting frequency ($M = -1.14 \pm 0.45$) and shallow and positive for the 262-Hz starting frequency ($M = 0.54 \pm 0.38$). Intercepts differed significantly for

the frequency change condition at the 263-Hz starting frequency ($M = 18.29 \pm 4.09$) and the 415-Hz starting frequency ($M = 27.37 \pm 5.77$), $t(13) = -2.43$, $p < .05$, and for the time change condition at the 263-Hz starting frequency ($M = 46.57 \pm 4.51$) and the 415-Hz starting frequency ($M = 72.21 \pm 5.75$), $t(13) = -5.15$, $p < .001$.

In sum, the results for the frequency change task show that participants were able to judge frequency change in discrete tone sequences. This result is important because it shows that the lack of an effect of frequency change on velocity estimates in the velocity task is not simply due to the inability to reliably judge frequency change in tone sequences.

Evaluation of four models of velocity perception

Finally, we used participants' estimates of total frequency change, total time change, and frequency velocity to evaluate four models of velocity perception for each participant. Here, we evaluated goodness-of-fit for models that predicted velocity estimates from perceived time change (duration; $\Delta t'$) and the ratio of perceived frequency change to perceived time change (i.e., $\Delta f'/\Delta t'$).[4] Specifically, we used linear regression to predict individual velocity estimates for each of the seven velocity conditions from participants' estimates of total duration and total frequency change divided by total duration of each sequence, respectively. In addition, we examined a model that predicted velocity estimates from estimates of frequency change ($\Delta f'$).

A fourth model we considered used multiple regression to predict individual velocity estimates for the seven velocity conditions from a linear combination of time change estimates and frequency change estimates. Standardized coefficients (β) for regression lines for velocity estimates as a function of each predictor are shown in Table 1 for each participant and for the overall data. Values of the Bayesian information criterion (BIC; Schwartz, 1978) were used to evaluate the goodness of fit of each model for each participant (see Table 1); lower values of BIC indicate a better fit of the model to the data.

Strybel et al. (1998) found that, overall, velocity estimates for auditory apparent motion were best predicted by the total duration of the sequence; the ratio of spatial separation (analogous to Δf here) to either total duration or SOA (analogous to $\Delta f/\Delta t$) failed to account for as much variance as the model based on total duration. In contrast, in the present study, the overall data (including all the participants) were best fit by the model predicting velocity

---

[4] Strybel et al. (1998) regressed velocity estimates on the physical separation and duration values, whereas we used participants' estimates of frequency separation and duration as predictors.

estimates from a linear combination of time change and frequency change cues (BIC = 456.9). However, the substantial individual differences in the best-fitting model warrant some caution in interpreting the model fits to the overall data. In this regard, inspection of Table 1 reveals that the sign of beta weights (β) was not always consistent across participants for a predictor; in particular, this was true for the frequency change model and the frequency change component of the additive model, where negative β values indicate that 6 participants provided higher velocity estimates for sequences with smaller values of frequency change.[5] This inconsistency serves to highlight the large individual differences in the contribution of frequency change information to perceived velocity

Finally, to evaluate the extent to which individual differences in goodness of model fits were potentially due to musical training, we calculated correlation coefficients for years of musical training and β values for each of the predictors. Years of musical training was significantly correlated with β weights for the frequency change predictor in the frequency model (Spearman's ρ = -0.55, p < .05) and the frequency change predictor in the additive model (Spearman's ρ = -0.61, p < .05). Interestingly, both correlations were negative, indicating that participants with more musical training were *less* likely to use frequency change information when judging velocity than were individuals with less musical training.

General discussion

Three experiments examined velocity perception for sounds moving in frequency space. The first two experiments assessed the relative contribution of frequency change (Δf) and time change (Δt) to participants' velocity estimates for tone glides that moved continuously in frequency space (Experiment 1) and for tone sequences that moved in frequency space in discrete steps (Experiment 2). In order to compare the results of the two experiments, stimulus velocities were matched across experiments and ranged from 2 to 14 ST/s in 2-ST/s steps. For tone glides (Experiment 1), we found that time change (Δt) and frequency change (Δf) cues contributed approximately equally to perceived velocity. For tone sequences, in contrast, perceived stimulus velocity was based primarily on Δt, with little to no contribution from Δf.

To address the possibility that the somewhat surprising lack of contribution of frequency change to judgments about the velocity (rate of frequency change) of tone sequences was simply due to participants' poor ability to

judge frequency change in tone sequences, the third experiment had participants make separate estimates of total frequency change (Δf) and total time change (duration; Δt), in addition to velocity. The availability of separate Δf, Δt, and $\Delta f/\Delta t$ estimates permitted us to evaluate a number of models of velocity perception; specifically, we used estimates of total duration and frequency change to predict perceived velocity. The results of Experiment 3 conclusively showed that participants were able to make frequency change estimates for tone sequences, yet their velocity estimates still relied almost exclusively on time change cues. Model fits revealed substantial individual differences, with overall data (including all the participants) best fit by an additive model that predicted velocity estimates from a linear combination of time change and frequency change estimates, where time change cues contributed more strongly to perceived velocity (β = -0.57) than did frequency change cues (β = .33).

One unexpected finding was the influence of starting frequency on velocity estimates in Experiment 3, which was not present in the first two experiments. Specifically, the slope of the regression line relating perceived velocity to $\Delta f/\Delta t$ in the time change condition was steeper for the 415-Hz starting frequency condition than for the 263-Hz starting frequency condition. Post hoc tests revealed that the starting frequency effect was influenced by task order (see Table 2), such that the slope for velocity estimates was steeper for the 415-Hz starting frequency when participants estimated velocity or frequency change first, but not when the time change task was completed first. This result is difficult to interpret practically but likely reflects carryover effects resulting from early attention to different aspects of the stimuli; specifically, there was no influence of starting frequency when participants completed the time change task first, for which ignoring frequency was likely a strategy. However, when the listener completed a task that required attention to frequency first (either the velocity or the frequency change task), slopes were steeper for the higher starting frequency.

Overall, the present findings support a parallel between velocity perception in physical space and frequency space. For tone glides moving continuously in frequency space, the results are consistent with work in vision showing that observers jointly make use of both spatial and temporal cues when judging the velocity of continuously moving objects (Algom & Cohen-Raz, 1984, 1987). Moreover, velocity perception of visual and auditory objects moving continuously in physical space has been shown to be remarkably consistent (Waugh et al., 1979). The present findings are also consistent with work by Strybel and colleagues (1998) for discrete auditory stimuli presented at a series of loudspeaker locations, whereby total time change (Δt) contributed strongly to perceived velocity, but spatial

---

[5] It is important to note, however, that no negative β values for frequency predictors differed significantly from 0, ps $\geq$ .28.

Table 1 Standardized coefficients (β) and goodness of fit (*indexed by BIC*) for regression lines through velocity estimates as a function of the perceived duration ($\Delta t'$), perceived frequency change ($\Delta f'$), the ratio of perceived frequency change to perceived duration ($\Delta f'/\Delta t'$), and a linear combination of perceived duration and perceived frequency change. Model fits are shown for individual participants and for the overall data.

ID	Model								
	Duration ($\Delta t'$)		Frequency Change ($\Delta f'$)		Velocity ($\Delta f'$) / ($\Delta t'$)		Additive $a(\Delta f') + b(\Delta t')$		
	β	BIC	β	BIC	β	BIC	$\beta\Delta t$	$\beta\Delta f$	BIC
1	-0.73	32.45	0.25	36.35	0.56	34.57	-0.71	0.14	33.33
2	-0.84	31.46	-0.31	37.76	0.54	36.38	-0.84	-0.01	32.57
3	-0.91	26.59	-0.10	36.24	0.58	33.96	-0.97	0.21	26.34
4	-0.66	33.84	0.41	36.07	0.39	36.19	-0.69	0.45	32.46
5	-0.04	32.23	0.68	28.70	0.02	32.23	-0.20	0.73	29.42
7	-0.78	32.61	0.16	37.68	0.48	36.36	-0.77	0.05	33.69
8	-0.13	32.90	0.34	32.29	0.23	32.70	-0.59	0.72	31.90
9	-0.34	32.55	0.49	31.70	0.57	31.07	-0.35	0.50	31.83
10	-0.92	24.39	-0.38	33.87	0.54	32.78	-0.90	-0.06	25.40
11	-0.02	33.57	0.66	30.34	0.61	31.00	0.20	0.72	31.07
12	-0.91	24.21	-0.03	34.02	0.50	32.39	-0.91	0.00	25.32
13	-0.62	33.25	0.48	34.51	0.76	31.23	-0.82	0.71	26.48
14	-0.82	24.20	-0.03	30.44	0.34	29.74	-0.83	-0.13	25.01
15	-0.84	28.91	-0.26	35.54	0.54	33.99	-0.82	-0.15	29.59
Overall	-0.49	466.84	0.27	486.74	0.77	486.86	-0.54	0.36	456.92

cues (Δ) were largely ignored. One important difference between the stimuli in the present study and those used by Strybel et al. (1998) was that Strybel's stimuli were experienced as moving despite being made up of discrete elements, whereas in the present study, motion in frequency space was only implied. Finally, the present results are consistent with other related work on auditory apparent motion for discrete auditory sequences, which has shown that whether listeners perceived a sound sequence as moving or not depended on the temporal, but not the spatial, characteristics of the sequences (Strybel, Manlingas, Chan, & Perrott, 1990; Strybel & Neale, 1994; Strybel, Witty, & Perrott, 1992).

Strybel and colleagues have proposed several possible explanations for the relatively weak contribution of spatial cues to perceived velocity for discrete sounds moving in physical space that have the potential to be applicable here. Thus, we will consider these in relation to estimates of velocity for sounds moving in frequency space. One possibility is that the use of a relatively small number of levels of Δs may have limited the extent to which listeners used spatial change cues to estimate velocity in the study by Strybel et al. (1998). Supporting this possibility, Melara and Mounts (1994) demonstrated that changing the number of levels of one dimension of a multidimensional stimulus modulates selective attention to individual dimensions; Strybel et al. (1998) implemented only two levels of Δs, in contrast with ten levels of Δt, which could have reduced attention to the spatial properties of auditory stimuli. However, in the present study, we used an equal number of levels of Δf and Δt but still observed a pattern of results similar to that in Strybel et al. (1998).

A second possibility concerns differences in specialization of the auditory and visual modalities for temporal and spatial processing, respectively (Hay, Pick, & Ikeda, 1965; Jack & Thurlow, 1973; Shams, Kamitani, & Shimojo, 2000). Specifically, Strybel et al. (1992) hypothesized that motion perception in the auditory and visual

Table 2 Experiment 3: Average Slope (±SEM) for the Regression Line for the Time Change Condition Shown Separately for the 262-Hz and 415-Hz Starting Frequencies as a Function of Task Order

		Task Order		
		VPT	PTV	TVP
Starting Frequency	262 Hz	5.09 (1.05)	2.86 (1.05)	4.88 (1.17)
	415 Hz	6.40 (0.61)	4.21 (0.61)	4.67 (0.69)

modalities is affected most by the dimension to which that modality is the most sensitive: time for audition and space for vision. This hypothesis has received support from studies demonstrating that judgments of the spatial location of an auditory stimulus are influenced by visual information, such as in the ventriloquism effect (Jack & Thurlow, 1973); conversely, judgments about the timing of a visual stimulus are affected by auditory information, such as in auditory driving (Rencanzone, 2002). However, this explanation is not sufficient to account for the relatively equal weighting of frequency and time cues for continuous tone glides.

Finally, Strybel et al. (1990) considered the possibility that the perception of velocity in an auditory apparent motion paradigm is dependent on the distance between receptors on the tonotopically arranged basilar membrane in the same way that perception of velocity for visual apparent motion is dependent on *retinal* separation. Thus, perception of apparent velocity in discrete auditory sequences moving in physical space would depend on *frequency* separation of sequence elements, rather than on spatial separation. However, a subsequent study failed to support this hypothesis. Strybel and Menges (1998) found that for relatively low frequencies (500 Hz, 1000 Hz), auditory apparent motion was reported only when both tones were within the same critical band and that the frequency separation required for perception of apparent motion did not increase with increasing SOA. Moreover, this proposition is not supported by the results of the present study; participants' estimates of velocity for discrete sequences were unchanged for values of Δf spanning between 3 and 21 ST.

More broadly, the present results contribute to work on attentional tracking of auditory stimuli. It has been proposed that time-frequency regularities of an auditory stimulus can effectively tune attention to time-frequency of future information-carrying events (Jones et al., 2006); listeners tend to extrapolate auditory patterns along the trajectory of motion or implied motion on the basis of time-frequency information (Crum & Hafter, 2008; Johnston & Jones, 2006). One implication of the present study is that attention may be tuned differently by discrete and continuous sounds moving in frequency space. Specifically, discrete auditory patterns may promote tuning of future-oriented attention in response primarily to the temporal characteristics of the pattern, whereas attention may be tuned on the basis of joint time-frequency information for continuous sounds.

One potential caveat for the conclusion that tone sequences promote attentional tuning primarily on the basis of temporal structure is that time changes (Δt) and frequency changes (Δf) were not matched for salience in the present study. Thus, it is possible that the greater

contribution of time change cues to perceived velocity in tone sequences may be partly due to differences in the salience of the time change versus frequency change information. The results for the frequency change task in Experiment 3 partially address this possibility by demonstrating that listeners had no trouble discriminating frequency changes in discrete sequences. Nonetheless, previous studies have demonstrated that when two stimulus dimensions are not matched for salience, the more salient dimension can affect the perception of the less salient dimension, thereby potentially qualifying conclusions regarding perception of the less salient dimension (Garner & Felfoldy, 1970; Melara & Mounts, 1993, 1994). In this regard, Ellis and Jones (2009) recently showed that when melodic accents (frequency cues) and temporal accents (time cues) are matched for salience, melodic accents make a larger contribution to perceived meter in music than had previously been observed in studies that did not equate the salience of the two accent types.

Summary

Frequency change (Δf) and time change (Δt) cues were shown to make different contributions to the perceived velocity of tone glides that moved continuously in frequency space and tone sequences that moved in discrete steps. For tone glides, both time change (duration) cues and frequency change cues contributed approximately equally to perceived velocity. In contrast, for tone sequences, which only provide inferred information about velocity, estimates were based almost entirely on temporal information, with surprisingly little contribution of frequency information. Separate judgments about frequency change (Δf) and time change (Δt) ruled out the possibility that the lack of contribution of frequency information to perceived velocity of tone sequences was simply due to poor ability to judge total frequency change in the tone sequences. Overall, the present results generalize previous findings of Strybel and colleagues for the perception of sounds moving in physical space to the perception of sounds moving in frequency space and contribute to the growing body of work supporting a relationship between motion in frequency space and motion in physical space.

References

Ahissar, M., Ahissar, E., Bergman, H., & Vaadia, E. (1992). Encoding of sound-source location and movement: Activity of single neurons and interactions between adjacent neurons in the monkey auditory cortex. *Journal of Neurophysiology, 67*, 203–215.

Algom, D., & Cohen-Raz, L. (1984). Visual velocity input–output functions: The integration of distance and duration onto subjective velocity. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 486–501.

Algom, D., & Cohen-Raz, L. (1987). Sensory and cognitive factors in the processing of visual velocity. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 3–13.

Altman, J. A., Syka, J., & Shmigidina, G. N. (1971). Neuronal activity in the medial geniculate body of the cat during monaural and binaural stimulation. *Experimental Brain Research, 10*, 81–93.

Antis, S. M., & Saida, S. (1985). Adaptation to auditory streaming of frequency-modulated tones. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 257–271.

Bill, J. C., & Teft, L. W. (1969). Space–time relations: Effects of time on perceived visual extent. *Journal of Experimental Psychology, 81*, 196–199.

Boersma, P., & Weenik, D. (2005). Praat: Doing phonetics by computer (Version 4.4.14).

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century-Crofts.

Bregman, A. S. (1990). *Auditory scene analysis.* Cambridge: MIT Press.

Burns, E. M. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music* (Vol. 2, pp. 215–264). San Diego: Academic Press.

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

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

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

Crum, P. A. C., & Hafter, E. R. (2008). Predicting the path of a changing sound: Velocity tracking and auditory continuity. *The Journal of the Acoustical Society of America, 124*, 1116–1129.

Dilley, L. C., & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language, 59*, 294–311.

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

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

Frick, R. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin, 97*, 412–429.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1*, 225–241.

Goldstein, A. G. (1957). Judgments of visual velocity as a function of length of observation time. *Journal of Experimental Psychology, 54*, 457–461.

Grantham, D. W., & Wightman, F. L. (1978). Detectability of varying interaural temporal differences. *The Journal of the Acoustical Society of America, 63*, 511–523.

Grondin, S., & Plourde, M. (2007). Discrimination of time intervals presented in sequences: Spatial effects with multiple auditory sources. *Human Movement Science, 26*, 702–716.

Hafter, E. R., Schlauch, R. S., & Tang, J. (1993). Attending to auditory filters that were not stimulated directly. *The Journal of the Acoustical Society of America, 94*, 743–747.

Hay, J. C., Pick, H. L., & Ikeda, K. (1965). Visual capture produced by prism spectacles. *Psychological Science, 2*, 215–216.

Henry, M. J., & McAuley, J. D. (2009). Evaluation of an imputed pitch velocity model of the auditory kappa effect. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 551–564.

Henry, M. J., McAuley, J. D., & Zaleha, M. F. (2009). Evaluation of an imputed pitch velocity model of the auditory tau effect. *Perception & Psychophysics, 71*, 1399–1413.

Howard, J. H., O'Toole, A. J., Parasuraman, R., & Bennett, K. B. (1984). Pattern-directed attention in uncertain-frequency detection. *Perception & Psychophysics, 35*, 256–264.

Howard, J. H., O'Toole, A. J., & Rice, S. E. (1986). The role of frequency versus information cues in uncertain frequency detection. *The Journal of the Acoustical Society of America, 79*, 788–791.

Jack, C. E., & Thurlow, W. R. (1973). Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Perceptual and Motor Skills, 37*, 967–979.

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, M. R. (1976). Time, Our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*, 323–355.

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., Johnston, H. M., & 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: Oxford University Press.

Krumhansl, C. (1991). Music psychology: Tonal structures in perception and memory. *AnnualReview of Psychology, 42*, 277–303.

Krumhansl, C. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*, 159–179.

Lappin, J. S., Bell, H. H., Harm, O. J., & Kottas, B. (1975). On the relation between time and space in the visual discrimination of velocity. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 383–394.

Lund, J. (1988). Anatomical organization of macaque striate visual cortex. *Annual Review of Neuroscience, 11*, 253–288.

Maunsell, J. H. R., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey: I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology, 49*, 1127–1147.

Melara, R. D., & Mounts, J. R. W. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition, 21*, 627–645.

Melara, R. D., & Mounts, J. R. (1994). Contextual influences on interactive processing: Effects of discriminability, quantity, and uncertainty. *Perception & Psychophysics, 56*, 73–90.

Melara, R. D., & O'brien, T. P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General, 116*, 323–336.

Merzenich, M. M., Colwell, S. A., & Anderson, R. A. (1982). Auditory forebrain organization. In C. N. Woosely (Ed.), *Cortical sensory organization: 3. Multiple auditory areas* (pp. 43–50). Clifton: Humana.

Orban, G. A., Kennedy, H., & Maes, H. (1981). Response to movement of neurons in areas 17 and 18 of the cat: Velocity selectivity. *Journal of Neurophysiology, 45*, 1043–1058.

Perrott, D. R., Buck, V., Waugh, W., & Strybel, T. Z. (1979). Dynamic auditory localization: Systematic replication of the auditory velocity function. *The Journal of Auditory Research, 19*, 277–285.

Posner, M. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32*, 3–25.

Posner, M., Snyder, C., & Davidson, B. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*, 160–174.

Rencanzone, G. H. (2002). Auditory influences on visual temporal rate perception. *Journal of Neurophysiology, 89*, 1078–1093.

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., & Pittenger, J. B. (2007). Tau and kappa effects in physical space: The case of audition. *Psychological Research, 71*, 201–218.

Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature, 408*, 788.

Smeets, J. B. J., & Brenner, E. (1995). Perception and action are based on the same visual information: Distinction between position and velocity. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 19–31.

Strybel, T. Z., & Menges, M. L. (1998). Auditory apparent motion between sine waves differing in frequency. *Perception, 27*, 483–495.

Strybel, T. Z., & Neale, W. (1994). The effect of burst duration, interstimulus onset interval, and loudspeaker arrangment on auditory apparent motion in the free field. *The Journal of the Acoustical Society of America, 96*, 3463–3475.

Strybel, T. Z., Manlingas, C. L., Chan, O., & Perrott, D. R. (1990). A comparison of the effects of spatial separation on apparent motion in the auditory and visual modalities. *Perception & Psychophysics, 47*, 439–448.

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.

Strybel, T. Z., Witty, A. M., & Perrott, D. R. (1992). Auditory apparent motion in the free field: The effects of stimulus duration and separation. *Perception & Psychophysics, 52*, 139–143.

Stumpf, E., Toronchuk, J. M., & Cynader, M. S. (1992). Neurons in cat primary auditory cortex sensitive to correlates of auditory motion in three-dimensional space. *Experimental Brain Research, 88*, 158–168.

Waugh, W., Strybel, T. Z., & Perrott, D. R. (1979). Perception of moving sounds: Velocity discrimination. *The Journal of Auditory Research, 19*, 103–110.

Werner, S., & Keller, E. (1994). Prosodic aspects of speech. In E. Keller (Ed.), *Fundamentals of speech synthesis and speech recognition: Basic concepts, state of the art, and future challenges* (pp. 23–40). Chichester: Wiley.