

NO SUSTAINED SOUND ILLUSION IN RHYTHMIC SEQUENCES

BRUNO H. REPP
Haskins Laboratories

RACHEL J. MARCUS
Yale University

RECENT RESEARCH HAS SHOWN THAT SUBDIVISION OF intervals between beats makes the beat tempo seem slower—a “divided time illusion” (DTI) in music. Another temporal illusion described in the psychophysical literature is that a sustained sound seems longer than a silent interval of the same duration. This “sustained sound illusion” (SSI) may be due to acceleration of an internal pacemaker by continuous sound, or it may result from slower perception of sound offsets than of sound onsets. Experiment 1 tested the pacemaker acceleration hypothesis in a rhythmic context by asking musicians to compare or reproduce the tempi of isochronous tone sequences played *legato* (“filled”) or *staccato* (“unfilled”). There was no indication that *legato* sequences were perceived as slower than *staccato* sequences. Experiment 2 tested the delayed offset perception hypothesis by asking musicians to judge the relative time of occurrence of abrupt or decaying tone offsets in the inter-onset intervals of isochronous sequences. There was no evidence of delayed perception of abrupt offsets, and decaying offsets were perceived only slightly late. These results suggest that the SSI, unlike the DTI, does not occur in rhythmic contexts and thus is probably not of musical relevance. More generally, the results challenge some proposed explanations of this illusion and call for further research on the conditions under which it does occur.

Received May 6, 2009, accepted March 11, 2010.

Key words: time perception, filled duration illusion, interval subdivision, tempo perception, sound offset perception

IT HAS LONG BEEN KNOWN THAT THE PERCEIVED duration of a temporal interval depends on what happens during that interval. Such effects are familiar

from everyday experience with time spans in the range of minutes and hours, but they occur in the perception of short durations as well, albeit with reversed polarity. While long durations usually seem longer when nothing is happening, Hall and Jastrow (1886) were among the first to report that a short (< 3 s) interval filled with clicks seems longer than an equally long silent interval. Meumann (1896), among others, conducted detailed research on this phenomenon, which later became known as *filled duration illusion* (Buffardi, 1971; Thomas & Brown, 1974). Another, perhaps better, name for it is *divided time illusion* (DTI; ten Hoopen, Miyauchi, & Nakajima, 2008), to distinguish it from another filled duration illusion, discussed below. DTIs also have been found in the visual and tactile modalities (Buffardi, 1971; Meumann, 1896; Roelofs & Zeeman, 1951).

Two recent studies have demonstrated that the DTI extends to rhythmic sequences (Repp, 2008; Repp & Bruttomesso, 2009): When the intervals between isochronous “beat” tones are subdivided by additional tones, the beat tempo appears to be slower, presumably because the intervals are perceived as longer. This effect was found in several different tasks (synchronization-continuation tapping, reproduction, perceptual comparison, and music performance), but it was considerably smaller than the illusion found in psychophysical studies, which typically have focused on the judgment of single intervals. The timing constraints imposed by a musical rhythm were likely responsible for this difference, as Meumann (1896) already noted. The DTI has the interesting implication that musicians may need to play music containing many notes slightly faster than music containing few notes, in order to perceive themselves (or be perceived by others) as playing at the same beat tempo. Repp and Bruttomesso report some data supporting this prediction.

Psychophysical studies also have demonstrated another kind of filled duration illusion: When the duration of a continuous sound (marked by sound onset and offset) is compared with the duration of an equally long silence (marked by a sound offset and a subsequent sound onset, or by two very brief sounds), the continuous sound seems longer than the silence (Meumann, 1896). We will

call this effect *sustained sound illusion* (SSI), to distinguish it from the DTI. A possibly related finding is that sounds are judged to last longer than visual stimuli of the same duration (Behar & Bevan, 1961; Goldstone, Boardman, & Lhamon, 1959; Goldstone & Goldfarb, 1963, 1964a, 1964b; Goldstone & Lhamon, 1974; Walker & Scott, 1981; Wearden, Edwards, Fakhri, & Percival, 1998). Because visual stimuli are silent, this crossmodal illusion may be essentially the same auditory phenomenon. Indeed, it is even conceivable that the DTI and the SSI have the same underlying causes. The general principle may be that silence seems shorter than non-silence, for durations up to a few seconds.

There are relatively few studies of the SSI, and their results are not very consistent. Craig (1973) presented two successive tones and asked participants to adjust the silent interval between them until its duration matched that of the first tone. For tone durations ranging from 100 to 1200 ms, he found that the silent interval had to be longer by a constant amount to be perceived as equally long; that constant was 657 ms on average, a huge difference. He also found similar, though somewhat smaller, differences with visual and vibrotactile stimuli. Auditory research conducted at about the same time in Germany by Zwicker (1969/1970) and Burghardt (1972, 1973) led to different results. (See Fastl & Zwicker, 2007, for a brief summary.) Their participants adjusted the duration of a sound to equal the duration of the silence between two later sounds or vice versa; Burghardt also used two other paradigms, with similar results. For sound durations ranging from 5 ms to about 900 ms, Zwicker found that the silence had to be about twice as long as the sound to be perceived as equally long. Thus, he found a constant ratio, not a constant difference. Burghardt's findings replicated Zwicker's results up to sound durations of about 400 ms. Unlike Zwicker, however, he found that the difference decreased for sound durations greater than 400 ms and disappeared around 1 s. The magnitude of the sound-silence difference also depended on the nature of the sound (noise versus tone, and pitch of tone).

Different explanations have been proposed for the two types of filled duration illusion. The DTI has been attributed to an increase in the amount of temporal information presented during an interval (Ornstein, 1969), to more complex encoding and decoding operations (Thomas & Brown, 1974), or to processing time required for each subinterval (Nakajima, 1978). Because a sustained steady-state sound does not differ from silence in temporal information content, these hypotheses cannot explain the SSI. One proposed explanation for the latter is that continuous sound makes an internal clock or pacemaker run faster, which leads to a shrinkage of subjective time and consequently to an overestimation

of duration (Penney, Gibbon, & Meck, 2000; Wearden et al., 1998; Wearden, Norton, Martin, & Montford-Bebb, 2007; Wearden, Todd, & Jones, 2006). This hypothesis implies a proportional or at least increasing difference between the judged durations as mean duration increases, which was indeed observed by Wearden and colleagues for durations ranging from 300 to 1200 ms, although they found no difference at durations below 300 ms. These results are quite different from those of Burghardt (1972, 1973), which may have been due to differences in method. The *pacemaker acceleration hypothesis* could in principle also account for the DTI.

Another explanation of the SSI is that perception of (abrupt) sound offset may have a longer latency than perception of sound onset (Grondin, 1993, 2001, 2003). According to this *delayed offset perception hypothesis*, the subjective duration of a continuous sound (which starts with an onset and ends with an offset) will be longer than its physical duration, whereas the subjective duration of a silence (which starts with a sound offset and ends with a sound onset) will be shorter than its physical duration. Thus, a silence must be lengthened for its subjective duration to equal that of a sound. This hypothesis predicts a constant difference between the subjective durations of sound and silence, at least beyond a certain minimal duration, as was found by Craig (1973). Burghardt (1972, 1973), who instead found a proportional or decreasing difference, proposed a more complex model according to which the sensory magnitude of a sound grows exponentially from its onset up to an asymptote and decays exponentially after its offset. The same threshold is assumed to govern sound onset and sound offset detection, but a lower threshold is assumed for the detection of the onset and offset of silence. Consequently, the silence following a sound is perceived to begin some time after the sound is perceived to end, and the onset of the next sound is perceived some time after the silence preceding it has been perceived to end. With these assumptions of "dead time" between sound and silence, Burghardt was able to achieve a reasonable fit of his model to his empirical data. Figure 12.4 in Fastl and Zwicker (2007) illustrates the offset delay hypothesis, though without the assumption of different thresholds for sound and silence.

In the present study we asked whether the SSI, like the DTI (Repp, 2008; Repp & Bruttomesso, 2009), is found in rhythmic contexts and thus potentially has implications for music performance. Both the pacemaker acceleration hypothesis and the offset delay hypothesis refer to basic perceptual and neurophysiologic processes that should transcend any details of materials, whether musical or nonmusical. Therefore, we found it quite reasonable to suppose that the SSI might occur in music. As in Repp's earlier studies of the DTI, we focused on interval

durations that could plausibly constitute interbeat intervals in music (Parncutt, 1994). Experiment 1 addressed the pacemaker acceleration hypothesis, whereas Experiment 2 tested the delayed offset perception hypothesis.

Experiment 1

In music, a contrast between filled and empty intervals occurs between a sequence of long notes and a sequence of short notes followed by rests, or between *legato* and *staccato* articulation. In *legato* playing, successive notes are connected without intervening silence, whereas in *staccato* playing each note is played much shorter than notated, so that the produced tone ends long before the next tone begins. Depending on the instrument, the tone offset may be more or less abrupt. Psychophysical experiments generally use artificial tones with abrupt onsets and offsets. We first conducted the present experiment with piano tones, whose offset is rather gradual, but later replicated it with artificial tones that had abrupt offsets, to enhance the contrast between sound and silence. Our prediction was that *legato* sequences should be perceived as slower (i.e., containing longer interonset intervals) than *staccato* sequences if there is an SSI.

The experiment included two tasks, perceptual judgment and reproduction, which had been used previously to demonstrate the DTI in rhythmic sequences (Repp, 2008; Repp & Bruttomesso, 2009). In the perception task, participants listened to two tone sequences (S1, S2) and judged whether the tempo of S2 was faster, the same as, or slower than the tempo of S1. S1 was played either *legato* or *staccato*, whereas S2 was always played *staccato*. If *legato* sequences are perceived as slower than *staccato* sequences, then when S1 is played *legato* participants should judge S2 as faster when its tempo is in fact the same, and as the same when its tempo is in fact slower. When S1 is played *staccato*, there should be no such bias. In the reproduction task, participants listened to a sequence that was played either *legato* or *staccato* and tried to reproduce its tempo exactly by tapping. Their taps produced auditory feedback in the form of *staccato* tones. If *legato* sequences are perceived as slower than *staccato* sequences, participants should tap slower when reproducing the tempo of a *legato* sequence than when reproducing the tempo of a *staccato* sequence.

Method

PARTICIPANTS

Nine graduate students from the Yale School of Music and both authors (BHR, RJM) participated, RJM only in the piano tone version of the experiment. The graduate student musicians (6 women and 3 men, ages 22–28) had

been trained on their primary instruments (piano-2, violin-2, viola, double bass, clarinet, bassoon, and harp) for 13–22 years. BHR (age 64) is a life-long amateur pianist with 10 years of instruction in childhood, and RJM (age 19) has had music training in voice (9 years) and violin (3 years).

MATERIALS AND EQUIPMENT

In the perception task, each trial consisted of two isochronous 5-tone sequences (S1, S2). S1 had one of seven possible interonset interval (IOI) durations: 660, 690, 720, 750, 780, 810, or 840 ms. S2 had a fixed IOI duration of 750 ms. The IOI between the last tone of S1 and the first tone of S2 was twice the IOI of S1. S1 consisted of low-pitch tones that formed an ascending chromatic scale starting on E2 (82.4 Hz, MIDI pitch 40) and ending on G#2 (103.8 Hz, MIDI pitch 44). The pitch change was necessary to distinguish successive *legato* tones, and low pitches were chosen to minimize the decay of sustained piano tones (Martin, 1947; Repp, 1995). S2 always consisted of a constant repeated pitch, F#2 (92.5 Hz, MIDI pitch 42). A constant intensity (MIDI velocity) was specified for all tones. S1 was played either *legato* or *staccato*, whereas S2 was always played *staccato*. *Legato* tones were specified to last as long as the IOI, whereas *staccato* tones were specified to last 40 ms. In the reproduction task, only S1 was presented in each trial, but participants heard tones contingent on their taps, so that their reproduction sounded like S2 in the perception task.

The artificial tones were synthesized using the signal-processing component of MAX/MSP 4.0.9 software. Each tone consisted of a fundamental frequency and the next three harmonics. The amplitudes of all four partials were specified to be equal. The amplitude of each tone was constant over its duration, except for 5-ms linear ramps at onset and offset to prevent audible clicks from occurring.

The piano tones were produced by a Roland RD-250s digital piano. Figure 1 shows waveforms and intensity contours of a *legato* and a *staccato* sequence. We generated these displays using Praat software (Boersma & Weenink, 2007) after recording the sounds by holding one headphone against the computer microphone. Sustained *legato* tones decayed by about 1 dB per 100 ms, whereas *staccato* tones reached the same peak amplitude during their short sustained duration but then decayed by nearly 10 dB per 100 ms, reaching the ambient noise floor within about 400 ms. Thus, at least the second half of the IOI between *staccato* tones was effectively silent, but perceptually the tone offset probably occurred early in the IOI (see Experiment 2). The fast decay simulates the effect of the dampers on the piano string vibrations; it can also be seen at the end of the last tone in the *legato*

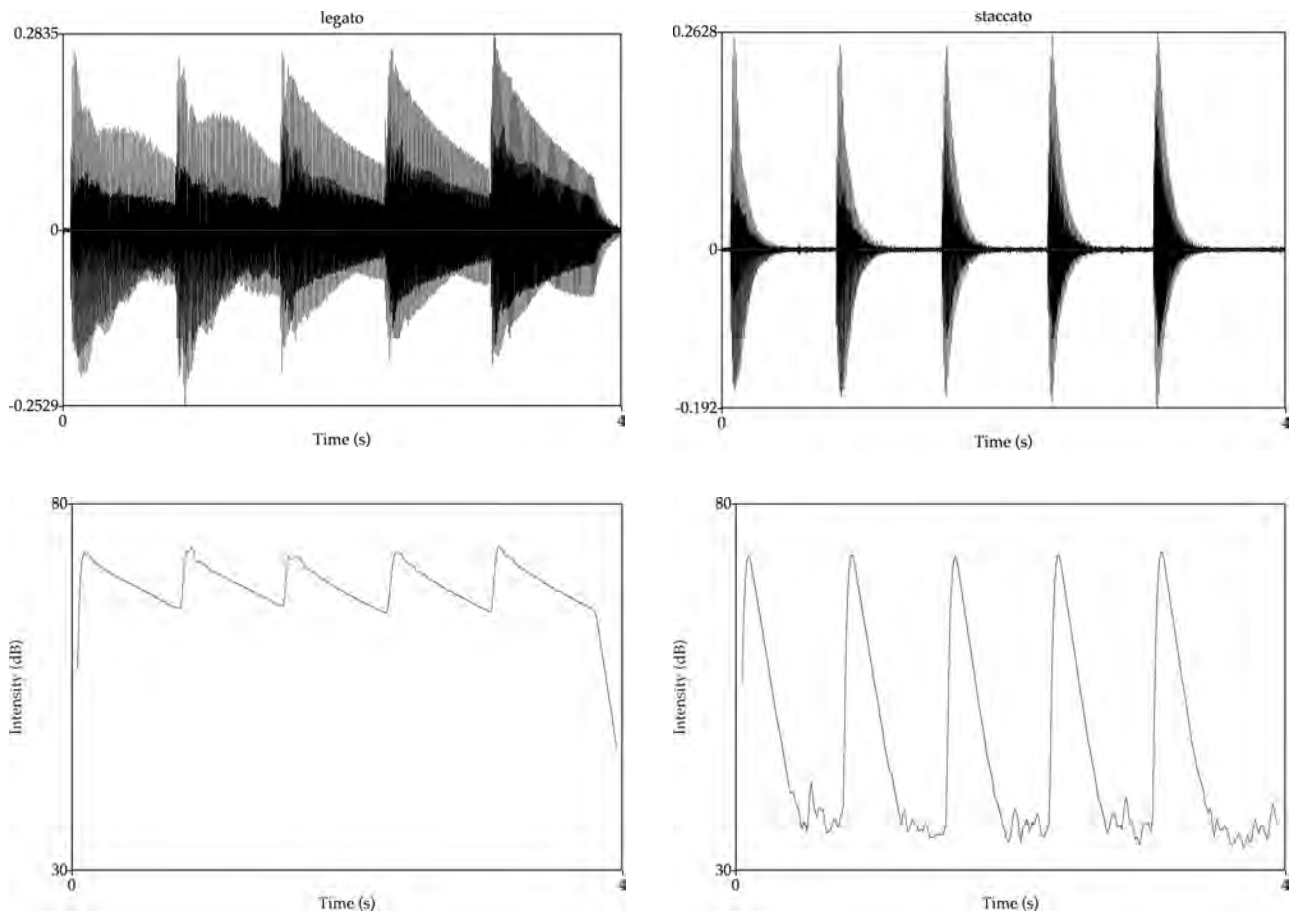


FIGURE 1. Examples of piano tones in Experiment 1: Waveforms (upper panels) and intensity contours (lower panels) of a legato sequence (left-hand panels) and a staccato sequence (right-hand panels) with interonset intervals of 750 ms. The x-axis ranges from 0 to 4 s, and the y-axis in the lower panels ranges from 30 to 80 dB. The y-axis of the upper panels is auto-scaled and different in the left and right panels.

sequence. (For a detailed analysis of the acoustic decay properties of tones from the same digital piano, see Repp, 1995.)

In each task, one block consisted of 14 trials (7 IOI durations \times 2 articulations for S1) presented in random order, and a total of 8 blocks was presented. A program written in MAX/MSP controlled the experiment. The computer (Macintosh Intel) was connected to the digital piano via a MOTU Fastlane-USB MIDI translator. Participants listened to the sequences over Sennheiser HD540 reference II earphones at a comfortable level and, in the reproduction task, tapped with their index finger on a Roland SPD-6 percussion pad held on their lap. Feedback tones from the taps were subject to electronic processing delays of approximately 30 ms for artificial tones and 15 ms for piano tones. These delays were not noticeable, and if they affected tapping tempo these effects were independent of the *legato-staccato* manipulation in S1.

PROCEDURE

The experiment with artificial tones was conducted about six months after the experiment with piano tones. In each experiment, the two tasks were completed in a single session lasting about one hour. Approximately half the participants started with the perception task; the others, with the reproduction task.

In the perception task, participants were asked to judge the tempo of S2 to be slower, the same, or faster relative to the tempo of S1 by clicking one of three response buttons displayed on the computer screen. The fact that S2 always had the same tempo from trial to trial was not mentioned. Each trial started 2 s after the previous response. Participants were able to listen to a trial again if their attention had wandered. Each block had a new randomization provided by the computer.

In the reproduction task, participants listened to S1 and then attempted to reproduce its tempo by tapping on the electronic percussion pad. Participants were asked

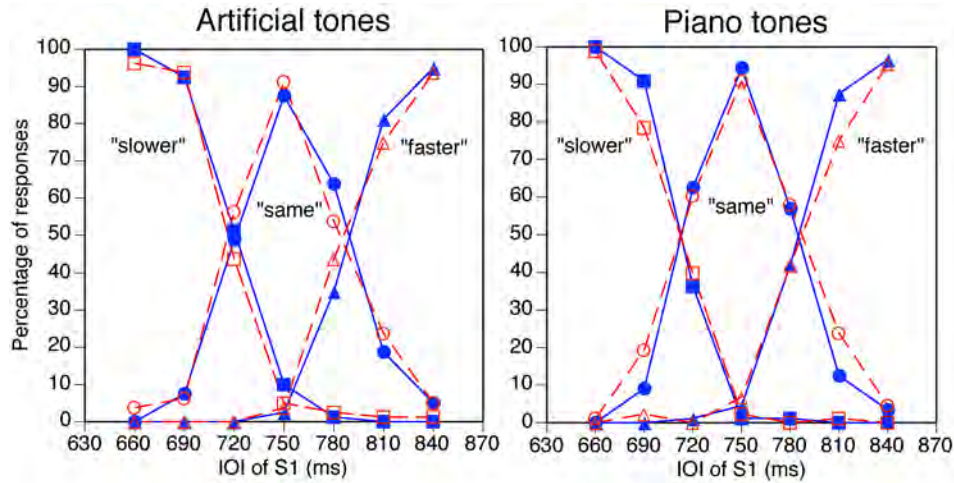


FIGURE 2. Results of the perception task in Experiment 1: Mean percentages of "slower," "same," and "faster" judgments of the second sequence (S2) relative to the first sequence (S1), when S1 was played either staccato (solid lines) or legato (dashed lines). IOI = interonset interval. The IOI of S2 was always 750 ms.

to skip one beat (i.e., to create an IOI of twice the duration of the IOI of the first sequence) before starting to tap. Participants could repeat a trial, if necessary. Each trial started 2 s after the space bar of the computer keyboard was pressed.

Results

The results of the perception task are shown in Figure 2. Mean response percentages are plotted as a function of the IOI of S1. In neither version of the experiment was there any effect of *staccato* versus *legato*. If *legato* sequences had been perceived as slower than *staccato* sequences, the whole response distribution (dashed

lines) should have been shifted to the left, indicating a relative decrease in "slower" responses, an increase in "faster" responses, and a shift in the peak of "same" responses to an IOI shorter than 750 ms. There was no indication of any such shift in the data, and the results for the two types of tone were similar. Not surprisingly, *t*-tests on mean "faster" response percentages (with "same" responses counting as half a response) revealed no difference between *staccato* and *legato* conditions: artificial tones, 47.1 vs. 48.0, $t(9) = -0.52$, $p = .622$; piano tones, 50.2 vs. 50.1, $t(10) = -0.05$, $p = .964$.

The results of the reproduction task are shown in Figure 3. The mean reproduction IOI (the mean of the four intertap intervals, averaged across repetitions and

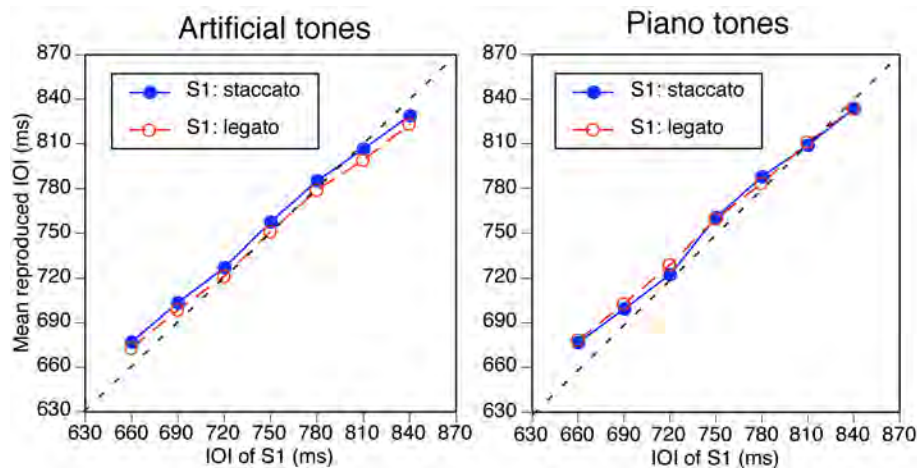


FIGURE 3. Results of the reproduction task in Experiment 1: Mean reproduced interonset interval (IOI) duration as a function of target (S1) IOI duration. The dashed diagonal line indicates perfect performance.

participants) is shown as a function of the IOI of S1, for S1 played *staccato* or *legato*. It can be seen that reproduction was very accurate, with only a slight tendency to tap too slow when S1 was relatively fast (IOI of 660 ms). If *legato* sequences had seemed slower than *staccato* sequences, the dashed line should have been higher than the solid line. However, there was no such difference between the *staccato* and *legato* conditions with either type of tone. In fact, there was a tendency toward a difference in the opposite direction with artificial tones, mean reproduction interval 755.1 vs. 749.3 ms, $t(10) = -1.95$, $p = .083$, whereas there was no hint of a difference with piano tones, 755.6 vs. 756.8 ms, $t(9) = 0.27$, $p = .793$.

Discussion

The results of Experiment 1 indicate that a *legato* sequence is not perceived as slower than a *staccato* sequence. If this negative result had been obtained only with piano tones, it could have been argued that their damped decay, by filling part of the IOI in *staccato* sequences, attenuated the effect, or that the sustained portion of *legato* tones was not effective as filler because it decayed. Therefore, it was important to replicate the results with artificial tones that clearly separated sound from silence. Of course, the stimuli represented only a small selection from a wide range of possible pitches, intensities, and IOI durations, and it is possible that an SSI would be obtained with different kinds of sequences. However, the present sequences were fairly representative of music with regard to intensity (comfortable) and tempo (somewhat slower than a typical beat; Parncutt, 1994). Only the pitch of the tones was rather low compared to typical melody tones. However, Burghardt (1973) found that the SSI was about as large with 200-Hz as with 1000-Hz pure tones, and the spectrum of the present complex tones extended up to about 460 Hz in artificial tones and to much higher frequencies in piano tones. Therefore, the low fundamental frequency of our tones was probably not a limiting factor.

The most likely limiting factor, if any, was IOI duration. Burghardt (1973) found that the SSI decreased with increasing interval duration and disappeared around 1 s. It is difficult to tell from his data whether there was still a significant illusion in the range of the present IOIs (around 750 ms), though it seems so. Craig (1973), however, found a huge illusion that was constant across a wide range of IOIs (200–1200 ms). The present findings are certainly inconsistent with Craig's results, as are Burghardt's. Also, Wearden et al. (2007) found a robust illusion with single intervals in the range of 400 to 600 ms.

Psychophysical studies typically concern the judgment of single intervals by participants who do not have extensive music training. Our use of multiple-interval rhythmic sequences and musician participants certainly could have reduced the magnitude of the SSI, but it should not have eliminated it completely. Note that a small but quite reliable DTI was obtained in the same tasks with similar materials (Repp, 2008; Repp & Bruttomesso, 2009). The absence of an SSI in the present experiment is surprising from the perspective of the pacemaker acceleration hypothesis. If an internal pacemaker is stimulated by sustained sound input, such stimulation should occur regardless of whether a single filled interval or a sequence of filled intervals is presented. If anything, the effect should be larger with multiple intervals because the pacemaker is being stimulated continuously. There is also no reason why pacemaker acceleration should not occur in musicians, given that it is a rather low-level psychophysiological effect. Therefore, our results cast some doubt on the pacemaker acceleration hypothesis as an explanation of the SSI generally, not just in the present context.

However, our negative results are still compatible with the delayed offset perception hypothesis, which was not addressed in Experiment 1. Because we presented filled intervals in sequence, the offset of each interval was marked by the onset of the next (i.e., by a change in pitch and, in piano tones, also by a change in amplitude). Therefore, any effect of delayed offset perception could not manifest itself, except in the final tone of the sequence, which probably had little effect on the perceived sequence tempo. A different paradigm is required to address the delayed offset perception hypothesis, which obviously applies only when tone offsets are not immediately followed by tone onsets. Craig (1973) asked participants to compare the duration of a tone with that of a following silence that was terminated by a second tone onset. Such judgments are not particularly relevant to music. Instead, we used in Experiment 2 a method of direct judgment of relative offset timing that, as far as we know, has not been used previously and that, unlike the direct comparison of sound and silence durations, has some musical relevance, for rhythm perception in particular.

Experiment 2

In this experiment we presented isochronous sequences of identical tones whose articulations (i.e., durations) varied from *staccato* (fairly short) to *portato* (fairly long but ending before the next tone onset). The participants' task was to judge when the tone offsets occurred relative to the midpoints of the IOIs between tone onsets. Another way of describing this task is that participants

were to judge whether the rhythm composed of alternating tone onsets and offsets was isochronous or not, and if not, whether the offsets occurred early or late. Although tone offsets do not normally function as rhythmic elements, it seemed to us that they could be heard as such if a mental effort were made to integrate them into the overall rhythm. The prediction was that tone offsets, if they are associated with longer perceptual latencies than tone onsets, would be judged as occurring late relative to the IOI midpoints when they were in fact on time, and would have to occur before the IOI midpoints in order to be judged as occurring at the midpoints.

Indeed, if the constant 657 ms difference between subjective sound and silence durations found by Craig (1973) were taken literally as indicating a long delay in offset perception, participants should not be able to judge offset timing at all if the IOI were shorter than 700 ms or so, because then the offset would be perceived as occurring after the next tone onset. Craig himself found this literal interpretation implausible, but as a precaution we employed three different IOI durations in Experiment 2, including two that are substantially longer than 700 ms.

We again used two types of tone, artificial tones with abrupt offsets and piano tones that decay gradually when damped. We wondered whether the offset of piano tones would be perceived to occur much later than that of artificial tones, and whether it would be more difficult to locate in time.

We have not found much previous literature relating to the question asked here. Vos and Ellermann (1989) investigated participants' ability to reproduce exactly sequences of electronic organ sounds, which have abrupt onsets and offsets. IOI and tone duration were varied. Reproduction by means of key presses was quite accurate, though more so for tone durations than for silence durations. However, these results provide no information about the question we are interested in because any perceptual illusions would have affected perception of both the models and their reproductions. More relevant is a study by Warm and Foulke (1970), who investigated whether reaction time (RT) to tone onset or offset is affected by rise or decay time, respectively. They found that RT to tone offset increased linearly with decay time, whereas RT to tone onset increased but little with rise time. However, they did not specify the shape of their rise and decay functions, which may have been linear rather than exponential. In any case, their findings suggest that the offset of a decaying sound may be perceived to occur much later than that of a sound that has an abrupt offset. By contrast, studies comparing RTs to abrupt tone onsets and offsets have not found any difference (Woodrow, 1915), which

seems inconsistent with the delayed offset perception hypothesis.

Method

PARTICIPANTS

The participants were the same as in Experiment 1 (including author RJM), except for one graduate student (a violinist) who was not available at the time.

MATERIALS AND EQUIPMENT

The equipment was the same as in Experiment 1. Artificial tones and piano tones were also generated in the same ways. All tones had a constant pitch of C2 (65 Hz, MIDI pitch 36). The piano tones decayed rapidly after their nominal offset, which is the time of key release signaled to the digital piano by a MIDI instruction from the computer. Figure 4 illustrates the different waveforms and intensity envelopes of the two types of tone, in particular the gradual offsets of the piano tones and the abrupt offsets of the artificial tones. Spectrograms indicated that some spectral splatter occurred at the offset of artificial tones, despite a 5-ms amplitude ramp. This may have increased the perceptual salience of the offset.

Each trial consisted of an isochronous sequence of five identical tones with IOI durations of 600, 1000, or 1400 ms. For each IOI duration, there were seven possible tone durations. The nominal tone offsets occurred at time points corresponding to 35% to 65% of the IOI, in steps of 5%. So, for example, for the IOI of 600 ms the nominal tone durations were 210, 240, 270, 300, 330, 360, and 390 ms. There were 21 randomized trials (3 IOI durations \times 7 tone durations) per block, and five consecutive blocks of trials were presented with each type of tone.

PROCEDURE

The experiment was run in a single one-hour session, about two months after the piano-tone condition of Experiment 1. Half the participants started with the piano tones, the other half with the artificial tones. Playback levels were adjusted to be comfortable and similar for the two types of tone. Instructions on the screen asked participants to judge whether the tone offsets were early, on time, or late relative to the midpoints of the intervals between tone onsets. Participants chose an answer and then clicked another button to proceed to the next trial. They could listen to a trial repeatedly if they wished.

Results

Figure 5 shows the percentages of "early," "on time," and "late" responses in separate panels for the three IOI duration conditions, and for the two types of tone in

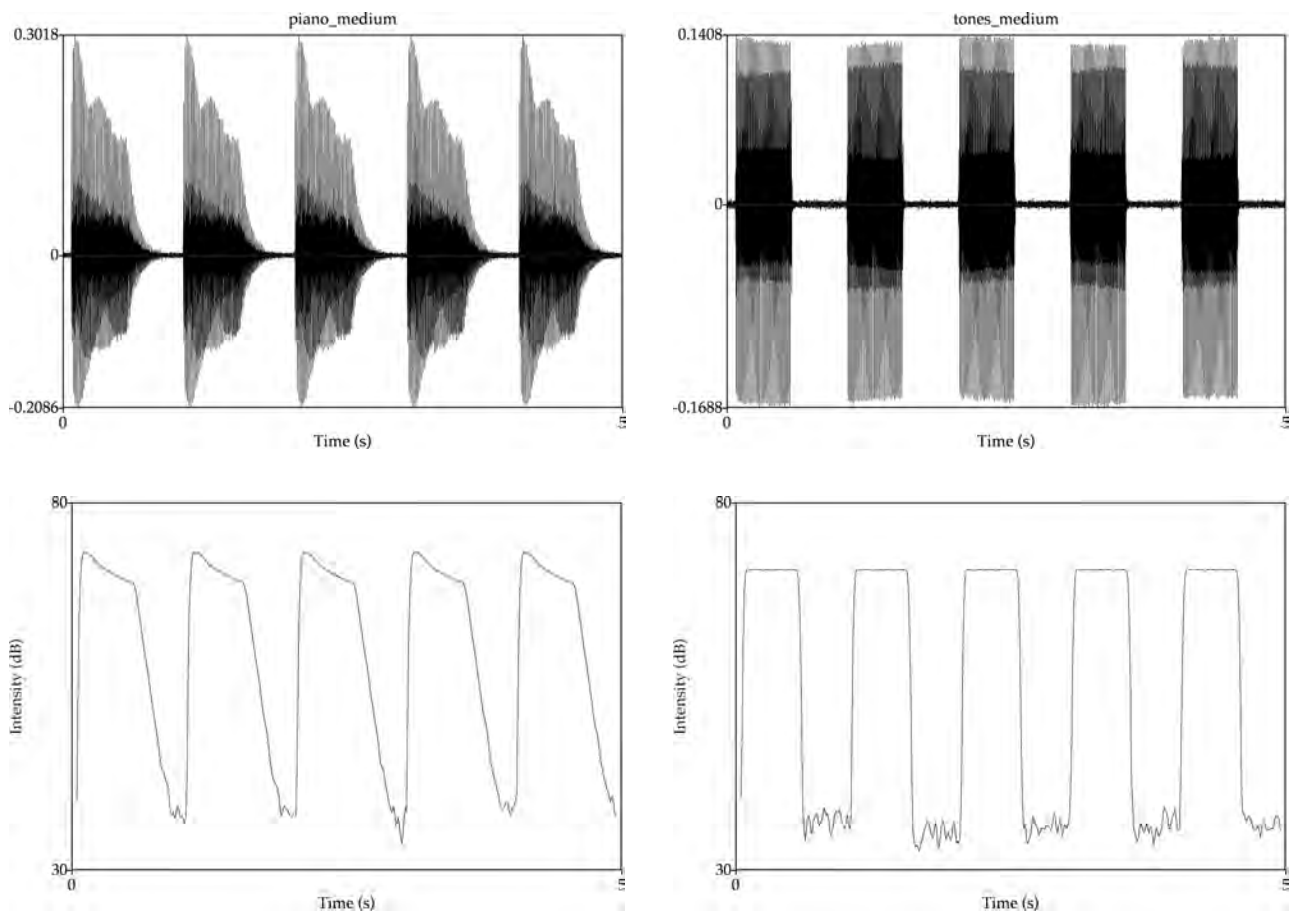


FIGURE 4. Examples of materials in Experiment 2: Waveforms (upper panels) and intensity contours (lower panels) of a sequence composed of piano tones (left-hand panels) and a sequence composed of artificial tones (right-hand panels), with IOIs of 1000 ms and nominal tone durations of 500 ms. The x-axis ranges from 0 to 5 s, and the y-axis in the lower panels ranges from 30 to 80 dB. The y-axis of the upper panels is auto-scaled and different in the left and right panels.

each panel. It can be seen that “on time” responses were generally centered at the IOI midpoint, or at least not far off. The results for the two types of tone look similar. The response functions at the short IOI (600 ms) are broader and seem more variable than those at the long IOI (1400 ms). This indicates that the judgments became easier to make at longer IOI durations. However, this could easily be due to the fact that the differences of the offset times from the IOI midpoint grew larger as IOI duration increased. The x-axes of the three panels in Figure 5 represent the same relative scale (percentage of IOI duration), but different absolute scales.

For easier comparison, the response percentages were transformed into “late” scores by adding half the percentage of “on-time” responses to the percentage of “late” responses. This yielded response functions that increased monotonically as a function of tone offset time. In Figure 6, “late” scores are plotted on a common absolute time scale. It is evident that the slopes of the functions are very

similar when considered in terms of absolute temporal differences. Moreover, their 50% intercepts are very close to the IOI midpoint in each case.

A 2 (types of tone) \times 3 (IOI durations) repeated-measures ANOVA on “late” scores averaged across tone durations revealed a significant main effect of type of tone, $F(1, 9) = 12.36$, $p = .007$, with higher scores for piano tones. The main effect of IOI duration and the interaction were not significant. To quantify the difference between the two types of tone, the 50% crossovers of the average response functions (Figure 6) were estimated by linear regression. At the 1000 ms and 1400 ms IOIs, the crossover for artificial tones was almost exactly at the IOI midpoint, at 502 ms and 699 ms, respectively. At the 600 ms IOI, it was slightly after the midpoint, at 314 ms. At all three IOIs, the crossover for piano tones was earlier than that for artificial tones, the difference being 11 ms on average. At the 600 ms IOI, the crossover for piano tones was exactly at the IOI midpoint.

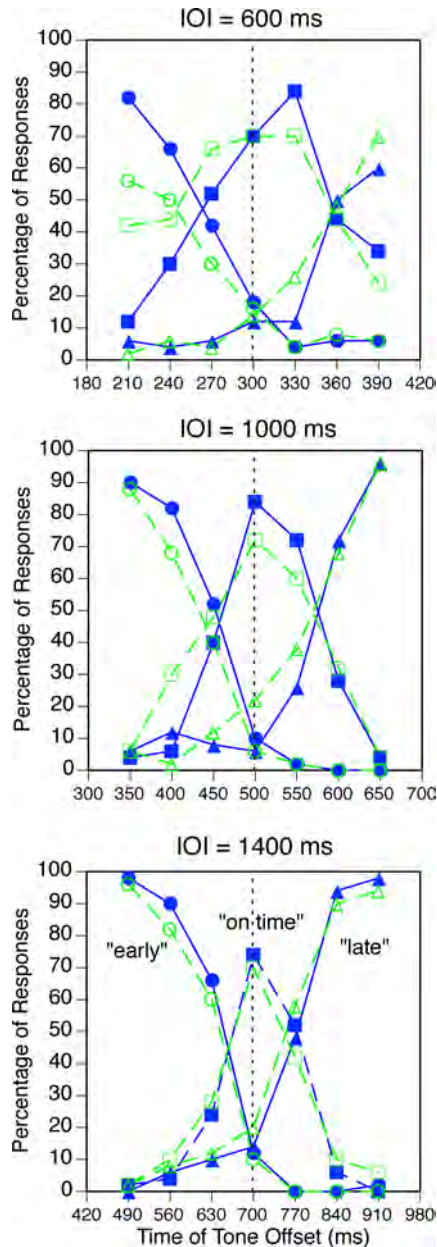


FIGURE 5. Results of Experiment 2: Percentages of "early," "on time," and "late" responses as a function of nominal tone offset time for artificial tones (solid lines) and piano tones (dashed lines) in the three IOI duration conditions. IOI midpoints are marked by vertical dotted lines.

To compare the slopes of the response functions, we calculated them for each participant's "late" scores in each condition by using linear regression on all seven data points at IOI = 600 ms, the central five data points at IOI = 1000 ms, and the central three data points at IOI = 1400 ms, thereby restricting analysis to the linear portion of each response function. A 2 (types of tone) \times 3 (IOI durations) repeated-measures ANOVA yielded no

significant effects; in particular, the main effect of IOI duration was far from significance, $F(2, 18) = 0.46$, $p = .565$ (Greenhouse-Geisser corrected).

Discussion

The results of Experiment 2 are interesting in several ways. First, they provide no evidence whatsoever for a delay in the perception of abrupt tone offsets relative to the perception of abrupt tone onsets. The timing of the offsets of artificial tones was judged quite accurately; offsets did not have to occur before the IOI midpoint in order to be perceived as occurring at the IOI midpoint. This result refutes the hypothesis that sound offsets take longer to detect than sound onsets.

Second, the results for piano tones were surprisingly similar to those for artificial tones. Although the nominal offset of piano tones did have to occur slightly earlier than the offset of artificial tones to be judged as occurring at the same time, the difference was remarkably small in view of the gradual offset of the piano tones. As described in Experiment 1, the damped decay rate was about 10 dB per 100 ms. The mean difference of 11 ms in the response function crossovers for artificial tones and piano tones implies that the offset of piano tones was perceived after they had decayed by only about 1 dB at the damped decay rate. This basically means that the offset of a piano tone was perceived as soon as the change from the slow decay rate of the sustained tone to the fast damped decay rate was detected. In other words, the results suggest that the offset of a piano tone is perceived very soon after the damper touches the strings (the nominal offset). The decaying energy following the nominal offset seems to have had no perceptual impact. It is possible that this finding is specific to low-pitched piano tones. With high-pitched piano tones, whose sustained decay rate is faster, the change in decay rate may be harder to detect, resulting in a longer delay of offset detection. However, because the damped decay rate, too, is faster at high pitches (Repp, 1995), there may not be much difference in the end.

Third, the similarity of the slopes of the response functions when plotted on the same absolute temporal scale (Figure 6) indicates that judgments of offset timing do not follow Weber's law. If they did, the slope of the response function should get shallower as the IOI gets longer. It appears that the offsets were not judged relative to the interval within which they occurred (i.e., they were not judged in terms of relative phase) but as absolute deviations from the interval midpoints. Since interval midpoints were not marked in any way, participants must have generated them internally. Oscillator models

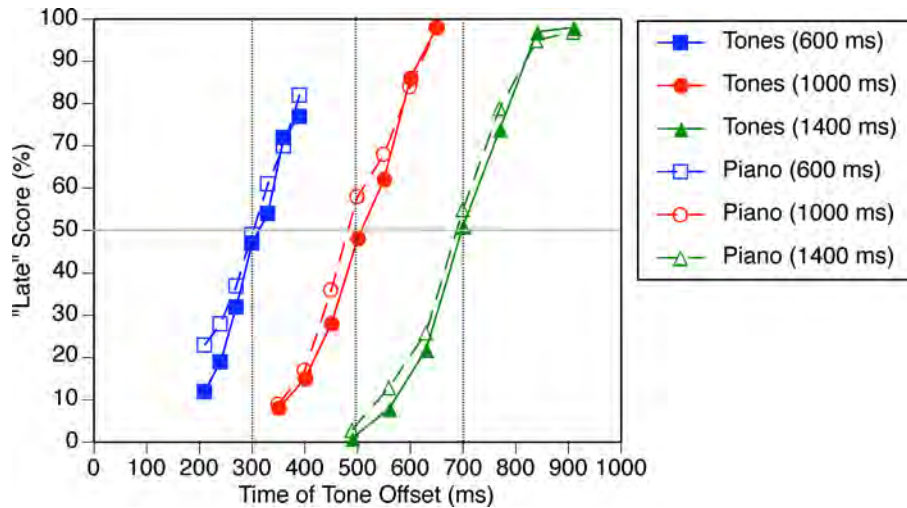


FIGURE 6. Results of Experiment 2 plotted on a common abscissa for artificial tones ("Tones") and piano tones ("Piano"). The "late" score represents the percentage of "late" responses plus half the percentage of "on time" responses. IOI midpoints are marked by vertical dotted lines.

of rhythmic entrainment (e.g., Large & Kolen, 1994; Tomic & Janata, 2008) predict that a periodic pacing frequency also entrains higher harmonics, and the second harmonic is the one that bisects the main period. Judging offset timing against these bisection points would constitute a simultaneity/successiveness judgment task, and therefore would not depend on IOI duration. Nevertheless, it is surprising that IOI duration had no effect because variability of internal oscillators, and hence uncertainty about the IOI midpoint, should increase with interval (period) duration, due to accumulation of biological noise.

General Discussion

The purpose of the present study was to determine whether an SSI occurs in rhythmic sequences. Experiment 1 used the methods with which Repp (2008) and Repp and Bruttomesso (2009) had demonstrated a DTI in similar sequences. However, unlike discrete subdivision, sustained sound did not make intervals seem longer: *Legato* sequences were not perceived as slower than *staccato* sequences. Consistent with this result, we are also unaware of any data or anecdotal reports suggesting that musicians exhibit a tendency to play faster when they play *legato* than when they play *staccato* at the same intended tempo. Such a compensatory tendency would be expected if an SSI occurred in music. By contrast, musicians readily acknowledge a (usually to-be-resisted) tendency to speed up when the note density in music increases, and Repp and Bruttomesso documented this tendency in the laboratory and by analyzing commercially

recorded performances. *Legato* playing is associated with fast tempo only in that *staccato* articulation becomes difficult at fast tempi. Thus, a fast tempo encourages *legato* playing, but *legato* articulation does not encourage a fast tempo.

Experiment 1 was intended to address the hypothesis that the SSI is due to acceleration of an internal pacemaker or clock. Since this acceleration presumably occurs at an early level in auditory processing, it is difficult to see why it should not have occurred also in our materials and with musically trained participants. Therefore, our results cast some doubt on this explanation of the SSI obtained in psychophysical studies (Penney et al., 2000; Wearden et al., 2006, 2007). There are a number of methodological differences between those studies and the present one that, it could be argued, might account for the difference in results. First, earlier studies concerned single intervals, whereas our experiment used sequences of four intervals. This is unlikely to explain the difference in results because the SSI has been shown to increase with interval duration (Wearden et al., 2007), which suggests that the hypothetical pacemaker acceleration is not transitory. We are quite confident that, had we conducted Experiment 1 with sequences of two instead of five tones (i.e., with single intervals), we would have obtained the same negative results, although probably with increased variability. Second, earlier studies often used verbal judgments of interval duration, whereas we used reproduction and direct comparison. Wearden et al. (2007), however, used a direct comparison method in their first experiment (albeit with a remembered standard) and found a large SSI. Third, unfilled

intervals, if they were delimited by short sounds, were typically delimited by clicks in earlier studies. It is possible that the 40-ms complex tones used in our *staccato* sequences were sufficient to accelerate the internal pacemaker as much as did sustained tones, but again this seems unlikely in view of the previously demonstrated increase in the SSI with interval duration. Fourth, it could be argued that our presentation of standard and comparison stimuli or reproductions within the same metrical framework (i.e., separated by an IOI that was exactly or approximately twice the IOI in the first sequence) eliminated the illusion. However, here again we need to cite Repp (2008) and Repp and Bruttomesso (2009), who found a reliable DTI with very similar sequences. At the very least, then, the SSI seems less robust than the DTI and dependent on temporal unpredictability. Judgments of single intervals, presentation of clicks, and random interstimulus intervals are not characteristic of musical contexts, to which the present research is intended to be relevant.

Experiment 2 was intended to test an alternative explanation of the SSI, namely that it is due to slower perceptual registration of sound offsets than sound onsets. However, the results showed that abrupt tone offsets were perceived almost exactly at the time at which they occurred relative to tone onsets, which implies that neural processing delays for onsets and offsets were of the same magnitude. Perception of gradual tone offsets was delayed only slightly, the delay probably reflecting the time it took to detect a sudden acceleration of the decay rate. These results argue strongly against models that predict longer perceptual delays for offset detection than for onset detection (Burghardt, 1972, 1973; Grondin, 1993, 2001, 2003).

The results of Experiment 2 do not contradict previous empirical findings showing that, under certain conditions, an interval filled with sound is perceived as longer than a silent interval. Our participants were not asked to compare filled and silent interval durations directly; rather, they were asked to judge the relative temporal location of sound offsets between sound onsets. What our results demonstrate is that a rhythm formed by alternating onset and offset events can be perceived accurately even though the relative durations of the intervals they delimit might be misestimated. This is consistent with recent evidence that perceived interval duration is not simply the difference between perceived offset and onset times (Bendixen, Grimm, & Schröger, 2006). Intervals are second-order temporal properties of time series of events, and their perception and judgment may be subject to influences of variables that play no role in the perception of time points. Even so, Craig's (1973)

dramatic results seem rather unusual and await replication. Evidently, his participants did not realize that the relative durations of contiguous sound and silence could be judged by gauging the relative temporal location of the first sound offset between the two sound onsets.

Even though tone offsets normally do not function as components of a rhythm (Vos & Ellerman, 1989), some researchers have assumed that regularly timed tone offsets can contribute to rhythmic entrainment. For example, Blaschke, Hass, Herrmann, and Rammsayer (2008) predicted (and found) better detection of changes in the duration of one IOI in tone sequences when the tone offsets occurred consistently at the IOI midpoint than when they did not. Our results lend some support to their assumption of entrainment to offsets by showing that tone offsets that occur at the IOI midpoint are indeed perceived as occurring at that time, not later. It thus seems possible that rhythmic tone offset timing exerts subtle effects on action timing and temporal judgment. This warrants further investigation.

If the SSI in psychophysics is due neither to pacemaker acceleration nor to delayed offset perception, what might be its cause? One possibility is that auditory foreground/background relationships play a role. Just as viewers cannot see two faces and a vase simultaneously in the familiar ambiguous visual figure, so sound and silence may be mutually exclusive objects of temporal perception. To perceive the duration of silence, the silence has to be foregrounded intentionally, and this may take considerable switching time if a preceding sound has just been focused on, as in Craig's (1973) experiments. When the silence does not have to be compared to the immediately preceding sound, as in Burghardt's (1972, 1973) studies, perceptual foregrounding of silence may be more successful. By postulating different thresholds for sound offset and silence onset detection, Burghardt effectively suggested that sound offsets may not be able to do double duty as silence onsets, just as a visual contour cannot simultaneously outline a face and a vase. Ultimately, however, the difference between sound and silence may be qualitative in nature (Meumann, 1896), and this non-temporal difference, rather than any difference in actual perceived duration, may bias perceptual judgment and account for the SSI in psychophysical experiments.

In conclusion, the present results suggest that, unlike the DTI, the SSI does not occur in rhythmic contexts. Thus, these two types of illusion probably reflect different underlying mechanisms and, in the case of the SSI, perhaps not those proposed in the literature. However, further research is certainly needed to close the methodological gap between the present experiments and previous psychophysical studies, to explore a wider range of

stimulus parameters, and to investigate the role of music training. The main reason why no SSI was obtained in the present research may be that no direct judgments of sound and silence duration were required. Rather, and consistent with musical practice, participants judged tempo (Experiment 1) and rhythmic timing (Experiment 2). These judgments could well have been affected by perceived interval duration, but they were not. This may mean that interval perception, when not explicitly judged, is not subject to the SSI. Instead, the SSI may occur only in explicit comparisons of sound and silence duration, in which cognitive factors also play a role.

Author Note

This research was supported by National Science Foundation Grant BCS-0642506. BHR collected the data, performed statistical tests, and prepared the manuscript for publication. RJM helped design the experiments, analyzed the data, and wrote a term paper about this research during her sophomore year at Yale.

Correspondence concerning this article should be addressed to Bruno H. Repp, Haskins Laboratories, 300 George Street, New Haven, CT 06511-6624. E-MAIL: repp@haskins.yale.edu

References

- BEHAR, I., & BEVAN, W. (1961). The perceived duration of auditory and visual intervals: Cross modal comparison and interaction. *American Journal of Psychology*, *74*, 17–26.
- BENDIXEN, A., GRIMM, S., & SCHRÖGER, E. (2006). The relation between onset, offset, and duration perception as examined by psychophysical data and event-related brain potentials. *Journal of Psychophysiology*, *20*, 40–51.
- BLASCHKE, S., HASS, J., HERRMANN, J. M., & RAMMSAYER, T. (2008, March). *Pacemaker-counter vs. oscillator-based timing mechanisms: New evidence for a pacemaker-counter system from sequence tasks with filled intervals*. Paper presented at the Eighteenth Annual New England Sequence and Timing Meeting, New Haven, CT.
- BOERSMA, P., & WEENINK, D. (2007). *Praat: Doing phonetics by computer* (Version 4.6.24) [Computer program]. Retrieved September 24, 2007 from <http://www.fon.hum.uva.nl/praat/>
- BUFFARDI, L. (1971). Factors affecting the filled-duration illusion in the auditory, tactual, and visual modalities. *Perception and Psychophysics*, *10*, 292–294.
- BURGHARDT, H. (1972). Einfaches Funktionsschema zur Beschreibung der subjektiven Dauer von Schallimpulsen und Schallpausen [A simple model for the description of the subjective duration of sound impulses and pauses]. *Kybernetik*, *12*, 21–29.
- BURGHARDT, H. (1973). Über die subjektive Dauer von Schallimpulsen und Schallpausen [On the subjective duration of sound impulses and pauses]. *Acustica*, *28*, 284–290.
- CRAIG, J. C. (1973). A constant error in the perception of brief temporal intervals. *Perception and Psychophysics*, *13*, 99–104.
- FASTL, H., & ZWICKER, E. (2007). *Psychoacoustics: Facts and models*. Berlin: Springer-Verlag.
- GOLDSTONE, S., BOARDMAN, W. K., & LHAMON, W. T. (1959). Intersensory comparisons of temporal judgments. *Journal of Experimental Psychology*, *57*, 243–248.
- GOLDSTONE, S., & GOLDFARB, J. L. (1963). Judgment of filled and unfilled durations: Intersensory factors. *Perceptual and Motor Skills*, *17*, 763–774.
- GOLDSTONE, S., & GOLDFARB, J. L. (1964a). Auditory and visual time judgment. *Journal of General Psychology*, *70*, 369–387.
- GOLDSTONE, S., & GOLDFARB, J. L. (1964b). Direct comparison of auditory and visual durations. *Journal of Experimental Psychology*, *67*, 483–485.
- GOLDSTONE, S., & LHAMON, W. T. (1974). Studies of auditory–visual differences in human time judgment. 1. Sounds are judged longer than lights. *Perceptual and Motor Skills*, *39*, 63–82.
- GRONDIN, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Perception and Psychophysics*, *54*, 383–394.
- GRONDIN, S. (2001). From physical time to the first and second moments of psychological time. *Psychological Bulletin*, *127*, 22–44.
- GRONDIN, S. (2003). Sensory modalities and temporal processing. In H. Helfrich (Ed.), *Time and mind II: Information processing perspectives* (pp. 75–92). Göttingen: Hogrefe & Huber.
- HALL, G. S., & JASTROW, J. (1886). Studies of rhythm. *Mind*, *11*, 55–62.
- LARGE, E. W., & KOLEN, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, *6*, 177–208.
- MARTIN, D. W. (1947). Decay rates of piano tones. *Journal of the Acoustical Society of America*, *19*, 535–541.
- MEUMANN, E. (1896). Beiträge zur Psychologie des Zeitbewußtseins: Dritte Abhandlung [Contributions to the psychology of time consciousness: Third essay]. *Philosophische Studien*, *12*, 128–254.
- NAKAJIMA, Y. (1987). A model of empty duration perception. *Perception*, *16*, 485–520.

- ORNSTEIN, R. E. (1969). *On the experience of time*. Harmondsworth, UK: Penguin.
- PARNCUTT, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, 11, 409–464.
- PENNEY, T. B., GIBBON, J., & MECK, W. H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1770–1787.
- REPP, B. H. (1995). Acoustics, perception, and production of *legato* articulation on a digital piano. *Journal of the Acoustical Society of America*, 97, 3862–3874.
- REPP, B. H. (2008). Metrical subdivision results in subjective slowing of the beat. *Music Perception*, 26, 19–39.
- REPP, B. H., & BRUTTOMESSO, M. (2009). A filled duration illusion in music: Effects of metrical subdivision on the perception and production of beat tempo. *Advances in Cognitive Psychology*, 5, 114–134.
- ROELOFS, C. O., & ZEEMAN, W. (1951). Influence of different sequences of optical stimuli on the estimation of duration of a given interval of time. *Acta Psychologica*, 8, 89–128.
- TEN HOOPEN, G., MIYAUCHI, R., & NAKAJIMA, Y. (2008). Time-based illusions in the auditory mode. In S. Grondin (Ed.), *Psychology of time* (pp. 139–187). Bingley, UK: Emerald.
- THOMAS, E. A. C., & BROWN, I., JR. (1974). Time perception and the filled-duration illusion. *Perception and Psychophysics*, 16, 449–458.
- TOMIC, S. T., & JANATA, P. (2008). Beyond the beat: Modeling metric structure in music and performance. *Journal of the Acoustical Society of America*, 124, 4024–4041.
- VOS, P. G., & ELLERMANN, H. H. (1989). Precision and accuracy in the reproduction of simple tone sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 179–187.
- WALKER, J. T., & SCOTT, K. J. (1981). Auditory–visual conflicts in the perceived duration of lights, tones, and gaps. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 1327–1339.
- WARM, J. S., & FOULKE, E. (1970). Effects of rate of signal rise and decay on reaction time to the onset and offset of acoustic stimuli. *Perception and Psychophysics*, 7, 159–160.
- WEARDEN, J. H., EDWARDS, H., FAKHRI, M., & PERCIVAL, A. (1998). Why “sounds are judged longer than lights”: Application of a model of the internal clock in humans. *Quarterly Journal of Experimental Psychology*, 51B, 97–120.
- WEARDEN, J. H., NORTON, R., MARTIN, S., & MONTFORD-BEBB, O. (2007). Internal clock processes and the filled-duration illusion. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 716–729.
- WEARDEN, J. H., TODD, N. P. M., & JONES, L. A. (2006). When do auditory/visual differences in duration judgements occur? *Quarterly Journal of Experimental Psychology*, 59, 1709–1724.
- WOODROW, H. (1915). Reactions to the cessation of stimuli and their nervous mechanism. *Psychological Review*, 22, 423–452.
- ZWICKER, E. (1969/1970). Subjektive und objektive Dauer von Schallimpulsen und Schallpausen [Subjective and objective durations of sound impulses and pauses]. *Acustica*, 22, 214–218.

