Perceived empty duration between sounds of different lengths: Possible relation with repetition and rhythmic grouping

Tsuyoshi Kuroda¹ · Erika Tomimatsu² · Simon Grondin³ · Makoto Miyazaki¹

Published online: 5 July 2016
© The Psychonomic Society, Inc. 2016

Abstract We investigated how perceived duration of empty time intervals would be modulated by the length of sounds marking those intervals. Three sounds were successively presented in Experiment 1. Each sound was short (S) or long (L), and the temporal position of the middle sound's onset was varied. The lengthening of each sound resulted in delayed perception of the onset; thus, the middle sound's onset had to be presented earlier in the SLS than in the LSL sequence so that participants perceived the three sounds as presented at equal interonset intervals. In Experiment 2, a short sound and a long sound were alternated repeatedly, and the relative duration of the SL interval to the LS interval was varied. This repeated sequence was perceived as consisting of equal interonset intervals when the onsets of all sounds were aligned at physically equal intervals. If the same onset delay as in the preceding experiment had occurred, participants should have perceived equality between the interonset intervals in the repeated sequence when the SL interval was physically shortened relative to the LS interval. The effects of sound length seemed to be canceled out when the presentation of intervals was repeated. Finally, the perceived duration of the interonset intervals in the repeated sequence

was not influenced by whether the participant's native language was French or Japanese, or by how the repeated sequence was perceptually segmented into rhythmic groups.

Keywords Time perception · Empty duration · Marker length · Rhythmic grouping · Repetition

There is an increasing body of literature in the field of psychophysics regarding how the duration processing of empty intervals that are marked by successive signals is influenced by physical structure, such as the sensory modality and the size, of those signals (e.g., Grondin, 2010; Ono & Kitazawa, 2009; Xuan, Zhang, & Chen, 2007). Some of these studies focus on the length of sounds marking empty intervals, examining one of the following issues: (1) how the listener's ability for detecting a change in the duration of empty intervals, as expressed by the just-noticeable difference or the slope of psychometric functions, is modulated by the lengthening of each marker (Grondin, Roussel, Gamache, Roy, & Ouellet, 2005; Grose, Hall, & Buss, 2001; Kuroda, Hasuo, & Grondin, 2013; Penner, 1976; Rammsayer & Leutner, 1996), and (2) how the perceived duration of empty intervals, as expressed by a point of subjective equality, is modulated by the lengthening of each marker (Hasuo, Nakajima, Osawa, & Fujishima, 2012; Woodrow, 1928). The former issue is concerned with the sensitivity (or variability) of the timekeeping process, and the latter one with its accuracy. The present study focused on the accuracy issue; in other words, we investigated the effects of sound length on the perceived duration of empty intervals.

Woodrow (1928) and Hasuo et al. (2012) examined how the perceived duration of single empty intervals, marked by two successive sounds, is modulated by the lengthening of each sound. Woodrow fixed an interstimulus (offset to onset) interval between two sounds at 500 ms and found that this

Electronic supplementary material The online version of this article (doi:10.3758/s13414-016-1172-x) contains supplementary material, which is available to authorized users.

✉ Tsuyoshi Kuroda
tkuroda@inf.shizuoka.ac.jp

¹ Faculty of Informatics, Shizuoka University, 3-5-1 Johoku, Naka-ku, Hamamatsu, Shizuoka 432-8011, Japan

² Department of Human Science, Kyushu University, 4-9-1 Shiobaru, Minami-ku, Fukuoka 815-8540, Japan

³ École de psychologie, Université Laval, 2325 rue des Bibliothèques, Quebec, QC G1V 0A6, Canada

interval was perceived as longer when either of the two sounds was made longer. Hasuo et al. fixed an interonset (onset to onset) interval between two sounds at 120, 240, or 360 ms, and found that lengthening the second marker resulted in an overestimation of this interval. In brief, single empty intervals are perceived as longer when these are marked by longer sounds. This finding was also obtained by Grondin, Ivry, Franz, Perreault, and Metthé (1996) who used intermodal intervals (250, 500, and 750 ms).

Hasuo, Nakajima, and Hirose (2011) asked listeners to compare two neighboring intervals marked by three successive sounds' onsets (in their first experiment). This procedure may be regarded as a type of the AXB paradigm, as used in the studies relevant to the kappa effect (Alards-Tomalain, Leboe-McGowan, & Mondor, 2013; Henry & McAuley, 2009; ten Hoopen, Miyauchi, & Nakajima, 2008). Hasuo et al. (2011) indicated that an interval was perceived as longer when its terminating marker was lengthened. For example, when the middle sound was made longer, this sound's onset had to be presented earlier so that the first interval (between the initial and middle markers) and the second interval (between the middle and last markers) were perceived as equal. Similar results were reported by Kuroda, Hasuo, Labonté, Laflamme, and Grondin (2014) using two neighboring inter- and intramodal intervals.

However, no studies have yet investigated the effects of marker length on perceived duration when the number of presentation of intervals is further increased (i.e., is more than two). Let us suppose that a short sound (S) and a long sound (L) are alternated repeatedly (as SLSLSLSL... or LSLSLS...). It may be difficult to predict the perceived duration of empty intervals in this repeated sequence from the findings in the previously mentioned literature because this perceived duration may be influenced by not only the length of the two consecutive sounds but also another factor—rhythmic grouping. Bolton (1894) conducted a phenomenological experiment in which a sound was repeated at a constant interval. Listeners spontaneously perceived the sequence as segmented into rhythmic groups consisting of several sounds (subjective rhythmization). Moreover, some of the listeners reported that they perceived a longer interval between the rhythmic groups. In other words, between-groups intervals were perceived as longer than within-groups intervals. The similar effects of rhythmic grouping on perceived duration were demonstrated in a recent psychophysical experiment (Geiser & Gabrieli, 2013), but for only nonmusician listeners.

There is a widely accepted principle that, when successive sounds are of different lengths and are segmented into rhythmic groups, the resulting groups are likely to have longer sounds on their termination (Bolton, 1894). According to this principle, the previous sequence (SLSLSL...) would be segmented into groups consisting of a short sound followed by a long sound ([SL][SL][SL]...; SL grouping), instead of groups

consisting of a long sound followed by a short sound (S)[LS][LS][L...; LS grouping). If between-groups intervals are perceived as longer, the interval between the long and short sounds (LS interval) should be perceived as longer than the interval between the short and long sounds (SL interval). Thus, the perceived duration of empty intervals in the repeated sequence may be attributed to the rhythmic grouping that listeners perceive according to the relative length of the alternated sounds.

Additionally, previous studies have indicated that a change in the duration of between-groups intervals is more difficult to detect than a change in the duration of within-groups intervals (Fitzgibbons, Pollatsek, & Thomas, 1974; Geiser & Gabrieli, 2013; Thorpe & Trehub, 1989; Trainor & Adams, 2000). Some kinds of perceptual grouping could thus influence the ability for detecting a change in the duration of empty intervals (sensitivity) as well as the perceived duration of empty intervals (accuracy). As noted previously, this study focused on the accuracy issue and examined whether the effects of rhythmic grouping could interact with those of sound length on perceived duration.

There are two experiments in our study. Stimulus sequences consist of three successive sounds in Experiment 1, where we examine the effects of marker length on the perceived duration of the two neighboring intervals, as in Hasuo et al. (2011). There is no reason to assume that only three sounds are perceived as two beats leading to SL or LS grouping. However, either SL or LS grouping should occur when a short sound and a long sound are alternated repeatedly in Experiment 2. We estimate the temporal conditions enabling participants to perceive equality between the interonset intervals in this repeated sequence. Moreover, an additional session is conducted in Experiment 2 to establish whether individuals perceive SL or LS grouping for the repeated sequence.

Two previous studies measured both the individual tendency to perceive some types of perceptual grouping and the ability of detecting a change in the duration of temporal gaps, but failed to find a systematic relation between the results at these tasks (Kuroda et al., 2013; Neff, Jesteadt, & Brown, 1982). Neff et al. (1982) used a sequence in which a low-frequency sound and a high-frequency sound were alternated repeatedly, and examined whether auditory stream segregations according to frequency proximity would modulate the ability for detecting a change in the duration of temporal gaps between the two consecutive sounds of different frequencies. In addition to conducting a discrimination experiment, they conducted a stream-segregation experiment, measuring whether participants perceived the stimulus sequence as integrated into one stream or segregated into two streams. The results indicated no signs of relation between sequential grouping and gap discrimination. It therefore seemed necessary to establish the individual tendency of rhythmic (SL or

LS) grouping in our study, which aimed to examine the relation between perceived duration and rhythmic grouping. Indeed, there is a technical difference between our study and the previous studies that tested the rhythmic grouping of repeated sequences (Hay & Diehl, 2007; Iversen, Patel, & Ohgushi, 2008; Kuroda et al., 2013). When the length of markers was manipulated, physical interstimulus intervals were fixed in the previous studies, but physical interonset intervals are fixed in this study. In other words, interonset intervals were physically varied when the length of markers was manipulated in the previous studies, whereas these intervals are manipulated independently from the length of markers in this study. We thus examine whether the same rhythmic grouping as in the previous studies would occur in Experiment 2. Consequently, Experiment 2 consists of two sessions, one for measuring perceived duration and the other for measuring rhythmic grouping.

A previous study by Kuroda et al. (2013) partly addressed an issue similar to the one at the heart of our study, examining the effects of rhythmic (SL or LS) grouping on the listener's ability of detecting a change in the duration of temporal gaps in the repeated sequence. However, their discrimination experiment was designed to measure only how well listeners could detect a change in the gap duration, but not how long the gaps were perceived.¹ In other words, that previous study tested temporal sensitivity but not perceived duration. Consequently, our study is the first one testing the perceived duration of empty intervals in terms of the relation with the individual's tendency to perceive rhythmic grouping, measuring *both* performances in a single experiment (Experiment 2). Furthermore, when the presentation of gaps was not repeated in Kuroda et al. (2013), participants compared two gaps, each of which was marked by two sounds (i.e., Sound 1—gap—Sound 2 ... Sound 3—gap—Sound 4). However, in our Experiment 1, participants compare two neighboring intervals that are marked by three sounds (Sound 1—gap—Sound 2—gap—Sound 3). This structure may more often be found in daily situations where one listens to music and speech.

In summary, the purpose of this study was to investigate how the perceived duration of interonset intervals would be modulated by the length of sounds marking those intervals. The effects of marker length on perceived duration were also examined in terms of the relation with the repetition of

intervals and with the listener's tendency to perceive rhythmic (SL or LS) grouping.

Experiment 1

Three sounds were successively presented in Experiment 1. The temporal position of the middle sound's onset, as well as the length of each sound, was manipulated. Participants judged whether the middle sound's onset appeared too early (close to the initial sound's onset) or too late (close to the last sound's onset) to perceive the three sounds as presented at equal interonset intervals. They could also respond that the middle sound's onset appeared exactly halfway between the initial and last ones. From the psychometric function of the "too early" and that of the "too late" probability, we estimated a point of subjective equality (PSE), expressing the position of the middle sound's onset that made participants perceive the three sounds as presented at equal interonset intervals.

Method

Participants

Sixteen participants (12 females), self-reporting having normal hearing, were recruited. They were students and employees at Laval University, aged 19 to 37 years. They consented to their participation by signing a form approved by the ethics review board of this institution. One more participant was recruited, but the data from this participant were not kept for the analysis.²

Stimuli

Digital signals of stimuli were sampled at 44100 Hz and quantized to 16 bits. Each trial consisted of three sounds that were successively presented from headphones (Sennheiser HD 477) at about 66 dBA. These sounds were square-like waves generated by mixing the fundamental of 500 Hz and the first three odd sinusoidal components. The component's amplitude decreased in proportion with their harmonic number; for example, the third harmonic (1500 Hz) had one third of the fundamental's amplitude.

Each of the three sounds was either 150 (short—S) or 262.5 ms (long—L), while the initial and last sounds were always of the same length, resulting in four length conditions: SSS, SLS, LSL, and LLL (see Fig. 1). SLS and LSL could be regarded as nonrepeated versions of sequences in which a short sound and a long sound were alternated repeatedly in

¹ In their experiment, each trial consisted of two repeated sequences, and participants compared these sequences in terms of the duration of the target gap included in each sequence. For example, participants judged whether the gap between the short sound and the long sound was shorter or longer in the first sequence than in the second sequence. Because the target gap was marked by the same pair of two sounds (either the pair of the short and then the long sound or that of the long and then the short sound) in both sequences, the results did not tell anything about the perceived-duration differences between the gap between the short and the long sound and the gap between the long and the short sound.

² This participant showed the "exactly halfway" response in more than 95 % trials for two conditions, resulting in almost no "too early" or "too late" responses to construct psychometric functions.

Fig. 1 Stimulus patterns of Experiment 1. T was 280, 320, 360, 400, 440, 480, or 520 ms

Experiment 2. The parameters of 150 ms and 262.5 ms were also used in Iversen et al. (2008) and Kuroda et al. (2013). Kuroda et al. reported that listeners' ability for detecting a change in the duration of single gaps was severely impaired when the first marker was lengthened from 150 to 262.5 ms. Rammsayer and Leutner (1996) also indicated that the ability for detecting a change in the duration of gaps is shifted when sound length exceeds a critical point of 200 ms. Thus, the lengthening of the marker from 150 to 262.5 ms might also influence perceived duration in this study. To avoid spectral splatter, amplitude rose and decayed during 20 ms at the beginning and end of each sound with raised-cosine ramps, the ramps being included in the sound length.

The interonset interval between the initial and the last sounds was fixed at 800 ms whereas the position of the middle sound's onset was varied; its onset was placed at 0, 40, 80, and 120 ms before or after the physical bisection point (400 ms), resulting in seven temporal positions (-120, -80, -40, ± 0 , +40, +80, +120 ms). Empty intervals around 400 ms were also used in Geiser and Gabrieli (2013).

Procedure

Participants were instructed to judge whether three sounds were presented at equal interonset intervals with three alternatives. They responded "exactly halfway" when the middle sound's onset was located exactly halfway between the initial and last onsets or when the three sounds' onsets appeared at exactly equal intervals. They responded "too early" when the middle sound's onset appeared too early to perceive the three sounds' onsets as presented at equal intervals or when the middle sound's onset was closer to the initial one. They responded "too late" when the middle sound's onset appeared too late to perceive the three sounds' onsets as presented at equal intervals or when the middle sound's onset was closer to the last one.

Participants listened to the stimulus pattern by clicking on the "play" pane. A 2-s silent interval began after clicking, and then the pattern was presented. Participants could listen to the pattern only once in each trial, but when listening was disturbed for some specific reason (e.g., yawning or coughing), they were allowed to listen again by clicking on the "replay" pane.

Four experimental conditions (SSS, SLS, LSL, and LLL) were presented in separate sessions. The order of these

sessions was counterbalanced. Each session was divided into three blocks. In each block, the seven patterns for the seven middle positions were presented 10 times each in a random order, resulting in 70 trials. Each block was preceded by two practice trials where randomly selected stimuli were presented. Each session took about 30 minutes.

Data analysis

The probability of responding "too early" and of responding "too late" was calculated against the middle sound's position for each experimental condition for each individual. Each probability was based on 30 responses ($=3$ blocks $\times$ 10 trials). The point of subjective equality (PSE) was defined as the midpoint of *uncertainty interval* (Guilford, 1954; see also Kuroda & Hasuo, 2014) that was estimated from the psychometric function of the "too early" and that of the "too late" probability. Technically, the cumulative normal distribution was fit to the "too late" function and to the reversed "too early" function (1 minus the probability) with a nonlinear least-square method (Levenberg-Marquardt algorithm). The lower limit of the uncertainty interval was an x -axis value at which the curve fit to the reversed "too early" function crossed .50 probability, whereas the upper limit was an x -axis value at which the curve fit to the "too late" function crossed .50 probability. PSE was given by dividing the sum of the lower and the upper limit by 2.

Results

The mean probability of responding "too early" and of responding "too late" as a function of the middle sound's position is shown in Fig. 2. The goodness of fitting the cumulative normal distribution to each function was generally high: For "too late," the R^2 value was above .90 in 63 cases out of 64 trials (4 patterns $\times$ 16 participants) and was .43 in one case.³ For "too early," the R^2 value was above .90 in 61 cases and was between .90 and .50 in three cases.

³ The data from the participant showing this low R^2 were kept in the following analysis. The results did not much change even when the data from this participant were removed, $F(3, 42) = 6.019$, $p = .002$, $\eta_p^2 = .301$. Tukey HSD analysis showed a significant difference between SLS versus LSL ($p < .001$, $d = .738$).

Fig. 2 Mean probability of responding “too early” and of responding “too late” for each experimental condition in Experiment 1. Bars represent 95 % confidence intervals

The mean PSE for each experimental condition is shown in Fig. 3a. The 95 % confidence intervals (CIs) are also apparent in this figure and indicate whether the mean PSE significantly differed from zero. The mean was significantly lower than zero in SSS, $t(15) = 2.348$, $p = .033$, $d = .587$, in SLS, $t(15) = 2.929$, $p = .010$, $d = .732$, in LLL, $t(15) = 3.395$, $p = .004$, $d = .849$, but not in LSL, $t(15) = 1.491$, $p = .157$, $d = .373$. A one-way repeated-measures analysis of variance (ANOVA) showed a significant effect for the four experimental conditions, $F(3, 45) = 5.447$, $p = .003$, $\eta_p^2 = .266$. Multiple comparisons based on the Tukey HSD method showed that the mean was significantly lower in SLS than in LSL ($p = .001$, $d = .650$).

Fig. 3 Mean point of subjective equality for each experimental condition in Experiment 1. Bars represent 95 % confidence intervals. **a** The y-axis is in the scale of the position of the middle sound's onset relative to the physical bisection point, corresponding to the x-axis in Fig. 2. **b** The y-axis is in the scale of the SL interval relative to the LS interval, corresponding to the one used in Experiment 2

An additional analysis was conducted to predict from the current results what duration conditions would enable listeners to perceive equality between interonset intervals when a short sound and a long sound were alternated repeatedly in Experiment 2. Because this repeated pattern includes both the SLS and LSL structures as presented in Fig. 4, a possible indicator for predicting the results of Experiment 2 would be the simple average of the results of SLS and of LSL in this experiment. However, the order of the SL interval and LS interval for SLS was opposite to that for LSL. In other words, PSE for SLS represents the SL interval minus 400 ms (the bisection point) when the SL and LS intervals were perceived as equivalent, whereas PSE for LSL represents the LS interval minus 400 ms when the SL and LS intervals were perceived as equivalent. To make it possible to average the results of SLS and LSL, we recalculated PSE in the scale of the relative duration of the SL to the LS interval for each sequence (see Fig. 3b, SLS and LSL). This variable is called PSE_{SL-LS} . For example, the PSE of 40 ms for LSL (= LS interval of 440 ms – bisection point of 400 ms) corresponds to the PSE_{SL-LS} of -80 ms (= SL interval of 360 ms – LS interval of 440 ms). The results of SLS and LSL were then averaged for each individual. The 95 % CIs in Fig. 3b indicate that the mean result for the average of SLS and LSL, that is, $(SLS + LSL)/2$ in the figure, was significantly lower than zero, $t(15) = 2.602$, $p = .020$, $d = .650$. Therefore, it would be reasonable to posit that participants would perceive the equality between the

Fig. 4 A stimulus sequence used in Experiment 2 and its relation with SLS and LSL of Experiment 1. I_{SL}/I_{LS} was 280/520, 320/480, 360/440, 400/400, 440/360, 480/320, or 520/280 ms

interonset intervals in the repeated pattern when the SL interval was physically shortened relative to the LS interval, rather than when these intervals were exactly equivalent.

Discussion

The mean PSE (not PSE_{SL-LS}) was about -5 ms when three sounds were of equal length, that is, for the SSS and LLL sequences, indicating that the three sounds were perceived as presented at equal interonset intervals when the middle sound's onset was located 5 ms before the physical bisection point. This could be regarded as a type of time order error (Eisler, Eisler, & Hellström, 2008). Because participants tended to perceive the first interval (between the initial and the middle onsets) as longer than the second interval (between the middle and the last onsets), the first interval had to be physically shortened relative to the second interval so that participants perceived these intervals as equal.

However, such bias (PSE significantly lower than zero) disappeared in the LSL sequence, and, moreover, this sequence led to significantly higher PSE than the SLS sequence. This SLS versus LSL difference was consistent with that reported by Hasuo et al. (2011) and could be attributed to delayed perception of the onset with the lengthening of the sound (p center shift; see Gordon, 1987; Morton, Marcus, & Frankish, 1976). Because the onset was perceived later for the long sound than for the short sound, the middle sound had to be presented physically earlier for SLS than for LSL so that participants perceived the three sounds as presented at equal interonset intervals.

Experiment 2

The lengthening of each sound resulted in delayed perception of the onset in Experiment 1. We examined whether the same effects of marker length would occur when a short sound and a long sound were alternated repeatedly (SLSLSL...or

LSLSLSL...) in Experiment 2. If the onset is perceived later for the long sound than for the short sound, participants should perceive the equality between the interonset intervals in this repeated sequence when the SL interval is physically shortened relative to the LS interval. However, as mentioned earlier, the perceived duration of the interonset intervals in the repeated sequence might be influenced by whether participants would perceive SL or LS grouping for the sequence. In other words, between-groups intervals might be perceived as longer than within-groups intervals. If individuals perceive SL grouping for the current sequence, the LS interval should be perceived as longer than the SL interval; then, participants should perceive the equality between the interonset intervals in the repeated sequence when the LS interval is physically shortened relative to the SL interval.

It might also be possible that the effects of rhythmic grouping would be working together with the delayed perception of the onset for the long sound relative to the short sound. If SL grouping occurs, for example, the effects of rhythmic grouping on perceived duration (resulting in the $SL < LS$ interval in perception) may be canceled out by the onset delay for the long sound relative to the short sound (the $SL > LS$ interval in perception). Then, participants might perceive the equality between the interonset intervals when the SL and LS intervals were physically equivalent.

The experiment consisted of two sessions, one for discrimination and one for grouping. In the discrimination session, we estimated the relative duration of the SL to the LS interval that made participants perceive the equality between the interonset intervals in the repeated sequence. In the grouping session, we estimated whether participants perceived the repeated sequence as segmented into SL or LS groups. The correlation between the results in these sessions was examined.

Japanese-speaking participants, as well as French-speaking participants, were recruited in this experiment. Japanese-speaking participants might have a different tendency to perceive rhythmic groupings than French-speaking participants. Iversen et al. (2008) presented stimulus sequences in which a short sound and a long sound were alternated repeatedly to American-English-speaking and Japanese-speaking participants. Most of the English-speaking participants perceived SL grouping, whereas about half the Japanese-speaking participants perceived LS grouping. The similar sequences were used by Kuroda et al. (2013) for French-speaking participants, who tended to perceive SL grouping. Hay and Diehl (2007) also indicated no difference between English-speaking versus French-speaking participants. In brief, French-speaking participants tended to perceive SL grouping while Japanese-speaking participants were almost equally split on whether each individual tended to perceive SL or LS grouping. Although it remains debatable which factor causes the difference in the tendency of rhythmic grouping (Hay & Diehl, 2007; Iversen et al., 2008; Patel, 2008), the perceived duration of interonset intervals in

the current sequence might also differ between French-speaking and Japanese-speaking participants if it is influenced by the tendency to perceive rhythmic grouping.

Method

Participants

Forty-five participants, self-reporting having normal hearing, were recruited: Twenty (12 females) were from Laval University, Québec, Canada; 13 (five females) from Kyushu University, Fukuoka, Japan; and 12 (five females) from Yamaguchi University, Yamaguchi, Japan. The reason of recruiting participants from two universities in Japan was just to accelerate data collection, and the data from these universities were combined. Participants consented to their participation by signing a form approved by the ethics review board of their institution.

Although English is a language spoken by many residents in Québec and by most university students, French is legally recognized as the official language of that province. French is also the official language of Laval University, and participants from that university self-reported that their native language was French. Hay and Diehl (2007) reported no rhythmic-grouping difference between English versus French speakers, and thus, there was no reason to posit that the tendency of rhythmic grouping changes even if a participant can speak English as well as French. One participant also took part in Experiment 1.

Participants from Kyushu and Yamaguchi University self-reported that their native language was Japanese. English classes are included in the compulsory education of Japan, whereas no participants reported being highly proficient at English.

Because the effects of rhythmic grouping on perceived duration may depend on the listener's expertise in music performance (Geiser & Gabrieli, 2013), we asked participants about the length of music training they took for private lessons or club activities. The mean period of training (minimum–maximum value) was 5.1 (0–17) years for French-speaking participants and 3.2 (0–15) year for Japanese-speaking participants. A one-way between-participants ANOVA showed that the mean training period did not significantly differ between French-speaking and Japanese-speaking participants, $F(1, 43) = 2.026$, $p = .162$, $\eta_p^2 = .045$. In addition, the mean age was 22 (19–32) years for French-speaking participants and 22 (21–32) years for Japanese-speaking participants. The ANOVA showed that the mean age did not significantly differ between French-speaking and Japanese-speaking participants, $F(1, 43) = .073$, $p = .788$, $\eta_p^2 = .002$.

Stimuli

Each of the two (discrimination and grouping) sessions consisted of 168 trials (=4 blocks $\times$ 42 trials), as explained in the following section. In each trial, a 150-ms sound and a 262.5-ms sound,

which corresponded to the short and the long sound of Experiment 1, respectively, were alternately presented 12 times, resulting in 24 sounds. On half trials, the initial sound of the pattern was short (i.e., the pattern was SLSLSLSL...), and on the other half, it was long (i.e., the pattern was LSLSLSL...). The data from these two types of patterns were collapsed. In addition, to prevent participants from distinguishing these patterns with the length of the first few sounds, intensity slopes were added at the beginning and end of each pattern. The level of the initial sound was attenuated by 32 dB. The level was then increased by 4 dB per sound and reached at a plateau (66 dBA) at the ninth sound. The level was constant until the 16th sound, and then decreased by 4 dB per sound.

The interonset interval between the short sound and the following long sound (I_{SL}) was varied from 280 to 520 ms whereas the interonset interval between the long sound and the following short sound (I_{LS}) was varied from 520 to 280 ms in steps of 40 ms (see Fig. 4). For example, I_{SL} was 320 ms when I_{LS} was 480 ms. These intervals were alternated repeatedly, like 320/480/320/480 ... when the initial sound was short, and like 480/320/480/320 ... when the initial sound was long. Consequently, $I_{SL} + I_{LS}$ was fixed at 800 ms.

Stimulus patterns were presented from headphones (Sennheiser HD 650) connected to an amplifier (Teac A-H01). This amplifier included a function of digital-to-analogue conversion and was directly (USB) connected to a computer in Kyushu University. In the other universities, however, the amplifier was connected to a computer via an USB audio processor (Onkyo SE-U33GXV) that converted digital to analogue signals, because of problems in the compatibility with the operating system. We recorded some outputted sounds and checked that the physical timing of these sounds was reasonably correct in these institutions.

Procedure

Participants completed the two sessions within one day. The discrimination session always preceded the grouping session in order to avoid the possibility that participants would utilize the knowledge regarding rhythmic grouping as a cue for equality discrimination.

In the discrimination session, participants were instructed to judge whether the onset of each sound appeared at “equal” or “unequal” intervals. The task was based on two-alternative choice instead of three-alternative choice used in Experiment 1. We avoided asking participants to judge whether the SL or the LS interval was longer because asking which interval was longer might encourage participants to pay attention to only one pair of SL and LS intervals included in the pattern, and this selective attention might influence the occurrence of rhythmic grouping.

In the grouping session, participants were instructed to judge whether the stimulus pattern was perceived as segmented into

groups consisting of the short sound followed by the long sound (“short–long” grouping) or of the long sound followed by the short sound (“long–short” grouping). Participants could also respond that any “other” grouping (or no grouping) occurred, whereas they were instructed to choose this alternative only when definitely necessary. The similar procedures were adopted by Kuroda et al. (2013).

For both sessions, when perception changed within one pattern (there was more than one possible response), participants were instructed to respond the one that they thought as the most dominant. Each session was divided into four blocks. In each block, 14 patterns (seven durations for SLSLSL... and LLSLSL...) were presented three times each in a random order, resulting in 42 trials. Each block was preceded by two warm-up trials, where randomly selected stimuli were presented. Each session took about 1 hour.

Data analysis

In the equality discrimination session, the probability of responding “equal” is calculated against I_{SL} minus I_{LS} for each individual. Since the two patterns (SLSLSL... and LLSLSL...) were collapsed as mentioned earlier, this probability was based on 24 responses (=2 patterns $\times$ 4 blocks $\times$ 3 responses). It was difficult to adopt a curve-fitting procedure (e.g., utilizing Gaussian function) for the estimation of PSE_{SL-LS} because the results obtained from 10 (four from French-speaking and six from Japanese-speaking) participants did not conform to a bell-shaped function (i.e., showed no clear peak).⁴ Alternatively, we estimated PSE_{SL-LS} with the following formula according to the arithmetic average method (Guilford, 1954), which was equivalent with the estimation of the first moment in the waveform moment analysis (Birngruber, Schröter, & Ulrich, 2015; Cacioppo & Dorfman, 1987):

$$PSE_{SL-LS} = \frac{\sum_{i=1}^7 p_i d_i}{\sum_{i=1}^7 p_i} \quad (1)$$

d_i represents the relative duration of I_{SL} to I_{LS} ($d_1, d_2, d_3, \dots, d_7 = -240, -160, -80, \dots, +240$ ms). The “equal” probability for the corresponding relative duration is represented by p_i . A negative PSE_{SL-LS} indicates that participants more frequently responded “equal” when the SL interval was physically shortened relative to the LS (the relative duration was below zero) than when the SL was lengthened relative to the LS interval (the relative duration was above zero); also, a positive PSE_{SL-LS} indicates that

participants more frequently responded “equal” when the relative duration was above zero than when it was below zero.

In the grouping session, there were 169 “other” responses (2.2 %) out of 7,560 trials (=45 participants $\times$ 7 durations $\times$ 24 responses). Because this number was reasonably small, half of the “other” responses were allocated to the “short–long” (i.e., SL grouping) responses and the remaining to the “long–short” (i.e., LS grouping) responses at each relative duration (I_{SL} minus I_{LS}) for each individual. Then, the probability of responding that “long–short” grouping occurred is calculated against the relative duration. The cumulative normal distribution was fit to the resulting psychometric function. We defined the point of grouping change (PGC) as an x -axis value at which the curve crossed 0.50. Because the function is shifted leftward if participants more frequently respond “long–short” grouping, a negative PGC indicates that participants tended to perceive LS grouping while a positive PGC indicates that participants tended to perceive SL grouping.

Results

Equality discrimination

The mean probability of responding “equal” as a function of the relative duration (I_{SL} minus I_{LS}) and the mean PSE_{SL-LS} for each participant group are shown in Fig. 5a and b, respectively. A one-way between-participants ANOVA showed that the mean PSE_{SL-LS} did not significantly differ between French-speaking and Japanese-speaking participants, $F(1, 43) = 1.857$, $p = .180$, $\eta_p^2 = .041$.⁵ Indeed, as presented in Fig. 5b, the y -axis value (PSE_{SL-LS}) of zero is included in the 95 % CIs, indicating that the mean PSE_{SL-LS} did not significantly differ from zero, for French-speaking participants, $t(19) = .886$, $p = .387$, $d = .198$, Japanese-speaking participants, $t(24) = 1.240$, $p = .227$, $d = .248$, and the data in which these groups were combined, $t(44) = .886$, $p = .380$, $d = .132$.

Rhythmic-grouping judgment

The mean probability of responding that “long–short” grouping occurred as a function of I_{SL} minus I_{LS} is shown in Fig. 6a. The R^2 value for the fitting of the cumulative normal distribution was above .90 in 32 participants, between .90 and .50 in 10 participants, and between .50 and .25 in three participants.⁶

⁴ We attempted to fit the Gaussian function to individual data, but for these participants the peak was estimated outside the range between -240 and +240 ms, or the goodness of fit was poor ($R^2 < .34$). In the supplement document, the results when they were removed so that the fitting procedure was made possible are reported.

⁵ The variance between participants was quite high in Japanese-speaking participants relative to French-speaking participants, but the Kolmogorov–Smirnov test showed that the data conformed to a normal distribution in Japanese-speaking participants ($p = .605$, two-tailed) as well as French-speaking participants ($p = .506$). See also Fig. 7 for the data distribution.

⁶ The data from the participants showing R^2 below .50 were kept in the following analysis. The results when these data were removed are reported in the supplement document.

Fig. 5 Results in the equality discrimination of Experiment 2. **a** Mean probability of responding “equal” as a function of I_{SL} minus I_{LS} . **b** Mean point of subjective equality (PSE_{SL-LS}) for each participant group. Bars represent 95 % confidence intervals

The mean PGC is shown in Fig. 6b. A one-way between-participants ANOVA showed that the mean PGC did not significantly differ between French-speaking and Japanese-speaking participants, $F(1, 43) = .342$, $p = .562$, $\eta_p^2 = .008$. The 95 % CIs indicate that the mean was significantly lower than zero, indicating the tendency to perceive LS grouping, for French-speaking participants, $t(19) = 3.250$, $p = .004$, $d = .727$, Japanese-speaking participants, $t(24) = 2.111$, $p = .045$, $d = .422$, and the data in which these groups were combined, $t(44) = 3.636$, $p < .001$, $d = .542$.

Relation between temporal equality perception and rhythmic grouping

The scatter plots of PSE_{SL-LS} versus PGC are shown in Fig. 7. If between-groups intervals are perceived as longer, a positive correlation should be found between PSE_{SL-LS} versus PGC. In other words, PSE_{SL-LS} should be a positive value (equality perception when $I_{SL} > I_{LS}$) if PGC is a positive value (SL grouping tendency); also, PSE_{SL-LS} should be a negative value (equality perception when $I_{SL} < I_{LS}$) if PGC is a negative value (LS grouping tendency). However, as presented in Fig. 7, there were no systematic relationships between these dependent variables. Indeed, no cases led to significant correlation

according to Pearson’s coefficient ($p = .602$, two-tailed, for French-speaking participants; $p = .086$ for Japanese-speaking participants; and $p = .139$ when these groups were pooled).

Discussion

The first purpose of Experiment 2 was to examine the temporal conditions that made participants perceive the equality between the interonset intervals in the repeated sequence. For both French-speaking and Japanese-speaking participants, the mean PSE_{SL-LS} did not significantly differ from zero, indicating that the current sequence was perceived as consisting of equal interonset intervals when the onsets of all sounds were aligned at physically equal intervals. This finding is inconsistent with what we predicted from the data of Experiment 1, where only three sounds were presented. Indeed, in Experiment 1, the mean PSE_{SL-LS} was around -32 and -26 ms for the SLS and LSL sequence, respectively, and the average was -29 ms, indicating that the lengthening of each marker resulted in delayed perception of the onset. If the same onset delay had occurred in Experiment 2, participants should have perceived the equality between the

Fig. 6 Results in the rhythmic-grouping judgment of Experiment 2. **a** Mean probability of responding that “long-short” grouping occurred as a function of I_{SL} minus I_{LS} . **b** Mean point of grouping change (PGC) for each participant group. Bars represent the 95 % confidence intervals

Fig. 7 Scatter plots of the point of grouping change (PGC) versus the point of subjective equality (PSE_{SL-LS}) in Experiment 2

interonset intervals when the SL interval was physically shortened relative to the LS interval.

The second purpose was to examine the relation between perceived duration and rhythmic grouping. If between-groups intervals had been perceived as longer than within-groups intervals, as predicted from the previous studies (Bolton, 1894; Geiser & Gabrieli, 2013), a positive correlation should have been found between PSE_{SL-LS} versus PGC. However, no significant correlation was found between these variables, and thus, we could not find evidence supporting the notion that between-groups intervals are perceived as longer.

As mentioned earlier, it might have been possible that the effects of rhythmic grouping interacted with the delayed perception of the onset for the long sound compared with the short sound. If participants had perceived SL grouping, the effects of rhythmic grouping on perceived duration (resulting in the $SL < LS$ interval in perception) should have been canceled out by the delayed perception of the onset (resulting in the $SL > LS$ interval in perception). The mean PSE_{SL-LS} was indeed almost zero for both French-speaking and Japanese-speaking participants. However, this explanation is not plausible because the mean PGC was significantly lower than zero in the grouping session, indicating the tendency to perceive LS grouping over SL grouping, for both French-speaking and Japanese-speaking participants. In this case, LS grouping should have yielded a longer perceived duration for the SL (between-groups) interval than the LS (within-groups) interval and, moreover, the delayed perception of the onset should also have yielded a longer perceived duration for the SL interval than the LS interval. Then, participants should have perceived the equality between the interonset intervals when the SL interval was physically shortened relative to the LS interval, resulting in the PSE_{SL-LS} lower than zero. However, the mean PSE_{SL-LS} was indeed almost zero.

Finally, the grouping session of the present experiment did not replicate the finding of the previous studies that French-speaking participants tended to perceive SL grouping (Hay & Diehl, 2007; Kuroda et al., 2013). The results of this session

(PGC) rather indicated that they tended to perceive LS grouping over SL grouping. One might argue that this result could be attributed to the intensity slope added at the beginning of the sequence (the level increased by 4 dB per sound). As mentioned in the stimulus section, this slope was added in order to reduce the effects of the initial sound of the sequence on rhythmic grouping. However, the long sound might be detected more easily than the short sound because of a higher sound energy (temporally integrated power) of the long sound. The long sound might thus likely become an initial sound in perception, facilitating LS grouping. This explanation would be less plausible given that the similar intensity slopes were adopted in the previous studies (Hay & Diehl, 2007; Kuroda et al., 2013) that demonstrated the SL-grouping tendency in French-speaking participants. However, the exact experimental settings differed between this study and the previous studies, making it difficult to rule out completely the possibility of a better detection of the long sound. Note that, despite this potential problem, we could keep the proposition that rhythmic grouping and perceived duration did not correlate with each other in this experiment. Indeed, LS grouping should have yielded a longer perceived duration for the SL interval than for the LS interval, even though this grouping was caused by some artifacts, if between-groups intervals had been perceived as longer. However, participants perceived the equality between the interonset intervals when all sounds were aligned at physically equal interonset intervals, rather than when the SL interval was physically shortened relative to the LS interval.

General Discussion

The purpose of our study was to investigate how the perceived duration of empty intervals is modulated by the length of sounds marking those intervals. Stimulus sequences consisted of three sounds in Experiment 1 and consisted of a short sound and a long sound that were alternated repeatedly in Experiment 2. The main results can be summarized as follows: (1) When three sounds were successively presented, the middle sound had to be presented earlier in the SLS than in the LSL sequence so that participants perceived these sounds as presented at equal interonset intervals. (2) Such length effects disappeared when a short sound and a long sound were alternated repeatedly—these sounds were perceived as presented at equal inter-onset intervals when the SL and LS intervals were physically equivalent. Indeed, the mean PSE_{SL-LS} was -29 ms when the results of the SLS and LSL sequences were averaged in Experiment 1, whereas the mean PSE_{SL-LS} was -3 and +11 ms for French-speaking and Japanese-speaking participants, respectively, in Experiment 2. (3) The perceived duration of interonset intervals in the repeated sequence was not influenced by whether the participant's native language was French or Japanese, or by whether

participants tended to perceive SL or LS grouping (no correlation between PSE_{SL-LS} vs. PGC).

These results (especially 1 and 2 in the previous paragraph) indicate that the lengthening of each marker resulted in delayed perception of the onset when only three sounds were successively presented, consistent with Hasuo et al. (2011), whereas such delay did not occur when a short sound and a long sound were alternated repeatedly. The marker-length effects seemed to be canceled out when the presentation of intervals was repeated. The similar repetition effects were reported by Kuroda et al. (2013) that examined the listener's ability of detecting a change in the duration of temporal gaps (instead of perceived duration). They found that, when the presentation of a gap marked by two sounds was not repeated, the ability of detecting a change in the gap duration was impaired by lengthening the first marker more than by lengthening the second marker. However, when a short sound and a long sound were alternated repeatedly, the performance of detecting a change in the gap duration remained identical regardless of whether participants discriminated the gaps between the short and the long sound or those between the long and the short sound.

The results of this study indicate the difficulty demonstrating any systematic relations between rhythmic grouping and perceived duration. Previous studies have posited that a change in the duration of between-groups intervals is more difficult to detect than a change in the duration of within-groups intervals (Fitzgibbons et al., 1974; Geiser & Gabrieli, 2013; Thorpe & Trehub, 1989; Trainor & Adams, 2000) and that between-groups intervals are perceived as longer than within-groups intervals (Bolton, 1894; Geiser & Gabrieli, 2013). We tested the latter proposition and found no significant correlation between PSE_{SL-LS} versus PGC. Kuroda et al. (2013) focused on the former issue but failed to find any relations between the listener's tendency to perceive rhythmic (SL or LS) grouping and the ability of detecting a change in the duration of gaps. We must be careful, however, in arguing whether these null results could be evidence against the proposition of the dependency between duration processing and rhythmic grouping. Indeed, several stimulus parameters differed between our and the previous studies. There is clearly a need here for additional data in which both the performance of some temporal tasks (such as accuracy or sensitivity) and the tendency of rhythmic grouping are measured.

The results of the present study may be compared with those of Vos, Jirí, and Kruysbergen (1995). In their experiments, participants synchronized their tapping with the onset of a sound that was repeated at a constant interval. The length of the sound was manipulated across trials but it was fixed within each trial (sequence). These authors reported that the timing of the tapping was delayed when a longer sound was repeated, indicating the delayed perception of the onset with the lengthening of the sound. This finding did not seem consistent with what we found in Experiment 2. That is, there were no signs

indicating delayed perception of the onset when a short sound and a long sound were alternated repeatedly. Vos et al. and our study differed in the experimental task (reproduction vs. discrimination), but they were similar in terms of the fact that the stimulus sequences consisted of many sounds. However, a single sound, of which length was fixed, was simply repeated in Vos et al., whereas two sounds of different lengths were alternated repeatedly in our study. Our study thus used more complicated structures of sequences. Although speculative, the effects of marker length might not be working well with more complicated structures like the ones used in our study. This speculation could also be applied to the results of Kuroda et al. (2013), where the effects of marker length on the detection of a change in the gap duration disappeared when a short sound and a long sound were alternated repeatedly. Alternatively, the cancellation of the marker-length effects for the present repeated sequences might be attributed to the fact that temporal discrimination was based on the local perception in Experiment 1 but on the global perception in Experiment 2. Participants focused on the temporal location of the middle sound in Experiment 1, which might facilitate the effects of the length of this sound. However, participants had to perceive the global structure of the sequence in Experiment 2, which might diminish the effects of the length of each individual sound.

Many physical parameters including length and intensity are changing within each sequence in music and speech. The perceptual system is engaged to process duration or rhythm from the sequences of such complicated structures. A further examination of duration processing with sound sequences within which the physical parameters of markers are changed, as in the ones used by the present study, may give new insights into understanding the mechanisms processing temporal information in music and speech.

Abbreviations *SL grouping* Perceiving a sequence of successive sounds as segmented into rhythmic groups consisting of a short sound followed by a long sound, *LS grouping* Perceiving a sequence of successive sounds as segmented into rhythmic groups consisting of a long sound followed by a short sound, *SL interval* (I_{SL}) An interonset time interval between a short sound and a long sound, *LS interval* (I_{LS}) An interonset time interval between a long sound and a short sound, *SSS* A sequence consisting of three short sounds, *SLS* A sequence consisting of a short, a long, and a short sound, in this order, *LSL* A sequence consisting of a long, a short, and a long sound, in this order, *LLL* A sequence consisting of three long sounds, *PSE* Point of subjective equality, *PGC* Point of grouping change

Authors' Notes This research was made possible by research grants awarded to T. K. by JSPS KAKENHI Grant Number 13J03275, 15K21195, 16H01502; to E. T. by JSPS KAKENHI Grant Number 24730623, 16J40101; to M. M. by JSPS KAKENHI Grant Number 25242058, 26119521, 16H01866; and to S. G. by the Natural Sciences and Engineering Council of Canada. We would like to express our gratitude

to Félix Désautels, Guillaume Plante, Tomomi Funai, Takuya Kishida, Tomomi Koizumi, and Ryuta Fukunaga for their help in data collection and to Dr. Doug Alards-Tomalín and one anonymous reviewer for their comments on previous versions of this article. All correspondence should be addressed to Tsuyoshi Kuroda, Faculty of Informatics, Shizuoka University, 3-5-1 Johoku, Naka-ku, Hamamatsu, Shizuoka, 432-8011, Japan. E-mail: tkuroda@inf.shizuoka.ac.jp

References

- Alards-Tomalín, D., Leboe-McGowan, L. C., & Mondor, T. A. (2013). Examining auditory kappa effects through manipulating intensity differences between sequential tones. *Psychological Research*, 77(4), 480–491. doi:10.1007/s00426-012-0438-8
- Bimgruber, T., Schröter, H., & Ulrich, R. (2015). The influence of stimulus repetition on duration judgments with simple stimuli. *Frontiers in Psychology*, 6(1213). doi:10.3389/fpsyg.2015.01213
- Bolton, T. L. (1894). Rhythm. *American Journal of Psychology*, 6(2), 145–238.
- Cacioppo, J. T., & Dorfman, D. D. (1987). Waveform moment analysis in psychophysiological research. *Psychological Bulletin*, 102(3), 421–438. doi:10.1037/0033-2909.102.3.421
- Eisler, H., Eisler, A. D., & Hellström, Å. (2008). Psychophysical issues in the study of time perception. In S. Grondin (Ed.), *Psychology of time* (pp. 75–109). Bingley: Emerald Group.
- Fitzgibbons, P. J., Pollatsek, A., & Thomas, I. B. (1974). Detection of temporal gaps within and between perceptual tonal groups. *Perception & Psychophysics*, 16(3), 522–528. doi:10.3758/BF03198581
- Geiser, E., & Gabrieli, J. D. E. (2013). Influence of rhythmic grouping on duration perception: A novel auditory illusion. *PloS One*, 8(1), e54273. doi:10.1371/journal.pone.0054273
- Gordon, J. W. (1987). The perceptual attack time of musical tones. *Journal of the Acoustical Society of America*, 82(1), 88–105. doi:10.1121/1.395441
- Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, 72(3), 561–582. doi:10.3758/APP.72.3.561
- Grondin, S., Ivry, R. B., Franz, E., Perreault, L., & Metthé, L. (1996). Markers' influence on the duration discrimination of intermodal intervals. *Perception & Psychophysics*, 58(3), 424–433. doi:10.3758/BF03206818
- Grondin, S., Roussel, M. E., Gamache, P. L., Roy, M., & Ouellet, B. (2005). The structure of sensory events and the accuracy of time judgments. *Perception*, 34(1), 45–58. doi:10.1068/p5369
- Grose, J. H., Hall, J. W., & Buss, E. (2001). Gap duration discrimination in listeners with cochlear hearing loss: Effects of gap and marker duration, frequency separation, and mode of presentation. *Journal of the Association for Research in Otolaryngology*, 2(4), 388–398. doi:10.1007/s101620010067
- Guilford, J. P. (1954). *Psychometric methods*. New York: McGraw-Hill.
- Hasuo, E., Nakajima, Y., & Hirose, Y. (2011). Effects of sound-marker durations on rhythm perception. *Perception*, 40(2), 220–242. doi:10.1068/p6846
- Hasuo, E., Nakajima, Y., Osawa, S., & Fujishima, H. (2012). Effects of temporal shapes of sound markers on the perception of interonset time intervals. *Attention, Perception, & Psychophysics*, 74(2), 430–445. doi:10.3758/s13414-011-0236-1
- Hay, J. S. F., & Diehl, R. L. (2007). Perception of rhythmic grouping: Testing the iambic/trochaic law. *Perception & Psychophysics*, 69(1), 113–122. doi:10.3758/BF03194458
- 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(2), 551–564. doi:10.1037/0096-1523.35.2.551
- Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America*, 124(4), 2263–2271. doi:10.1121/1.2973189
- Kuroda, T., & Hasuo, E. (2014). The very first step to start psychophysical experiments. *Acoustical Science and Technology*, 35(1), 1–9. doi:10.1250/ast.35.1
- Kuroda, T., Hasuo, E., & Grondin, S. (2013). Discrimination of brief gaps marked by two stimuli: Effects of sound length, repetition, and rhythmic grouping. *Perception*, 42(1), 82–94. doi:10.1068/p7186
- Kuroda, T., Hasuo, E., Labonté, K., Laflamme, V., & Grondin, S. (2014). Discrimination of two neighboring intra- and intermodal empty time intervals marked by three successive stimuli. *Acta Psychologica*, 149, 134–141. doi:10.1016/j.actpsy.2013.11.013
- Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers). *Psychological Review*, 83(5), 405–408. doi:10.1037/0033-295X.83.5.405
- Neff, D. L., Jesteadt, W., & Brown, E. L. (1982). The relation between gap discrimination and auditory stream segregation. *Perception & Psychophysics*, 31(5), 493–501. doi:10.3758/BF03204859
- Ono, F., & Kitazawa, S. (2009). The effect of marker size on the perception of an empty interval. *Psychonomic Bulletin & Review*, 16(1), 182–189. doi:10.3758/PBR.16.1.182
- Patel, A. D. (2008). *Music, language, and the brain*. New York: Oxford University Press.
- Penner, M. J. (1976). The effect of marker variability on the discrimination of temporal intervals. *Perception & Psychophysics*, 19(5), 466–469. doi:10.3758/BF03199409
- Rammsayer, T. H., & Leutner, D. (1996). Temporal discrimination as a function of marker duration. *Perception & Psychophysics*, 58(8), 1213–1223. doi:10.3758/BF03207554
- 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: Emerald Group.
- Thorpe, L. A., & Trehub, S. E. (1989). Duration illusion and auditory grouping in infancy. *Developmental Psychology*, 25(1), 122–127. doi:10.1037/0012-1649.25.1.122
- Trainor, L. J., & Adams, B. (2000). Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. *Perception & Psychophysics*, 62(2), 333–340. doi:10.3758/BF03205553
- Vos, P. G., Jirí, M., & van Kruysbergen, N. W. (1995). The perceptual centre of a stimulus as the cue for synchronization to a metronome: Evidence from asynchronies. *Quarterly Journal of Experimental Psychology Section A*, 48(4), 1024–1040. doi:10.1080/14640749508401427
- Woodrow, H. (1928). Behavior with respect to short temporal stimulus-forms. *Journal of Experimental Psychology*, 11(3), 167–193. doi:10.1037/h0074282
- Xuan, B., Zhang, D., & Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of Vision*, 7(10). doi:10.1167/7.10.2