

# Does a Rhythmic Context Have an Effect on Perceptual Weights in Auditory Intensity Processing?

Daniel Oberfeld

Johannes Gutenberg–Universität Mainz

The effects of a rhythmic context on auditory intensity processing were studied. Experiment 1 tested the hypothesis that the involuntary temporal direction of attention by a rhythmic context sequence influences the temporal weighting of loudness. Perceptual weight analysis was used to measure the attention directed to individual temporal portions of a longer stimulus by estimating the importance of individual temporal segments for global loudness judgements. A rhythmic context resulted in a sharper temporal weighting profile, but the expected rhythmic pattern of weights was not observed. In Experiment 2, the accuracy for detecting a peak in the temporal loudness profile of a level-fluctuating noise was measured for the level increment presented at expected and unexpected times. The rhythmic context again caused a more pronounced temporal weighting profile. However, the accuracy was neither significantly increased for expected target segment positions nor decreased for unexpected target segment positions, relative to the condition without a rhythmic context sequence.

*Keywords:* rhythmic attention, perceptual weights, temporal weighting, loudness, intensity resolution

The work of Mari Riess Jones and coworkers has demonstrated that temporal expectancies generated by a rhythmic context (e.g., an isochronous tone sequence) influence the processing of future events (e.g., Jones, 1976; Jones & Boltz, 1989). Several studies showed that a rhythmic context has an effect on accuracy in inherently temporal tasks like the discrimination of temporal intervals (e.g., Barnes & Jones, 2000; Jones & McAuley, 2005) or the production of melodic durations (Jones & Boltz, 1989). For example, Large and Jones (1999) found that time intervals ending at a rhythmically expected time are more accurately judged than time intervals ending unexpectedly, indicating that attention can be directed not only spatially (cf. Posner, 1980) but also temporally (Barnes & Jones, 2000; Boltz, 1993). The temporal dynamics have been described as an “attentional pulse” or “attending rhythm” (Large & Jones, 1999; McAuley & Kidd, 1998) that directs attention to future points in time. Barnes and Jones (2000) argued that the direction of attention by a rhythmic context can be described as a stimulus-driven, bottom-up involuntary process, much as the spatial direction of attention by so-called peripheral visual cues (e.g., Posner, 1980; for a recent discussion, see Jones, Johnston, & Puente, 2006).

Given the claim that temporally directed attention is a rather universal phenomenon (Barnes & Jones, 2000; Jones, Moynihan, MacKenzie, & Puente, 2002; Large & Jones, 1999), the question arises as to whether evidence for rhythmic attention can also be found for tasks not involving time judgements. In fact, Jones et al.

(2002) reported such evidence for a pitch discrimination task. Participants compared the pitch of a comparison tone with the pitch of a standard tone, with a rhythmically structured tone sequence interpolated between the standard tone and the comparison tone. Accuracy was superior if the comparison tone was presented at a rhythmically expected time, that is, if the interonset interval (IOI) between the last interpolated tone and the comparison tone was identical to the IOIs in the interpolated sequence. Compatible results were obtained by Jones et al. (2006).

According to the attentional pulse hypothesis, rhythmic attention is not restricted to the time and pitch dimensions. The first aim of the present study was to answer the question of whether the findings of Jones et al. (2002) can be generalised to auditory intensity, which besides pitch is one of the most important perceptual dimensions of auditory events.

The second issue addressed by the present study was the method for measuring attention. In the experiments by Jones et al. (2002, 2006), the accuracy for detecting a difference in pitch was measured for brief tones presented at various temporal positions, and differences in the amount of attention directed to the individual temporal positions were deduced from differences in accuracy. The point in time receiving the largest amount of attention was defined as the point in time for which the accuracy was maximal. Common as this method is in the study of attention, there exists an alternative approach, which seems more appropriate for studying the direct perceptual consequences of rhythmic attention. In this study's Experiment 1, the influence of individual temporal portions of a longer sound on global loudness judgements (i.e., subjective magnitude) was measured by estimating the perceptual weight (Berg, 1989) each element received. Participants evaluated the overall loudness of a noise fluctuating in level. In the rhythm condition, the noise was preceded by a short isochronous sequence of pure tones (rhythmic context sequence). Perceptual weight analysis was used (e.g., Ahumada & Lovell, 1971; Berg, 1989;

---

I am grateful to Tina Plank for valuable suggestions concerning the experiments. I thank Simon Grondin, Sylvia Jankowiak, J. Devin McAuley, Yoshitaka Nakajima, and Tina Plank for helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Daniel Oberfeld, Department of Psychology, Johannes Gutenberg–Universität Mainz, 55099, Mainz, Germany. E-mail: oberfeld@uni-mainz.de

Lutfi, 1992), for which a random variation is imposed on the individual components of a complex stimulus on each trial. In the present study, the components were the temporal segments constituting the level-fluctuating noise. If the participant assigns a large weight to a particular component—that is, if attention is directed to the component (Berg, 1990)—there will be a strong correlation between the component's random level perturbation and the participant's response. If on the other hand the component is unimportant for the decision, the responses will be statistically independent of the random variation (Kortekaas, Buus, & Florentine, 2003; Lutfi, 1995; Richards & Zhu, 1994). According to the attentional pulse hypothesis, temporal segments in an auditory stimulus to which attention is directed should have a larger influence on loudness perception than temporal segments presented at rhythmically unexpected points in time. Estimating the weights assigned to the individual temporal segments allowed testing of the hypothesis that the involuntary direction of attention by a rhythmic context sequence results in a rhythmic pattern of weights because the weights assigned to temporal segments presented at rhythmically expected times (i.e., at points in time lying on a continuation of the rhythmic context sequence) should be larger than the weights assigned to neighbouring “off-beat” segments.

It should be noted that the “maximum accuracy = maximum attention” and the “maximum weight = maximum attention” approaches will not necessarily produce the same results. For example, Kortekaas et al. (2003) found that in an intensity discrimination task involving multitone complexes, participants placed a larger weight on the level information from the higher frequency components. In contrast, Green (1993) reviewed studies in which the accuracy for detecting an intensity increment on one component of a multitone complex was measured and concluded that accuracy is highest for the central frequency component. This means that the participants in the experiments by Kortekaas et al. attended to frequencies for which their sensitivity was lower than for some of the remaining frequencies. In other words, if the amount of attention had been deduced from accuracy rather than from perceptual weights, a different pattern of results would have emerged. Therefore, measuring the influence of temporal segments coinciding with expected and unexpected time points on subjective magnitude (loudness) rather than measuring attention on the basis of accuracy in a discrimination task can provide new information on rhythmic attention.

Experiment 2 complemented Experiment 1 in also studying the effect of a rhythmic context on the processing of auditory intensity, but with a focus on accuracy, as in previous studies. The experimental manipulation involved both a condition in which the participant could profit from using the rhythmic context sequence because it was a valid cue (Posner, 1980) to the temporal position of the target segment and a condition in which the rhythm was an invalid cue. To make the stimuli in the two experiments comparable, a target element embedded in a longer sound was presented rather than an isolated target element. The task was to detect a peak in the loudness profile of a level-fluctuating noise, that is, a level increment imposed on the target segment (temporal profile listening task; cf. Plank & Ellermeier, 2003). Two rhythmic context sequences, either correctly indicating the temporal position of the target segment (on-beat condition) or priming a point in time 100 ms earlier than the onset of the target segment (off-beat condition),

were contrasted. Relative to the control condition without rhythmic context, the accuracy was predicted to be superior if the onset of the target segment coincided with a rhythmically expected time and to be impaired if the onset of the target segment occurred earlier than the expected time. Perceptual weights were also estimated. In the on-beat condition, the temporal cue to the onset of the target segment provided by the rhythm was predicted to result in a pattern of weights closer to the optimal weights than in the control condition. In the off-beat condition, the nontarget segment coinciding with the expected time point was predicted to receive a larger, nonoptimal weight than in the control condition because of the involuntary direction of attention to its onset.

### Experiment 1: Global Loudness Judgements

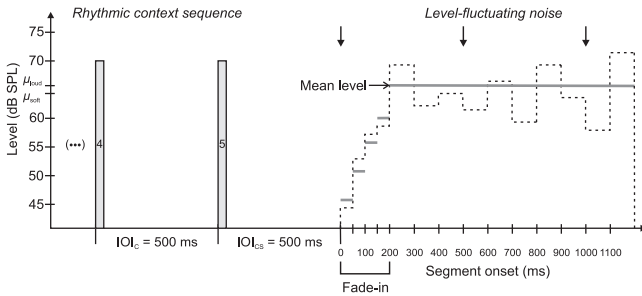
Experiment 1 studied the effects of a rhythmic context on the temporal weighting of loudness. In the control condition, participants judged the overall loudness of a level-fluctuating noise consisting of 14 contiguous temporal segments. In the rhythm condition, the noise was preceded by an isochronous rhythmic context sequence, which was predicted to modulate the temporal weighting of loudness, inducing a rhythmic pattern of weights with peaks at expected times.

It is known that for global loudness judgements, the initial portion of a longer stimulus receives a greater weight than the later parts (Ellermeier & Schrödl, 2000; Pedersen & Ellermeier, 2004). This pattern is observed even if each segment provides the same amount of information concerning the correct response, so that the optimum response strategy would be to weight the level information provided by each temporal segment uniformly (Berg, 1989). Such a primacy effect could reduce the sensitivity of the procedure for detecting a change in the weights caused by the rhythmic context. Therefore, a gradual fade-in was imposed on the first four noise segments, which has been shown to result in participants weighting the level information from the unattenuated segments more uniformly (Oberfeld & Plank, 2005). The weights assigned to the fade-in segments were expected to be close to zero, and the maximum weight was expected for the first unattenuated segment (delayed primacy effect; Oberfeld & Plank, 2005).

### Method

*Participants.* Seven psychology students (4 women, 3 men; age range = 21–24 years) participated in the experiment voluntarily. They either received course credit or were paid for their participation. All participants reported normal hearing; detection thresholds in the right ear were better than 10 dB HL (hearing level) at all octave frequencies between 0.5 and 8 kHz. Only 2 participants had experience in comparable psychoacoustic tasks.

*Stimuli and apparatus.* The level-fluctuating noise consisted of a sequence of 14 wide-band noise segments, which were presented contiguously (i.e., with no pauses between them). The duration of the first four segments was 50 ms, and the duration of the remaining segments was 100 ms, resulting in a total duration of 1,200 ms. On each trial, the sound pressure levels (SPLs) of the 14 noise segments were drawn independently from a normal distribution with mean  $\mu = 65$  dB SPL and standard deviation  $\sigma = 2$  dB. To produce the fade-in, the levels of the first four segments were subsequently attenuated by subtracting 20, 15, 10, and 5 dB,



**Figure 1.** Experiment 1: Schematic depiction of a trial. Participants judged the overall loudness of a 1,200-ms level-fluctuating wide-band noise composed of 14 temporal segments (dashed line). On each trial, the level of each segment was drawn independently from either a soft distribution (mean  $\mu_{\text{Soft}} = 64.5$  dB SPL) or a loud distribution (mean  $\mu_{\text{Loud}} = 65.5$  dB SPL). Subsequently, a fade-in was imposed on the first four segments. The thick grey lines show the mean segment levels for a loud trial. In the rhythm condition, the noise was preceded by a rhythmic context sequence consisting of five tones (gray rectangles; only Tones 4 and 5 are shown) presented with an interonset interval of 500 ms. The arrows indicate the rhythmically expected times. In the control condition, no rhythm was presented.

respectively (see Figure 1). Each of the so-constructed stimuli was then randomly chosen to be a soft or a loud trial. A level increment of  $\Delta L/2 = 0.5$  dB was added to each segment in a loud trial. The same value was subtracted from each segment in a soft trial.

The rhythmic context sequence consisted of five 1-kHz tone bursts with a duration of 30 ms including 5-ms  $\cos^2$  ramps. The reason for using pure tones rather than noise bursts was to prevent the participants from judging the loudness of the context sequence instead of judging the loudness of the noise only. On the basis of loudness matches between the pure-tone context sequence and the level-fluctuating noise obtained in a pretest, a level of 81.6 dB SPL was used for the context sequence for all participants. The interonset interval in the rhythmic context sequence ( $\text{IOI}_C$ ) was 500 ms. The interval between the onset of the last tone of the context sequence and the onset of the first noise segment ( $\text{IOI}_{CS}$ ) was also 500 ms. Thus, the rhythmically expected times were located at the onsets of the 1st segment, the 8th segment (500 ms after onset of the first segment), and the 13th segment. Because of the predicted very small weight assigned to the first, attenuated segment, the rhythmic context was expected to have no effect on this particular weight, even though the segment coincided with a rhythmically expected time point.

The stimuli were generated digitally, played back via two channels of an RME ADI/S D/A converter ( $f_s = 44.1$  kHz, 24-bit resolution), attenuated (two TDT PA5s), buffered (TDT HB7), and presented diotically via Sennheiser HDA 200 headphones. The experiment was conducted in a single-walled sound-insulated chamber. Participants were tested individually.

**Procedure.** The stimuli were presented in a one-interval, two-alternative forced-choice procedure (absolute identification task; see Macmillan & Creelman, 2005). Each trial was randomly chosen to be a soft or a loud trial with equal probability. The participants' task was to decide whether they had been presented a soft or a loud noise. Responses were collected via two buttons on a numeric keypad. No feedback was provided. The next trial followed the response after an intertrial interval of 2 s.

There were two experimental conditions. In the rhythm condition, the noise was preceded by the rhythmic context sequence. In the control condition, the context sequence was omitted. Each participant received 1,000 trials in each of the two conditions (within-subjects design). As in previous studies (Barnes & Jones, 2000; Jones et al., 2002), participants were instructed to ignore the rhythm because the aim of the experiment was to study an involuntary direction of attention by the rhythmic context. Only one condition was presented within a block of 50 trials. Blocks containing the control condition and blocks containing the rhythm condition were presented in alternating order. The experiment started with a practice session, followed by four experimental sessions. A testing session was completed in approximately 1 hr, with a brief rest period.

**Data analysis.** The trial-by-trial data obtained for each participant in each of the two experimental conditions were analysed separately to estimate the relative perceptual weight with which each of the temporal segments had contributed to the decision of the participant. Multiple logistic regression was used to estimate the influence of the level of each individual temporal segment on the response of the listener (cf. Agresti, 2002; Hosmer & Lemeshow, 1989; Pedersen & Ellermeier, 2004).<sup>1</sup> The dichotomous response ("loud" or "soft") served as the dependent variable, and the randomly varied SPLs of the 14 temporal segments served as the independent variables (predictors). In the regression model, the dichotomous response was predicted by a linear combination of the 14 segment levels. Each trial entered the analysis as one observation. For each of the 7 (participants)  $\times$  2 (condition) logistic regression models, the fit was better than the fit of the null model containing only the intercept (likelihood-ratio chi-square test,  $p < .0001$ ; cf. Agresti, 2002). A measure for the predictive power of a logistic regression model is the area under the receiver operating characteristic (ROC) curve. This measure provides information about the degree to which the predicted values are concordant with the observed outcome (cf. Agresti, 2002; see Hosmer & Lemeshow, 1989, for a critical discussion). An area of 0.5 and 1.0 means no or a perfect accuracy of the prediction, respectively. Across the 7 (participants)  $\times$  2 (condition) fitted logistic regression models, the area under the ROC curve ranged between 0.75 and 0.90 ( $M = 0.82$ ,  $SD = 0.040$ ), indicating a reasonably good predictive power.

The regression coefficients were taken as the weight estimates. For a given segment, a regression coefficient equal to zero means that the level of the segment had no influence at all on the decision to judge the noise as being either soft or loud. A regression coefficient greater than zero means that the probability to respond that the loud noise was presented increased with the level of the segment. A regression coefficient smaller than zero indicates the opposite relation between the level of the segment and the prob-

<sup>1</sup> Note that several alternative methods have been used to estimate relative perceptual weights, all of which are based on a similar stochastic decision model (Berg, 1989; Lutfi, 1995; Richards & Zhu, 1994). The logistic multiple regression approach adopted in the present study uses the same rationale as trial-by-trial correlational analysis (Lutfi, 1995; Richards & Zhu, 1994), but a logistic link function (McCullagh & Nelder, 1989) and a maximum-likelihood method instead of a least-squares approach for fitting the model.

ability to respond that the loud noise was presented. The 14 weights estimated for each participant in each experimental condition were normalised such that the sum of the absolute values was unity (see Kortekaas et al., 2003). Note that the analysis is concerned with relative weights, for example, the question of whether the level of the first noise segment has a greater influence on the decision than the level of the fifth segment, regardless of the general strength of the association between the segment levels and the decision (cf. Berg, 1989).

Sensitivity and response bias were analysed in terms of a signal detection theory model assuming equal-variance Gaussian distributions (Green & Swets, 1966).<sup>2</sup>

For the statistical analysis, repeated-measures analyses of variance (ANOVA) based on a univariate approach with Huynh-Feldt adjusted degrees of freedom were conducted (cf. Keselman, Algina, & Kowalchuk, 2001). An alpha level of .05 was used for all statistical tests.

## Results

**Perceptual weights.** As can be seen in Figure 2, the mean weights in both the control condition and the rhythm condition reflected the expected delayed primacy effect (Oberfeld & Plank, 2005). The weights assigned to the four segments constituting the fade-in were small, and the maximum weight was observed for the first unattenuated segment. All participants showed this pattern. No recency effect was observed.

Contrary to the predictions of the rhythmic attention hypothesis, the rhythmic context resulted in a rhythmic pattern of weights for only 2 of the 7 participants. In the mean data (see Figure 2), there was no evidence for peaks at expected times. However, in the rhythm condition the peak at the first unattenuated segment (onset = 200 ms) was sharper than in the control condition. An ANOVA with the within-subject factors segment and experimental condition provided partial support for this more pronounced delayed primacy effect by indicating a marginally significant Segment  $\times$  Condition interaction,  $F(13, 78) = 1.8, p = .062, \eta^2 = .24$ . As expected, there was also a significant effect of segment,  $F(13, 78) = 30.1, p = .001, \eta^2 = .83$ . Because of the normalisation of the weights, there was no effect of experimental condition,  $F(1, 6) = 2.3$ .

**Sensitivity and bias.** Mean sensitivity as measured by the signal detection theory index  $d'$  (cf. Macmillan & Creelman, 2005) was 1.05 ( $SD = 0.20$ ) in the control condition and 1.06 ( $SD = 0.20$ ) in the rhythm condition. The rhythmic context had no significant effect on sensitivity,  $t(6) = 0.35$ .

The response bias was also unaffected by the rhythmic context: The mean value of the bias parameter  $c$  was  $-0.20$  ( $SD = 0.24$ ) in the control condition and  $-0.21$  ( $SD = 0.27$ ) in the rhythm condition,  $t(6) = 0.13, p > .8$ . Averaged across conditions,  $c$  was  $-0.206$  ( $SD = 0.24$ ); that is, the participants tended to respond that the loud noise had been presented. A one-sample  $t$  test showed that mean  $c$  was significantly smaller than 0,  $t(6) = -2.26, p = .032$  (one-tailed).

## Discussion

The rhythmic context had an effect on the temporal weighting of loudness: In the rhythm condition, participants tended to produce

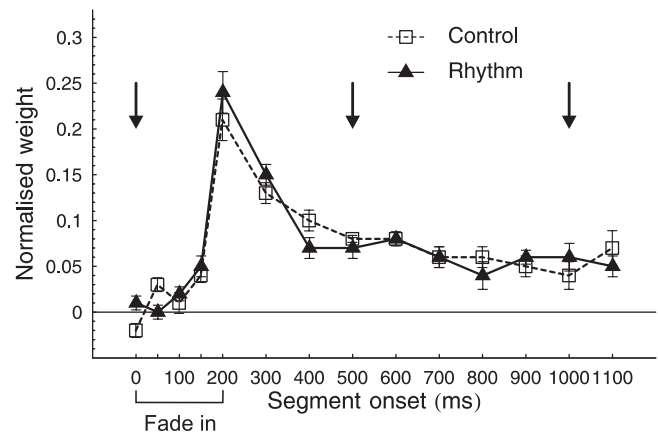


Figure 2. Experiment 1: Mean normalised perceptual weights for the 14 segments. Boxes = control condition; triangles = rhythm condition. Arrows indicate rhythmically expected times. Error bars represent  $\pm 1$  standard error of the mean of the seven individual estimates.

a sharper weighting profile. A potential explanation for this effect is that the participants used the rhythm to direct their attention to the first unattenuated segment. Even though the latter segment was not presented at an expected time, the temporal interval between the onset of the last tone of the context sequence and the onset of the first unattenuated segment was constant. Therefore, some timing mechanism may have been used for directing attention to the latter segment. To examine this possibility, a single context tone could be presented, either at a fixed position relative to the onset of the first noise segment or with its temporal position varied randomly. In addition, presenting a temporally irregular sequence of context tones should reduce the effect, as in the study by Jones et al. (2002). An alternative explanation for the sharper temporal weighting profile is proposed in the General Discussion.

Contrary to the hypothesis based on dynamic attending theory (Large & Jones, 1999), the rhythmic context did not result in a significant rhythmic modulation of the temporal weights. A potential reason is that the optimum strategy in Experiment 1 was to weight the level information from each segment uniformly because

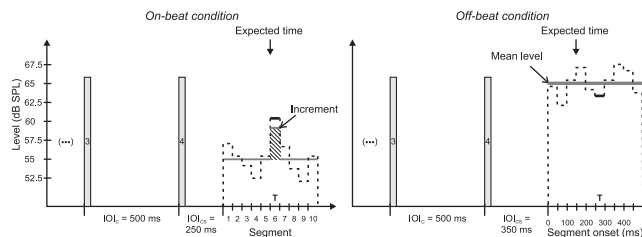
<sup>2</sup> Trials in which the arithmetic mean of all segment levels before the attenuation of the first four segments was greater than the expected value of the mean segment level across soft and loud trials (65 dB SPL) were taken as the “signal” and the remaining trials as the “noise.” For each observer, the proportion of hits, that is, the proportion of “loud” responses on signal trials, and the proportion of false alarms (FA) were computed for each of the two conditions, aggregated across blocks. The proportions were converted to standard normal deviates ( $z_{Hit}$  and  $z_{FA}$ ) using the relation  $z = F^{-1}(P)$ , where  $z$  is the standard normal deviate corresponding to the proportion  $P$  and  $F^{-1}(x)$  is the inverse cumulative density function of a normal distribution with mean 0 and standard deviation 1, evaluated at the point  $x$ . The sensitivity index is then given by  $d' = z_{Hit} - z_{FA}$ . It denotes the standardized difference between the means of the two hypothesized distributions. Response bias was analyzed in terms of  $c = -1/2(z_{Hit} + z_{FA})$ , which denotes the location of the decision criterion relative to the cross-over point of the signal and the noise distribution (Macmillan & Creelman, 2005). Unbiased responding corresponds to  $c = 0$ . Positive values of  $c$  indicate that the participant tended to respond that the soft noise had been presented.



each segment provided the same amount of information concerning the correct response (Berg, 1989). Adopting a rhythmic pattern of weights would thus have impaired performance. It is therefore conceivable that the participants ignored the rhythm to maximise performance. Still, the results are not in accordance with the assumption brought forward by Barnes and Jones (2000) that the temporal direction of attention by a rhythmic context sequence is an involuntary process.

## Experiment 2: Temporal Profile Listening

As discussed above, the absence of a rhythmic pattern in the perceptual weights observed in Experiment 1 may have been due to the participants successfully ignoring the rhythm. In addition, it is possible that although a rhythmic context has no effect on the perceptual weights, it does influence the accuracy in an intensity discrimination task, analogously to the effects on frequency resolution reported in previous studies (Jones et al., 2002, 2006). For these reasons, a temporal profile listening task (cf. Plank & Ellermeier, 2003) that allowed measurement of both accuracy and perceptual weights was used in Experiment 2. The task required detecting a peak in the temporal loudness profile of a level-fluctuating noise, that is, a level increment imposed on one of the temporal segments (see Figure 3). To prevent the participants from basing their responses on the overall level, a random level rove of  $\pm 10$  dB was introduced (cf. Green, 1988). The optimum strategy in such a task is to respond “increment present” if the level of the target segment is higher than the average level of the nontarget segments, ignoring the variability in overall level (Berg & Green, 1990). This means that the participant is required to differentially process the level information from the target segment and the level information from all remaining segments. Although it has been demonstrated that the temporal resolution of the human auditory system is sufficiently high to analyse individual segments with respect to their loudness even if the segment duration is only 10 ms (Plank & Ellermeier, 2003, 2004), the task should become easier if a rhythmic context can be used to determine the temporal position of the target segment. Therefore, the task was comparable to the pitch discrimination experiment by Jones et al. (2002) in that a participant could profit from using contextual timing information in the on-beat condition in which the onset of the target segment coincided with an expected time point (left panel in Figure 3). In the on-beat condition, the sensitivity was thus expected to be higher than in the control condition without rhythmic context. In the off-beat condition in which the onset of the target segment occurred 100 ms later than the expected time point (right panel in Figure 3), the sensitivity was expected to be impaired. The on-beat condition additionally provides a test for the hypothesis that in Experiment 1 the pattern of weights was incompatible with an effect of rhythmic attention only because the listeners ignored the rhythm. Participants were expected to be more successful in adopting the optimum set of temporal weights than in the control condition. Although the results from Experiment 1 indicated that a context rhythm does not result in an involuntary direction of attention in a global loudness judgement task, the off-beat condition can be used to examine whether an involuntary direction of rhythmic attention plays a role in a temporal profile listening task. In this condition, the direction of attention to the segment on the rhythmically expected time point (Segment 4, occurring 100 ms



**Figure 3.** Experiment 2: Schematic depiction of a trial. The task was to detect a level increment on the centre segment of a 500-ms level-fluctuating wide-band noise composed of 10 segments (dashed line). A random-level perturbation was imposed on each segment. The thick grey lines show the mean segment levels. In the left panel, the noise contains the increment (hatched rectangle) on the target segment (Segment 6; denoted by the  $T$  and the thick black line). A random-level rove was used to prevent the participants from responding on the basis of overall loudness. In the two rhythm conditions, the noise was preceded by a rhythmic context sequence (CS) consisting of four tones (grey rectangles) presented with an interonset interval (IOI) of 500 ms (only Tones 3 and 4 are shown). Left panel: On-beat condition ( $\text{IOI}_{\text{CS}} = 250$  ms). The target segment onset coincided with a rhythmically expected time point (indicated by the arrow). Right panel: Off-beat condition ( $\text{IOI}_{\text{CS}} = 350$  ms). The onset of the target segment onset occurred 100 ms later than the rhythmically expected time.

earlier than the increment) should result in a stronger deviation from the optimum set of weights.

## Method

**Participants.** Six volunteers (4 women, 2 men; age range = 22–33 years) participated in the experiment. I was one of the participants. The remaining participants were psychology students participating for partial course credit or payment. All participants reported normal hearing; detection thresholds in the right ear were better than 15 dB HL at all octave frequencies between 0.5 and 8 kHz. Three of the participants had already participated in Experiment 1.

**Stimuli and apparatus.** The level-fluctuating noise consisted of 10 50-ms wide-band noise segments presented contiguously. The shorter overall duration than in Experiment 1 was selected so that the target segment temporally centred in the noise could be presented on the first rhythmically expected time point following the last tone of the context sequence in the on-beat condition. On each trial, the SPLs of the 10 noise segments were drawn independently from a normal distribution with  $\mu = 57$  dB SPL and  $\sigma = 2$  dB. With an a priori probability of .5, a 4-dB level increment was added to the level of the sixth segment, the target segment (see Figure 3). A random level rove of  $\pm 10$  dB was imposed on all segments: On each trial, a random number was sampled from a uniform distribution with a minimum of  $-10$  dB and a maximum of 10 dB. This value was added to each of the 10 segment levels.

In the two rhythm conditions, an isochronous rhythmic context sequence preceded the noise. The sequence consisted of four 1-kHz tone bursts with a duration of 30 ms including 5-ms  $\cos^2$  ramps. The  $\text{IOI}_{\text{C}}$  was 500 ms. In the on-beat condition (left panel in Figure 3), the  $\text{IOI}_{\text{CS}}$  between the onset of the last tone of the rhythm and the onset of the first noise segment was 250 ms. Thus, the onset of the target segment coincided with a rhythmically expected time point. In the off-beat condition (right panel in Figure

3),  $IOI_{CS}$  was 350 ms, such that the target segment occurred 100 ms later than the expected time. The same apparatus as in Experiment 1 was used.

**Procedure.** The stimuli were presented in a one-interval, two-alternative forced-choice procedure. The participants' task was to decide whether the level increment on the target segment had been present or absent. They were informed that the increment would be present or absent with equal probability. Responses were collected via two buttons on a numeric keypad. Visual trial-by-trial feedback was provided throughout the experiment. A response was evaluated as being correct if the participant correctly indicated whether the increment had been present or absent. Participants were informed that because of the random sampling of the segment levels, feedback could be counterintuitive in some trials (cf. Kortekaas et al., 2003). The next trial followed the response after an intertrial interval of approximately 2 s.

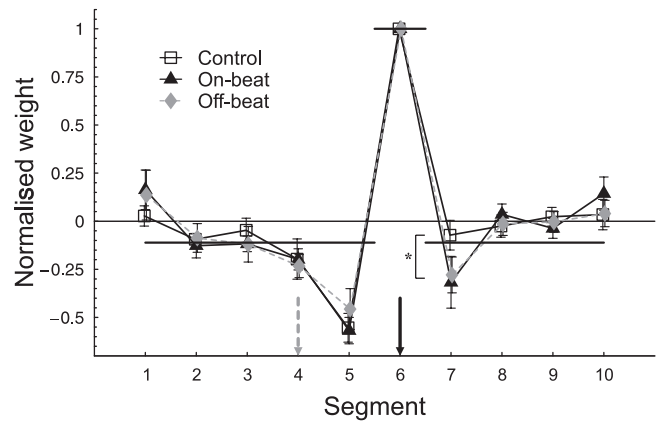
Each participant received 1,000 trials in each of the three experimental conditions (within-subjects design). Participants were informed that in some of the blocks, a rhythm consisting of four tones would precede the noise, and they were instructed to ignore the rhythm. Following two practice sessions, five experimental sessions were run in which participants first received a practice block. Four consecutive blocks of each of the three experimental conditions were then presented. Only one condition was presented in each block of 50 trials. The reason for not interleaving the three conditions within a block was to maximise the probability that participants would use the rhythm to direct their attention to the target segment in the on-beat condition. The order of conditions was balanced across sessions. A testing session was completed in approximately 1 hr, with a brief rest period.

**Data analysis.** A multiple logistic regression model was used to estimate the perceptual weights for each participant in each of the three experimental conditions. The dichotomous response ("increment present" or "increment absent") served as the dependent variable, and the randomly varied SPLs of the 10 temporal segments served as the independent variables. For each of the 6 (participants)  $\times$  3 (condition) models, the fit was better than the fit of the null model containing only the intercept (likelihood-ratio chi-square test,  $p < .0001$ ). The area under the ROC curve ranged between .63 and .81 ( $M = .74$ ,  $SD = .064$ ).

The regression coefficients were taken as the weight estimates. The resulting weights obtained for each participant and each experimental condition were normalised by setting the weight of the signal component to unity and scaling the magnitude of the remaining weights relative to this unit weight (Berg & Green, 1990).

## Results

**Perceptual weights.** In the profile listening task used in Experiment 2, if the signal component is assigned a weight of unity and if the variance of the random perturbation imposed on each segment is small relative to the variability in overall level induced by the level rove, then the optimum weight for each nontarget segment is approximately  $-1/(n - 1)$ , where  $n$  is the number of segments (Berg & Green, 1990). As can be seen in Figure 4, the mean weights were close to the ideal weight of  $-1/9$  for Segments 2–4. The weights of the two segments directly neighbouring the target segment were more negative than the optimum weights.



**Figure 4.** Experiment 2: Mean normalised perceptual weights. The weight of the target segment (Segment 6) was constrained to equal unity. Boxes = control condition; triangles = on-beat condition; grey diamonds = off-beat condition. The asterisk indicates the significantly more negative weights assigned to Segment 7 in the two rhythm conditions, relative to the control condition. The arrows indicate the expected times (solid arrow = on-beat; dashed grey arrow = off-beat). The horizontal lines show the ideal weights. Error bars represent  $\pm 1$  standard error of the mean of the six individual estimates.

Segment 1 and the final three segments received weights close to zero. This pattern of weights indicates that the participants used a "profile listening" strategy (cf. Green, 1988); it is compatible with previous results for the same type of task (Plank & Ellermeier, 2003).

The average weight of Segment 7 directly following the target segment was significantly more negative in the two rhythm conditions ( $M = -0.30$ ,  $SD = 0.27$ ) than in the control condition ( $M = -0.07$ ,  $SD = 0.19$ ),  $t(5) = -2.71$ ,  $p = .042$  (two-tailed). In the off-beat condition, the weight assigned to Segment 4, which coincided with the expected time point, did not differ from the weight in the control condition,  $t(5) = 0.42$ .

An ANOVA with the within-subjects factors segment and condition was conducted for the weights estimated for the nine nontarget segments. The reason to exclude the target segment was that because of the normalisation, this weight was unity for all conditions. There was a significant effect of segment,  $F(8, 40) = 8.9$ ,  $p = .001$ ,  $\eta^2 = .64$ . This result confirms the observation of nonuniform weights assigned to the nine nontarget segments. Neither the effect of condition,  $F(2, 10) = 1.7$ , nor the Segment  $\times$  Condition interaction,  $F(16, 80) = 1.1$ , was significant.

**Sensitivity and bias.** Mean sensitivity ( $d'$ ) was 0.69 ( $SD = 0.25$ ), 0.69 ( $SD = 0.25$ ), and 0.73 ( $SD = 0.30$ ) in the control condition, the on-beat condition, and the off-beat condition, respectively. A repeated-measures ANOVA confirmed that the experimental condition had no significant effect on sensitivity,  $F(2, 10) = 0.56$ , unlike in previous studies measuring accuracy as a function of the temporal position of the target element relative to a rhythmic context.

There was virtually no response bias, presumably because of the availability of trial-by-trial feedback. The mean value of the bias parameter  $c$  was 0.03 ( $SD = 0.11$ ) in the control condition, 0.06 ( $SD = 0.17$ ) in the on-beat condition, and 0.06 ( $SD = 0.22$ ) in the

on-beat condition. The effect of the experimental condition on  $c$  was not significant,  $F(2, 10) = 0.125$ .

### Discussion

The rhythmic context induced a systematic change in the perceptual weights. In the two rhythm conditions, participants adopted a stronger, nonoptimal contrast between the level of the target segment and the level of the directly following segment: The weight assigned to the latter segment was more negative in the two rhythm conditions than in the control condition. In other words, both rhythmic context sequences resulted in a more pronounced temporal weighting profile, just as in Experiment 1.

Despite the fact that the structure of the task was comparable to the pitch discrimination experiment by Jones et al. (2002) in that a target item occurred either at an expected or at an unexpected time, the sensitivity was not significantly greater in the on-beat condition than in the control condition. There was also no significant decrease in sensitivity in the off-beat condition. The hypothesis that in the off-beat condition a nonoptimal set of weights would be induced because the segment presented at the expected time would be mistaken as the target segment was not supported, either. An explanation for these findings could be that the participants used the trial-by-trial feedback to compensate for the wrong temporal position indicated by the rhythmic context sequence in the off-beat condition, for example by adopting a strategy of “lagging,” that is, of voluntarily attending to a point in time somewhat later than the rhythmically expected time. Such a strategy may have been advocated by the blocked presentation of the three different conditions.

### General Discussion

Previous work on rhythmic attention was extended in two aspects. First, the effects of a rhythmic context were studied for another, not inherently temporal dimension: auditory intensity. Second, although previous studies showed an effect on accuracy compatible with the attentional pulse hypothesis, Experiment 1 examined the perceptual consequences of rhythmic attention more directly by measuring the influence of temporal portions of a longer sound coinciding with expected and unexpected time points on global loudness. Perceptual weight analysis was used (Berg, 1989; Richards & Zhu, 1994). In Experiment 2, the accuracy in a temporal profile listening task was measured as a function of the target segment position relative to the rhythmic expectancies induced by a rhythmic context sequence. In addition, perceptual weights were estimated.

In both experiments, an isochronous rhythmic context sequence preceding the to-be-judged stimulus induced a change in the perceptual weights assigned to the temporal segments constituting the stimulus. In Experiment 1, the rhythm resulted in a sharper peak at the first unattenuated segment. In Experiment 2, participants adopted a nonoptimally strong contrast between the target segment containing the increment and the following segment, regardless of whether the increment was presented at an expected or an unexpected time. An—admittedly speculative—explanation for these findings would be that the presentation of a rhythmically structured context induces a mode of intensity processing in which the listener pays special attention to the temporal dynamics of a sound,

that is, the evolution of the sound over time. Put differently, it could be argued that the listeners were searching for a continuation of the rhythm, that is, for “accented” events standing out from their vicinity (Fraisse, 1982), as for instance onsets or peaks in level. In Experiment 1, the listeners could therefore have attended to the systematic increase in level occurring during the first 300 ms because of the fade-in, and the first one or two unattenuated segments could have sounded especially salient, explaining the more pronounced delayed primacy effect. In Experiment 2, the temporal dynamics of the noise were most pronounced in the vicinity of the level increment imposed on the target segment, whereas the beginning and ending of the noise sounded more uniform. The level increment represented an accent embedded within the noise. Again, if a context rhythm induced attention to the temporal dynamics (e.g., accented events), it is conceivable that the listeners directed their attention to the target segment and to the two neighbouring segments. In fact, if a listener bases his or her decision on the levels of only the latter three segments, and if the weight assigned to the target segment is again set to unity, then the optimal weight for the two neighbouring segments would be  $-1/(3 - 1) = -0.5$  (Berg & Green, 1990). Therefore, the significant shift of the weight assigned to the segment directly following the target segment induced by the two context sequences in Experiment 2 (see Figure 4) would have resulted in a set of weights closer to the optimum weights. This “rhythm mode” explanation could be tested by using the same procedure as in Experiment 1 and varying the delay ( $IOI_{CS}$ ) between the context sequence and the noise. If the sharper temporal weighting profiles observed in the presence of a rhythmic context were indeed because of the induction of a rhythm mode of intensity processing, then a rhythmic context sequence should cause the effect regardless of its timing relative to the noise onset. Therefore, the delayed primacy effect should generally be more pronounced than in the control condition without rhythmic context, independent of the duration of the critical interval ( $IOI_{CS}$ ).

The results of the present study are in several aspects incompatible with the hypothesis that a rhythmic context sequence directs attention to rhythmically expected time points (Large & Jones, 1999). In Experiment 1, no rhythmic pattern of weights was observed. Thus, the perceptual effects of a rhythmic context sequence, as reflected in the perceptual weights, seem to be smaller than the effect on accuracy reported in previous studies (Jones et al., 2002, 2006). This conclusion needs to be qualified, however, because adopting a rhythmically structured set of weights would have resulted in a drop in performance, so that it is conceivable that the participants successfully tried to ignore the rhythm. Still, the results are reason to question Barnes and Jones’s (2000) assumption that rhythmic attention is a stimulus-driven process that participants should not be able to control voluntarily.

Even more surprising was the absence of an effect of the rhythmic context on the accuracy in the profile listening task in Experiment 2. The hypothesis was that sensitivity would be greater if the target segment was presented at a rhythmically expected time point, just as frequency resolution was higher in this situation in the experiments by Jones et al. (2002, 2006). A potential explanation for the lack of an effect in Experiment 2 could be that the onset of the noise was temporally unexpected because unlike in Experiment 1 it did not coincide with a continuation of the rhythmic context sequence but occurred earlier in both the on-beat and

off-beat conditions. As noted above, the stimuli were constructed in this way to present the target segment on the first expected time point primed by the context sequence in the on-beat condition. Now, according to the theory of dynamic attending (Large & Jones, 1999; McAuley & Jones, 2003), an unexpected event is likely to perturb the ongoing attentional rhythm and to modify the temporal expectations established by the context rhythm. McAuley and Jones (2003) showed that listeners may come close to phase resetting in response to temporally unexpected onsets, especially for onsets that are early. In the on-beat condition, however, the phase modification should have been moderate because the onset of the noise coincided with a double metrical subdivision of the beat induced by the context sequence (cf. Large & Kolen, 1994; Repp, 2006). For this reason, the relative phase of the noise onset was  $\Phi = 0.5$  (antiphase). Now, in the dynamical oscillator models formulated by Large and Kolen (1994) and Large and Jones (1999), the effect of an out-of-phase event on the phase of the oscillator (termed *phase coupling*) is zero in the antiphase case (see Figure 7 in Large & Jones, 1999). According to the theory of dynamic attending by Large and Jones (1999), the “syncopated” onset of the noise in the on-beat condition of Experiment 2 should thus have had no or only a small effect on the temporal expectations established by the rhythmic context sequence. Note that Repp (2004) reported results indicating that the critical variable may be not the relative phase but the absolute temporal separation between the expected onset and an interpolated distractor event. Still, Repp (2004) found virtually no effect if the delay between the onset of the distractor tone and the expected onset exceeded 200 ms. Therefore, his data also suggest only a small effect of the early noise onset on the attentional rhythm in the on-beat condition, where the delay was 250 ms. Nevertheless, it cannot be precluded that the unexpected noise onset in Experiment 2 weakened the temporal expectations induced by the rhythmic context.

A potential issue in both experiments was the use of a rhythmic context sequence consisting of pure tones combined with noise target stimuli. The spectral difference might have reduced the effect of the rhythm, for instance because participants organised the two different sounds into two perceptual streams (cf. Bregman, 1990). In the on-beat condition in Experiment 2, however, in which using the rhythmic context sequence to determine the position of the target segment would have resulted in an increase in accuracy, it seems unlikely that the participants were not capable of adopting this strategy only because the timbre of the context sequence was different from the timbre of the target stimulus.

Taken together, the results do not support the prediction derived from the attentional pulse hypothesis (Large & Jones, 1999) that a rhythmic context should induce a rhythmic pattern in the perceptual weights effective for auditory intensity processing. Additional experiments are necessary to decide whether the accuracy of auditory intensity processing is generally unaffected by a rhythmic context. For example, an intensity discrimination paradigm more closely resembling the pitch discrimination task used by Jones et al. (2002) could be studied. For the temporal profile listening task, different increment onsets could be interleaved within one block to prevent the participants from using the constant temporal interval between noise onset and increment onset to direct their attention to the target segment in the profile listening task, thereby increasing the utility of contextual timing information.

## Résumé

Le présent article porte sur les effets du contexte rythmique sur le traitement de l'intensité auditive. L'expérience 1 tentait de vérifier l'hypothèse selon laquelle la direction temporelle involontaire engendrée par une séquence présentée dans un contexte rythmique a une incidence sur la pondération temporelle de l'intensité sonore. L'analyse du coefficient de pondération perceptuelle a servi à mesurer l'attention dirigée vers les portions temporelles individuelles associées à un stimulus prolongé, par le biais de l'appréciation de l'importance des segments temporels individuels dans les jugements globaux de l'intensité sonore. Le contexte rythmique a entraîné un profil de pondération temporelle plus prononcé, mais on n'a pas observé le modèle rythmique attendu des coefficients de pondération. Au cours de l'expérience 2, l'exactitude de la détection d'une crête dans le profil temporel de l'intensité sonore d'un son de niveau variable a été mesurée par rapport au niveau d'augmentation présenté à des moments planifiés et non planifiés. Encore une fois, le contexte rythmique a donné lieu à un profil de pondération temporelle plus prononcé. Toutefois, l'exactitude de la détection n'a pas augmenté lorsque la position du segment cible était prévue, ni diminué lorsque la position du segment cible était imprévue, par rapport à la condition où aucune séquence n'était présentée dans un contexte rythmique.

## References

- Agresti, A. (2002). *Categorical data analysis* (2nd ed.). Hoboken, NJ: Wiley.
- Ahumada, A., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, 49, 1751–1756.
- Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology*, 41, 254–311.
- Berg, B. G. (1989). Analysis of weights in multiple observation tasks. *Journal of the Acoustical Society of America*, 86, 1743–1746.
- Berg, B. G. (1990). Observer efficiency and weights in a multiple observation task. *Journal of the Acoustical Society of America*, 88, 149–158.
- Berg, B. G., & Green, D. M. (1990). Spectral weights in profile listening. *Journal of the Acoustical Society of America*, 88, 758–766.
- Boltz, M. G. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics*, 53, 585–600.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Ellermeier, W., & Schrödl, S. (2000). Temporal weights in loudness summation. In C. Bonnet (Ed.), *Fechner Day 2000* (pp. 169–173). Strasbourg, France: Université Louis Pasteur.
- Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). New York: Academic Press.
- Green, D. M. (1988). *Profile analysis. Auditory intensity discrimination*. New York: Oxford University Press.
- Green, D. M. (1993). Auditory intensity discrimination. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Human psychophysics* (pp. 13–55). New York: Springer.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. Hoboken, NJ: Wiley.
- Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323–335.
- Jones, M. R., & Boltz, M. G. (1989). Dynamic attending and responses to time. *Psychological Review*, 96, 459–491.
- Jones, M. R., Johnston, H. M., & Puente, J. (2006). Effects of auditory



- pattern structure on anticipatory and reactive attending. *Cognitive Psychology*, 53, 59–96.
- Jones, M. R., & McAuley, J. D. (2005). Time judgments in global temporal contexts. *Perception & Psychophysics*, 67, 398–417.
- Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, 13, 313–319.
- Keselman, H. J., Algina, J., & Kowalchuk, R. K. (2001). The analysis of repeated measures designs: A review. *British Journal of Mathematical and Statistical Psychology*, 54, 1–20.
- Kortekaas, R., Buus, S., & Florentine, M. (2003). Perceptual weights in auditory level discrimination. *Journal of the Acoustical Society of America*, 113, 3306–3322.
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106, 119–159.
- Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, 6, 177–208.
- Lutfi, R. A. (1992). Informational processing of complex sound: III. Interference. *Journal of the Acoustical Society of America*, 91, 3391–3401.
- Lutfi, R. A. (1995). Correlation coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks. *Journal of the Acoustical Society of America*, 97, 1333–1334.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 1102–1125.
- McAuley, J. D., & Kidd, G. R. (1998). Effect of deviations from temporal expectation on tempo discrimination of isochronous tone sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1786–1800.
- McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). London: Chapman & Hall.
- Oberfeld, D., & Plank, T. (2005). Temporal weighting of loudness: Effects of a fade in. In Deutsche Gesellschaft für Akustik (Ed.), *Fortschritte der Akustik—DAGA 2005* (pp. 227–228). Berlin: DEGA.
- Pedersen, B., & Ellermeier, W. (2004). Individual differences in integrating loudness over time. In Société Française d'Acoustique & Deutsche Gesellschaft für Akustik (Eds.), *Proceedings of the Joint Congress CFA/DAGA '04* (pp. 887–888). Paris: SFA.
- Plank, T., & Ellermeier, W. (2003). Discrimination of temporal loudness profiles. In B. Berglund & E. Borg (Eds.), *Fechner Day 2003. Proceedings of the 19th annual meeting of the International Society for Psychophysics* (pp. 241–244). Stockholm: ISP.
- Plank, T., & Ellermeier, W. (2004). Discriminating temporal loudness patterns in the absence of overall level cues. In Société Française d'Acoustique & Deutsche Gesellschaft für Akustik (Eds.), *Proceedings of the Joint Congress CFA/DAGA '04* (pp. 397–398). Paris: SFA.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25.
- Repp, B. H. (2004). On the nature of phase attraction in sensorimotor synchronization with interleaved auditory sequences. *Human Movement Science*, 23, 389–413.
- Repp, B. H. (2006). Multiple temporal references in sensorimotor synchronization with metrical auditory sequences. *Psychological Research*, 72, 79–98. Retrieved March 2, 2007 from <http://www.springerlink.com/content/t287537511747001/>
- Richards, V. M., & Zhu, S. (1994). Relative estimates of combination weights, decision criteria, and internal noise based on correlation coefficients. *Journal of the Acoustical Society of America*, 95, 423–434.

Received December 6, 2006

Accepted March 30, 2007 ■