UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Probing attentive and preattentive emergent meter in adult listeners without extensive music training

Ladinig, O.; Honing, H.; Háden, G.; Winkler, I.

DOI
10.1525/MP.2009.26.4.377

Publication date
2009

Document Version
Final published version

Published in
Music Perception

Link to publication

Citation for published version (APA):
Ladinig, O., Honing, H., Háden, G., & Winkler, I. (2009). Probing attentive and preattentive emergent meter in adult listeners without extensive music training. *Music Perception*, *26*(4), 377-386. https://doi.org/10.1525/MP.2009.26.4.377

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:11 Oct 2026

Probing Attentive and Preattentive Emergent Meter in Adult Listeners Without Extensive Music Training

Olivia Ladinig and Henkjan Honing
Universiteit van Amsterdam, Amsterdam, The Netherlands

Gábor Háden and István Winkler
Hungarian Academy of Sciences, Budapest, Hungary

Beat and meter induction are considered important structuring mechanisms underlying the perception of rhythm. Meter comprises two or more levels of hierarchically ordered regular beats with different periodicities. When listening to music, adult listeners weight events within a measure in a hierarchical manner. We tested if listeners without advanced music training form such hierarchical representations for a rhythmical sound sequence under different attention conditions (Attend, Unattend, and Passive). Participants detected occasional weakly and strongly syncopated rhythmic patterns within the context of a strictly metrical rhythmical sound sequence. Detection performance was better and faster when syncopation occurred in a metrically strong as compared to a metrically weaker position. Compatible electrophysiological differences (earlier and higher-amplitude MMN responses) were obtained when participants did not attend the rhythmical sound sequences. These data indicate that hierarchical representations for rhythmical sound sequences are formed preattentively in the human auditory system.

Received June 22, 2008, accepted January 12, 2009.

Key words: rhythm, emergent meter, syncopation, event-related potentials, attention

The concepts of *beat* and *meter* are well-established terms in music production and perception (Clarke, 1999; London, 2004). Most authors agree that *beat induction*, the cognitive ability that allows one to infer a regular beat (or pulse) from a musical excerpt, is universal in and unique to humans, enabling us to entrain to music, and coordinate our movements with others (Honing, 2002). Meter can be defined as being composed of at least two levels of beat with different periodicities. However, there is little agreement in the literature regarding the perceptual/cognitive reality of meter. Is meter simply a concept facilitating the structuring of written musical scores, introduced by composers and performers, or are there indeed some cognitive faculties reflected in the concept of meter? Beat induction can be considered the simplest case of meter, and refers to the subjective emphasis of certain elements of a rhythm (but also in an isochronous stream of clicks), making some elements more salient than others; the beat or *tactus* (Lerdahl & Jackendoff, 1983) is usually equally spaced in time, and is reflected in spontaneous tapping and dancing, usually with an interbeat interval close to 600 ms (Bolton, 1894; Brochard, Abecasis, Potter, Ragot, & Drake, 2003; London, 2004; Yeston, 1976). Meter, seen here as a more fine-grained differentiation of the elements of a rhythm due to multiple levels of hierarchically ordered regular beats, requires the specification of a fixed entity of duration, in this case one *musical measure*. Theoretical models (Lerdahl & Jackendoff, 1983; Longuet-Higgins & Lee, 1984) specify *metric salience,* a value assigned to each sequential position of a rhythmic sound pattern regarding to its position within that measure, by recursively breaking down a musical pattern (with an initially specified length) into subpatterns of equal length (see the top of Figure 1).

The number of recursive subdivisions needed to arrive at a given point (event) in a rhythmic pattern governs the salience of that point: the more subdivisions needed, the lower the salience of the point. The first position in the measure (referred to as the *downbeat*) receives the highest salience in any pattern. In other words, meter reflects the fact that different events in a musical pattern have different importance for the listener. In general, it holds that the higher the salience of an event compared to other events within the same measure, the more listeners expect it to occur. A high-salience event is more important for processing the measure, as indicated for example by the fact that it gets memorized and recalled easier, and, if it is absent, the measure will be perceived as being more complex (Fitch

Music Perception volume 26, issue 4, pp. 377–386, issn 0730-7829, electronic issn 1533-8312 © 2009 by the regents of the university of california. all rights reserved. please direct all requests for permission to photocopy or reproduce article content through the university of california press's rights and permissions website, http://www.ucpressjournals.com/reprintinfo.asp. DOI:10.1525/mp.2009.26.4.377

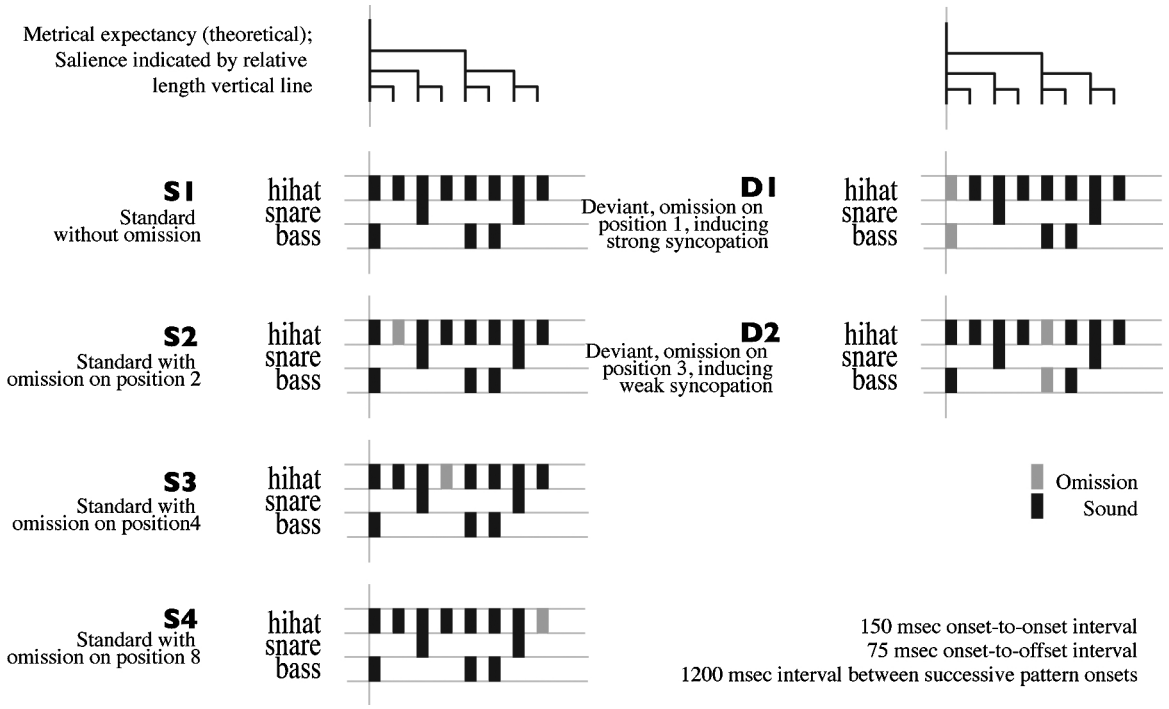

FIGURE 1. Schematic illustration of the stimuli used in the experiment. The top of the figure represents the recursive subdivision of a rhythmic pattern (cf. Martin, 1972) with eight equidistant grid points. The horizontal dimension represents the subdivisions of one musical measure; the vertical dimension represents event salience (i.e., increasing salience with longer lines).

& Rosenfeld, 2007; Pressing, 2002). Supporting this notion, Palmer and Krumhansl (1990) showed, for a corpus of Western classical music, that the average distribution of event occurrences within a measure was highly correlated with the theoretical model proposed by Lerdahl and Jackendoff (1983).

Existing theories disagree whether or not sensitivity to meter is prevalent in all listeners, and where such sensitivity, if any, would come from. Specifically, the question is, whether or not listeners form multilevel hierarchical representations for rhythmic sequences. Expectations in adult listeners with formal music training suggest that they weight events within a measure in a hierarchical manner (Jongsma, Desain, & Honing, 2004; Palmer & Krumhansl, 1990). A study by Ladinig and Honing (2009) shows that this holds irrespective of listener's musical expertise. Furthermore, recent evidence suggests that already at a very early age (e.g., at seven months of age), human infants are sensitive to metric violations (Hannon & Johnson, 2005). Thus it is possible that humans possess some processing predisposition to extract hierarchically structured regularities from complex patterns. Lower–level chunking processes are usually more or less automatic (i.e., they proceed even when one does not attend the given stimuli; e.g., temporal integration, see Cowan, 1984). In contrast, higher-level chunking processes typically require attention to be focused on the stimuli, because they rely on voluntary allocation of limited–capacity resources (e.g., finding sentences in continuous speech). The crucial question is whether or not the hierarchical representation characterizing meter emerges when the rhythmical sound sequence falls outside the focus of attention.

In the current study, we tested whether meter (hierarchical representation for a rhythmical sound sequence) emerges in adults with no extensive music training, and whether meter emergence is modulated by attention. To this end, reactions to meter violations were assessed using behavioral and electrophysiological measures. Reaction time (RT) and discrimination sensitivity (*d'*) measurements served to characterize active detection of meter violations, whereas event-related brain potentials (ERP) were used to assess the detection of meter violations under different task loads while the rhythmic sound sequences were not relevant to the participants' task. The mismatch negativity (MMN) ERP component (Näätänen, Gaillard, & Mäntysalo, 1978; for a recent review, see Näätänen, Paavilainen, Rinne, & Alho, 2007) can be used as a sensitive tool for determining which regular features of a sound sequence the

brain has detected, because MMN is elicited by sounds violating detected auditory regularities. Furthermore, MMN is elicited even when participants perform a task that is unrelated to the test sound sequence (for a review of the effects of attention on MMN, see Sussman, 2007).

MMN has been shown to reflect violations of musical regularities and the effects of music training (for a review, see Tervaniemi & Huotilainen, 2003). For example, Trainor, McDonald, and Alain (2002) showed that participants with no formal music training detected occasional pitch interval changes within transposed melodies in the absence of focused attention. Other studies showed sensitivity to musical key (e.g., Brattico, Tervaniemi, Näätänen, & Peretz, 2006), mistuning of chords (Leino, Brattico, Tervaniemi, & Vuust, 2007), etc. Although fewer previous investigations addressed rhythm processing with the MMN method (the exceptions are Pablos Martin, Deltenre, Hoonhorts, Markessis, Rossion, & Colin, 2007; Vuust et al., 2005), the representation of simpler temporal features has been studied in more detail. For example, it was found that occasionally shortening the interstimulus interval in an otherwise isochronous sequence of sounds elicits the MMN (Nordby, Roth, & Pfefferbaum, 1988). Omitting a sound from a sequence delivered at a fast presentation rate also triggers the MMN response (Yabe, Tervaniemi, Reinikainen, & Näätänen, 1997). Regarding more complex temporal patterns, Pablos Martin and colleagues (2007) found faster processing of binary (e.g., 1:2) as opposed to nonbinary (e.g., 1:3) interval ratios. Finally, music training effects have been shown for both melodic (e.g., Brattico & Näätänen, 2002; Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004) and rhythmic patterns (Vuust et al., 2005; van Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2005). The current interpretation of MMN generation suggests that this ERP component is elicited in response to deviations from expected sounds (Baldeweg, 2007; Winkler, 2007). This makes MMN especially appropriate for testing the emergence of musical meter, because it allows one to compare the strength of expectations between violations at different positions of a rhythmical pattern. The strength of expectation is a prime behavioral correlate of the hierarchical metric structure and more salient deviations trigger earlier and possibly larger-amplitude MMN responses (for a review, see Näätänen & Alho, 1997).

Based on these principles, we presented participants with sound sequences consisting of four sound patterns (Figure 1) having strictly metrical rhythms of the same type (Standard patterns; 90% of the patterns overall), and two patterns that were syncopated variants of the same rhythm (Deviant patterns; 10% overall). One deviant violated the standard pattern at the downbeat position (strong syncopation), and the other at the second most salient position (weaker syncopation). If the brain creates a hierarchical representation for the rhythm of the sound sequences, syncopation at the downbeat is expected to elicit stronger responses from participants than syncopation at the metrically less salient position. "Stronger" response means better detection performance when syncopated patterns are designated as targets and earlier and possibly higher-amplitude MMN response when participants ignore the rhythmic sequence. If, however, the sound sequence is represented in terms of a single-level structure, then sounds in all positions are equally expected by the brain and, therefore, the responses to syncopation will not be stronger at the downbeat than in the metrically less salient position.

Effects of attention were tested at three levels: (1) meter violations are task-relevant (Behavioral Experiment); (2) meter violations are task-irrelevant: participants perform an easy concurrent task (watching a muted movie with subtitles; Electrophysiological Experiment, "Passive Condition"); and (3) meter violations are task-irrelevant: participants perform a difficult concurrent task (detecting unpredictable slight intensity changes in a noise stream; Electrophysiological Experiment, "Unattend Condition"). If forming a hierarchical representation of the rhythmical sound sequence required focused attention, then the strength of expectation should only depend on the position of the syncopation within the pattern when participants focus their attention on the sound sequence. If, however, a hierarchical representation of the rhythmical sound sequence is formed even without focused attention, then syncopation is expected to elicit a stronger response at the downbeat than in the metrically less salient position, irrespectively of the attention condition.

Method

Participants

Twelve healthy volunteers (seven male, $M = 22.83$, $SD = 3.93$) participated in the experiment. Participants gave informed consent after the procedures and aims of the experiments were explained to them. The study was approved by the Ethical Committee (institutional review board) of the Institute for Psychology, Hungarian Academy of Sciences. All participants had frequency thresholds not higher than 20 dB SPL in the

250-4000 Hz range and no threshold difference exceeding 10 dB between the two ears (assessed with a Mediroll, SA-5 audiometer). All participants reported to have received less than one year of music training (i.e., playing an instrument, or singing in a choir) after the obligatory music lessons in primary/secondary school in the past, and did not perform music regularly (defined as once a month) for the past two years. Each participant was tested in both experiments (behavioral and electrophysiological), which were carried out in one session on the same day. One participant's (male, age 20) data was excluded from the analyses because of measurement errors. Throughout the experiments, participants sat in a comfortable chair in the sound-attenuated experimental chamber of the Institute for Psychology, Budapest.

Stimuli

Six different sound patterns were constructed (see Figure 1), which were variants of a rhythmic rock pattern (base-pattern, S1) with eight grid points. The rhythmic patterns were presented by a typical rock-drum accompaniment using snare and bass, and with a hihat on every grid point. The base pattern and the three variants (containing omissions on the lowest metrical level) were *strictly metrical*; that is, they contained no syncopation or slurred notes throughout the pattern. Together, these four metric patterns formed the set of standard patterns (S1-S4). In order to avoid the confound of finding responses resulting from simple pattern matching, a set of sound patterns that share the characteristic of being strictly metrical and regular rhythms, instead of a single sound pattern, was employed to constitute the standard ("abstract MMN"). Two deviants were constructed by omitting events on metrically salient positions in the base-pattern, which lead to syncopated patterns: A strongly syncopated pattern was created by omitting the downbeat (D1), and a slightly weaker syncopation by omitting the second most important beat (D2). Sounds were generated using QuickTime's *drum* timbres (Apple Inc.). Sound duration was 50 ms for *hihat*, 150 ms for *snare*, and 100 ms for *bass* sounds. The interval between grid points (onset-to-onset interval) was 150 ms. Thus each pattern lasted 1200 ms, with no extra silence between patterns (i.e., they formed a continuous stream of rhythm).

Procedures for the Behavioral Experiment

In the behavioral experiment, we assessed the effects of different metrical positions on deviance detection by asking participants to listen to two blocks of 300 continuously presented patterns and to indicate when they felt that there was a break in the rhythm by pressing a response button placed in their dominant hand. The instructions given to participants were as follows:

> You will be presented with sequences of a continuous, regular rhythm. From time to time, the rhythm will be disrupted by some irregularity. This irregularity can be described as if the rhythm appeared to break, or stumble, or get syncopated for a moment. Please indicate by pressing the button as soon as you think such an event occurred.

Two stimulus blocks with 90% standard patterns (S1, S2, S3, and S4 with equal probabilities of 22.5% each) were presented. In one block, D1 was the deviant rhythmic pattern (10%) and in the other block, D2 was the deviant rhythmic pattern (10%). Randomization was constrained so that at least three standard patterns intervened between successive deviants and with S4 never preceding a deviant. The latter constraint was necessary to avoid concatenating two gaps, because S4 had an omission at the last grid position, whereas D1 at the first. The stimuli were presented binaurally using MATLAB via headphones (Sennheiser HD-430), 60 dB over the individual hearing threshold. The order of the two stimulus blocks (differing in the deviant pattern) was balanced across participants.

Data Analysis for the Behavioral Experiment

For each participant, *d'* values (a measure of discrimination sensitivity; see Macmillan & Creelman, 1991) and average reaction-times (RT) for correct responses were computed using MATLAB. The *d'* values were calculated separately from the hit rates for the D1 and D2 deviants and the overall false alarm rate. Responses given within 200-2000 ms from the target (omission) onset were regarded as hits; all other responses as false alarms. Paired two-sample *t*-tests were performed to compare *d'* and RT between the two deviants.

Procedures for the Electrophysiological Experiment

The electrophysiological experiment was conducted always before the behavioral experiment. The fixed order was necessary to avoid drawing participants' attention to the rhythmic deviations. Electrodes were removed between the two experiments, thus giving participants approximately 30 minutes of break time between the two experiments.

The rhythmic stimulus sequences were constructed from the same sound patterns as in the behavioral experiment, but they were delivered by two loudspeakers positioned 0.40 m from the side and 0.15 m behind the participants' head. Sound intensity was again 60 dB above the participant's hearing threshold. A continuous white noise with its intensity alternating between 52 and 54 dB above the participant's hearing threshold was presented concurrently with the rhythmic sound sequences. The noise stream was used to direct attention away from the rhythmic sound sequence in the Unattend condition (see below). Intensity changes occurred randomly with 1.5–32.0 s ($M = 16.75$ s) between them. The noise stream was delivered by a third loudspeaker placed directly in front of the participant at a distance of 1.35 m. During the stimulus blocks, participants also watched a self-selected muted movie with subtitles.

Two attention conditions were employed with identical auditory stimulation (rhythmic sequence and continuous noise). In the Unattend Condition, participants were asked to press a response button to the intensity changes in the noise stream. Performance in the intensity change detection task (group-average hit rate $HR = 0.78$, standard deviation $SD = 0.12$, and reaction time $RT = 1035$ ms, $SD = 77$ ms) showed that the task was difficult but possible to perform at a relatively high level. In the Passive Condition, participants were instructed to ignore all sounds (both the rhythmic sequence and the continuous noise) and to follow a muted self-selected movie. Each condition received 10 stimulus blocks of 300 continuously presented rhythmic patterns. Stimulus blocks consisted of 90% standard patterns (S1, S2, S3, and S4 with equal probabilities of 22.5%, each), 5% of the D1, and 5% of the D2 pattern. Presenting both types of deviants within the same stimulus block ensured that they appeared within exactly the same context and thus the deviance-related ERP responses could be compared directly. Randomization was constrained so that at least three standard patterns intervened between successive deviants and, for the same reasons as mentioned above, the S4 pattern never preceded a deviant pattern. Constructing 90% of the sequence from four different frequent patterns was necessary to avoid MMN being elicited by simple pattern deviation and thus to allow us to interpret the ERP responses specific to the D1 and D2 deviants as related to rhythm violations. Occasional changes of a single repeating pattern are known to elicit MMN even when rhythm is not violated (e.g., Winkler & Schröger, 1995). In the current design, the "standard" (the sequences made up of S1, S2, S3, and S4) is the rhythm, not any given sound pattern, and the deviants are the rhythmic violations caused by D1 and D2. In order to be able to directly compare the deviance-related responses elicited by D1 and D2, these responses were derived by separately subtracting the response elicited by the D1 and D2 patterns when they were regular (standard) within a sequence from when they were violating the rhythm of the sequence (deviant). Thus the pattern-specific responses were eliminated from the difference waveforms, which could then be compared with each other. To this end, participants also were presented with two control stimulus blocks of 300 patterns presenting sequences composed of either the D1 or the D2 pattern alone. The responses recorded to the D1 and D2 patterns in the control stimulus blocks (i.e., when they are standard patterns) served to derive the MMN response (see the EEG data analysis section below). The order of the two attention conditions was balanced across participants. Stimulus blocks usually were separated by short 1-2 minutes breaks, with longer breaks allowing the participant to leave the experimental chamber inserted at need.

EEG Recording

The electroencephalogram (EEG) was recorded at the F3, Fz, F4, C3, Cz, and C4 scalp locations (according to the international 10-20 system) and the left and right mastoids (A1 and A2, respectively), with the common reference electrode attached to the tip of the nose. The ground electrode was placed on the forehead. Eye movements were monitored by recording the electrooculogram (EOG) between two electrodes placed above and below the left eye (vertical EOG) and between two electrodes placed lateral to the outer canthi on both sides (horizontal EOG). EEG was recorded with 32 bit resolution at a sampling rate of 250 Hz by a Neuroscan, NuAmps amplifier (Compumedics Neuroscan Inc.). The signals were on-line low-pass filtered at 40 Hz.

EEG Data analysis

EEG was filtered off-line between 0.1 and 20 Hz. For each D1 and D2 pattern (experimental and control stimulus blocks, separately), an epoch of 1200-ms duration was extracted from the continuous EEG record. The epoch started 600 ms before the onset of the deviation. Epochs with a voltage change below 0.1 μV or above 100 μV on any EEG or EOG channel within the—100 to 500 ms time window (relative to the deviation onset) were rejected from further analysis. Epochs were baseline-corrected by the average voltage of the whole analysis period and averaged separately for

the two deviants and identical control patterns and in the two attention conditions. Using the whole analysis period as baseline balances possible slow shifts that may appear in the long analysis period. The mean number of artifact-free deviant trials per participant was 130.

MMN peak latencies were established as the central (Cz) negative maximum of the average deviant-minus-control difference waveform in the 100-250 ms *post-deviance* time-range, separately for each participant, deviant, and condition. Peak latencies were established automatically in the target latency range. In cases where two or more negative peaks fell within the 100-250 post-deviance time-window and the amplitude difference between the peaks was small (<0.5 μV), selection of the latency was aided by visual inspection of waveforms recorded by the C and F electrodes. The effects of attention and deviance position were analyzed by a repeated-measure analysis of variance (ANOVA) with the structure Attention (Unattend vs. Passive) × Position (Strong vs. Weak).

MMN mean amplitudes were averaged from 60 ms time windows centered on the central (Cz) negative MMN peaks observed from the group-averaged deviant-minus-control difference waveforms, separately for the two deviants and two attention conditions. Thus MMN was derived by subtracting between responses elicited by identical sound patterns presented in different sequences (i.e., when D1 and D2 are deviants among standards and when D1 and D2 form homogenous control sequences). Responses elicited by the standard patterns were not used in the MMN measurements. This derivation of MMN prevents the emergence of confounding differences stemming from pattern-specific ERP responses. The group-averaged central MMN peak latencies were: 160, 140, 196, and 176 ms *from deviation (omission) onset* for the Unattend-Strong, Passive-Strong, Unattend-Weak, and Passive-Weak deviant responses, respectively. The effects of attention, deviance position, and the scalp distribution of the MMN amplitudes were analyzed with a repeated-measure ANOVA of the structure Attention (Unattend vs. Passive) × Position (Strong vs. Weak) × Frontality (Frontal vs. Central electrode line) × Laterality (Left vs. Middle vs. Right). All significant effects and interactions are reported below. Greenhouse-Geisser correction of the degrees of freedom was applied where appropriate and the ε correction factor as well as η^2 effect size are reported.

Behavioral Data

Discrimination sensitivity was significantly higher for Strong than for Weak deviants, $t(10) = 2.80$, $p < .05$; d'(Strong) = 2.77, d'(Weak) = 2.13. There was also a tendency for faster RT's for Strong than for Weak deviants, $t(10) = 1.85$, $p < .10$; RT(Strong) = 536.69 ms, RT(Weak) = 585.68 ms.

Discussion of the Behavioral Data

Higher sensitivity and shorter RT's for Strong as compared to Weak deviants suggest that theoretical metrical salience affected the processing of rhythmic patterns in our participants when they attended the stimulus sequence.

Electrophysiological Data

The D1 and D2 patterns elicited a fronto-centrally more negative response between 100 and 250 ms from the onset of the omissions when the patterns violated the rhythmic context set up by the frequent standard patterns (S1-S4) than when the same patterns were presented alone in the homogeneous control stimulus blocks (Figure 2). The difference between the ERP responses elicited by the deviant and the identical control stimuli can be identified as an MMN response (cf. below). Significantly shorter MMN peak latencies (measured from the onset of deviation; see Figures 2 and 3) were obtained for Strong as compared to Weak deviants, $F(1, 10) = 20.69$, $p < .01$, $\eta^2 = 0.67$ (average peak latencies: Passive[Strong] = 145.45 ms, Passive[Weak] = 165.45 ms, Unattend[Strong] = 149.09 ms, and Unattend[Weak] = 190.18 ms). The ANOVA of MMN amplitudes (see Figures 2 and 3, and Table 1 for mean MMN amplitudes) yielded main effects of Position, $F(1, 10) = 5.62$, $p < .05$, $\eta^2 = 0.36$, Frontality, $F(1, 10) = 10.56$, $p < .01$, $\eta^2 = 0.51$, and Laterality, $F(2, 20) = 13.86$, $p < .001$, $\varepsilon = 0.83$, $\eta^2 = 0.58$. Strong deviants elicited higher-amplitude MMN responses as compared to Weak deviants. MMN was larger over central than frontal electrodes and over midline than lateral electrodes. There was also a significant interaction between Attention and Frontality, $F(1, 10) = 35.24$, $p < .001$, $\eta^2 = 0.78$, stemming from lower frontal MMN amplitudes in the Passive condition than in any other combination of these two factors (Tukey HSD posthoc test with $df = 10$, $p < .001$ for all of the referred comparisons). This result rules out the possibility that the deviant-minus-control difference waveform would contain significant contribution from the N2b ERP component. This is because N2b is elicited only when participants actively detect a stimulus (Novak, Ritter, Vaughan, & Wiznitzer, 1990). Furthermore, the ERP difference cannot reflect difference between two

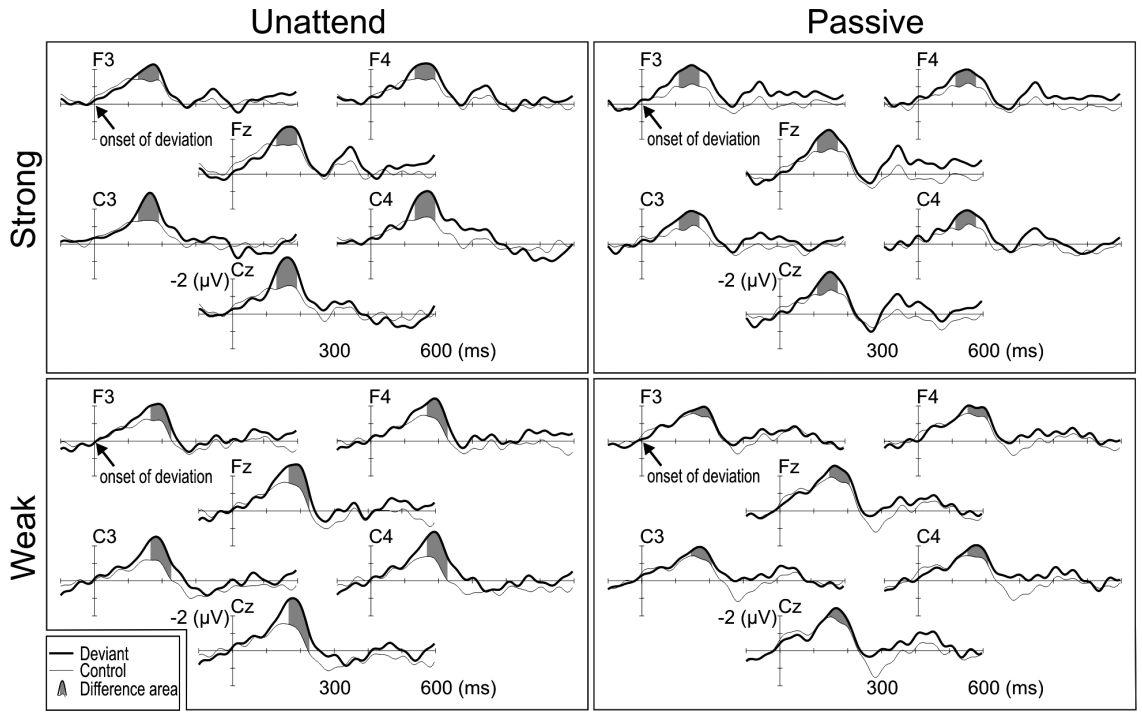

FIGURE 2. Group-averaged ($n = 11$) ERP responses elicited by deviant patterns (experimental stimulus blocks; thick lines) and identical control patterns (control stimulus blocks; thin lines). Left: Unattend condition; right: Passive condition. Upper panels show the responses to Strong, lower panels to Weak metrical position deviants. The area between deviant and control responses within the measurement window is marked by grey shading. Responses are aligned at the onset of deviation (the time point at which the omitted sound appears in the S1 pattern).

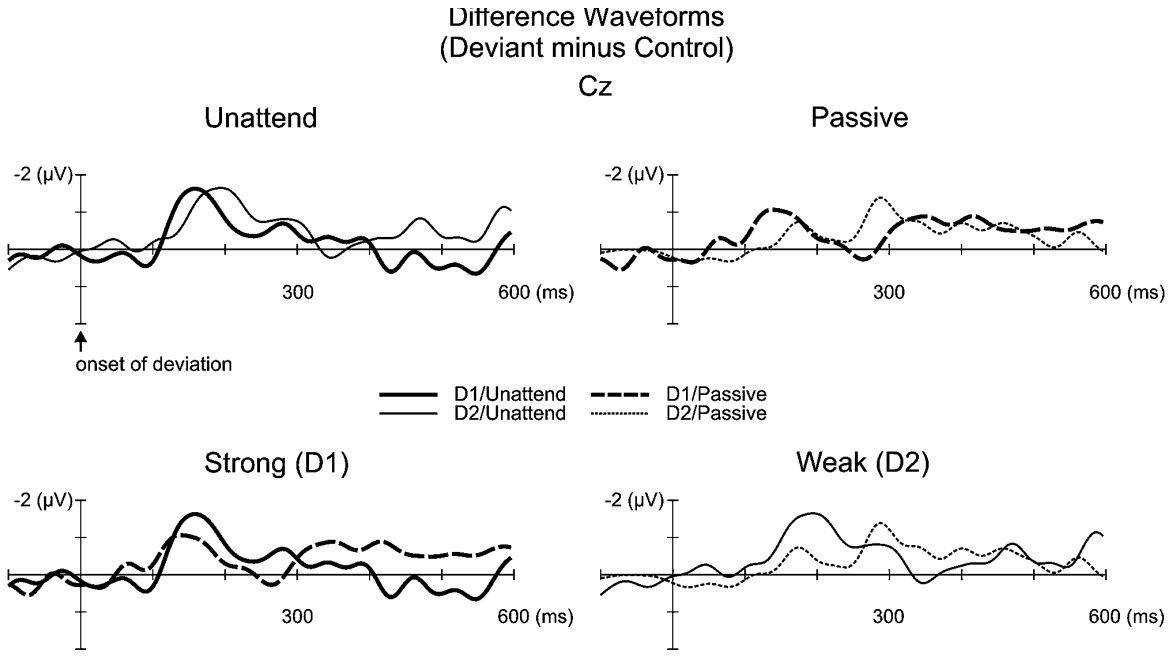

FIGURE 3. Group-averaged ($n = 11$) deviant-minus-control difference waveforms (thick lines for Strong, thin lines for Weak deviants; continuous lines for the Unattend, dashed lines for the Passive condition). Top panels: Comparison between responses elicited by Strong and Weak deviants, separately for the Unattend (left) and Passive (right) conditions. Bottom panels: Comparison between the two attention conditions, separately for Strong (left) and Weak (right) deviants. Responses are aligned at the onset of the deviation.

TABLE 1. Group-Averaged MMN Amplitudes in μV with Standard Errors of the Mean (SEM) in Parentheses.

Attention	Passive		Unattend	
Electrode/Position	Strong	Weak	Strong	Weak
F3	−2.23 (0.40)	−1.20 (0.29)	−2.00 (0.19)	−1.53 (0.40)
Fz	−2.62 (0.47)	−1.70 (0.38)	−2.58 (0.28)	−1.99 (0.47)
F4	−1.93 (0.41)	−1.27 (0.45)	−2.10 (0.31)	−1.68 (0.41)
C3	−2.03 (0.37)	−1.42 (0.35)	−2.72 (0.34)	−2.15 (0.37)
Cz	−2.57 (0.47)	−1.71 (0.41)	−3.29 (0.30)	−2.49 (0.47)
C4	−2.08 (0.41)	−1.48 (0.40)	−2.99 (0.35)	−2.38 (0.41)

N1 components, because it is elicited by sound omissions, which do not elicit the N1 component. Very importantly, the Attention factor did not significantly interact with the Position factor for either peak latencies or MMN amplitudes. This means that Strong deviants elicited earlier and higher-amplitude MMN responses than Weak deviants irrespective of the attention conditions.

Discussion of the Electrophysiological Data

MMN responses were elicited by deviations in both metrical positions and in both attention conditions. This suggests that rhythmic violations are detected even when attention is not focused on the sound sequence. Furthermore, Strong deviants elicited a stronger (earlier and higher-amplitude) response than Weak ones. This result corroborates the behavioral data in suggesting that metric salience affected the detection of rhythm violations. Stronger MMN responses are usually recorded to perceptually more salient deviations (Näätänen & Alho, 1997). Since the amount of raw acoustic deviation did not differ between the two deviant positions, larger perceived deviations suggest sharper (more precise) memory representations for metrically salient elements of rhythmic patterns (a similar effect on the sharpness of the memory representations underlying MMN has been demonstrated by masking studies; see Winkler, Reinikainen, & Näätänen, 1993). Modulation of the memory representations by metric salience strongly argues for the conclusion that the brain formed hierarchical representations for the rhythmic stimulus sequences.

The only effect of attention was lower frontal MMN amplitudes in the Passive compared with the Unattend condition. This effect was not significantly different between MMNs elicited by Strong and Weak deviants. Rather, it probably reflects differences in the general activity of the frontal cortex in the two attention conditions (e.g., difference in the arousal level or between processing simple sound change as opposed to following a movie). Thus it appears that the processing of meter (forming hierarchical representations for rhythmical sound sequences) does not require significant amounts of limited higher-level capacities, a sign that meter may be processed at lower levels of auditory perception. The picture emerging from the electrophysiological data is that meter is extracted more or less automatically from rhythmic sequences, suggesting that it is an "intelligent" low level auditory processing capability, of which more and more are discovered by recent research (Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001).

General Discussion and Conclusion

The behavioral detection of syncopated rhythms data as well as the ERPs recorded under two different attention conditions concerning meter induction were consistent in our group of participants. They were able to detect syncopated rhythms in an active behavioral task (indicated by the accuracy and speed of detection), as well as passively in the ERP experiment when they focused their attention on a task unrelated to the rhythmic sound sequences. Not only did participants distinguish syncopated patterns from strictly metrical ones, but they also showed sensitivity to the position (metric salience) or in other words, to the strength of the syncopation. This result is in full accordance with the Longuet-Higgins and Lee (1984) model, which predicts that the most salient position elicits a significantly stronger response than syncopation on any lower salient position of the rhythm. Furthermore, this result suggests that meter is not only a concept facilitating the structuring of written musical scores, but it corresponds to the structure of memory representations in the human brain.

These results suggest that beat induction, which according to Povel (1981) is an essential first step in the

perception of temporal sequences, is functional both in active and passive listening situations. Furthermore, our participants clearly were sensitive to the hierarchical ordering in beat perception (as revealed by the difference in responses between D1 and D2; cf. Figure 3). This provides further evidence for the general perceptual/cognitive capability based interpretation of meter. While earlier research showed only a marginal sensitivity to meter in listeners with little or no formal music training (e.g., Jongsma et al., 2004; Palmer & Krumhansl, 1990), the current study demonstrated that meter is a mental representation that does not require advanced formal music training. This conclusion does not rule out the possibility that, similarly to other music-related processing capabilities, the representation of rhythmic structures can be improved by music training (see for example van Zuijen et al., 2005). It remains a question for future research whether basic sensitivity for meter is a result of learning by exposure to one's musical environment (Huron, 2006), or, as hinted by the current results as well as studies showing sensitivity to meter at a very early age (Hannon & Johnson, 2005), whether it stems from a general cognitive predisposition of the human brain for breaking down complex patterns recursively into equal sized subpatterns (Martin, 1972).

Author Note

Correspondence concerning this article should be addressed to Olivia Ladinig, Music Cognition Group, Institute for Logic, Language and Computation (ILLC), Universiteit van Amsterdam (UvA), The Netherlands. E-MAIL: olivia.ladinig@gmail.com

References

Baldeweg, T. (2007). ERP repetition effects and mismatch negativity generation: A predictive coding perspective. *Journal of Psychophysiology, 211*, 204-213.

Bolton, T. (1894). Rhythm. *The American Journal of Psychology*, *6,* 145-238.

Brattico, E., & Näätänen, R. (2002). Context effects on pitch perception in musicians and nonmusicians: Evidence from event-related-potential recordings. *Music Perception, 19*, 199-222.

Brattico, E., Tervaniemi, M., Näätänen, R., & Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain Research, 1117*, 162-174.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The 'ticktock' of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science, 14,* 362-366.

Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *Psychology of music* (2nd ed., pp. 473-500). New York, NY: Academic Press.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin, 96,* 341-370.

Fitch, W. T., & Rosenfeld, A. J. (2007). Perception and production of syncopated rhythms. *Music Perception, 25,* 43-58.

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience, 16,* 1010-1021.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology, 50,* 354-377.

Honing, H. (2002). Structure and interpretation of rhythm and timing. *Dutch Journal of Music Theory*, *7*, 227-232.

Huron, D. (2006) *Sweet anticipation. Music and the psychology of expectation*. Cambridge, MA: The MIT Press.

Jongsma, M. L. A., Desain, P., & Honing, H. (2004). Rhythmic context influences the auditory evoked potentials of musicians and non-musicians. *Biological Psychology, 66,* 129-152.

Ladinig, O., & Honing, H. (2009). *Complexity judgments as a measure for event salience.* Manuscript submitted for publication.

Leino, S., Brattico, E., Tervaniemi, M., & Vuust, P. (2007). Representation of harmony rules in the human brain: Further evidence from event-related potentials. *Brain Research, 1142,* 169-177.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

London, J. (2004). *Hearing in time: Psychological aspects of musical meter.* New York, NY: Oxford University Press.

Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception, 1*, 424-441.

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

Martin, J. G. (1972). Rhythmic (hierarchic) versus serial structure in speech and other behaviour. *Psychological Review*, *79*, 487-509.

Näätänen, R., & Alho, K. (1997). Mismatch negativity (MMN)—the measure for central sound representation accuracy. *Audiology and Neuro-otology, 2,* 341-353.

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective attention effect on evoked potential reinterpreted. *Acta Psychologica, 42*, 313-329.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118,* 2544-2590.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). 'Primitive intelligence' in the auditory cortex. *Trends in Neurosciences, 24,* 283-288.

Nordby, H., Roth, W. T., & Pfefferbaum, A. (1988). Event-related potentials to time-deviant and pitch-deviant tones. *Psychophysiology, 25,* 249-261.

Novak, G. P., Ritter, W., Vaughan, Jr., H. G., & Wiznitzer, M. L. (1990). Differentiation of negative event related potentials in an auditory discrimination task. *Electroencephalography and Clinical Neurophysiology, 75,* 255-275.

Pablos Martin, X., Deltenre, P., Hoonhorst, I., Markessis, E., Rossion, B., & Colin, C. (2007). Perceptual biases for rhythm: The mismatch negativity latency indexes the privileged status of binary vs non-binary interval ratios. *Clinical Neurophysiology, 118,* 2709-2715.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 728-741.

Povel, D. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 3-18.

Pressing, J. (2002). *Cognitive complexity and the structure of musical patterns.* Retrieved June 2, 2006, from University of Melbourne web site: http://psy.uq.edu.au/CogPsych/Noetica/OpenForumIssue8/Pressing.html

Sussman, E. S. (2007). A new view on the MMN and attention debate: The role of context in processing auditory events. *Journal of Psychophysiology, 21*, 164-175.

Tervaniemi, M., & Huotilainen, M. (2003). The promises of change-related brain potentials in cognitive neuroscience of music. *Neurosciences and Music, 999,* 29-39.

Trainor, L. J., McDonald, K. L., & Alain, C. (2002). Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *Journal of Cognitive Neuroscience, 14,* 1-13.

van Zuijen, T. L., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2005). Auditory organization of sound sequences by a temporal or numerical regularity: A mismatch negativity study comparing musicians and non-musicians. *Cognitive Brain Research, 23,* 270-276.

Vuust, P., Pallesen, K. J., Bailey, C., van Zuijen, T. L., Gjedde, A., Roepstoff, A., & Østergaard, L. (2005). To musicians, the message is in the meter. Pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage, 24,* 560-564.

Winkler, I. (2007). Interpreting the mismatch negativity (MMN). *Journal of Psychophysiology, 21,* 147-163.

Winkler, I., Reinikainen, K., & Näätänen, R. (1993). Event related brain potentials reflect traces of the echoic memory in humans. *Perception and Psychophysics, 53,* 443-449.

Winkler, I., & Schröger, E. (1995). Neural representation for the temporal structure of sound patterns. *Neuroreport, 6,* 690-694.

Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *NeuroReport, 8,* 1971-1974.

Yeston, M. (1976). *The stratification of musical rhythm.* New Haven, CT: Yale University Press.