Modeling the Tendency for Music to Induce Movement in Humans: First

Correlations with Low-Level Audio Descriptors across Music Genres

Guy Madison[1] , Fabien Gouyon[2], Fredrik Ullén[3], and Kalle Hörnström[1]

[1]Umeå University, [2]INESC Porto, [3]Karolinska Institutet

Author Note

Guy Madison and Kalle Hörnström, Department of Psychology Umeå University, Sweden

Fabien Gouyon, INESC Porto, Portugal.

Fredrik Ullén, Department of Women's and Children's Health and Stockholm Brain Institute, Karolinska Institutet, Sweden.

Correspondince concerning this article should be sent to Guy Madison, Department of Psychology, Umeå University, SE-901 87 UMEÅ, E-mail: guy.madison@psy.umu.se.

Abstract

Groove is often described as the experience of music that makes people tap their feet and want to dance. A high degree of consistency in ratings of groove across listeners indicates that physical properties of the sound signal contribute to groove [Madison (2006) *Music Perception, 24,* 201-208]. Here, correlations were assessed between listeners' ratings and a number of quantitative descriptors of rhythmic properties for one hundred music examples from five distinct traditional music genres. Groove was related to several different rhythmic properties, some of which were genre-specific and some of which were general across genres. Two descriptors corresponding to the density of events between beats and the salience of the beat, respectively, were strongly correlated with groove across domains. In contrast, systematic deviations from strict positions on the metrical grid, so-called microtiming, did not play any significant role. The results are discussed from a functional perspective of rhythmic music to enable and facilitate entrainment and precise synchronisation among individuals.

Keywords: adjective ratings; audio analysis; coordination; computational modelling; evolution; dance; groove; movement; music; preferences; entrainment; synchronisation

Music often induces spontaneous movements in people, such as rhythmic nodding or tapping of the feet. We call the experience that motivates or induces such movement groove (Madison, 2006). Indeed, much music is intended for synchronised movement in the form of dance, drill, and ritual behaviours (McNeil, 1995). To facilitate entrainment or coordinated action is therefore one function of many kinds of music. There is accumulating evidence to indicate that the connection between movement and the rhythmic component of music is biologically determined. This suggests that this connection might have had an adaptive function at some point in human phylogeny (Merker, Madison, & Eckerdal, 2009). First, music is a human universal (Pinker, 2003), and co-ordinated dance to rhythmically predictable music presumably occurs in all cultures (Nettl, 2000). Second, passive listening to rhythmic sound sequences activates brain regions in the motor system, e.g. the supplementary and pre-supplementary motor areas and lateral premotor cortex, even in tasks without any reference to movement (e.g.Bengtsson et al., 2008; Chen, Penhune, & Zatorre, 2008; Grahn & Brett, 2007). Third, experiencing rhythmic music is associated with pleasure, as indicated by self-ratings (Madison, 2006; Todd, 2001), by activation of brain areas associated with reward and arousal, such as the amygdala and orbitofrontal cortex (Blood & Zatorre, 2001), and by psychophysiological measures, including respiration rate (Khalfa, Roy, Rainville, Dalla Bella, & Peretz, 2008) and biochemical markers (Möckel et al., 1994).

Music in general and rhythmic predictability in particular is thus associated with behavioural and physiological correlates that one would expect from a phylogenetic trait. It has been proposed that entrainment among individuals is or has been adaptive (e.g. Roederer, 1984; Hodges, 1989; Merker, 1999; 2000), which could have endowed us with a motivational apparatus to engage in such behaviour. While such ultimate explanations are outside the scope of the present article, we retain their functional predictions as a useful working hypothesis:

signal properties that facilitate synchronisation are preferable, and groove might reflect an assessment of this utility.

What predictions can be made about physical properties of the sound signal that facilitate synchronisation? Human synchronisation is based on predictive timing, since reacting to each other's actions would have a lag of at least 100 ms (Kauranen & Vanharanta, 1996). Predictive timing largely relies on the signal to be periodic, that is, to feature a regular beat. Prediction is most accurate when the period is in the range of 300-1,000 ms (Fraisse, 1982), and the beat-to-beat variability must be no larger than a few percent of the beat interval (Madison & Merker, 2002). These conditions are met by most music and certainly by all dance music, and correspond to a wide range of tempi from 60 to 200 beats per minute (BPM). According to so-called BPM lists provided by the disc jockey community, music suitable for dancing exhibits a pronounced peak close to 125 BPM, and a less pronounced peak close to 100 BPM (van Noorden & Moelants, 1999). Thus, one might assume that a metronome set at 125 BPM would be an ideal stimulus for entrainment, since it exhibits the preferred tempo without any variability or other events that may distract from its simple and efficient predictive time structure. Few people would consider that a particularly motivating stimulus for dancing, however, and would probably rate it low on groove. What else in the musical signal might then conceivably be related to groove?

Most music demonstrates a rich web of rhythmic patterns. It is notable that even temporally regular melodies are often imbedded in embellished and rhythmically more complex accompaniments. The overall assumption of the present study is that several features of such rhythmic patterns facilitate synchronization, and we test several hypotheses concerning which specific properties of the music increase groove. Before laying out the background to our hypotheses in detail, we must briefly review the fundamental characteristics of the so-called metrical structure of music.

A metrical structure, or metrical grid, is characteristic of music cross-culturally. It is reflected in the well-known small-integer subdivisions of larger units: half notes, fourth notes, eighth notes, and so forth. The metrical structure can be described as hierarchical with lower levels of shorter intervals being subordinate to higher levels of longer intervals. Lower levels are typically represented in the rhythmic accompaniment, while higher levels are constituted by the measure and even larger structures defined by melodic or rhythmic patterning. Consequently, different metrical levels provide redundant representations of the beat, and reinforce each other.

Enter the fact that human timing is non-linear with respect to time. As mentioned above, intervals in the range 300-1,000 ms are favoured for the beat, the primary temporal level for entrainment and synchronisation. On the one hand, intervals shorter than the beat may be favoured for achieving high temporal precision. Temporal variability in human performance is essentially a constant proportion (around 0.03-0.05) of the interval to be timed, at least in the range 300-900 ms (Madison, 2001). A relatively slow musical tempo of, say, 80 BPM (750 ms between beats) would thus yield a standard deviation on the order of 40 ms in the onset of sounds produced by a human voice (cf. Hibi, 1983). For such a tempo, a fair proportion of the sounds of two or more sequences of sounds produced by humans would be perceptually asynchronous and would therefore not lead to signal summation (cf. Merker, 2000b), whereas a very fast tempo of, say, 240 BPM would yield very few, if any, events that are asynchronous. Accordingly, temporal subdivision is found to facilitate precise synchronisation (Repp, 2003), probably because rhythmical levels faster than the beat provide richer temporal information (Repp, 2005).

On the other hand, intervals longer than the beat may be favoured for co-ordination on a time scale of up to a few seconds, moving a particular limb or the whole body in a particular direction as is required in dance. At the upper end of the tempo range, events tend to be

perceived as members of a group or sequence of events rather than separate events when their interval is shorter than ~330 ms (240 BPM) (Kohno, 1993; Riecker, Wildgruber, Mathiak, Grodd, & Ackermann, 2003). At the lower end, longer intervals enable the identification of specific points in temporal patterns time so that particular movements can be correctly assigned in time and space. Rhythmic patterning is found to considerably improve synchronisation to events with long intervals, demonstrating that temporal information provided between movements is indeed used by the auditory system to improve the timing of these movements (Madison, 2009).

When there is only one temporal level of information like that of a metronome, there is hence obviously a tradeoff between short intervals that provide high temporal precision and long intervals that correspond to actual movement and movement patterns. This might be a functional explanation for the metrical structure in music. While this could form a theoretical discussion in its own right, we only touch on it briefly here for the sake of argument and the hypotheses it generates with respect to groove: Inasmuch as both segmentation and subdivision of the beat into larger and shorter units facilitates different aspects of synchronisation behaviour, it seems likely that such redundant rhythmical patterning contributes to the experience of groove. The rhythmic patterning discussed up to this point is accommodated within the idealised metrical structure – in other words a perfectly isochronous segmentation of time. In contrast, the small literature on groove has almost exclusively focused on microtiming as the factor underlying groove, that is, on deviations from isochrony (see e.g., Keil, 1995; Keil & Feld, 1994; Iyer, 2002; McGuiness, 2005; Waadeland, 2001). Because the ubiquitous deviations from canonical time values found in human performance of music are typically smaller than the smallest canonical time-value used in a given musical context (e.g. 16th or 32nd notes), they are often referred to as microtiming (Gouyon, 2007). While some amount of variability is inherently unsystematic and related to human limits in

perception and motor control, there is also systematic microtiming, as defined by its consistency within (Shaffer & Todd, 1994) or across performers (Repp, 1998). One reason for the focus on microtiming as the vehicle for groove is probably that both groove and microtiming are known to differ between performances of the same musical piece. Since the musical structure is assumed constant in this case, differences in microtiming would appear to be a likely explanation.

In conclusion, groove appears to reflect the music's efficiency for entrainment. The physical correlates of groove might, we propose, include (1) the degree of repetitive rhythmical patterning around comfortable movement rate, on the time scale up to a few seconds (henceforth *Beat Salience*), (2) the relative magnitude of *periodic* sound events at metrical levels faster than the beat (henceforth *Fast Metrical Levels*), and (3) the density of sound events between beats generally (henceforth *Event Density*), because they may also increase the temporal information, and (4) systematic (i.e., to some extent repetitive) micro-timing of events *between* beats (henceforth *Systematic Microtiming*), because that may increase the predictability on a time horizon of multiple beats, useful for the more complex co-ordination typical of dance and drill. In addition to this, we also considered (5) unsystematic (i.e. non-repetitive) micro-timing *around* beats (henceforth *Unsystematic Microtiming*), because it has been suggested that such deviations may be a correlate of groove (Keil et al., 1994; Keil, 1995).

In music, variables tend to form clusters of properties that we call styles or genres, whose perceptual significance is so powerful that they often can be discriminated after hearing less than one second of a music example (Gjerdingen & Perrott, 2008). The properties themselves are largely unknown or unmeasurable by known methods, however. For a correlational design this poses a risk for confounds, in that correlations between observed variables might be driven by unobserved variables that are in turn correlated with the observed ones. Consider

for example the hypothetical case that one genre always features a high-pitched rhythm instrument on every beat and that the examples of this genre also yield high ratings of groove, although this happens to be unrelated to the presence of this instrument. When this rhythm instrument due to its high spectral power yields higher values in one of the rhythmic descriptors (such as beat salience, described in the method section), there is a risk for a spurious correlation between this descriptor and the groove ratings. This kind of risk can be decreased by a careful choice of music examples. Since music within a particular genre is more homogeneous in a large number of (unknown) properties than is music across genres, correlations between sound descriptors and groove among music examples within the same genre are less likely to be a side effect of confounding variables.

Another issue related to genre is that a common function, such as facilitating synchronisation, might be realised by different means across the great diversity of style elements among the world's many musical traditions. We have already identified four different physical properties that conceivably should facilitate synchronisation and therefore induce groove. Inasmuch as these different properties may to some extent achieve the same perceptual effect independently of each other, different musical traditions might have employed each of them to different extents. Comparisons across genres could therefore lend stronger credibility to our functional hypothesis if it is found that groove is equally relevant but induced by different means in musical traditions that have developed relatively independently of each other.

In order to address these questions, we selected five distinct music genres on their likelihood of having developed independently of each other. To this end, we favoured traditional music, which is likely to have maintained some of its characteristics over time. In addition to jazz, we chose genres coming from well-defined and non-overlapping geographical regions that have had a relatively small influence of Western or other music

heavily dispersed by mass media. The minimal number of examples that could yield meaningful correlations being about twenty, one hundred examples in all were sampled from recordings of Greek, Indian, Jazz, Samba, and West African music.

We predicted that listeners' ratings of groove would be correlated with (1) Beat Salience, (2) Fast Metrical Levels, (3) Event Density, and (4) Systematic Microtiming. It was further predicted that ratings would not be correlated with Unsystematic Microtiming, because we cannot think of a plausible functional link for such a relation. No particular predictions were made with respect to music genres, except that they might differ in their patterns of correlations.

Materials and methods

Participants

Seven female and twelve male native Swedes acted as listeners. Apart from obligatory recorder lessons in primary school, none had participated in formal music or dance training, or had sung or played a musical instrument in a systematic fashion. Their musical preferences were not considered because the design asked for correlations across the sample of participants, and because preferences were assumed to play a minor role for these correlations anyway. Participants were recruited by advertisements on the university campus, ranged from 19 to 32 years in age, and were paid for their participation.

Stimuli

Twenty music examples were selected from each of five music genres, namely traditional folk music from a certain region, here referred to as Greek, Indian, Jazz, Samba, and West African, making a total of 100 examples. The examples were taken from web sites and commercially available CDs (see Appendix for artists and titles). They were copied from positions within the original sound tracks that were representative for the track as a whole. This typically

meant at least one complete musical phrase, beginning on the first beat in a measure. As a consequence, the duration of the examples ranged from 9.06 to 14.55 s. All examples were subjected to equal amplitude normalisation. The tempi of the examples ranged from 81 to 181 BPM as determined by tapping to the music using two different methods. The first was to tap a metronome with a tempo gauge function (Boss DB-66). The other was to tap a computer key while the music was played by the Sonic Visualiser[1] software, and then carefully aligning the resulting graphical representations of the taps with the sonogram representation of the music with a precision better than 5 ms. This was done independently by authors KH and GM, who both have extensive experience of ensemble music performance and music teaching. Both found the task simple and unambiguous, and did not experience that genres differed in how difficult it was to find the most salient beat level. This is often the case for popular music (e.g., Levitin & Cook, 1996; van Noorden et al., 1999). These four tempi determinations differed less than 2 BPM for any music example.

Rating scales

Three words were subjected to ratings of their appropriateness for describing each music example. Groove was carefully defined prior to the experiment; the literal translation from Swedish was "evokes the sensation of wanting to move some part of the body" and was represented by the shorter word *rörelseskapande* (Madison, 2006). The other two words *välbekant* (*familiar*) and *bra* (*good*) were defined as "you have listened to similar music before" and "you like the music and wish to continue listening". The scales appeared as horizontal lines divided by 11 equidistant short vertical lines marked with the numbers 0 through 10, anchored "not at all appropriate" (0) and "very appropriate" (10).

It is often the case that individuals differ in how they use the response space offered by a scale, both in terms of central tendency (i.e., generally low or high ratings) or in variability

---

[1] http://www.sonicvisualiser.org/

(i.e., use the full range or a limited part of the range). A range-correction procedure was therefore applied, in which the minimum and maximum ratings across all 100 responses were obtained for each participant, and each of that participant's rating x_i was transformed to the quotient between $x_i - x_{(min)}$ and $x_{(max)} - x_{(min)}$ (Lykken, Rose, Luther, & Maley, 1966).

Design

Music Genre was the independent variable. Each genre featured 20 music examples in order to provide some naturally occurring variability. Dependent variables were responses to the three rating scales. *Groove* was the main dependent variable and *Familiarity* was a post-hoc control of listeners' previous experience with the different genres and whether they had heard any of the music examples before. *Good* was included for possible post-hoc evaluation of the amount of variability within the music samples and of the listeners' rating consistency, should it prove to be poor for any of the rating scales. All music examples were presented in a different random order for each listener. The rating scales also appeared in a different random order on the computer display.

Procedure and Apparatus

The experiment was administered by a custom-made computer program, which played the sound files through the built-in sound card of a PC and a pair of headphones, and collected responses by means of the computer's mouse or keyboard. Each listener individually attended one session, lasting between 41 and 56 minutes, which began with thorough written instructions of the task ahead. Part of the instruction was (translated from Swedish) "You will hear a large number of music examples. For each example you are to rate how well you think each of three different adjectives corresponds with your experience of the music". Listeners were asked to note on a notepad if they recognized the example. They were also encouraged to work in a calm and concentrated fashion, to rate each example spontaneously, and to take a break when feeling fatigued or inattentive. The written instructions included definitions of the

rating words (stated in the previous section), and the listeners were told to use the words accordingly.

The first block consisted of 10 music examples, two from each of the five genres. These examples were taken from other positions in the tracks from which some of the actual examples were taken. Listeners were told that this was the start of the experiment proper, but its purpose was in fact to orient participants about the type of music and the range of the properties to rate in the experiment. Ratings from the first block were not included in the analysis. Each session was terminated with a brief interview concerning the listener's musical habits and assessment of the rating task.

Sound descriptors

The purpose of the sound descriptors was to measure, as well as possible, the magnitude of physical properties of the sound signal corresponding to the psychological effects or functions outlined in the introduction. Thus, each descriptor can be seen as a probe, like a litmus paper, sensitive to a particular, pre-defined property. A thorough survey of computational models of tempo and beat perception, meter perception and timing perception can be found in Gouyon (2005). Computational models of microtiming have rarely been applied (Bilmes, 1993; Iyer, 2002; Iyer, 1998). Seppanen (2001) and Gouyon et al. (2002) have reported automatic determination of fast metrical levels, and Busse (2002) proposed the computation of a groove factor of MIDI signals. Computational models have also been proposed for the determination of rhythm patterns (Dixon, Gouyon, & Widmer, 2004; Wright & Berdahl, 2006).

Before the computation of all descriptors the audio data were pre-processed into a representation of lower dimensionality that highlights energy changes (cf. Klapuri, Eronen, & Astola, 2006). More precisely, we computed on short consecutive chunks of the signal (of about 10 ms) the energy of half-wave rectified subband signals, as follows. First, the audio

signal was filtered in eight non-overlapping frequency bands by means of eight 6th-order Butterworth filters: a first low-pass filter with a cut-off frequency of 100 Hz, 6 band-pass filters and 1 high-pass filter distributed uniformly on a logarithmic frequency scale (i.e. passbands are approximately [100 Hz – 216 Hz], [216 Hz – 467 Hz], [467 Hz – 1009 Hz], [1009 Hz – 2183 Hz], [2183 Hz – 4719 Hz], [4719 Hz – 10200 Hz] and [10200 Hz – 22050 Hz]). Second, the signal was half-wave rectified, squared, and downsampled to a sampling frequency of 86 Hz in each band, after group delay had been taken into account. Third, signals in each of the eight bands were summed into a single time series over which we then computed the degree of change, as the differential normalized with its magnitude. This is supposed to provide a good emulation of human audition (indeed, according to Weber's law, for humans, the just noticeable difference in the increment of a physical attribute depends linearly on its magnitude before incrementing). The resulting time series is denoted x(n) in the remainder of this paper (see Figure 1 for an illustration and Gouyon (2005) for further implementation details).

Some of the descriptors required knowing which time points in the music signal correspond to the perceived beat, typically the onsets of quarter-notes. These time points were determined by means of tapping to the music followed by visually guided alignment, as described above. It should be noted that possible confusions in the choice of metrical level for the tempo, either by this procedure or by the listeners, is unlikely to have a significant effect. This is because (1) the important factor is that reference time points are precisely aligned with the sound; (2) the sound descriptors entail averaging values computed on individual beats, hence the relatively small influence of having twice or half as many data points; (3) tempo *per se* is only used in a minor additional correlation of groove versus tempo, unrelated to the descriptors.

Beat Salience

This descriptor was designed to measure the degree of repetitive rhythmical patterning around confortable movement rate. It was based on the estimation of self-similarity of patterns in signal magnitude, as highlighted in a particular representation of the data: the Rhythm Periodicity Function (RPF). This function measures the amount of self-similarity as a function of time lag, and is computed as the autocorrelation function r(τ) of x(n), as follows:

$$r(\tau) = \sum_{n=0}^{N-\tau-1} x(n)x(n+\tau), \forall\, \tau \in \{0 \cdots U\}$$

where N is the number of samples of x(n) and U is the upper limit for the autocorrelation lag τ. We normalized the function so that r(0) = 1 and used a maximum lag U of 5 seconds (i.e. a frequency of 0.2 Hz, or 12 Beats Per Minute). See Figure 2 for an example of RPF.

Self-similarity is an assumption-free approach to detect recurrent periodicities, regardless of where in the signal they may appear (in terms of phase). The Beat Salience was computed as follows: (1) detect peaks in the RPF, (2) consider only peaks corresponding to multiples of subdivisions of the tempo, (3) select the peak closest to 600 ms (i.e. preferred tempo of 100 BPM), (4) retrieve its amplitude.

Fast Metrical Levels

This descriptor was designed to measure the relative magnitude of periodic sound events at metrical levels faster than the beat. It is computed as follows: (1) detect peaks in the RPF, (2) retrieve magnitudes of peaks corresponding to tempo and faster levels, (3) compute the difference between the average magnitudes of tempo and faster level peaks.

Event Density

With a descriptor called *Event Density*, we assessed local energy variability, in our sense a convenient proxy for the perceptual salience of sound events that occur at small temporal

scales, faster than the beat level. Event Density was computed as the x(n) variability per beat, averaged piecewise (see examples of beats over x(n) in Figure 3).

Systematic Microtiming

We measured microtiming deviations of sound events between beats, i.e. *within* the time span of Inter-Beat Intervals (IBIs). For this, the audio signal was first segmented into beat units (see Fig 3) which were subsequently resampled to the same duration to cater for potential variations of IBIs - we chose to resample to 40 points per IBI. To each IBI correspond a particular amplitude pattern. We then computed an average pattern for each excerpt (examples of average patterns are given in Fig 4). Local maxima in the close vicinity of specific positions in the pattern, e.g. strict 16th-notes, indicate systematic timing deviations. For instance, in Figure 4 it can be seen that both the third and fourth 16th-note beats are slightly ahead of their strict positions on the metrical grid (by up to 2.5 percent of the IBI in the case of a, i.e. almost 20 ms at a tempo of 90 BPM).

Given a specific excerpt, Systematic Microtiming was computed as follows: (1) compute the average IBI amplitude pattern (as above), (2) retrieve deviations from strict positions on the metrical grid, (3) weight each deviation by the height of the corresponding peak, (4) normalize these values between 0 and 1, (5) select the maximum.

Unsystematic Microtiming

For a given excerpt, we defined Unsystematic Microtiming as the mean absolute deviation of each beat in this excerpt from its nominal position. The deviation from nominal position is computed in a constant time window of 80 ms[2] centered around each beat and is defined as:

---

[2] window lengths relative the IBI did not produce significant differences

$$dev. = \left| \left(\frac{\sum\limits_{i=1}^{N} i \cdot x(i)}{\sum\limits_{i=1}^{N} x(i)} \right) - \frac{N}{2} \right|$$

where N is the length of a beat segment and $x = \{x(1) \mathsf{K} \ x(N)\}$ the samples of x(n) in the time window around the beat. This constant time window is more appropriate than one proportional to the IBI because humans' temporal perception is not proportional to the tempo when the signal is metrical (i.e. multi-level) as in music (Madison & Paulin, 2010).

Note the important differences between the computations of Systematic and Unsystematic Microtiming (MT). Systematic MT is computed *between* beats while Unsystematic MT is based on deviations *around* beats. For Systematic MT, computing the deviations of the average pattern guarantees that they are not incidental, while Unsystematic MT consists of averaged absolute values of deviations.

Results and discussion

According to the interviews, the listeners were comfortable with the task, although a few indicated that it became somewhat taxing towards the end of the session. Five listeners said they recognized some music examples but it turned out that only three music examples could be correctly identified across all trials. No one found the ratings particularly difficult. Two listeners commented that the word *good* was more difficult to rate than the others because of being "too subjective". Two other listeners commented that they were uncertain whether their groove ratings always followed the given definition "evokes the sensation of wanting to move some part of the body", either because one "didn't know how to move" or because it was "hard to disregard the mental image of people dancing". No data were excluded from analysis based on these observations. Listeners' reports of preferred music to listen to varied quite

naturally, but appeared on a whole to be representative for this age segment with a strong dominance of hard rock, pop, rock, techno, and to some extent world music. There was only occasional mention of (classical) art music, jazz, Latin, or folk music. In other words, all listeners were almost equally unfamiliar with the music presented in the experiment.

One precondition for obtaining a correlation between two variables is that both of them vary. Since the music examples were unsystematically sampled from each population of music genre, we cannot take it for granted that they actually vary either in their sound properties, as measured by the sound descriptors, or in their perceived groove. The confidence intervals of both listeners' ratings of groove and values for each descriptor showed substantial variability in these variables among the 20 music examples within each genre.

Listener ratings

Listeners' consistency was assessed by Cronbach's alpha, both within each rating scale and within and across each genre, as shown in Table 1. Consistency was generally highest for groove, in agreement with previous studies (e.g., Madison, 2006; Madison & Merker, 2003). Low alphas ($< .70$) were found only for Familiar and Good, for some genres, which is to be expected due to individual differences in listening experience and preferences.

A two-way (5 genres × 20 music examples) repeated measures analyses of variance (ANOVA) was used for assessing main effects of genre on groove ratings. We applied it both to the raw ratings and the range-corrected ratings. The range-corrected ratings yielded slightly smaller error terms for all scales, and were therefore used in subsequent analyses. Range-corrected groove ratings showed a significant effect of genre ($F_{4, 72} = 18.56$, $p < 0.00001$). One-way repeated measures ANOVAs were used for assessing both the difference in groove ratings among examples within each genre and their consistency across listeners in terms of separate effect and error variance estimates, which is summarised in Figure 5. F-values ($df = 19, 342$) ranged from 2.92 for Jazz to 8.98 for Samba, demonstrating that music examples did

differ in groove within each genre, and that listeners could consistently rate these differences on the group level. The smaller differences in error variance than in effect variance among genres demonstrate that rating consistency is more a function of the perceived differences among music examples than of the differences among listeners. Figure 5 also depicts mean range-corrected ratings, which show a weak, if any, relation to the variance components, indicating that the level of groove has little correspondence to the perceived differences in groove among examples or to the consistency in groove ratings.

The smaller F for Jazz and to some degree Indian could be an effect of either lesser variability in these samples, or of listeners' lesser ability to discriminate groove in jazz and Indian music due to unfamiliarity with these genres. The mean familiarity ratings were highest (0.56) for Jazz, followed by Samba (0.44), West African (0.37), Indian (0.35), and Greek (0.34). This indicates that low familiarity was not the cause of the small F value for Jazz, a conclusion also supported by the small MS error for Jazz. The smaller MS effect for Jazz is naturally reflected in the smaller Cronbach's alpha, however. Given that the mean rating for Jazz was quite high, it would seem that the present examples of Jazz were quite homogeneous in their level of groove compared to the other genres. Indeed, the sample did not include slower or mellow examples of jazz, such as ballads. In contrast, the low ratings of familiarity and the high MS error for Indian suggests that poorer discriminability may underlie the relatively small F value for this genre. Nevertheless, the generally high level of F values across genres indicates that the ratings were sufficiently consistent to serve as a basis for computing correlations with the sound descriptors.

As mentioned in the introduction, groove is found to be correlated with preference, which might therefore be a likely confound of groove. The rating scale intercorrelations were, across all genres and examples, 0.78 between *groove* and *good*, 0.49 between *groove* and *familiar*, and 0.60 between *good* and *familiar*. By performing the same analyses for *good* (ratings of

preference) as for groove, we assessed the likelihood that the correlations in Table 3 are in fact driven by music preferences. A two-way ANOVA (5 genres × 20 music examples) with range-corrected ratings of *good* as dependent variable showed no significant effect of genre ($F_{4, 72} = 0.79$, $p = 0.53$), in contrast to groove. Effects of music example within each genre were, according to one-way repeated measures ANOVAs, somewhat smaller than for groove, with *F* values ranging from 1.67 to 4.90. These results do not contradict that preference acts as a confounding variable within genre. In subsequent computations of correlations between ratings of groove and the sound descriptors we therefore controlled for *good*, which only moderately reduced correlations.

Audio descriptors

Figure 6 shows descriptive statistics for each combination of the five descriptors and five genres. The computations of the descriptors yields values on different orders of magnitude, and Beat Salience was therefore multiplied with 7.0, Event Density with 15.0, and Unsystematic Microtiming with 50.0 to yield comparable scales. Note that both the means and ranges of descriptor values differ between some genres, which provides a possible basis for spurious correlations across all music examples pooled that may not be valid for any genre separately. Table 2 shows the inter-correlations between the descriptors across all 100 music examples. The highest correlations are on the order of 0.4, indicating a moderate covariation between Beat Salience on the one hand, and Event Density and Fast Metrical Levels on the other. It is of course an open question to which degree this is caused by a covariation of the measured properties in this sample of music or by dependencies between the descriptors.

Correlations between descriptors and ratings

In this section, we examine correlations between the ratings and the six audio signal properties, namely Beat Salience, Event Density, Fast Metrical Levels, Systematic and Unsystematic Microtiming, and tempo. Table 3 shows the correlations between groove ratings and descriptors, both for each genre separately and for all genres pooled together. All

correlations were computed both as-is and controlled for *good*, in which case all except three remained significant (these are indicated by parentheses in Table 3).

The most conspicuous pattern is, firstly, that the strongest correlations are found for Beat Salience and Event Density among the descriptors and for Greek, Indian, and Samba among the genres. Jazz exhibits very small correlations overall, and correlations with West African examples become non-significant when controlled for ratings of *good*. In spite of this discrepancy among genres, many descriptors seem to be able to predict groove across genres, due to a combination of 3-4 relatively high and one or two small or nil correlations. This means that if we had not considered genre, we would have been inclined to think that Beat Salience, Event Density, and Unsystematic Microtiming all generally underlie the experience of groove.

The second most salient observation is that the present rhythmical descriptors generally seem to play a substantially greater role than the present microtiming descriptors, both across and within genres. Note that Systematic Microtiming was negatively correlated with groove for Greek, for which - in other words - non-isochronicity is associated with less groove.

Thirdly, there is an interaction between descriptor property and music genre, in that Systematic Microtiming seems to play no role at all for Indian, Jazz, and West African, but a substantial role for Samba. However, this correlation was absorbed by Beat Salience and Event Density in a multiple regression, reported below, which suggests the possibility that it is an artefact related to interdependencies between these descriptors.

Fourthly, Unsystematic Microtiming seems not to play any role for groove in this sample of music, since these per-genre correlations are non-significant and are furthermore absorbed by other descriptors, as seen in Figure 8. The correlation across genres is significant, but it seems to be larger than one would expect from combining the per-genre correlations and we

therefore suspect that it is in part inflated by the mean genre differences exhibited in Figure 6e.

Finally, tempo seems to play a minor role in this data set. Mean and range of tempi were for Greek 116.35 BPM (range 61-182), Indian 143.1 (104-180), Jazz 128.1 (81-175), Samba, 99.2 (76-160), and for West African 125.4 (90-157), which means that there was ample tempo variability within each genre. The grand mean was 122.4 BPM, which is in the centre of the range of maximally preferred tempo across many different genres (Moelants, 2002). As seen in Table 3, all correlations between groove and tempo were non-significant, and the correlation across genres was furthermore negative, in contrast to a previously observed trend for a positive correlation (Madison, 2006). This difference may be explained by a more heterogenous sample of music in that study, including both up-tempo jazz, ballads, and several other genres, and that in such a wide sample music intended to induce groove tends to be faster than ballads and the like. Another possibility is that groove is related to the examples' proximity to a general preferred tempo. We therefore computed the distance for each example from 100 BPM according to $|tempo - 100|$. This distance was barely significantly correlated with groove (-0.22, $p < .05$), meaning that groove ratings tended to decrease as the tempo moved away from 100 BPM. No per-genre correlations ($r = .04 - .41$) were significant. These inconsistent patterns of correlations show that neither absolute nor preferred tempo has any simple relation to groove.

As mentioned, it is possible that the descriptors are to some extent intrinsically dependent. In addition to this, the measured properties in the music examples may covary, and they may also be redundant in their contributions to groove. To assess their unique contributions, a multiple regression was performed for each genre and for all genres together. The multiple R^2 for all rhythmic descriptors was .537 for all genres pooled, and was surprisingly large for three genres, ranging from .841 for Samba, .709 for Greek, and .631 for Indian. Figure 8

summarises the multiple regression analysis in terms the amount of change in R^2 given by adding descriptors in the order of their relative contribution, that is, according to a stepwise forward entry model. For Jazz and West African only one descriptor passed the entry criterion ($F > 1.0$), and these are therefore not shown in the figure. For Jazz, removing Fast Metrical Levels from the model subtracts 6.43 percent of the total explained variance by all descriptors (14.2%), and removing Beat Salience likewise subtracts 25.65 percent from the total 28.8 percent for West African. A two regressor best subset analysis confirmed that Event Density and Beat Salience were, either together or separately, the best predictors (lowest Mallow's Cp) for all genres except Jazz.

Thus, the multiple regression results largely confirm the pattern seen in Table 3, but show in addition that the contribution of Beat Salience is to a substantial part absorbed by Event Density. The two microtiming descriptors are likewise redundant in relation to the rhythmical descriptors, which can also be said about Fast Metrical Levels for Samba. Note that the highest contribution of Systematic Microtiming, found for Greek, emanates from a negative correlation. A two regressor best subset analysis confirmed that Event Density and Beat Salience were, either together or separately, the best predictors (lowest Mallow's Cp) for all genres except Jazz. When pooling all genres, partial correlations for Event Density (beta = .43) and Beat Salience (beta = .38) remained highly significant (p < .00003) in a model that included all five descriptors as covariates. No significant partial correlations were found for the other three descriptors.

As a general note, one should be aware that correlations across genres might differ in their interpretation from those within genres, because the former may not simply be an aggregate of the latter. Possible mean differences between genres might for example inflate correlations across genres, which can be one explanation for what might be perceived as

disproportionalities in the contributions among Event Density, Fast Metrical Levels, and Unsystematic Microtiming across genres.

To further assess the relation between preference and groove, we inspected correlations between ratings of *good* and the sound descriptors, again both within and across genres, and found that the correlations were consistently smaller for *good*, while the patterns of correlations were largely similar for *groove* and *good*. The exception was Jazz, which both exhibited larger correlations and a different pattern of correlations between *good* and the descriptors than between *groove* and the descriptors. When controlling for groove, correlations with *good* remained significant for both Event Density ($r = .735$, $p < .00001$) and Fast Metrical Levels ($r = .501$, $p = .029$). We note that the correlation between good and groove ratings might reflect either that certain musical properties cause both high *good* and *groove* ratings (i.e. these properties makes the music both groovy and attractive) or that a high *groove* rating causes a high *good* rating (i.e. that groove in itself is attractive). A scenario where ratings of *good* confound the relation between audio signal properties and groove ratings appears implausible, since groove is but one out of many properties that make people appreciate music. In conclusion, preference – as estimated by ratings of good - might underlie ratings of groove for Jazz, but this contribution is considerably smaller for the other genres and could reasonably account for only fractions of the correlations between descriptors and groove.

General discussion

In this study, we asked whether the experience of groove is related to physical properties of the music signal that may be predicated by its function to enable and facilitate entrainment and precise synchronisation among humans. The results indicated ubiquitous and surprisingly strong relations between groove and a number of rhythmical descriptors developed for this

particular purpose, given that this was our first take on descriptors and that the sample of music was arbitrary (except for the choice of genres). The results being exhaustively described above and the implications largely stated in the introduction, we focus this general discussion on the differences between genres, possible caveats, and prospects for future research along these lines.

West African, Samba, and Jazz evoked the highest mean ratings of groove, followed by Greek and Indian, while the variability in ratings among music examples within genres was largest for Samba and Greek and smallest for Jazz. Given the unsystematic sample of music examples, this cannot in any way be generalised to these genres at large. It may however be important for interpreting the differences in relations between groove and descriptors found among genres.

Descriptors were equally non-systematically sampled from the infinite population of possible descriptors, but their design was informed by a set of relatively well-defined acoustic-perceptual demand characteristics. The main caveat here is that there may be other descriptors that tap these characteristics even better than the present ones, and that the present descriptors might unintentionally tap other, unforeseen characteristics. This might in future research be addressed by two main approaches: comparing large numbers of descriptors for their ability to predict groove ratings on large and very homogenous samples of music, and optimizing descriptors with respect to synthetic sound examples with known physical properties.

With the present descriptors, however, Beat Salience and Event Density explained substantially more of the groove ratings than did the remaining descriptors. We could find no support for the idea that microtiming contributes to groove, although this may be due to limitations in the present design, for example in the sampling of music examples. Nevertheless, the results show that that (1) correlations with microtiming descriptors were

generally small and non-significant; (2) For genres exhibiting substantial correlations with microtiming descriptors it was either negative (Greek and Indian) or absorbed by other descriptors (Samba); (3) Although Unsystematic MT was significant for all genres pooled, it seems to be inconsistent with the per-genre correlations and is more likely attributed to differences in desriptor means among genres. This should be interpreted in the light of the present design, which was concerned with general effects according to our assumptions of a phylogenetic trait. The situation we envisioned to be relevant was thus that people without particular knowledge or training encounter music they have not heard before, as may be the case when attending a ritual in a foreign tribe. Groove should be induced even in this situation in order to reflect its functional significance.

It is possible that appreciating microtiming requires learning by repeated listening to a particular piece of music, or possibly considerable experience with music in a particular genre. The present participants were not selected for this. However, that would indicate that microtiming has relevance for the suggested functional perspective. One could also envision that some aspects of our inborn abilites become exploited for different purposes, a so-called spandrel. One apparent and dramatic change to consider in this context is the unlimited possibility to record and reproduce music that has emerged in only 50 years. In addition to a drastic increase in the total amount of music exposed to, it provides us with the artificial and uncanny experience of hearing the exact same performance umpteen times.

Further assessment of the role of microtiming would therefore require quite different types of designs, in which one attempts to disentangle effects of microtiming per se from confounding factors that ensue from listeners' high familiarity and expertise.

On a related note, Samba featured a substantial zero-order correlation between groove and Systematic Microtiming, although it was largely absorbed by the rhythmic descriptors. Samba has a rather fixed rhythmical structure that repeats every measure and is highly similar to the

metrical structure, a characteristic likely to invite performers to "do something more". Vienna Waltz is similar in these respects, and is known to feature very large deviations from isochrony on the level of the measure (Gabrielsson, 1985). It is also known among musicians that it is difficult to perform Samba correctly; even the most fundamental rudiments require substantial training to reach an acceptable result. If this difficulty stems from learning the stylistically prescribed microtiming patterns, then a considerable amount of learning would analogously be required to perceive the same patterns. This is true of course whether they actually facilitate groove or not.

Samba was characterised by very high correlations with no less than four descriptors, which naturally proved to be highly redundant. This suggests that Samba employs more means to induce groove than the other genres in this study, as might be expected by music dedicated to dancing. The overwhelming contribution of Event Density to groove is accounted for by the tendency to accentuate Fast Metrical Levels, defined by high-pitched drums (e.g. pandeiro).

West African is also designed for movement, but seems to employ the strategy to focus on one means, namely Beat Salience. In other words, different means may be used to reach the same goal, and the two means found most important in the present study are both cenceivably in agreement with the suggested function to facilitate synchronisation and co-ordination.

Jazz is strongly associated with groove, and was indeed rated relatively high in groove across music examples. However, Jazz exhibitied the smallest correlations between ratings and descriptors, which might suggest that we did not manage to "break the code" for the Jazz groove. In jazz, groove has in particular been attributed to the swing ratio, the relative duration of the two "swung" notes in the rhythmic ostinato so characteristic for much classical jazz music (Keil et al., 1994; Busse, 2002). Not all jazz features swing in this sense, but 15 of the present Jazz examples did. Differences in swing ratio should naturally manifest

themselves in the microtiming descriptors but, contrary to popular belief, they exhibited no significant correlation with groove. However, this corresponds perfectly with the finding that the swing ratio is trivially related to tempo (Friberg & Sundström, 2002), suggesting that its purpose is merely to make the two intervals discriminable and in effect create a rhythmical pattern.

Jazz is also commonly associated with expressive performance - that the listeners' experience depends on how you play rather than on what you play. Expressive timing should also manifest itself in one of the microtiming descriptors, but no such correlation did materialise. What did stand out, however, was the relatively small variability in groove ratings among the Jazz examples, as seen in Figure 5. Apart from the possibility that the present descriptors fail to measure some essential rhythmic property related to the experience of groove in jazz music specifically, other reasons might therefore be that the present examples of Jazz were quite homogeneous in their (high) level of groove. It might even be that this is typical for jazz in general, so long as slower or mellow examples of jazz, such as ballads, are not included.

Greek is the only genre in this study that exhibits negative correlations between groove and Systematic Microtiming deviations. In other words, these examples feature systematic temporal deviation patterns whose presence decreases the experience of groove. The most obvious interpretation is that this genre requires precise metrical performance, and that poor performance in this regard diminishes its tendency to induce movement. Greek was also the only genre to demonstrate a unique contribution of Fast Metrical Levels (cf. Drake, Gros, & Penel, 1999) to the experience of groove. It can be argued that Fast Metrical Levels is a special case of Event Density, since both pertain to events faster than the tempo. It could not be predicted beforehand which of these varieties would be most psychologically relevant for groove; recurrent events that belong to a particular metrical level for the former (e.g. 8th or

16th notes) or any events that occur between quarter notes for the latter (including, e.g., syncopations or sounds with continuosly varying metrical positions). As the more inclusive descriptor turned out to play a major role, it is likely that it absorbs some of the variance that is in fact due to Fast Metrical Levels.

Indian music also exhibits a trend for negative correlations with Systematic Microtiming deviations, which is quite conceivable given the strong emphasis on isochrony and a strict adherence to metrical subdivisions, perfected through many years of devoted training by traditional Indian musicians. Indeed, their musical communication is to a large extent focused explicitly on rhythmical patterns, in agreement with the pattern of correlations.

Finally, the lack of a consistent relation between beat tempo and groove may seem surprising, given that tempo is considered such a critical and powerful aspect among musicians. There were trends for negative correlations between tempo and groove for Indian and across all genres, but since this is in the opposite direction from what has been previously found (Madison, 2003; 2006), they serve rather as further argument against any general effect of tempo. Indeed, the descriptors that were most clearly associated with groove - Beat Salience, Event Density, and to some extent Fast Metrical Levels – are all defined and measured independently from tempo (except that that the peak closest to 100 BPM was selected for Beat Salience). This is of course consistent with the notion emphasized here that metrical structure and rhythmical patterning provide multiple temporal levels across a wide range of intervals surrounding the beat. More importantly, the temporal space in the musical range seems to be relatively evenly filled, regardless of the beat tempo: Music with slow tempi have more levels faster than the beat compared to music with fast tempo, which correspondingly has more levels slower than the beat. This leads to the effect that the perceived speed of music in original tempi exhibits a shallower function of beat tempo than does the perceived speed of music with artificially changed tempi (Madison and Paulin,

2010). This is best explained by the fact that if we take a piece of music recorded in 100 BPM and increase it to 120 BPM, then all the faster and slower rhythmical levels will also be increased in rate. Music that is played in its natural tempo, however, tends to have optimized the fastest and slowest levels to create a moderate impression of speed.

The results suggest a number of hypotheses for future research. For example, should we dismiss tempo as a relevant factor behind groove, or might there be an optimal individual tempo for groove? This might be addressed by assessing the correspondence between the most salient periodicities in real music, as measured by the present descriptors, and anthropometric factors (cf. Todd, Cousins, & Lee, 2007). Could there be a relation between the number of devices – as tapped by different descriptors - and the extent to which the music is intended for dancing? Can microtiming induce groove? If so, does that require more prior knowledge or information than does induction by means of the descriptors found most effective here? If so, what knowledge or information is required? Some of these questions require systematic experimental manipulation of the variables of interest, using synthesised music examples.

A final issue is how music genres and their respective features might further contribute to interpreting the results in future studies. Although genre is a fuzzy concept, in particular with respect to recent, urban subcultures, listeners' ratings of genre nevertheless account for a considerable amount of the variance among music examples (cf. Aucouturier & Pachet, 2003). That genre can be discriminated based on very brief examples (down to fractions of a second) demonstrates that low-level acoustic properties differ between genres (Gjerdingen et al., 2008). These facts buttress our assumption that the present genres, presumably largely due to their different geographical areas of origin, differ in their fundamental structural as well as their low-level acoustic properties. In this light, it seems all the more remarkable that they all induce groove to a comparable extent, and that the underlying acoustic properties overlap for

some genres. Nevertheless, these observations are in agreement with the status of groove as a human universal.

In all, the results are well in agreement with our predictions based on a functional role of rhythmic music for entrainment and synchronisation among individuals. They demonstrate a highly significant role of physical properties; the feasability of the methods applied, and support for the underlying functional perspective. The present study strengthens the position of groove as a perceptually salient dimension of music, and shows that groove may be mediated in similar ways in different genres of music. Computational modelling of rhythmic and other acoustic features along the lines demonstrated here may profitably be applied to increasing precision in defining the structural properties associated with groove. In addition to developing the understanding of the functions of groove, it may ideally provide new knowledge about the cognitive and perceptual processes underlying timing in general.

References

Aucouturier, J.-J. & Pachet, F. (2003). Representing musical genre: a state of the art. Journal of New Music Research, 32, 83-93.

Bengtsson, S. L., Ullén, F., Ehrsson, H. H., Hashimoto, T., Kito, T., Naito, E., Forssberg, H., & Sadato, N. (2008). Listening to rhythms activates motor and premotor cortices. Cortex, 45, 62-71.

Bilmes, J. (1993). Timing is of the essence: Perceptual and computational techniques for representing, learning, and reproducing timing in percussive rhythm. Media Lab, Massachusetts Institute of Technology.

Blood, A. J. & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. Proceedings of the National Academy of Sciences, 98, 11818-11823.

Busse, W. G. (2002). Toward objective measurement and evaluation of jazz piano performance via MIDI-based groove quantize templates. Music Perception, 19, 443-461.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. Cerebral Cortex, 18, 2844-2854.

Dixon, S. E., Gouyon, F., & Widmer, G. (2004). Towards characterisation of music via rhythmic patterns. In Proceedings of the 5th International Conference on Music Information Retrieval Barcelona, Spain.

Drake, C., Gros, L., & Penel, A. (1999). How fast is that music? The relation between physical and perceived tempo. In S.W.Yi (Ed.), Music, Mind, and Science (pp. 190-203). Seoul: Seoul National University Press.

Fraisse, P. (1982). Rhythm and tempo. In D.Deutsch (Ed.), The psychology of music (pp. 149-180). London: Academic Press.

Friberg, A. & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. Music Perception, 19, 333-349.

Gabrielsson, A. (1985). Interplay between analysis and synthesis in studies of music performance and music experience. Music Perception, 3, 59-86.

Gjerdingen, R. O. & Perrott, D. (2008). Scanning the dial: the rapid recognition of music genres. Journal of New Music Research, **37**, 93-100.

Gouyon, F. (2005). A computational approach to rhythm description. PhD Universitat Pompeu Fabra, Barcelona.

Gouyon, F. (2007). Microtiming in "Samba de Roda" - Preliminary experiments with polyphonic audio. In Brazilian Symposium on Computer Music.

Gouyon, F., Herrera, P., & Cano, P. (2002). Pulse-dependent analyses of percussive music. In AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio Espoo, Finland.

Grahn, J. A. & Brett, M. (2007). Rhythm perception in motor areas of the brain. Journal of Cognitive Neuroscience, 19, 893-906.

Hibi, S. (1983). Rhythm perception in repetitive sound sequence. Acoustical Society of Japan (E), 4, 83-95.

Hodges, D. (1989). Why are we musical? Speculations on the evolutionary plausibility of musical bahavior. Bulletin of the Council for Research in Music Education, 99, 7-22.

Iyer, V. S. (1998). Microstructures of feel, macrostructures of sound: Embodied cognition in West African and African-American Musics. Unpublished PhD thesis. University of California.

Iyer, V. S. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. Music Perception, 19, 387-414.

Kauranen, K. & Vanharanta, H. (1996). Influences of aging, gender, and handedness on motor performance of upper and lower extremities. Perceptual and Motor Skills, 82, 515-525.

Keil, C. (1995). The theory of participatory discrepancies: a progress report. Ethnomusicology, 39, 1-19.

Keil, C. & Feld, S. (1994). Music grooves. Chicago: University of Chicago Press.

Khalfa, S., Roy, M., Rainville, P., Dalla Bella, S., & Peretz, I. (2008). Role of tempo entrainment in psychophysiological differentiation of happy and sad music? International Journal of Psychophysiology, 68, 17-26.

Klapuri, A., Eronen, A. J., & Astola, J. T. (2006). Analysis of the meter of acoustic musical signals. IEEE Transactions on Audio, Speech and Language Processing, **14**, 342-355.

Kohno, M. (1993). Perceptual sense unit and echoic memory. International Journal of Psycholinguistics, 9, 13-31.

Levitin, D. J. & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. Perception & Psychophysics, 58, 927-935.

Lykken, D. T., Rose, R., Luther, B., & Maley, M. (1966). Correcting psychophysiological measures for individual differences in range. Psychological Bulletin, 66, 481-484.

Madison, G. (2001). Variability in isochronous tapping: higher-order dependencies as a function of inter tap interval. Journal of Experimental Psychology: Human Perception and Performance, 27, 411-422.

Madison, G. (2003). Perception of jazz and other groove-based music as a function of tempo. In R. Kopiez, A. C. Lehmann, I. Wolther, & C. Wolf (Eds.), Proceedings of the 5th triennial ESCOM conference (pp. 365-367). Hannover, Germany: School of music and drama.

Madison, G. (2006). Experiencing groove induced by music: consistency and phenomenology. Music Perception, 24, 201-208.

Madison, G. (2009). An auditory illusion of infinite tempo change based on multiple temporal levels. PLoS ONE, 4, e8151.

Madison, G. & Merker, B. (2002). On the limits of anisynchrony in pulse attribution. Psychological Research, 66, 201-207.

Madison, G. & Merker, B. (2003). Consistency in listeners' ratings as a function of listening time. In R. Bresin (Ed.), Proceedings of the Stockholm music acoustics conference (pp. 639-642). Stockholm, Sweden: Royal College of Technology.

Madison, G. & Paulin, J. (2010). Ratings of speed in real music as a function of both original and manipulated beat tempo. Journal of the Acoustical Society of America, 128, 3032-3040.

McGuiness, A. (2005). Microtiming deviations in groove. Master's thesis Centre for New Media Arts, The Australian National University.

McNeil, W. H. (1995). Keeping together in time. Dance and drill in human history. Cambridge, MA: Harvard University Press.

Merker, B. (1999). Synchronous chorusing and the origins of music. Musicae Scientiae, Special issue 1999-2000, 59-73.

Merker, B. (2000). Synchronous chorusing and human origins. In N.L.Wallin, B. Merker, & S. W. Brown (Eds.), The origins of music (pp. 315-327). Cambridge, MA: MIT Press.

Merker, B., Madison, G., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. Cortex, 45, 4-17.

Moelants, D. (2002). Preferred tempo reconsidered. In C.Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), (pp. 580-583). Adelaide: Causal Productions.

Möckel, M., Röcker, L., Störk, T., Vollert, J., Danne, O., Eichstädt, H., Müller, R., & Hochrein, H. (1994). Immediate physiological responses of healthy volunteers to different types of music: cardiovascular, hormonal and mental changes. European Journal of Applied Physiology, 68, 451-459.

Nettl, B. (2000). An Ethnomusicologist contemplates universals in musical sound and musical culture. In N.L.Wallin, B. Merker, & S. W. Brown (Eds.), The origins of music (pp. 463-472). Cambridge, MA: MIT Press.

Pinker, S. (2003). The blank slate: The modern denial of human nature. London: Penguin Books.

Repp, B. H. (1998). A microcosm of musical expression I: Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. Journal of the Acoustical Society of America, 104, 1085-1100.

Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. Journal of Motor Behavior, 35, 355-370.

Repp, B. H. (2005). Rate limits of on-beat and off-beat tapping with simple auditory rhythms: 2. The roles of different kinds of accent. Music Perception, 23, 165-188.

Riecker, A., Wildgruber, D., Mathiak, K., Grodd, W., & Ackermann, H. (2003). Parametric analysis of rate-dependent hemodynamic response functions of cortical and subcortical brain structures during auditorily cued finger tapping: a fMRI study. Neuroimage, 18, 731-739.

Roederer, J. (1984). The search for a survival value of music. Music Perception, 1, 350-356.

Seppanen, J. (2001). Computational models of musical meter recognition. Master's thesis Tampere University of Technology.

Shaffer, L. H. & Todd, N. P. M. (1994). The interpretative component in musical performance. In R.Aiello (Ed.), Musical perceptions (pp. 258-270). New York: Oxford University Press.

Todd, N. P. M. (2001). Evidence for a behavioral significance of saccular acoustic sensitivity in humans. Journal of the Acoustical Society of America, 110, 380-390.

Todd, N. P. M., Cousins, R., & Lee, C. S. (2007). The contribution of anthropometric factors to individual differences in the perception of rhythm. Empirical Musicology Review, 2, 1-13.

van Noorden, L. & Moelants, D. (1999). Resonance in the perception of musical pulse. Journal of New Music Research, 28, 43-66.

Waadeland, C. H. (2001). "It don't mean a thing if it ain't got that swing" - Simulating expressive timing by modulated movements. Journal of New Music Research, 30, 23-37.

Wright, M. & Berdahl, E. (2006). Towards machine learning of expressive microtiming in brazilian drumming. In Proceedings of the International Computer Music Conference New Orleans, Louisiana, USA.
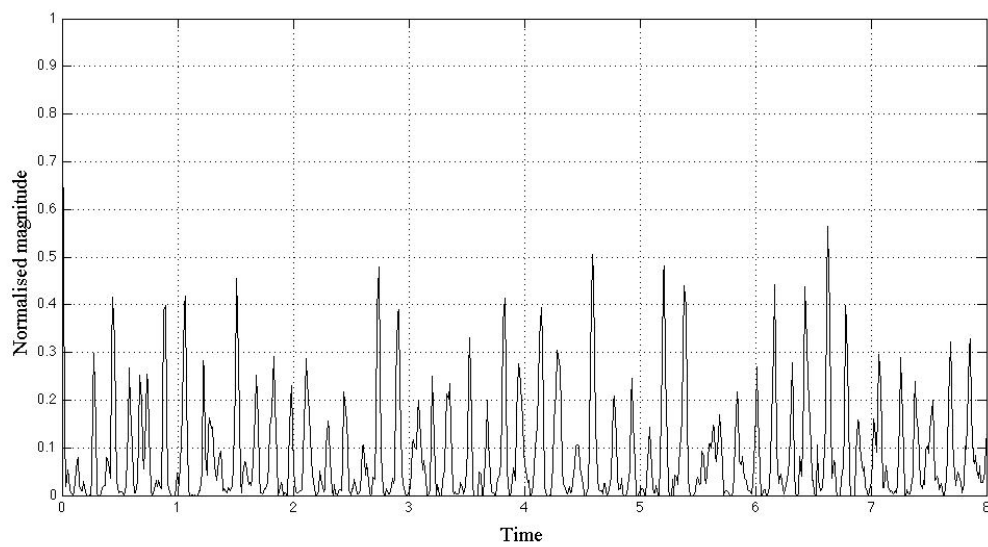
Figure 1. Example of pre-processed signal representation x(n) on a short excerpt of Samba music. This pre-processing is common for all descriptors. Axes represent normalised magnitude vs time (in seconds).
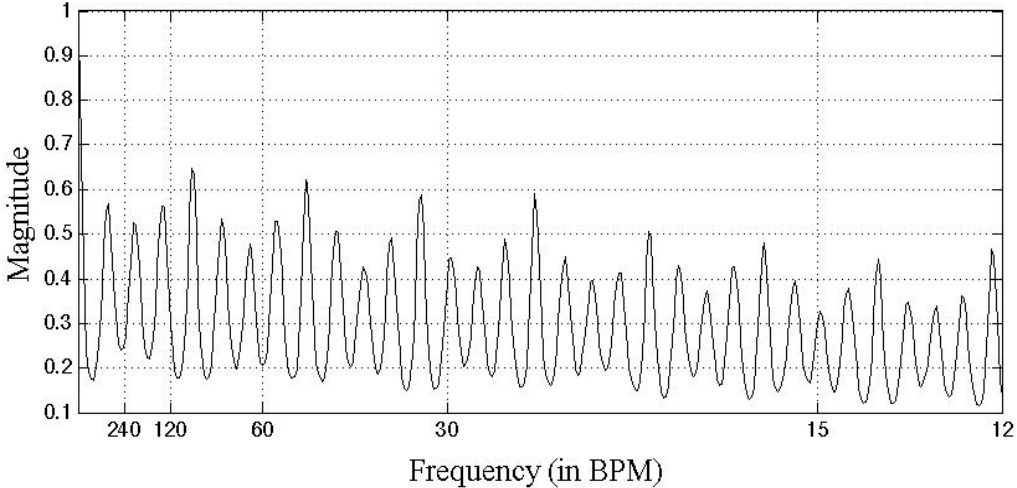
Figure 2. Example of Rhythm Periodicity Function, used to compute the descriptors Beat Salience and Fast Metrical Levels. It represents pulse magnitude vs pulse frequency (in BPM), and is computed on the same sound excerpt as in Figure 1.
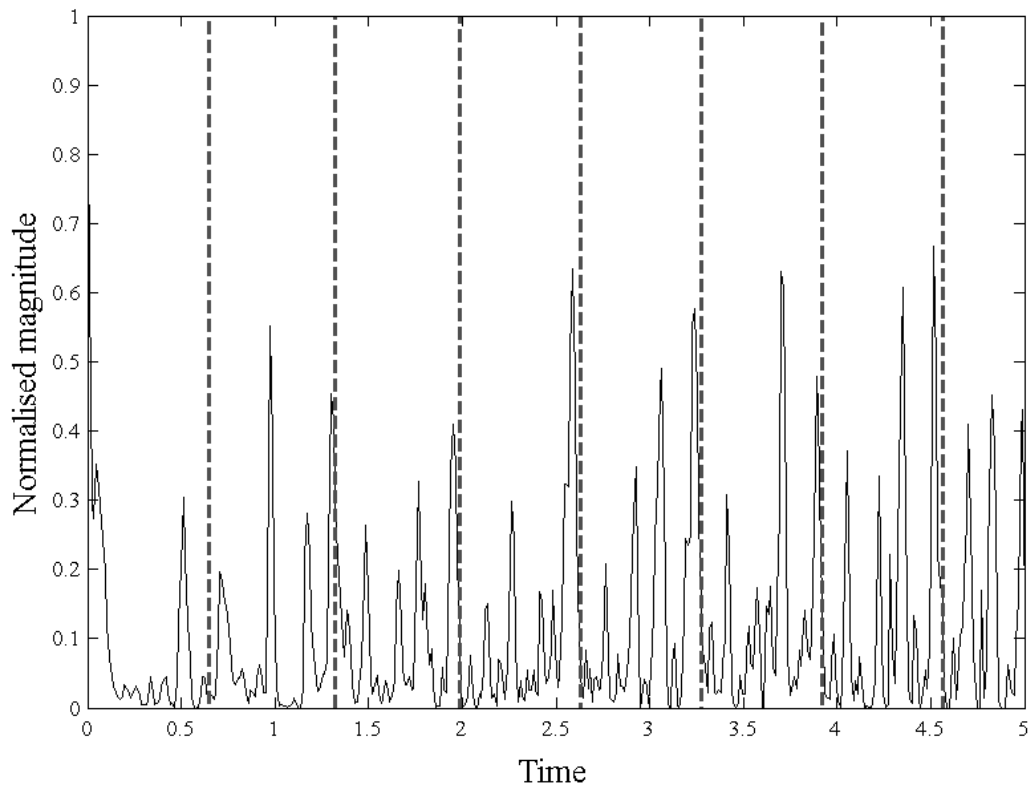
Figure 3. Example of pre-processed signal representation x(n), and quarter note beats (vertical dashed lines), used for the computation of Event Density, Systematic Microtiming and Unsystematic Microtiming. This example comes from a different Samba excerpt than in previous figures.
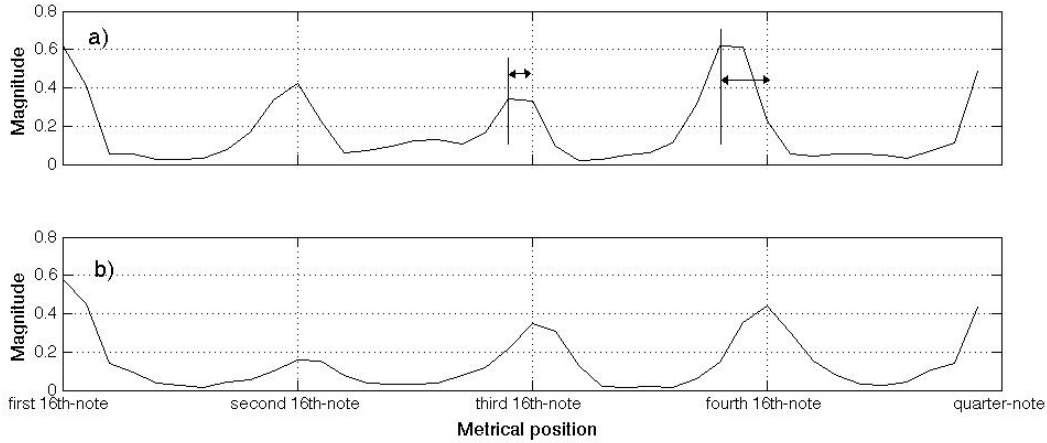
Figure 4. Illustration of the computation of Systematic Microtiming and Unsystematic Microtiming. Average amplitude patterns in two different excerpts, a) and b). Note that in a) there are deviations from metrical positions, indicated by arrows, around the third and fourth 16th-note, while in b) there are none.
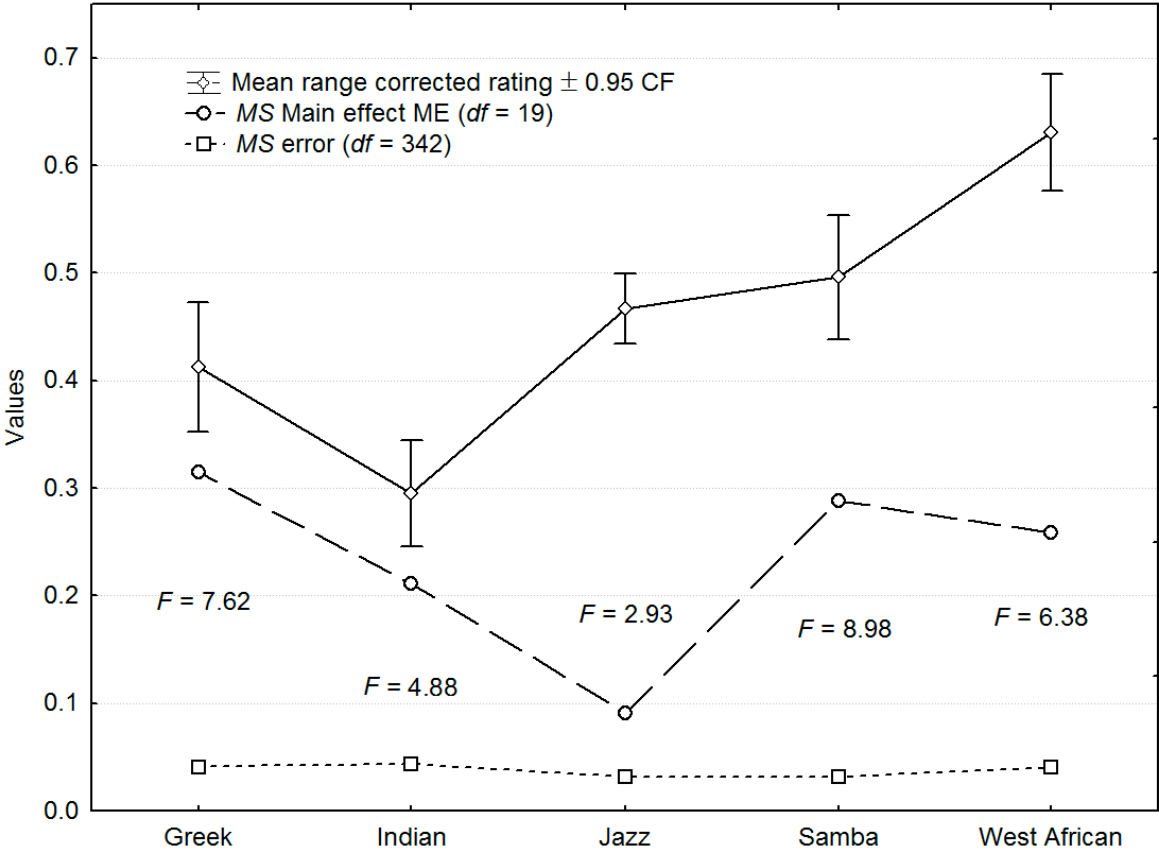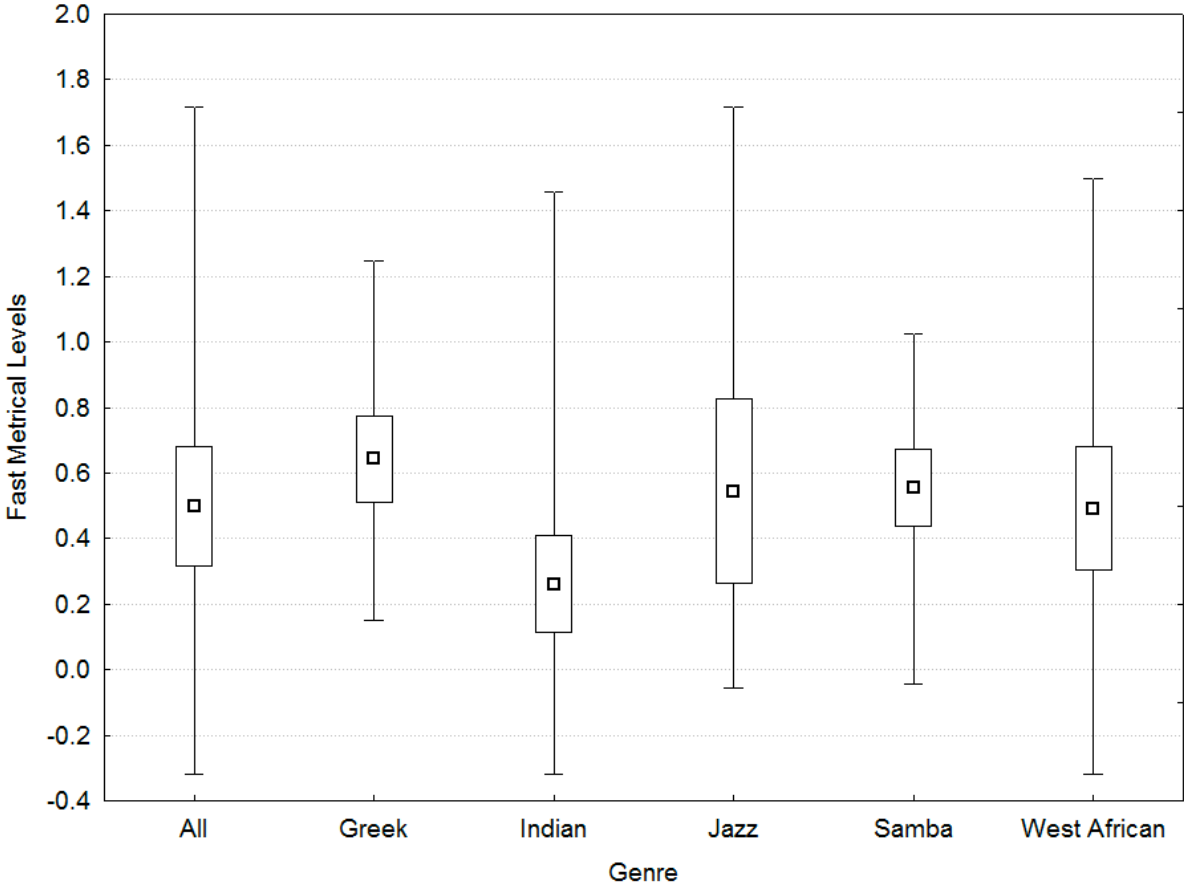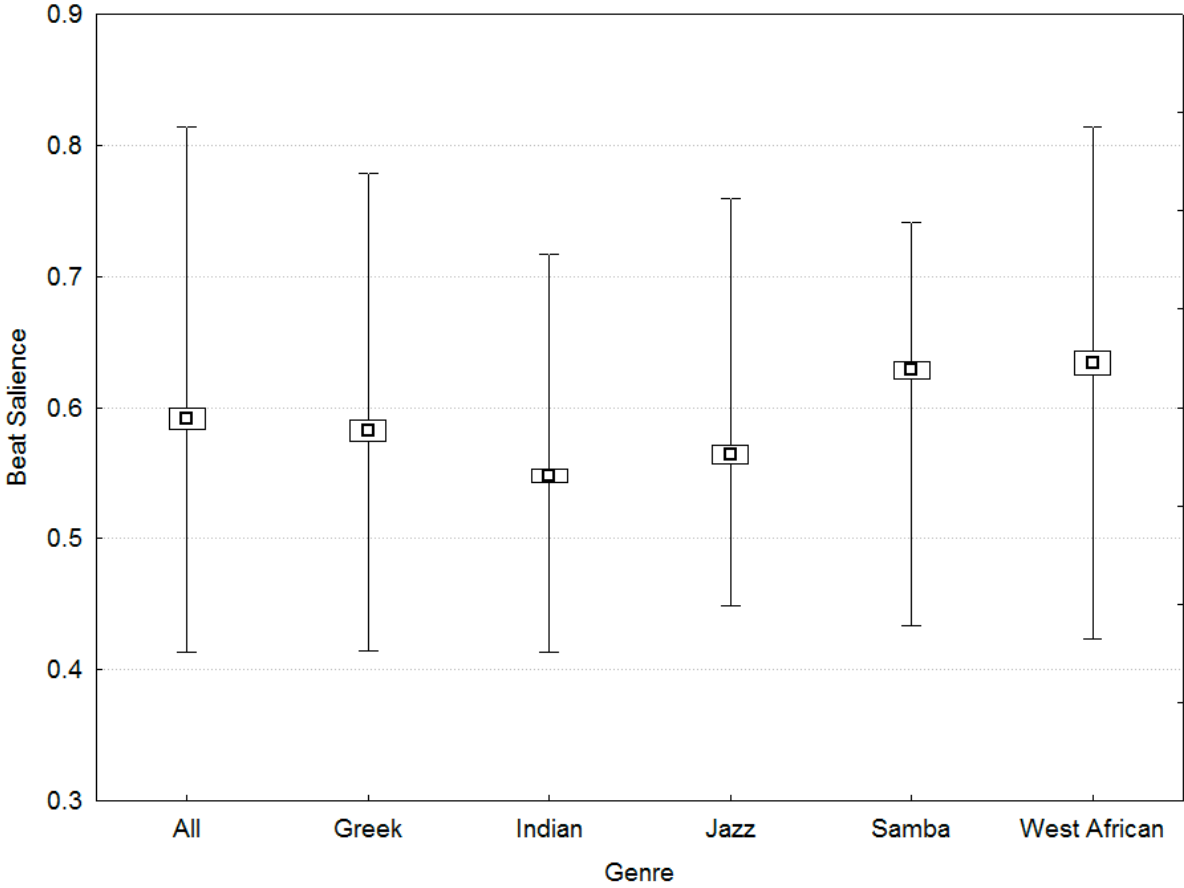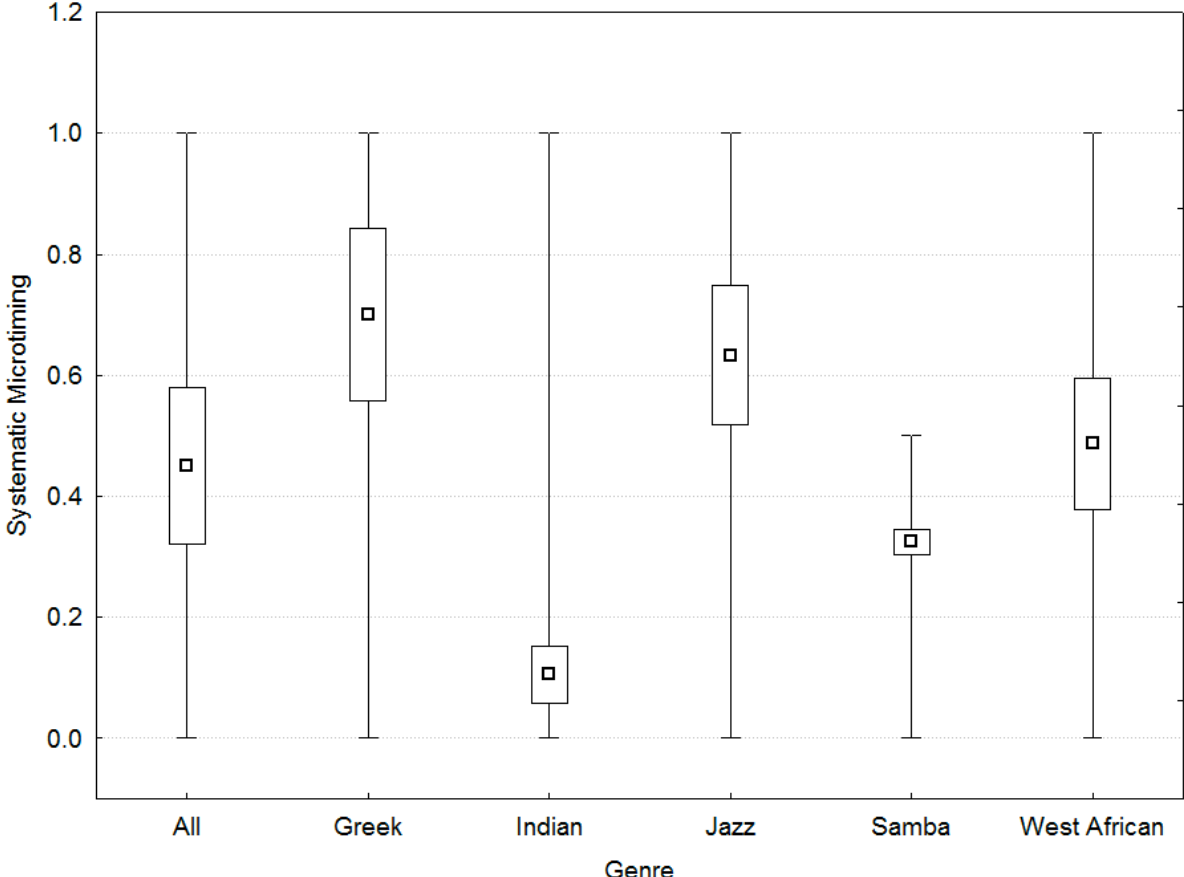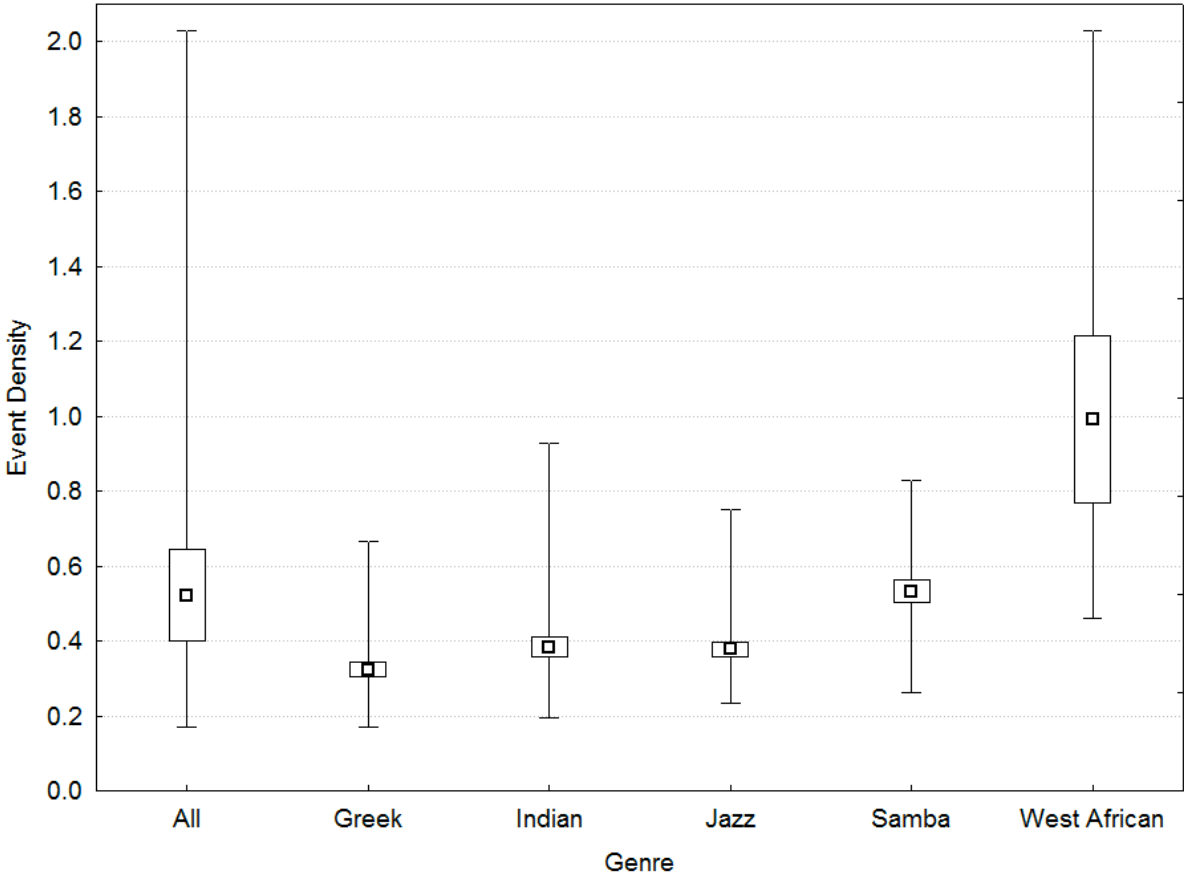
Figure 5. Mean ratings of groove (range-corrected) and their ANOVA variance components as a function of genre. Error bars denote 0.95 confidence intervals.
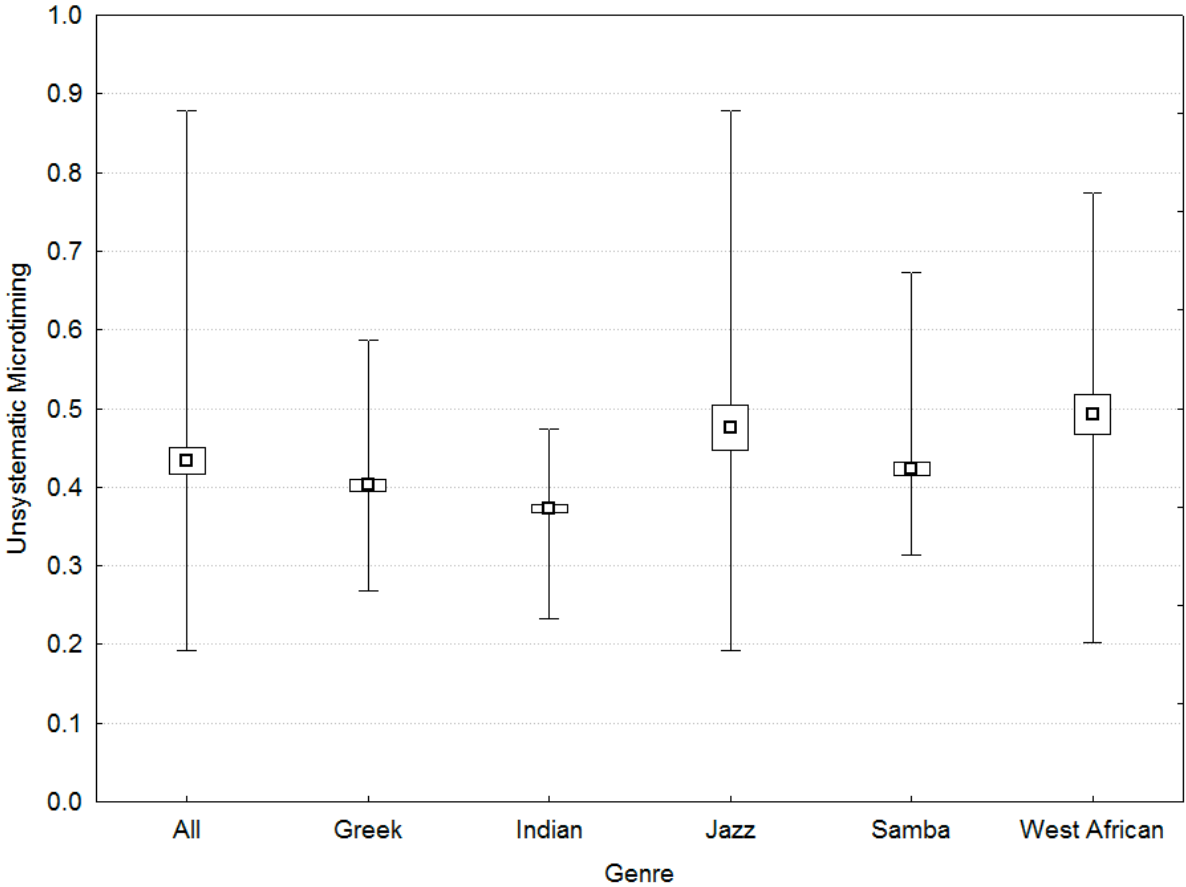
Figure 6. Descriptive statistics for each genre and descriptor. Each descriptor is shown in a separate panel, in which squares indicate the mean, boxes the variance, and whiskers the minimum and maximum descriptor values for the 20 music examples in each genre. Beat Salience was multiplied with 7.0, Event Density with 15.0, and Unsystematic Microtiming with 50.0 to yield comparable scales.
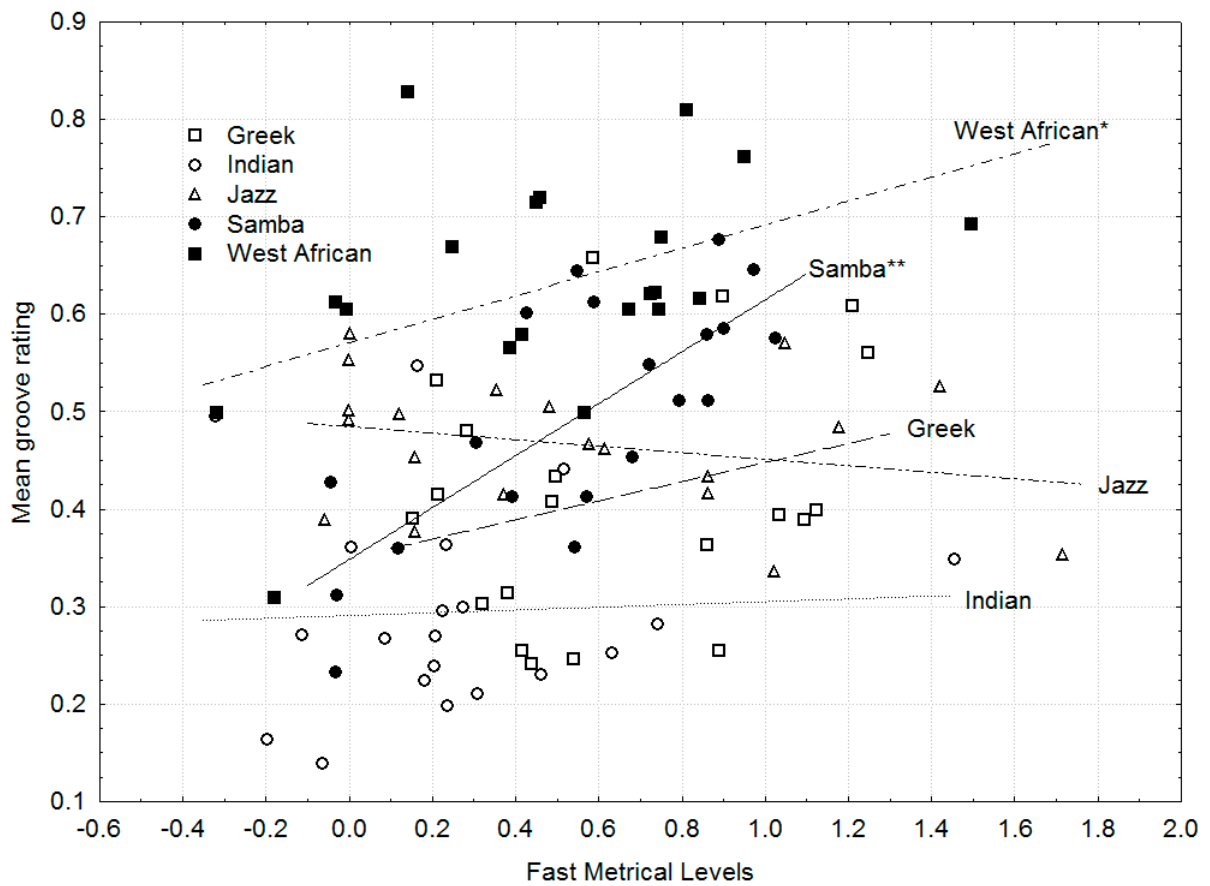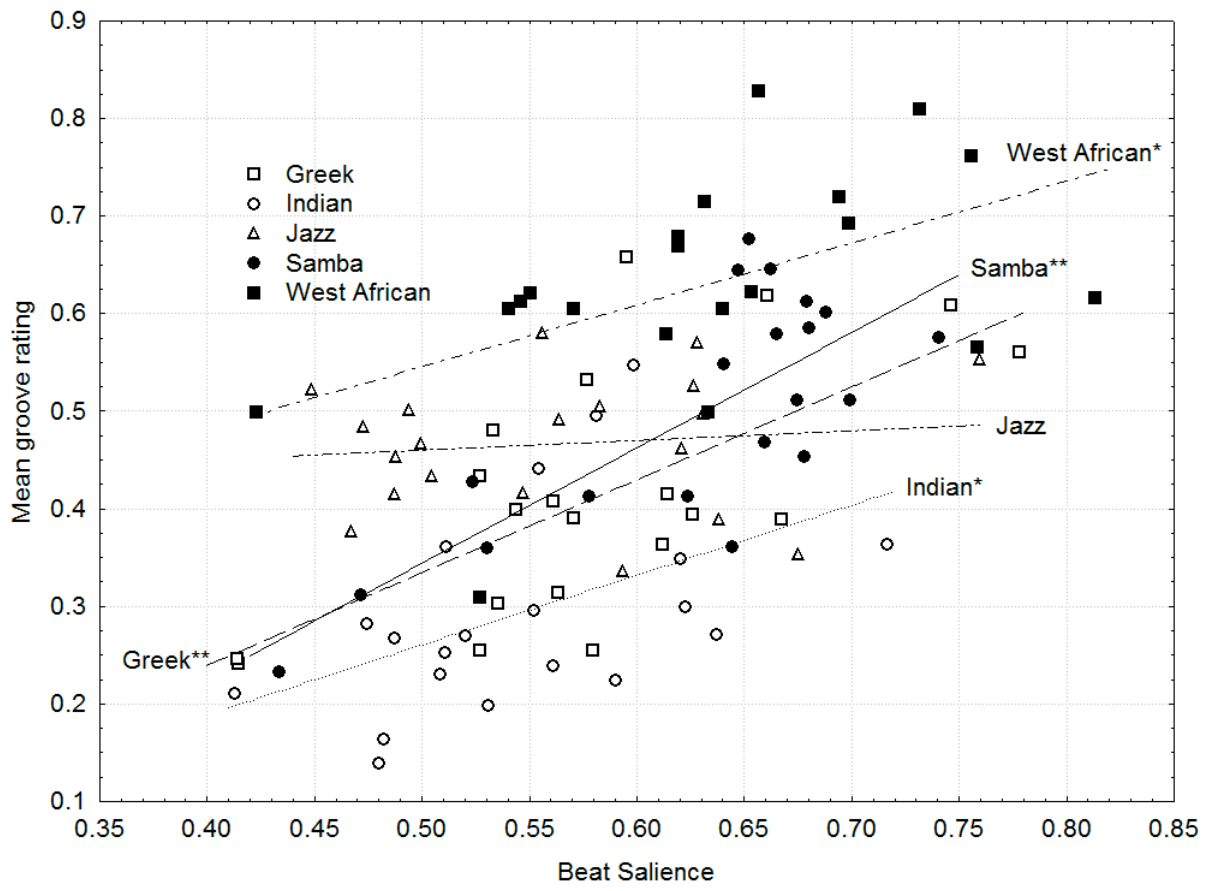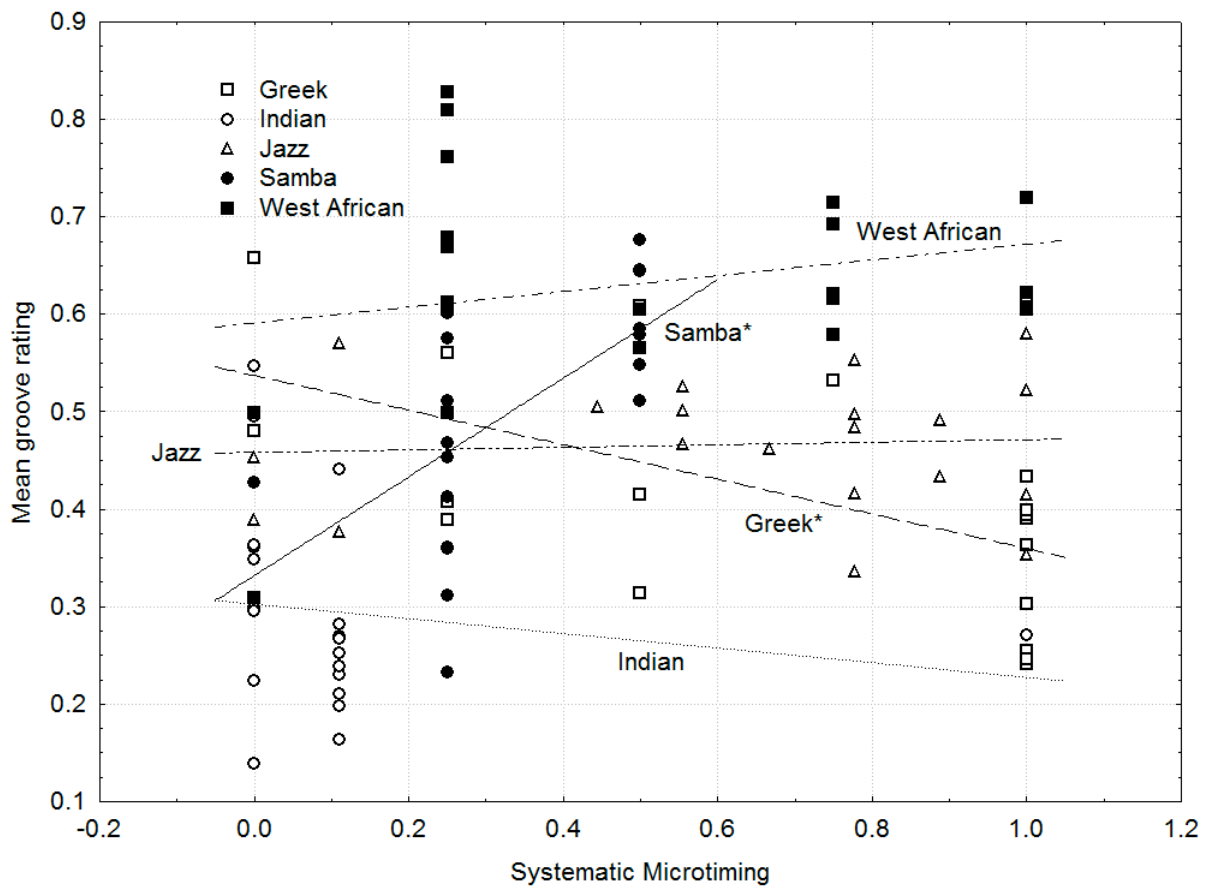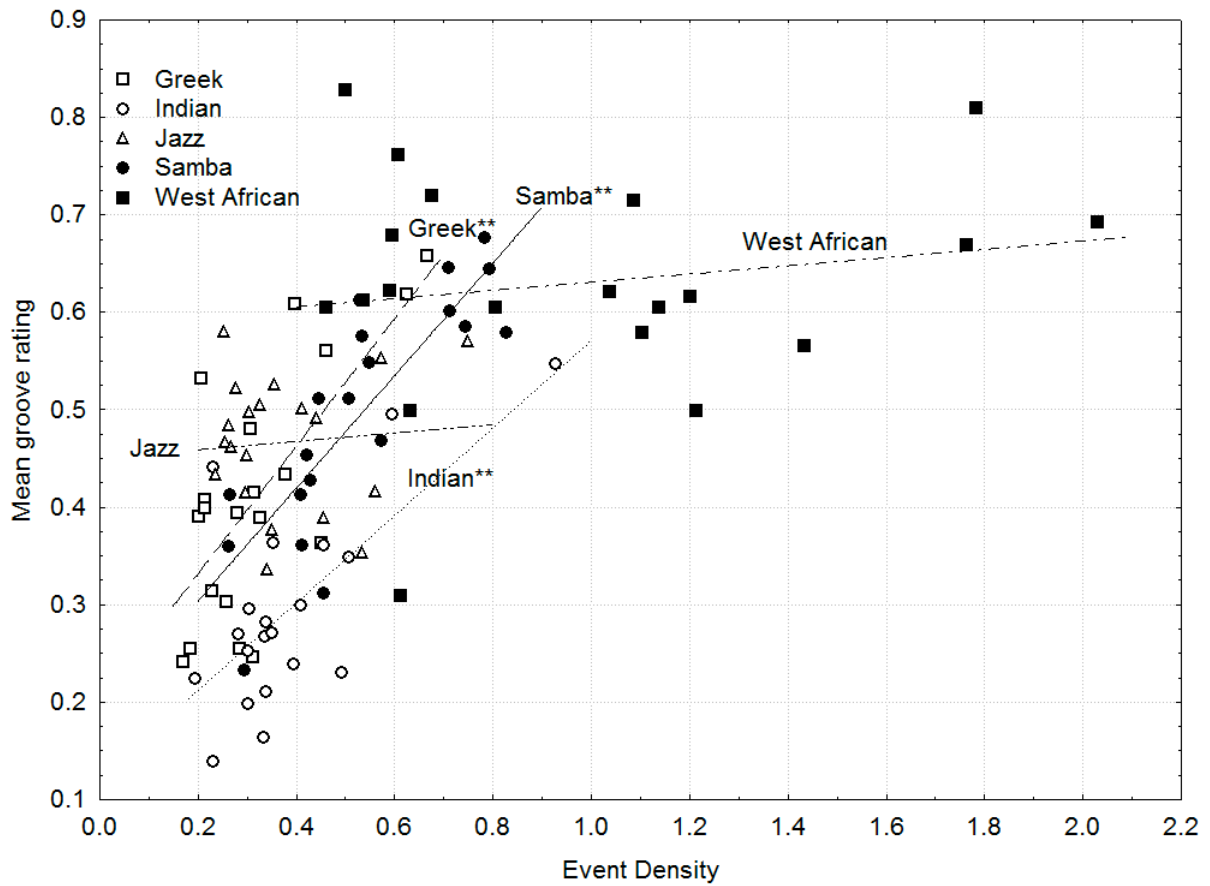
Figure 7. Scatterplots of the relation between descriptor values and mean groove ratings (across participants), for each genre separately (marked with different symbols). The top left panel shows this relation for Beat Salience, top right for Fast Metrical Levels, bottom left for Event Density, and bottom right for Systematic Microtiming. Regression lines are fitted to the points representing the music examples of each genre separately, and asterisks in the legends indicate the alpha level (*$p < 0.05$; **$p < 0.001$). Beat Salience was multiplied with 7.0 and Event Density with 15.0 to yield comparable scales.
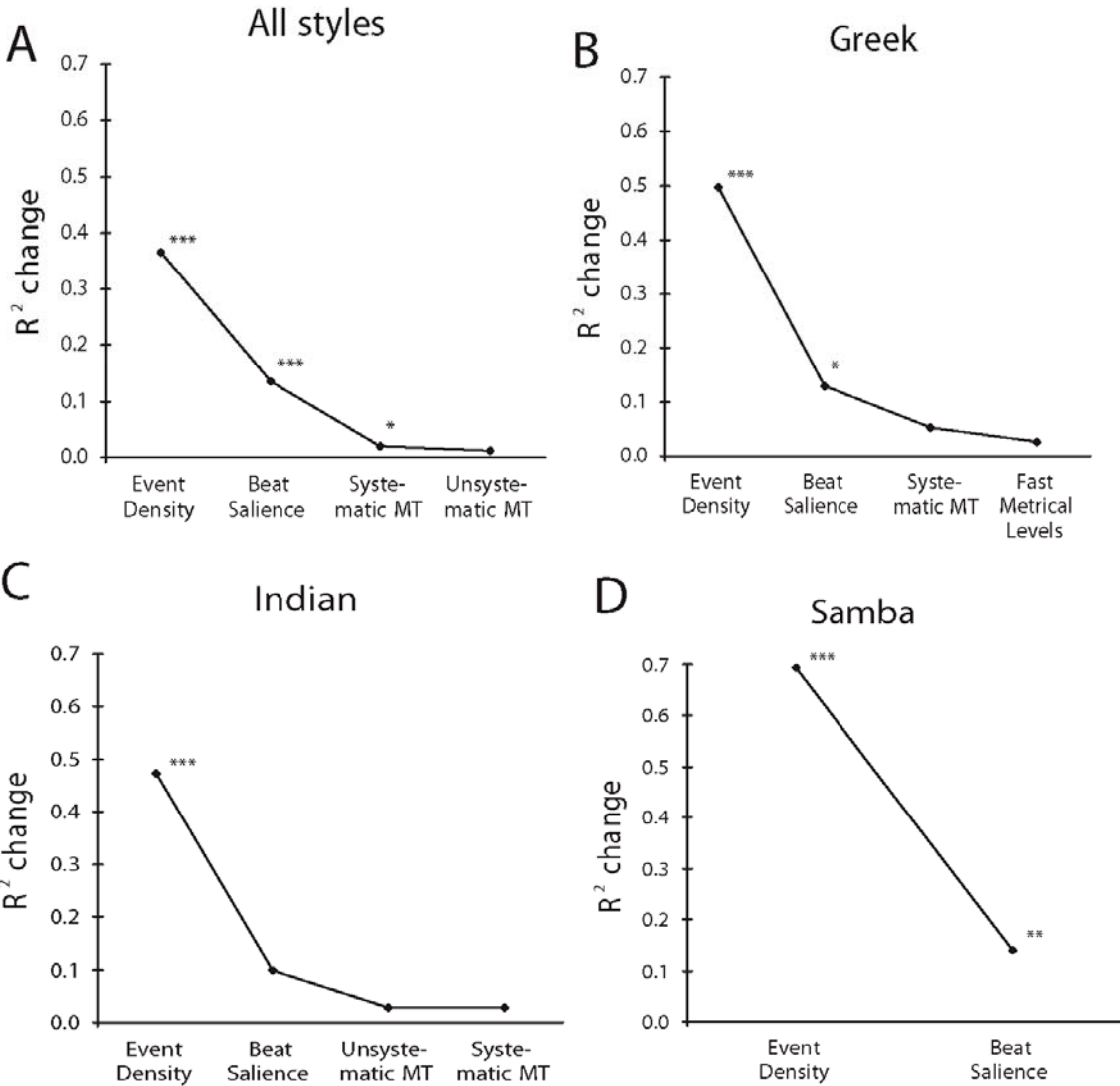
Figure 8. Changes in explained variance according to stepwise linear regression of groove ratings on the five descriptors. Note. *** = p < .001, ** = p < .01, * = p < .05. Note that the contribution of Systematic MT for Greek refers to a negative correlation.

Table 1. Interrater Reliability. Cronbach's alpha for the three scales and the five genres both together and separately.

	Groove	Familiar	Good
All genres	.884	.772	.708
Greek	.861	.460	.707
Indian	.789	-	.858
Jazz	.787	-	.586
Samba	.880	-	.800
West african	.843	-	.808

Note. Alpha could not be computed for the familiarity of most genres because of too little variance (many ratings were zero).

Table 2. Correlations Between Descriptors. Pearson *r* values are given for zero-order correlations found when pooling all music examples (n = 100). Significant correlations are indicated with asterisks (* p < 0.05; ** p < 0.001). For significant correlations, the values in parentheses show the corresponding *r* values found when splitting the sample into even (n = 50) and odd (n = 50) music examples.

	Fast Metrical Levels	Event Density	Systematic Microtiming	Unsystematic Microtiming
Beat Salience	.44**	.41**	.05	.23*
	(.42, .46)	(.45, .37)		(.20, .25)
Fast Metrical Levels		.16	.27*	.09
			(.23, .31)	
Event Density			-.04	.25*
				(.21, .28)
Systematic Microtiming				.16

Table 3. Correlations Between mean Groove Ratings and Descriptors. Pearson *r* values are given for zero-order correlations for all genres pooled (n = 100 examples), and for each of the five genres separately (n = 20 examples/genre). Significant correlations are indicated with asterisks (*, p < 0.05; **, p < 0.001). Correlations with groove ratings remained significant when controlling for Good ratings, except in three cases. In these cases, the *r* value is given in parentheses.

	Beat Salience	Fast Metrical Levels	Event Density	Systematic Microtiming	Unsystematic Microtiming	Tempo
All genres	.57**	(.25*)	.61**	.15	.34**	-.18
Greek	.67**	.27	.72**	-.51*	.19	.13
Indian	.48*	.03	.67**	-.15	-.21	-.34
Jazz	.09	-.29	.02	.08	.27	.08
Samba	.75**	.73**	.84**	.60*	.39	-.04
West african	(.52*)	(.45*)	.18	.23	.19	.14

Appendix 1. The 100 music examples used as stimuli in the listening experiment.

	Artist	Album	Title/track
	Greek		
1	Manos Hatzidakis	Athanasia	O Giannis o fonias
2	Manos Hatzidakis	Athanasia	I mpalanta tou Ouri
3	Sotiria Mpellou	H arhontissa	Nyxtose xoris feggari
4	Various	Velventina tragoudia	12 evzonakia
5	Various	Velventina tragoudia	Paisa sty vrysi
6	Manos Hatzidakis	To hamogelo tis tzokonta	Oi dolofonoi
7	Babis Gkoles	Hatzikyriakeio mia zoi rempetika	Zoi mou, farmakothikes
8	Aleka Mavili	To panorama tou ellinikou kinimatografou epohi defteri	Afto t'agori me ta matia ta r
9	Babis Gkoles	Hatzikyriakeio mia zoi rempetika	Atimi tyxi
10	Manos Hatzidakis	Athanasia	Melagholiko emvatirio
11	Melina Kana	Tis agapis gerakaris	Malamatenia mou zoi
12	Manos Hatzidakis	To hamogelo tis Tzokonta	I parthena tis geitonias mou
13	Giorgos Ntalaras	Afieroma sto Marko Vamvakari zontani ihografisi	Ta matoklada sou lampoun
14	Grigoris Mpithikotsis	Gia ton Grigori	Mana mou kai Panagia
15	Sotiria Mpellou	24 megales epityhies	Kardia paraponiara
16	Marinella	Ena tragoudi ein' i zoi mou	Koita me sta matia
17	Dimitris Mitropanos	40 hronia tragoudia tha leo mia zoi	Thessaloniki
18	Vicky Mosholiou	40 xronia Viky Mosholiou	Horismos
19	Vicky Mosholiou	40 xronia Viky Mosholiou	Pira ap' to xeri sou nero
20	Sotiria Mpellou	Meta to rempetiko	O kosmos einai san mpakse
	Indian		
1	Suranjan Choudhury	Various	Ae Mere Zohra Jabeen
2	Sanjoy Bandopadhyay	Various	Anandi-Kalyan
3	Suranjan Choudhury	Various	Banshi Shuney Aar Kaaj Na
4	Murali Krishna	Various	Vijaya Gopala Te Mangalar
5	Murali Krishna	Various	Itaramulu Eruganayya
6	Murali Krishna	Various	Kalaya Yashode Tava Baala
7	Shakthidhar	Various	Bhatiyali Dhun
8	Sanjoy Bandopadhyay	Various	Kafi Thumri
9	Jayateerth Mevundi	Various	Raga Kedar
10	Roshan Jamal Bhartiya	Various	Raga Joge
11	Pandit Jasraj	Various	Rag Gujari Todi
12	Yesudass	Various	Nagumomu Ganaleni
13	Amjad Ali Khan and Sons	Various	Raga Bapu Kauns
14	Dr. M. Balamuralikrishna	Various	Pallavi
15	Dr. M. Balamuralikrishna	Various	Pallavi

16	Amjad Ali Khan and Sons	Various	Bengali and Assamese Folk Songs
17	Amjad Ali Khan and Sons	Various	Tigalbandi in Rag Khamaj
18	Raga Kafi	Various	Raga Mood
19	Rajeev Taranath	Various	Bhairavi
20	John McLaughlin & Shakti	Remember Shakti	The wish

	Jazz		
1	Charles Mingus	White Box of jazz	So long Eric
2	Eddie Lockjaw Davis	Triumvirate	Lester Leaps In
3	Stéphane Grapelli	It might as well be swing	Have you met Miss Jones
4	Lionel Hampton	Black and White Box of jazz	Jeepers Creepers
5	Doncaster Jazz Orchestra	Flight of fancy	Marie's Shuffle
6	Gerry Mulligan	Black Box of jazz	Limelight
7	Buddy Rich	White Box of jazz	Moments notice
8	Paul Horn	Jazz Masterpieces	Work Song
9	Herbie Hancock	Cantaloupe Island	And what if I don't
10	Selaelo Selota	Black and White Box of jazz	Nathi
11	Lionel Hampton	White Box of jazz (CD 1)	On green dolphin street
12	Phil Woods	Jazz Masterpieces	Caravan
13	Winton Marsalis	Black and White Box of jazz	One by one
14	Jimmy Hamilton	Black and White Box of jazz	Satin Doll
15	Art Blakey and The Jazz Messengers	Live At Bubba's Jazz Restaurant	Soulful Mr Timmons
16	Teddy Wilson	Black Box of jazz	One O'Clock Jump
17	Stan Getz	Strike Up The Band	Heartplace
18	Miles Davis	Kind of Blue/Porgy and Bess/Sketches of Spain	All Blues
19	Herbie Hancock	Cantaloupe Island	Cantaloupe Island
20	Miles Davis	Kind of Blue/Porgy and Bess/Sketches of Spain	Freddie Freeloader

	Samba		
1	Teresa Cristina and Grupo Semente	A vida me fez assim	Acalanto
2	Teresa Cristina and Grupo Semente	A vida me fez assim	Viver
3	Paulinho da Viola and Elton Medeiros	Samba na madrugada	Maioria sem nenhum
4	Elton Medeiros, Nelson Sargento, and Galo Preto	So Cartola	Divina Dama
5	Elton Medeiros, Nelson Sargento, and Galo Preto	So Cartola	A mesma historia
6	Elton Medeiros, Nelson Sargento, and Galo Preto	So Cartola	Ciumente doentio
7	Paulinho da Viola and Elton Medeiros	Samba na madrugada	Depois de tanto amor
8	Paulinho da Viola and Elton Medeiros	Samba na madrugada	Sofreguidao
9	Teresa Cristina and Grupo Semente	A vida me fez assim	O passar dos anos
10	Elton Medeiros, Nelson Sargento, and Galo Preto	So Cartola	Tive sim
11	Paulinho da Viola and Elton Medeiros	Samba na madrugada	Samba original
12	Teresa Cristina and Grupo Semente	A vida me fez assim	Agua do rio
13	Paulinho da Viola and Elton Medeiros	Samba na madrugada	Minha confissao

14	Elton Medeiros, Nelson Sargento, and Galo Preto	So Cartola	Peito vazio
15	Teresa Cristina and Grupo Semente	A vida me fez assim	Ja era
16	Paulinho da Viola and Elton Medeiros	Samba na madrugada	Perfeito amor
17	Teresa Cristina and Grupo Semente	A vida me fez assim	Um calo de estimacao
18	Teresa Cristina and Grupo Semente	A vida me fez assim	Embala eu
19	Teresa Cristina and Grupo Semente	A vida me fez assim	Portela
20	Elton Medeiros, Nelson Sargento, and Galo Preto	So Cartola	Velho estacio

	West African		
1	Amadu Bamba	Drums of the Firda Fula	Track # 5
2	Tama Walo	Keepers of the Talking Drum	Track # 6
3	Omar Thiam & Jam Bugum	Sabar: The Soul of Senegal	Track # 3
4	Mamadou Ly	Mandinka Drum Master	Track # 4
5	Doudou N'Diaye Rose	Djabote	Track # 2
6	Saikouba Badjie & Modibo Traore	Babu Casamance!	Track # 9
7	Aja Addy	The Medicine Man	Track # 1
8	Traditional	Ewe Drumming From Ghana: The Soup Which is Sweet Draws the Chairs in Closer	Track # 2
9	Traditional	Children's Songs from Around the World: Volume 1: Guinea and Senegal	Track # 3
10	Obo Addy	Okropong	Track # 1
11	Mustapha Tettey Addy	The Royal Drums of Ghana	Track # 1
12	Meni Nonn Ni OB	Ka Amehewo Kule	Track # 4
13	Traditional	Dagbamba Masters: Traditional Drumming from Tamale, Ghana	Track # 3
14	Babatunde Olatunji	Drums of Passion	Track # 1
15	Aja Addy	Power and Patience	Track # 4
16	Akom	The Art of Possession	Track # 2
17	Traditional	Master Drummers of the Dagbon: Volume 2	Track # 2
18	Mustapha Tettey Addy	Secret Rhythms	Track # 3
19	Ade Olumoko and African Spirit	Musique Apala Du Peuple Yoruba	Track # 12
20	Mamady Keita	Sila Laka	Track # 7