RELATIONSHIPS BETWEEN SPECTRAL FLUX, PERCEIVED RHYTHMIC STRENGTH, AND THE PROPENSITY TO MOVE

Birgitta Burger **Riikka Ahokas** **Aaro Keipi** **Petri Toiviainen**
Department of Music, University of Jyväskylä, Jyväskylä, Finland
{birgitta.burger@, j.riikka.ahokas@student., aaro.j.keipi@student., petri.toiviainen@}jyu.fi

ABSTRACT

The tendency to move to music seems to be built into human nature. Previous studies have shown a relationship between movement and the degree of spectral flux in music, particularly in the lower sub-bands. In this study, listeners' perceptions of a range of frequency-restricted musical stimuli were investigated in order to find relationships between perceived musical aspects (rhythm, melody, and fluctuation) and the spectral flux in three different frequency bands. Additionally, the relationship between the perception of features in specific frequency bands and participants' desire to move was studied. Participants were presented with clips of frequency-restricted musical stimuli and answered four questions related to musical features. Both perceived strength of the rhythm and the propensity to move were found to correlate highly with low-frequency spectral flux. Additionally, a lower but still significant correlation was found between these perceived musical features and high-frequency spectral flux. This suggests that the spectral flux of both low and high frequency ranges can be utilized as a measure of perceived rhythm in music, and that the degree of spectral flux and the perceived rhythmic strength in high and low frequency bands are at least partly responsible for the extent to which listeners consciously desire to move when listening to music.

1. INTRODUCTION

When listening to rhythmic music we tend to move our bodies with it. Movements induced by music might be subconscious, with almost indistinguishable trappings, or deliberate, strong and intentional. The proclivity to move with music seems to be built into human nature, which is described in the literature as groove (see, e.g., [1]-[3]). These studies propose that the functional role of rhythmic music and the construct of groove are related to the evolution of entrainment and social behavior and state that synchronizing is the simplest form of entrainment from a psychological point of view. "Synchronization" is also the concept that Leman [4] suggests as the most fundamental component in bodily engagement with music. He proposed three concepts of (co-existing) corporeal articulations – "Synchronization", "Embodied Attuning", and "Empathy" – that differ in the degree of musical involvement and in the kind of action-perception couplings involved. "Synchronization" forms the fundamental component, as synchronizing to a beat is easy and spontaneous. As the first step in engaging with the music, movements could be used for imitation and prediction of beat-related features in the music. The second component, "Embodied Attuning", concerns the linkage of body movement to musical features more complex than the basic beat, such as melody, harmony, rhythm, tonality, or timbre. Following this idea, movement could be used to reflect, imitate, and navigate within the musical structure. Finally, "Empathy" is seen as the component that links musical features to expressivity and emotions.

Thus, music-induced movements seem to be associated with rhythmic features of music, such as periodic and regular patterns of beats and pulses. In basic western popular music settings the rhythm section (the drummer and bass player) are responsible for providing the rhythm. Van Dyck et al. [5] studied the effect of the dynamics of the bass drum on dancers in order to find if the bass drum is a feature that dominates music-induced movements. The authors concluded that the dynamic changes of the bass drum have an underlying effect on the intensity of movement while dancing. Burger and colleagues (see [6] and [7]) conducted a motion capture study, in which participants were asked to move to various pop music stimuli. They performed computational feature extraction on both the movement and the music data and found several relationships between movement characteristics and rhythm- and timbre-related musical features. Their results indicate that clear pulses in the music encouraged participants to move their whole body with low spatial complexity, while spectral flux in the low and high frequency ranges was more distinctly related to certain body parts. With an increasing amount of flux in the low and high frequencies, the authors discovered an increase in head and hand movement as well as an increase in temporally regular movement synchronized to different metrical levels, whereas more complex, irregular rhythmic structures resulted in temporally less regular movement. The authors concluded that spectral flux was related to the perception of rhythm of the music – flux of the low frequencies being associated with kick drum and bass guitar and high frequency flux being influenced by hi-hat and cymbal sounds – and therefore considered important for inducing movement. However, the perceptual dimension of spectral flux of restricted frequency bands has only been studied so far in connection to polyphonic timbre (see [8] and [9]) and music information retrieval related applications, such as automatic classification (see [10] and [11], both slightly differing in their technical implementation), but not strictly in relation to rhythm perception.

The spectral flux of restricted frequency bands, or sub-band flux, is a computational measure indicating the extent to which the spectrum changes over time. When computing this feature (see [8]), the stimulus is divided into 10 frequency bands, each band containing one octave in the range of 0 to 22050 Hz. After that the sub-band flux is calculated for each of these ten bands by taking the average of the Euclidean distances of the spectra for each pair of two consecutive frames of the signal (for more information about the derivation of the feature, see [8]). Two spectrograms of sub-band no. 2 (50-100 Hz) are displayed in Figure 1 to show the difference between high and low amounts of sub-band flux.

The purpose of this study was to investigate listener's perception of a range of frequency-restricted musical stimuli. The original versions of the stimuli have already been used in the movement studies cited previously (see [6] and [7]), however the present study included both the original version and three different frequency restricted versions of the original stimuli (low, mid, and high frequencies). We aimed to find relationships between specific musical aspects (such as rhythm, melody, and fluctuation) and the spectral flux in the different frequency bands. Additionally, we were interested in the relationship between the perception of musical features in specific frequency bands and participants' desire to move. We hypothesized that the perceived strength of the rhythm is positively correlated with the spectral flux, especially for the stimuli restricted to low frequencies (sub-band 2), so participants would perceive the low-frequency sub-band flux as being related to rhythm. Furthermore, we assumed positive correlations between the desire to move and the spectral flux for sub-band 2 and 9, as the spectral flux in these bands was found to be related to several characteristics of human movement (see [6] and [7]).

2. METHOD

2.1 Participants

A total of 38 participants (26 females; average age: 26.42, SD of age: 4.95) took part in the experiment. Participants were international students from the University of Jyväskylä, Finland. Participants were compensated with a movie ticket.

2.2 Stimuli

The stimuli consisted of 30-second segments from 30 different popular songs from various genres including Techno, Pop, Rock, Latin, Funk, and Jazz (the same musical stimuli as in [6] and [7] – a list of stimuli is included in these publications). They were all non-vocal and in 4/4 time, but differed in their rhythmic complexity and pulse clarity. In order to present participants with frequency-restricted stimuli, each clip was modified using MATLAB MIRToolbox 1.4 (see [12]): The clip was first divided into ten frequency bands, each band containing one octave in the range of 0 to 22050 Hz. Then the sub-bands of interest (sub-band 2: 50-100 Hz, sub-band 6: 800-1600 Hz, and sub-band 9: 6400-12800 Hz) were extracted and saved as .wav files.

2.3 Apparatus

To gather the perceptual ratings, a special patch was created in Max/MSP 5, a graphical programming environment, running on Max OS X. The setup enabled the participants to repeat excerpts as often as they wished and to move forward at their own speed. The stimuli were played back through active studio monitors (Genelec 8030A). The participants could themselves adjust the volume to a preferred level.

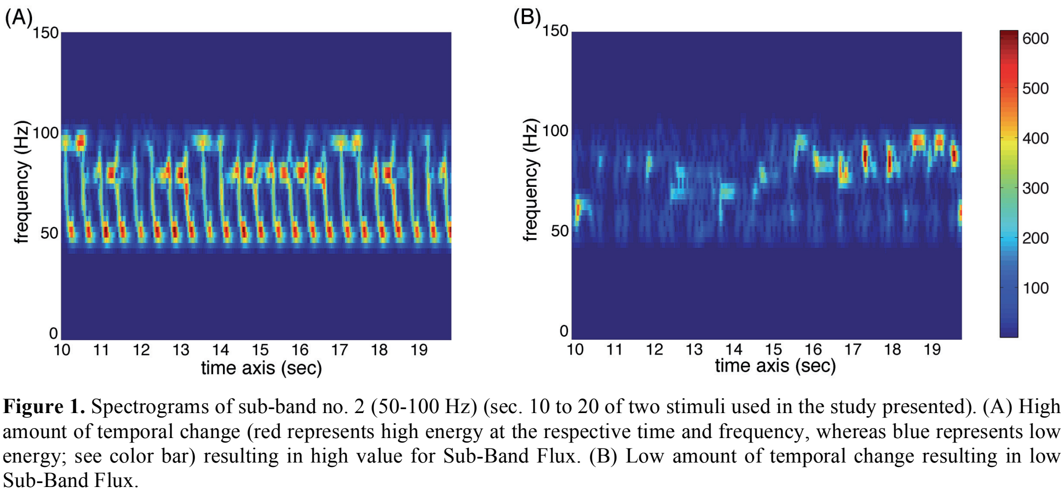

Figure 1. Spectrograms of sub-band no. 2 (50-100 Hz) (sec. 10 to 20 of two stimuli used in the study presented). (A) High amount of temporal change (red represents high energy at the respective time and frequency, whereas blue represents low energy; see color bar) resulting in high value for Sub-Band Flux. (B) Low amount of temporal change resulting in low Sub-Band Flux.

2.4 Procedure

The experiment was divided into four sections: one section containing the stimuli restricted to sub-band 2, a second section containing the stimuli restricted to sub-band 6, a third section containing the stimuli restricted to sub-band 9, and a fourth section containing the original stimuli. Each section was presented separately; first, the three sub-band restricted sections in random order, followed by the section containing the original clips. This section was always presented last to avoid a biased rating due to knowing the whole stimuli. The stimuli were also randomized within each section. Participants accomplished the experiment individually. They were asked to answer four questions, rating each on a seven-step scale (from "not at all" to "very much"):

1. *How prominent is the rhythm?*
2. *How prominent is the melody?*
3. *How much fluctuation is there in the music (How much is "going on" in the music)?*
4. *How strongly does it make you want to move?*

Preceding the experiment, there was a practice section with one example to allow participants to become familiar with the interface and the questions. The average duration of the experiment was 75 minutes.

2.5 Spectral flux extraction

For each of the four versions of each stimulus (three frequency-restricted and the original stimuli), the spectral flux was computed (using MATLAB MIRToolbox 1.4 [12]) by calculating the Euclidean distances of the spectra for each pair of consecutive frames of the signal, using a frame length of 25 ms and an overlap of 50% between successive frames. Subsequently, we averaged across the resulting time-series of flux values to receive one value for each of the four versions of the stimuli.

3. RESULTS

The first step of the analysis comprised checking the consistency of the ratings of the participants by calculating intraclass correlations (cf., [13]) for each question and stimulus type separately. The results are presented in Table 1.

	SB 2	SB 6	SB 9	Orig.
Question 1: ***rhythm?***	.95 ***	.97 ***	.96 ***	.96 ***
Question 2: ***melody?***	.94 ***	.95 ***	.93 ***	.95 ***
Question 3: ***fluctuation?***	.87 ***	.90 ***	.88 ***	.93 ***
Question 4: ***movement?***	.93 ***	.93 ***	.91 ***	.94 ***

*** $p < .001$

Table 1. Intraclass correlations for each question and stimulus type.

As these correlation coefficients indicate sufficiently high inter-participant consistency, we averaged the ratings across participants to receive one value per stimulus. Such high intraclass correlations, especially for question 1 ("*How prominent was the rhythm?*"), also suggest that the concepts of rhythm and melody were understood in a coherent way by the participants (despite findings related to cultural dependencies of rhythm perception, see [14] and [15]). Worth noting is that the correlation coefficient for question 3 ("*How much fluctuation is there in the music*") showed the lowest value for each stimulus type.

To investigate the relationship between the spectral flux data (calculated for each of the four versions of the stimuli as described in section 2.5) and the perceptual evaluations of the stimuli, we correlated the rating scores of the four questions for the music clips per stimulus type (averaged across participants) with the respective flux data. The results of the correlations are displayed in Table 2. Correlations with significance values less than p < .01 are indicated with asterisks.

	SB 2	SB 6	SB 9	Orig.
Question 1: ***rhythm?***	.79 ***	.20	.47 **	.54 **
Question 2: ***melody?***	-.01	-.28	.27	.03
Question 3: ***fluctuation?***	-.15	.01	.37	.11
Question 4: ***movement?***	.65 ***	.20	.52 **	.57 ***

** $p < .01$, *** $p < .001$

Table 2. Correlations between spectral flux and ratings on questions 1-4 for each stimulus type.

The strongest correlation ($r(30) = .79$, $p < .001$) for question 1 ("*How prominent is the rhythm?*") was found for the sub-band 2 stimuli. Question 4 ("*How strongly does it make you want to move?*") was also relatively highly correlated ($r(30) = .65$, $p < .001$) to these stimuli. Not quite as strong – though still significant – were the correlations between the same two questions and sub-band 9 flux ($r(30) = .47$, $p < .01$, for question 1, and $r(30) = .52$, $p < .01$, for question 4, respectively) and between these two questions and the flux of the original stimuli ($r(30) = .54$, $p < .01$, for question 1, and $r(30) = .57$, $p < .001$, for question 4, respectively). As all correlations were positive, these results suggest that participants rated stimuli with an increasing amount of flux in both low and high frequency ranges and overall flux with higher prominence of rhythm and with higher desire to move to the presented stimuli.

Meanwhile, the values for question 2 ("*How prominent is the melody?*") showed non-significant correlations with flux data of all stimulus types, suggesting that there is no relationship between the perceived melody prominence and the amount of (sub-band) flux in the stimuli.

Interestingly, the values for question 3: "*How much fluctuation is there in the music (How much is "going on" in the music)?*" showed no significant correlation to the (sub-band) flux data. This suggests that there was no relation between the fluctuation participants perceived in the stimuli and the computationally extracted flux in the sub-bands and the original clips.

Subsequently, we performed correlations between the ratings, segregated by each stimulus type. The results are shown in Figure 2.

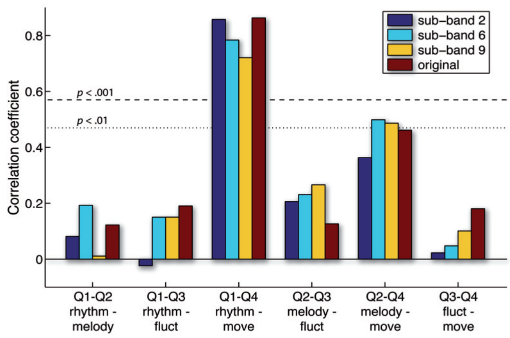

Figure 2. Correlations between the four ratings for each stimulus type separately.

The correlations between the questions show a similar pattern across the different stimulus types: for all stimulus types, question 1 ("*How prominent is the rhythm?*") correlated positively with question 4 ("*How strongly does it make you want to move?*") (sub-band 2: $r(30) = .86$, $p < .001$; sub-band 6: $r(30) = .78$, $p < .001$; sub-band 9: $r(30) = .72$, $p < .001$; original: $r(30) = .86$, $p < .001$) suggesting that the perception of a prominent rhythm was related to participants' eagerness to move to such stimuli regardless of its frequency range. For sub-band 6 and 9, additionally, question 2 ("*How prominent is the melody*") correlated moderately high with question 4 ("*How strongly does it make you want to move?*") (sub-band 6: $r(30) = .50$, $p < .01$; sub-band 9: $r(30) = .49$, $p < .01$). Thus, in the stimuli restricted to both the mid and high frequencies, the melody also appears to contribute to the willingness to move to such stimuli. The remaining correlations were non-significant.

4. DISCUSSION

We conducted an experiment to investigate participants' perceptions of rhythm, melody, fluctuation, and the desire to move in full-frequency and frequency-restricted musical stimuli. The participants' answers were consistent with our hypothesis: for stimuli restricted to the low frequency band (sub-band 2), stimuli having a higher amount of sub-band flux were perceived as being stronger related to the rhythm of the music than stimuli with a lower amount of sub-band flux, as suggested by the high correlations of flux data in sub-band 2 with the question "*How prominent is the rhythm?*". These correlations are likely due to the frequency range of specific rhythmic instruments in this sub-band, such as kick drum and low bass notes. Additionally, a greater amount of low frequency spectral flux would induce the desire of movement in participants, as was suggested by the positive correlation with the question "*How strongly does it make you want to move?*".

In addition, these two questions were also highly correlated to the spectral flux in sub-band 9. These correlations are likely due to the frequency range of specific rhythmic instruments in sub-band 9, such as hi-hat and some snare drum partials. These findings corroborate the results reported in [6] and [7], which showed higher amounts of specific bodily movements related to the amount of spectral flux in sub-bands 2 and 9. Spectral flux in sub-bands 2 and 9 may therefore be potentially more effective than spectral flux in other sub-bands in encouraging people to move to music.

The low correlations for question 3 ("*How much fluctuation is there in the music (How much is "going on" in the music)?*") with all sub-bands, however, showed that participants could not perceive the amount of fluctuation in the stimuli. There are two possible explanations for this: 1) the participants were simply unable to accurately and consistently hear the amount of fluctuation in the individual sub-bands; or 2) participants did not understand the concept of fluctuation in this context. The intraclass correlation results for this question (see Table 2) showed that the answers for this question were less consistent compared to the other questions, so the latter explanation seems likely. Interestingly, none of the participants asked about the term during the data collection, thus it could be assumed that the participants understood the term fluctuation, but that this notion differed across participants. Future studies on a similar topic should make certain that the concept of fluctuation is clearly understood by participants.

Rhythmic content has been found to be strongly related to movement (see [5]-[7]). The high correlations between the questions "*How prominent is the rhythm?*" and "*How strongly does it make you want to move?*" for all the four stimulus types suggests that the perception of a prominent rhythm is related to participants' eagerness to move to stimuli regardless of its frequency range. Future studies could analyze the actual differences of participants' movement for specific frequency ranges, for instance in a motion capture setting.

The relationship found between perceived rhythm and desire to move (see previous paragraph) could also serve as support for Leman's theory of corporeal articulations [4]. Stimuli that participants rated to contain a strong rhythm, were also rated high on desire to move, suggesting a connection between both. That could be seen as being in line with the concepts of "Synchronization" and "Embodied Attuning", in which beat/musical features, such as the rhythm, are proposed to induce body movement.

There was a weaker – but still apparent – correlation between the two questions "*How prominent is the melody*" and "*How much does it make you want to move?*", which points to a relationship between melodic strength and music-induced desire to move. The effect of melodic content on movement could also be the subject of future studies.

It could be argued that some of the participants' answers (especially to question 4: "*How much does it make* *you want to move?*") may have been skewed by prior experience with the particular stimuli since the stimuli used were clips from western popular music. However, it could be assumed that the wide range of backgrounds of participants, as well as the modification of the sub-band clips to within a certain frequency range (which often made the source music difficult to distinguish), helped to minimize the possible effects of familiarity on the ratings. Nevertheless, it might be valuable for future data acquisitions to include collecting both familiarity and preference ratings of the stimuli. This would give more insight into relationships between music characteristics and movement propensity, as it could be assumed that participants still rate certain stimuli high on "desire to move" although they do not like them.

We excluded extreme genres in our stimuli selection – such as death metal for example – as such music might be more prone to familiarity and preference than other popular music. Although musical styles such as death metal could contain high spectral flux and be considered rhythmic, not everybody would probably feel the urge to dance to such music. Thus, if such extreme genres were considered in the stimuli selection, relationships between spectral flux and movement propensity might be less linear than presented in this paper.

The presented analysis utilized one value – the average of the flux time-series – as measure for the spectral flux. As such, this could be regarded as an over-generalization, since taking the mean disregards information about the temporal regularity that the flux series should exhibit in order to induce movement. In general, a random use of, for instance, the kick drum would also result in a high amount of low-frequency flux, but it would fail to evoke a sensation of rhythm or movement in the listener. However, our stimuli were throughout the whole stimulus duration all metrically regular, had a sensation of pulse, and were steady in most of the musical characteristics. Thus, we believe that temporal averaging of the flux time-series was a suitable way to receive a relevant measure of spectral flux for each stimulus.

In conclusion, the results of this study show that for stimuli being restricted to low frequencies and, to a lesser extent, for stimuli being restricted to high frequencies, a high amount of spectral flux was perceived as having a more prominent rhythm. This suggests that the sub-band flux of both low and high frequency ranges can be utilized as a possible measure of perceived rhythm in music. Furthermore, the significant correlation between the answers to the questions "*How prominent is the rhythm*" and "*How much does it make you want to move*" to the spectral flux in sub-bands 2 and 9 point to an important role of the spectral content in these sub-bands; in essence, it suggests that the degree of flux and the perceived rhythmic strength in sub-bands 2 and 9 are at least partly responsible for the extent to which listeners consciously desire to move when listening to music. This is consistent with previous research (see [6] and [7]) that identified spectral flux in these particular sub-bands as correlating with various characteristics of bodily movements.

Acknowledgments

This study was supported by the Academy of Finland (project 118616).

5. REFERENCES

[1] G. Madison, "Experiencing groove induced by music: Consistency and phenomenology," in Music Perception 24(2), 2006, pp. 201-208.

[2] G. Madison, F. Gouyon, F. Ullén, and K. Hörnström, "Modeling the tendency for music to induce movement in humans: first correlations with low-level audio descriptors across music genres," in Journal of Experimental Psychology: Human Perception and Performance 37(5), 2011, pp. 1578-1594.

[3] P. Janata, S. T. Tomic, and J. M. Haberman, "Sensorimotor coupling in music and the psychology of the groove", in Journal of Experimental Psychology: General 141(1), 2011, pp. 54-75.

[4] M. Leman, Embodied Music Cognition and Mediation Technology. Cambridge, MA/London, UK: MIT Press, 2007.

[5] E. van Dyck, D. Moelants, M. Demey, P. Coussement, A. Deweppe, and M. Leman, "The impact of the bass drum on body movement in spontaneous dance," in Proceedings of the 11th International Conference in Music Perception and Cognition, Seattle, WA, 2010, pp. 429-434.

[6] B. Burger, M. R. Thompson, G. Luck, S. Saarikallio, and P. Toiviainen, "Music moves us: Beat-related musical features influence regularity of music-induced movement," in Proceedings of the 12th International Conference in Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences for Music, Thessaloniki, Greece, 2012, pp. 183-187.

[7] B. Burger, M. R. Thompson, S. Saarikallio, G. Luck, and P. Toiviainen, "Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement," in Frontiers in Psychology, 4:183, 2013.

[8] V. Alluri, and P. Toiviainen, "Exploring perceptual and acoustical correlates of polyphonic timbre," in Music Perception 27(3), 2010, pp. 223-242.

[9] V. Alluri, and P. Toiviainen, "Effects of enculturation of the semantic and acoustic correlates of polyphonic timbre," in Music Perception 29(3), 2012, pp. 297-310.

[10] D.-N. Jiang, L. Lu, H.-J. Zhang, J.-H. Tao, and L.-H. Cai, "Music type classification by spectral contrast feature," in *Proceedings of the IEEE International*

Conference on Multimedia and Expo, 2002, pp. 113-116.

[11] R. Cai, L. Lu, A. Hanjalic, H.-J. Zhang, L.-H. Cai, "A flexible framework for key audio effects detection and auditory context inference," in IEEE Transactions on Audio, Speech and Language Processing 14(3), 2006, pp. 1026-1039.

[12] O. Lartillot, and P. Toiviainen, "A Matlab toolbox for musical feature extraction from audio," in Proceedings of the 10th International Conference on Digital Audio Effects, Bordeaux, France, 2007, 1-8.

[13] P. E. Shrout, and J.L. Fleiss, "Intraclass correlations: uses in assessing rater reliability," in Psychological Bulletin 86(2), 1979, pp. 420-428.

[14] A. D. Patel, J. R. Iversen, and K. Ohgushi, "Cultural differences in rhythm perception: what is the influence of native language?" In Proceedings of the 8th International Conference on Music Perception and Cognition, Evanston, IL, 2004, 88-89.

[15] G. Soley and E. E. Hannon, "Infants Prefer the Musical Meter of Their Own Culture: A Cross-Cultural Comparison," in Developmental Psychology 46(1), 2004, 286-292.