PAD AND SAD: TWO AWARENESS-WEIGHTED RHYTHMIC SIMILARITY DISTANCES

Daniel Gómez-Marín
Universitat Pompeu Fabra
daniel.gomez@upf.edu

Sergi Jordà
Universitat Pompeu Fabra
sergi.jorda@upf.edu

Perfecto Herrera
Universitat Pompeu Fabra
perfecto.herrera@upf.edu

ABSTRACT

Measuring rhythm similarity is relevant for the analysis and generation of music. Existing similarity metrics tend to consider our perception of rhythms as homogeneous in time without discriminating the importance of some regions from others. On a previously reported experiment we observed that measures of similarity may differ given the presence or absence of a pulse inducing sound and the importance of those measures is not constant along the pattern. These results are now reinterpreted by refining the previously proposed metrics. We consider that the perceptual contribution of each beat to the measured similarity is non-homogeneous but might indeed depend on the temporal positions of the beat along the bar. We show that with these improvements, the correlation between the previously evaluated experimental similarity and predictions based on our metrics increases substantially. We conclude by discussing a possible new methodology for evaluating rhythmic similarity between audio loops.

1. INTRODUCTION

Rhythm similarity is an important problem for both music cognition and music retrieval. Determining which aspects of the musical flow are used by musical brains to decide if two musical excerpts share similarities with respect to rhythm, would make it possible to build algorithms that approximate human ratings about such relatedness. The applications of such algorithms in MIR contexts should be obvious and some have already been addressed [34] [14] [6] [21]. Unfortunately, there is a gap between the knowledge provided by cognitive models and engineering models with respect to similarity in general, and rhythm similarity in particular. Rhythm similarity metrics used in MIR are frequently based on superficial information such as inter-onset intervals, overall tempo or beat rate, onset density, and consider full-length songs to derive a single similarity value. Contrastingly, rhythm similarity models developed by cognitive scientists insist on the importance of syncopation, beat salience, periodicities and shorter timescales to determine similarity. In this paper we address the above-mentioned gap and propose two rhythm similarity distances that refine those currently available (and probably rougher than desirable). The proposed distances have been derived from music cognition knowledge and have been tuned using experiments involving human listeners. We additionally show that they can be adapted to work (at least) in a music-loop collection organization context, where music creators want to organize their building blocks in rhythm contrasting or rhythm flowing ways where similarity would provide the criterion for such concatenation of elements.

Previous work has used rhythmic descriptors, computed from audio signals, to analyze song databases. A common collection used for testing genre classification methodologies, The Ballroom dataset, has been sorted automatically using different rhythmic descriptors and methodologies [4] [30] [10] [25]. Out of the ballroom dataset very few authors have addressed rhythm in electronic music with rhythmic descriptors [11] [24] [2]. The logic behind most of these research is the assumption that if a corpus is classified according to annotated labels, the features used for that clustering are somehow related to the phenomena that generates the clustering. In other words, a correct classification implies that the features used are perceptually relevant.

Using symbolic representations of music, other authors propose metrics to evaluate rhythmic similarity that have shown to be useful in melody classification [34] or have proven correlation with cognitive judgements in rhythmic similarity experiments [13] [26] [1].

However, neither the audio-based methodologies or the symbolic metrics for rhythm similarity ([24] being an exception) has been designed for exploring short audio segments such as loops. Moreover, methodologies to evaluate rhythmic similarity between two audio loops and retrieve a value that can be analogous to a human rating are not yet available. Therefore we want to develop perceptually grounded rhythm similarity metrics to be used with short audio loops.

Our effort throughout this paper is to present two new rhythmic similarity metrics derived from revisiting the results of our cognitive experiments on rhythm similarity perception [9]. After revisiting our previous experiments, two metrics arise as useful in similarity prediction tasks. Based on those metrics we then introduce a new methodology to explore rhythmic similarity between audio loops.

The metrics proposed are based on the requirement that rhythmic similarity must be rooted in current knowledge of rhythm perception, where the notions of beat entrainment, reinforcement and syncopation are fundamental. We hypothesize that a proper rhythmic similarity measure can be built upon those perceptual considerations, emphasizing the idea that our attention when judging the similarity between two rhythms is not evenly distributed in time. We specifically propose that we are more aware of certain regions of a rhythm than others, affecting the way in which we measure their similarity [7]. To test our hypothesis we use the results of perceptual experiments published previously [9], where subjects are asked to evaluate the similarity between two monotimbral and monotonic rhythmic patterns, inducing a beat before the rhythms are presented. Similarity ratings obtained from the experiment are compared with predictions computed with our metrics for the same rhythmic patterns in order to analyze their correlation.

High correlation values between the similarity ratings of our previous experiment and the metrics presented here are found, suggesting that blending awareness and syncopation is important for accurately predicting rhythmic similarity. Although results are promising, there is still a need to expand the proposed metrics to include dynamics, duration and timbre, which also have a high impact on our perceptual judgments. Finally we want to explore if the measures we propose, besides providing good fits and predictions of human judgements, can be used to organize loop collections. The use of our metrics in audio analysis will be discussed in the last sections of the paper, where we propose a methodology and evaluate it using audio loops of drum break patterns. Our results for this pilot validation present significant correlations between the similarity judgements of the subjects and the predictions resulting from our methodology.

2. STATE OF THE ART

2.1 Beat Induction

The fact that us humans induce a pulse sensation when listening to music is by no means trivial and it seems to be an innate and involuntary process [35]. It is known that the mechanisms that favour our acquisition of a beat when listening to music can also be triggered by any sequence of onsets [27]. This emergent beat entrainment is a cognitive process that can be divided two stages: first, we try to infer a metrical structure either by computing distances from intervals of the musical surface, where at least 5 to 10 notes are needed [3], or just try to match the incoming sound to an internal repertoire of known rhythms. Finally, once a meter has been hypothesized, it is maintained in the form of expectancies that interact with the new incoming sounds [18]. During this interaction, the expected pulse can be reinforced or disconfirmed. When challenged, brain rejection signals have been measured by means of EEG [16]. The occurrence of a disconfirmation is often referred to as syncopation, indicating notes that were expected on the beat but were presented on a previous metrical position [19].

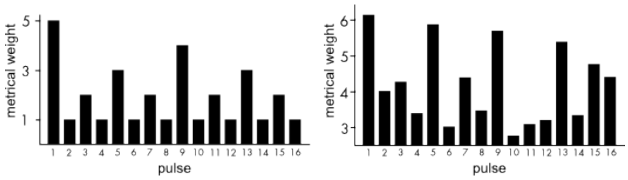

Figure 1. Lerdhal and Jackendorf′s [17] metrical weight profile (left) and the experimentally revised version of Palmer and Krumhansl [23] measured for musicians(right).

In order to represent the variability of expectancies along a rhythmic pattern, researchers use profiles that indicate the metrical weight of a note depending on its position. Different profiles that highlight the importance of a beat reinforcement or a syncopated event, depending on its occurrence within a full metrical period, have been proposed. A main theoretical profile [17] and an updated version experimentally revised with musicians [23] are presented in Figure 1. These profiles stress the existence of a perceptual hierarchy of sound events depending on their occurrence, suggesting that some reinforcements or syncopations are perceptually more relevant than others. These ideas have led to algorithms that measure the syncopation of a monotimbral unaccented phrase [31]. Moreover, these algorithms have been used to correlate syncopation with the difficulty to tap along rhythms [5], musical complexity [32] [8] [28] and musical pleasure and desire to dance [36] stressing the idea that syncopation has a powerful effect on our perception of music.

2.2 Rhythmic Similarity

Once we can extract a numerical value from a pattern of onsets such as its syncopation value, comparing patterns and establishing distances between them is mathematically possible. One main approach, proposed by Johnson-Laird [15], is to analyze the onsets present on every beat of a rhythmic pattern and subscribe the beat to a category depending if it reinforces the beat, challenges the beat or does nothing to the beat . This approach has been modified [29] and successfully tested in experimental conditions [1]. These ideas will be further expanded throughout this paper.

As most proposed similarity metrics are measured on monotimbral, monotonal and unaccented symbolic representations of rhythm, there are others who have explored the use of string similarity techniques as the swap distance or the edit distance [20] [22] to measure similarity between patterns. The edit distance has proven to be a useful predictor of human similarity judgements [33] [12] [26].

Here we use similarity metrics based on syncopation, specifically a variation of the theory of Johnson-Laird in which we expand the possible groups that a beat can be subscribed to (syncopation, reinforcement or nothing). In the following sections we present, test and discuss an improvement over a previously published metric and explore the possibility of using these symbolic metrics in rhythmic analysis of audio signals.

3. METHOD

In this section we present different concepts that are the building blocks of our rhythmic similarity algorithms. We have to make some simplifying assumptions, considering one bar, monotimbral, monotonal, percussive patterns with 4/4 time signature and a minimum resolution of a sixteenth note. The symbolic representation of such patterns is binary, where a 1 indicates an onset and 0 indicates a silence. Therefore the patterns used throughout this work are 16 digit sequences of zeroes and ones.

3.1 Beats to syncopation groups

Rhythms are split in beats, in our case each beat has four steps (four digits). Each beat of a rhythm is classified into a group according to its relation with the pulse, either a reinforcement or a challenge (See table 1). This method is a variation of Johnson-Laird′s method [15], in which beats are clustered in three broad categories: syncopation, reinforcement or nothing depending if the elements of the beat are a reinforcement, a challenge or have no interaction with the pulse. We have expanded Johnson-Laird′s method by splitting syncopation into three possible groups (groups 5 to 7, Table 1), reinforcement is split in three groups (groups 1 to 3, Table 1) and adding a new category where a syncopation and a reinforcement are both present (group 8, Table 1).

The procedure to classify each beat is to compute its syncopation value using the beat profile 2 0 1 0. This profile is derived from Lerdahl and Jackendorf′s [17] in which weights are proportional to the duration of the note each accent represents: an accent of a whole note has a higher weight than an accent on a half note, which is higher than an accent on a quarter note, and so forth. In our beat profile the first onset, that is coincident with the pulse, has a higher weight than the third onset which is coincident with an eighth note.

It is important to note that an onset on the fourth step of a beat generates a syncopation only if the first step of the next beat is a silence. Therefore to calculate the appropriate syncopation values for every beat, the first step of the following beat has to be considered. The syncopation value for each beat is the sum of each onset′s metrical weights.

Each beat can then be assigned to one out of eight syncopation categories, but we have considered the case of a reinforcement on the first step and a syncopation on the fourth step 1001_ (total syncopation value = 0) as special cases belonging to syncopation group #8.

3.2 Coincidence

We propose here two metrics, one that explores if two patterns have the same onsets and silences on a specific beat, which we call pattern coincidence distance (PD) and the other one, named syncopation coincidence distance (SD) which explores if a specific beat of two patterns belong to the same syncopation group (see previous Table 1).

Group	value	Patterns
1	3	1010_ 1010x
2	2	1000_1000x 1001x 1011x
3	1	0010_ 0010x 0110_ 0110x 1110_ 1110x
4	0	0000_ 0000x 1111x 0011x 0001x 0111x
5	-1	0100_ 0100x 1100_ 1100x 0101x 1101x
6	-2	0001_ 0011_ 0111_ 1111_
7	-3	0101_ 1101_
8	0	1001_ 1011_

Table 1. Relation between syncopation group, syncopation value and beat patterns. The symbol '_' indicates a silence at the beginning of the next beat and the symbol 'x' indicates an onset at the beginning of the nest beat.

Here we give an illustrative example to understand PD and SD. The two first beats of a given pattern A have the following onset/silence configuration 1001 0110 and another pattern B has 1100 0010. Their respective syncopation groups are #8 #3 and #5 #3. The pattern coincidence (PD) is computed by looking at the percentage of coincident onsets and silences on the same beat of each pattern. Their coincidence values would be (2+3)/8 = 0.625 because for the first beat there are 2 out of four notes coincident between 1001_ and 1100; and for the second beat, there are 3 coincidences between 0110 and 0010. In total there are 2+3 coincidences out of 8 possible. On the other hand, to measure the syncopation coincidence (SD), for the first beat of patterns A and B, we get that 1001_ belongs to family #8 and 1100 belongs to family #5. Clearly 8 is different from 5. But if we look at the second beat, 0110 and 0010 belong to the same group #3, thus group coincidence is 0+1=1. With these metrics we obtain two methods for measuring a numerical value of the coincidence between two coincident beats of different patterns. If the coincidence between all the beats of two patterns is computed, this value can be used as a measure of similarity between the two patterns. However, we might consider that, as different onsets have different metrical weights depending on their position within a pattern, beats can also have different perceptual relevance depending on their position within the pattern. In this paper we have conceptualized this factor as awareness.

3.3 Awareness as an effect of metrical hierarchy

Our previously published results [8] suggest a difference in the relevance of each beat when measuring similarity between two patterns based on coincidence. This awareness has proven important when exploring correlations between our experimental results of similarity and the rhythmic patterns compared. Thus we propose each beat to have different relevance when evaluating similarity between two patterns in the presence of a pre defined metrical context. Awareness is conceived as weight factors applied to each beat′s coincidence metric (either PD or SD). These weights emphasize or moderate each beat′s importance on the final distance value. This concept will be addressed in the following section and is decisive for explaining our results.

3.4 Rhythmic Similarity Metrics

Our metrics are straightforward and are based on computing any of the two types of coincidence (either beat or syncopation group), and using them directly or with an awareness-based weighting. We finally have four metrics, two non-weighted ones. Pattern coincidence Distance (PD) and Syncopation group coincidence Distance (SD), Pattern coincidence and Awareness Distance (PAD) and Syncopation group coincidence and Awareness Distance (SAD). The weights of the PAD and SAD metrics will be explored on the following sections.

$$PD = pc1 + pc2 + pc3 + pc4 \tag{1}$$

$$SD = sc1 + sc2 + sc3 + sc4 \tag{2}$$

$$PAD = pc1w1 + pc2w2 + pc3w3 + pc4w4 \tag{3}$$

$$SAD = sc1w1 + sc2w2 + sc3w3 + sc4w4 \tag{4}$$

Where pc(n) is pattern coincidence, sc(n) is syncopation group coincidence, w(n) is the weighting of each beat, n is the order of the beat within a full metric cycle.

4. EXPERIMENT

In previously published paper [9] we performed two rhythmic similarity experiments, one inducing the beat and another without inducing the beat. In this paper we are revisiting the beat-induced experiment to test our new metrics with the similarity ratings obtained in the previous one.

In one of the experiments, twenty one subjects (recruited among the MTG staff and UPF pool of students, all of them with musical experience of more than 5 years as amateur performers) rated different rhythm pairs in the presence of a beat-inducing kick drum. The rhythm pairs were constructed by making variations of a main pattern as shown in Table 2. A region of the base pattern was shifted progressively generating new patterns. Nine different main patterns were designed and the length and origin of the region was varied systematically. Thirty six rhythm pairs plus a control pair were tested by all the subjects who rated similarity using a Likert scale. To promote rhythm entrainment, a kick drum, coincident with the start of every beat, was presented before and simultaneously with the tested rhythms.

5. RESULTS

The mode of the similarity ratings for each pair of patterns was used as the value capturing their similarity. All the pairs of patterns presented to the subjects are analyzed with the metrics described in section 3, exploring the correlations with the similarity ratings reported for each pair.

In our previously reported experiment, we computed the PD distance for every tested pair and observed a Spearman Rank correlation with the subjects similarity ratings

Base Pattern	variation
1010 1110 1000 1010	1101 0110 1000 1010
1010 1110 1000 1010	1010 1011 1000 1010
1010 1110 1000 1010	1001 0101 1000 1010
1010 1110 1000 1010	1010 1010 1100 1010

Table 2. Example of four stimuli pairs used in the experiment. The left column has the base pattern and the derived variations are on the right column. The similarity measures of the subjects are between the base pattern and each variation. The underlined portion of the base pattern is repeated in the variations.

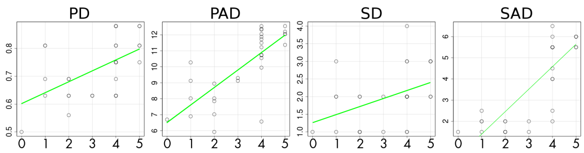

Figure 2. PD, PAD, SD and SAD predictions correlated with similarity ratings. X axis: similarity ratings, y axis PD, PAD, SD and SAD predictions from left to right.

of 0.54 (p-value < 0.005). We also computed the SD distance which has a Spearman rank correlation value of 0.46 with the similarity ratings (p-value < 0.01).

Here we calculate our newly introduced metrics PAD and SAD. To calculate PAD, a linear regression between the coincidence result of each beat and the similarity ratings is computed. The normalized weights obtained for beats 1 to 4 are 1, 0.27, 0.22 and 0.16 respectively. We take the weights of the linear regression as indications of the awareness for each beat. Using those weights we get the PAD distance with a Spearman Rank correlation value of 0.76 (p-value < 0.001).To calculate SAD a linear regression between each beats coincidence and similarity ratings generated the following normalized weights for beat 1 to 4: 1, 0.075, 0.14 and 0.12 respectively. Again, we take the weights of the linear regression and use them as indications of the awareness for each beat. Applying those weights we get the SAD distance which has a Spearman Rank correlation value of 0.81 (p-value < 0.001) which is above significance.

The resulting awareness profiles of both PAD and SAD metrics have a similar behaviour (see Figure 3). In both cases the importance of the first beat is almost 5 times larger than the other beats. Our experimental hypothesis is that this phenomena evidences a hierarchical organization of rhythmic elements in time where the first element of a rhythmic sequence is of greater importance than the rest.

These correlation values suggest that the PAD and SAD metrics are reasonable candidates to predict rhythmic similarity between two patterns of onsets in the presence of a beat, the way in which most of the music is experienced. The PAD and SAD metrics surpass the results found and reported in our previous experiment, which makes them suitable to be used in real life scenarios.

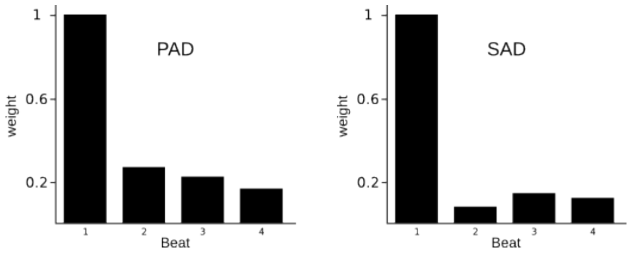

Figure 3. Awareness profiles of the PAD and SAD distances that generated best correlations with rhythm similarity ratings.

6. DISCUSSION

It can be seen that the SAD metric has the highest correlation values with human similarity, rating slightly above the PAD metric, while the non-weighted metrics PD and SD are barely significant. This suggests that the concepts of syncopation groups and beat awareness are perceptually relevant.

The drop in correlation values when there is no awareness weighting validates the idea of each beat having a different importance when beat induced subjects try to make sense of them. It appears that the first beat is the most important followed by the third, the fourth and the second.

The SAD metric is based on comparing if syncopation groups are coincident between different patterns (see section 3.2). This means that a change from one family to any other family is penalized by our algorithm despite if the change is between syncopation to syncopation (groups 5 to 7 in Table 1) or reinforcement to reinforcement (groups 1 to 3 in Table 1) or if it is a change from a syncopation to a reinforcement group or to the nothing group (or vice versa). Since the SAD metric has a positive correlation with similarity ratings, this suggests that any change between groups decreases our perception of similarity. On the other hand, perception of rhythmic similarity is highly influenced with the coincidence between syncopation groups or patterns and the position of those coincidences within the pattern.

7. PILOT VALIDATION

Direct MIR applications from PAD and SAD metrics can be the exploration of loop databases, based on rhythm similarity. The simplest approach would be to use an onset detector to the loop signal and extract a general onset pattern. This would lead to a single-level pattern deprived of any instrumental information where all musical interplay, the main information, would be lost. On the other hand, a robust source separation system would be ideal, where an audio loop could be completely split into its different instrumental components and then converted to a symbolic representation. But the technologies to perform such a task are not yet reliable. An alternative would be to extract onset patterns from meaningful frequency bands that could preserve spectral information present on the audio loop.

We propose a methodology where a sound loop, of known metric length, is segmented every sixteenth note value and filtered in 23 Bark bands. This is a typical spectral representation which approximates frequency resolution of the human hearing. The energy peaks in each band are considered as onsets and the rest as silences. In this way we convert an audio loop into a binary matrix of onset and silences of 23 bands times the number of analysis windows. An audio loop is then decomposed in 23 parallel rhythmic patterns that can be compared with the 23 patterns of another audio loop measuring PAD and SAD distances between bands. The sum of the band to band distances is the overall PAD or SAD distance between two audio loops. Note that this methodology is tempo independent if the loops compared have the same known metrical length.

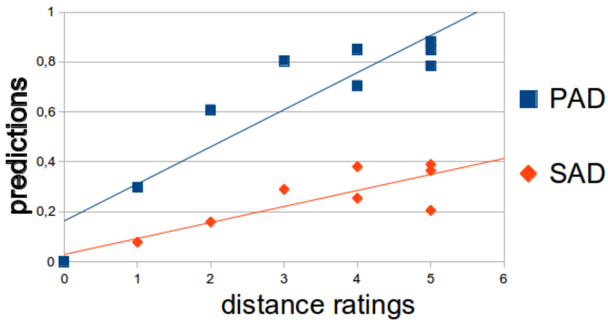

Figure 4. Predicted similarity vs similarity ratings of ten audio loops using our methodology with PAD and SAD metrics.

As a pilot validation for our methodology, an experiment was carried out using nine different drum break loops in audio format (downloaded from http://rhythm-lab.com). All loops were post processed to have a metrical length of two bars. Fifteen musically trained subjects were invited to rate the rhythmic similarity between one audio loop and the rest using a Likert scale divided in 5 steps, from ”very similar” to ”very different”. The mode of the results for each pair was used as the representative similarity value and the correlations with PAD and SAD distances were measured. The awareness profile used for both PAD and SAD was 1 0.075 0.14 0.12 extracted from the results presented in section 5 (see Figure 3, right).

The obtained results present (p-value < 0.001) significant correlation between the similarity reported by the fifteen subjects and the PAD and SAD distances (Figure 4). The PAD distance has a 0.80 Spearman rank correlation value (p-value < 0.01). The SAD distance has a Spearmann correlation value of 0.75 (p-value < 0.05).

It is quite interesting that PAD and SAD distances provide reliable similarity predictions, given the subjectivity of the task and the fact that the breaks come form very different recordings with an obvious difference in timbre and dynamics. For this pilot validation The PAD has a higher correlation value with the similarity ratings than the SAD metric.

8. CONCLUSION AND FUTURE WORK

Based on these results, we propose that measuring the PAD and SAD distance between two rhythms with an induced beat is an effective way to predict human rhythmic similarity ratings. Perceptually motivated rhythm similarity measures that are applied to MIR problems should take into account both the syncopation groups and a beat-awareness measure, in order to match subjective appreciations of rhythm similarity.

The rhythms used in the foundational experiments of our metrics are only limited to a 4/4 time signature, a 16 step length, sixteenth note resolution and binary dynamics. Expanding the signature to other common signatures, smaller note resolutions and subtler dynamics is important in order to broaden the validity and usefulness of our metrics and methodology.

Even though our methodology for measuring similarity among loops has slightly significant correlation values, both with PAD and SAD, it is important to consider the scale of the pilot validation is limited. New experiments with a higher amount of loops should be carried out in order to explore the real advantages and limitations of our methodology.

9. ACKNOWLEDGEMENTS

We would like to thank Julián Burbano for his help in the analysis of the data. This research has been partially supported by the EU funded GiantSteps project (FP7-ICT-2013-10 Grant agreement nr 610591).

10. REFERENCES

[1] Erica Cao, Max Lotstein, and Philip N. Johnson-Laird. similarity and families of musical rhythms. *Music Perception: An Interdisciplinary Journal*, 31(5):444–469, 2014.

[2] Nick Collins. Influence in early electronic dance music: an audio content analysis investigation. In *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*, number Ismir, pages 1–6, 2012.

[3] Peter Desain and Henkjan Honing. Computational models of beat induction: The rule-based approach. *Journal of New Music Research*, 28(1):29–42, 1999.

[4] Simon Dixon, Fabien Gouyon, and Gerhard Widmer. Towards characterisation of music via rhythmic patterns. In *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR 2004)*, volume 5, pages 509–516, 2004.

[5] W. Tecumseh Fitch and Andrew J. Rosenfeld. Perception and Production of Syncopated Rhythms, 2007.

[6] Jamie Forth. *Cognitively-motivated geometric methods of pattern discovery and models of similarity in music*. PhD thesis, Goldsmiths, University of London, 2012.

[7] Murray Glanzer and Anita R. Cunitz. Storage mechanisms in free recall. *Journal of verbal learning and verbal behavior*, 6:351–360, 1966.

[8] Francisco Gómez, Eric Thul, and G Toussaint. An experimental comparison of formal measures of rhythmic syncopation. *Proceedings of the International Computer Music Conference*, pages 101–104, 2007.

[9] Daniel Gómez-Marín, Sergi Jordà, and Perfecto Herera. Strictly Rhythm: Exploring the effects of identical regions and meter induction in rhythmic similarity perception. In *11th International Symposium on Computer Music Multidisciplinary Research (CMMR)*, Plymouth, 2015.

[10] Fabien Gouyon, Simon Dixon, Elias Pampalk, and Gerhard Widmer. Evaluating rhythmic descriptors for musical genre classification. In *AES 25th International Conference*, pages 1–9, LONDON, 2004.

[11] Matthias Gruhne, Christian Dittmar, and Daniel Gaertner. Improving Rhythmic Similarity Computation by Beat Histogram Transformations. *10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, (Ismir):177–182, 2009.

[12] Catherine Guastavino, Francisco Gómez, Godfried Toussaint, Fabrice Marandola, and Emilia Gómez. Measuring Similarity between Flamenco Rhythmic Patterns, 2009.

[13] Ludger Hofmann-Engl. Rhythmic similarity: A theoretical and empirical approach. *7th International Conference on Music Perception and Cognition*, (c):564–567, 2002.

[14] Henkjan Honing. Lure(d) into listening: The potential of cognition-based music information retrieval. *Emperical Musicology Review*, 5(4):121–126, 2010.

[15] Philip N. Johnson-Laird. Rhythm and meter: A theory at the computational level., 1991.

[16] Olivia Ladinig, Henkjan Honing, Gábor Háden, and István Winkler. Probing Attentive and Preattentive Emergent Meter in Adult Listeners without Extensive Music Training, 2009.

[17] Fred Lerdahl and Ray Jackendoff. *A generative theory of tonal music*. MIT Press, 1985.

[18] Justin London. *Hearing in Time*. Oxford University Press, 2012.

[19] H. Christopher Longuet-Higgins and Christopher S. Lee. The Rhythmic Interpretation of Monophonic Music. *Music Perception: An Interdisciplinary Journal*, 1(4):424–441, 1984.

[20] Marcel Mongeau and David Sankoff. Comparison of musical sequences. *Computers and the Humanities*, 24(3):161–175, 1990.

[21] Alberto Novello, Martin M.F. McKinney, and Armin Kohlrausch. Perceptual Evaluation of Inter-song Similarity in Western Popular Music, 2011.

[22] Keith S. Orpen and David Huron. Measurement of similarity in music: A quantitative approach for non-parametric representations. *Computers in music research*, 4:1–44, 1992.

[23] Caroline Palmer and Carol L. Krumhansl. Mental representations for musical meter. *Journal of experimental psychology. Human perception and performance*, 16(4):728–741, 1990.

[24] Maria Panteli, Bruno Rocha, Niels Bogaards, and Aline Honingh. Development of a Rhythm Similarity Model for Electronic Dance Music. Number Ismir, pages 537–542, 2014.

[25] Geoffroy Peeters. Spectral and Temporal Periodicity Representations of Rhythm for the Automatic Classification of Music Audio Signal. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(5):1242–1252, 2011.

[26] Olaf Post and Godfried Toussaint. The Edit Distance as a Measure of Perceived Rhythmic Similarity. *Empirical Musicology Review*, 6(3):164–179, 2011.

[27] Dirk-jan Povel and Peter Essens. Perception of Temporal Patterns. *Music Perception1*, 2(4):411–440, 1985.

[28] Jeffrey Pressing. Cognitive complexity and the structure of musical patterns. *Noetica*, 3(8):1–8, 1998.

[29] Jasba Simpson and David Huron. The Perception of Rhythmic Similarity: A Test of a Modified Version of Johnson-Lairds Theory. *Canadian Acoustics*, 21(3):89–94, 1993.

[30] L. M. Smith. Rhythmic similarity using metrical profile matching. In *International Computer Music Conference*, 2010.

[31] Leigh M Smith and Henkjan Honing. Evaluating and extending computational models of rhythmic syncopation in music. In *Proceedings of the International Computer Music Conference*, pages 688–691, 2006.

[32] Eric Thul, Godfried T. Toussaint, and Godfried T.Toussaint. Rhythm complexity measures: a comparison of mathematical models of human perception and performance. In *Proceedings of the International Society for Music Information Retrieval (ISMIR) Conference, Section 5c - Rhythm and Meter*, number 8, pages 663–668, 2008.

[33] Godfried T. Toussaint, Malcolm Campbell, and Naor Brown. Computational Models of Symbolic Rhythm Similarity: Correlation with Human Judgments. *Analytical Approaches to World Music*, 1(2), 2011.

[34] Anja Volk, Jörg Garbers, Peter Van Kranenburg, Frans Wiering, Remco C Veltkamp, and Louis P Grijp. Applying rhythmic similarity based on inner metric analysis to folksong research. In *Eighth International Conference on Music Information Retrieval. Austrian Computer Society*, 2007.

[35] István Winkler, Gábor P Háden, Olivia Ladinig, István Sziller, and Henkjan Honing. Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America*, 106(7):2468–2471, 2009.

[36] Maria A G Witek, Eric F. Clarke, Morten L. Kringelbach, and Peter Vuust. Effects of Polyphonic Context, Instrumentation, and Metrical Location on Syncopation in Music. *Music Perception: An Interdisciplinary Journal*, 32(2):201 – 217, 2014.