Strictly Rhythm: Exploring the effects of identical regions and meter induction in rhythmic similarity perception

Daniel Gómez-Marín[1], Sergi Jordà[1], and Perfecto Herrera[1]

Music Technology Group, UPF, Barcelona, Spain
daniel.gomez@upf.edu,
WWW home page: http://mtg.upf.edu/home

Abstract. This paper is inspired in the ideas of rhythmical variation and evolution, which are connected to similarity and contrast. Two experiments on rhythm similarity are presented that examine the possible relations between objective metrics and human similarity ratings. We wanted to test the possible differences in similarity ratings when a beat was induced and when it was not. The experimental design is based on identical regions inserted in the rhythmic stimuli which are progressively shifted. Twentyone subjects participated in 2 experiments devised to calibrate the effect of identical regions and beat induction in similarity ratings. Results show that identical regions can influence similarity ratings more likely when there is not a meter induced. On the other hand, the induction of a pulse is prone to elicit an attention to coincidences between rhythms. It is also observed that coincidences in the first region of a rhythmic pattern have more importance than coincidences on other regions in order to be correlated to human similarity ratings. Practical consequences of these findings are discussed in the context of tools and agents for music creation.

Keywords: rhythm, similarity metrics, syncopation, edit distance

1 Introduction

Rhythmic similarity is a research topic of of rhythm perception, that links the knowledge of cognitive mechanisms with musical practice. Knowing how alike two rhythms will appeal to a human listener is connected with the tasks of rhythmic variation and development, which are fundamental for music composition and performance. To predict how similar two rhythms are requires the comprehension of the processes that a sound undergoes from its acoustical stimulus until it can be described by its rhythmical characteristics. Some authors have used this knowledge of rhythm perception to elaborate theories and define metrics that can somehow answer the question of rhythmic similarity. For being considered relevant, these metrics have to be contrasted with perceptual similarity experiments. The most developed similarity metrics can be divided on two main categories, either in the computation of string dissimilarity (i.e. the edit distance) or in pulse induction and an implied set of temporal hierarchies (i.e. syncopation-based metrics).

The main goal of this paper is to study the relation between current similarity metrics and subjective similarity measures, and the context in which these measures are undertaken. Particularly there is an interest in understanding if the metrics are consistent both in isolated scenarios (such as an experimental settings) as to common music listening, performing and composing scenarios where a pulse is present. Two experiments were carried out, one where similarity ratings were given to a set of pairs of rhythms without any beat induction and a second experiment where the same pairs of rhythms were rated with an artificial beat induction. The difference in similarity ratings between the two experiments is explored and the results show that very often ratings differ when the same pair of patterns is compared with and without an induced beat. The discrepancy in the results suggests that different mechanisms operate in each case, each focusing in different elements of the rhythms. This different rating behaviours are related to the fact that rhythms autonomously induce the pulse by which they are going to be measured afterwards [1].

After a concise state of the art on rhythm perception and similarity metrics (section 2), we present in section 3 the methods and techniques used to test our hypotheses. Sections on results (4 and 5) and its relevance (section 6) complete this paper.

2 State of the art

2.1 Rhythm Perception

Making sense of a musical rhythm implies the ability to feel a pulse when presented with a sequence of sound events. One theory suggests that different rhythmic patterns have different induction strengths, which is the ease by with which subjects generate an internal clock commonly embodied by tapping or nodding when listening to music [2]. It is found experimentally that when a constant time cue or a rhythmic context is given, subjects are more accurate at reproducing an induced pattern. Further research suggests that this internal clock is dynamic and can adapt to temporal changes in the stimulus [3]. There are several approaches to beat induction modeling using different techniques, methods and results [4].

Lerdhal and Jackendorf propose that the pulse or tactus is used to determine the temporal structure of a musical pattern, based on isochronous subdivisions of the beat that have different salience weights or hierarchies [5]. When a musical phrase is analyzed using these hierarchies, or salience profiles, the notion of syncopation, and particularly the notion of syncopation level emerges [6]. The syncopation level is based on assigning weights to the notes on a musical phrase based on their rhythmic relations with the beat. The syncopation salience of a sound event is related to its challenging or reinforcement of the beat, where a syncopation is determined by presence of an onset anticipating the beat (or an important subdivision of it), filling the expectancy of a beat with a silence or a tie challenging the continuity of the beat [1].

Different authors use and expand the concepts of syncopation and syncopation level. Experiments have suggested a correlation between the syncopation level of a musical phrase and the difficulty for subjects to reproduce a rhythm [7]. The hierarchies established for the subdivisions and repetitions of the pulse are further explored, suggesting the existence of different weight profiles for musicians and non musicians [8]. Syncopation has also been considered when exploring the desire to move and the experience of pleasure in subjects exposed to drum breaks, suggesting that intermediate degrees of syncopation elicit the strongest desire to move and pleasure in music associated with groove. The authors propose extension of the syncopation level from the monophonic version on Longuet-Higgins and Lee to a new polyphonic measure [9]. Syncopation is a concept based on the perceptual phenomena of beat expectancy, caused by measurable features of a rhythmic pattern. This concept is used successfully in different studies associated to rhythm perception and cognition.

2.2 Rhythmic similarity

What does it mean that two rhythms feel similar? what are the criteria involved in such a judgement? what can we generalize from that claim? Many authors have explored the idea of formally modelling and predicting human subjects' ratings by means of a similarity value computed from rhythm-derived numerical features. The notions of syncopation, rhythmic salience and weight profiles have all played a role in this rhythmic similarity exploration. The simplest syncopation measure, by which two patterns can be compared, is a single number, resulting of accumulating the syncopations encountered in a musical phrase according to one of the weight profiles exposed above. More elaborated values establish how similar is the syncope of two phrases by means of a syncopation histogram (counting the different levels of syncopation found a musical phrase). Some authors [10] have found some correlations between human similarity judgements and syncopation histograms. Others such as Cao et al. [11] propose the grouping of rhythms by families that share a same level of syncope within a given window of analysis. This measure is based on dividing a phrase in isochronous fragments and labeling each fragment depending on its relation with the beat (N if the event reinforces the beat, S if it is syncope, O if it is neither). A rhythmic pattern can then be represented by a sequence of N, S or O. The resulting sequences constitute the different rhythmic families. Although the experiments show that families are a powerful discriminator for distances, the authors also report on another factor that has a higher relation with similarity which is the pattern of onsets. The authors propose that if two patterns share an identical region (IR) their similarity ratings are likely to be high. Although the authors do not explicitly define the concept of IRs on their paper, some inferences can be made from the stimuli they use to test them:

- The size of the identical regions is always longer than a beat and no longer than two and a half beats.

– Each identical region has at least three onsets.
– The identical region is always at the beginning of at least one of the tested rhythms.
– Shifts between identical regions in from pattern to pattern are always one beat or one and a half beats.
– Rhythms that are compared among them have the same amount of onsets. They have 6 or 7 onsets.
– Only in one case there is a repetition of a small fragment of one pattern present in two regions of another pattern.

The edit distance is a metric commonly used for comparing two strings, measuring the amount of transformations that one string must have in order to become the other. For example, "the eight-pulse rhythm [x x . x x . x .] may be obtained from the seven-pulse rhythm [x x . x x . .] by inserting the symbol x between the sixth and seventh pulses in the seven-pulse rhythm. A deletion is the inverse operation of an insertion. A substitution replaces one symbol for another. For instance, the eight-pulse rhythm [x x x x x x . .] may be converted to the six-pulse rhythm [x x x . . .] by changing the sixth symbol in the eight-pulse rhythm from x to . (a substitution) and deleting the first two x symbols. Thus, the edit distance simply permits the comparison of rhythms that have different numbers of pulses as well as onsets, since deletions shorten the duration of a rhythm, and insertions lengthen it"[12].

The edit distance has been used as a tool for establishing similarity ratings between musical phrases [13] [14]. Recent experiments show how the edit distance between rhythms fairly correlates with human judgements [15] [12][16]. Experiments based on a prototype rhythm, the clave son, and variations of it show that the higher the edit distance, the further the perceived closeness between both patterns [17].

2.3 Precision of metrics

As presented above, some of the rhythmic similarity metrics correlate with human similarity ratings but none has the ability to distinguish between two closely similar rhythms. The edit distance for example, seems to be a good method for clustering patterns in similarity groups, but fails at establishing further differences among patterns, bringing patterns rhythmically different into the same category. One case could be these three different rhythms,

1 0 0 0 1 0 0 1 0 0 1 0 1 0 0 0
1 0 0 1 0 0 1 0 1 0 0 0 1 0 0 0
1 0 0 1 0 1 0 0 0 1 0 0 1 0 0 0

all located at edit distance 2 from the "clave son"

1 0 0 1 0 0 1 0 0 0 1 0 1 0 0 0

but to a listener they are clearly different [17].

The shortcoming of the edit distance to establish differences among clearly different rhythms opens the door for new analyses. The question of how can a group of different rhythms that are equidistant to a reference (i.e. at a low edit distance values) be differentiated and ranked by human listeners becomes relevant. The following experiments deal with rhythmic patterns that are located at a same or close edit distance from a source, but that are recognized as different. The experiments will try to reveal relevant aspects for the definition of similarity metrics that allow a more precise and subtle classification.

Two experiments are designed to explore the characteristics of the IRs, specifically to test the relation between their sizes, locations and shift size with experimental rhythmic similarity ratings. Furthermore, we want to understand the implications of IRs when comparing two rhythms in the presence and in absence of a pulse.

3 Method

3.1 Materials

Measures The algorithms used to measure similarity between pairs of rhythmic phrases are:

The edit distance (ED): it is used only to pre-select the patterns that will be used as stimuli of the experiments. All the pairs, as will be presented on the next section, are sought to have a constant edit distance of 2.

Pattern Coincidence (PC): is measured as the percentage of equal onsets present on both patterns on the same specific region. It can be measured for the complete pattern as a single distance value, or as a set of distance values representing the coincidences found in between beats, named Pattern Coincidence by Beat (PCB).

Family Difference (FD): This measure is based on replacing the letters of the syncopation families by numbers. If there is a syncopation it is quantified as 1, if there is nothing it is marked as 0, if there is a reinforcement of the beat it is quantified as -1. Then, when comparing between two families a difference between values for each beat is computed. It can be summed up as a single value or it can be left as discrete distance values between each intra-beat region, named Family Difference by Beat (FDB).

Family Coincidence (FC): As discussed in the previous section, syncopation families are strings of letters indicating if there is a syncopation, a reinforcement or nothing within each two beats of a musical phrase. Family coincidence is measured as the percentage of syncopation coincidences found between two complete phrases. Family coincidence by beat (FCB) is a group of boolean values representing if there was a coincidence or not in between each two beats of the rhythms being measured.

To compute the PCB, the PC between each beat of the two patterns is measured obtaining a 4 number vector. Then a linear regression between all the similarity ratings and all the 4-number-vectors is computed in order to obtain the best weights for each element of the vector. The same procedure is used to obtain the FDB and the FCB.

Stimuli We selected 9 patterns as bases to create variations by shifting them. This way we created 4 variations per base pattern, yielding the total of 36 patterns used as stimuli. The variation patterns were created so that a small fragment of the base pattern (Identical Region or IR) was displaced 1 to 4 positions. Performing this shift both base and variation patterns contain the same IR but located at a certain distance from the original position. The original position of the IR was also controlled, each group had an IR selected from position 1,2,3,4 and 6. The size of all the IR is 6 steps measured from the first onset to the last onset. There are 3 or 4 onsets present on each IR. The edit distances of each main pattern to their variations is 2 (17 patterns), 3 (15 patterns), 4 (3 patterns) and 5 (1 pattern). Ideally, a fixed edit distance between main patterns and their variations would have been set; but given the need to shift the IR with precision, edit distance between patterns is 2 or 3 in most pairs. A 37th pair, consisting of two identical patterns, was added for controlling the consistency in the answers. Rhythms are reproduced with a clave sound sampled from the Roland TR-727 with no dynamic changes.

System The system used to carry out the experiment was implemented in Pure Data Extended. It consists of two play buttons to reproduce each rhythm of the pair. After listening to a pair of patterns subjects rate similarity on a 7-step Likert scale. Levels 0, 2, 4 and 6 of the scale were labeled as "The same ", "quite similar", "not very similar", "not similar at all". Levels 1,3 and 5 were not labeled.

3.2 Subjects

Twenty-one subjects (19 males, 2 females), recruited among the MTG staff and UPF pool of students, participated as subjects in this experiment. All of them with musical experience of more than 5 years at least as amateur performers. Two of the subjects were formed in non western musical traditions. The subjects were invited to participate freely in the experiment and there was no reward for their participation.

4 Experiment 1 - Stimuli without rhythmic context

The 37 rhythm pairs were presented in a subject-specific randomized order and without any possibility to listen to them more than once. On the interface each pattern is played by pressing its corresponding button (labeled as pattern A and pattern B) which is disabled after clicked. Once both rhythms have been played, the 7 step Likert scale is enabled for the subject to rate similarity. Once one pair of rhythms is rated, a button to go to the next pair is enabled. When the next button is pushed a 4-seconds pause is started and then the next pair is loaded and the play buttons are active again. This procedure is repeated until all the 37 pairs are ranked. After finishing the experiment a questionnaire is presented verbally to the subjects in order to asses particular aspects of the activity. The questions asked are: How difficult do you think the experiment was? Did you use any physical action (nodding, tapping, other) to help you measure similarity? How did you asses the similarity? Did you use any algorithm to define similarity? Do you think any part of the pattern is more important than other in your ratings? The data collected from the experiment is the index of the pair of rhythms, the similarity rating given by the subject and the order in which the pairs were assigned to pair A and pair B.

The analysis of the results is based on the mode of the ratings for every stimuli pair. To get a better picture of the results of experiment 1, pairs with more than 50% of the results scattered over 2 Likert scale marks were discarded as inconsistent between subjects. Only 16% of the pairs were removed, namely pairs 1, 3, 4, 19, 23 and 35. The dispersions of all pairs is shown in Figure 1. Each pair

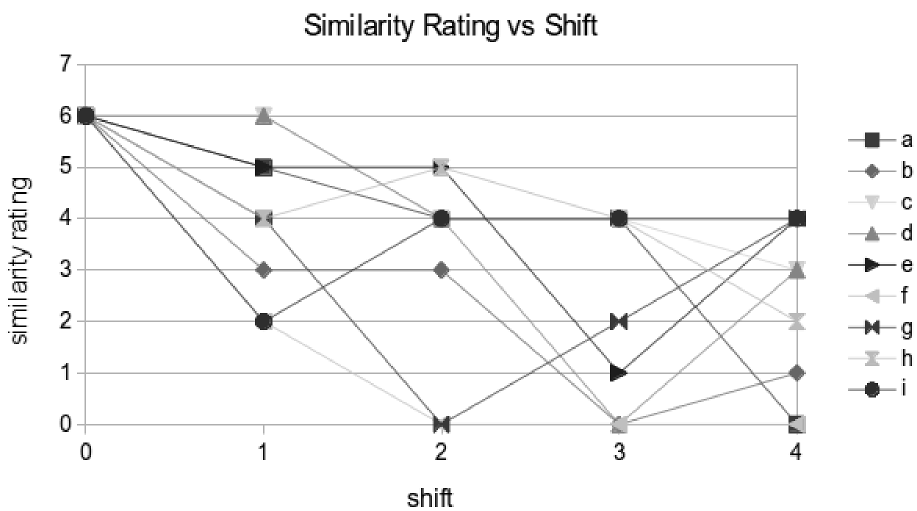

Fig. 1. Similarity rating vs shift without the presence of a rhythmical context discriminated by groups, from a to i. 6:the same, 0:not similar at all

of stimulus, plotted by its similarity ratings and the shift of the IR, is organized by groups from a to i on figure 3. There is a trend that suggests that an increase in the shift reduces the similarity rating (Friedman chi-squared = 23.878, df = 4, p-value = 8.45e-05) with significant Spearman rank order correlations (a: -0.97, b: -0.87, c: -0.95, d: -0.87, e: -0.87, f: -0.89, g: -0.46, h: -0.82, i: -0.22. P-Values a:

0.0048, b: 0.0539, c: 0.0138, d: 0.0539, e: 0.0539, f: 0.0405, g: 0.4338, h: 0.0886, i: 0.7177). Low P-values and high negative Spearman rank order correlations suggest a negative correspondence between shift and similarity, the further the IR is shifted the lower the similarity rating. Possible relations between objective measures (see 3.1) and the similarity results obtained in the absence of a rhythmical context are presented on Table 1. All Spearman Correlation values are all below significance, suggesting that similarity ratings are not correspondent to any of the objective measures from the rhythmic patterns.

To calculate the PCB, FCB and FDB the PC, FC and FD are computed between each of the four beats of every pair. Then a linear regression between the four computed values (all patterns are 4 beats long) and the prediction is computed to extract the weights for each beat. These weights will be discussed in section 6.

Table 1. Spearman Rank correlation values for each objective metric and the similarity ratings without a rhythmic context

Objective metric	Spearman correlation
PC	0,114158372
PCB	0,230098172
FCB	0,426401379
FDB	0,316359972
FD	0,093654438

These results suggest a relation between the IR and similarity ratings when the rhythms are presented to the subjects without beat induction. It also gives clues to the features of the IRs such as the size and shift, complementing the results of Cao et al.

5 Experiment 2 - Stimuli with rhythmic context

In this experiment we wanted to test if a rhythmic context, given the presence of a metric-inducing sound (i.e. a kick drum), could modify (and in which sense) the similarity ratings given to the same pairs used in experiment 1. In every stimuli of this experiment a kick drum is played four times on the start of every beat at a tempo of 120 beats per minute, then the kick drum and one of the patterns of the pair is played simultaneously, then just the kick drum again four times and finally the kick drum simultaneously with the remaining pattern. Again, the 37 rhythm pairs were presented in a subject-specific randomized order. The same interface was used as in experiment 1, but this time the play button reproduced the whole sequence of kick, kick + rhythm A, kick, kick + rhythm B. Two play buttons were active so this time subjects could listen twice to the whole identical sequence. After finishing the experiment the same questionnaire as in experiment 1 is presented verbally to the subjects in order to asses particular aspects of the activity. The data collected from the experiment is the index of the pair of rhythms, the similarity rating given by the subject and the order in which the pairs were presented during the whole sequence.

As in experiiment 1, the analysis of the results is based on the mode of the ratings for every stimuli pair. In order to get a better picture of the results obtained in experiment 2, the stimuli pairs are analyzed in search for the most consistent inter-subject ratings (Figure 4). Ratings with 50% of the results spread out three or more perceptual scale values are removed, namely pairs 23, 24, 26, 27, 28, 35, being the 16% of the original set.

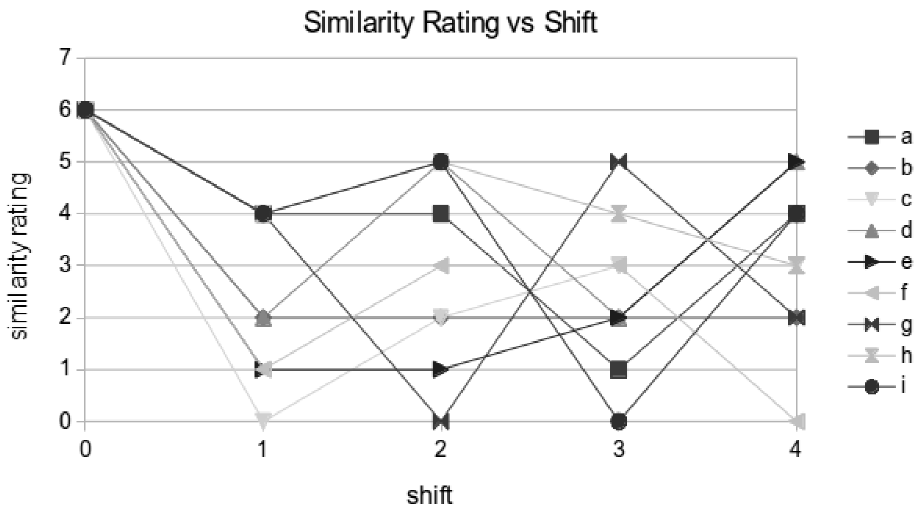

Fig. 2. Similarity rating vs shift with the presence of a rhythmical context discriminated by groups, from a to i

Every stimulus pair is compared with the shift of the IR from one pattern to the other and with the similarity rating obtained when a rhythmic context was present (Figure 2). These results show some signifficant correlation values correlation values (a: -0.67, b: -0.71, c: -0.05, d: -0.32, e: -0.05, f: -0.67, g: -0.50, h: -0.82, i: -0.67) but very high P-values on the Spearman Correlation test (a: 0.2152, b: 0.1817, c: 0.9347, d: 0.6042, e: 0.9347, f: 0.2189, g: 0.3910, h: 0.0886, i: 0.2189). Friedman rank sum test has also an elevated P-value = 0.5834. These values strongly suggest that no aspect of the IR is relevant to asses rhythmic similarity when a pulse is induced.

Possible relations between similarity ratings and the different objective measures extracted from the patterns were also analyzed (Figure 3). As in experiment 1, the PCB, FDB and FCB are computed after the PC, FC and FD are computed between each of the four beats of every pair. Alinear regression between the four computed values (all patterns are 4 beats long) and the prediction is calculated to extract the weights for each beat. These weights will be discussed in the next section.

The PC measure has a Spearman correlation value just above significance with the similarity ratings (0.54365, p-value = 0,001902). Measures in which beats are weighted independently also exhibit just above significance Spearman correlation values: PCB (0.76171, p-value = 1,01E-003), FCB (0.52926, p-value = 0,002635). These values, opposed to the analysis of the IR, evidence how different objective measures respond accordingly to similarity ratings when a pulse is present.

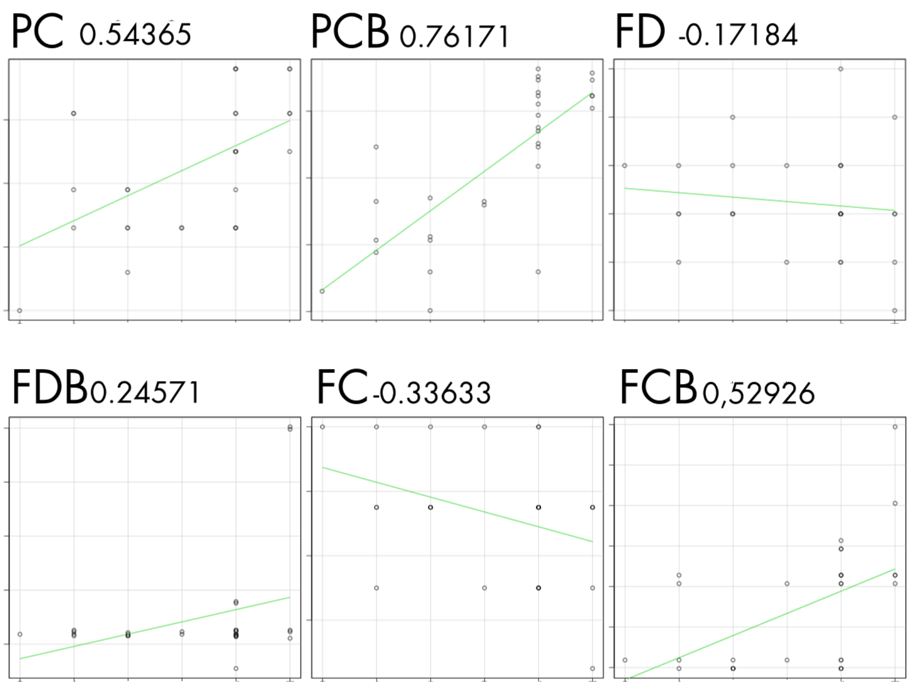

Fig. 3. Objective similarity measures plotted against human similarity ratings of experiment 2. Spearman correlation value on top of each plot. Pattern Coincidence (PC), Pattern Coincidence by beat (PCB) and Family Coincidence by beat (FCB) have an above significant Spearman correlation.

6 Discussion and conclusions

A general view of the data shows a clear difference between the similarity results obtained for the same pairs of rhythms depending whether they are presented within a rhythmic context or not. This can be seen on Figure 4, where the between-subject similarity obtained for the 37 pairs in both experiments is not convergent. In some cases it is the same (pairs 2, 12, 24, 25, 26 29, 30, 31, 36) in some cases highly contradictory (pairs 3, 4, 9, 13, 17, 18, 22, 23, 27, 35) but generally in disagreement (75% of the pairs). The same pairs of rhythms are rated differently depending on the presence or absence of a rhythmic context.

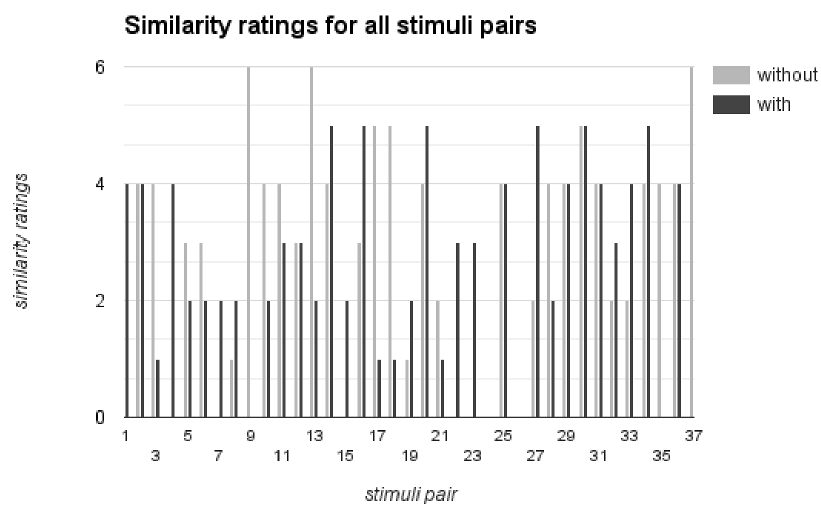

Fig. 4. Similarity ratings for all stimuli pairs. Results with rhythmic context dark gray, without rhythmic context light gray.

The influence of shift in similarity ratings in experiment 1 and experiment 2 also differ in tendency. While on experiment 1 (no rhythmic context) shift seems to have an inverse correspondence with similarity, for most of the groups on experiment 2 (with rhythmic context) no direct relation with the shift is appreciated. Presumably the emergence of IRs and their shift as a relevant factor for rhythmic similarity only in the case where there is no rhythmical context, could be related to an alternative perceptual mechanism triggered when no metrical cues are offered to decypher a musical sequence in terms of its rhythmical properties. This mechanism could be analogous to comparing the similarity between two words by looking at the letters and their order and not by their meaning. It seems to be clear that a shallow similarity computation may happen based on superficial features and no rhythm context whereas a more abstract and layered mechanism operates when a metric context is set.

This previous observation is aligned with the fact that the amount of notes needed for a beat to be induced is from 5 to 10 [4]. This could lead to conclude that in experiment 1 a sense of beat was acquired just as every phrase was ending and therefore no metrical structure was ever induced during this experiment. Nevertheless as 36 pairs of rhythms at 120 BPM were listened during experiment 1, a reminiscent notion of the tempo after each exposition could be accumulated and influenced further comparisons. This observation is out of the analysis and all results of experiment 1 are treated as non beat inducing. Another factor that is left out is the possible use of memory to recall a rhythm that just finished with the late acquired meter, so the rhythm is evoked with a meter although such meter was not originally present.

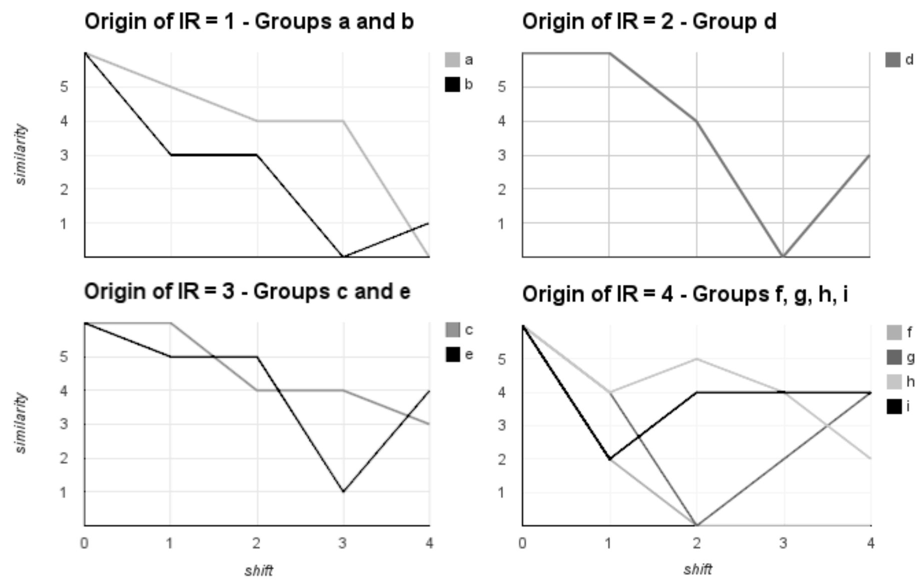

Fig. 5. Relationship between origin and shift for experiment 1.

Groups that have lower correlation between the shift and the similarity ratings are groups in which the IR had an origin closer to the start of the rhythm. The farthest the origin, the least correlation between shift and similarity ratings (see Figure 5). Spearman correlations for different origins are origin 1: -1.0000, origin 2: -0.8721, origin 3: -0.9000, origin 4: -0.6669. Their respective pairwise two-sided p-values are 0.0001, 0.0539, 0.0374, 0.2189. Although all Spearman correlations are high, their significance decreases progressively as the origin of the IR increases. This decline in significance can be related to a loss of strength in the alternative perceptual mechanism mentioned above. In other words the IR effect is stronger when the IR is on the first steps of a sequence. The matter of what happens when the IR is further than 4 steps was not considered while preparing the stimuli, therefore is out of the range of the experiments.

The same analysis of the influence of the IRs origin and shift with similarity in experiment 2 yields the following Spearman correlation values : origin1: -0.6708204, origin 2: -0.3162278, origin 3: 0.00000001, origin 4: -0.97467943. The corresponding pairwise two-sided p-values are 0.2152, 0.6042, 1.0000, 0.0048 respectively. None of the p-values accounts for significance, again showing a disengagement between inducing a meter and the relevance of having an IR in two patterns. This can lead to the conclusion that a mechanism based on IRs is not relevant when an induced meter is present.

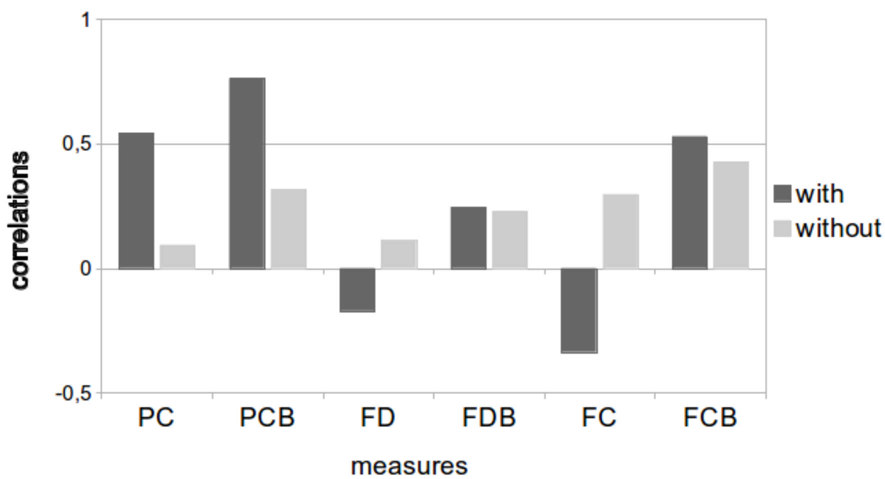

Fig. 6. Spearman correlation between different objective measures extracted from the patterns and the resulting perceptual results of each experiment. Light gray: without beat induction (Experiment 1), dark gray: with beat induction (Experiment 2)

A different phenomena occurs with rhythms assessed in a metrical context. Figure 6 shows that the highest correlations between objective measures and subjective similarity ratings are obtained when subjects were exposed to a rhythmic context. Correlation indexes of the data collected without a rhythmic context, on the other side, are always below significance. Measures with correlations above significance are the pattern coincidence (PC), pattern coincidence by beat (PCB) and family coincidence by beat (FCB). The best Spearman correlation ranks are obtained with measures computed by beat and weighted independently. One way to interpret the weights is to think of them as analogous to importance and could be paired with a notion of awareness in a perceptual sense. These weights could possibly be giving clues of hierarchies imposed on different sections of a pattern when a subject evaluates similarity in a rhythmic context.

Looking at Table 2, we see that coefficients for beat 1 are very prominent both for the PCB (pattern coincidence by beat) and the FCB (family coincidence

Table 2. Weights for each beat for the different measures correlated above significance in experiment 2

metric	beat1	beat2	beat3	beat4
PCB	4.21520	1.15980	0.94951	0.70146
FCB	-1.47619	0.10389	-1.04761	-1.93506

by beat) measures. Although the weight curve is different for both of them. A progressive descending curve for PCB, suggests a decline in importance of the beats as time advances. An n-shaped curve for the FCB suggest a higher relevance for the first and last beats of the rhythms.

As a final summary five observations regarding both experiments can be presented.

– Similarity ratings of patterns change depending on the presence or absence of a pulse which metrically coincides with the onsets of the patterns being measured.
– In the absence of a pulse, a mechanism based on searching identical regions (IR) of one pattern into the other one is predominant over coincidences and syncopation for giving a similarity rating.
– Similarity ratings without a rhythmical context are inversely related with the shift in steps of the IR from one pattern to the other.
– In the presence of a pulse, a mechanism based on coincidences and syncopation is more relevant for predicting human similarity ratings.
– Similarity ratings in the presence of a pulse weight particular differently the regions of the rhythms being compared.

Our studies on rhythm similarity, in addition to providing hints on the working of the cognitive musical processes, is leading towards interesting practical applications. These correlated measures and their coefficients can be the source of algorithms used in musical classification and creation scenarios. Further research could use implementations of our results into metrics to analyze melodic data sets. Generative algorithms can also make use of these results as rules to create new rhythms based on a seed pattern controlling the desired similarity level.

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).

References

1. Honing, H. (2012). Without it no music: beat induction as a fundamental musical trait. Annals of the New York Academy of Sciences, 1252(1), 85-91.

2. Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. Music perception, 411-440.
3. Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. Psychological review, 106(1), 119.
4. Desain, P., & Honing, H. (1999). Computational models of beat induction: The rule-based approach. Journal of new music research, 28(1), 29-42.
5. Lerdahl, F., & Jackendoff, R. (1985). A generative theory of tonal music. MIT press.
6. Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. Music Perception, 424-441.
7. Fitch, W. T., & Rosenfeld, A. J. (2007). Perception and production of syncopated rhythms.
8. Ladinig, O. (2009). Temporal expectations and their violations. PhD thesis.
9. Witek, Maria AG, Eric F. Clarke, Mikkel Wallentin, Morten L. Kringelbach, and Peter Vuust. "Syncopation, body-movement and pleasure in groove music." PloS one 9, no. 4 (2014): e94446.
10. Tutzer, F. (2011). Drum rhythm retrieval based on rhythm and sound similarity. Master's thesis, Departament of Information and Communication Technologies Universitat Pompeu Fabra, Barcelona.
11. Cao, E., Lotstein, M., & Johnson-Laird, P. N. (2014). Similarity and Families of Musical Rhythms. Music Perception: An Interdisciplinary Journal, 31(5), 444-469.
12. Toussaint, G. T., Campbell, M., & Brown, N. (2011). Computational models of symbolic rhythm similarity: Correlation with human judgments. Analytical Approaches to World Music, 1 (2).
13. Mongeau, M., & Sankoff, D. (1990). Comparison of musical sequences. Computers and the Humanities, 24(3), 161-175.
14. Orpen, K. S., & Huron, D. (1992). Measurement of similarity in music: A quantitative approach for non-parametric representations. Computers in music research, 4, 1-44.
15. Thul, E., & Toussaint, G. T. (2008, September). Rhythm Complexity Measures: A Comparison of Mathematical Models of Human Perception and Performance. In ISMIR (pp. 663-668).
16. Guastavino, C., Gomez, F., Toussaint, G., Marandola, F., & Gomez, E. (2009). Measuring similarity between flamenco rhythmic patterns. Journal of New Music Research, 38(2), 129-138.
17. Post, O., & Toussaint, G. (2011). The Edit Distance as a Measure of Perceived Rhythmic Similarity. Empirical Musicology Review, 6(3).