QUANTIFYING METRICAL AMBIGUITY

Patrick Flanagan

ABSTRACT

This paper explores how data generated by meter induction models may be recycled to quantify metrical ambiguity, which is calculated by measuring the dispersion of metrical induction strengths across a population of possible meters. A measure of dispersion commonly used in economics to measure income inequality, the Gini coefficient, is introduced for this purpose. The value of this metric as a rhythmic descriptor is explored by quantifying the ambiguity of several common clave patterns and comparing the results to other metrics of rhythmic complexity and syncopation.

1 INTRODUCTION

This study initially grew out of an interest in modeling listener responses of a tapping study in which listeners tapped to the beat of sections of Steve Reich's *Piano Phase* [17]. For some sections, tapping rate and phase varied widely among listeners, and for others they displayed relative consistency. Accordingly, the goal of the study was to model metrical ambiguity as the degree of variety of listener responses to a given rhythm.

This task, however, would be complicated by a lack of simplified experimental data; the experiments of [17] involved pitched music, which added a layer of complexity I wanted to avoid. I therefore turned my attention to a corollary kind of metrical ambiguity, which is the amenability of a rhythm to gestalt flip, in which a listener reevaluates the location of the beat. This amenability has been convincingly advanced as one reason for the pull of African and African-derived rhythms on listeners [7]. Data on metrical gestalt flips is no easier to come by, but orientating the study around the proverbial average listener instead of groups of listeners allowed it to interface with the voluminous literature on syncopation. While metrical ambiguity is not equivalent to syncopation, the results of this study suggest that it roughly correlates to syncopation and offers an interesting rhythmic descriptor in its own right that could be used for theoretical analysis, genre classification, and the production of metrically ambiguous music.

To describe the ambiguity of the metrical scene, the ambiguity model recycles data that existing models of meter and beat induction often discard, the data about the induction strengths of *all* potential meters. The central idea of the model is that a more even distribution of induction strengths across a group of potential meters will cause greater ambiguity in the perception of meter. The model of ambiguity, therefore, requires a model of meter induction that, for a given rhythm, produces numerical values for an array of potential meters, with each meter defined by its period and phase relation to some arbitrary point in the rhythm. For the sake of simplicity, this study uses a hybrid model based on older models of meter induction that operate on quantized symbolic data, but the theory of ambiguity presented here is extensible to any model, including those that work with audio, that produces such an array.

In this paper the term "meter" refers to an isochronous, tactus-level pulse, characterized by a period and phase expressed in discrete integer multiples of some smallest unit of musical time. The more common term for this isochronous pulse is, of course, "the beat." "Meter," however, is the preferable term because this paper uses the conventionally understood downbeat of musical examples as a reference point for the tactus-level pulse.

2 METER INDUCTION MODEL

2.1 Influences on the Model

While the intent of this paper is not to advance a new model of meter induction, the existing models are either too complex for the purposes of this study or otherwise flawed. I therefore employ a hybrid meter induction model that calculates an induction strength for each candidate meter by summing the values of phenomenal accents that a given meter encounters, a strategy inherited from [12; 9; 2; 10]. Parncutt's model is the most sophisticated of the four in that it incorporates durational accent and tempo preference. For this reason, I largely adopt Parncutt's model.

Parncutt's model, however, has one flaw that prevents its full deployment in this study. In calculating what he called "pulse-match salience," which is the degree of match between phenomenal accents and an isochronous pulse stream, Parncutt summed the *product* of the phenomenal accents that occur on neighboring pulses. The problem with multiplying the phenomenal accents that occur on neighboring pulses is that some rhythms contain no inter-onset intervals (IOIs) that map onto candidate meters that we know to be viable meters. In this case, the metrical induction score, or pulse-match salience, of a viable meter turns out to be 0. For example, consider the Bossa Nova pattern in Figure 2.

The rhythm contains no inter-onset intervals of 4 tatums that begin on what, according to musical practice, is the beat. There is one IOI of 4 tatums in the middle of the pattern, but it is shifted by 2 tatums off the beat. The pulse-match salience as calculated by Parncutt's formula for what is conventionally understood as the meter of this rhythm is 0, which clearly is not a desirable result. This problem did not crop up in Parncutt's article because the rhythmic patterns he investigated are relatively simple and unsyncopated and therefore have at least one IOI that coincides with a reasonable metrical candidate. As discussed below, the model presented here avoids this problem by summing the phenomenal accents that coincide with a candidate meter regardless of their adjacency.

2.2 The Meter Induction Algorithm

The model first describes a rhythmic surface as a series of values indexed to the tatums of the rhythm, with 0s signifying a rest or continuation of a previously attacked note and non-zero values signfiying phenomenal accents. The model follows that of Parncutt and uses only durational accent in modeling phenomenal accent and is calculated as

$$A_d\left(T\right) = \left(1 - \exp\left(\frac{-IOI\left(T\right)}{r}\right)\right)^i \tag{1}$$

where T refers to the index of an event in tatums, exp is the natural exponential function, *IOI(T)* is the duration in milliseconds of the IOI from the onset at tatum T to the next onset, *r* is the saturation duration, and *i* is the accent index. Calculating *IOI(T)* requires fixing the speed (i.e., tatum duration), which is set to 200 ms for the examples in this paper. Saturation duration is the threshold of short term auditory memory, beyond which point lengthening the duration of an IOI will increase only marginally its durational accent. Parncutt, by optimizing the value of *r* to fit his experimental results, arrived at 500 ms, and I adopt this value in the model. The free parameter *i* magnifies the difference in accent between long and short durations. I follow Parncutt in setting it to 2.

The next step is generating the meters that will be tested against the accent vector. The model follows [11] in using only those meters with a period that is a factor of the length of the accent vector and excludes meters with extremely fast or slow tempos. Each candidate meter is further characterized by its phase, which is measured from index 0 of the accent vector. The preliminary induction strength *S* for meter *m*, characterized by period *P* and phase *Q*, over A_d, can then be calculated as

$$S_m = \frac{1}{n}\left(\sum_{i=0}^{n-1} A_d\left(Pi + Q\right)\right) \tag{2}$$

where *n* is the number of beats of the meter that occur in the pattern. It is equal to the pattern length divided by the period of the meter. Multiplying the sum of accents by the reciprocal of *n* normalizes the induction strengths with respect to period length. As Parncutt noted, normalizing for period length allows us to explicitly introduce a model of human tempo preference, which numerous researchers have fixed at 600 ms [5]. Parcutt formalized tempo preference as

$$S_p = \exp\left(-\frac{1}{2}\left(\frac{1}{\sigma}\log_{10}\left(\frac{p}{u}\right)\right)^2\right) \tag{3}$$

where S_p is the induction strength for period *p* of the meter and *u* and σ are the free parameters representing the preferred tempo and the standard deviation of its logarithm. Here I depart from Parncutt and use the commonly observed values of 600 ms and 0.2, respectively.

Incorporating the model of tempo preference into the induction strength model yields

$$S'_m = S_m S_p \tag{4}$$

Applying (4) to the Bossa Nova pattern of Figure 2, assuming a speed of 200 ms, yields induction strengths for 16 different candidate meters, which are presented in the bar chart of Figure 1, ordered from lowest induction strength to highest. The next task is to quantify the evenness of this distribution.

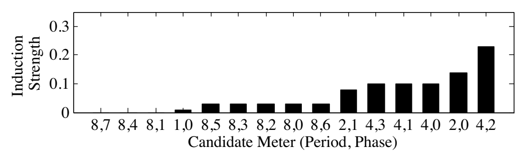

Figure 1. Induction strengths of the Bossa Nova pattern, ordered lowest to highest, with speed set to 200 ms.

3 THE GINI COEFFECIENT

Statisticians have developed a number of tools for quantifying the evenness of a distribution of values. One of them, the Gini coefficient, has gained popularity in economics as a measure of income inequality because it is easy to interpret and its value is independent of the scale of the data (i.e., unit of measurement) and the sample size. For this study, the inequality of the dispersion of induction strengths across a population of candidate meters is calculated. In the formula's discrete version, for a population *P* of n candidate meters with induction strengths calculated from (4), the Gini coefficient can be calculated as

$$G(P) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \mid x_i - x_j \mid}{2n^2\mu} \tag{5}$$

where μ is the mean induction strength, and x_i and x_j are induction strengths of the candidate meters. It produces values with a range of 0 to 1, where 1 signifies perfect inequality and 0 signifies perfect equality. In music theoretical terms, lower values denote greater metrical ambiguity, and vice versa. In practice, values for metrical ambiguity range from the high 0.20s for high metrical ambiguity up to the mid 0.70s for low metrical ambiguity. With the necessary tools in hand, we may now proceed to assess the metrical ambiguity of some rhythms.

4 METRICAL AMBIGUITY AND CLAVE RHYTHMS

4.1 Comparison to Syncopation Measures

Thus far the Gini coefficient has been described as a measure of metrical ambiguity, which is not necessarily synonymous with rhythmic complexity or syncopation. A fast isochronous pulse stream is metrically ambiguous because it is amenable to subjective rhythmization, but it is also rhythmically simple. Nevertheless, it is fruitful to compare the Gini coefficient to existing measures of syncopation.

Toussaint has tested several metrics of complexity and syncopation against clave rhythms in [16; 15], with particular attention paid to six of the most popular 4/4 timelines: the Bossa Nova, Gahu, Rumba, Soukous, Son, and Shiko patterns (Figure 2). Histograms of the induction strengths according to (4), ordered lowest to highest, of each rhythm with a moderate tatum duration of 200 ms are presented in Figure 3 (except for the Bossa Nova pattern, the histogram of which is presented in Figure 1). Note that the bars are labelled by their period and phase, and that the candidate meter that is most commonly understood to orgranize the rhythm in musical practice is 4,0. A simple visual inspection of the histograms reveals variation in the evenness of the distribution of induction strengths; to the eye, Shiko is the most uneven, and Bossa Nova the most even. As one would expect, these differences are reflected in the Gini coefficients of each rhythm, which are displayed in the first row of Table 1 alongside several metrics of rhythmic complexity and syncopation. In reading Table 1, be mindful that the Gini coefficient measures inequality of metrical induction strengths and therefore is inversely related to metrical ambiguity. A low value implies higher metrical ambiguity.

A word about the different metrics of complexity in Table 1. "Gini Syncopation" is derived from the Gini coefficient as explained in section 4.2. Michael Keith's measure of syncopation assigns each note to one of four categories: unsyncopated, hesitation, anticipation, and syncopation. Each category is associated with a different value or degree of syncopation, and the total syncopation of a passage equals the sum of its constituent notes' values. The reader is referred to [15] for more information. Pressing's measure [13] is similar in that it parses a rhythm into six basic units and sums their values.

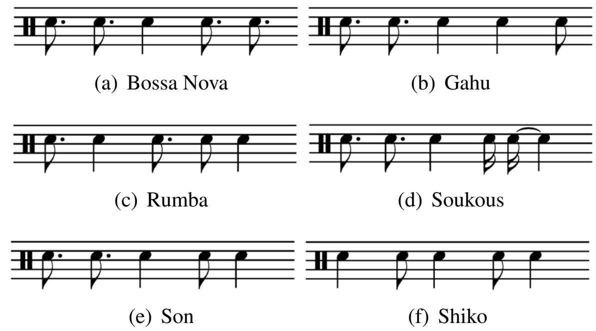

Figure 2. Six common 4/4 timelines.

	Bossa Nova	Gahu	Rumba	Soukous	Son	Shiko
Gini	0.583	0.719	0.679	0.683	0.679	0.803
Gini Syncopation	0.894	0.841	0.733	0.876	0.571	0.492
Keith	3	3	2	3	2	1
Off-Beatness	2	1	2	2	1	0
Pressing	22	19.5	17	15	14.5	6
Metrical Complexity	6	5	5	6	4	2
LHL	6	5	5	6	4	2
WNBD	4	3.6	3.6	3.6	2.8	1.2

Table 1. A comparison of measures of rhythmic complexity and syncopation for six clave rhythms.

Two different measures of rhythmic syncopation were proposed in [15]. The first is off-beatness, which is equal to the number of onsets that occur at indices of the measure that share no factors other than 1 with the length of the measure (a measure in usually indexed in 16th notes). The second measure of syncopation, called "metrical complexity," employs the metrical accent grid of [6]. In this scheme, the level of syncopation is related to the coincidence of onsets and shallow levels in Lerdahl and Jackendoff's metrical accent grid. The Longuet-Higgins Lee syncopation measure (LHL) also employs a metrical accent grid while considering whether an onset is followed a rest (or continuation) or another onset [8]. Finally, the Weighted Note-to-Beat Measure (WNBD) relates syncopation to the distance from each onset to the nearest beat [4].

The extreme Gini values in Table 1 correspond well to the other metrics of syncopation and to intuition; Bossa Nova is the most syncopated pattern and Shiko the least. Most of the metrics, however, have different interpretations of Gahu, Soukous, Son, and Rumba. According to [16], the patterns should be ordered in decreasing degree of syncopation as Bossa Nova, Gahu, Rumba/Soukous, Son, and Shiko, which is exactly the order of Pressing's model. I would have

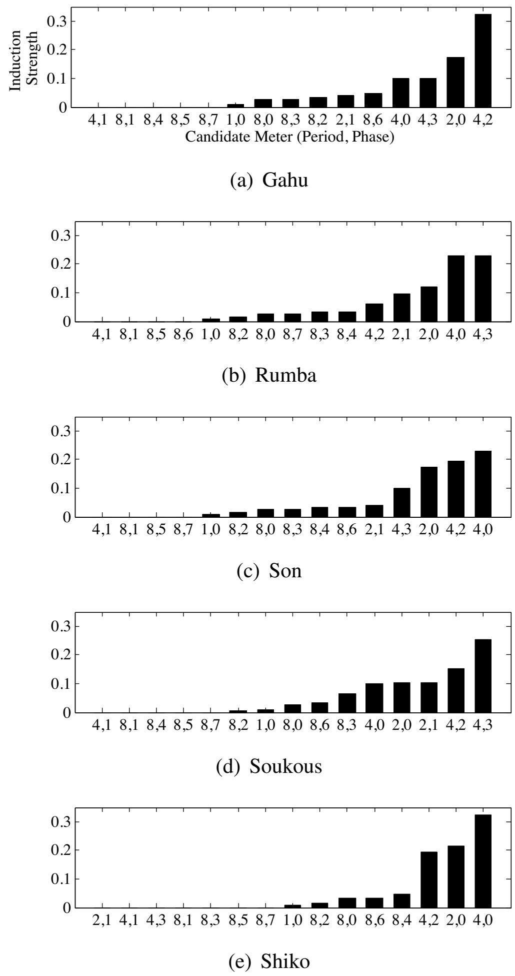

Figure 3. Histograms of the induction strengths of potential meters of five additional clave patterns. The two numbers below each bar represent the period and phase of a candidate meter.

to quibble with the assesment of Gahu as the second most syncopated pattern. The adjacency of the last two notes of Soukous, a uniquely tight spacing among the patterns, makes the pattern sound more syncopated by accenting the last onset, which is not on the beat, in two ways. First, the tight spacing of the fourth and fifth onsets and the relatively long IOI before the fourth onset makes the fourth onset function as a kind of pick-up to the fifth. Second, the last onset receives additional accent by virtue of the long inter-onset interval that follows it, which at five tatums, is the longest IOI of any of the patterns. The combined effect of the fourth onset functioning as a pick-up and the durational accent of the first onset lend strength to hearing the fifth onset as on the beat, an effect which, because the fifth onset is not on the beat, gives the rhythm a strongly syncopated feel. Gahu, by contrast, confirms the location of its “true” meter by placing its shortest IOI of two tatums between the last and first onsets. This eighth note IOI provides a pick-up into the real downbeat.

Clearly, experiments should be carried out to settle the matter, but for the sake of argument, let us bump Soukous ahead of Gahu on the syncopation scale. It would then read Bossa Nova, Soukous, Gahu, Rumba, Son, and Shiko. This slight rearrangement does nothing to help the Gini coefficient as a model of syncopation because it still ranks Son and Rumba as more complex than Soukous and Gahu. The reason it does this, however, is that it does not measure the metrical ambiguity of these particular rhythms but of the beat-class set types to which they belong. The Gini coefficient has no conception of a rhythm in a particular metrical context, whereas the other metrics assess a rhythm against a predetermined metrical framework, which listeners in real musical situations may infer from cues such as harmony and melody or other instruments. To this extent, comparing the Gini coefficient to the other metrics is comparing apples and oranges. As we shall see, using the Gini coefficient to describe beat-class sets is illuminating in certain contexts, but a metric of rhythmic complexity or syncopation should be able to distinguish a series of quarter notes that are on the beat from a series of quarter notes that are off the beat. The Gini coefficient does not do this.

4.2 From Metrical Ambiguity to Syncopation

There is a solution to this problem that, while adding additional mathematical complexity to an equation that is already far removed from the surface phenomena of the music, does at least correspond to our intuitions while offering several advantages over the existing metrics of rhythmic complexity and syncopation. The reason Gahu and Soukous are incorrectly assessed by the model is that it pays no heed to the relatively low induction score given to the meter that contextualizes the rhythm in musical practice (by defininition, the candidate meter 4,0). Indeed, one striking feature of Figures 2 and 4 is that for three of the rhythms (Bossa Nova, Gahu, and Soukous), the meter 4,0 is not even close to having the highest induction strength. A metric of syncopation should reflect this.

What is needed, then, to transform the Gini coefficient into a metric of syncopation in the mold of the others is a way of incorporating information about the strength of the meter understood to be the real meter relative to the other candidate meters. Furthermore, to aid comparison to other metrics, the score should increase with syncopation. The relative strength of the actual meter, 4,0, in comparison to the others can be neatly measured as its ranking, r, in the order list of induction strengths, and is set equal to the number of other meters that have an induction strength greater than or equal to that of the actual meter. Thus for Bossa Nova and Soukous, for example, r equals 5. We can use the reciprocal of r and the Gini coefficient to calculate a measure of syncopation. For an accent vector A_d that produces a population *P* of meter induction strengths, its syncopation *Sync* may be described as

$$Sync(A_d) = 1 - \left(G(P)(1 - e^{-\frac{1}{r}})\right) \quad (6)$$

Inputting the values for the Gini coefficient and meter induction strengths calculated above yields the scores in the second row of Table 1.

These values correspond with what I have argued above is the most reasonable ranking of the patterns. This metric enjoys two other advantages over the others. The first is that it is invariant to scale. The other metrics invariably increase with the length of the rhythm being measured, making it impossible to compare rhythms of different lengths. The syncopation measure proposed here, on the other hand, always produces a value between 0 and 1. Second, it has a much finer grain of discrimination than the others. On the downside, it depends on a large number of free parameters, a circumstance that prevents fixing a final value for the syncopation of a rhythm conceived in the abstract; there is no single score for the Bossa Nova pattern. This fault could, however, be turned into an advantage if it allowed analysts to probe the effects of tempo and various kinds of phenomenal accent on syncopation.

4.3 Correspondence to Experimental Results

In [14], Spearman rank correlations were calculated between the experimental data of [12] and [3], which measured the difficulty of performing a set of rhythms, and the values produced by various models of syncopation for the same rhythms. The correlations give an indication of how well the models predict performance difficulty, which is assumed to be a reasonable proxy for syncopation. For the models tested, which were those listed in Table 1 (excluding the Gini coefficient models presented here), correlations ranged from 0.224 to 0.787. The procedure of [14] was repeated for the Gini coefficient and the Gini syncopation metrics, with the following results: The Gini coefficient showed correlations of 0.51 and 0.42 for the data of [12] and [3], respectively. The Gini syncopation measure produced correlations of 0.56 and 0.51.

These figures should treated cautiously. The Essens study [3] was a small experiment with only six participants and high *p*-values (i.e., low confidence). That said, the results indicate that the Gini coefficient and Gini syncopation measure are decent but not good predictors of performance difficulty. Considering that the best performing models in [14], the LHL and metrical complexity models, operate in reference to a full metrical hierarchy, one could speculate that the performance of the Gini measures could be improved by replacing the modified Parncutt beat induction model with one that uses a metrical hierarchy. It is also noteworthy that these same two models produced rankings of the six clave rhythms in Table 1 that agree with my own assessment and that of the Gini syncopation measure.

5 METRICAL AMBIGUITY OF BEAT-CLASS SET TYPES

The concept of the beat-class set was (re)introduced in [1] in order to analyze the properties and transformations of rhythmic material in Steve Reich's phase-shifting music (the concept originated in Babbitt's timepoint system). At issue in [1] was the transformational potential of rhythmic material, a topic that the Gini coefficient addresses because it describes how amenable a beat-class (bc) set type might be to various metrical interpretations. To investigate what the Gini coefficient might reveal about metrical ambiguity and beat-class sets, Gini coefficients for all mod 12 bc set types at three different speeds were calculated.

The most striking feature of the results is that the bc types with the highest ambiguity are those formed by one of the generator cycles of the modulo 12 system, the 5- (or 7-) cycle, the generator of the diatonic scale. In another interesting twist, if one skips applying rules for phenomenal accent and calculates induction strengths over an accent vector of 1s and 0s, the Gini coefficients of bc set types are invariant under M operation. Under M operation, the 5 cycle transforms into the 1-cycle, which means that, at least when accent rules are ignored, the two most famous scales of Western music, when transformed into rhythms, also turn out to be, at least according to the model developed here, the most metrically ambiguous bc set types. It might seem surprising that the dispersion of values produced by this model of meter induction, which was designed to reflect human preference for aligning isochronous downbeats with note onsets, would intersect so directly with set theory, but the explanation is almost intuitive.

By definition, the candidate meters have periods that are factors of the length of the meter. The generators cycles, by definition, have periods that are prime relative to the length of the meter and therefore prime relative to the periods of the candidate meters. The relative primeness of the candidate meters and the cycle generators ensures that the generator will cycle through all of the meters of a different phase and the same period before coinciding with the same meter twice. This cycling over candidate meters produces the most even distribution of onsets among the candidate meters, which in turn assures a low Gini coefficient.

It should be obvious from the status of cycle 1 generated beat-class sets as highly ambiguous that maximal ambiguity is hardly a guarantor of musical interest. As Cohn [1] noted in a parenthetical aside, these sets are "inherently less interesting for other reasons as well, although to articulate these reasons is not an easy task." One possible reason is that while bc sets generated from the 1-cycle are metrically ambiguous, their grouping structure is all too obvious. The maximally even dispersion of onsets in the 5-cycle generated sets creates the potential for ambiguity of grouping as well as metrical structure, allowing grouping and metrical structures to interact in interesting ways. Alternatively, perhaps the maximally even sets are more interesting because they constantly flip the beat around, forcing the listener to retrospectively reevaluate the location in the metrical grid of what has already been heard. In this sense, the metrical ambiguity of maximally even rhythms is like that of the familiar visual phenomenon, in which the image of two faces opposite one another can be flipped to the background to produce an image of a vase. The perceptual gestalt is fragile but clearly delineated. The 1-cycle generated rhythms are more like looking at a fuzzy image. The gestalt is not so much ambiguous as amorphous. To be precise, their metrical structure is amorphous while the larger scale grouping structure is clear. One is confronted with a blob of notes followed by silence, a sonic image both blunt and vague.

6 CONCLUSION

The use of the Gini coefficient to quantify metrical ambiguity holds promise as a means to assess the syncopation of short, repeated rhythmic patterns and the perceptual qualities of beat-class set types. While a model of meter induction, based on older models, was introduced here for the purpose of creating input data for the Gini coefficient, the principle of using a dispersion of probabilities over candidate meters to calculate metrical ambiguity is in no way dependent on this model of meter induction. It may be fruitful to apply this principle to data from the more recent models that work with audio in order to use metrical ambiguity as a rhythmic descriptor of audio tracks. As it stands, the Gini coefficient may help analysts discuss rhythmic material in repetitive compositions and possibly serve as a utility for composers and other creators of rhythmically complex music.

7 REFERENCES

[1] R. Cohn, "Transpositional Combination of Beat-Class Sets in Steve Reich's Phase-Shifting Music," *Perspectives of New Music*, vol. 30, no. 2, pp. 146–177, 1992.

[2] D. Eck, "A Positive-Evidence Model for Rhythmical Beat Induction," *Journal of New Music Research*, vol. 30, no. 2, pp. 187–200, 2001.

[3] P. Essens, "Structuring temporal sequences: Comparison of models and factors of complexity," *Perception & Psychophysics*, vol. 57, no. 4, pp. 519–532, 1995.

[4] F. Gomez, A. Melvin, D. Rappaport, and G. Toussaint, "Mathematical Measures of Syncopation," *Proc. BRIDGES: Mathematical Connections in Art, Music and Science*, pp. 73–84, 2005.

[5] S. Handel, *Listening*. MIT Press Cambridge, Mass, 1989.

[6] F. Lerdahl and R. Jackendoff, *A Generative Theory of Tonal Music*. MIT Press, 1996.

[7] D. Locke, *Drum Gahu!: A Systematic Method for an African Percussion Piece*. White Cliffs Media Co., 1987.

[8] H. Longuet-Higgins and C. Lee, "The rhythmic interpretation of monophonic music," *Music Perception*, vol. 1, no. 4, pp. 424–441, 1984.

[9] J. McAuley and P. Semple, "The effect of tempo and musical experience on perceived beat," *Australian Journal of Psychology*, vol. 51, no. 3, pp. 176–187, 1999.

[10] R. Parncutt, "A perceptual model of pulse salience and metrical accent in musical rhythms," *Music Perception*, vol. 11, no. 4, pp. 409–464, 1994.

[11] D. Povel and P. Essens, "Perception of temporal patterns," *Music Perception*, vol. 2, no. 4, pp. 411–440, 1985.

[12] D. Povel and H. Okkerman, "Accents in equitone sequences." *Perception and Psychophysics*, vol. 30, no. 6, pp. 565–72, 1981.

[13] J. Pressing, "Cognitive complexity and the structure of musical patterns," *Proceedings of the 4th Conference of the Australasian Cognitive Science Society, Newcastle, Australia*, 1997.

[14] E. Thul and G. T. Toussaint, "On the relation between rhythm complexity measures and human rhythmic performance," in *C3S2E '08: Proceedings of the 2008 C3S2E conference*. New York, NY, USA: ACM, 2008, pp. 199–204.

[15] G. Toussaint, "Mathematical features for recognizing preference in Sub-Saharan African traditional rhythm timelines," *Proceedings of the 3rd International Conference on Advances in Pattern Recognition*, pp. 18–27, 2005.

[16] ——, "A mathematical analysis of African, Brazilian, and Cuban clave rhythms," *Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science*, pp. 157–168, 2002.

[17] S. Wilson, "Pattern perception and temporality in the music of Steve Reich: An interdisciplinary approach," Ph.D. dissertation, University of New South Wales, 1999.