Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

# The Pairwise Variability Index as a Tool in Musical Rhythm Analysis

Godfried T. Toussaint[*1]

[*]Faculty of Science, New York University Abu Dhabi, United Arab Emirates
[1]gt42@nyu.edu

## ABSTRACT

The normalized pairwise variability index (nPVI) is a measure of the average variation of a set of distances (durations) that are obtained from successive ordered pairs of events. It was originally conceived for measuring the rhythmic differences between languages on the basis of vowel length. More recently, it has also been employed successfully to compare large-scale rhythm in speech and music. London & Jones (2011) suggested that the nPVI could become a useful general tool for musical rhythm analysis. One goal of this study is to determine how well the nPVI models various dimensions of musical rhythmic complexity, ranging from human performance and perceptual complexities, to mathematical measures of metric complexity and rhythm irregularity. A second goal is to determine to what extent the nPVI is capable of discriminating between short, symbolically notated, musical rhythms across meters, genres, and cultures. It is shown that the nPVI suffers from several shortcomings, when it comes to modeling metric complexity and rhythm complexity, in the context of short symbolic rhythmic patterns, such as Sub-Saharan African bell patterns, Arabic rhythms, Rumanian dance rhythms, and Indian *talas*. Nevertheless, comparisons with previous experimental results reveal that the nPVI correlates mildly, but significantly, with human performance complexity. It is also able to discriminate between most of the families of rhythms tested. However, no highly significant differences were found between the nPVI values of binary and ternary musical rhythms, partly mirroring the findings by Patel & Daniele (2003) for language rhythms.

## I. INTRODUCTION

The normalized pairwise variability index (nPVI) is a measure of the average variation of a set of distances (durations) that are obtained from successive adjacent ordered pairs of events. It was originally conceived for measuring the rhythmic differences between languages on the basis of vowel length (Grabe & Low, 2002), and several successful applications in this domain have been realized (Gibbon & Gut, 2001). It has also been applied to the determination of the cognitive complexity of using text-entry systems (Sandness, F. E. & Jian, H.-L., 2004). A review of the history, rationale, and application of the nPVI to the study of languages is given by Nolan & Asu (2009). More recently, the measure has also been employed successfully to compare speech rhythm with rhythm in music (McGowan & Levitt, 2011; London & Jones, 2011; Patel & Daniele, 2003; Huron and Ollen, 2003; Daniele & Patel, 2004). It has been suggested by London & Jones (2011) that the nPVI could become a useful tool for musical rhythm analysis as such. Indeed, it has already been used successfully to compare rhythm in musical scores and their performances (Raju, Asu & Ross, 2010), as well as to distinguish between compositional styles in 19th Century French and German art song (VanHandel, 2006; VanHandel & Song, 2009).

One goal of the present study is to determine how well the nPVI models various dimensions of musical rhythmic complexity, ranging from human performance and perceptual complexities to mathematical measures of metric complexity and rhythm irregularity. A second goal is to determine to what extent the nPVI is capable of discriminating between short, symbolically notated, musical rhythms across meters, genres, and cultures. It is shown that the nPVI suffers from several shortcomings when it comes to modeling metric complexity and rhythm complexity, in the context of short symbolic rhythmic patterns such as Sub-Saharan African bell patterns, Arabic rhythms, Rumanian dance rhythms, and Indian *talas*. Nevertheless, comparisons with previous experimental results reveal that the nPVI correlates mildly, but significantly, with human performance complexity. The index is also able to discriminate between most of the families of rhythms tested. However, no highly significant differences were found between the nPVI values for binary and ternary musical rhythms, partly mirroring the findings by Patel & Daniele (2003) for language rhythms

## II. CHANGE AS A MEASURE OF COMPLEXITY

Nick Chater, (1999, p. 287) suggests that judgments of complexity are akin to judgments of irregularity. Since a rhythm consists of a pattern of inter-onset intervals, one way to measure the irregularity of the rhythm is by measuring the irregularity of the intervals that make up the rhythm. The *standard deviation* is a widely used measure of the dispersion of a random variable, used in statistics and probability. When the random variable is the size of the inter-onset intervals it becomes a measure of irregularity, and has thus been used frequently in speech and language studies. However, musical rhythm and speech are not static; they are processes that unfold in time, and the standard deviation measure "lifts" the intervals out of their original order, disregarding the relationships that exist between adjacent intervals. A better measure of variability across time should incorporate change. Measures of change have been used to characterize the complexity of binary sequences in the visual perceptual domain. Psotka, J., (1975), designed a measure called *syntely* to gage how much the structure of the early portions of a sequence influence the terminal sections, or "the strength of stimulus continuation." Indeed, Aksentijevic & Gibson (2003, 2011), characterized psychological complexity as change: "Structural information is contained in the transition from one symbol (or element) to another and not in the symbols themselves." The "normalized Pairwise Variability Index" is a measure of variability that attempts to capture this notion of change. Grabe and Low (2002) define the "normalized Pairwise Variability Index" (nPVI) for a rhythm with adjacent inter-onset intervals (IOI) as:

$$nPVI = \left(\frac{100}{m-1}\right)\sum_{k=1}^{m-1}\left|\frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2}\right|$$

where $m$ is the number of adjacent vocalic intervals in an utterance, and $d_k$ is the duration of the $k$th interval. Translating this terminology to the musical rhythmic domain converts the vocalic intervals to the adjacent inter-onset intervals (IOI). Notably, Patel (2008, p. 133) laments the use of the term "variability" for the nPVI because, as he rightly points out, it is fundamentally a measure of temporal *contrast* between adjacent durations, rather than variation, and the two measures are not monotonically related. Indeed, consider the two durational patterns A = [2-5-1-3] and B = [1-3-6-2-5]. The rhythm A has a lower standard deviation than B (1.708 versus 2.074), whereas it has a higher nPVI value (106.3 versus 88.1). In spite of anomalies such as these, the results presented below indicate that in many musical contexts the two measures are highly and significantly correlated.

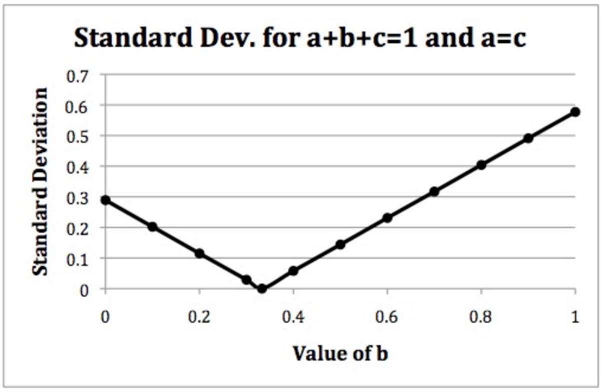


**Figure 1. The standard deviation as a function of *b* for *a*+*b*+*c*=1 and *a*=*c***

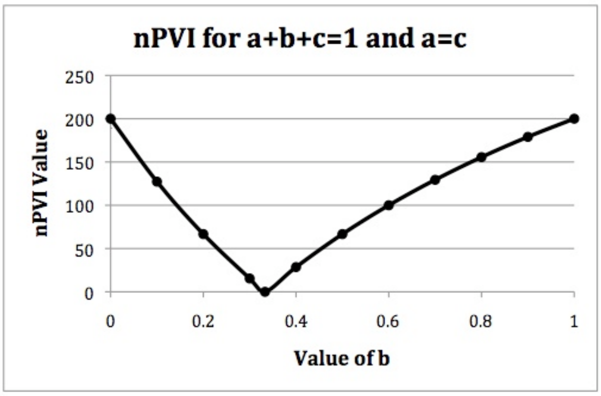


**Figure 2. The nPVI as a function of *b* for *a*+*b*+*c*=1 and *a*=*c***

The two curves in Figures 1 and 2 help to clarify the relationship between the standard deviation and the nPVI. Consider a sequence of three adjacent inter-onset intervals $a$, $b$, $c$ such that $a + b + c = 1$, and $a = c$. The two measures are plotted as a function of $b$, the duration of the middle interval. Both measures take on their minimum value (zero) when the three intervals are equal, and both take on their maximum values when $b$ is either zero or one. Also both functions decrease monotonically from $b$ = 0 to 1/3, and increase monotonically from $b$ = 1/3 to 1. However, whereas the standard deviation varies linearly, the nPVI does not. The nPVI function is slightly convex from $b$ = 0 to 1/3, and slightly concave from $b$ = 1/3 to 1. The only substantial difference between the two curves is in the regions near their maximum values at the extremes of $b$. The nPVI takes on the same value (200) for $b$ = 0 and $b$ = 1, whereas the standard deviation is almost twice as large for $b$ = 1 than for $b$ = 0. Consider two IOI sequences: A = (0.15, 0.7, 0.15) and B = (0.45, 0.1, 0.45). The nPVI regards these two sequences as having almost equal variability: nPVI(A) = 130 and nPVI(B) = 127. However their standard deviations are quite different: SD(A) = 0.32 and SD(B) = 0.20.

# III. METHOD AND DATA

Several sets of rhythms, notated as binary sequences, were collected, including various families of synthetic rhythms (random and systematic), rhythms from India, the African diaspora, the Arab world, and Rumania. Some had been previously evaluated experimentally according to measures of human performance and perceptual complexities. The nPVI values and standard deviations of the IOIs of all rhythms were calculated. These values yielded, for each data set, two orders of the rhythms. These orders were then compared with those produced by the remaining measures, using Spearman rank correlation coefficients.

## A. The Data Sets Used

Several researchers have carried out listening experiments with collections of artificially generated rhythms, in order to test hypotheses about the mental representations of rhythms. The experiments done by Povel & Essens (1995), Shmulevich & Povel (2000), Essens (1995), and Fitch & Rosenfeld (2007), used three data sets of rhythms that provided measures of human perceptual and performance complexities for each rhythm. The three data sets are briefly described in the following. For further details and listings of all the rhythms (in box notation) the reader is referred to the original papers. All the rhythms in these three data sets consist of sixteen unit pulse time spans.

*1) The Povel-Essens Data.* This data comprise 35 rhythms, all of which contain 9 attacks. They all start with an attack on the first pulse, and end with a long interval [x . . .]. The rhythms are made up of all possible permutations of the 9 inter-onset durations {1,1,1,1,1,2,2,3,4}, and do not resemble the rhythmic timelines used in traditional music. Every rhythm in the collection has five intervals of duration 1, two of duration 2, and one of duration 3 and 4, each.

*2) The Essens Data.* This data consists of 24 rhythms, in which the number of onsets varies between 8 and 13, and is generally greater than that of the Povel-Essens rhythms. All the rhythms also start with an attack on their first pulse. Like the rhythms in the Povel-Essens data, these rhythms bear little resemblance to the rhythms used as timelines in musical practice.

*3) The Fitch-Rosenfeld Data.* This data consists of 30 rhythms, in which the number of onsets is smaller than in the rhythms of the other two data sets; six rhythms have four onsets and the rest have five. Also noteworthy is that unlike the other two data sets, 17 rhythms start on a silent pulse. Furthermore, in contrast to the two data sets described above, these rhythms were generated in such a way as to vary the amount of syncopation present in the rhythms, as measured by Fitch and Rosenfeld's implementation of the syncopation measure of Longuet-Higgins & Lee, (1984). These authors did not appear to realize that most of the rhythms generated (or their cyclic rotations) are in fact rhythmic patterns found in Sub-Saharan African and Indian music.

*4) Random Rhythms.* For part of his study of rhythm complexity measures, Thul, E., (2008), generated 50 random 16-pulse rhythms listed in his Table 4.7 on p. 58. The rhythms were obtained by programming a random number generator to simulate "flipping an unbiased coin" sixteen times. Then "heads" was associated with an onset, and "tails" with a silent pulse. The 27 rhythms from this list that started with an onset were chosen for comparison with the other data sets in this study.

*5) North Indian Talas*. In Indian classical music a *tala* (also taal ot tal) is a cyclically recurring clap pattern of fixed length that corresponds somewhat to the concept of meter in Western music or *compás* in the flamenco music of southern Spain. Clayton, M., (2000), provides an in-depth analysis of talas and their role in North Indian classical music. The twelve North Indian talas used here were taken from Example 5.1 on pages 58-59 of Clayton's book. For a comparison of North Indian talas and Sub-Saharan African timelines see Thul, E., & Toussaint, G. T., (2008).

*6) South Indian Talas*. The Carnatic music of South India is also based on rhythmic cycles. One of these systems consists of 35 *sulaadi talas* (Morris, R., 1998). The five *eka* talas are made up of single durations, and therefore the nPVI and standard deviation are not defined for them. The data used in this study consisted of the other 30 talas.

*7) Decitalas.* A thirteenth century Indian manuscript written by Sarngadeva lists 130 talas called *decitalas*, which may be found in the book by R. S. Johnson (1975), in Appendix II, on page 194. See also Morris, R., (1998). Fourteen of these are regular rhythms, which yield nPVI and standard deviation values of zero. Since these entries create ties between the two measures and inflate the Spearman rank correlation coefficients, they were removed, leaving 116 irregular decitalas for the comparisons listed in Table 1. The decitalas have the greatest range of pulses ($3 \leq n \leq 71$) and onsets ($2 \leq k \leq 19$) of all the data sets used in this study. Thus they provide a good data set with which to study how the nPVI varies as a function of $k$ and $n$. Not surprisingly, higher values of $n$ tend to have more onsets. The Spearman rank correlation bears this out: $r = 0.68$ with $p < 0.000001$. The number of onsets relative to the number of pulses, also called the note density (Cerulo, K. A., 1988), may be considered to be a mild contributing factor of rhythm complexity. However, for the decitalas the nPVI and $k$ yield a slight but statistically significant negative correlation: $r = -0.2$ with $p < 0.015$. It appears that an increase in the number of onsets tends to decrease the contrast between successive adjacent IOI's. On the other hand, the nPVI is not correlated with the number of pulses $n$ ($r = 0.08$, $p < 0.18$).

*8) Sub-Saharan African Timelines.* Agawu, K., (2006), p. 1, defines a timeline as a "bell pattern, bell rhythm, guideline, time keeper, topos, and phrasing referent," and characterizes it as a "rhythmic figure of modest duration that is played as an ostinato throughout a given dance composition." Most timelines used in Sub-Saharan African Diaspora music (shortened here to simply African music) employ a cycle of twelve (ternary) or sixteen (binary) pulses. Perhaps the most distinctive of these are the binary and ternary "signature" timelines with durational patterns [3-3-4-2-4] and [2-2-3-2-3], respectively. The fourteen ternary and sixteen binary timelines used in this study were taken from the papers by Rahn, J., (1987), Rahn, J., (1996), Toussaint, G. T., (2005), and the references therein.

*9) Euclidean Rhythms.* One of the oldest and most well known algorithms in the field of computer science, identified in Euclid's Elements (circa 300 B.C.) as Proposition II in Book VII, is known today as the Euclidean algorithm. It was designed to compute the greatest common divisor of two given integers; see Franklin, P. (1956). Donald Knuth (1998) calls this algorithm the "granddaddy of all algorithms, because it is the oldest nontrivial algorithm that has survived to the present day." The idea behind the Euclidean algorithm is remarkably simple: repeatedly replace the larger of the two numbers $k$ and $n$ by their difference, until both are equal. The final number thus obtained is the greatest common divisor of $k$ and $n$. Consider the two numbers $k = 5$ and $n = 8$. Subtracting 5 from 8 yields 3; 5 minus 3 equals 2; 3 minus 2 equals 1; and finally 2 minus 1 equals 1. Therefore the greatest common divisor of 5 and 8 is 1. In Toussaint, G. T., (2005) it was shown that by associating $n$ with the number of pulses in the cycle, and $k$ with the number of onsets (attacks), the Euclidean algorithm may be used to generate most rhythm timelines used in traditional music found all over the world. The key lies in extracting, not the answer to the original problem, but the structure observed in the repeated subtraction process used to obtain the answer. This process is illustrated in Figure 3 with $k = 5$ and $n = 7$. The figure is self-explanatory. First the onsets and silent pulses are ordered one after the other as in (a). Then in the repeated subtraction phase, the silent pulses are moved to the positions as shown in (b). If there are more silent pulses than onsets then the number of silent pulses moved is the same as the number of onsets. This process is continued until there is only one (or zero) remaining columns, as in (c). The columns are then concatenated as in (d) and (e) to obtain the final rhythm in (f). This particular Euclidean rhythm denoted by E(5, 7) = [× · ×× · ××] = (21211) is the *Nawakhat* pattern, a popular Arabic rhythm (Standifer, J. A.,1988). In Nubia it is called the *Al Noht* rhythm (Hagoel, K., 2003). In theoretical computer science Euclidean rhythms, also known as Euclidean strings (Ellis, J., Ruskey, F., Sawada, J., & Simpson, J., 2003), have been discovered independently in different contexts, such as calendar leap year calculations, drawing digital straight lines, and word theory. In music theory they are called *maximally even* sets. Here the terms maximally even and Euclidean are used interchangeably. For a sampling of this literature the reader is referred to Amiot, E., (2007), Clough, J. & J. Douthett, J., (1991), Douthett, J. & Krantz, K., (2007), and Toussaint, G. T., (2005). In the experiments described here the 46 Euclidean rhythms used were taken from Toussaint, G. T., (2005), and the references therein. Since Euclidean rhythms may be generated by such a simple rule, it follows that they have very low Kolmogorov (or information theoretic) complexity. See Chaitin, G. J., (1974), and Lempel, A. & Ziv, J., (1976). Furthermore, by their property of maximally evenly distributed onsets, the rhythms tend to maximize repetitiveness, and minimize contrast. They are therefore expected to have relatively low nPVI values, and thus furnish an extreme data set for comparison with the other data sets. This property also permits the robustness of the nPVI to be tested across widely differing data sets.

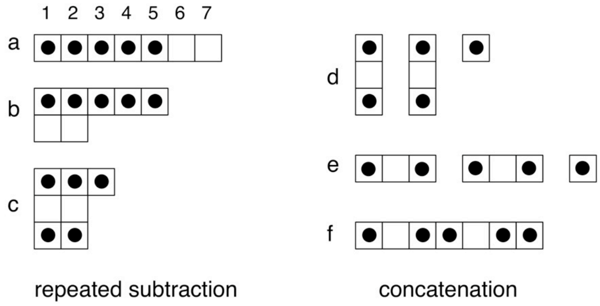


**Figure 3. Generating a Euclidean rhythm with *k* = 5 and *n* = 7**

*10) Rumanian Folk Dance Rhythms.* Proca-Ciortea (1969), investigated over 1,100 Rumanian folk dances, from which fifty-six rhythms were extracted. The rhythms, which are mostly binary, consist mainly of eighth and quarter notes. Six of these rhythms were either regular or exhibited anacrusis, and were deleted, leaving fifty rhythms used in this study.

*11) Arabian Wazn.* Rhythmic patterns in Arabian music are known as *wazn*. They may be compared to Sub-Saharan African timelines in structure and function, although their IOI patterns are quite different. Whereas the African timelines use time spans (measures) that are composed predominantly of twelve and sixteen pulses, the Arabian wazn employ a wide variety of different values. The data used here were composed of nineteen wazn taken from the book by Touma, H. H. (1996). The longest wazn was the samah consisting of nineteen onsets in a time span of thirty-six pulses, with an IOI pattern given by [2-1-1-4-1-1-1-1-2-4-2-1-1-1-1]. Compared to the African timelines, the Arabian wazn are more irregular.

*12) Optimal Golomb Rulers.* A Golomb ruler is a ruler that has "few" marks, and which permits measuring distances only between pairs of these marks. Furthermore, it is desirable to measure as many distinct distances as possible. See Alperin, R. C. & Drobot, V., (2011) for a clear introduction to Golomb rulers. The problem arises in the need for distributing expensive radio telescope elements across a stretch of land so as to better receive signals from outer space. An optimal Golomb ruler with *k* marks is one such that no other shorter Golomb ruler with *k* marks exists. Furthermore, if the Golomb ruler measures all the distances ranging from one to the length of the ruler it is called *perfect*. For example the ruler with marks at points 0, 1, 4, and 6 is a perfect Golomb ruler. The length of the ruler is six, the pairwise distances realized are {1, 2, 3, 4, 5, 6}, and it yields the "rhythm" [x x . . x . ] with a durational IOI pattern [1-3-2]. The twenty shortest optimal Golomb rulers starting with (0, 1, 3) were obtained from Shearer, J. B., (2012). The motivation for using these rulers in the experiments is that, by the nature of their design to have all their pairwise distances distinct, they tend to yield "rhythms" that are highly irregular. (See Freeman, D. K., (1997) for a discussion on this topic.) Therefore Golomb rulers provide an another extreme data set useful for te comparative analysis of the nPVI. The twenty rulers used here had IOIs that varied in number between two and twelve. The ruler with the highest nPVI value of 106.3 in Figure 4 was the fourth in the list, with marks at (0, 2, 7, 8, 11) yielding IOIs [2-5-1-3].

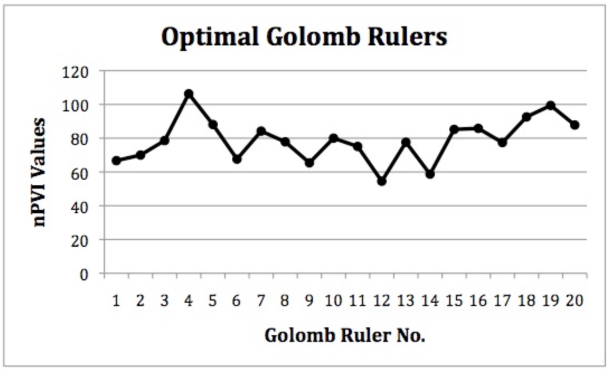


**Figure 4. The nPVI for the 20 shortest optimal Golomb rulers**

*13) Mathematical Measures of Complexity*. In the study of speech rhythm, a measure of variability that has been often used is the classical statistical measure of the standard deviation of the vocalic or consonantal interval durations. See for example the papers by Ramus, F., Nespor, M., & Mehler, J. (1999), and Yoon, T.-J., (2010). Since the nPVI was originally proposed as a means of avoiding the drawbacks of the standard deviation, it was decided to compare the two measures to determine the extent of their differences. The nPVI was also compared with Keith's measure of metrical complexity. In the musical domain Michael Keith (1991), proposed a measure of meter complexity, based on a hierarchical partition of a meter into sub-meters, and on the frequency of alternations between binary and ternary units within different levels of this hierarchy. In Table 5.4 of Chapter 5 of his book, he lists seventy-six metrical patterns with the number of pulses as high as twenty, along with their complexity values.

For a string of symbols *S*, Keith's measure of meter complexity, denoted by *C*(*S*), is defined for meters consisting of durational patterns made up of strings of 2's and 3's, such as the African signature pattern, the *fume-fume* [2-2-3-2-3] and the *guajira* flamenco compás [3-3-2-2-2]. Keith first partitions *S* into any string of disjoint *subunits*. At the lowest level these two rhythms are partitioned into the subunits [2][2][3][2][3] and [3][3][2][2][2], respectively. At this level, the complexity of an individual [2]-unit is 2, and that of a [3]-unit is 3. For a given partition the complexity of the string *C*(*S*) is the sum of the complexities of the subunits. Thus at this level the complexities of the fume-fume and guajira are the same: 2+2+3+2+3 = 3+3+2+2+2 = 12. Keith defines a *unit* as one or more *identical contiguous* subunits. Thus another possible pair of partitions for these two rhythms consists of [2-2][3][2][3] and [3-3][2-2-2], respectively. The complexity value of a unit $U$ consisting of a number of identical subunits $H$ is defined as $C(U) = max\{\#\text{subunits}, C(\text{H})\}$, where #subunits denotes the number of subunits. For example, the complexity of the unit [2-2] is $max\{2, 2\} = 2$, and that of the unit [2-2-2] is $max\{3, 2\} = 3$. If we denote a partition of $S$ by $S_U$, then the complexity of a given partition of $S$ into units, denoted by $C(S_U)$, is the sum of the complexities of the units, $C(U)$. Therefore for this partition $C(S_U)$(*fume-fume*) = 2+3+2+2 = 9, and $C(S_U)$(*guajira*) = 3+3 = 6. Finally, the complexity of the sequence $C(S)$ is the minimum complexity, minimized over all possible partitions; $C(S) = min\{C(S_U)\}$. In our example the guajira admits several other partitions such as [3-3][2-2][2], [3-3][2][2-2], [3][3][2-2][2], and [3][3][2][2-2], the complexities of which are, respectively, 7, 7, 10, and 10. Therefore the final complexities are $C$(*fume-fume*) = 9, and $C$(*guajira*) = 6.

A scatter-plot of the nPVI values as a function of Keith's complexity measure, for the seventy-six metric patterns taken from Michael Keith's book is shown in Figure 5. Although the two measures are highly and significantly correlated ($r$ = 0.65 $p$ < 0.01), the relationship between the two measures is dominated by vertical and horizontal portions. Most noticeable is that for a complexity value of 5, there are numerous rhythms with nPVI values ranging from zero to 40.

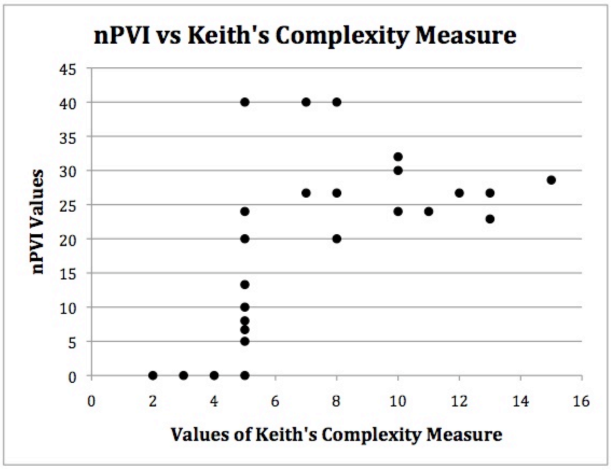


**Figure 5. nPVI as a function of Keith's metric complexity**

## IV. RESULTS

The Spearman rank correlation coefficients and statistical significance values obtained by comparing the various complexity measures with the nPVI for all the data sets are listed in Table 1. Those with statistically significant $p$-values are highlighted in bold.

A comparison of the nPVI with the human performance and perceptual complexities obtained with the three data sets of Povel-Essens, Essens, and Fitch-Rosenfeld, yields curious mixed results. The performance and perceptual complexity judgments in the Povel-Essens and Essens data sets do not correlate at all with the nPVI values. Furthermore, in the case of the Povel-Essens data, the nPVI cannot even be compared with the standard deviation because the standard deviation of the IOI's is the same for all the rhythms, due to the fact that they consist of permutations of the intervals in the set {1,1,1,1,1,2,2,3,4}. Thus the Spearman rank correlation coefficient is not computable for this type of data. The Fitch-Rosenfeld rhythms tell a different story however. Here the human performance complexity is mildly but significantly correlated with the nPVI ($r$ = 0.40 with $p$ < 0.02), as is the standard deviation ($r$ = 0.57 with $p$ < 0.01). For the Essens rhythms the standard deviation is also highly and significantly correlated with the nPVI ($r$ = 0.67 with $p$ < 0.01).

**Table 1. Spearman rank correlations between the nPVI values of various complexity measures**

| Complexity Measure | nPVI |
|---|---|
| Performance Complexity (Povel-Essens) | $r = -0.006$ $p < 0.5$ |
| Perceptual Complexity (Povel-Essens) | $r = 0.100$ $p < 0.27$ |
| Standard Deviation (Povel-Essens) | Stan. Dev. = 0 |
| Performance Complexity (Essens) | $r = 0.047$ $p < 0.42$ |
| Perceptual Complexity (Essens) | $r = 0.006$ $p < 0.49$ |
| Standard Deviation (Essens) | **$r = 0.67$ $p < 0.01$** |
| Performance Complexity (Fitch-Rosenfeld) | **$r = 0.40$ $p < 0.02$** |
| Standard Deviation (Fitch-Rosenfeld) | **$r = 0.57$ $p < 0.01$** |
| Keith's Complexity Measure - C(S) | **$r = 0.65$ $p < 0.01$** |
| Stan. Dev. (12 North Indian Talas) | **$r = 0.73$ $p < 0.01$** |
| Stan. Dev. (30 South Indian Talas) | **$r = 0.67$ $p < 0.01$** |
| Stan. Dev. (116 Irregular Decitalas) | **$r = 0.77$ $p < 0.01$** |
| Stan. Dev. (14 African Timelines; $k$=5, $n$=12) | **$r = 0.80$ $p < 0.01$** |
| Stan. Dev. (16 African Timelines; $k$=5, $n$=16) | **$r = 0.86$ $p < 0.01$** |
| Stan. Dev. (46 Euclidean Rhythms) | **$r = 0.55$ $p < 0.01$** |
| Stan. Dev. (28 Random Rhythms; $n$=16) | **$r = 0.55$ $p < 0.01$** |
| Stan. Dev. (50 Rumanian Folk Rhythms) | **$r = 0.35$ $p < 0.01$** |
| Stan. Dev. (19 Arabian *Wazn*) | $r = 0.31$ $p < 0.1$ |
| Stan. Dev. (20 Golomb Rulers) | $r = 0.31$ $p < 0.1$ |

Patel and Daniele (2003) compared the rhythm of English and French music and language using the nPVI of the lengths of the notes in music, and the vocalic durations in speech, respectively. In both cases they found that the nPVI values for British English were greater than for French. To test the influence of musical meter on the nPVI values of note durations, to account for the differences observed, they separated their musical themes into those with binary and ternary meters, and found that the nPVI values of the binary and ternary music corpora did not differ significantly. Their corpora consisted of 137 English musical themes from composers such as Elgar, Delius, and Holst, and 181 French musical themes from composers that included Debussy, Honegger, and Ravel. These results motivated the exploration of whether there is any significant difference between the nPVI values in a completely different musical context: African rhythm instead of classical European music, and note durations replaced by inter-onset intervals. For this purpose two groups of rhythms were compiled from a collection of books and journals, consisting of 34 binary sixteen-pulse timelines, and 39 ternary twelve-pulse timelines. The nPVI values are shown in Figure 6, where the error bars indicate one standard deviation above and below the mean. The most notable aspect is that the two means differ little in absolute terms compared to the variability within each group ($t$ = 1.542, with $p$ = 0.0636, one-sided test) mirroring somewhat the findings by Patel & Daniele (2003). However, the Kolmogorv-Smirnoff test rejects the null hypothesis that the two distributions of nPVI values are the same ($D = 0.3385$ with $p < 0.02$).

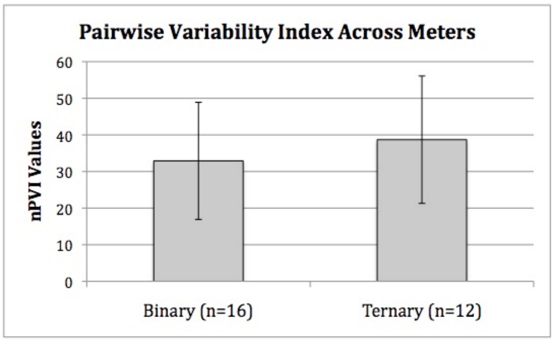


**Figure 6. The nPVI values for the binary and ternary African diaspora rhythm timelines**

As mentioned in the introduction, it has been shown in previous research that the nPVI could be used to discriminate among different compositional styles in 19th Century French and German art song (VanHandel, 2006; VanHandel & Song, 2009). These results motivated the testing of whether the nPVI is able to discriminate between rhythms of different styles, as well as cultures. To this end several corpora of rhythms were compiled: Euclidean rhythms, Arabian rhythms, Rumanian rhythms, African timelines, and Indian talas. The average values of the nPVI scores for six of these corpora are shown in the graph in Figure 7, in increasing order, along with error bars indicating plus and minus one standard deviation. Again, the variation within each group is much greater than the differences between the means. Nevertheless, the general trend is that Euclidean rhythms appear to be the simplest, and North Indian talas the most complex.

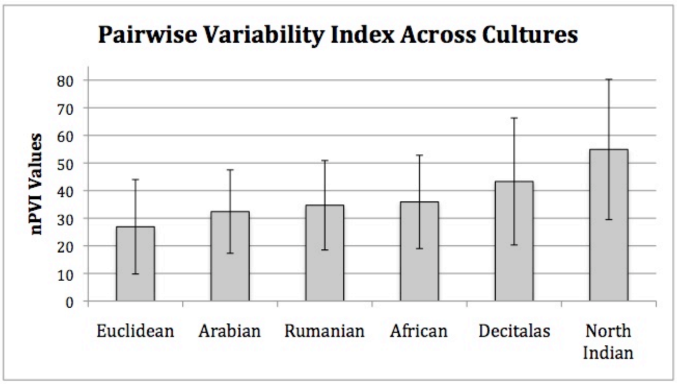


**Figure 7. The nPVI averages for rhythms from different genres**

To obtain some insight into the sensitivity of the variability of the nPVI for different genres of rhythms, the nPVI was also computed for two sets of random rhythms that were generated as described in *Subsection 4).* The nPVI values for the binary 16-pulse and ternary 12-pulse rhythms are shown in Figure 8, alongside the value for the Euclidean rhythms. As might be expected, the nPVI values for the random rhythms are much higher than for the Euclidean (maximally even) rhythms. A Kolmogorov-Smirnov test of the difference between the two distributions yields a distance of $D = 0.65$ with $p < 0.001$, confirming the intuition that the maximally even rhythms are highly non-random. However, the variability for the 16-pulse random rhythms is not greater than that for any of the other families of rhythms tested. Indeed, from Figure 7 it may be observed that the North Indian rhythms and Decitalas have the highest variability of all. Does this mean that the Indian talas are more random than the rhythms of other genres? On the contrary, one might expect that variability is greater in specifically designed rhythms than in random rhythms. To test this conceivable hypothesis Kolmogorov-Smirnov tests were performed with the nPVI values of the random rhythms (12-pulse and 16-pulse rhythms combined) and the three types Indian talas separately: North Indian talas, South Indian Sulaadi talas, and Decitalas. The results are listed in Table 2. For the South Indian and Decitalas we may reject the hypothesis that they come from the same distribution as the random rhythms. On the other hand, for the North Indian talas this is not the case, although this may be due to the small sample size of the North Indian talas. Comparing the nPVI values of the three systems of Indian talas with each other using Kolmogorov-Smirnov tests suggests that the Decitalas are significantly different from the North Indian talas ($D = 0.389$ with $p < 0.05$), as well as the South Indian talas ($D = 0.378$ with $p < 0.001$), but the North Indian talas are not significantly different from the South Indian talas ($D = 0.316$ with $p < 0.3$).

**Table 2. Kolmogorov-Smirnov distances and significance tests for comparison of the nPVI values of the Indian talas with random rhythms**

| Kolmogorov-Smirnov Tests Against Random Rhythms | | |
|---|---|---|
| **North Indian Talas** | **South Indian Talas** | **Decitalas** |
| $D = 0.328$ | $D = 0.327$ | $D = 0.364$ |
| $p < 0.21$ | $p < 0.03$ | $p < 0.001$ |

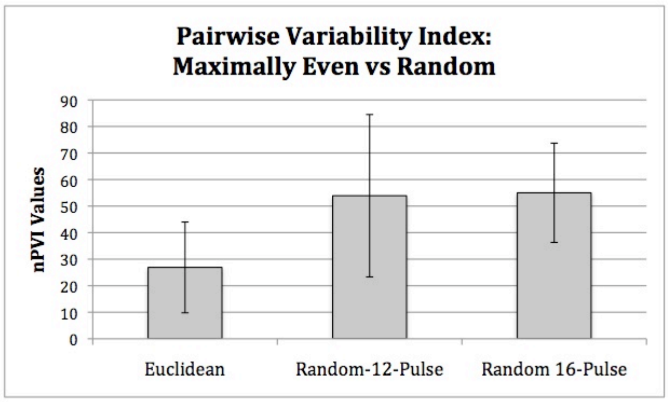


**Figure 8. Comparing maximally even and random rhythms**

From Figure 7 it may be observed that the Arabian, Rumanian, and African rhythms have almost equal average nPVI values, as well as similar degrees of variation. Furthermore all three genres of rhythms have similar short IOI's with a generous supply of IOI's made up of durations of two and three pulses. Therefore these data sets provide strict tests of the power of the nPVI to discriminate between these genres of rhythms coming from different cultures. Table 3 shows the Kolmogorov-Smirnov distances and significance tests of the nPVI values obtained across cultures for the six families of rhythms of Figure 7. The results with statistically significant *p*-values (less than 0.05) are highlighted in bold. Of the fifteen pairwise comparisons, ten are statistically significant. The nPVI can easily distinguish between the North Indian talas and all the remaining data sets. The Euclidean rhythms are significantly different from all other rhythms except the Arabian rhythms. The African timelines and Indian Decitalas are distinguishable from the Arabian rhythms at the 0.06 level. The rhythm families hardest to distinguish are the Arabian from the Rumanian and African rhythms as well as the Rumanian rhythms from the African rhythms.

**Table 3. Kolmogorov-Smirnov distances and significance tests for comparison of the nPVI values across cultures**

| Kolmogorov-Smirnov Tests Across Cultures | | | | | |
|---|---|---|---|---|---|
| | Arabian | Rumanian | African | Decitalas | North Indian |
| Euclidean | $D = 0.23$ $p < 0.40$ | **$D = 0.28$ $p < 0.04$** | **$D = 0.35$ $p < 0.01$** | **$D = 0.38$ $p < 0.01$** | **$D = 0.59$ $p < 0.01$** |
| Arabian | - | $D = 0.14$ $p < 0.94$ | $D = 0.19$ $p < 0.60$ | $D = 0.31$ $p < 0.06$ | **$D = 0.61$ $p < 0.01$** |
| Rumanian | | - | $D = 0.13$ $p < 0.66$ | **$D = 0.24$ $p < 0.03$** | **$D = 0.56$ $p < 0.01$** |
| African | | | - | **$D = 0.2$ $p < 0.05$** | **$D = 0.52$ $p < 0.01$** |
| Decitalas | | | | - | **$D = 0.39$ $p < 0.05$** |

# V. CONCLUSION

The goal of the research described here was to explore the efficacy of the nPVI as a tool for the analysis of the complexity of short musical rhythms. The main results of this study follow two general trends that depend on the types of data used. The first type of data comprised mainly human judgments of perceptual and performance complexities obtained from listening tests performed with rhythms that were artificially generated in the laboratory. The second kind of data consisted of mathematical measures of complexity (nPVI, standard deviation, and Keith's metric complexity) that were computed on rhythms from traditional musical practices of several different cultures. On the first kind of data the nPVI measure performed poorly, with one exception. On the second kind it fared much better.

The correlation results in Table 1 exhibit a pattern of nPVI values that depends on the type of rhythm to which the nPVI is applied. The correlations between the complexity measures and the nPVI values tend to be, either non-existent, low or statistically insignificant, for extreme rhythms, or for those artificially generated by means of combinatorial methods, whereas the correlations are high and statistically significant for rhythms that are used in practice in traditional music. The standard deviation of the IOI's correlates most highly for the Sub-Saharan African timelines: $r = 0.80$ with $p < 0.01$ for the ternary timelines, and $r = 0.86$ with $p < 0.01$ for the binary timelines. Thus it may function also as a measure of contrast. These results indicate that the nPVI may be a promising and powerful tool in certain contexts, although the precise nature of these contexts has yet to be determined. The results also indicate that the nPVI, at least for some rhythms, is not too different from the standard deviation of their IOI's.

One observation that is common to all the results obtained here, is that the nPVI measure has a high variance. From Figures 6, 7, and 8, it is evident that the standard deviations of the nPVI values for all the data are sensitive to outliers. This behavior has been previously noted in the research on language (Wiget, L., et al., 2010). To combat this sensitivity Jian, H.-L., (2004), proposed a modification of the original nPVI formula, that incorporates the *median* rather than the mean of the adjacent interval differences. It has yet to be determined whether this variant of the nPVI would improve the results obtained here.

It is an obvious fact that the order of the duration intervals of a rhythm influences the rhythm's perceived complexity, by creating contrast between intervals and their adjacent intervals, as well as between intervals and the underlying meter. However, the standard deviation of the IOI's is by definition blind to this order. Not surprisingly it is not difficult to create examples that show that the standard deviation fails to completely characterize the complexity of rhythm timelines. Consider the well-known Cuban clave son and Brazilian clave bossa-nova rhythms with IOI durational patterns [3-3-4-2-4] and [3-3-4-3-3], respectively (Toussaint, G. T., 2002). The bossa-nova is more complex (and syncopated) than the son, but the standard deviation of the former is 0.837, whereas for the latter it is only 0.447. Although the nPVI takes order into account and is thus contrast-sensitive, it is still oblivious to the underlying meter. Hence the nPVI fares no better, yielding a value of 40.47 for the son and 7.14 for the bossa-nova.

Barry, W., Andreeva, B., & Koreman, J., (2009), expose additional limitations of the nPVI in its ability to capture perceived rhythm in the Bulgarian, English, and German languages. These results support the thesis of Arvaniti, A., (2009), who argues that metrics such as the standard deviation as well as the nPVI are unreliable predictors of rhythmic types in languages. Nolan & Asu, (2009) conclude from their study that in language, duration cannot be "assumed to be either the exclusive correlate of perceived rhythm nor to act independently of other cues in perception," and according to Royer, F. L. & Garner, W. R., (1970), "pattern organizations are wholistic." The present study suggests similar conclusions with respect to short rhythms in the musical domain. Nevertheless, as the results in Table 3 attest, the nPVI is successful at discriminating between most of the pairs of rhythm families tested.

To be more generally useful for both theoretical and practical musical rhythm analysis, a suitable modification of the nPVI that takes metrical information into account, is probably necessary. What the exact nature of such a modification should entail is as yet an open question. However, investigations along these lines have already started. London, J. & Jones, K., (2011), propose several modifications of the nPVI for its specific application to musical rhythm. These refinements include the application the nPVI hierarchically to higher levels of rhythmic structure, analysing binary and ternary (duple and triple) rhythms separately, and using alternative codings of the IOI durations, such as rounding durations to the nearest beat. In addition to incorporating meter, refinements of the nPVI that take grouping into account provide further alternatives.

# ACKNOWLEDGMENT

This research was supported by a research grant from the Provost's Office of New York University Abu Dhabi, in Abu Dhabi, United Arab Emirates, administered through the Faculty of Science.

# REFERENCES


Agawu, K. (2006). Structural Analysis or Cultural Analysis? Competing Perspectives on the "Standard Pattern" of West African Rhythm. *Journal of the American Musicological Society*, *59*(1), 1–46.

Aksentijevic, A. & Gibson, K. (2011). Complexity Equals Change. *Cognitive Systems Research,* doi:10.1016/j.cogsys.2011.01.002.

Aksentijevic, A. & Gibson, K. (2003). Towards a Universal Measure of Structural Complexity. *Proceedings of the 8th International Conference of Cognitive Science*, Univ. of New South Wales, Sydney, Vol 1, pp 12-18.

Alperin, R. C., & Drobot, V. (2011). Golomb Rulers. *Mathematics Magazi*ne, *84*, 48-55.

Amiot, E. (2007). David Lewin and Maximally Even Sets. *Journal of Mathematics and Music*, *1*(3), 157-172.

Arvaniti, A. (2009). Rhythm, Timing and the Timing of Rhythm. *Phonetica*, *66*, 46-63.

Barry, W., Andreeva, B., & Koreman, J. (2009). Do Rhythm Measures Reflect Perceived Rhythm? *Phonetica*, *66*, 78-94.

Cerulo, K. A. (1988). Analyzing cultural products: A new method of measurement, *Social Science Research*, *17*, 317-352.

Chaitin, G. J. (1974). Information Theoretic Computational Complexity. *IEEE Transactions on Information Theory*, *20*, 10–15

Chater, N. (1999). The Search for Simplicity: A Fundamental Cognitive Principle? *The Quarterly Journal of Experimental Psychology, 52A* (2), 273-302.

Clayton, M. (2000). *Time in Indian Music*, Oxford University Press.

Clough, J. & J. Douthett, J. (1991). Maximally Even Sets. *Journal of Music Theory*, *35*, 93–173.

Daniele, J. R. & Patel, A. D. (2004). The Interplay of Linguistic and Historical Influences on Musical Rhythm in Different Cultures. *Proceedings of the 8th International Conference on Music Perception and Cognition*, Evanston, Illinois, (August): 759-762.

Douthett, J. & Krantz, K. (2007). Maximally Even Sets and Configurations: Common Threads in Mathematics, Physics, and Music. *Journal of Combinatorial Optimization*, *14*(4), 385-410.

Ellis, J., Ruskey, F., Sawada, J., & Simpson, J. (2003). Euclidean Strings. *Theoretical Computer Science*, *30*(1), 321–340.

Essens, P., (1995). Structuring Temporal Sequences: Comparison of Models and Factors of Complexity. *Perception and Psychophysics*, *57*(4), 519–532.

Fitch, W. T. & Rosenfeld, A. J., (2007). Perception and Production of Syncopated Rhythms. *Music Perception*, *25*(1), 43–58.

Franklin, P. (1956). The Euclidean Algorithm. *The American Mathematical Monthly*, *63*(9), 663–664.

Freeman, D. K. (1997). Maximizing Irregularity and the Golomb Ruler Problem. Internal Report. C.N.R.S. Luminy, Centre de Physique Theorique, Marseille, France.

Gibbon, D. & Gut, U. (2001). Measuring Speech Rhythm. *Eurospeech* 2001 *Scandinavia*: *7th European Conference on Speech Communication and Technology*, Aalborg, Denmark (September).

Grabe, E., & Low, E. (2002). Durational Variability in Speech and the Rhythm Class Hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Papers in Laboratory Phonology 7.* Cambridge: Cambridge University Press.

Hagoel, K. (2003). *The Art of Middle Eastern Rhythm*. OR-TAV Music Publications, Kfar Sava, Israel

Huron, D. and Ollen, J. (2003). Agogic Contrast in French and English Themes: Further Support for Patel and Daniele (2003). *Music Perception*, *21*, 267-272.

Jian, H.-L. (2004). An Improved Pair-Wise Variability Index for Comparing the Timing Characteristics of Speech. *Proceedings of the 8th International Conference on Spoken Language Processing*, Jeju Island, Korea, October 4-8.

Keith, M. (1991). *From Polychords to Pólya: Adventures in Musical Combinatorics*. Princeton: Vinculum Press.
London, J. & Jones, K. (2011). Rhythmic Refinements to the nPVI Measure: A Reanalysis of Patel and Daniele. *Music Perception: An Interdisciplinary Journal, 29*(1), 115-120.

Knuth, D. E. (1998). *The Art of Computer Programming*, Vol. 2, 3rd edition. Addison Wesley, Reading, Massachussets.

Lempel, A. & Ziv, J. (1976). On the Complexity of Finite Sequences. *IEEE Transactions on Information Theory*, *22*(1), 75–81.

London, J. & Jones, K. (2011). Rhythmic Refinements to the nPVI Measure: A Reanalysis of Patel & Daniele (2003a), *Music Perception: An Interdisciplinary Journal*, *29*(1), 115-120.

Longuet-Higgins, H. C. and Lee, C. (1984). The Rhythmic Interpretation of Monophonic Music. *Music Perception*. *1*, 424–441.

McGowan, R. W. & Levitt, A. G. (2011). A Comparison of Rhythm in English Dialects and Music. *Music Perception: An Interdisciplinary Journal*. *28*(3), 307-314.

Morris, R. (1998). Sets, Scales, and Rhythmic Cycles: A Classification of Talas in Indian Music. In *21st Annual Meeting of the Society for Music Theory*, Chapel Hill, NC.

Nolan, F. & Asu, E. L. (2009). The Pairwise Variability Index and Coexisting Rhythms in Language. *Phonetica*, *66*, 64-67.

Patel, A. D. (2008). Music, Language, and the Brain. Oxford University Press, New Yorrk.

Patel, A. D. & Daniele, J. R. (2003). An Empirical Comparison of Rhythm in Language and Music. *Cognition*, *87*, B35-B45.

Patel, A. D. & Daniele, J. R. (2003). Stress-Timed vs. Syllable-Timed Music? A Comment on Huron and Ollen (2003). *Music Perception: An Interdisciplinary Journal*, *21*(2), 273- 276.

Povel, D. & Essens, P., (1985). Perception of Temporal Patterns. *Music Perception*, *2*, 411– 440.

Proca-Ciortea, V. (1969). On Rhythm in Rumanian Folk Dance. *Yearbook of the International Folk Music Counci*l, *1*, 176-199.

Psotka, J. (1975). Simplicity, Symmetry, and Syntely: Stimulus Measures of Binary Pattern Structure. *Memory & Cognition*, *3*(4), 343-444.

Rahn, J. (1987). Asymmetrical Ostinatos in Sub-Saharan Music: Time, Pitch, and Cycles Reconsidered. *In Theory Only: Journal of the Michigan Music Theory Society, 9*(7), 23– 37.

Rahn, J. (1996). Turning the Analysis Around: African-Derived Rhythms and Europe-Derived Music Theory. *Black Music Research Journal*, *16*(1), 71–89.

Raju, M., Asu, E. L. & Ross, J. (2010). Comparison of Rhythm in Musical Scores and Performances as Measured with the Pairwise Variability Index. *Musicae Scientiae*, *14*(1), 51-71.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of Linguistic Rhythm in the Speech Signal. *Cognition 73*, 265–292.

Royer, F. L. & Garner, W. R. (1970). Perceptual Organization of Nine-Element Auditory Temporal Patterns. *Perception & Psychophysics*, *7*(2), 115-120.

Sandness, F. E. & Jian, H.-L. (2004). Pair-Wise Variability Index: Evaluating the Cognitive Difficulty of Using Mobile Text Entry Systems. In *Mobile Human-Computer Interaction: Proceedings of the 6th International Symposium*, Brewster, S. A. & Dunlop, M. D. (Eds.), Glasgow, UK, September 13-16,pp. 347-350.

Shearer, J. B. (2012). Golomb Rulers. http://www.research.ibm.com/people/s/shearer/gropt.html

Shmulevich, I. & Povel, D.-J., (2000). Measures of Temporal Pattern Complexity. *Journal of New Music Research*, *29*(1), 61–69.

Standifer, J. A. (1988). The Tuareg: Their Music and Dances. *The Black Perspective in Music*, *16*(1), 45–62.

Thul, E. (2008). Measuring the Complexity of Musical Rhythm. M.Sc. Thesis, School of Computer Science, McGill University, Montreal, Canada.

Thul, E., & Toussaint, G. T. (2008). A Comparative Phylogenetic Analysis of African Timelines and North Indian Talas. Proceedings of *BRIDGES: Mathematics, Music, Art, Architecture, and Culture*.

Touma, H. H. (1996). *The Music of the Arabs*. Portland, Oregon: Amadeus Press.

Toussaint, G. T. (2005). The Euclidean Algorithm Generates Traditional Musical Rhythms. *Proceedings of BRIDGES: Mathematical Connections in Art, Music, and Science*, Banff, Alberta, Canada, pp. 47-56.

Toussaint, G. T. (2002). A mathematical analysis of African, Brazilian, and Cuban clave rhythms. *Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science,* Townson University, Towson, MD, July 27-29, pp. 157-168

VanHandel, L. & Song, T. (2009). Influence of Linguistic Rhythm on Individual Compositional Style in 19th Century French and German Art Song, *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music*, Jyväskylä, Finland: 553-5557.

VanHandel, L. (2006). Trends in/over Time: Rhythm in Speech and Musical Melody in 19th-Century Art Song. *Proceedings of the 9th International Conference on Music Perception and Cognition*, University of Bologna, (August): 1153-1158.

Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Sven L. Mattys, S. L. (2010). How Stable are Acoustic Metrics of Contrastive Speech Rhythm? *Journal of the Acoustical Society of America*, 127/3:1559-1569.

Yoon, T.-J. (2010). Capturing Inter-speaker Invariance Using Statistical Measures of Rhythm. *Proceedings of the Speech Prosody Conference*, Chicago, 100201: 1-4.