An Alternative Postulate to see Melody as “Language”

Hokky Situngkir
<hs@compsoc.bandungfe.net>
Dept. Computational Sociology
Bandung Fe Institute

May 25th 2007

Abstract

The paper proposes a way to see melodic features in music/songs in the terms of “letters” constituting “words”, while in return investigating the fulfillment of Zipf-Mandelbrot Law in them. Some interesting findings are reported including some possible conjectures for classification of melodic and musical artifacts considering several aspects of culture. The paper ends with some discussions related to further directions, be it enrichment in musicology and the possible plan for musical generative art.

Keywords: melody in music, Zipf-Mandelbrot Law, musical categories, generative art.

Music charms us,
although its beauty consists only in the agreement of numbers and in the counting.
Gottfried Wilhelm Leibniz (1646 - 1716)

"Is there a meaning to music?"
My answer would be, "Yes."
And "Can you state in so many words what the meaning is?"
My answer to that would be "No."
Aaron Copland (2002)

1. Background Insights

To some artists, music is a kind of language although it 'paradoxically' cannot be translated literally into our daily words (*cf.* Copland, 2002). Yet, music can be seen as one media plays role in communications and messaging among human being. Musical artifacts are often related to social, pragmatic, and emotional constitution of linguistic structures emerging in discourse, performance, textuality, and poetics (*cf.* Feld & Aaron, 1994). Anthropologically, music brings the systematic and complex information related to culture as an artifact arouse in civilizations or ethnicities while in contrast, psychologically music plays a major role in perceptions, emotional states, etc.

Roughly speaking, a composer or performer may communicate with her listener through musical gestures or by lyrics. Here we recognize that music is mostly a kind of harmonic mixtures between melodic harmony brought by different audio frequencies presented along with the lyrical message expressed in natural language. Even further, complexities of musical artifacts are also raised by the way they are presented, be it live performance, radio broadcastings, video clips in audio-visual media, or possibly from digital files in the form of compact discs, mp3s, iPod, or simply just a slight cut of polyphonic melodies from cellular phone ringtones.

Despite the poetic beauty may arouse from the sound, the rhythmic and harmonious melodies as the core of music have always been interesting thing for deep scrutinizing. Why certain songs may represent some ethnicities and or cultures, why our preferences to music sometimes bounded by certain genres of music, and why popular music is always distinguishable from time to time, are few from many inquiries regarding music.

This fact invigorates the inquiries reported in the paper. The work of Voss & Clarke (1975) has shown an interesting aspect to begin with as they showed that the pitch and loudness of audio resulted from musical arts and speeches follows the Zipf's distribution of *1/f*. This was followed by further enhancing the analytical tools used in the analysis of Su & Wu (2006) showing the multifractality in musical arts. Furthermore, the work of Manaris *et.*

al. (2003) reflecting further work presenting endeavor to find the concept of pleasantness of music while practically showing the possible generation of music by using the Zipf-Mandelbrot law.

The Zipf and Zipf-Mandelbrot law are linguistic statistical properties of texts while the latter is a mathematical improvement to the previous one (Situngkir, 2007). The paper intends to present the persistence of the Zipf-Mandelbrot law in musical art but it is not the physical factors *e.g.:* amplitudes, frequencies, of any recordings. Our analytical approach is laid upon our understanding that as music can be represented in certain notations, they can be regarded as text/corpus, as well as emerging the Zipf-Mandelbrot law.

However, we recognize that there has been some previous report presenting the Zipf-Mandelbrot law in musical structures and notations *e.g.:* Manaris, *et. al.* (2001) tried to see the possible classifications of musical arts by observing the distinctive exhibited fitting variables of musical metrics from various musical genres: baroque, 20th century, blues, and jazz music. They proposed some metrics including the product of duration and pitch, melodic intervals, and melodic intervals.

Regarding to this, the short paper presented here attends a slightly different proposal. The paper begins with discussions on how to see musical art as kind of text/corpus along with introducing a metrics that would be available for further statistical approaches. The paper ends with a bridge to further direction on an alternative steps in musical generative art.

2. Reading Musical Works

Can we see music as a kind of “language”? First of first, the work on cognitive science has nonetheless, given sign that there should be the grammar of music even though it would be hard to be followed by Chomskian linguists while it is also not impossible to. Some work has in fact, been conducted regarding to incorporating structural and grammatical pattern in music or in roughly speaking, serious and careful standard linguistic-style observation and approach to some musical genres (see for instance Steedman, 1996).

As it has been discussed in the introductory section of the paper, music is somewhat a complex artifact that a lot of properties comprised it. Some aspects or dimensions (Levitin, 2006: 19-53) can measure most music, *i.e.*:

- ✓ *Tone/Note*: the discrete musical sound. The tone is the sound that we hear and the note is what we see on a musical score.
- ✓ *Pitch*: the psychological construct that is relating both the actual frequency of a particular tone and to its relative position in the musical scale.
- ✓ *Rhythm*: the duration of a series of notes.

- ✓ *Tempo*: the overall speed of pace of a musical piece.
- ✓ *Contour*: the overall shape of a melody.
- ✓ *Timbre*: the color of the sound – a different physical vibration of different musical instruments or human's vocal cords yielded the distinguishable timbre.
- ✓ *Loudness*: the psychological construct that relates to the physical amplitude of a tone.
- ✓ *Spatial Location*: where the sound is coming from.
- ✓ *Reverberation*: the perception of how distant the source is from the hearer in combination with how large a room or a hall the music is being played.

Those are the dimensions frequently used to distinct one musical sound to another. However, variety may be introduced by contrasts in musical elements such as *melody*, the prominent musical sequences within a piece; *rhythm*, the durations of patterns and notes that create forward movement; *harmony*, the combination of notes sounded at the same time; and *key*, the tonality produced by seven tones in a recognizable relationship to a central tone. Repetition of these elements may make the music more unified or coherent, as with the return of a melodic or rhythmic pattern heard earlier in the piece.

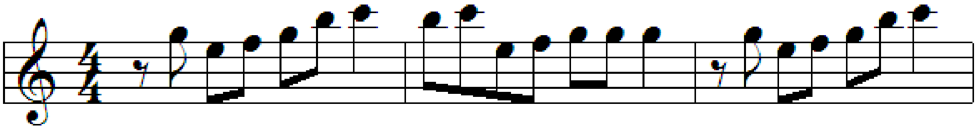
However, the rest of the paper would approach melody as the sole observed element of music. Melody is the most important for it enables us to distinguish one work from another, and it is melody that human beings are innately able to reproduce by singing, humming, and whistling. Furthermore, melody makes music memorable to us (Selfridge-Field, 1998).

We use MIDI (musical instrument digital interface), the serial interface standard that allows the connection of music synthesizers, musical instruments, and computers (Heckroth, 1995). By using MIDI we could analytically correspond the note in the observed musical artifact and the sound that supposed to be played[1]. Thus, our focus would be in the pitches and its respective duration as the two primary elements comprising the melody of a song.

The paper presents an innovation on doing the observation that is distinctive with those similar previous approaches on music and the stylized (statistical) properties. The musical elements we use in our Zipfian analysis is the pitches and the durations comprised a song simultaneously and a unity. We realize that when we talk about a song or a melody as a musical artifact, a pitch cannot always be separated with the respective duration.

[1] Our computational model, technically, is regarding to MIDI standard digital music and some function used in Eerola & Toiviainen (2004).

In minimal point of view[2], MIDI files are constituted from the pitches with integers (*i.e.*: the note C as 64) and the duration (in beats or seconds) of each note in the score. For example, the cut from musical score of the Sundanese national-themed folk song *Manuk Dadali* below,



is read by the MIDI sequencer as of 19-notes along with the corresponding duration as we see in table 1. Then here is the tricky part. Both variables of pitch, φ, and the duration, δ, is transformed into its binary and we combine the binary form of φ_b and δ_b as a single long binary of θ_b.

Table 1
The pitches and the corresponding duration of *Manuk Dadali* Song

pitch φ	duration (in beats)	duration × 100 δ
79	0.53333	533.33
76	0.60833	608.33
77	0.65833	658.33
79	0.51667	516.67
83	0.6	600
84	0.94167	941.67
83	0.43333	433.33
84	0.55	550
76	0.61667	616.67
77	0.575	575
79	0.35	350
79	0.33333	333.33
79	1.2917	1291.7
79	0.475	475
76	0.53333	533.33
77	0.60833	608.33
79	0.63333	633.33
83	0.39167	391.67
84	0.98333	983.33

For example, as we have

$$\varphi_d(t=\tau) = 74 \rightarrow \varphi_b(t=\tau) = 0001001010$$

and

$$\delta_d(t=\tau) = 1000 \rightarrow \delta_b(t=\tau) = 1111101000$$

then we have

[2] In fact, MIDI files contained information more than just the pitches and the durations of a song.: the tempo, velocity, the copyright information, etc.

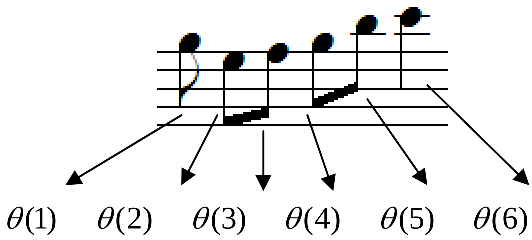
$\theta_b(t=\tau) = 0001001010111110100$

as the letter (with index number showing the sequence in the score $t=\tau$) combined with other $\theta_b(t)$, $t \in T$ resulting words with length *m* letters. This long binary sequence of θ_b is now the elements of the set of all the "letters" used in the song that is now we treated as text, $\theta_b \in \Theta_b$. Thus, if θ_b, we would have words used in the musical-"corpus" as they are comprised from the set Θ_b.

Simply speaking, this would remind us to the chromosome model widely used in genetic algorithm-like model (*cf.* Gen & Cheng, 1997:16-31). This is really not a new, since our model is incorporating one that has been previously used in financial time series (Situngkir & Surya, 2005) based on the pioneered work of Ausloos & Ivanova (1999). The latter apparently is inspired by the work on the analysis for DNA sequence which also exhibiting the Zipf's Law.

Thus, the numbers of the words in the musical corpus would be the function of *m*. If we have *m=1*, then it would similar to the standard analysis used in some previous Zipfian analytical work. Nonetheless, as we changed the value of *m* used in the analysis, we would have put into account the correlations between one note (constituted by the pitches and the corresponding durations) with another in the corpus we are observing.

As an example, the first measure of the song *Manuk Dadali* is formed by notes



and for *m=2*, we have words $\psi \in \Psi^{m=2}$ of

$$\Psi^{m=2}_{manuk\_dadali} = \left\{\psi^2(1), \psi^2(2), \psi^2(3), \ldots\right\}$$

where

$$\psi^2(1) = \{\theta(1), \theta(2)\}, \psi^2(2) = \{\theta(2), \theta(3)\}, \ \psi^2(3) = \{\theta(3), \theta(4)\}, \ \cdots \quad (1)$$

and for *m=3*, we have words $\psi \in \Psi^{m=3}$ of

$$\Psi^{m=3}_{manuk\_dadali} = \left\{\psi^3(1), \psi^3(2), \psi^3(3), \ldots\right\} \quad (2)$$

where

$$\psi^3(1)=\{\theta(1),\theta(2),\theta(3)\}\,,\ \psi^3(2)=\{\theta(2),\theta(3),\theta(4)\}\,,\ \psi^3(3)=\{\theta(3),\theta(4),\theta(5)\}\,,\ \cdots \qquad (3)$$

$$\vdots$$

$$\psi^{m=M}(1)=\{\theta(1),\theta(2),\ldots,\theta(M)\}\,,\ \cdots\ \psi^{m=M}(N)=\{\theta(N),\theta(N+1),\ldots,\theta(N+M-1)\} \qquad (4)$$

and so on, the greater *m* we used in our observation, the longer the words we have which consequently the longer correlation among notes we put into account.

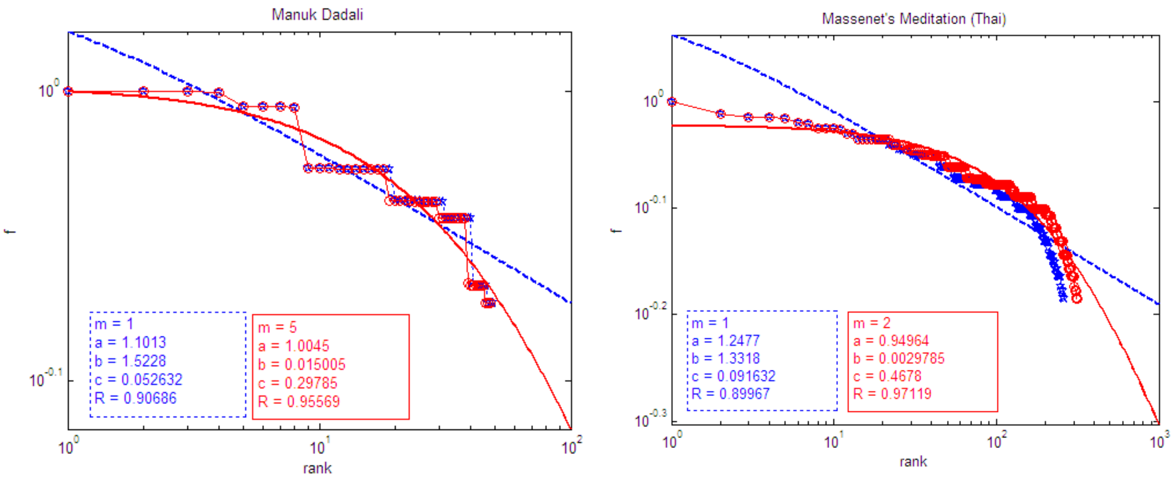


Figure 1. Two very different music: Sundanese *Manuk Dadali* and a masterpiece from French componist, Massenet, the Meditation (from Thai).

3. The Zipf-Mandelbrot Law in Music

Zipf-Mandelbrot Law has been confirmed in a lot of corpus and even cross-languages texts (Situngkir, 2007). By using our model as it has been discussed in the previous section, music could also be seen as corpus. Some work on Zipfian analysis on musical works (Manaris, *et. al.*, 2001) has discovered this and the platform we propose in the paper is supposed to confirm this interesting statistical properties of music. The notion of words in our analysis is obviously useful since it covers the existence of sequential correlation within musical works. The example of obtained better fit is clearly shown in figure 1 for the Sundanese music, *Manuk Dadali* and French componist's masterpiece, the Massenet from Thai. The fitting of the Zipf-Mandelbrot Law can be written as

$$f(r)=\frac{a}{(1+br)^c} \qquad (5)$$

or in the logarithmic form of,

$$\log f(r) \approx \log \mathrm{a} - c \log(1 + \mathrm{br}) \tag{6}$$

where *f* is the frequency with the corresponding rank (*r*), while *a, b,* and *c* are the fitting parameters of the Zipf-Mandelbrot Law (Mandelbrot, 1983).

We do analysis with some of Indonesian songs. It is worth noting that the data we used in this report is the "cleaned" MIDI files, means that we use the single stave data and we omit the staves that are apparently not the main melody of the song. Thus, roughly speaking we may call the data we use hereas the jingle of the song since we must admit that some music is relied upon the harmony yielded by simultaneously played instruments[3].

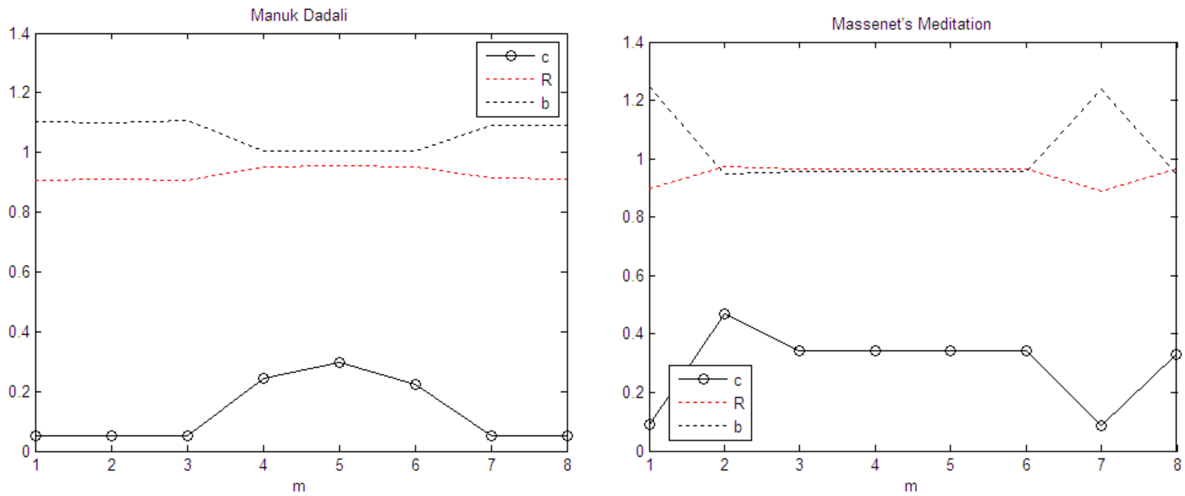


Figure 2. The comparison of Zipf-Mandelbrot fitting parameters (*b* and *c*) along with the *R* showing the different properties of the two distinctive song.

From the fitting process, we could see clearly that the sequential correlation is an important aspect if we want to analyze any musical product especially as we focus on the melody constituting a song. Furthermore, we could see that all musical products has the same best value of *m.* Obviously from figure 1, we can see that *Manuk Dadali* is better to be fitted with *m=5* while Massenet's Meditation with *m=2.* This fact is shown more obviously in figure 2.

[3] The MIDI files used in this report are downloadable in http://www.bandungfe.net/data/midis01.zip or by contacting author directly.

Table 2

Various *m* and respective fitting parameters on various musical artifacts

SONGS		*R*	*m*	*c*	*b*
Classical	*Ode To Joy (from Beethoven's 9th Symphony)*	0.9798	4	0.1799	0.05
	The First Violin of Mozart's Symphony 40	0.9606	1	0.7308	0.0024
Indonesian Anthem Songs	*Syukur*	0.9901	3	0.2927	0.0172
	Gugur Bunga	0.9763	1	0.9151	0.0105
	Halo-halo Bandung	0.9767	1	0.4736	0.0236
	Dari Sabang sampai Merauke	0.9614	1	0.7486	0.0147
	Gugur Bunga	0.9763	1	0.9151	0.0105
	Satu Nusa Satu Bangsa	0.9744	1	1.2725	0.0079
	Indonesia Pusaka	0.985	1	1.472	0.0063
Malay & Keroncong Songs	*Selendang Sutera*	0.953	1	1.5955	0.0063
	Teluk Bayur	0.9843	1	0.6055	0.0129
Indonesian Popular Songs	*Kugadaikan Cintaku*	0.9795	1	1.0871	0.0024
	Kaulah Segalanya	0.9638	1	0.3052	0.0057
	Kemesraan	0.9819	2	0.359	0.0085
	Gang Kelinci	0.966	7	0.693	0.0035
	Buku ini Aku Pinjam	0.9714	3	0.361	0.0057
Indonesian Folk (ethnic) Songs	*Suwe Ora Jamu*	0.9935	1	1.374	0.0102
	Bubuy Bulan	0.9734	1	1.7792	0.0054
	Tokecang	0.9876	1	0.3993	0.0261
	Ayam den Lapeh	0.9849	2	0.3134	0.0123
	Buka Pintu	0.99	1	0.5523	0.0201
	Dago Inang Sarge	0.9066	2	0.5182	0.0071
	Hela Rotan	0.992	1	2.3617	0.006
	O Ina Ni Keke	0.9909	1	0.889	0.0128

However, one "letter" constituting element (that we assumed as word) is the dominance in our various data. This is shown by the result shown in table 1 from various songs but also some reference data of the classical western music. We might want to classify the songs used in some categories, *i.e.:*

- ✓ *The Indonesian Anthem Songs*: the songs that mainly understood as songs known to be reflected the Indonesian historical moments, *e.g.:* independence, patriotism.
- ✓ *The Indonesian Popular Songs:* modern songs from Indonesian modern artists with the influence with some western popular music from various periods of time in which the songs broadcasted and becoming well known (70's, 80's, and 90's). In addition, this category as shown in the table is also enriched by some popular music

known as *keroncong* and Malay songs. *Keroncong* is actually a modern popular music but it has been inherently acculturated become the part of Javanic ethnic culture.

- ✓ *Indonesian (Traditional) Folk Songs:* the songs with lyrical and (some of them) the melody reflecting some ethnic groups in Indonesia. However, musicologist has shown that the "genuine" Indonesian folk songs turn out to be pentatonic (comprised by only 5 notes rather than the 7 notes per scale known in modern western music tradition).

The fitting with the Zipf-Mandelbrot of equation [6] seems to be good enough for the average $R \approx 0.9$, nonetheless with various values of *m* used in our analysis.

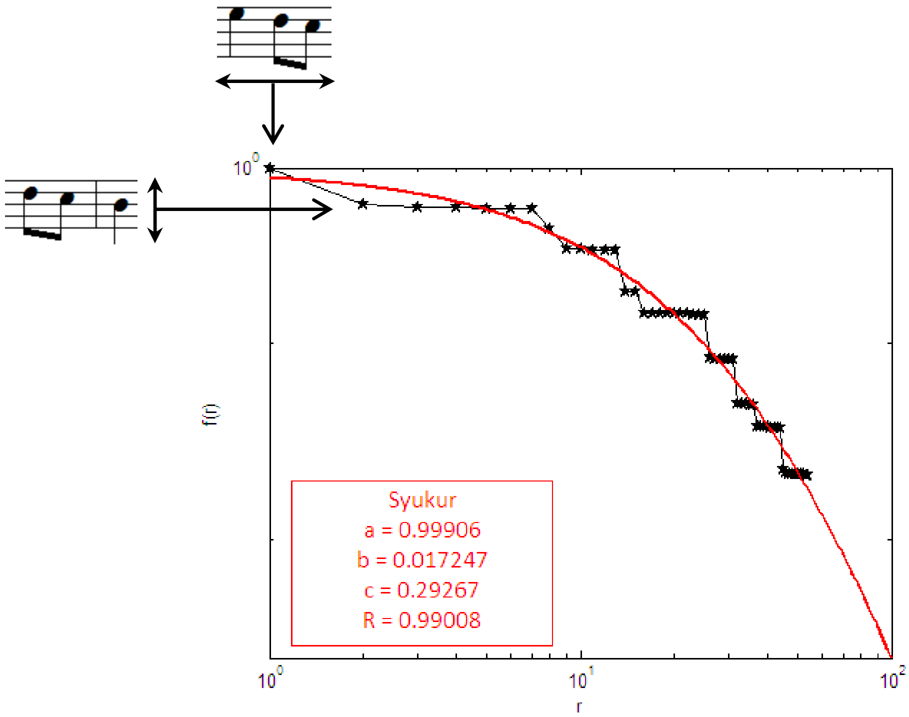


Figure 3. The two most frequent words (comprised by three notes) used in Indonesian anthem song, *Syukur*

4. Discussions & Further Directions

It is tempting to compare the fitting parameters found in songs and their categories. From the various songs we observed, we found interestingly that there are actually some conjectures could be used for our further research directions. First of first, we should remember that the Zipf-Mandelbrot Law is, in fact, showing the dominance of particular (or correlated sequence) of notes in the observed song. For example, figure 3 shows the Zipf-Mandelbrot fit to an Indonesian anthem song, *Syukur*, along with the obtained fitting parameters. It is clear that the most frequent notes would be placed as the first rank and so on, the second most used notes becomes the second rank, *etc.* Thus, roughly speaking, the plot shows us the relative dominance among notes (or notes) that comprises a song.

Information we obtained from the fitting parameters depict how this “dominance” occurs among the notes. Qualitatively speaking, the steeper the resulting curve would show more dominance over one note to another. The steepness is shown by the parameter *b* and *c* in the fitting process. Different with the original Zipf’s power function, the Zipf-Mandelbrot is ruled by at least two parameters when it says about the sloppiness of the frequency of each statistical event. The bigger the value of *b* and *c*, the steeper the yielded curve. Specifically, while the power exponent of *c* is ruling the slope of the curve, the values of *b* rules whether or not there are some sort of sloppiness occurs among the most frequent events.

Intuitively we can imagine the curve representing the each song in table 1. Roughly speaking, the greater *c*, thus the frequency distinction of several notes (or consecutive sequences of notes) tends to bigger in logarithmical proportion. Nonetheless, the lesser *b* yields circumstances in which among some frequent events of the first rank notes are having relatively similar frequencies.

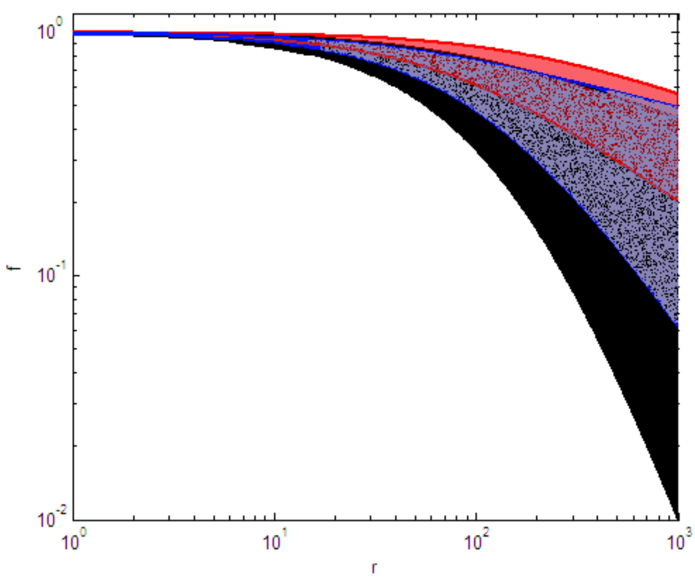


Figure 4. The area of the Zipf-Mandelbrot fits area of three big categories from various songs analyzed: Indonesian ethnic folk songs (*black*), modern Indonesian popular songs (*red*), and the Indonesian anthem songs (*blue*)

Beyond various musical scores we observed, we could see some interesting different areas in which a melody of a song laid within as shown in figure 4. At some cases, we could see that folk songs are more likely laid in the steeper areas. Here the used of some particular notes are depicting big gaps with those of the rarer used notes. However, from our observation on the western classical music that we used merely as reference, the modern popular songs are more likely showing the more similar pattern, the sloppier curve. Here, the employment of some notes (consecutive sequence of notes) are relatively depicting smaller frequency gaps. Furthermore, the Indonesian anthem songs are influenced by both traditional folk music and the western-style music. Some songs *e.g.*: *Satu Nusa Satu Bangsa* are relatively more "ethnic" while some other *e.g.: Syukur* are relatively showing more modern style. Nevertheless, in the figure, we could also see that some of Indonesian ethnic folk songs are also having been in the influence of the modern western musicality. Of course, since Indonesian artists or musicians have actually been interacting with some of modern musical works for not at all short period, the signatures of modern music should have been present even in Indonesian ethnic folk songs – at least, this occurred, for instance, as the modern musical instruments like *ukulele* was being introduced in Indonesian musical realms.

Obviously, further discussions about this finding could lead us to more philosophical discourses about general Indonesian musical works. For example, the discourse on how we could recognize a melodic artifact being influenced by ones that are socially constructed from very different cultures. Ethnomusicologists would be able to have this kind of discussions in further impact. Further directions of investigations that could be able to be conducted are things like the dynamics and the evolution of the signatures of modern popular musical from time to time. It is instinctive, that as we heard a sort of melody it is tempting to categorize in which period of modern popular music it may appear and widely accepted whether it is the 60's, 70's. 80's, or even 90's. This work may give conjectures to scrutinize this scientifically and could free us from subjectivity in our investigation and observation.

In the other hand, as we have the understanding about such things, it is easier for us to a more comprehensive research on musical generative art. In fact, this is the very first motive of research reported in the paper. The use of genetic algorithm or and other optimization models may lead us since the quantitative parameters we obtain can be useful for the technical fitness function as a melody is being generated. Those issues are left as further works and directions of research.

5. Conluding Remarks

A model to see melodies in a song as computationally constructed similar with those understood as in language is proposed – even though it modestly should be noted that there is no intention to have this proposition in the term of formal definition of language itself that has, in fact, been quite matured in computational linguistics. The technique is by representing the pitches and the respective durations in the musical score as letters constituting correlated sequences that could be regarded as words. This brought us to further top-down observation by investigating the rank-plot, known as Zipfian plot. Thus, the apparent Zipf-Mandelbrot Law is shown with relatively good result.

Further discussions brought us to both more philosophical and technical discussions about melodic features in music/songs. Reading statistical properties of melodies may bring us to the discourses related to the categorizations of music as anthropologically constructed in cultures and civilizations while in the other hand may show us conjectures of their classifications regarding to musical styles, places, time-period, and even the scope of the musical creations. In the other hand, technical works might use the result for further analysis on creativity in the musical context and the entertaining yet challenging approaches for the directions of future musical generative art.

Acknowledgement

This work is part of the research in Dept. Computational Sociology CS07001a. Author thanks Surya Research International for the support.

Works Cited

Ausloos, M., and Ivanova, K. (1999). "Precise (m,k)-Zipf diagram analysis of mathematical and financial time series when m = 6, k=2". *Physica A* 270:526-542.

Copland, A. (2002). *What to Listen For in Music*. Signet Classics.

Eerola, T. & Toiviainen, P. (2004). “MIR in Matlab: The Midi Toolbox”. *Proceedings of 5th International Conference on Music Information Retrieval (ISMIR 2004) Barcelona, 2004* pp. 22-27. Universitat Pompeu Fabra.

Feld, S., Aaron, A. F. (1994). "Music and Language". *Annual Review Anthropology* 23: 25-53.

Gen, M. & Cheng, H. (1997). *Genetic Algorithms & Engineering Design*. John Wiley & Sons, Inc.

Heckroth, J. (1995). Tutorial on MIDI and Music Synthesis. The MIDI Manufacturers Association. URL: http://www.harmony-central.com/MIDI/Doc/tutorial.html

Levitin, D. J. (2006). *This is Your Brain on Music: The Science of Human Obsession*. Dutton.

Manaris, B., McCormick, C., Purewal, T. (2001). “Progress Towards Recognizing and Classifying Beautiful Music with Computers: MIDI Encoded Music and the Zipf-Mandelbrot Law”. *Technical Report CoC-CS* TR#2001-7-2. College of Charleston.

McCormack, J. (2006). “New Challenges for Evolutionary Music and Art”. In Lanzi, P. L. (eds.), *ACM SIGEVOlution Newsletter* 1(1): pp. 5-11. ACM Special Interest Group on Genetic and Evolutionary Computation.

Manaris, B., Vaughan, D., Wagner, C. (2003). “Evolutionary Music and the Zipf-Mandelbrot Law: Developing Fitness Functions for Pleasant Music. In Cagnoni, S. *Lecture notes in computer science* 2611 pp. 522-34. Springer.

Mandelbrot, B. B. (1983). *The Fractal Geometry of Nature*. Freeman.

Selfridge-Field, E. (1998). "Conceptual and representational issues in melodic comparison". In Hewlett,W.B. & Selfridge-Field, E. (eds.), *Melodic Similarity: Concepts, Procedures, and Applications*. MIT Press.

Situngkir, H., Surya, Y. (2005). “What can we see from Investment Simulation based on Generalized (m,2)-Zipf Law?”. *BFI Working Paper Series* WPE2005.

Situngkir, H. (2007). “Regimes in Babel are Confirmed: Report on Findings in Several Indonesian Ethnic Biblical Texts”. *BFI Working Paper Series* WPC2007.

Su, Z-Y., Wu, T. (2006). "Multifractal Analyses of Music Sequences". *Physica D* 221: 188-194.

Steedman, M. (1996). “The Blues and the Abstract Truth: Music and Mental Models”. In A. Garnham and J. Oakhill, (eds.), *Mental Models In Cognitive Science* pp.305-18. Erlbaum.

Temperley, D. (2001). *The Cognition of Basic Musical Structures*. MIT Press.

Voss, R.F., Clarke, J. (1975). “*1/f* Noise in Music and Speech”. *Nature* 258: 317 – 318.