

Rhythmic Elements in Jeru the Damaja's "Come Clean"

Introduction

One of the most compelling elements of rap music is the rhythm, specifically the way that the live vocal track shifts rhythmically against the (usually) looped drumbeat. If this drumbeat is programmed into a sequencer or drum machine (rather than a 2 or 4 bar looped musical *sample*, which is common), it can be very regular (even if *inéga*le, that is, even if the groove of the beat features slightly *swung* durational values). This regular groove pattern can set up a strong rhythmic underpinning upon which the rhythmic essence of the vocal line can dance.

From a purely *mechanical* standpoint (ignoring sociological or political relevance and poetic savvy), there are four rhythmic elements that help give a vocal rap performance its panache: syncopation, *latency* (see definition below), the treatment of pitch, and the treatment of prosody (which includes treatment of pitch). Using Jeru the Damaja's 1995 song "Come Clean" as an example, we shall see some of the ways that all of these elements are intricately entwined and dependent on each other to create a rhythmically sophisticated vocal performance.

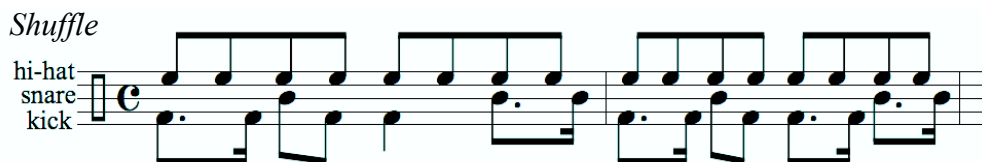
Background

In the process of focusing on such a narrow topic I have made several assumptions during the analysis of this music within a prescribed paradigm. Some of these choices may seem more intrusive than others; however, I will try to qualify each

one. This does not mean, of course, that other methods could be used or assumptions made. Indeed, any musical transcription assumes that the method of notation is a valid one: who is to say that standard western notation could be a meaningful method of understanding non-western music? What you see through your microscope is only as clear as your lenses or your observational techniques. That said, the validity of the transcriptional and analytical choices and methods in this paper should be made clear as each of these elements comes in to focus.

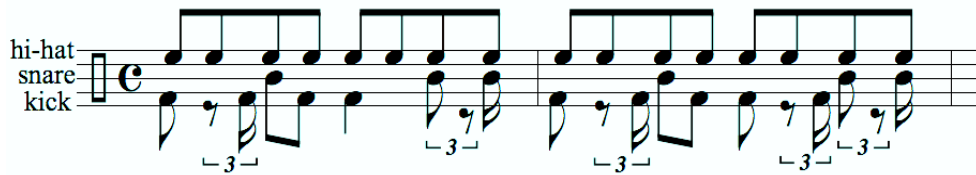
The Beat

As the drumbeat provides the fundamental rhythmic background in “Come Clean”, it is the first important element to analyze if you are to understand the rhythmic context of the vocal performance. The drumbeat is a two bar loop, and for practical purposes in a *musical performance setting* (rather than a purely analytical one) would likely be notated as:



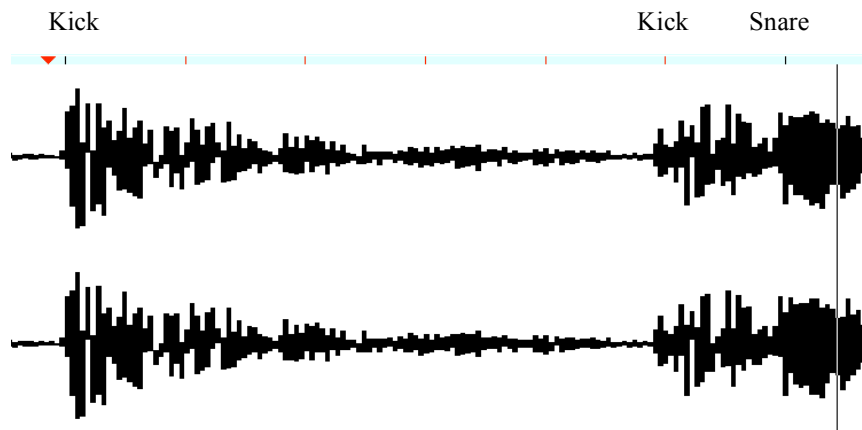
Example 1. Transcription of the beat for “Come Clean” with squared-off sixteenths and the instruction of “shuffle.”

However, the rhythmic feel of the piece is swung sixteenths, or triplet sixteenths. Therefore, any event which occurs on sixteenths 2, 4, 6, 8, 10, 12, or 16 will likely occur closer to, (but usually not exactly on), the third note of a triplet sixteenth group, such as:



Example 2. A more accurate, although less practical, transcription of the beat.

We can see from the close view of the waveform of the beat that the second kick drum event (see Example 1 above) occurs closer to the 6th triplet sixteenth by mapping a series of tick marks to the beat which divides it evenly into six parts:



Example 3. Close view of the drumbeat and back ground for “Come Clean.” The first black tick mark shows the downbeat and the second black tick mark shows the first snare hit, or beat 2. The beat is divided evenly into 6 parts, which equal triplet sixteenths. You can see that the second kick drum event from the drumbeat occurs close to the 6th triplet sixteenth of the measure.

Because the exactly placement of those even-numbered sixteenths varies from piece to piece, no *standardized* precise notation (which could be predictably used for a live performance) for them can exist. Indeed, the “feel” of the music can be said to exist in the placement of these events – some music will be more swung and other music will be less swung. Therefore, the standardized notation for this beat would usually use straight sixteenths (as in Example 1 above) with the instruction to “swing sixteenths “or” shuffle, leaving the exact feel to the performer. This allows each group or performer to conceive of his or her own relationship to the swing.

Appendices and Definition of Terms

Appendix I shows the lyrics for the song written in phrases as if it were a poem, with each logical phrase on its own line (numbered for future use). There are, several interesting elisions in the structure of the lines where the length of the lines varies to create an unpredictable rhyming structure and rhythmic feel. An example is:

*Got a freaky, freaky, freaky, freaky flow¹
Control the mic like Fidel Castro¹ locked Cuba²
So deep you can scuba dive³
My jive's³ origin is unknown like the Jubas²*

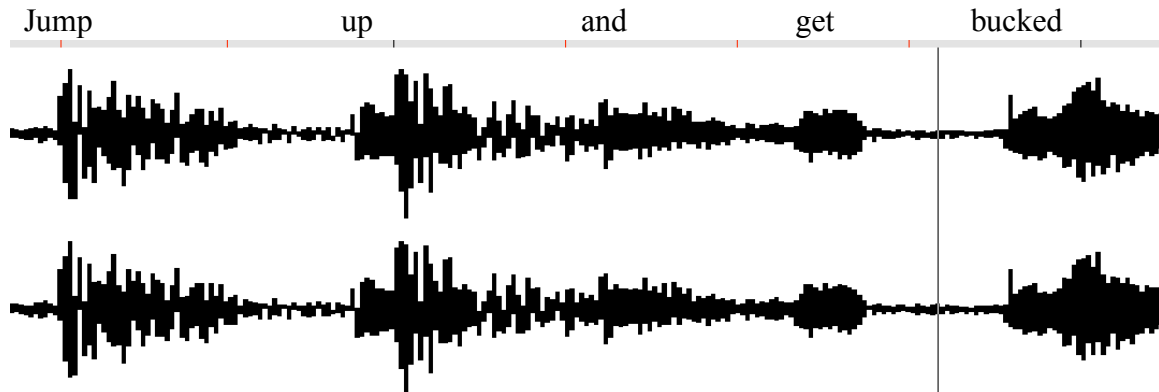
The rapper rhymes “*flow*” with “*Ca-stro*,” “*Cuba*” with “*Jubas*,” and “*dive*” with “*jive's*” (each rhyme group is given a number above). Therefore, he has intermingled the rhyme groups, some of which have a longer duration than others. The effect is one of upsetting the regular flow of rhetorical meaning and adding an element of surprise to the vocal performance. As we get further into the other elements, the phrases and rhyming groups will be examined again.

Appendix II shows a transcription of the piece using a straight sixteenth grid as a basis, and ignores the swung feel of the sixteenths. The reason I have temporarily ignored the swung feel of the sixteenths is that, when looking at the rhythmic feel of the vocal performance in milliseconds, it is actually quite irregular, and to transcribe the piece using swung sixteenths would be just as misrepresenting as one with straight sixteenths. Therefore, I chose to use the standardized notation, which *assumes* swung sixteenths, with the understanding that the details will be more closely examined later.

This assumption is one which is similar to that used in the transcription of the beat above. Therefore a phrase which is notated as:

this *one example*: if we want to find a generalized notation which can apply to more than one example, then one of these transcriptions would likely be used (in fact this type of rhythmic figure does appear in almost exactly the same way several times – more on this below).

If we take a close look at the waveform:



Example 8. The first black tick mark is the kick drum on beat 3, and the second black tick mark is the snare drum on beat 4. The words “jump up and get bucked” can be seen as the other waveform shapes, which are evenly spaced.

several things are made clear. First, the word “*jump*” appears almost exactly on the eighth note before the beat, and therefore it will be placed on that eighth note without the conflict of triplet vs. no triplet. Deciding on a notation for the words “*up and get bucked*” however, is tricky: while the *temporal placement* of these words in relation to the underlying beat sounds like something close to the that of Example 7 above (but in which the words “*up and get bucked*” are *not* evenly spaced in respect to each other), it is clear from viewing the waveform and from listening that the words *are evenly spaced*. If we measure the distance between the quarter-note beat to “*and*,” “*and*” to “*get*,” and “*get*” to “*bucked*,” we can see that each distance is almost exactly .213 seconds. This confirms their even spacing, and implies some kind of triplet.

However, it does not satisfy the observation that these 4 words do not align very closely with the beat, and, if grouped as a triplet and notated as in *Example 6* above,

would appear quite *early*. Therefore, a *latency value* can be assigned to these notes (more below) to properly represent the temporal displacement of these events. *It is like an eighth-note triplet which has been shifted early*. Again, a transcription paradigm with an asterisk is being adopted here for a general understanding the music.

Appendix III attempts to pull together all of the elements I am studying: latency, pitch, syncopation, and syllabic stress or *prosody* (borrowing a term from the study of linguistics). To understand how syncopation is used in this piece, I devised the scheme in which each word or syllable is given a level of syncopation based on the perceived durational transcription. Here is the breakdown of how these values are conceived:

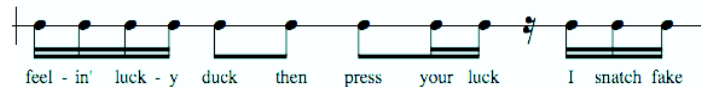
Syncopation Level	Event occurs:
0	directly on beat 1
1	directly on beat 3
2	directly on beat 2
3	directly on beat 4
4	on eighth notes 2 or 4 (or triplet eighths 2,3, 5 or 6)
5	on eighth notes 6 or 8 (or triplet eighths (8,9, 11 or 12)
6	on sixteenth notes 2, 4, 6 or 8
7	on sixteenth notes 10, 12, 14 or 16
8	on an off-beat note from beats 1-2 (see levels 4 and 6) without a preceding event of the same or shorter duration
9	on an off-beat note from beats 3-4 (see levels 5 and 7) without a preceding event of the same or shorter duration

Example 9. Syncopation levels of any possible event

Each word or syllable is given a level number, as shown in Appendix III, and a shaded cell ranging from white to dark brown.

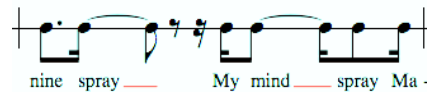
In a somewhat unscientific attempt to measure the “unpredictability” of events, this system makes many assumptions. It could therefore be seen to be highly problematic. For example, it assumes that beats 1 and 3 are the strong accented beats of this music while beats 2 and 4 are the weak beats (and assumes that beat 1 is stronger than beat 3, and beat 2 stronger than beat 4). While this may be true for Western European

classical music, does this apply to this Hip Hop style, which, it has been argued, is heavily influenced by African rhythmic sensibility? After all, most rock, funk, and pop music places the accents on 2 and 4. Further, this system also makes the assumption that *an event which is not prepared by a preceding event of equal or longer duration is more syncopated than one which is*. For example:



Example 10. The syncopation of the words “*I snatch fake*”

The word “*I*” is highly syncopated because it is not prepared with a sixteenth note directly before it. The word “*snatch*” is also highly syncopated because it falls on an eighth note but there is no event on the eighth note before it (on beat 4 of the measure). The word “*fake*,” however, is less syncopated than “*I*” because it has been prepared with the sixteenth before it. Furthermore, in the following example:



Example 11. The syncopation of the words “*spray, my mind spray*”

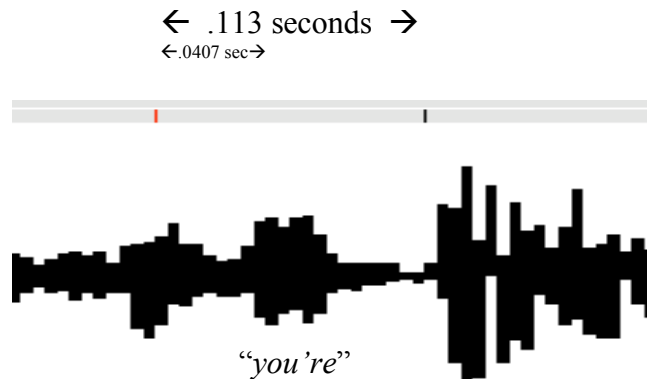
the second occurrence of the word “*spray*” is highly syncopated because it is not prepared with a preceding sixteenth, even though the word “*mind*” does last for a dotted eighth and lasts up until the beginning of the word “*spray*.” Finally, it may seem like a grand omission that the 1st and 2nd beat bias does not come into play regarding these unprepared durations.

In defense of this system I would offer that, quite simply, syncopation is a complicated phenomenon which cannot be understood merely with a prescriptive notation on paper and which must be understood within the context of the *live*

performance which uses the background rhythm and music. Therefore, this system is merely a glance into the rich world of syncopation for use only in this paper. It will be used here mostly to understand how the lyrics are placed within the context of the beat and measure (and with the other elements that Appendix III is examining), and will come in handy later for our analysis.

Appendix III also maps the *latency* of each syllable or word. For the purposes of this paper, I am defining *latency* as *the value of the distance of a word or musical event from its closest perceived durational placement, within a predetermined durational or quantization grid*. Rap music is “groove” music, which features a regular flow of regular pulses, and if the music features a regular occurrence of swung sixteenths, (as does “Come Clean”), then we can say that the duration grid of this piece will be a *swung sixteenth note quantization grid*. It should be noted here that, while the *straight* sixteenth was adopted in the transcription, the *swung* sixteenth is used here for an analysis of latency, since this is a closer approximation of the actual sound of the music, and because the beat lays down a steady flow of swung sixteenths (see Example 3 above). If an event occurs close to one of these swung sixteenths, then the latency of the event will be given a value of its distance between that “home” sixteenth and the next *marker* that occurs within this grid. It is important to understand that the next *marker* is not necessarily the next swung sixteenth – it may be the next *eighth*: In Example 5 above we can see how it is inherent in the natural flow of the swung sixteenths to place the *swung* event on the third sixteenth of a triplet sixteenth group. Therefore, in the following examination, all events are measured by their proximity to the closest marker, as if they were to occur in a perfect swung sixteenth grid, which uses triplet sixteenth groups.

Example 12 from the song shows how each latency value was achieved:



Example 12. The word “*you’re*” is given a latency value of 36 with the following formula: $(1 - ((.113 - .0407)/.113))$

Each swung sixteenth note takes approximately .113 seconds to occur. If an event sounds in a surface listening as if it occurs directly on that swung sixteenth (and would therefore be transcribed that way), but, upon closer viewing, actually occurs .407 seconds later (as in the word “*you’re*” in Example 12), then the word would get a latency value of 36. This is achieved with the following formula: $(1 - ((.113 - .0407)/.113)) = .36$ or 36%. The word “*you’re*” occurs 36% of the way to the next marker. The latency of each *late* event has been mapped out in this way in Appendix III where green bars represent late events.

What happens, however, when an event is *early*? This does not happen as often, but I believe the rapper uses this rhythmic effect in some interesting ways. Since an event which is early can sound alarming and exciting in his predominately late vocal performance, I chose to reverse the intensity of early events, so that the earlier an event is, the shorter the orange or red bar in the downward direction. Therefore, an event which occurs almost on following marker will get a very tall bar, while one which occurs earlier in the measure (which would not sound so alarming to the ear and might even sound like a late event which belongs on the previous marker) would get a shorter, more

lightly colored bar. This helps to chart the some of the exciting events which throw the rhythmic feel off in yet another interesting way.

This latency system may seem problematic in ways similar to that of the syncopation and transcription issues, as it assumes that events must occur close to these markers merely *because* I am assigning them to the piece. Latency here, as with syncopation, is a complicated phenomenon which deals with the underlying regularity of the drum beat – without it (and with a varying live drum track) latency, as it is described and charted here, might be less meaningful, as the exact placement of any event could be seen to be as *variable* as the underlying beat. *It is, however, designed to describe something which is clearly audible*, and, as it makes the best attempt possible at showing the placement of these events, it will prove useful for our analysis below.

In addition to the latency, Appendix III shows several elements of the prosody of the vocal performance: pitch, syllabic stress, and phrase intonation. First, each phrase, as parsed in Appendix I, is marked with a vertical double bar. This helps to align all the elements that Appendix III is examining. Next, an approximation of the pitch of each syllable is shown overlaid under each syllable, using A3 (the A below middle C) as the mean pitch around which the vocal performance hinges. It was difficult to assign a note to each syllable, since very often a single syllable would glide up or down and cover a whole step or more. Therefore, in an attempt to show the pitch contour of the lyrics, I generally chose the most extreme pitch that a syllable covered to contrast it to the neighboring syllables.

Next on Appendix III, the intonation or shape of each phrase is drawn with arrows across the cells representing the pitches. And finally, the syllabic stress of each word

within the pattern of everyday speech is given. Generally, speech is grouped into *metrical feet*, a term borrowed from linguistics and poetry analysis. A *metrical foot* consists of two syllables – one weak and one strong. If the stress pattern goes weak→strong, then it is called *iambic*, and if the stress pattern goes strong→weak, then it is called *trochaic*. If the foot has only one syllable (as in the expression “Hey!”), then the stress on that syllable must be strong. If a one-syllable word occurs, and then there is a pause before the next word, then these two words would not be of the same foot.

Normal spoken English does not necessarily have any *regular* pattern of syllabic stresses, although some languages do. Also, classical poetry makes use of different patterns (hence the *iambic pentameter* of poetry by Shakespeare and others, in which the syllables all follow an iambic pattern and each line has ten syllables). This vocal performance, however, is not spoken, but *declaimed*, and therefore will have a different pattern of stresses. As far as I can see, it does not adhere to any regular pattern within each foot, and therefore, to aid in visual clarity, I chose not to group the syllables by feet, but to retain the phrase groupings as mentioned above. This method will prove useful in the section below.

Analysis

Now that we have defined all the terms and approaches, tried temporarily to toss out the objections, and described some of the individual phenomena found in this piece, we can begin to look at how these phenomena might be related to each other using the multi-dimensional Appendix III as a guide. Here are some observations:

Latency: there is a regular flow of *late* events throughout the entire verse. There are sporadic *early* events throughout the verse, but more at the beginning of the verse.

The rapper treats these problematic eighth-note triplets (mentioned in the section on Appendix II) in an interesting way: on the words “*jump up and get bucked,*” and “*but if you persist*” he shifts the triplet grouping early by nearly the same amount! The latency values of the words “*up and get bucked*” are 65, 78, 53, and 34, while that of “*but if you persist*” are 33, 86, 47, and 35. These are not exact (and some error should be assumed in measurement – in fact it was often difficult to decide exactly where the beginning of each quarter-note beat occurs), but they do confirm what is audible, that the rhythmic sound of these two phrases are very similar.

Other observations: there are 5 events which are off the chart in lateness and are given a latency value of 100 or higher – these still seemed worthy of placing on their marker, rather than ascribing them to the next marker with a different latency value. Also, late events very often occur near the end of the phrases. There are exceptions, but it seems that the syllables seem to get later as the phrase continues, and then the timing is reset for the next phrase.

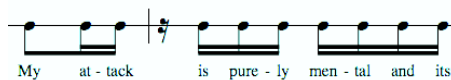
Pitch: 20 of the 36 phrases end on an upswing of pitch. This is completely uncommon in every day speech, where this rising intonation would be reserved for a question. It seems to go well with this kind of declaimed speech, in which each phrase is like a demand or statement of power or skill. Very often a phrase will start with a more regular contour and then fan out in register so that it becomes like an increasing tremolo.

Also interesting is the overall structure of the use of high notes, specifically the C’s. There are two, on the words, “*what*” and “*gang bang*” near the beginning of the verse. We then do not hear any other high C’s until near the end of the verse, on the words, “*tote,*” “*mental,*” “*nonsense,*” “*persist,*” “*flip,*” “*dip,*” and “*spit*” (not actually quite

a C, but very close to it and sounds to my ear as if it is in this same register). This is not to say that these are the most important words in the verse. What can be seen here is a kind of introduction of that high register, and then a driving home of it later in the verse: We get two occurrences within the first 22 seconds of the rhyme, and then 6 occurrences in the last 20 seconds, without any other occurrences in between (although “*mist*,” and “*jungles*” come pretty close).

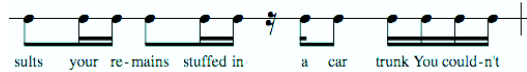
Another observation in the world of pitch is the way that the rapper ends each multi-phrase group (as parsed on Appendix II – usually two phrases, although sometimes more if there are internal rhyming groups such as that noted above the group starting on “*got a freaky, freaky, freaky freaky flow*”): generally in the same manner that he ended the first half of the multi-phrase group. For example, he ends both the phrases “*I’m a true master you can check my credentials*” and “*cause I choose to use my infinite potential*” on a downward glide. This is quite common in the multi-phrase groups here and helps to give the phrases coherence and to help parse them for the listener.

Syncopation and rhythm: Often the downbeat is evaded with a sixteenth note right before and a sixteenth note right after, such as in the phrase:



Example 13.

This makes for an accented downbeat as the ear fills in the accent, and it is an exciting way to present the rhythmic flow. Also, if an event occurs on the downbeat, it is often not followed by a sixteenth note, but will get a full eighth note value as in the phrase:



Example 14.

This helps to accent the downbeat in another way.

There are also syncopated events usually at the beginning and ending of each phrase. This is in part because *the rapper begins nearly every phrase on the last eighth or three swung sixteenths of the measure!* This has the effect of accenting these words which occur on the following downbeat from where the phrase began:

Front, feelin', gangster, nine, malignant, results, come, survive, gang bang, relentless, true, choose, freaky, control, locked, jives, accumulated, rather, back, (mac), attack, wake, stagnate, snot-box, hoes, know, Dirty (from *Dirty Rotten Scamp* - a nickname for himself - also, *Dirty Rotten Scoundrel*)

What is fascinating here is that almost all of these words are arguably the most important words in the phrases in which they occur (notice there are no “the’s” or “and’s” which are not important words for dramatic import). For example, in the phrase “*you couldn’t come to the jungles of the East poppin’ that yang*” the imagery here is of a foreign rapper coming to Jeru’s home turf where the other rapper could not match Jeru’s rhyming skill. Or in the phrase, “*I’ve accumulated honies all across the map,*” the rapper emphasizes his prowess with the opposite sex with the word “*accumulated.*” This implies that these women are accessories to him, which can be accumulated like so much detritus. Further analysis into each phrase reveals that the important words are often on the downbeats, and using his methods of accenting the downbeats, he is able to make these moments highly memorable.

Another fascinating observation is that the second half of the multi-phrase rhyming groups generally have more late events, such as in the two-bar phrase: “*you couldn’t come to the jungles of the East poppin’ that yang, you won’t survive, get live,*

catching wreck is our thing.” This has the effect of making you pay attention to the second half of the phrase since it is the one with the rhythmic surprise in it. Since the second half the multi-phrase group is the “2” in the rapper’s “1-2 punch,” this latency technique helps emphasize the import of the text and drive home the meaning. An interesting exception to this pattern is the phrase “*got a freaky, freaky, freaky, freaky flow*” which features the most latent event (*with latency value of 112!*) of the entire verse: the word “*flow*” (arguably, but certainly in the context of this paper, the most important word of the verse!). I believe that the rapper uses a great deal of late events here combined with the repeated word “*freaky*” to help drive home the import of this concept.

Syllabic stress: There appear to be no patterns of syllabic stress, however, the rapper does frequently shift from an iambic pattern to a trochaic pattern, which helps give the rhythmic flow a degree of liveliness. Furthermore, there are several places where he presents a word with a stress that is very uncharacteristic for that word, such as “*gangster*” (usually GANG-ster, here it is gang-STER), “*lyrics*,” (here it is ly-RICS) “*origin*” (here it is OR-i-GIN) “*attack*,” (here it is AT-tack) “*stagnate*,” (here it is stag-NATE) “*nonsense*,” (here it is non-SENSE) and “*persist*” (here it is PER-sist). These disruptions of the predicted stress of these words aid in giving the vocal performance an unpredictability and surprising character.

Conclusion

We have taken a close look into some of phenomena of this music and techniques the rapper employs to give an unpredictable and lively vocal performance. I have no doubt that if one were to stare at Appendix III for a while other things might pop out, and

certainly different viewers might have different perceptions of the graph when viewed while listening to the music. And perhaps other methods of analysis could be used, and then a completely different conclusion would be reached. I have attempted, however, to be as faithful to this music as possible, as its rhythmic influence on me as a composer cannot be overstated. The most interesting element that seems to be untouched here is the element of the meter: are beats 1 and 3 really the strong beats in the rapper's mind, and is beat 1 stronger than beat 3? These are assumptions that it is too easy to make when coming from a Western European classical background, but do they apply here? These, as well as other assumptions, upon which I have just barely touched, could connect to a whole world of historical and ethnomusicological meaning.

Appendix I – Lyrics

Grouped in phrases by logic, with underlines parsing the rhyming groups

Verse 1

1. You wanna front?
2. What?
3. Jump up and get bucked
4. If you're feelin' lucky duck then press your luck
5. I snatch fake gangster emcees and make them faggot flambé
6. Your nine spray
7. My mind spray
8. Malignant mist that'll (unclear)*
9. The results your remains stuffed in a car trunk
10. You couldn't come to the jungles of the East poppin' that yang
11. You won't survive
12. Get live
13. Catchin' wreck is our thang
14. I don't gang bang or shoot our bang bang
15. The relentless lyrics the only dope I slang
16. I'm a true master: you can check my credentials
17. 'Cause I choose to use my infinite potential
18. Got a freaky, freaky, freaky, freaky flow
19. Control the mic like Fidel Castro locked Cuba
20. So deep that you can't scuba dive
21. My jive's origin is unknown like the Jubas
22. I've accumulated honeys all across the map
23. 'Cause I'd rather bust a nut than bust a cap in your back
24. In fact my rap snaps your sacroiliac
25. I'm the mac so I don't need to (unclear)*
26. My attack is purely mental and it's nature's not hate
27. It's meant to wake you up out of your brainwashed state
28. Stagnate
29. Nonsense
30. But if you persist
31. You'll get your snotbox bust you press up on this
32. I flip
33. Hoes dip
34. None of the real niggers skip
35. You don't know enough math to count the mics that I've ripped
36. Peep the dirty rotten scamp as his verbal weapons spit

*I contacted Jeru the Damaja via his *MySpace* page, but was unable to get a clear answer to these two passages (I don't believe it is managed by Jeru himself, but by a clever fan)