The role of emotion in AAVE pronunciation: Mumble Rap as a phenomenon of Language Evolution

Bob Rossen
Advised by Ans van Kemenade

Abstract

This study will focus on the effect of emotion on pronunciation in the specific case of Mumble Rap, a relatively new phenomenon in rap music introduced by speakers of African American Vernacular English (AAVE). It is hypothesised that mumble rapping is the result of emotion in speech and hence intonation, pitch range, and enunciation energy are higher in Mumble Rap tracks than in Lyrical Rap tracks. Five tracks of both styles are transcribed and analysed by measuring the vocal parameters using Praat computer software. A significant difference in pitch levels as well as pronunciation accuracy between the two rap styles provide evidence to suggest that speech in Mumble Rap is more emotional than Lyrical Rap. This provides insight into the development of AAVE as a variety of standard American English (AmE).

Key words: emotion, pronunciation, AAVE, rap music, voice quality
Teachers who will receive this document: A. Van Kemenade & J. Geenen
Title of document: Bachelor Thesis
Name of course: BA Thesis Taalkunde
Date of submission: 13 June 2018

The work submitted here is the sole responsibility of the undersigned, who has neither committed plagiarism nor colluded in its production.

Signed

Name of student: Bob Rossen

Student number: s4156854
# Table of Contents

Abstract 1

1. Introduction 4
2. Background / Literature 5
   2.1.1 The History of Hiphop 5
   2.1.2 Controversy 8
   2.1.3 Defining Mumble Rap 9
   2.2.1 Emotion in Speech 11
   2.2.2 Intonation, pitch, and pitch range 12
   2.2.3 Intensity 14
   2.2.4 Pronunciation 16
   2.3 Expected Findings 17
3. Method 18
   3.1 Content Analysis 18
   3.2 Eliminating Variables 18
   3.3 Sample Selection 19
   3.4 Procedure 20
   3.5 Phonetic Analysis 24
4. Results 25
   4.1 Pronunciation 25
   4.2 Intensity 25
   4.3 Pitch 25
   4.3 Pitch Range 27
   4.4 Other Findings 28
5. Conclusion 28
6. Discussion 29
References 31

Appendix #1 36
1. Introduction

The unintelligible pronunciation of lyrics in music is no new phenomenon. Jimi Hendrix sang that he “kissed the sky” in his famous *Purple Haze* (Hendrix, 1970), but more than one fan was convinced that he said “kissed the guy” instead. As he mumbled the words people failed to agree on what Hendrix actually sang. Since the 2010s several American rappers, specifically from the South Eastern States, started uttering indecipherable vocals in a similar fashion, and were hence labelled Mumble Rappers. This style is revolutionary for a music genre that is known for its lexical dexterity and creativity with wordplay. The popularity of Mumble Rappers is exploding and so is the art form, with numerous mumble rap tracks by big artists still coming out every week. The rapping style is characterised by laxations and omissions in the pronunciation of the words, which are known in linguistics as assimilations and segmental deletions. This type of speech thus differs from the pronunciation in other rap music, which is more in line with the standard pronunciation of African American Vernacular English (AAVE). As hiphop language is a representation of AAVE speech (Richardson, 2006), this new way of rapping seems to reflect a development in the use of the English language by a community of native speakers.

AAVE is a language that is systematically different from AmE. This shines through in the lexicon with new vocabulary or in the grammar with constructions such as double negation, but also in the pronunciation system. There is no standard of pronunciation for AAVE, but it is clear that it features a lot of phonetic reduction (Wolfram & Schilling, 2016). This can be observed in the pronunciation of words such as ‘alright’, where there is /r/ and /l/ dropping in the comparison between standard AmE /ɔˈlət/ and AAVE /ɔˈət/. It is suggested that this difference in pronunciation between AmE and AAVE has sociolinguistic grounds, as Schwartz (1978) found that AAVE speakers regard emotional factors as more important in speech than linguistic factors. More recently, Polzin and Waibel (1998) found evidence to suggest that the emotional state of the speaker affects the pronunciation accuracy. These findings, however, do not necessarily suggest that AAVE speakers are more emotional than AmE speakers at all times, but they do provide a possible explanation for the decrease in pronunciation accuracy in AAVE speech.

If Mumble Rap is more emotional than the other ways of rapping, this should be perceptible in the voice of mumble rappers, as research shows that an increase of emotion in speech has consequences for the voice quality (Sherer, 1984). More
specifically, emotional speakers have been found to speak louder, more intensely and with a higher intonation than emotionally neutral speakers. This raises the question what the influence is of emotion on the pronunciation accuracy in AAVE speech. If Mumble Rap is indeed more emotional speech than lyrical rap, this should be observable in higher scores on values of vocal parameters.

This study will investigate the influence of emotion on the pronunciation of AAVE conducting a linguistic analysis of rap tracks. First, the pronunciation accuracy will be measured with the help of phonetic transcriptions to distinguish the two categories. This will form as the basis of the second analysis, which will be an in-depth voice analysis with Praat software. The average pitch levels, pitch range, and speech intensity will be computed and subsequently analysed to see if mumble rapping can be regarded as more emotional than standard rapping on the basis of these values. The results will provide an answer to the question if mumbling in rap is based on an increase of emotion, which would explain the decrease in pronunciation accuracy. This will help understand the mechanisms underlying changes in the language of a group of AAVE speakers.

2. Background / Literature

2.1.1 The History of Hiphop

Hiphop has been said to be conceived in The Bronx, New York around 1973 (Cole, 2015). DJ Kool Herc, known as the founding father of hiphop, was hosting parties where he would start experimenting with different records playing concomitantly on multiple turntables. This created music that could play for as long as he wanted. Hence, the first beat production was created from this process. Herc started noticing that rhythmically speaking into the microphone sparked something in the crowd, so he invited his friend Coke La Rock to be the stage host so Herc could focus on mixing the records. During one of these parties, Coke rapped his first line “There’s not a man that can’t be thrown, not a horse that can’t be rode, a bull that can’t be stopped, there’s not a disco that I Coke La Rock can’t rock”. Rap music was born, and Kool Herc and Coke la Rock are since known as the first MCs.

The style’s popularity spread rapidly around the underground scene of New York (Adaso, 2017). Many new MCs and DJs started experimenting with records and raps. Three artists by the name of Africa Bambaataa, Grandmaster Caz, and Grandmaster Flash acquired
local fame by performing at parties in the neighbourhoods of New York. During one of the
parties, Africa Bambaataa and Disco King Mario directed their raps and beats towards each
other, which marks the first battle rap that the style would become famous for. This is also
the time when the term hiphop popped up to describe the rhythm. When the Sugarhill Gang
created the first hiphop track ever in 1979, called ‘Rappers Delight’, the first lines of the
lyrics read ”I said a hip hop, Hippie to the hippie, The hip, hip a hop, and you don't stop, a
rock it out” (Sugarhill Gang, 1979). The term hiphop has been used interchangeably with
rap music ever since.

The style of hiphop has always been up for refurbishment. Whereas 1970s and early
80s hiphop was typical of its light-hearted lyrics and dito sounds, with the Beastie Boys and
Salt ’n Peppa as shining examples, hiphop of the 1980s was rawer. West Coast rap music
was emphatic of the gangster lifestyle, with pioneers as Ice-T and Los Angeles’ N.W.A..
They got nationwide attention and commercial success but also received considerable
critique. Tracks like ‘F*ck the Police’ were on top of the charts but were criticised at the
same time for their aggressive lyrics which even evoked attention of the FBI at some point.
The image and the tone with which people talked about rap music was far from positive, but
many others new artists were inspired by the reality raps and started rapping about their own
lives, and experiences of social inequity and political oppression. The following decade saw
record after record being put out and is now seen as the golden era of hiphop (Duinker &
Martin, 2016), with artist like Run-DMC, the Wu Tang Clan, A Tribe Called Quest, Slick
Rick, and Outkast innovating the style.

Hiphop reached to new heights in the 1990s with artists like 2Pac, the Notorious
B.I.G., and Snoop Dogg. They helped giving a voice to black communities that suffered
from political oppression (Iton, 2008). Unfortunately, this was also the time when a rivalry
started between artist from the East Coast and West Coast which resulted in the death of
2Pac and the Notorious B.I.G. among many others. Up to this day feuds, or beefs as they are
called by the artists, are a part of hiphop. Artists like Jay-Z, Kanye West, and Nas also
experienced this, but when they first came up in the 1990s they were praised for their
creative wordplay and revolutionary sound. The genre penetrated the mainstream radio
stations in the United States and was spread across other parts of the world by the end of the
century. Eminem was put on stage by his mentor Dr. Dre a few years later. His mesmerising
raps made him the most acclaimed white artist up to today.
The new millennium began as a logical follow up to the post Gangsta Rap era. Artists like 50 Cent and Puff Daddy successfully promoted their hardship lifestyle linking their lyrics to catchy beats. The genre continued to gain popularity as more and more artists were being booked by European festivals, with Jay-Z even becoming the first rapper to headline Glastonbury, the worlds’ largest festival at the time. At the same time, there was room for an alternative kind of hiphop that did not revolve around the gangster lifestyle. Artists like Lupe Fiasco and Lil Wayne were pioneers of the skateboard genre. They represented a subculture with clever music about every day struggles playing with heavy imagery their lyrics. Whereas the alternatives in hiphop emerged, mainstream hiphop was losing support from the fans. The industry began experimenting with new sounds using technologies to make beats and alter vocals. The attempt to modernise hiphop was not appreciated by everyone. Jay-Z responded with a track called Death Of Autotune, referring to the automated rap voices of artists like T-Pain. Nas even named his new album *Hip hop is Dead* as a comment on the development in the genre.

Despite the alleged critical state of the genre, the commercialisation continued successfully. This time, new artists were responsible for this development rather than Gangster rappers. Kanye West mixed light-hearted lyrics with complex productions and vice versa on his *Graduation*, and Lil Wayne sold over one million copies in the first week with *Tha Carter III*. They excelled with creative and innovative sound productions using numerous short samples from other tracks to produce beats. A war about copyrights was lurking around the corner but after several lawsuits it was decided that rappers could now buy samples they want to use. They were no longer limited in their creativity, albeit serious questions were raised about the originality of the music. The discussion ended when hiphop legend KRS One concluded the following: “hiphop didn’t invent anything, but it reinvented everything”. In the Southern states, artists invented the subgenre of Trap music, which centred around gritty and rebellious lyrics guided by heavy kicks and drums (Stelios, 2013).

Since the 2010s, mainstream hiphop has an enormous fanbase and hence all kinds of new artists pop up and help dominate the world’s charts. The genre has even become the number one music genre in the US, surpassing rock music (Nielsen, 2017). The big artists of today are Drake, Kendrick Lamar, Pharrell Williams, and J. Cole among many others. They continue to innovate the music by creating their own style and as a result they have acquired worldwide fame. From the Southern States, specifically in Georgia, Atlanta, a new subgenre popped up around 2012. Artists like Migos, Future, and Young Thug delivered rap
music that did not seem to be based on lyricism. Their tracks featured lyrics that were almost impossible to understand because the pronunciation was arguably poor. This did not seem to have any effect on the popularity of these artists though, as hits like ‘Bad and Boujee’ and ‘Mask Off’ reached top positions in the charts despite of the fast and warbling raps. Many other Southern rappers like 21 Savage, Lil Uzi, and Lil Yachty came up with similar rap styles and continue to deliver hits ever since. Mumble Rap has thus found its place in hiphop.

2.1.2 Controversy

No issue in contemporary hiphop has sparked off more debate than Mumble Rap. Critics are often supportive of the art or completely against it. On the one hand there are people who see Mumble Rap as an unpleasant break with the conventions of hiphop, as rap music traditionally revolves around being creative with words rather than with articulation. Lyrical Rap has received most appreciation not just for its lexical dexterity, but also because the music gave a voice to oppressed blacks in several states of the United States during the 1990s. Rap legends like Jay-Z, Tupac Shakur, and Nas were viewed as ‘‘poets that spoke for a generation of young black Americans’’ (Kane, 2018) and this gave them something that came close to biblical esteem. This is not the case for mumble rappers when hearing the majority of the critique on them. One of hiphop’s pioneers, Eminem, has said more than once that he principally refuses to listen to Mumble Rap. California’s Hopsin even took it a step further when he made a track called No Words, in which he says ‘‘These fools ain’t spittin’ no type of dope shit, but that’s not even the bad part: they’re not even saying the words anymore.’’ Many rappers as well as fans have voiced a similar opinion. Hopsin and Eminem represent an audience who believe hiphop is turning into a rudiment of what was once a precious art.

The other side of the debate tends to see Mumble Rap as something more complex, with mumble rappers being innovators of an evolving music genre. Although the warbling in songs does not seem to carry a lot of meaning, this is not necessarily true. Mumble rappers still use lyrics to express themselves, but they put more energy in being creative with their voice. This makes it harder for listeners to decipher the lyrics, but that does not say anything about the lyrical dexterity. Although not a mumble rapper himself, current rap star Kendrick Lamar called Mumble Rap ‘‘an evolution of hiphop’’ in an interview with Forbes (O’Malley Greenberg, 2017). Many people, such as acclaimed writer Shea Serrano
shares this perspective when he talks about the Atlanta artists Young Thug and Future, perhaps the most forefronted mumble rappers. He states that their music indeed fails to follow the tradition of self-expression through wordplay, but instead it represents “...the latest step in the genre’s linguistic evolution: Young Thug expresses his feelings more purely through sounds” (Locke, 2015). Lyrical articulation in contemporary hiphop seems to be diminishing then, with more and more focus given to creative articulation.

2.1.3 Defining Mumble Rap

Whereas critique on Mumble Rap is widespread, a clear definition remains absent. The term was introduced by rapper Wiz Khalifa early 2016 (Landoll, 2016), but a blueprint of what constitutes Mumble Rap has never created. This means that, at best, a description can be coined with the help of several sources as well as a phonetic analysis. First, there are several definitions of mumbling as a speech phenomenon. The Oxford Dictionary defines a mumble as “saying something indistinctly and quietly, making it difficult for others to hear” (Mumble, n.d.). Linguist Darin Flynn called it “speech that can be considered as informal, authentic, and natural” (Flynn & Morel, 2017). Academic studies often define mumbling as having to do with a reduced voice loudness (Mattys & Liss, 2008; Singer, 2009: Arvola, Arvid, Tholander, 2011). These descriptions do not cover the full phenomenon of Mumble Rap though, as artists do not necessarily differentiate in the volume of their voices other than lyrical rappers or regular speech in general. The term that is used in discussions around the hiphop style seems to deal with the pronunciation aspect of rapping only. It can be concluded that there is a mismatch in the term and the phenomenon, which could be the reason why there is still no comprehensive definition of Mumble Rap.

Mumbling in rap music mainly revolves around faded syllables and deleted segments, making the lyrics potentially unintelligible. Segmental deletion occurs frequently in regular speech (Davidson, 2006), and is thus not something unique to music. The best known example of segmental deletion in the English language is the elision of the schwa, for example in the pronunciation of ‘semester’ as /sˈmɛstər/. Davidson (2006, p.3) points out that the faster people speak the more likely it is that they do not pronounce all segments, which could mean that rapping naturally evokes segmental deletion. Yasim (1995) constructed a corpus of rap songs, and concluded that a rap track contains an average of 144 beats per minute, and each beat indicates one stressed syllable. He points out that unstressed syllables are not counted in this calculation, which indicates that the estimation of a minimal
utterance of the double amount of 144 syllables per minute is accurate, and this number is possibly even higher. This exceeds the speech rate of 220 syllables per minute that is average in spontaneous speech of speakers of American English (Michael, Maclagan, Chen, 2004, p.2). It can thus be expected that syllable- and segmental deletion is more likely to occur in rap music because of the speech rate of rapping.

A typical example of a Mumble Rap song comes from Lil Uzi’s ‘‘XO Tour Life’’, in which he raps several lines that are almost undecipherable without reading the lyrics. This becomes clear when listening to the song, but also when looking at the phonetic transcription of Uzi’s rapping. Table 1 shows the phonetic transcription of a line from the song, in which segmental deletion and fading syllables can be observed.

<table>
<thead>
<tr>
<th>Phonetic Transcription Standard AmE</th>
<th>I</th>
<th>Might</th>
<th>Blow</th>
<th>My</th>
<th>Brains</th>
<th>Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcription of the Mumble Rap</td>
<td>a</td>
<td>mar</td>
<td>blo</td>
<td>ma</td>
<td>bre</td>
<td>a:</td>
</tr>
<tr>
<td># Phoneme loss</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1. Pronunciation of a chorus line from XO Tour Life

The loss of phonemes is called deletion, and this occurs frequently in AAVE (Labov, 1995). There are two ways in which deletion occurs. First, there is vowel- or consonant deletion, which happens with the /ə/ in fear [fir] or the /ɔ/ in four [fo]. This can be the result of progressive or regressive assimilation, but also because of omissions when segments are syllable-initial or syllable-final. Deletion also occurs with syllables, changing the number of syllables in words. An example of this is the pronunciation of ‘probably’ [probbl] as [proli].
2.2.1 Emotion in Speech

The term emotion has been used interchangeably with arousal and sensation in contemporary science. That is why it has been defined and redefined throughout the years by many scholars, predominantly from the social sciences. The definition that is frequently used in linguistics is formulated by Izard (1977), who describes emotion as “… (a) the experience or conscious feeling of emotion, (b) the processes that occur in the brain and nervous system and (c) the observable expressive patterns of emotion (p.4)”. Emotions come and go, and the intensity as well as the duration can change instantly. Whereas most of the research has focused on the experience of emotions, this study will look at the expression of emotions, i.e. behavioural activity. There are several ways to express emotions, such as the use of gestures, face expression, and speech. The latter is the primary indicator of the emotional state of the speaker, and will therefore by analysed in this research.

Research into the role of emotion in speech can be done in two ways. The lexicon can be taken as an indicator of affective changes as people typically use different words when they experience arousal (Mohammad & Turney, 2010). It is also possible to look at paralinguistic features when studying emotion, which has been the main focus of research on emotion in speech since the 1980s. The seminal researches by Sherer (1984, 1986) have aimed at exploring the voice quality in the expression of basic emotions like anger, fear, and happiness. He found that loudness, fundamental frequency and voice quality change when speakers become aroused. More specifically, he showed that the tension of voice relates to the expression of emotion. Evidence suggests that a tense voice is indicative of either anger, joy, or fear, whereas a lax voice often indicates sadness. There is a lot of overlap in this though, as two opposite emotions like anger and happiness have similar voice qualities. He stressed that additional work into vocal expression and the role of emotion is required to identify patterns in voice quality. His research, however, functioned as the basis of research in emotion analysis and the role of voice and voice quality in emotional speech.

The prosodic domain has many more factors which have been researched in further studies on emotional speech. El Ayadi, Kamel, and Karray (2011) have constructed an overview of speech features. In an attempt to design a speech recognition machine, they looked at several speech studies to pinpoint the pivotal features of emotional speech. They analysed the methodology and results of 17 studies, including Chinese, German, Danish and Mandarin studies apart from English. Most of them used professional actors, or at least some sort of stimulated speech in their design. Emotions such as anger, fear, joy, and
sadness were acted out to investigate the voice characteristics. Looking at the results, El Ayadi, Kamel, and Karray (2011) found that emotions could be distinguished by their values on features related to pitch, formants, energy, timing, and articulation. This gave them the opportunity to construct a set of features which consisted of all the vocal parameters that were necessary for the design of their machine. They listed intonation, pitch levels, speech rate, speech volume and enunciation energy as the key features of emotional speech.

2.2.2 Intonation, pitch, and pitch range

Many researchers have also investigated voice characteristics of emotional speech. Mozziconacci and Hermes (1999) focussed on the role of intonation, specifically when people convey emotions and attitudes to one another. They were able to measure intonation looking at the pitch levels speakers reached with their voice. There are two ways in which they investigated intonation patterns. They first analysed the audio waves of emotional speech fragments. Three Dutch speakers were asked to enact six emotions and a neutral utterance of preselected sentences. They had intonation experts listen to these fragments next, and categorise the pitch waves using the standard of the Dutch grammar of intonation by ’t Hart, Collier, and Cohen (1990). For example, an the sounds could be labled as, ‘an early prominence-lending rise’ (1), with ‘a late prominence-lending fall’ (A). Although they did not find a clear one-to-one pattern for every emotion, there were patterns that were found significantly more often in expressing emotions than in expressing non-emotional speech, such as the patterns that end with either ‘a very late non-prominence lending rise’ (2) or ‘a very late non-prominence lending fall’ (C). This was taken as evidence of the claim that speakers use intonation patterns to convey emotions.

The second part of the analysis was done with the help of an experiment designed to test this relation. Mozziconacci and Hermes (1999) modified neutral sentences, such as ‘Zij hebben een nieuwe auto gekocht’ into an emotional sentence with a voice synthesiser, copying the patterns they found in the first test. Participants were then asked to assign emotions to varieties of the same sentence. They correctly identified the pre-intended emotion in 22.1% of the cases, which is just above chance level. The researchers explain this low score by stating that pitch levels and pitch range, which are the key components of intonation, may well indicate other emotions than the one they had constructed. They suggest that other aspects of voice such as voice quality and speaking rate contain the necessary information about emotions that participants need to correctly identify them.
These were kept constant in this study, which suggests that the assessors had too little information available to perform at a higher standard. Although these results show that there is no clear one-to-one relation between emotions and specific intonation patterns, evidence from the first analysis suggests that intonation plays a key role in expressing emotion through speech.

Rodero (2002) conducted a similar experiment with participants to see if intonation patterns link up with the perception of emotions. She had speakers read out the sentence: “It was a morning just like any other. However, the meeting would be a way to talk about life. He did not know if it would be good or bad” four times enacting four different emotions. This resulted in a total of 64 recordings which were analysed using Praat. Without any instructions about speaking at certain pitch levels, the speakers all systematically used different pitch levels for different emotions. This is the first indication of the pivotal role of pitch levels in expressing emotions. To further explore this relation, she had students without prior knowledge of voice analyses listen to the recordings and recognise the emotions. They correctly identified all four emotions well above chance level, with sadness being the emotion that was recognised most accurately. It is hence concluded that the emotional load of speech is conveyed by movements in the intonation curve and pitch levels. These finding also indicate that listeners successfully extract the information a speaker put in his voice, which explains why people use different pitch levels in communicating messages. It is thus argued that intonation and pitch levels are the pivotal factors in studying the emotional load of speech.

Further evidence on intonation and pitch levels in emotional speech comes from the medical sciences. Möbes, Joppich, Stiebritz, Dengler, and Schröder (2008) investigated voice fragments of Parkinson’s patients, as they typically produce monotonous speech as a result of their deficient nervous system. They questioned the idea that motor impairment is the reason for their way of speaking, so they designed an experiment to test this notion. Sixteen patients were included in their study who all still had a phonation capacity that was similar to a healthy person’s capacity. The patients were asked to once say and once imitate the word “Anna” in a happy, sad or neutral emotional state. The control group, which consisted of people healthy people of roughly the same age, followed the same procedure but without the imitation task. The results showed that patients spoke with a significantly lower pitch level and loudness than healthy people in the production task. When asked to imitate, however, the patients produced speech with almost equal values to healthy patients.
in all three emotional states. The results not only indicate that motor impairment cannot explain the monotony of PD speech, but they also suggest that people naturally use pitch levels and loudness to distinguish emotional speech from neutral speech.

These studies show that emotional speech is different from neutral speech, and this is best observed by looking at the pitch levels of speech utterances. The sudden drops and rises in the audio waves are claimed to be direct indicators of a change in the emotional state of the speaker. Results from Mozziconacci and Hermes (2002) suggest that the identification of emotions is a process that can be analyzed when incorporating the full range of voice characteristics. This is, however, beyond the scope of this study.

### 2.2.3 Intensity

Another way to express emotion in speech is by putting in effort in terms of intensity and loudness. Intensity relates to the amount of energy someone uses to speak, and is therefore a reflection of the affective state of the speaker (Sherer, 2000, p.225). Intensity is easy to measure but there may be some confounding variables that intervene in the measurement process, such as the acoustic characteristics of the recording environment. The importance of the methodology is something to be taken into account in the following discussion of the studies regarding speech intensity.

Bacharowski and Owen (1995) investigated speech intensity by evoking emotions on the participants’ side. They had 120 participants perform a lexical decision task while wearing a headset. They were recorded as the computer gave feedback after every block of answers, such as ‘Good Job’ or ‘Try Harder’ accompanied with an emoji. Participants in the reward group got 75% percent positive feedback and 25% negative feedback whereas participants in the failure group got feedback in a reversed ratio. The feedback was thus either congruent or incongruent with the answer, irrespective of the performance of participants. The intention was to induce negative and positive responses so the voice characteristics of their comments could be measured. They measured the fundamental frequency (pitch levels) and the perturbation of the voices by looking at the values on jitter and shimmer. These are variables related to voice quality, and thus correlate with the intensity of speaking. The results of the experiment showed a significant effect of fundamental frequency, but no effects for jitter and shimmer were found. The researchers consider the technological limitations of the experiment in their review of the unexpected insignificant results. They emphasize that jitter and shimmer are still highly likely to be
indicators of intense emotional speaking, but further research is needed to legitimise this claim. They do conclude on the basis of pitch levels that the intensity of speech is a representation of the emotional state of the speaker.

Laukka, Juslin, and Bresin (2004) further explored the acoustical properties of emotional speech. Among other factors, such as fundamental frequency and speech rate, they looked at voice intensity as a characteristic of emotion. This was operationalised as loudness and measured by using the perception of students. In line with previous research, they distinguished between several emotions (anger, fear, disgust, happiness and sadness), but they did not compare these with phrases of neutral emotion. They asked eight actors to enact the emotions either with weak intensity or strong intensity. They expected to find higher rates of loudness in strong emotion intensity speech than in weak intensity speech. 176 recording were acoustically analysed on a number of vocal cues by two groups of listeners. One group consisted of students whereas the other was formed by six experts on speech analysis. They were asked to listen to the fragments and label each fragment as either low or high in emotional intensity. Both groups consistently scored the strong intensity fragments as more emotional than the weak intensity fragments for all five emotions. The conclusion follows that voice intensity is a key factor in expressing and perceiving emotion in speech.

Cowie and Douglas-Cowie (1996) looked at energy levels in emotional speech in an attempt to explain intensity variations in the prosodic domain. They constructed five passages displaying four emotions (anger, fear, joy, sadness) and a neutral tone to function as a baseline. Forty volunteers were asked to read the passages out loud so the researchers could measure the speech volume as an indication of intensity. An ASSESS analysis showed systematic differences between intensity levels in all five passages. The passages evoking anger, fear, and happiness were read out at a significantly higher mean and median volume than the sad and neutral passage. These results provide further evidence of the relation between emotional speech and loudness as an indicator of intensity.

The results of these studies suggest that the intensity of speech correlates with the emotional state of the speaker. An emotional speaker naturally speaks louder than a calm speaker, as is the claim. The studies conducted laboratory experiments with actors or pre-selected speakers in contrast to this study, which will use audio extracts from music tracks to analyse the intensity in the voices. It is yet unclear if these voice fragments can be analysed
in a similar way, but if there are differences in intensity between the two rap styles, this should be observable in the decibel measurement of the software.

2.2.4 Pronunciation

Polzin & Waibel (1998) looked at the pronunciation in emotional expressions, as they were interested in the recognition accuracy of emotions in emotional speech. They opted to design a machine that could recognise emotions comparable to human performance, for which they needed to investigate acoustic information of emotional speech. This kind of information is different than prosodic information as it centers around influences on pronunciation rather than suprasegmental influences. They had drama students utter a number of sentences imitating four different emotions (anger, sad, happy, afraid). They had to say the sentences twice, once acting emotionally neutral, and once with the expression of an emotion. These were then transcribed to measure the number of segmental deletions and establish the relation between pronunciation accuracy and emotional speech. Results showed that the word accuracy was highest in the utterance of no emotion, with 71.9%. This differed with the word accuracy of the emotional fragments, which ranged from 45.6% in the expression of sadness until 64.2% in the angry sentences. They did not perform any statistical tests on these data as they were primarily interested in the design of the machine, but they did make claims concluding that emotion in speech has a negative effect on the word accuracy. They also conducted a small experiment asking participants to recognise the emotion and interpret the meaning of the words. The expectation followed that the word recognition rate would drop as a result of a decrease of acoustic information, i.e. the segmental deletions in emotional speech. A pretest showed that they had no problem identifying emotions and words in the neutral pronunciation. However, participants scored significantly lower in recognising the words when the sentences were uttered in an emotional state. When given the emotion beforehand, the recognition rate increased significantly, up to 70.1% for the sad utterances compared to 77.6% for the neutral. This difference is still significant, but it does indicate that knowledge of the speaker can positively affect the intelligibility of the speech utterance. That information, however, does not contradict the notion that word accuracy is negatively affected by emotional speech.

This is supported by similar research conducted by Kienast & Sendmeier (2000), who found other evidence to suggest that emotion affects articulation. They looked at consonant articulation, as they claim that the pronunciation accuracy of consonants suffers
most from an emotionally aroused speaker. Apart from the four basic emotions that Polzin & Waibel (1998) included in their experiment, the researchers added boredom, disgust, and neutral as versions of emotion. They were interested in four aspects of articulation; progressive and regressive assimilations, segmental reduction, and energy distribution in voiceless fricatives. It was hypothesised that these processes occur frequently in emotional speech after Kohler (1990), who found pronunciation effects in an investigation into German. Although the idea was attested, the researchers were interested if this evidence would hold up for other languages so they created an English corpus of audio fragments for their study. They conducted an experiment in which they asked five male and five female actors to enact the seven different emotions in a number of sentences, and analysed these fragments with the help of professional phoneticians. It was found that the average number of assimilations is significantly lower in the expression of fear, sadness, boredom, and anger than in neutral and happy utterances. Segmental reduction occurred most frequently in the sentences expressing joy, sadness, and fear, and less frequently in the neutral and happy versions. The vowel quality was also measured, and found to be better in the angry utterance than in the other utterances. These results were all significant. They suggest that the fast speech rate of the actors is responsible for the assimilations and deletions, as this is one of the most important reasons for reduction in speech (Kohler, 1990). These results are especially interesting because this study involves rap music, and the nature of rap music involves producing a lot of words in a relatively short time.

2.3 Expected Findings

Following Kienast & Sendmeier (2000) and Polzin & Waibel (1998) raises the expectation that Mumble Rap tracks have a lower pronunciation accuracy than Lyrical Rap tracks. It is also expected, after Mozziconacci and Hermes (1999) and Rodero (2002), that Mumble Rap tracks have a higher average intonation than Lyrical rap tracks. This translates into the hypotheses that Mumble Rap is higher in average pitch, and has higher average pitch ranges. Following research by Cowie and Douglas-Cowie (1996), it is hypothesised that Mumble Rap is more emotional than Lyrical Rap because of a higher intensity that is put into speech.

Emotional speech can be distinguished from neutral speech with the help of the literature. The findings provide insight into the role of pronunciation, intensity, and pitch levels in emotional speech. Most of the studies, however, base their conclusions on findings
in experiments. To see if hiphop music can function as a new corpus in emotional speech research, a selected sample of tracks are first to be analysed on the vocal parameters discussed in the previous sections. This study will thus make an attempt in finding differences in emotion in rap styles conducting a linguistic analysis. If the hypotheses are confirmed, the results could function as a basis of experimental research using different rap styles.

3. Method

3.1 Content Analysis
According to Wichmann (2000), the most accurate way to detect emotion in speech is via a linguistic analysis. The decision was hence made not to use participants’ judgements of emotional content but to perform a quantitative analysis of rap tracks to be able to make stronger claims about the emotional load. The speech in Mumble Rap tracks and Lyrical Rap tracks was compared in two ways. First, phonemic transcriptions of the texts were constructed with the help of www.topphonetics.com. This gave the transcriptions of the lyrics according to standard American English (AmE). The vocals of the raps were phonetically transcribed next to see the difference with the model pronunciation and count the number of segmental deletions. These were analysed with the help of SPSS to test the word- and sentence accuracy of the raps, i.e. the pronunciation accuracy. Second, the vocals of the tracks were analysed on several factors of emotion using the computer software Praat. This programme has proved to be an effective method in exploring the prosodic features of English speech (Gorjian, Hayati, Pourkhoni, 2013). The tracks were segmented according to the number of lines to be analyse. The average intensity, average pitch, and pitch range were the three factors extracted from Praat. These were assembled in a dataset which was analysed with a within-subjects dependant T-test in SPSS.

3.2 Eliminating Variables
Tracks were matched on several criteria to eliminate possible disturbing factors in the analysis when selecting the sample. The tracks were eligible for the sample only if their duration was between 3 and 4 minutes, as this ensured that rappers had equal time to divide their enunciation energy. As research by Heffernan (2010) found that clarity of speech is based on gender, i.e. men naturally use more voice reduction than women (p.83), this
analysis only included tracks from male rappers. Similarly, every artist was included maximally one time in the sample to have as many different voices as possible to analyse. This ensured that the results did not reflect individual speech styles, and conclusions could be drawn based on the speech style of a group of AAVE speakers. The tracks were also matched on theme/genre to limit the likelihood that the difference in lyrical content determine the emotional load of the tracks. This possible interference was negligible by selecting a sample of 10 tracks.

3.3 Sample Selection
The tracks that were included in the analysis were selected from two lists on Spotify, called ‘mumble rap’, and ‘lyrical rap’, which gave a combined list of 200 tracks. Tracks that featured female vocals in any form were eliminated from the sample, leaving 189 tracks. In order to match theme/genre, the tracks were categorised with the help of the song-text database Genius as this website gives short thematic descriptions of the tracks. After an analysis of the lyrical content, the songs were categorised in four themes: sex/love, money, introspection, and self-promotion. This gave a total of 110 tracks. To test if the tracks were rightfully labelled as either Mumble Rap or Lyrical Rap, they were subjected to a phonetic analysis. The first verse (16 lines) of these tracks were transcribed and scored on the number of instances of segmental- and syllable deletion, thereby determining if a track could be categorised as a Mumble Rap, Lyrical Rap, or Indecisive. When Mumble Rap tracks showed a minimum average of 2 instances of segmental deletion per line, they were categorised as such. Lyrical Rap tracks were selected and similarly analysed, and categorised as such when no more than an average of 1 instance of segmental deletion occurred per line.

Tracks falling under the ‘Indecisive’ category were not included in the analysis. This selection process resulted in 5 representative tracks for both categories that were matched on duration, gender, and content, making up a total of 10 tracks for the sample. The following table shows the tracks that were taken for the sample.
**Table 2.** The selected tracks from the two categories

<table>
<thead>
<tr>
<th>Lyrical Rap</th>
<th>Mumble Rap</th>
<th>Genre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outkast - Rosa Parks</td>
<td>Young Thug - No Limit</td>
<td>Sex/Love</td>
</tr>
<tr>
<td>Drake - The Zone</td>
<td>Lil Yachty - I Spy</td>
<td>Self-Promotion</td>
</tr>
<tr>
<td>Dr. Dre ft. Snoop Dogg - The Next Episode</td>
<td>Future - Mask Off</td>
<td>Drugs</td>
</tr>
<tr>
<td>Eminem - Mockingbird</td>
<td>Chief Keef - Champagne</td>
<td>Introspection</td>
</tr>
<tr>
<td>J. Cole - Forbidden Fruit</td>
<td>Migos - T-Shirt</td>
<td>Self-Promotion</td>
</tr>
</tbody>
</table>

**3.4 Procedure**

The first step of the dual analysis of the tracks involved an analysis of the phonetic transcriptions to see if there was a difference in pronunciation accuracy between Mumble Raps and Lyrical Raps. This was done by comparing every track’s standard phonetic transcriptions of American Standard English (AmE), which were provided by Top Phonetics, to the transcription of the pronunciation of the lyrics. Every sentence was scored on the number of occurrences of segmental deletion, by which the division could be made between Lyrical Rap tracks and Mumble Rap tracks. Below are two representative samples of the analysis of Lyrical Raps tracks and two examples of the analysis of Mumble Rap tracks.

*Example 1. Outkast - Rosa Parks (Lyrical Rap)*

Lyrics:
I met a gypsy and she hipped me to some life game 
To stimulate then activate the left and right brain 
Said baby boy you only funky as your last cut 
You focus on the past your ass'll be a has what 
that's one to live by or either that one to die to 
I try to just throw it at you determine your own adventure 
Andre, got to her station here's my destination
She got off the bus, the conversation lingered in my head for hours
Took a shower kinda sour cause my favorite group ain't coming with it

Standard AmE: Pronunciation:

ai met e 'dʒiːpsi ænd fi hipt mi tu 'sæm laɪf geɪm ≠
tu 'strɪmjuː,kεɪt ðen 'æktə,vert ðe lɛft ænd ræt breɪn ≠
səd 'bɛrbi bɔr ju 'ounli 'fæŋki æz juər lɛst kæt kə [1s]
ju 'fɔukəs æn ðə pæst juər æz'l bi e hæz wæt wʌ [1s]
ðæts wæn tu lɪv bær cr 'ɪdər ðæt wæn tu dær tu wʌ [1s]
ai trɑɪ tu dʒæst ðrəʊ æt ju 'dɛtərmən juər əʊn æd'ventʃər θʊʊ [1s]
'ɒn,dreɪ, gæt tu hɑr 'stɛɪfən hɜːr zæ mat ,dɛstə'neɪfən ≠
'ʃi gæt ɔ ðə bɑs, ðe,kæn'vɜːsəfɛn 'lɪŋɡərd ɪn mæt hɛd fr 'lɔvərz ≠
tʊk æ 'ʃɔvr 'kɪnda 'sɔvər kæz mæ tɪ vɪrnɪt ɡrʊp eɪnt 'kæmɪŋ wið ɪt ≠

Example 2. Usher ft. Young Thug - No Limit (Mumble Rap)

Lyrics:
You finer than wine
Baby girl I ain’t lying
Make my homies drop a dime
Commit a crime
Jeopardize my lifeline
Just to see your vital sign
Ain’t no limit, babe we do it larger
Ain’t no limit babe when you a starter
Martyr outsmart the ’Rari, ‘Rari

Standard AmE: Pronunciation:

ju 'fæɪnər əʊn wæn ju 'fæɪnər əʊn wai [1s]
'bɛrbi ɡɜrl æiənt 'lərɪŋ bɛrbi ɡɜrl æiənt 'lərɪŋ [3s]
Example 3. J Cole - Forbidden Fruit (Lyrical Rap)

Lyrics:
Ey yo, I walked through the valley of the shadow of death
When niggas hold tec's like they mad at the ref
That's why I keep a cross on my chest, either that or a vest
Do you believe that Eve had Adam in check?
And if so, you gotta expect to sip juice
From the forbidden fruit and get loose
Cole is the king, most definite
My little black book thicker than the Old Testament
Niggas pay for head but the pussy sold separate

Standard AmE:       Pronunciation:

EI Jou, ai wokt θru de 'væli AV de 'fæ, dou AV deO
wen nigez hould teks lærk deI mæd æt de ref
ðæts wair ai kip ð kräz an mair tʃ'est, 'lðer ðæt or ø vest
du ju br'liv ðæt iv hæd 'ædem in tʃ'æk?
ænd if sou, ju 'ɡate ɪk'spekt tu sip dʒus

meɪk mare homɪs dræp æ dair
KE'mit æ kraɪm
'dɛpər, daɪr maɪ 'lair, flaɪn
dʒæst tu si juə 'vartə sai
ɛin nəu 'ɪmət, bɛəd wi du ɪt 'lɑrdʒə
ɛin nəu 'ɪmət bɛəd wɛn ju æ 'stɑrtə
'mɑrər 'ɔut, smɑrt ði 'Rari, 'Rari

22
from de 'förbiden' frukt änd get lus
kouлиз de knj, moust 'defenat
mar 'litel' blæk buk 'ôiker' öen öi ould 'testament
nagez per fr hed bät de 'pusi'ould 'sépret
≠

Example 4. Lil Yachty ft. Kyle - I Spy (Mumble Rap)

Lyrics:
She said she 21, I might have to I.D. that
All my bitches come in pairs like balls in my nutsack
I remember ridin' around the city in a hatchback
Lookin' for a problem with my young goblins
I'mma send a model home with her neck throbbin'
I done made so much money that it's non-stoppin'
Got my brothers on my back like the last name
I remember tellin' everyone I couldn't be tamed
Woah, six months later I had snapped and now I'm in the game

Standard AmE:                 Pronunciation:

'ji sed 'ji twenti wan, ar mait hæv       'ji se 'ji tweti wan, ar mai hæv
tu ar.di. ðæt               tu ar.di. ðæ [4s]
ol mar 'bitjiz kam in ðerz laiK              ol mar 'bitjiz kam in ðerz laiK
bolz in maï natsæk                 bolz in maï natsæ [1s]
ar ri'member 'raidin a'raund        ar ri'membe 'raidi 'raun
ðe 'siti in e 'hætʃbæk           ðe 'siti in e 'hætʃbæ [5s]
'lukin for e 'probleam            'lukin for e 'probleam
wɪð mar jæn 'gablɪnz            wɪð mar jæn 'gablɪ [2s]
l'me send e 'mɑdel hʊm            l'me sen e 'mɑdel hʊm
wɪð hær nek 'θræbin            wɪð hær nek 'θræbin [2s]
ar dæn meid sou mɑtʃ 'mæni          ar dæn meid sou mɑtʃ 'mæni
ðæt ðɪs 'hæn,stɑrɪn            ðæt ðɪs 'hæn,stɑrɪ [2s]
3.5 Phonetic Analysis

The second way of analysis involved a phonetic analysis using Praat software. For this to be successful, the vocals of each track had to be separated from the accompanying beats. The acapella version of several tracks were available on the website acapalle4u.com. Those which were not available were manually edited using the software from phonimind.com. This gave the vocals of ten tracks in total. These were put into Praat to be able to measure intonation, pitch, and enunciation energy (labled as intensity). The tracks were segmented in the number of lines they count, the chorus excluded. This gave the possibility to measure each factor for each line and hence get a number of scores for every track.

Fig 1. The audio waves of a line from Mockingbird by Eminem
The yellow line shows the average intensity of the line measured in decibels, which received a score of 76,57 dB. Pitch was measured in Hertz on the right hand side of the lower part of the figure. As observed, this can vary between 0 and 500 Hz, and received a score of 182,2 Hz for the first line of the rap verse. The maximum pitch (151,90 Hz) and the minimum pitch (215,98) of each line were extracted from the programme, which gave a pitch range of 64,08 Hz. This measurement process was done for every line of the selected tracks which resulted in three scores for each of the 197 lines of the sample.

4. Results

4.1 Pronunciation
The total of rap lines included in the analysis is 197. The category Lyrical Rap contains 113 lines which leaves 84 lines to be analysed in the category Mumble Rap. The maximum number of segmental deletions observed in a line is 5 whereas the minimum is 0. There is no line of the category Mumble Rap that shows no segmental deletions, whereas there are 83 lines of Lyrical Rap that show no segmental deletions. The mean score for Mumble Rap was higher (2.52, SD = 1.058) than the mean score for Lyrical Rap (0.27, SD = 0.468). The data were analysed with a Pearson Chi-Square test which showed a significant difference in pronunciation accuracy between the two groups (P = .00 < 0.05).

4.2 Intensity
The highest average intensity measured in decibels (dB) was 79,94 and the lowest was 64,12. The mean intensity was slightly higher for the category Lyrical Rap (73,72 dB, SD = 3,95) than the mean intensity for category Mumble Rap (72,57 dB, SD = 3,07). The data were analysed with a T-Test for independent samples which gave an F-value of 0.179. This turned out to be not significant (P = .672 > .05) and the direction was also against expectations.

4.3 Pitch
The highest pitch measured in Hertz (Hz) was 355,4 whereas the lowest was 106,8. The mean pitch was higher for the category Mumble Rap (216,40 Hz, SD = 48,61) than the mean pitch for category Lyrical Rap (165,35 Hz, SD = 28,08). The data were again analysed with an independent samples T-Test. The F-value was 25,83 with significance of .000 which is
below the standard of $p = 0.05$. There is thus a significant difference between the groups in the Pitch.

The following graphs show examples of the developments in pitch and intensity throughout the pronunciation of a line from two tracks. The black line in figure 2 shows the pitch fluctuations between the utterance of the last and last word of the line. The red line also shows the fluctuation in intensity and hence gives an indication of the enunciation energy by the artist. The line from *Mockingbird* by Eminem shows a more stable progression of the pitch and intensity levels then the line from *No Limit* by Usher and Young Thug. As the latter falls under the category Mumble Rap, this conclusion is in line with the expectation.

*Fig 2.* The pitch and intensity of the first line from *Mockingbird* by Eminem.
4.3 Pitch Range

The highest pitch range was measured in the track *Mask Off* by Future from the category Mumble Rap, with a score of 412.37 Hz. The lowest pitch range was found in a line from *The Zone* by Drake, with a score of 7.44 Hz. The mean pitch range for the category Mumble Rap was higher (222.58 Hz, SD = 96.33) than the mean pitch range for the category Lyrical Rap (118.29 Hz, SD = 82.63). Again, an independent samples T-Test was conducted which gave an F-value of 3.866. The results were not significant as the p-value (0.051) was just above the level of significance (0.05). Figure 3 shows a schematic representation of the results of the statistics test of pitch range as well as the other three variables.

**Table 3. Results of the Analyses**

<table>
<thead>
<tr>
<th>Variable</th>
<th>p.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pronunciation Accuracy.</td>
<td>.00*</td>
</tr>
<tr>
<td>Intensity</td>
<td>.672</td>
</tr>
<tr>
<td>Pitch</td>
<td>.00*</td>
</tr>
<tr>
<td>Pitch Range</td>
<td>.051</td>
</tr>
</tbody>
</table>

*p <.05
4.4 Other Findings

The researcher noticed some unexpected differences between the audio waves of tracks from different genres during the phonetic analysis. For example, the two tracks of the introspection genre, Eminem’s *Mockingbird* and by Chief Keef’s *Champagne*, seemed more monotonous than tracks from the other genres. This suspicion was explored by looking at the mean pitch range from these tracks and comparing it with the mean pitch range of the other three genres. Three independent samples T-Tests gave a significant difference in pitch range of Introspection tracks with Sex/Love tracks ($P = .00 < .05$), with tracks from the genre Self-Promotion ($P = .001 < .05$), and with tracks of the genre Drugs ($P = .022 < .05$). It could thus be concluded that there is a significant difference in pitch range between introspective tracks and tracks revolving around drugs, sex, and self-promotion.

5. Conclusion

The main goal of this research was to investigate if Mumble Rap is a representation of a new speech phenomenon in AAVE that allows speakers to put more emotion in their speech by dropping in pronunciation accuracy. Results from the analyses suggest that Mumble Rap is more emotional than Lyrical Rap on the basis of a decreased articulation and higher average pitch levels.

The first aspect of the study dealt with the distinction between the two styles of rapping. The expectation after listening to the tracks followed that Mumble Rap tracks show a great number of segmental deletions and syllable fading, and the phonetic analysis confirmed this hypothesis with a significant difference with Lyrical Rap tracks. The five rappers that were selected for this category showed a word accuracy that could almost serve as model pronunciation of standard AmE. This contrasts with the intelligibility of the Mumble raps, which is often arguably impossible to comprehend at the first time someone listens to the tracks. The results of the first analysis suggest that this way of rapping is more emotional than rapping with a relatively clear pronunciation, which could thus be explained by suggesting that the pronunciation accuracy drops as a result of emotional speech.

The second part of the study investigated intensity, pitch, and pitch range as the vocal parameters of emotional speech. Whereas the average pitch was significantly higher in the Mumble Rap sample than in Lyrical Rap sample, the pitch range and speech intensity were not. Although the pitch ranges showed greater variations in the first sample, this turned
out to be insignificant in the analysis because the standard deviation was relatively high. This statistical problem can be resolved in the future by taking a larger sample of tracks into account.

6. **Discussion**

This study shows that there may be a system to the way mumble rappers produce speech in their tracks. The significant difference with lyrical rappers in pitch levels and pronunciation accuracy suggest that they rap with more emotion in their voices. This conclusion is based on Polzin & Waibel’s (1998) findings related to the effect of emotion on the pronunciation accuracy, and the many intonation studies that have been done over the last decades. The result of this academic focus is a very detailed picture of what emotional speech sounds like and which vocal parameters give the best insight into this. The question that this research tries to answer is if Mumble Rap is a phenomenon that can be explained with this theoretical framework. Mumble rappers are maybe not necessarily in a state of more emotion when they rap, but it could well be that their way of speaking gives prominence to their voice expression over their lexicon. The rapid rise of the genre and growing number of mumble rappers suggest that they either all imitate each other, or that they portray a way of speaking that is natural in their life. The latter can explain the fact that the pioneers of Mumble Rap all come from the same region of the United States, and is thus more likely.

The results of this study give no unambiguous answer to the posed question, but they can function as a starting point of research into this development of AAVE. Hiphop fits in this field as it has proven to be a valid representation of speech and language in many studies into variances of AAVE (Richardson, 2006). Hiphop music gives insight into communities of black America, and has an endless repository of speech which is constantly updated with new music from new artists. The need for further research is strong when researchers want to keep up to date in explaining phenomena of a lexical or para-linguistic nature. The unexpected findings regarding the correlation between introspective rap tracks and monotonous intonation is an example of the opportunities hiphop studies offer. The limitations this study has encountered are hence relevant in light of providing insight into AAVE phenomena with future research.

A few obstacles may have had an influence on validity of the outcome of the study. The selection of the sample turned out to be complicated as there was a limited amount of
tracks available on the internet from which the beat was removed, i.e. had an acapella version. This caused a problem as not every preselected track had such a version. The selection then had to be changed and the approach was reversed. First, the search for acapella versions of hiphop tracks was done, and these were subsequently labelled as belonging to either one of the two categories. This process was very time-consuming albeit the results still gave a representative sample.

An obvious explanation can be given to account for the lack of a significant relation between the rap categories and the average intensity. Contemporary music is the result of a production process rather than an audio fragment that was recorded in the same stretch of time as actual tracks. Record companies and producers can alter voices and sounds in any way possible, and do this to such an extent that rappers and singers always sound clean and pure. This has no consequences for the measurement of the average pitch levels and pitch range, but it does negatively affect the validity of the intensity measurement. The tracks are all bordering on the same volume so that listeners do not have to change the radio volume every time a song switches. The voices are thus incomparable to recorded studio material of experiments, which are frequently conducted in studies on voice characteristics. Feasible results of the analysis of intensity levels are thus not expected when dealing with music in general.

The acapella versions of the selected tracks were also not entirely consistent in the quality of the audio. Record companies and artists do not always make vocals-only files available, which encourages fans to manually construct these themselves. Some vocals where less clear than others then, but this should not have had any effect on the analysis apart from a possible interference with the intensity measurement.
References


Appendix #1

The sample selection has resulted in ten tracks which will be analysed in two ways as stated before. This appendix will show the full analysis of the pronunciation of the ten rap tracks. The lyrics and pronunciation of five tracks from the category ‘Lyrical Rap’ and five tracks from the category ‘Mumble Rap’ are phonetically compared to the standard American pronunciation (AmE). The analysis will show the number of segmental deletions in the rap styles. Vowel relaxations such as \( \text{hed} \) into 'hedid' occur in both styles and are natural in informal speaking, so they will not be counted as segmental deletions. Similarly, instances of consonant-end dropping, such as \( \text{filn} \) into \( \text{film} \), will not be analysed as segmental deletion as \( \eta \)-dropping is common in AAVE. They are no indicators of mumbling for that reason. A standard AAVE transcriber is not available at this point, so the transcriptions of the verses will be according to Standard American English (AmE) with the help of https://topphonetics.com/.

#1 J Cole - Forbidden Fruit (Self-promotion)

Ey yo, I walked through the valley of the shadow of death
When niggas hold tec's like they mad at the ref
That's why I keep a cross on my chest, either that or a vest
Do you believe that Eve had Adam in check?
And if so, you gotta expect to sip juice
From the forbidden fruit and get loose
Cole is the king, most definite
My little black book thicker than the Old Testament
Niggas pay for head but the pussy sold separate
Same bitch giving brains to the minister
The same reason they call Mr. Cee "the finisher"
Forbidden fruit, watch for the Adam's apple
Slick with words don't hate me, son
What you eat don't make me shit
And who you fuck don't make me cum
Put a price on my head won't make me run
Try to kill me but it can't be done
Cause my words gon' live forever
You put two and two together Cole here forever

Standard AmE:       Pronunciation:
&i jou, ai wokt thu ðæ 'væli æv ðe 'fæ,tou æv ðæθ
wen nigaz hould tëks laik ðei mæd æt ðe ref
ðæts war ai kip æ kros æn mai tʃest, ðiðer ðæt æ æ vëst
wok [1s] ≠
≠
The pronunciation deviates from the model transcription with 10 segmental deletions. This, however, is below the standard of calling it a Mumble Rap, as there are 19 lines in this verse. The track is thus labeled as a Lyrical rap.

#2 Eminem - Mockingbird (*Introspection*)

Hailie I know you miss your mom and I know you miss your dad
Well I'm gone but I'm trying to give you the life that I never had
I can see you're sad, even when you smile, even when you laugh
I can see it in your eyes, deep inside you want to cry
Cause you're scared, I ain't there?
Daddy's with you in your prayers
No more crying, wipe them tears
Daddy's here, no more nightmares
We gon' pull together through it, we gon' do it
Laney uncles crazy, ain't he?
Yeah but he loves you girl and you better know it
We're all we got in this world
When it spins, when it swirls
When it whirls, when it twirls
Two little beautiful girls
Lookin' puzzled, in a daze
I know it's confusing you
Daddy's always on the move, mama's always on the news
I try to keep you sheltered from it but somehow it seems
The harder that I try to do that, the more it backfires on me
All the things growing up his daddy that he had to see
Daddy don't want you to see but you see just as much as he did
We did not plan it to be this way, your mother and me
But things have gotten so bad between us
I don't see us ever being together ever again
Like we used to be when we were teenagers
But then of course everything always happens for a reason
I guess it was never meant to be
But it's just something we have no control over and that's what destiny is
But no more worries, rest your head and go to sleep
Maybe one day we'll wake up and this will all just be a dream

Standard AmE:

Heaɪ ju mɪs juər mænd ju mɪs juər dæd
wel ɜm ɡɑn bát ɜm ˈtrɛɪəŋ tu ɡɪv jʊ ðə laɪf dæt ɜ ˈnɪp hæd
aryl ɛn si ʃəd, ˈɪvɪn wɛn ju smɑɭ, ˈɪvɪn wɛn ju læf
arl ɛn si it ɪn juər əɹz, ʤɪp ɪnˈsɑɪd ju wɒnt tu kraɪ
kɑz juə ˈskɜrd, ɜ ɪnt ɔər?
ˈdændʒ wɪd ju ɪn juər ˈprɛrz
nʊr ˈmɒr ˈkraɪŋ, waɪp ərn tɜrz
ˈdændʒ hɪr, nʊr ˈmɑɪt,merz
wi ɡon' pʊl ˈteɡədər θru it, wi ɡon' du it
ˈlɛni ˈæŋkəl 'kreɪzi, ɪnt hi?
ˈjæ bæt hi ˈlaʊz ju ɡɑrl ənd ju ˈbɛtər nʊu it
wɪr ɔl wi ɡɑt ɪn ˈdɪs wɜrld
wɛn it 'spɪnz, wɛn it swɔrlz
wɛn it wɜrlz, wɛn it twɔrlz
tu 'ɪtɛl 'bjʊtɛfəl ɡərlz
ˈlʊkɪn 'pæzəld, ɪn θ ə ˈdɛrz
ar nʊ u ɪts ˈkɛŋʃuˈzɪŋ ju
ˈdændʒ bəˌweɪz ən ðə mʌv, ˈmæməz ˈbɹˌweɪz ən ðə nʌz
ar ˈtɛɪ tʃɪp ju ˈʃɛltərd fræm it bɑt 'sæm,haʊ it si:mz
ðə ˈhɑrðər ˈdæt ar ˈtɛɪ ˈtju ˈdu ˈdæt, ðə mɔr it ˈbeɪʃ,feiərəz ən mi
ɔl ðə ˈθɪŋz ˈɡrʊɪn ər hɪz ˈdændʒ ˈdæt hi hæd tu si
ˈdændʒ ˈdɔʊnt ˈwɒnt ju tu si bɑt ju si ˈdʒæst æz mʌtʃ æz hɪ dɪd
wi ˈdɪd nɑt plɛn it tu bi ðɪs wɛr, juər ˈmædər ənd mi
bɑt ˈθɪŋz hæv 'ɡɑtən sʊ ˈbæd brɪˈtwɪn əz
ar ˈdɔʊnt si əs ˈɛvər ˈbɪn teɡəˈdər ˈɛvər ˈɛɡən
lɑt kɪ juzd tu bɪ wɛn wi wɜr 'tʃiˌnɪdʒərz
bɑt ðən əv kɔrs ˈɛvri ˈθɪŋ ˈbɹˌweɪz ˈhæpənʃz fɔr ə ˈrɪzən
ar ɡɛz it wɔz ˈnɪpər ˈmɛnt tu bi
bɑt ɪts ˈdʒæst ˈsæmθɪŋ wi hæv nʊ u ˈkɛnˈ tɾɒul ˈouvər ənd ˈdæts
wɑt ˈdɛsteni ɪz
bət nəʊ mər 'wɜːriz, rɛst jʊər hɛd aænd gəʊ tu slɪp
ˈmeɪbɪ wənd də rɪ ˈwɜrk əp aænd ə dɪz wɪl ɔl dʒæst bi ə dɹɪm

Conclusion:

The perceived pronunciation is close to the model AmE pronunciation. A few differences have been observed in the pronunciation of the word ‘yours’. These count as segmental deletions, and with the adding of the other few occurrences of this phenomenon, this verse of the song Mockingbird counts a total of 7 segmental deletions. The track is thus rightfully selected for the category of Lyrical Rap based on these findings.

#3 Dr. Dre ft. Snoop Dogg - The Next Episode (Dr. Dre’s verse only) (Drugs)

Straight off the fuckin streets of C-P-T
King of the beats you ride to em in your Fleet (Fleetwood)
Or Coupe DeVille rollin on dubs
How you feelin whoopy whoop nigga whut?
Dre and Snoop chronic'ed out in the 'llac
With Doc in the back, sippin on 'gnac (yeah)
Clip in the strap, dippin through hoods (what hoods?)
Compton, Long Beach, Inglewood!
South Central out to the Westside (wessyde)
It's California Love, this California bud got a nigga gang of pub
I'm on one, I might bail up in the Century Club
With my jeans on, and my team strong
Get my drink on, and my smoke on
Then go home with, somethin to poke on (whassup bitch?)
Loc it's on for the two-triple-oh
Comin real, it's the next episode…

Standard AmE:        Pronunciation:
street ɔf ə fəkɪn strɪts əv si-pi-ti
kiŋ əv ə bɪts ju ˈrænd tu ɛm ɪn jʊər flɪt (ˈfliˌtwʊd)
ɔr kʊp dəˈvɪl ˈrælin ən dʌbz
hæ ju ɪn wʊptɪ: wʊp nɪɡe wʊt?
Dreɪ ænd snʊp kroʊnɪk'd ɔut ɪn əi ˈlæk
wɔð dæk ɪn ə ˈbæk, sɪrɪn ən 'njæək (jæ)
klɪp ɪn ə stræŋp, dɪrɪn ˈθru hʌdz (wʌt hʌdz?)
'kæmpən, ɪ ˈbɪtʃ, 'ɪŋɡelˌwʊd!
səʊəˈsɛntrəl ɔut tu ə ˈwestˈsɜːrd (wesərd)
ɪts ˌkeɪləˈfɔrnjə lɑv, əɪs ˌkeɪləˈfɔrnjə bæd ɡæt ə nɪge ɡænə əv pæb
ɑɪm ən wʌn, ə ˈmæt ˈbɛrl əp ɪn ə ˈsɛntʃəri klæb
wɔd mər ˈdʒinz ən, ænd mər ˈtɪm strɔŋ
ɡet mər dɹɪŋk ən, ænd mər smʊək ən
ðən ɡu ˈhəʊm wɔd, ˈsæmθɪn tu pʊək ən (wəsʊp ˈbɪtʃ?)

mo [1s]
Lak its an for de tu-'tripel-ou
Comin real, its de nekt 'epa,soud...

Conclusion:

This track matches the model AmE pronunciation almost perfectly, apart from 2 instances of segmental deletion in the last three lines. It is categorised as a Lyrical Rap for this reason.

#4 Outkast - Rosa Parks (Andre 3000’s verse only) (Sex/Love)

I met a gypsy and she hipped me to some life game
To stimulate then activate the left and right brain
Said baby boy you only funky as your last cut
You focus on the past your ass' is a has what
that's one to live by or either that one to die to
I try to just throw it at you determine your own adventure
Andre, got to her station here's my destination
She got off the bus, the conversation lingered in my head for hours
Took a shower kinda sour cause my favorite group ain't coming with it
But I'm with ya cause you probably going through it anyway
But anyhow when in doubt went on out and bought it
Cause I thought it would be jamming but examine all the flawsky-wawsky
Awfully, it's sad and it's costly, but that's all she wrote
And I hope I never have to float in that boat
Up shit creek it's weak is the last quote
That I want to hear when I'm going down when all's said and done
And we got a new joe in town
When the record player get to skipping and slowing down
All yawl can say is them niggas earned that crown but until then

Standard AmE:  

Pronunciation:

'æt mət ə 'dʒɪpsi ənd fi hɪpt mi tu ˈsɑm laɪf ɡeɪm  ≠
'tu 'strɪnʤə,liət ən ˈækɪ,vert ə de ˈlɛft ənd rɑɪt ˈbreɪn  ≠
'sed ˈbeɪbi bɔɪ ju ˈounli ˈlɑŋki əz jʊər ˈleɪst ˈkæt  kaɪ [1s]
ju ˈfoukəs ən ə pɛəst jʊər əs'l bi ə hæz wət  waɪ [1s]
ˈdæts ˈwæn tu ˈɪv bɑɪ ɔr ˈlɪdər ˈdæt ˈwæn tu ˈdeɪr tu  waɪ [1s]
æ t ræɪ tu ˈdʒɑst θou ɪt æt ju ˈdɛtərmən jʊər ˈoun əɛd'ventʃər θou [1s]
'ɑn,dreɪ, ɡæt tu hɑr 'stiːfən hɜr ˈmɑɹ,ˈdɛstə'neɪʃən  ≠
'jɪ ɡæt ɔf ə bæs, ə ,ˈkænvrˌseɪʃən ˈlɪŋɡərd ɪn ˈmeɪ ʰæd fər ˈævərər  ≠
tu kə ˈfɔʊər ˈkɪnda ˈsɔʊər kɑz ˈmæt ˈfeɪvərɪt ˈgrʊp ˈɛnt ˈkæmɪŋ ˈwɪd ɪt  ≠
'bæt əɪm ˈwɪd ju ju ˈkeɪ ˈprəʊbəbli ˈɡəʊɪŋ ˈθru ɪt ˈɛnɪ,ˈwɛr  ≠
bæt ˈɛnɪˌhɔː ˈwɛn in ˈdɔt ˈwɜnt ən ˈaʊt əɛnd bæt ɪt  wɛn [1s]
ˈkɑz ər ˈθɒt ɪt ˈwʊd bi ˈdʒæmɪŋ bɑt ɪɡ'zæmɪŋ ɔl ə fɪske-ˈwɔskɪ  ≠
'ɑflɪ, ɪts ˈsæd ənd ɪts ˈkɑstli, bæt ˈdæts ɔl ˈfi ˈrʌt  ≠
Whoa, all these broken hearts on that pole
Man if pole dancers an art you know how many fuckin' artists I know
Got some new bills in the mail
Got some big favours I owe
Got some good things ahead of me
When these bad bitches let go
Well, girl let's go
Walk your broken heart through that door
Sit yo sexy ass on that couch
Wipe that lipstick off of your mouth
I take it slow
She in love with my crew
She said make enough so I can try some
I thought taking drugs jus' ain't you
Yeah girl, just be you
And I do this shit for my hometown
It being going down it ain't new
That's that north north, that up top
That O-V-O and that X-O
Your girlfriend at our next show
But it's all good, don't stress though
First night fuck, never really planned it
Take a deep breath, no need to panic
Lips so French, ass so Spanish
You don't really like attention, I don't know if she gon' manage out here
But she got me all up in my zone
Said she like the view I got in this place
Shit I did all of that on my own

Standard AmE:

<table>
<thead>
<tr>
<th>Word</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>whoa</td>
<td>wou, ɔl ðiz 'broukən harts an ədət pauəl</td>
</tr>
<tr>
<td>man if pole</td>
<td>'dænsərz ən ərt ju nɔu</td>
</tr>
<tr>
<td>ha</td>
<td>'mɛni 'fækin 'ɑrtɪsts æt nɔu</td>
</tr>
<tr>
<td>get sam nu bɪtʃ ɪn ədə mɛəl</td>
<td></td>
</tr>
<tr>
<td>get sam big feɪnərs æt ɔu</td>
<td></td>
</tr>
<tr>
<td>get sam gud ˈθɪŋz əˈhɛd æn mi</td>
<td>ʌ [1s]</td>
</tr>
</tbody>
</table>

#5 Drake ft. The Weekend - The Zone (Drake's verse only) (Self-Promotion)
Never been a gopher (no) but I always been a soldier
Stop all that flexin’, young nigga don’t wanna go there
I get high on my own, sir, heard you gon’ clone sir

Real whole pies (whoa), all time high (high)
Do it for the culture (culture), they gon’ bite like vultures (vultures)
Way back when I was trappin’ out Toyotas (skrr, skrr)
I’m a hit the gas (gas), twelve can’t pull me over (twelve)
Space coupe, Quavo Yoda, pourin’, drankin’ sodas

I get high on my own, sir, heard you gon’ clone sir
Stop all that flexin’, young nigga don’t wanna go there
Never been a gopher (no) but I always been a soldier
Young niggas in the cut, posted like a vulture
Divin' off the stage in the crowd, it's a mosh pit
Yeah, shawty bad but she broke cause she don't own shit
Mama asked me "Son, when the trappin' gon' quit?"
I been ridin' round through the city in my new bitch

Standard AmE

nain'tin 'hainti-fairv ('haintiz),
tu 'θæuzənd fai̯v (tu-'θæuzəndz)
sin ɪ wɜd mar aɪz (sin ɪ),
doup stil ə'laɪv (doup)
riel məb taɪz (məb),
riel frəɡ aɪz (frəg)
riel hoo̯l paɪz (wʊ),
ɔl tɔm hɛə (hət)
du ɪt fa ðə 'kɑltfə (ˈkɑltfə),
dəi gon' bæt laɪk ˈvɑltʃərz (ˈvɑltʃərz)

wei bæk wɛn aɪ wæz 'træpɪn ɔut ,
bɔ'outaz (skrr, skrr)
l'ma hɪt ðə ɡeəs (ɡeəs),
twɛlv kænt pɔl mi 'ˈouəər (twɛlv)
spərs kʊp, Quavo Yoda,
ˈpɔrɪn, 'draɛŋkɪn 'soudez
ar ɡet hər ən mar oun, sər,
hɑrd ju gon' kləʊn sər
stɑp ɔl ˈdæt 'fleksɪn,
jɑn ˈnɪɡɡə ˈdɑʊnt ˈwænə ˈgou ðər
ˈnɛvr ˈbɪn ə 'ɡoʊfər (nou)
bɑt əɪ ˈbɛlˌwiəz ˈbɪn ə 'souldʒər
jɑn ˈnɪɡɡəs ɪn ðə kæt,
ˈpɔʊstɪd lɑɪk ə 'vɑltʃər
ˈdɑrvɪn ʃɪ ðə stɛrdʒ ɪn ðə kraʊd,
ɪts ə mɒʃ pɪt
jɛ, shawty bæd
bɑt ʃi ˈbruk kæz ʃi ˈdʊnt oun fɪt

'mɑmə æskt mi "sɑn,
wɛn ðə 'træpɪn ˈɡɒn' kwɪt?"
ɑr ˈbɪn ˈraɪdɪn ˈraʊnd ˈθru
ðə 'sɪtɪ ɪn mɑr nu ˈbɪtʃ

Pronunciation

nain'ti 'haini-fairv ('hainiz)
tu 'θæuzən fai̯v (tu-'θæuzəndz) [3s]
sin ɪ wi mar aɪz (sin ɪ),
doup stil ə'laɪv (doup) [3s]
riel ma ˈtɜz (mə),
riel frəɡ aɪz (frə) [3s]
riel hoo̯l paɪz (wʊ)
ɔl tɔm hɛə (hə) [1s]
du ɪt ðə 'kɑltfə (ˈkɑltfə)
dəi gon' bæt laɪk ˈvɑltʃəz (ˈvɑltʃəz)
[s4s]
wei bæk wɛn aɪ wæz 'træpɪn ɔut
bɔ'outaz (skrr, skrr) [2s]
l'ma hɪt ðə ɡeəs (ɡeəs)
twɛl kænt pɔl mi 'ˈouəər (twɛl) [4s]
spərs kʊp, Quavo Yoda
ˈpɔrɪ, 'draɛŋkɪ 'soudez [3s]
ar ɡe hɑr ən mar oun, sər,
hɑrd ju gon' kləʊn sər [1s]
sta ɔl ˈdæt 'fleksɪ,
jɑn ˈnɪɡɡəs ˈdɑʊnt ˈwænə ˈgou ðər [4s]
ˈnɛvər ˈbɪn ə 'ɡoʊfər (nou)
bɑt əɪ ˈbɛlˌwiəz ˈbɪn ə 'souldʒər [2s]
jɑn ˈnɪɡɡəs ɪn ðə kɑ
ˈpɔʊstɪd lɑɪk ə 'vɑltʃər [1s]
ˈdɑrvɪn ʃɪ ðə stɛrdʒ ɪn ðə kraʊd,
ɪts ə mɒʃ pɪ [1s]
jɛ, shawty bæd
bɑt ʃi ˈbruk kæz ʃi ˈdʊnt oun fɪ
[3s]
'mɑmə æskt mi "sɑ,
wɛn ðə 'træpɪn ˈɡɒn' kwɪt?" [3s]
ɑr ˈbɪn ˈraɪdɪn ˈraʊnd ˈθu
ðə 'sɪtɪ ɪn mɑr nu ˈbɪtʃ [2s]
Conclusion:

The perceived pronunciation deviates from the model AmE pronunciation as expected from the selection. There are 16 lines in this verse that show 40 occurrences of segmental deletions in the pronunciation, which validates the track as a Mumble Rap.

#7 Chief Keef ft. Nervo - Champagne (Introspection)

You don't owe me anything, thought you should know that
All the shit you've put up with, girl, I deserve a toe tag
Sunny days, clear skies, rainy days, cold nights
You was there for me when I was snowed in, you know that I know right
But tonight's your night (bang, bang)
You was with me when the po put you and me in the coupe
You the only girl that can get a nigga in a suit
Drop top riding 'cause you're so exotic
You won't reply to gold 'cause you're like a diamond (skrt skrt)
We can go out shopping, pop tags
Since we got assistants we can leave the mall without bags
Take me to the time that we first met
Take me to the time that we first met

Standard AmE:

ju doun ou mi 'eniθiŋ,
θat ju fud nou dæt
ɔl de fıt juv put ap wið,
gɔrl, ai dr'zɔrv a tou tæŋ
'sæni dɛiz, klər skærz,
'reini dɛiz, kould narts
ju waz ðer fɔr mi wɛn ai waz snəud ɪŋ,
ju nɔu dæt ai nɔu ræt
bæt te'netis jʊər nart (bæŋ, bæŋ)
ju waz wið mi wɛn ə de pɔu
put ju ænd mi ɪn ə ku setUp ju ænd mi ɪn ə ku
ju ɔi 'ounli gɔrl dæt kæn
get a nɪgə ɪn ə su
drap tə 'raɪdɪŋ kez
jur sou ɪg'zɔtɪk
u wəʊnt nɪ'plær tu gould
ekæz jur lɑɪk ə 'dɑrmənd
wi kæn gou əut ʃɑpɪŋ, pæp tægz
sɪns wi gæt ə'sɪstənts
wi kæn liv ə de məl wɪˈθɔut bægz

Pronunciation:

ju doun ou mi 'eniθiŋ,
θat ju fud nou dæ [2s]
ɔl de fıt juv put ə wi,
gɔrl, ai dr'zɔrv a tou tæ [3s]
'sæni dɛiz, klər skærz,
'reini dɛiz, kould narts [1s]
ju waz ðer fe mi wɛn ai waz snəu ɪŋ
ju nɔu dæt ai nɔu ræt [3s]
bæt te'netis jʊər nær (bæŋ, bæŋ) [2s]
ju waz wið mi wɛn de pɔu
put ju ænd mi ɪn de kup [1s]
ju ɔi 'ounli gɔrl dæt kæn
get a nɪə ɪn ə su [3s]
drap tə 'raɪdɪŋ kez
jur sou ɪg'zɔtɪ [2s]
u wəʊnt nɪ'plær tu gould
kez jur lɑɪk ə 'dɑrmən [4s]
wi kæn gou əut ʃɑpɪŋ, pæp tæz [3s]
sɪns wi gæt ə'sɪstənts
wi kæn liv ə de məl wɪˈθɔu bæz [2s]
teik mi tu ðə taim ðæt wi fərst mæt

Conclusion:

The verse contains 13 rap lines in which a total of 30 segmental deletions occur. The track is thus labeled as a Mumble Rap track.

#8 Future - Mask Off (Drugs)

Two cups, toast up with the gang
From food stamps to a whole 'nother domain
Out the bottom, I'm a livin' proof (Super)
They compromising, half a million on the coupe
Drug houses, lookin' like Peru
Graduated, I was overdue
Pink Molly, I can barely move
Ask about me, I'm gon' bust a move
Rick James, 33 chains
Ocean air, crusin' Biscayne
Top off, that's a liability
Hit the gas, boostin' my adrenaline

Standard AmE:

tu kaps, toust ap wið ðə gæn
frəm fu stæmps tu ðə houl 'nʌðər dəʊ'meɪn
aut ðə 'bætəm, aɪm ðə 'ɪvɪn prʌf ('super)
dəi 'kæmprə,meɪzɪn, hæf ðə 'mɪljən ən ðə kup
dræg 'hɔsəz, 'lʊkɪn lærk pə'rʊ
'grædjuə'tɪd, æ zi wæz 'ouvər'du
prɪŋk 'mæli, æ zə ˈbeəli muv
æsk ə'bæut mi, aɪm gən' bæst ðə muv
rɪk dʒeɪmz, 33 fərnt
'ɒvən ər, ˈkruːzin brɪskeɪn
tɔp ɔɪ, əʊəst ə,laɪər'bɪlɪtɪ
hit ðə gæs, 'bʌstɪn maɪ ðrə'nələn

Pronunciation:

tu kaps, toust ə wi ðə gæ [4s]
frəm fu stæmps tu ðə houl 'nʌðə dəʊ'meɪn [3s]
aut ðə 'bætə, aɪm ðə 'ɪvɪ pru ('supe) [4s]
dəi 'kæmprəmaizi, hæf ðə 'mɪljən ə ðə ku [3s]
drə 'hɔsəz, 'lʊkɪ lærk pə'rʊ [s2]
grædjuə'tɪt, æ zi wæz 'ouvər'du [1s]
prɪŋk 'mæli, æ zə ˈbeəli mu [2s]
æsk ə'bæu mi, aɪm gən' bæst ðə mu [2s]
rɪk dʒeɪz, 33 fərnt [2s]
'ɒvən ər, ˈkruːzin brɪskeɪ [1s]
tɔp ɔ, əʊəs ðə, laɪər'bɪlɪtɪ [2s]
hit ðə gæs, 'bʌstɪn maɪ ðrə'nələ [2s]
Conclusion

The first verse of this tracks counts 13 lines in which 28 instances of segmental deletion occur. The pronunciation is significantly different from the model AmE pronunciation and is therefore rightfully labeled as a Mumble Rap track.

#9 Usher ft. Young Thug - No Limit (Young Thug’s verse only) (Sex/Love)

You finer than wine
Baby girl I ain’t lying
Make my homies drop a dime
Commit a crime
Jeopardize my lifeline
Just to see your vital sign
Ain’t no limit, babe we do it larger
Ain’t no limit babe when you a starter
Martyr outsmart the 'Rari, 'Rari
Fill the session with Bacardi Barbies
Kerosene, kerosene
Promise spin, washing machine
Thin waisted primadonna
Never limit, I’m a stoner
Tinted out, them never rentals
Fuck them boys, they always get us
Bottle and rag denim
I’ma spend my night with ‘dem
I could put karats all over you
Karats all over you
Never mind, we only poppin’ shit
Man I been getting high with these fools
And she said all her friends fake, was solo rocking
And she a real bad bad bitch, she ain’t gotta Photoshop it

Standard AmE:  

ju 'fɑɪnər ðæn wain  
'berbi gɜrl ər ɛnt 'lærən  
meɪk mer əʊ həməs ˈdræp ə dæm  
kə'mit ə ˈkrɑːm  
ˈdʒɛpə,deɪz ɹer ˈlɑr ili  
dʒʌst tu ɹi juer ˈvɑːtəl sɑːrn  
ɛɪnt nuu ˈɪmət, ˈbeɪb wi ðu ɪt 'lɑrdʒər  
ɛɪnt nuu ˈɪmət ˈbeɪb wɛn ju ə 'stɑːtər

Pronunciation:  

ju 'fɑɪnər ðæn wai [1s]  
beɪbi ɡɜrl ər ɛn 'lær [3s]  
meɪk ɹer həməs ˈdræp ə dər [2s]  
kə'mit ə ˈkrɑːm [1s]  
ˈdʒɛpə,dɛɹ ɹer ˈlɑr ʃlɑːn  
dʒʌst tu ɹi ju ɹi ˈvɑːtəl sɑːn  
ɛɪn nuu ˈɪmə, ˈbeɪb wi ðu ɪt 'lɑdʒə [4s]  
ɛɪn nuu ˈɪmə bei wɛn ju ə 'stætə [4s]
Went from "Damn this nigga lame" to remember my name

Woah, six months later I had snapped and now I'm in the game

I remember tellin' everyone I couldn't be tamed

Got my brothers on my back like the last name

I done made so much money that it's non-stoppin'

She said she 21, I might have to I.D. that

All my bitches come in pairs like balls in my nutsack

She said she 21, I might have to I.D. that

I remember riding around the city in a hatchback

Lookin' for a problem with my young goblins

I'mma send a model home with her neck throbbin'

I done made so much money that it's non-stoppin'

Got my brothers on my back like the last name

I remember tellin' everyone I couldn't be tamed

Woah, six months later I had snapped and now I'm in the game

Went from fake chains to diamonds in another lane

Went from "Can you take me here?" to screening out the lane

Went from "Damn this nigga lame" to remember my name
So I remember all the people who ain't fucked with me
They went to college, now all them niggas is history
Upgraded from gold to diamonds in my teeth
Riding deep in the van, like we lookin' for a mystery (Raggy!)
So don't fuck with me, no, lately I been livin' life luxury
Boat and Kyle stick together like piano keys
And on my mother's mother I won't fuck a bitch without a damn rubber

Wou, síks māṉs 'leite ar hæd
snæpt ðæd nau aɪm in ðə gɪrm
went fræm fɛək tʃeɪnt tu '
dəɪmændz in e'nædær lɛn
went fræm 'kænd ju tərk mi hɪʔ'
 tu 'skrɪnət ət ðə lɛɪn
went fræm "dæm ðɪs nɪɡə lɛɪm"
tu rɪ'mɛmbər maɪ nɛɪm
sou ar rɪ'mɛmbər ol ðə 'pɪpɛl
huænt fækt wɪð mi
ˈðɛɾ wɛnt tu 'kɔɪrdz,
nau ðə ʌm nɪɡə iz 'hɪstəri
ep'greیدed frɒm gʊld
tu 'dəɪmændz in mæɪ tɪð
ˈrɛɪnd ɪŋ ðə væn,
lærk wi 'lʊkɪŋ fɔr ə 'mɪstəri
sou dɔʊnt fæk wɪð mi, nʊu,

Wou, síks māṉs 'leite ar hæd
snæpt ðæd nau aɪm in ðə gɪrm [4s]
went fræm fɛək tʃeɪnt tu 'dəɪmændz in e'nædær lɛɪn [3s]
went fræm "kænd ju tərk mi hɪʔ?" tu 'skrɪnət ət ðə lɛɪn [3s]
went fræm "dæm ðɪs nɪɡə lɛɪm" tu rɪ'mɛmbər maɪ nɛɪm [2s]
sou ar rɪ'mɛmbər ol ðə 'pɪpɛl
huænt fækt wɪð mi [3s]
ˈðɛɾ wɛnt tu 'kɔɪrdz,
nau ðə ʌm nɪɡə iz 'hɪstəri [1s]
ep'greɪdəd frɒm gʊld
tu 'dəɪmændz in mæɪ tɪð [4s]
ˈrɛɪnd ɪŋ ðə væn,
lærk wi 'lʊkɪŋ fɔr ə 'mɪstəri [1s]
sou dɔʊnt fæk wɪð mi, nʊu,
Conclusion:

The 19 lines of Lil Yachty’s verse show 53 instances of segmental deletion and is thus categorised as a Mumble Rap track.