

# Regional Variation in West and East Coast African-American English Prosody and Rap Flows

Language and Speech  
2020, Vol. 63(4) 713–745  
© The Author(s) 2019



Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/0023830919881479  
journals.sagepub.com/home/las



**Steven Gilbers** 

Department of Applied Linguistics, Center for Language and Cognition Groningen, Research School of Behavioural and Cognitive Neurosciences, University of Groningen, the Netherlands

**Nienke Hoeksema**

Neurobiology of Language Department, Max Planck Institute for Psycholinguistics, the Netherlands

**Kees de Bot**

**Wander Lowie** 

Department of Applied Linguistics, Center for Language and Cognition Groningen, Research School of Behavioural and Cognitive Neurosciences, University of Groningen, the Netherlands

## Abstract

Regional variation in African-American English (AAE) is especially salient to its speakers involved with hip-hop culture, as hip-hop assigns great importance to regional identity and regional accents are a key means of expressing regional identity. However, little is known about AAE regional variation regarding prosodic rhythm and melody. In hip-hop music, regional variation can also be observed, with different regions' rap performances being characterized by distinct "flows" (i.e., rhythmic and melodic delivery), an observation which has not been quantitatively investigated yet. This study concerns regional variation in AAE speech and rap, specifically regarding the United States' East and West Coasts. It investigates how East Coast and West Coast AAE prosody are distinct, how East Coast and West Coast rap flows differ, and whether the two domains follow a similar pattern: more rhythmic and melodic variation on the West Coast compared to the East Coast for both speech and rap. To this end, free speech and rap recordings of 16 prominent African-American members of the East Coast and West Coast hip-hop communities were phonetically analyzed regarding rhythm (e.g., syllable isochrony and musical timing) and melody (i.e., pitch fluctuation) using a combination of existing and novel methodological approaches. The results mostly confirm the hypotheses that East Coast AAE speech and rap are less rhythmically diverse and more monotone than West Coast AAE speech and rap, respectively. They also

## Corresponding author:

Steven Gilbers, Department of Applied Linguistics, Center for Language and Cognition Groningen, Research School of Behavioural and Cognitive Neurosciences, University of Groningen, Oude Kijk in 't Jatstraat 26, Groningen, 9712EK, the Netherlands.

Email: s.r.k.i.gilbers@rug.nl

show that regional variation in AAE prosody and rap flows pattern in similar ways, suggesting a connection between rhythm and melody in language and music.

### Keywords

Prosody, rap flows, African-American English, hip-hop, regional variation

## Introduction

Though regional variation in African-American English (AAE) remains a highly understudied subject—especially considering that AAE is among the most frequently investigated varieties of English in sociolinguistic history (Schneider, 1996; Thomas & Carter, 2006)—several studies have found that AAE exhibits significant regional variation (Wolfram, 2007; Wolfram & Kohn, 2015). It has also been shown that regional variation in AAE is salient to its speakers, especially those involved with hip-hop culture, in which great importance is assigned to region and place (Forman, 2002; Hess, 2009; Morgan, 1993, 1998, 2001, 2002). In fact, speakers of AAE are aware of their regional accent features, and hip-hop aficionados make active use of them to express their regional identity and affiliations (Gilbers, 2018; Morgan, 2001, 2002). In hip-hop music, regional variation can also be observed with different regions' rap performances being characterized by distinct styles or "flows": the rhythmic and melodic aspects of a rap performance. As with regional accent features, rappers consciously use these regional flows to represent for their cities or larger regions (Alim, 2004; Forman, 2000, 2002), but how exactly these regional flows differ from each other has not been studied yet. From a linguistic perspective, what is striking about the concept of flow is how similar it is to the concept of prosody: the rhythmic and melodic (i.e., intonational) aspects of spoken language.

The present study, which operates on the intersection of variationist sociophonetics, musicology, and hip-hop studies, is concerned with regional variation in AAE speech prosody and rap flows, specifically regarding the United States' East and West Coasts. It is also concerned with how this relates to research on the connection between language and music. The study's main aims are to investigate whether East Coast and West Coast AAE differ regarding prosodic rhythm and melody, whether East Coast and West Coast rap flows are rhythmically and melodically distinct, and whether regional variation in speech and rap are related to each other in that they follow the same patterns, ideas which have not been (quantitatively) examined thus far. In addition, the article proposes a new methodology for studying melodic and rhythmic variation in speech and rap.

In the subsections below, an in-depth discussion of AAE prosody, rap flows, and the possible language–music connection in a hip-hop context will be presented to provide context for the study's aims and hypotheses. Throughout these discussions, particular attention will be paid to regional variation, especially with regards to the East and West Coasts.

### 1.1 AAE prosody

Prosody is a greatly understudied aspect of AAE (Thomas, 2015; Thomas & Carter, 2006). The main study on AAE prosodic rhythm was conducted by Thomas and Carter (2006), who examined the speech rhythms of African-Americans from North Carolina along the stress-timed/syllable-timed continuum. On this continuum, language varieties whose stressed and unstressed syllables are of substantially different lengths (e.g., Dutch) are considered to be on the stress-timed end of the spectrum, whereas language varieties whose stressed and unstressed syllables are of similar length (e.g., French) are considered to be on the syllable-timed end of the spectrum. Thomas and

Carter found that the speech of African-American North Carolinians, like that of European-American North Carolinians, was highly stress-timed in nature (i.e., their stressed syllables were much longer than their unstressed ones). They also compared these recordings with recordings of African-Americans born in the mid-19th century and found that the latter were much more syllable-timed (i.e., stressed and unstressed syllables' durations were relatively similar), suggesting that AAE's syllables had grown to become less isochronous over time. However, considering that only North Carolinian speakers were analyzed, the Thomas and Carter (2006) study presumably does not constitute a very representative sample of AAE speakers. Regarding prosodic melody, AAE speakers' pitch range has been found to be greater than that of European-American English speakers in free speech (e.g., Hudson & Holbrook, 1981, 1982; Jun & Foreman, 1996; Loman, 1975; Tarone, 1973) but smaller in read speech (e.g., Cole, Thomas, Britt, & Coggs, 2005; Goodwin, Goodwin, & Yaeger-Dror, 2002; Thomas, 1999), and it has also been shown that AAE intonation patterns within and at the end of utterances are distinct from those of other varieties of English, which provides listeners with auditory cues to distinguish African-Americans from European-Americans (e.g., Foreman, 2000; Holliday, 2016; Thomas, Lass, & Carpenter, 2010).

*1.1.1 Regional variation in AAE prosody.* Though AAE prosody has been studied in a wide variety of locales—for example, rural communities in Texas (Thomas, 1999), North Carolina (Cole et al., 2005; Wolfram & Thomas, 2002), South Carolina (Goodwin et al., 2002), and Louisiana (Green, 2002) as well as urban communities in Washington, DC (Loman, 1975), Raleigh (McLarty, 2011), Seattle (Tarone, 1973), Los Angeles (Foreman, 2000; Jun & Foreman, 1996), and Pittsburgh (Gooden, 2009)—regional variation has rarely, if ever, been the focal point of AAE prosody research. Moreover, because of diverging methodologies, comparing the findings of the aforementioned studies is difficult (Thomas, 2015), and major dialect regions have hardly been examined at all. As a result, while it is likely that “the intonational features that are diagnostically African American are subject to ( . . . ) regional variation” (Thomas, 2015, p. 420), we have no clear idea of how regional variation surfaces in AAE prosody.

## 1.2 Flow

“Flow” is to rap what prosody is to language; both are concerned with the rhythmic and melodic aspects of their respective domains. The concept of flow has many definitions (e.g., Adams, 2009, 2015; Connor, 2018; Edwards, 2013; Krims, 2000), but most rap scholars today agree that “flow” refers to the rhythmic and melodic aspects of a rap performance. As such, analyzing rap flows is essentially no different from analyzing any other type of musical expression that revolves around rhythm and pitch. However, compared to other, more traditional forms of music, rhythm and (especially) melody are treated somewhat differently in rap music. In addition, rhythm and melody are described using different terminologies in scholarly work on rap music compared to the literature on more traditional forms of music. For this reason, the present subsection is dedicated to outlining the main rhythmic and melodic aspects of flow. This is followed by a discussion of what is currently known regarding the differences between East Coast and West Coast rap flows.

*1.2.1 Flow's rhythmic dimension.* In terms of rhythm, many variables come into play when it comes to flow. For instance, although nearly all of rap music is performed in a 4/4-time signature, the number of syllables, the length of these notes (e.g., eighth notes, sixteenth notes, triplet notes, etc.), and the number of musical rests in a bar can vary greatly between flows. Moreover, flows are largely defined by musical microtiming: it makes a significant difference in terms of flow whether rappers perform syllables right on the beat or instead slightly before or after. Arguably the most

important difference in terms of rhythm between rapping and singing, however, is that rap vocals primarily serve a percussive function. In other words, rap is, generally speaking, more concerned with the temporal location of where notes land than it is with the duration of said notes (Edwards, 2013). As such, it is often more insightful to analyze rap vocals as percussion (like one would analyze drum patterns) than it is to analyze rap vocals in terms of, for instance, note duration (as is common in musicological work on most Western (vocal) music). As Edwards writes:

A lot of the rhythmic techniques in rapping are closer to percussion and drumming than they are to traditional poetic techniques; (. . .) rhythmically [rapping]’s very similar to percussion. (. . .) Therefore, most of the terms [used to describe rapping] come from percussion and drumming rather than poetry analysis. (Edwards, 2013, p. 3)

Finally, while this is outside of the scope of the current study, it is worth noting that rhythmic accents and the metrical characteristics of a rap performance can distinguish one flow from another as well. In this regard, rhyme schemes play an important part, as rhymed units add an extra level of rhythmic depth: recurring sequences of rhymes add rhythmic accents that distinguish the flow of bars that are otherwise identical in terms of musical notation.

*1.2.2 Flow’s melodic dimension.* Despite the fact that rap is usually distinguished from singing based on its lack of melody and its greater focus on rhythm (which is why it is considered more closely related to speech than singing is), pitch contours play a central role in flow (e.g., Adams, 2015; Connor, 2018). Most rap performances are not melodic in the traditional sense, meaning that rappers generally neither “hit” notes nor match their voice’s pitch to the song’s key. In fact, when people noticed how the rapper Rakim would rap in the same key as the music, they let him know this was highly unusual, as Rakim himself explained on *The Cipher Show* (Setaro, Kross, & Griffin, 2014):

I noticed that, like, whenever I wrote rhymes, I kind of wrote them in the key, (. . .) At first, I ain’t really know I was doing it. Then somebody, you know, brought it to my attention, like, “yo, man, you hear how every time the key of the track change, [so do] you?” And I’m like, “ain’t I supposed to do that?” (. . .) And then, you know, the more I started paying attention, really too many people don’t really do that.

While Rakim forms an exception to the rule, all rappers manipulate pitch in some shape or form on the speech–singing continuum, and the idea that rap is marked by monotony, which can often be found in criticisms of rap music (e.g., Kelley, 2011), is unequivocally false. To illustrate, consider the late rapper Guru of the group Gang Starr, who was endearingly called “The King of Monotone” by his fans because, compared to his peers, he displayed relatively little pitch inflection. While his nickname was technically inaccurate—Guru’s rapping was most definitely not purely monotone in an absolute, robot-like sense—the fact that he was even given this moniker implies that a flow marked by relative monotony is something highly unusual in rap. Scholars such as Adams (2015) and Connor (2018) have compared rap’s use of melody to the way artists such as Bob Dylan employ the technique of *Sprechstimme*, a vocal technique which does not emphasize any notes in particular in its pitch contours.

As in speech, raising or lowering one’s pitch at specific moments during a rap performance can serve many discursive purposes, such as emphasizing a punchline, building up tension, or asserting authority or dominance (Edwards, 2013). Relatively little pitch fluctuation throughout a verse often indicates a gritty, serious ambiance, whereas higher levels of pitch fluctuation tend to indicate levity or a more relaxed vibe. It should be mentioned here, however, that there is most definitely no one-to-one relationship between how much pitch fluctuation a rapper displays and the light-heartedness of the topics they address. For example, while Snoop Dogg is renowned for frequently

changing his pitch when he raps (Edwards, 2013, p. 96), he often does so while telling some of hip-hop's darkest stories of gang violence and police brutality, as can be observed from the following lines of the song "Murder was the case" (Broadus, Young, & Arnaud, 1993), which exhibit several dramatic changes in pitch:

Pumping on my chest and I'm screaming

I stop breathing, damn, I see demons

Dear God, I wonder can you save me?

I can't die, my boo boo's 'bout to have my baby (00:57–1:07)

*1.2.3 Regional variation in rap flows—East Coast vs. West Coast.* Very little research (be it qualitative or quantitative) has been conducted on how exactly East Coast and West Coast rap flows differ from each other. However, while Edwards does not offer a direct comparison of the two, his discussion of various rhythmic and melodic flow techniques provides some valuable insights in this respect. Regarding rhythm, he highlights the typically West Coast "lazy tails" syncopation technique, in which "the rapping is squarely in time to the beat, apart from certain syllables or words that are said so that they slide off the beat. This adds variation to the flow and gives it a laid-back feel" (Edwards, 2013, p. 34). Key in this quotation is that lazy tails, which usually involve elongation of a syllable's onset so that its nucleus is pronounced slightly after the beat rather than right on it, are used to add rhythmic variation. Their pervasiveness in the oeuvres of West Coast rappers such as Snoop Dogg, Ice Cube, Shock G, MC Eiht, and Nipsey Hussle indicates a strong tendency towards loose rather than strict rhythms, especially considering their near absence in the works of East Coast rappers.

In terms of melody, Edwards (2013) reports on several rappers whose flows are marked by a relatively monotone delivery, highlighting the works of five rappers in particular: Nas, Guru (of Gang Starr), Divine Styler, Large Professor (of Main Source), and Rakim. Though Edwards does not ascribe monotone rapping to any region in particular, the five rappers he uses to illustrate the phenomenon are all New York-based.<sup>1</sup> In contrast, when he discusses rappers and rap groups whose flows are known for their pitch inflections, Edwards points towards Snoop Dogg several times as well as to RBX, the South Central Cartel, The Pharcyde, E-40, The D.O.C., and Pharoahe Monch. Excluding Pharoahe Monch, who is from Queens, New York, all these artists are based on the West Coast.<sup>2</sup> These examples suggest that the typical East Coast rap flow is marked by relative monotony compared to the typical West Coast flow.

### *1.3 Hip-hop's language-music connection*

As of yet, it is unclear why different American regions' rap flows have developed in different ways. One possible explanation, which was suggested by Kautny (2015), is that certain languages might be better suited for certain styles of flow due to their intonational characteristics. Kautny suggested that Spanish, because it is relatively isochronous (i.e., syllable-timed), would probably lend itself better to "steady styles with crossing pulses" (Kautny, 2015, p. 114), whereas a less isochronous language such as English would lend itself better to flows with a looser rhythm. However, this hypothesis has not been tested yet in the context of different languages, let alone in the context of different regional varieties of one language's sociolect. It has not been investigated either whether Kautny's hypothesis would extend to the domain of pitch/melody as well. Coming from a linguistic

perspective, Thomas has suggested that “it could be fruitful to explore the links between [AAE prosodic] features and various styles of music, poetry, and public speaking that African Americans have bestowed upon American culture” (Thomas, 2015, p. 432).

While the connection between regional variation in speech prosody and music has not been studied in the context of rap flows yet, the claim that composers’ music reflects their native language’s prosodic patterns (e.g., Abraham, 1974; Hall, 1953) has been investigated in other contexts. Patel, Iversen and Rosenberg (2006), for instance, analyzed prosodic rhythm and melody for British-English and French and then compared their rhythmic and melodic characteristics to those of British and French classical music. Their findings indicated that British-English speech was more rhythmically variable than French speech and that it was also marked by greater levels of melodic interval variability. Correspondingly, British classical music was found to be more rhythmically diverse and to exhibit more melodic variation than French classical music. From these findings, Patel et al. concluded that the rhythms and melodies of music are informed by those of speech, suggesting that “statistical learning of the prosodic patterns of speech creates implicit knowledge of rhythmic and melodic patterns in language, which can in turn influence the creation of rhythmic and tonal patterns in music” (Patel et al., 2006, p. 3043). As such, language and music (especially regarding rhythm and pitch) are thought to be different yet closely related domains that operate in fundamentally similar ways. Comparable findings on the link between regional speech varieties and instrumental music were reported by, among others, Huron and Ollen (2003), McGowan and Levitt (2011), and Hansen, Sadakata, and Pearce (2016), both in similar and different contexts than the ones investigated by Patel et al. (2006). Later studies (e.g., Jekiel, 2014; Lee, Brown, & Müllensiefen, 2017; Van Handel & Song, 2010) have also looked at the link between prosody and vocal music rather than instrumental music with most of these studies’ outcomes supporting the idea that there is a connection between prosody and music. It should be noted here, however, that Temperley’s (2017) findings regarding rhythm in French and Italian vocal music as compared to British and German vocal music patterned in the opposite direction than was expected.

#### *1.4 Aims and hypotheses*

Inspired by Kautny’s (2015) idea that regional variation in rap flows may be informed by linguistic variation and by Thomas’s (2015) call for research exploring the link between AAE prosody and African-American styles of music, this study aims to establish a number of things that we currently know very little about: how East Coast and West Coast AAE differ regarding prosody; how East Coast and West Coast rap flows are distinct; and whether prosody and rap flows follow similar patterns when it comes to regional variation. To this end, the following hypotheses were formulated:

- (1) West Coast AAE exhibits more pitch fluctuation than East Coast AAE;
- (2) West Coast rap flows exhibit more pitch fluctuation than East Coast rap flows;
- (3) West Coast AAE exhibits more rhythmic variation than East Coast AAE; and
- (4) West Coast rap flows exhibit more rhythmic variation than East Coast rap flows.

The hypotheses relating to regional variation in rap flows (i.e., hypotheses 2 and 4) follow what little knowledge is available in the literature regarding East Coast and West Coast rap flows’ usage of pitch and rhythm (Edwards, 2013). Since no previous research comparing East Coast and West Coast AAE prosody has yet been conducted, hypotheses 1 and 3 are not directly based on previously reported findings. Instead, they are based on Kautny’s (2015) suggestion that regional variation in rap flows is informed by regional variation in speech and on Patel et al.’s (2006) finding that



**Table 1.** Selected West Coast and East Coast rappers.

West Coast	East Coast
Daz Dillinger, Dr. Dre, Ice Cube, Kendrick Lamar, Nipsey Huddle, Snoop Dogg, The Game, Warren G	50 Cent, Biggie Smalls (The Notorious B.I.G.), Fabolous, Jay-Z, Mos Def, Nas, Raekwon, Rakim

British and French prosody differ from each other in the same way that British and French classical music do. Following from these four hypotheses, a fifth, overarching hypothesis was formulated:

- (5) Regional variation in rap flows follows a similar pattern as regional variation in prosody.

To test these hypotheses, free speech and rap recordings of 16 prominent African-American members of the East Coast and West Coast hip-hop communities were phonetically analyzed regarding melody (pitch fluctuation) and rhythm (degree of syllable isochrony, variation across utterances, and musical (micro)timing).

## 2 Method

The materials, data, and several methodological tools described in this section are available at <https://osf.io/dzc3s> (Gilbers et al., 2019).

### 2.1 Prosody and flow—Melody

The present section describes the methodology used for the part of the study focusing on pitch contours and pitch fluctuation in West Coast and East Coast AAE speech and West Coast and East Coast rap flows.

#### 2.1.1 Materials

**2.1.1.1 Speech.** Audio from 16 interviews with prominent African-American rappers from the East Coast and West Coast—eight rappers per coast—was collected from various sources on the internet (e.g., YouTube videos and podcast episodes; see Appendix for an overview). To limit intracoastal variability, all selected speakers were from either the greater Los Angeles metropolitan area or the greater New York City metropolitan area, the two dominant hip-hop hubs for the West Coast and East Coast, respectively.<sup>3</sup> The aim was to only include rappers who are widely known among both fanatic and casual hip-hop listeners, and who are held in high regard by fans and critics alike. The “The best rappers from Los Angeles” (n.d.) and “The best New York rappers” (n.d.) lists published on *Ranker.com* were used as a starting point for artist selection, the rankings of which were democratically decided through thousands of online votes. For each list, the researchers took the top 50 and subsequently filtered out all artists that did not fit the study’s demographic criteria (i.e., artists that were not male, African-American, or not originally from the right city) or that were not actually rappers but rather singers (e.g., Nate Dogg) or deejays (e.g., DJ Yella), for instance. Finally, rap groups were filtered out so that only individuals remained on the list. The top 25 rappers per coast that remained after these filters had been applied were placed on a shortlist. From this shortlist, the first 16 artists (8 per coast) for who the researchers found interview recordings and three a capella rap verses of sufficient quality and length were selected.<sup>4</sup> All 16 artists that ended up being selected have had considerable commercial and/or critical success, and can undeniably be considered hip-hop ambassadors for their respective regions as a result. Table 1 presents an overview of the artists that were selected for analysis.

**Table 2.** Rap verses selected for pitch analysis.

Subject	Song title	Song artist	Verse (length)
Daz D.	"All I need" (2006)	Daz Dillinger	1 (0:41), 2 (0:42), 3 (0:43)
Dr. Dre	"Nuthin' but a 'G' thang" (1992)	Dr. Dre (with Snoop Dogg)	2 (0:30), 4 (0:40)
	"Still. D.R.E." (1999)	Dr. Dre (with Snoop Dogg)	1 (0:40)
Ice Cube	"It was a good day" (1991)	Ice Cube	1 (0:58), 2 (0:47), 3 (1:09)
Kendrick L.	"Backseat freestyle" (2012)	Kendrick Lamar	1 (0:37), 2 (0:37), 3 (0:40)
Nipsey H.	"Ocean views" (2016)	Nipsey Hussle	1 (0:46), 2 (0:47), 3 (0:48)
Snoop D.	"Nuthin' but a 'G' thang" (1992)	Dr. Dre (with Snoop Dogg)	1 (0:56), 3 (0:28)
	"The next episode" (1999)	Dr. Dre (with Snoop Dogg et al.)	1 (0:41)
The Game	"Dreams" (2004)	The Game	1 (0:49), 2 (0:49), 3 (0:48)
Warren G	"Regulate" (1994)	Warren G (with Nate Dogg)	1 (0:40)
	"I need a light" (2005)	Warren G (with Nate Dogg)	1 (0:40), 2 (0:41)
50 Cent	"Hustler's ambition" (2005)	50 Cent	1 (0:44), 2 (0:43), 3 (0:43)
Biggie	"Sky's the limit" (1997)	The Notorious B.I.G. (with 112)	1 (0:43), 2 (0:34), 3 (0:41)
Fabulous	"Baby" (2004)	Fabulous	1 (0:28), 2 (0:37), 3 (0:18)
Jay-Z	"Moment of clarity" (2003)	Jay-Z	1 (0:35), 2 (0:45), 3 (0:46)
Mos Def	"Summertime" (2007)	Mos Def (with Esthero)	1 (1:00), 2 (0:46)
	"Travellin' man" (1998)	DJ Honda (with Mos Def)	1 (0:38)
Nas	"It ain't hard to tell" (1994)	Nas	1 (0:33), 2 (0:32), 3 (0:36)
Raekwon	"Curious" (2006)	Raekwon	1 (1:14), 2 (1:05)
	"New York" (2005)	AZ (with Raekwon et al.)	2 (0:41)
Rakim	"Once upon a rhyme in Japan" (2000)	Nigo (with Rakim)	1 (0:50), 2 (0:49), 3 (0:49)

From the interviews, 15 excerpts of uninterrupted speech (10 to 15 seconds in length) per subject were selected for analysis. None of these excerpts featured other voices than the subjects', nor did they feature any background noise that could possibly interfere with the acoustic measurements.

**2.1.1.2 Rap.** For each of the 16 rappers listed in Table 1, three a capella<sup>5</sup> rap verses were collected for pitch analysis online. All of these a capella recordings were so-called "studio a capellas"—officially released a capella recordings by the artists or their labels themselves—or high quality a capellas created by skilled deejays/producers, not the typical "D.I.Y. a capellas" created by mere hobbyists (e.g., by filtering out certain frequencies in an attempt to "cut out" the sound of the instruments). This choice was made because the latter type is usually characterized by many artifacts in the signal which would likely interfere with the study's pitch analyses. Because the study's scope is limited to rap vocals, any instances of more traditional singing (i.e., controlled pitch contours of a musical melody) were manually removed from the recordings using Adobe Audition Version 3.0 software (Adobe Inc., 2007). None of the recordings featured voices that were artificially pitched up or down. Moreover, none of the voices analyzed were auto-tuned—that is, the pitch as performed by the artists was not rounded up or down to the nearest semitone—which is a common feature of contemporary rap (e.g., the work by artists such as Lil Wayne and Future). Whether recordings featured artificially manipulated voices was assessed by carefully examining semitone pitch measurements (i.e., assess whether the semitone pitch measurements featured a disproportionate number of integer values, which would indicate "rounded off," auto-tuned notes) and through auditory examination of the recordings (e.g., listening for the types of audio artifacts that are the natural, unavoidable result of digital pitch manipulation). Table 2 presents an overview of the verses used for this study's pitch analyses.



**Table 3.** Overview of the relationship between (hundredths of) semitones, hertz (Hz) values, and the new scale's values.

Semitone	Hz	Value
A <sub>0</sub>	55.000 Hz	0
A <sub>0.01</sub>	55.032 Hz	0.01
A <sub>0.02</sub>	55.064 Hz	0.02
A <sub>0.03</sub>	55.095 Hz	0.03
A <sub>1</sub>	110 Hz	12
A <sub>2</sub>	220 Hz	24
A <sub>3</sub>	440 Hz	36

**2.1.2 Procedure.** The speech excerpts and rap verses described above were phonetically analyzed using identical methods. Firstly, fundamental frequency was measured in hertz (Hz) every 10 milliseconds (ms) in Praat (Boersma & Weenink, 2016). These measurements were performed automatically using a script that was modified from Cook (2002) to measure pitch 100 times per second rather than only 10 times per second.

Next, the Hz values were converted to a novel logarithmic scale based on semitones because the Hz scale's linear nature does not accurately reflect pitch perception: while the difference between an A<sub>0</sub> note and an A<sub>1</sub> note is 55 Hz and the difference between an A<sub>1</sub> note and an A<sub>2</sub> note is 110 Hz, humans perceive these differences as being equal, namely a difference of 12 semitones (1 octave). Hz-to-semitone conversion was deemed necessary because otherwise the pitch data for two speakers whose voices have a different average fundamental frequency could not be “fairly” compared to each other. The resulting logarithmic semitone scale spans a frequency range of 55 Hz (A<sub>0</sub>) until 440 Hz (A<sub>3</sub>), which covers the range of male voices' pitch under normal circumstances (approximately one octave below the average until two octaves above the average male fundamental frequency). Because per definition, semitones and octaves can go lower *ad infinitum*, even well beyond the range of human perception, the scale's 0-point was arbitrarily selected to be A<sub>0</sub>. The following formula describes the relationship between a given hundredth of a semitone (i.e., cent) and the next hundredth of a semitone in terms of Hz, where  $x_{n+1}$  stands for the Hz value associated with the next cent, and  $x_n$  stands for the Hz value of the one preceding it:

$$x_{n+1} = x_n \times 2^{1/1200}$$

Every Hz measurement was rounded off to the nearest corresponding cent value and subsequently converted to the corresponding value on the new scale (see Table 3).

After conversion to the new semitone-based logarithmic scale, pitch contours for all speech excerpts and rap verses were plotted to allow for visual comparison of the degree of pitch fluctuation. Moreover, absolute delta scores (i.e., the difference between one measurement and the next measurement 10 ms later) were assessed to quantify the intensity of pitch fluctuation. Delta scores higher than five semitones per 10 ms were removed from the data, as pitch jumps of that size within a hundredth of a second are highly unlikely to be reflective of actual pitch production in speech. Rather, these are often the result of Praat erroneously measuring so-called octave jumps (Praat occasionally registers a pitch decrease of, say, two semitones as a 10-semitone increase) or Praat accidentally picking up the pitch from non-speech background noise.

**Table 4.** 16 bar (excerpts of) rap verses selected for rhythmic analysis.

Subject	Song title	Song artist	Verse (length)
Daz D.	"Got my mind made up" (1996)	2Pac (with Daz Dillinger et al.)	1 (0:41)
Dr. Dre	"Let me ride" (1992)	Dr. Dre (with Snoop Dogg et al.)	1 (0:41)
Ice Cube	"It was a good day" (1991)	Ice Cube	1 (0:46)
Kendrick L.	"m.A.A.d. city" (2012)	Kendrick Lamar (with MC Eiht)	2 (0:42)
Nipsey H.	"Last time that I checc'd" (2018)	Nipsey Hussle (with YG)	1 (0:40)
Snoop Dogg	"Gz and hustlas" (1993)	Snoop Dogg (with Nanci Fletcher)	1 (0:42)
The Game	"Dreams" (2004)	The Game	1 (0:49)
Warren G	"Party we will throw now" (2012)	Warren G (with Nate Dogg et al.)	2 (0:39)
50 Cent	"Hustler's ambition" (2005)	50 Cent	1 (0:43)
Biggie	"Machine gun funk" (1994)	The Notorious B.I.G.	1 (0:40)
Fabulous	"So N.Y." (2012)	Fabulous	1 (0:44)
Jay-Z	"Feelin' it" (1996)	Jay-Z (with Mecca)	1 (0:46)
Mos Def	"Mathematics" (1999)	Mos Def	1 (0:41)
Nas	"N.Y. state of mind" (1994)	Nas	1 (0:45)
Raekwon	"C.R.E.A.M." (1993)	Wu-Tang Clan	1 (0:42)
Rakim	"Don't sweat the technique" (1991)	Eric B. & Rakim	1 (0:36)

## 2.2 Prosody and flow—Rhythm

This subsection deals with the methodology used for the part of the study focusing on rhythmic variation in West Coast and East Coast AAE speech and West Coast and East Coast rap flows.

### 2.2.1 Materials

**2.2.1.1 Speech.** For each of the 16 speakers mentioned in Table 1, five excerpts (10 to 15 seconds in length) were selected for rhythmic analysis. For each rapper, these five excerpts were the same as their first five (out of 15) excerpts used for the pitch analyses.

**2.2.1.2 Rap.** For each of the 16 rappers, 16 musical bars from a single verse were selected for analysis.<sup>6</sup> All selected verses came from prominent songs in the rappers' careers. Table 4 provides an overview of the rap verses used for rhythmic analysis. Contrary to the rap verses selected for pitch analysis, full studio recordings rather than a capella verses were used because instrumental accompaniment allowed for more accurate division of the verses into separate bars (see the Procedure subsection below for details).

### 2.2.2 Procedure

**2.2.2.1 Speech.** Because of the inherent problems involved with assessing individual consonants in field recordings and determining where one syllable begins and the next one ends, it was decided to assess rhythmic variability in speech based on vowel duration, which is considered a generally robust variable that is relatively easy to measure (see Thomas & Carter, 2006 for a more extensive description of this). These measurements were performed manually, since a test run using several Praat scripts to automate measurements resulted in too many erroneous assessments. Following Thomas and Carter (2006), if a syllable's nucleus was not occupied by a vowel but by a syllabic consonant, the duration of this syllabic consonant was measured. A total of 4077 syllables were measured this way (2070 for West Coast AAE and 2007 for East Coast AAE).

The present study's approach to quantifying rhythmic variability in speech constitutes a modified version of the methods described in Thomas and Carter (2006), Patel and Daniele (2003), and Patel et al. (2006). As there are many dimensions to rhythm and rhythmic variability, it was decided to

analyze variability not through only one but through two different measurements that analyze variability from a different perspective. The first of these is the normalized pairwise variability index (nPVI; Patel & Daniele, 2003; Patel et al., 2006), a measurement based on the pairwise variability index measurement developed by Low, Grabe, and Nolan (2000). The nPVI measurement reflects variability across the stress-timed/syllable-timed continuum (i.e., how similar in length a language variety's stressed and unstressed syllables are) with higher values indicating relatively higher levels of stress-timedness. Patel et al. (2006, p. 3035) designed the formula to calculate nPVI scores as follows:

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

Here,  $m$  stands for the number of durational elements in a sequence and  $d_k$  represents the  $k$ th element's duration. As Patel et al. describe it, the “nPVI computes the absolute difference between each successive pair of durations in a sequence, normalized by the mean duration of the pair” (Patel et al., 2006, p. 3035), technically making it a contrastiveness index. As Thomas and Carter (2006) note, one of the advantages of working with this measurement is that it controls for overall speech rate, allowing for comparisons of speakers with different speech rates, or even for comparing multiple phrases by a single speaker who varies greatly regarding tempo.

The second measurement is the coefficient of variation (CV) measurement, which is equal to a data set's standard deviation divided by its mean. As such, it is a normalized version of the standard deviation measurement, and hence it allows for fair variation comparisons across subjects regardless of the size of their mean scores (Segalowitz & Hulstijn, 2005; Segalowitz, Poulsen, & Segalowitz, 1999). CV was chosen to be used alongside the nPVI measurement because unlike the nPVI measurement, it can describe overall degrees of variation across a whole utterance rather than merely compare adjacent syllables' durations.

**2.2.2.2 Rap.** When engaging in the analysis of rhythmic variability in rap flows, one inadvertently runs into a number of methodological issues that require creative solutions. For starters, it is highly difficult to apply the same methodology to rap flows as the one described above for speech. In order to get an accurate representation of the rhythms of rap flows, one needs to analyze them in conjunction with the beats they are rapped over (Adams, 2015; Krims, 2000). However, to accurately analyze vocals phonetically using Praat (Boersma & Weenink, 2016), recordings without any instrumentation are required, since the wave forms of the instruments would otherwise obscure those of the voice.

Considering the abovementioned, direct phonetic analysis of rap flow rhythms using Praat is not a viable option, which is why an alternative representation of rap flows is required: musical notation of some kind, be it traditional (i.e., sheet music) or a novel, less conventional method. The mechanics of traditional Western music notation—which is primarily concerned with melody instead of rhythm—are notoriously cumbersome to describe the rhythmic subtleties of rap flows, which is why the field of hip-hop musicology has seen myriad attempts to notate flows in non-traditional ways (Adams, 2015; Connor, 2018). To illustrate, Krims (2000) and Adams (2009) use a grid system which divides bars into sixteen subunits (at times adapting the grid to allow for the notation of, for instance, triplets). Kautny (2015), on the other hand, switches between more traditional notation and a more complicated version of Adams's grid system (which schematically reflects microtiming by placing words slightly before or after grid boundaries) depending on what he intends to illustrate. Edwards (2013) relies on a simpler grid that marks the location of a bar's four beats and relies mostly on deviating spelling to illustrate rhythmic abnormalities. Even when rap scholars do take a more traditional route to analyzing flow, they often have to resort to creative

methods in order to accurately represent rappers' rhythms. Connor (2018), for instance, represents rappers' microtiming using traditional Western music notation, but does so in a sophisticated yet highly unusual manner by analyzing all bars in terms of nonuplets.

While all these approaches offer great insights into rappers' rhythms (and to varying degrees into their microtiming), they were not designed and are hence not suited to describe the exact microtiming that rappers display, especially when the researcher's aim is to quantitatively compare the microtiming of different (groups of) rappers. In a sense, this is reminiscent of the distinction between the linguistic fields of phonology and phonetics: while phonology deals with the sound structure of speech, phonetics deals with the actual acoustic realizations of speech, with only the latter allowing for quantitative data analysis. What was needed for this study was to create an approach based on the actual realizations of rap rhythms rather than an approach based on rap rhythms' underlying structures. As the study is a quantitative one, representations of flow not only needed to be faithful to the rappers' microtiming, but they needed to be quantifiable as well. To this end, a new methodology was designed based on a system that will be greatly familiar to musicians involved with electronic music production: Musical Instrument Digital Interface (MIDI)—the music industry standard for communication between digital instruments and music software. Given that MIDI is not dependent on tempo (see below for a more detailed discussion of this), the newly designed method not only allows for highly accurate representations of rappers' rhythmic subtleties, it also allows for the quantification of these subtleties so that rappers' flows can be compared to those of others as well as be statistically analyzed. Below, a detailed description of the new method is presented.

*2.2.2.2.1 MIDI-based method of rhythmic flow analysis.* For each of the selected verses, the corresponding song was extracted in its entirety in WAVE format (.wav) and loaded into the digital audio workstation software Maschine version 2.3 (Native Instruments, 2015).<sup>7</sup> Using Maschine's on-board sampling interface, the exact start of the target verse's first bar and end of its sixteenth were determined through careful examination of the waveforms and the song's audio itself, and the rest of the song was deleted afterwards. Next, the song's beats per minute (BPM) value was determined manually through the software's "BPM tap" feature and further tweaked to ensure that the software would play back the 16 bar excerpt in the correct tempo (i.e., that it would play back the excerpt as a perfect musical loop without any "stuttering" in the rhythm as the sample would start its next loop). The following step involved automatically dividing the excerpt into 16 separate units (each forming a single, perfectly looping bar) using the "Slice" tool of Maschine's on-board sampler (see Figure 1 for a screenshot of this process). The final preparatory step involved loading sonically unobtrusive percussive sounds (heavily low-pass filtered snare drums) onto each drum pad of the Maschine Mikro Mk 2 controller (Native Instruments, 2015).<sup>8</sup>

Next, the rhythms of each bar's rapping were recreated in MIDI. The first step of this process, which was repeated for each bar, involved the researcher<sup>9</sup> carefully listening to the looping bar to get a good, general understanding of the artist's rapping style. Next, an approximated recreation of the rapper's microtiming was performed manually using a technique called "finger drumming" (rhythmically tapping electronic drum pads) and recorded in MIDI-format. Pads were played whenever the rapper made a syllable "land," which roughly corresponds to the moment when a syllable reaches the amplitude peak of its nucleus. Once the timing of all of a bar's syllables was recreated and recorded this way, the researcher would play back the MIDI recording using the previously mentioned unobtrusive percussive sounds alongside the original recording, all the while scrutinizing each and every MIDI note's timing and tweaking their temporal position if necessary (i.e., moving a note forward or backward so that it aligned with the rapper's timing as closely as possible). It should be noted here that Maschine's temporal MIDI resolution is extremely high,



**Figure 1.** Splicing Nipsey Hussle’s verse into 16 separate units using Maschine’s “Slice” tool.

**Table 5.** Overview of how Musical Instrument Digital Interface (MIDI) tick values correspond to musical temporal units and how they correspond to time in milliseconds (based on a typical 90 beats per minute (BPM) hip-hop tempo).

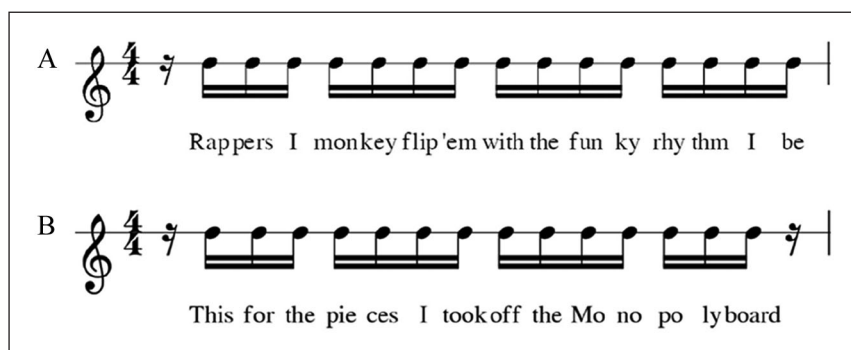
Note duration	Duration in MIDI ticks	Duration in seconds (90 BPM)
1 note (i.e., a single 4/4 bar)	3840	2.667
1/2 note	1920	1.333
1/4 note (i.e., one beat)	960	0.667
1/8 note	480	0.333
1/16 note	240	0.167
1/32 note	120	0.083
1/64 note	60	0.042
1/128 note	30	0.021
1/256 note	15	0.010

namely 960 “ticks” or “pulses” per quarter note, so these temporal tweaks can be highly precise.<sup>10</sup> Table 5 presents an overview of how MIDI tick values correspond to actual musical temporal units (i.e., note durations) and duration in seconds for a typical 90 BPM hip-hop tempo.

To illustrate the method’s precision, first consider Figures 2A and 2B, which show a single bar from Nas’s song “N.Y. state of mind” (Jones & Martin, 1994) and the late Nipsey Hussle’s song “Last time that I checc’d” (Asghedom & Jackson, 2018) represented in traditional music notation, respectively.

Both bars start with a single sixteenth rest, followed by a barrage of sixteenth notes (15 in the case of Nas, 14 in the case of Nipsey Hussle, who ends the bar with another sixteenth rest). These transcriptions suggest that rhythmically speaking, these bars are nearly identical and rhythmically invariable (i.e., rapid-fire flows without any syncopation). In contrast, Figure 3, which was created using this study’s MIDI-based approach, presents a representation of these bars along a grid with 16 subunits (cf. Adams, 2009; Krims, 2000, etc.) in which each black bar represents the temporal location of each of the verses’ notes rounded off to 256<sup>th</sup> notes for illustrative purposes.

Two things become apparent from inspecting Figure 3. Firstly, both rappers syncopate their notes—that is, none of the notes are performed straight on the beat, nor are the temporal distances between them the same each time—a rhythmic detail glossed over by the traditional music notation



**Figure 2.** (A) First bar of the first verse on Nas's "N.Y. state of mind" (1994); and (B) tenth bar of the first verse of Nipsey Hussle's "Last time that I checc'd" (2018).

of Figures 2A and 2B above. Secondly, the two rappers substantially differ from each other in terms of microtiming. In other words, the MIDI-based method revealed that two bars which seemed to be nearly identical at first glance were actually examples of two remarkably different flows.

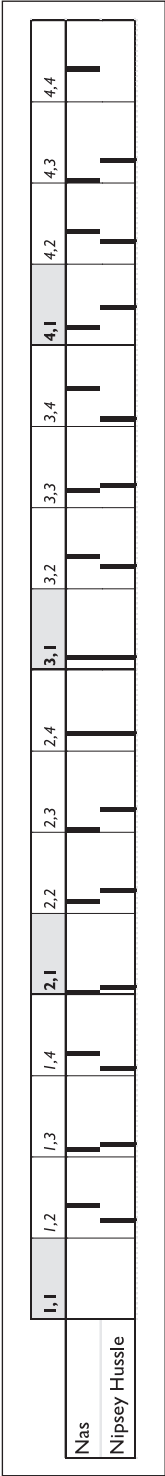
Once these steps were completed for all 16 bars, the MIDI recordings were exported in MIDI-format. Next, the binary MIDI files were converted to text files using a script capable of extracting numerical MIDI tick data available online (Flash Music Games, 2018). These data detail on which MIDI "tick" each syllable was performed. To illustrate, if a note was performed right on the first count of a bar, the MIDI tick value would be 0, if it fell exactly after one sixteenth rest, its value would be 240, etc. These data were entered into Excel, after which the interval distances between adjacent notes' temporal locations for each bar were calculated in terms of MIDI ticks. The measurement of choice for the rhythmic analysis of rap flows was the CV one described earlier. While Patel et al. (2006), in their study on the rhythms of French and British-English speech in relation to the rhythms of French and British classical music, opted to analyze musical rhythm using the nPVI measurement, it was decided not to take this approach for the present study (see the Discussion section below for a detailed explanation as to why).

To test the reliability of our rhythmic analyses, two additional raters were recruited to recode a quarter of the dataset (two randomly selected rappers per coast, namely Dr. Dre, Snoop Dogg, Jay-Z, and Nas). The first of these was a phonologist with a background in music but only a rudimentary knowledge/understanding of hip-hop and rap music. The second was a hip-hop producer, programmer, and radio host who is not involved with research or academia, but who does have extensive knowledge about hip-hop music and culture. Neither of these coders had been involved with the study so far, nor were they aware of the study's hypotheses at the time the inter-coder reliability tests took place. They were merely taught how the MIDI-based method worked and subsequently executed it. The results of this inter-coder reliability test are presented in the Results section below.

### 2.3 Statistical analysis

The data were statistically analyzed using one-sided independent samples *t*-tests and the Mann-Whitney *U* test. When two dependent variables correlated with each other and represented a part of the same theoretical construct, a multivariate analysis of variance was performed, as well as subsequent analyses of covariance (ANCOVAs). Inter-coder reliability for the MIDI-based method was assessed by calculating intraclass correlation coefficient estimates and their 95% confidence





**Figure 3.** Nas's first bar and Nipsey Hussle's tenth bar represented using the present study's MIDI-based approach—the bar grid is divided into four beats, with 1.1 standing for the bar's first beat, 2.1 for its second beat, etc., and further subdivided in sixteenth notes; the rappers' microtiming is represented with a temporal resolution of 256<sup>th</sup> notes.

intervals based on a single-measures, absolute-agreement, two-way mixed-effects model. All analyses were performed using SPSS 20 (SPSS Inc. Chicago, IL).

### 3 Results

In this section, the results of the comparisons between East Coast AAE and West Coast AAE speech prosody and East Coast and West Coast rap flows are presented, firstly regarding melody for speech and rap, and secondly regarding rhythm for speech and rap. They are visually presented together in Figure 4 with mean group values, individuals' mean values, and statistical significance indicated.

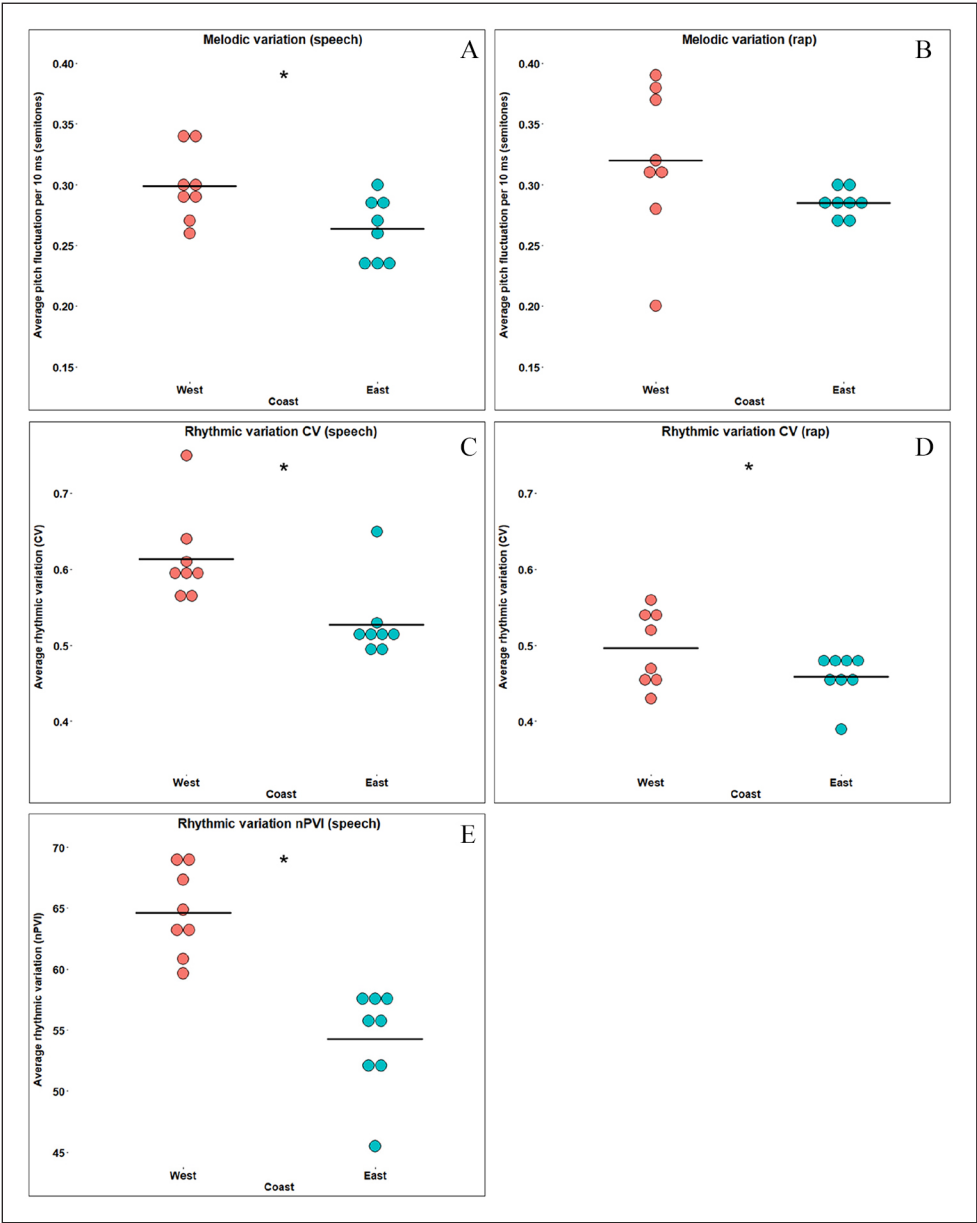
#### 3.1 Melody

**3.1.1 Melody—Speech prosody.** This subsection deals with pitch fluctuation in speech. As such, it is linked to the study's first hypothesis, that is, that West Coast AAE exhibits more pitch fluctuation than East Coast AAE.

To test the hypothesis that West Coast AAE exhibits more pitch fluctuation than East Coast AAE, a one-sided independent samples *t*-test was performed. In the current study, West Coast AAE, with a mean pitch fluctuation slope of 0.298 semitones per every 10 ms (standard deviation (*SD*) = 0.030), exhibited significantly more pitch fluctuation than East Coast AAE, with a mean pitch fluctuation slope of 0.262 semitones per every 10 ms (*SD* = 0.025),  $t(14) = 2.263$ ,  $p = 0.01$ ,  $d = 1.303$  (see Figure 4A). Figures 5A and 5B provide examples of these pitch contour plots for West Coast and East Coast AAE speech, respectively. These pitch contour plots, which were created using the semitone-based logarithmic scale designed for this study, illustrate the differences in West Coast and East Coast AAE pitch fluctuation reported above. Compared to the East Coast AAE examples, the peaks, valleys, and the magnitude of the slope of the West Coast AAE examples are visibly more extreme. Note that the vertical axes of Figures 5A and 5B both encompass the same pitch range (20 semitones) and that the horizontal axes both cover the same amount of time in seconds (12 seconds), allowing for "fair" visual comparison of the pitch contours.

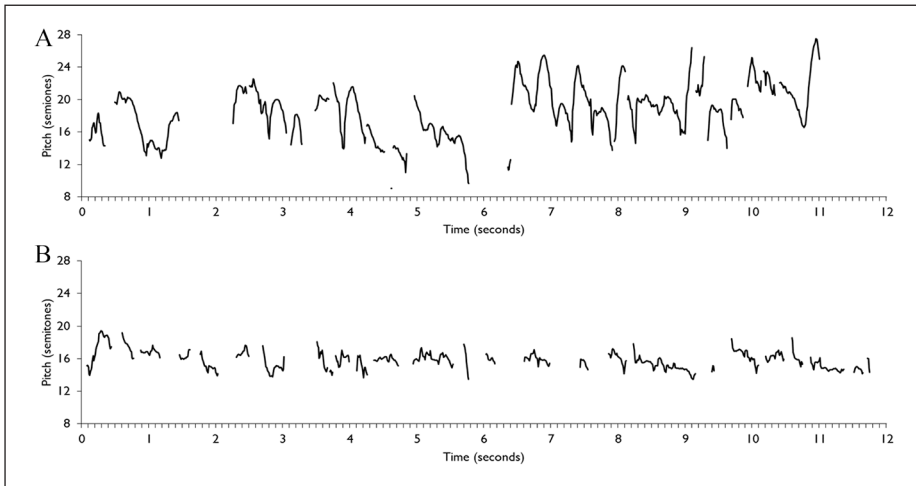
**3.1.2 Melody—rap flows.** This subsection deals with pitch fluctuation in rap flows. As such, it is linked to the study's second hypothesis, namely that West Coast rap exhibits more pitch fluctuation than East Coast rap.

To test the hypothesis that West Coast rap flows exhibit more pitch fluctuation than East Coast rap flows, a one-sided independent samples *t*-test was performed. In the present study, West Coast rap was found to exhibit an average pitch fluctuation slope of 0.320 semitones per every 10 ms (*SD* = 0.060), and East Coast rap a mean fluctuation slope of 0.285 (*SD* = 0.012; see Figure 4B). West Coast rap's greater pitch fluctuation compared to East Coast rap's is illustrated by the pitch contour plots of Figures 6A (Dr. Dre) and 6B (Biggie Smalls), respectively. From inspection of Figure 4B, it is immediately obvious that one subject of the West Coast group exhibited markedly different pitch fluctuation levels than his peers. This subject was Kendrick Lamar, and in fact, his average pitch fluctuation of 0.202 semitones per every 10 ms was not only the lowest of the West Coast group, but also far lower than the pitch fluctuation of any of the East Coast rappers. While the West Coast rap excerpts exhibited, on average, more pitch fluctuation than the East Coast rap excerpts, this was not a significant difference,  $t(14) = 1.582$ ,  $p = 0.068$ ,  $d = 1.303$ . An explorative one-sided independent samples *t*-test with Kendrick Lamar's rap flow data excluded was thus performed to gain a better understanding of how Lamar's data point affected the comparison between coasts. If Lamar's data had been excluded, West Coast rap's average pitch fluctuation



**Figure 4.** (A) Distribution of average pitch fluctuation per 10 milliseconds (ms) (semitones) for speech; (B) distribution of average pitch fluctuation per 10 ms (semitones) for rap; (C) distribution of average rhythmic variation (coefficient of variation (CV)) for speech; (D) distribution of average rhythmic variation (CV) for rap; and (E) distribution of average rhythmic variation normalized pairwise variability index (nPVI) for rap—horizontal lines indicate group mean values; asterisks indicate significant differences.

would have been 0.337 semitones per every 10 ms ( $SD = 0.040$ ) compared to East Coast rap's 0.285 ( $SD = 0.012$ ), which would have been a significant difference,  $t(13) = 3.241$ ,  $p = 0.007$ ,  $d = 1.723$ .



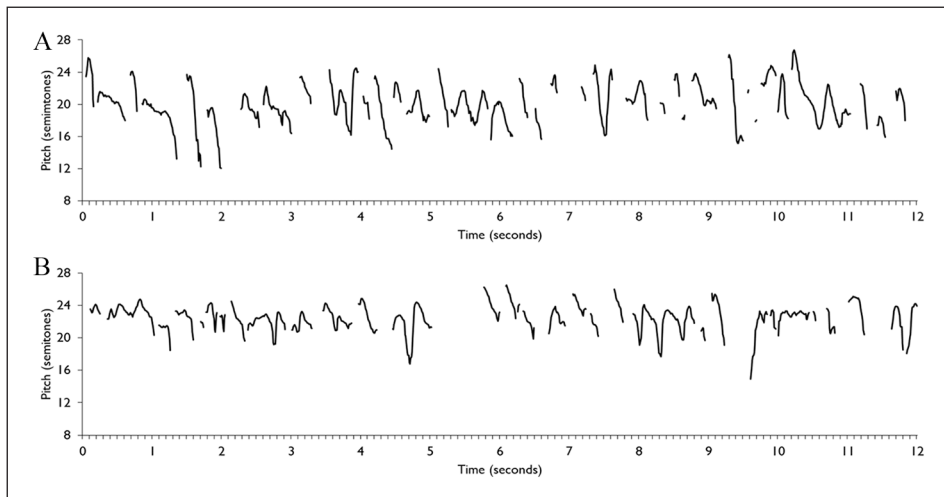
**Figure 5.** (A) West Coast African-American English (AAE) speech pitch contour example—Snoop Dogg (excerpt 5): “You got three homeboys, right? These two you grew up with, these two gangbain’, uhm, four years down the line. Them your homeboys, you love ‘em like brothers. They know your mom and daddy. When the—when the others said come by to ride. . .”; and (B) East Coast AAE speech pitch contour example—Jay-Z (excerpt 5): “And in between all that stress and angst and, you know, having to deal with one another in such close proximity, there’s so much love, and there was playing in the Johnny Pump, and there was, uh. . .”

## 3.2 Rhythm

**3.2.1 Rhythm—speech prosody.** This subsection deals with rhythmic variation in speech. As such, it is linked to the study’s third hypothesis, namely that West Coast AAE exhibits more rhythmic variation than East Coast AAE.

The hypothesis that West Coast AAE speech is characterized by more rhythmic variation than East Coast AAE was tested using two measurements, namely nPVI (a measurement which represents how stress-timed/syllable-timed a language variety is) and CV (a measurement which, as used in this study, reflects rhythmic variability across utterances). To test the hypothesis that West Coast AAE is more rhythmically variable than East Coast AAE along the stress-timed/syllable-timed continuum, a one-sided independent samples *t*-test was performed. West Coast AAE speech was found to exhibit a significantly higher nPVI score than East Coast AAE, namely 64.629 ( $SD = 3.565$ ) compared to 54.238 ( $SD = 4.211$ ) respectively,  $t(14) = 5.326$ ,  $p < 0.001$ ,  $d = 2.663$  (see Figure 4E). To test the hypothesis that West Coast AAE shows more rhythmic variation than East Coast AAE across entire utterances, a one-sided Mann–Whitney *U* test was performed. West Coast AAE speakers displayed significantly more rhythmic variation across utterances than the East Coast AAE speakers did in terms of syllable duration CV values, namely 0.614 ( $SD = 0.061$ ) and 0.527 ( $SD = 0.050$ ) respectively,  $U = 7.000$ ,  $p = 0.004$  (see Figure 4C).

“Speech rhythm nPVI” and “speech rhythm CV” were found to correlate with one another significantly,  $r = 0.627$ ,  $p = 0.009$ . Considering that these two variables both measure rhythmic variation in speech but from different perspectives, they were not just analyzed separately but also together by means of a multivariate one-way ANOVA. A statistically significant difference in speech rhythm based on coast was found when analyzing the two variables together,  $F(2,13) = 14.853$ ,  $p < 0.001$ , Wilks’ lambda = 0.304, partial  $\eta^2 = 0.696$ . In other words, roughly 70% of the



**Figure 6.** (A) West Coast rap pitch contour example—Dr. Dre’s “Nuthin’ but a ‘G’ thang” (Young, Broadus, Jr., & Curry, 1992; verse 1, excerpt): “Peepin’, and I’m creepin’, and I’m creepin’ / But I damn near got capped, ‘cause my beeper kept beepin’ / Now, it’s time for me to make my impression felt / So sit back, relax, and strap on your seatbelts / You never been on a ride like this before. . .”; and (B) East Coast rap pitch contour example—Biggie Smalls’s “Sky’s the limit” (Wallace et al., 1997; verse 3, excerpt): “After realizin’ to master enterprisin’ / I ain’t have to be in school by ten, I then / Began to encounter with my counterparts / On how to burn the block apart, break it down into sections. . .”

rhythmic variation observed in speech can be explained by which coast the speakers are from. Next, the unique contributions towards predicting which coast a speaker is from were teased apart for the two speech rhythm variables by means of two ANCOVAs. The result of the ANCOVA performed with “coast” as the independent variable, “speech rhythm nPVI” as the dependent variable, and “speech rhythm CV” as the covariate was significant,  $F(1,13) = 12.286, p = 0.04$ . The result of the ANCOVA performed with “coast” as the independent variable, “speech rhythm CV” as the dependent variable, and “speech rhythm nPVI” as the covariate was not significant,  $F(1,13) = 1.113, p = 0.311$ . This means that while both groups differ from each other significantly regarding rhythmic variability in speech on the basis of either nPVI or CV, some of the regional variation is only explained by “speech rhythm nPVI” and not by “speech rhythm CV,” whereas the opposite is not the case.

**3.2.2 Rhythm—rap flows.** This subsection deals with rhythmic variation in rap. As such, it is linked to the study’s fourth hypothesis, namely that West Coast rap exhibits more rhythmic variation than East Coast rap.

**3.2.2.1 Inter-coder reliability investigation.** The inter-coder reliability investigation for this study’s MIDI-based method of analyzing rhythm in rap revealed that the two re-coders’ measurements were, on average, slightly later than the main coder’s, namely 10.59 and 8.25 MIDI ticks, which corresponds to 7 ms and 6 ms, respectively. These are minuscule timing differences, below the just-noticeable-difference-levels for rhythmic timing reported by Friberg and Sundberg (1993), which suggests high agreement among the three coders.

The interval values that resulted from the main coder's as well as the first and second re-coders' coding were statistically analyzed by calculating intraclass correlation coefficient estimates and their 95% confidence intervals based on a single-measures, absolute-agreement, two-way mixed-effects model. An intraclass correlation of 0.95 (95% confidence interval: 0.94–0.96) was found,  $F(709, 1418) = 57.54, p < 0.001$ . Following the guidelines provided by Koo and Li (2016), this indicates that there was excellent inter-coder reliability, the highest reliability level listed in said guidelines.

**3.2.2.2 Rhythmic variation in West Coast and East Coast rap.** To test the hypothesis that West Coast rap flows exhibit more rhythmic variation than East Coast rap flows, a one-sided independent samples *t*-test was performed. The West Coast subjects displayed significantly more rhythmic variation in their rap verses than the East Coast subjects did in terms of note interval CV values, namely 0.495 ( $SD = 0.050$ ) and 0.457 ( $SD = 0.030$ ) respectively,  $t(14) = 1.863, p = 0.044, d = 0.932$  (see Figure 4D).

West Coast rap's higher degree of rhythmic variability as reported above is illustrated by the following representations of rappers' rhythms (Figure 7). They show how Jay-Z (East Coast) is quite steady regarding the note length (almost exclusively sixteenths) while Snoop Dogg (West Coast) switches between note lengths multiple times (sixteenths and thirty-seconds) while also using rests, leaving more "space" in the bar. Furthermore, despite some syncopation, Jay-Z's rapping is much more "straight on the beat" than Snoop Dogg's is, as the latter's delivery is "late" for nearly every note.

### 3.3 Melody and rhythm combined—speech prosody and rap flows

Following Patel et al. (2006), this subsection will examine melody and rhythm together in light of the study's fifth hypothesis (i.e., that regional variation in rap flows follows a similar pattern as regional variation in prosody). Figure 8 presents the two plotted together on a two-dimensional space for West Coast AAE, East Coast AAE, West Coast rap, and East Coast rap, with average melodic variation (pitch fluctuation per 10 ms in semitones) plotted along the horizontal axis and overall rhythmic variation (CV) plotted along the vertical axis. Interestingly, it can be observed that compared to speech, rap exhibited *more* melodic variation but *less* rhythmic variation.<sup>11</sup>

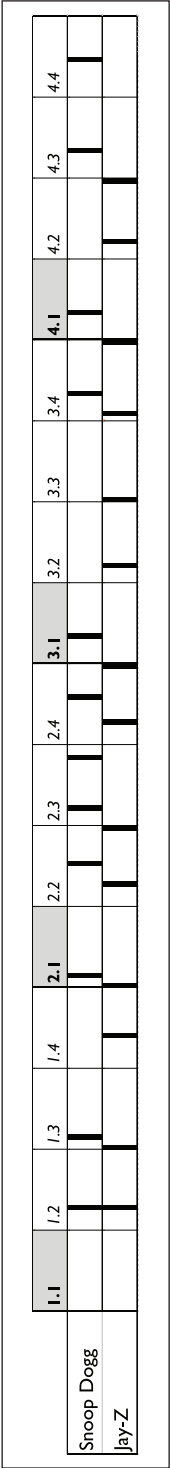
The prosodic distance (p.d.) between West Coast AAE and East Coast AAE prosody ( $WC_s, EC_s$ )—that is, the Euclidean distance between the points representing the mean melodic and rhythmic (CV) values—can be calculated using the following formula adapted from Patel et al. (2006):

$$p.d.(WC_s, EC_s) = \sqrt{\left( Melody_{WC_s} - Melody_{EC_s} \right)^2 + \left( Rhythm_{WC_s} - Rhythm_{EC_s} \right)^2}$$

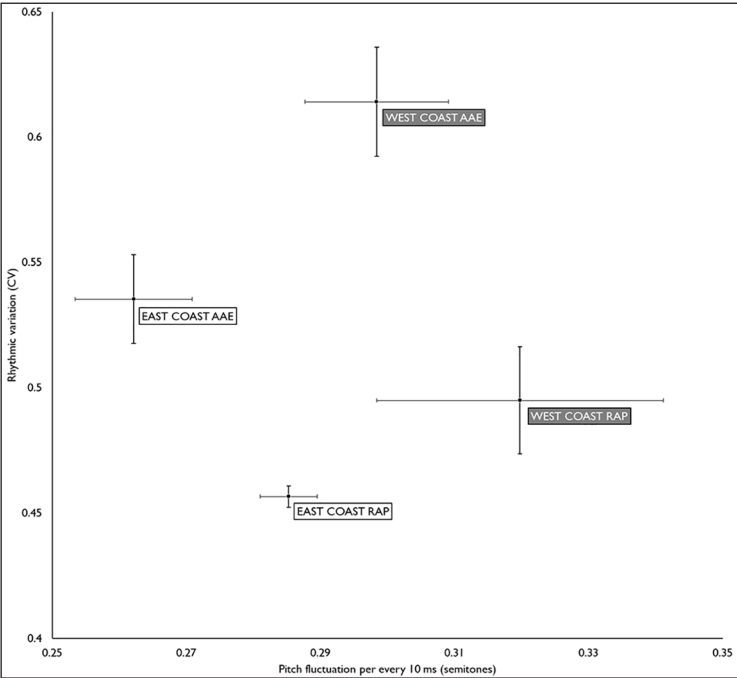
The p.d. between West Coast AAE and East Coast AAE speech prosody was found to be 0.094 rhythm–melody units. Applying the same formula to the rap data shows that the distance between West Coast and East Coast rap flows was 0.052 rhythm–melody units. The musical distance between the West Coast and East Coast was thus smaller than the linguistic distance (roughly 55% the size of the linguistic difference).

If one were to draw a line connecting East Coast AAE speech and West Coast AAE speech in the rhythm–melody space, this line would lie at a highly similar angle to a line connecting East Coast and West Coast rap flows. The two lines would move in the same direction (towards greater melodic variation and greater rhythmic variation). In fact, the angle between the two vectors (i.e., the vector for East Coast-to-West Coast AAE speech and the vector for East Coast-to-West Coast rap) is small: only 20.2° (calculated using standard trigonometric formulas).





**Figure 7.** Rap rhythm plot examples—Snoop Dogg (West Coast; ninth bar from “Gz & hustlas”, 1993) and Jay-Z (East Coast; seventh bar from “Feelin’ it”, 1996)—the bar grid is divided into four beats, with 1.1 standing for the bar’s first beat, 2.1 for its second beat, etc., and further subdivided in sixteenth notes; the rappers’ microtiming is represented with a temporal resolution of 256<sup>th</sup> notes.



**Figure 8.** (A) Melodic variation (pitch fluctuation slope per every 10 milliseconds [ms]; x-axis) and rhythmic variation (coefficient of variation; y-axis) for West Coast African-American English (AAE) and rap (gray) and East Coast AAE and rap (white)—error bars show standard errors.

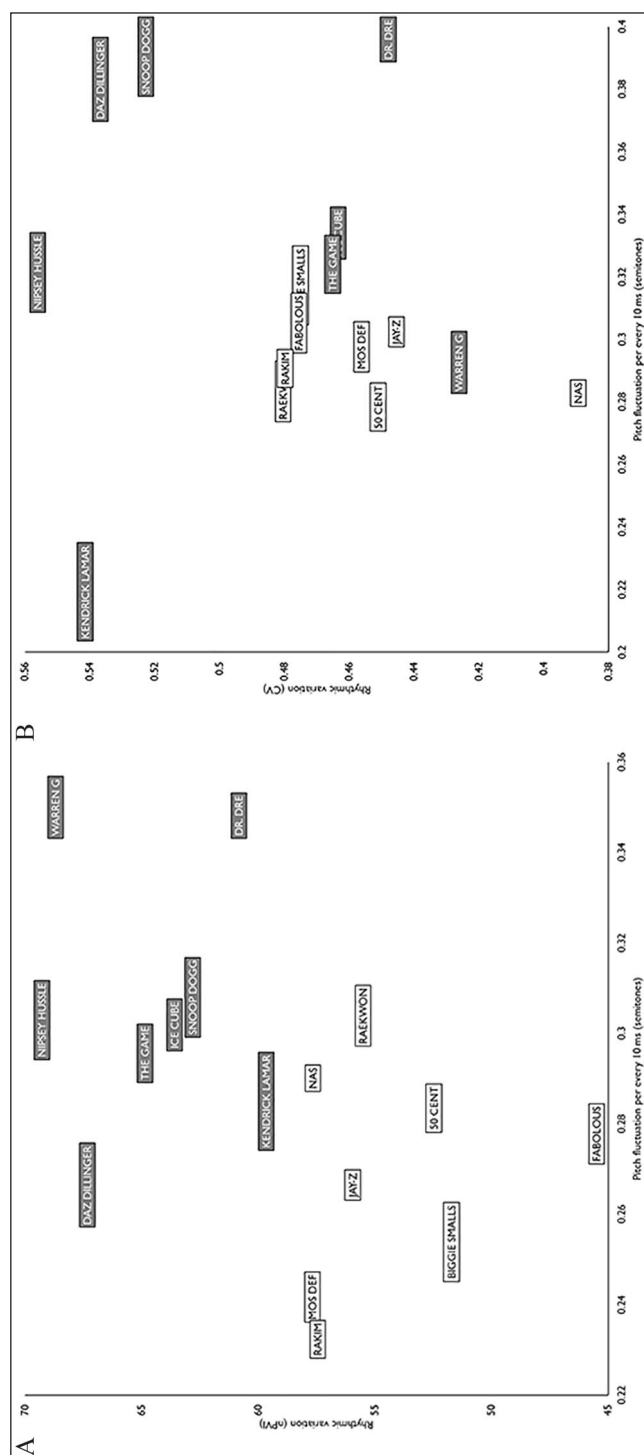
Plotting the individual subjects’ displays of melodic and rhythmic variation for speech (Figure 9A) and for rap (Figure 9B) reveals that the groups largely occupy distinct areas of the two-dimensional rhythm–melody space in both domains (although plenty of variation can be observed along each dimension). This is most clearly the case for speech, as Figure 9A shows a complete separation between the two groups in this space, but it can also be seen to be the case for the domain of rap, with one exception: Warren G, a West Coaster who finds himself right in the middle of the East Coast group. Figure 9B also clearly shows how Kendrick Lamar follows his West Coast peers in terms of rhythm but deviates from every other subject in terms of melody in rap. The rhythm–melody plots suggest that the regional characteristics of these regions’ prosody and flow may be defined not by melody and rhythm separately, but by their melodic and rhythmic characteristics together.

# 4 Discussion

In this section, the current study’s results will be discussed in light of its hypotheses and the existing literature. Moreover, the novel elements of its methodology will be evaluated, for instance its MIDI-based rhythmic flow analyses.

## 4.1 Discussion of method

In this study, both speech and rap were analyzed in terms of melodic and rhythmic variation using a combination of existing methods and methods newly designed for the present study. This



**Figure 9.** (A) Melodic variation in speech (pitch fluctuation slope per every 10 milliseconds [ms]; x-axis) and rhythmic variation normalized pairwise variability index [nPVI]; y-axis) for West Coast African-American English (AAE) prosody (gray) and East Coast AAE prosody (white), plotted per subject; (B) Melodic variation (x-axis) and rhythmic variation (note interval coefficient of variation values; y-axis) for West Coast rap flows (gray) and East Coast rap flows (white), plotted per subject.

subsection will discuss the study's methodology in light of previous research while also looking at possible future applications of its novel methods.

*4.1.1 Comparison with previous approaches.* Patel et al.'s (2006) methodology lies at the core of that of the present study. Nevertheless, our study's methodology differs from Patel et al.'s (2006) methodology in a number of ways (aside from the different linguistic context and the fact that the present study deals with rap vocals rather than classical music). For one, Patel et al. (2006) were unable to directly compare the music of composers such as Debussy, Holst, and Roussel to the way these composers spoke, since their subjects all passed away before speech recordings became possible or common. In contrast, the music and speech analyzed in the present study were produced by the same individuals, allowing for more direct comparisons of the domains of language and music, especially since instead of sheet music (for instrumental music), actual vocal music performances were analyzed in the present study for both the melodic and rhythmic domains. Since vocal rather than instrumental music was analyzed in this study, it was arguably more likely that a connection between speech and music would be found in the present study's context than in Patel et al.'s (2006) study, although it should be noted that, as was mentioned in subsection 1.3, the results of previous studies dealing with vocal music were not always entirely predictable.

It is worth noting here that Patel et al. (2006) and the present study differ regarding the degree to which the music that was analyzed is improvisational. While both classical compositions and rap compositions are often carefully constructed, classical compositions are completely formalized in the sheet music, whereas rap compositions often leave room for improvisations in terms of micro-timing, or may even be composed entirely in an improvisational fashion during the actual recording process. For example, artists such as Snoop Dogg are known for (partly) freestyling many of their verses. In fact, Snoop Dogg himself told *MTV News* that the lyrics and flow of the song we rhythmically analyzed for the present study—"Gz and hustlas"—were completely freestyled (Markman, 2013):

"Gz and hustlas" was easy, (. . .) it was supposed to be a mic check. I was just supposed to go in there and mic check. And when I mic checked, I freestyled, and whatever came out is the record that you hear.

As such, the present study's focus on rap means that it extends the existing literature on the language–music connection into the domain of improvisational music.

Regarding rhythm, Patel et al.'s (2006) methodology differs from the current study's in that Patel et al. (2006) used the nPVI measurement for both speech and music, whereas the current study only uses it for speech. The reason for this is that, as Thomas and Carter (2006) asserted, one of the characteristics of the nPVI measurement is that it normalizes for overall speech rate. While this is beneficial when analyzing rhythmic variation in speech, which is naturally marked by changes in tempo throughout, the same is not the case for rap. This is because in hip-hop music, speech rate is already normalized because the rhythms of rappers' flows are guided by a constant rhythm—the steady beat of the music. The nPVI's normalization of tempo changes is therefore: (1) redundant; and (2) potentially harmful as it could obscure meaningful aspects of rhythmic variation produced by rap artists. The nPVI's speech rate normalization also does not benefit comparability of rap verses of different tempos, since this is already taken care of by MIDI itself. MIDI automatically allows for a direct comparison between the timing in rap verses marked by different tempos in terms of BPM, as it keeps track of musical timing in terms of musical units rather than time in, say, seconds or milliseconds: a note starting after a sixteenth rest will have a tick value of 240 regardless of whether the music has a BPM value of 90, 127, or even 180.

Regarding melody, the present study's metrics differ from Patel et al.'s (2006) as well, as the latter did not analyze entire pitch contours but analyzed strong abstractions of these contours instead: Patel et al. (2006) only considered pitch in vowels and assigned either a level tone or a glide to each individual vowel using a glissando threshold based on perceptual research on pitch movement (Hart, 1976; Mertens, 2004). While insightful, this approach removes the subtle details of pitch contours, which is why a different approach was taken in the present study.

**4.1.2 Methodological innovations.** With regards to melody, our study proposes a semitone-based pitch scale that allows for perceptually accurate plotting of pitch contours. Additionally, a new way of articulating the intensity of pitch fluctuation was developed for this study, namely in terms of delta scores for pitch measurements in (hundredths of) semitones. The temporal resolution of these delta scores was 10 ms, meaning that pitch changes were assessed 100 times per second. This measurement was found to quantitatively reflect what could be visually observed from the pitch contour plots, namely that average delta scores were higher for pitch contours that were visibly less stable compared to more steadily developing pitch contours. Note that this measurement only captures one aspect of pitch fluctuation: its intensity. It does not, however, capture when rises and falls in pitch occur. It also does not say anything about larger “waves” of pitch fluctuation (i.e., whether a speaker's average pitch is consistent across phrases or whether some phrases are marked by higher mean pitch levels than others).

The study's rhythmic analyses of speech employ existing methods and parameters (i.e., nPVI and CV values based on vowel duration measurements; see the method of Patel et al., 2006, for example). However, for the rhythmic analyses of rap, no method existed yet that could sufficiently grasp the rhythmic nuances of rap performances while allowing for quantitative analyses. For this reason, a MIDI-based method of analyzing rhythmic variation and microtiming in rap flows was developed, which is arguably this study's biggest methodological innovation.<sup>12</sup> The approach allows for accurate representation of microtiming in rap flows due to MIDI's high temporal resolution, automatically allows for comparison of songs with different BPM values, and, unlike any other available method, allows for quantitative analysis of rap rhythms. It should be noted that due to the manual nature of the method, the researcher's sense of musical timing might interfere with the data during the “first take” of recreating a rapper's microtiming in MIDI to a certain extent (although no more so than any other existing form of rap flow notations). Moreover, humans are incapable of (consistently) recreating rhythms without error—see, for instance, Repp (2005) for a review of the literature on sensorimotor synchronization—and as a result of this, the “first take recreations” are likely to feature timing mistakes every once in a while. Because of this, each and every single note of the “first take recreations” needs to be manually double-checked to see if it accurately reflects artists' timing, and, if necessary, be manually adjusted to match the auditory signal (as was done for this study). This study's inter-coder reliability investigation showed that this method was both accurate and characterized by excellent reliability, making it suitable for application in future studies as well.

The applications of this method are not limited to those of the present study's as it can be used for other purposes as well. For instance, the method could be used for the creation of a big data rap flow corpus, which could in turn be used to create “heat map” visualizations of which notes of a bar are used most often by individual rappers or rappers from different regions or eras. Examples of such heat maps based on the present study's data are presented in Figure 10. Figures 10A and 10B show individual rappers' average flows (West Coast and East Coast, respectively). Figure 10C, which provides heat maps based on coastal averages, reveals that East Coast rappers generally tend to “hit” each eighth count of a bar, whereas West Coast rappers are less likely to do so (especially





near the end of bars) and show a more equal frequency distribution across all 64<sup>th</sup> notes, which indicates greater levels of rhythmic variation.

## 4.2 Discussion of results

Four hypotheses were formulated for this study, the first two of which are related to the domain of melody and the latter two of which are related to the domain of rhythm:

- (1) West Coast AAE exhibits more pitch fluctuation than East Coast AAE;
- (2) West Coast rap flows exhibit more pitch fluctuation than East Coast rap flows;
- (3) West Coast AAE exhibits more rhythmic variation than East Coast AAE; and
- (4) West Coast rap flows exhibit more rhythmic variation than East Coast rap flows.

Implied by these four hypotheses is the article's fifth, overarching hypothesis that regional variation in rap flows follows a similar pattern as regional variation in speech prosody in the context of West Coast and East Coast AAE and rap.

The study's results support the first hypothesis, as West Coast AAE's average pitch fluctuation slope per every 10 ms was found to be significantly greater than East Coast AAE's. As was the case with speech, the average pitch fluctuation slope of West Coast rap flows was also found to be greater than that of East Coast rap flows. These findings do not support the second hypothesis, though, as only a marginal effect was observed. Closer examination of the data revealed that this insignificant result was the result of the pitch fluctuation exhibited by Kendrick Lamar, an outlier in the West Coast group displaying a substantially lower degree of pitch fluctuation. For this reason, another explorative statistical analysis was performed with Lamar's data excluded. Without Lamar's data included, West Coast rap flows were found to exhibit significantly greater levels of pitch fluctuation than East Coast rap flows. This suggests that it would be valuable to replicate the present study in the future with a larger subject pool in order to see if West Coast rap flows are indeed marked by greater melodic variation than East Coast rap flows.

In terms of prosodic rhythm, rhythmic variation was examined along two parameters: nPVI scores (reflecting rhythmic variation along the stress-timed/syllable-timed continuum); and CV (reflecting variation across entire utterances). Regarding the former, it was found that West Coast AAE was characterized by significantly higher nPVI scores than East Coast AAE, meaning that West Coast AAE is markedly more stress-timed. In fact, no East Coast subject displayed a higher nPVI score than any single West Coast subject. In terms of variation across entire utterances, the results show that West Coast AAE speakers displayed significantly higher CV levels than East Coast AAE speakers, that is, displayed more rhythmic variation across whole utterances. Both the nPVI and CV findings therefore support the study's third hypothesis. Regarding rap flows, rhythmic variation was tested only in terms of note interval CV data. It was found that West Coast rap flows showed significantly higher CV values than East Coast rap flows. Hence, the findings support the fourth hypothesis as well. Interestingly, it was found that some of the regional variation is only explained by "speech rhythm nPVI" and not by "speech rhythm CV," whereas the opposite was not found to be the case. Future studies could investigate this further to gain a better understanding of how rhythmic variation in terms of nPVI and CV are related to each other.

In sum, hypotheses 1, 3, and 4 were all supported by significant results. This was not the case for hypothesis 2, however. Because visual inspection of Figure 4B revealed that one subject—Kendrick Lamar—behaved markedly different than the others, an explorative investigation was conducted in which Lamar's data point was excluded from analysis. It was found that, had Lamar's data point not been included, the difference between coasts would have been significant. Lamar's

display of pitch fluctuation while rapping seemingly formed the exception to the rule, and it might be the case that Lamar's use of pitch highlights an artistic motivation that was not considered before the hypotheses were formed: while rappers are on the one hand encouraged by their environment to represent for their home region by using the flows typically associated with said region, they might also have an incentive to deviate from these norms in order to stand out amongst thousands of other aspiring rappers fighting for attention in the local, national, and international markets. As a result, somebody such as Kendrick Lamar, whose use of pitch in speech is similar to that of his West Coast peers, perhaps decided to do something completely different with his flow than his colleagues on the West (and even the East) Coast. It is worth mentioning here that Lamar actually displayed some of the most rhythmic variation of all rappers analyzed in this study, so if he was indeed trying to stand out from the crowd with his rapping, he only did so in the domain of melody, not rhythm. Of course, this is all highly speculative, and further research needs to be conducted on West Coast rap flows in general and Kendrick Lamar's flow in particular to see if there is merit to this explanation of why Lamar deviated from his peers.

From these data, it can be concluded that West Coast AAE prosody and East Coast AAE prosody are distinct from each other for the domains of both rhythm and pitch, with West Coast AAE exhibiting more rhythmic and melodic variation. For rap flows, the same pattern of regional variation can be observed, although the differences are a little less clear-cut (see the discussion of Kendrick Lamar above). Hence, the findings support Kautny's (2015) idea that rap flows may be reflective of prosodic features. Moreover, the data suggest that Patel et al.'s (2006) findings that classical composers were influenced by their mother tongue's prosodic features may be extended to the domain of rap. These findings also suggest that this influence of prosody on music may be so strong that its effects cannot just be observed when comparing music from speakers of distinct languages, but of regional variants of the same language (variety) as well. More research needs to be conducted to verify these observations, however, and the present study should be seen as a starting point rather than an end station in this regard.

The present study found that the musical p.d. between the East Coast and the West Coast was roughly 55% the size of the linguistic prosodic difference, a larger percentage than was found by Patel et al. (2006) for English and French speech and music (about 30%). This finding would suggest that music and speech are more similar to each other in the West Coast/East Coast AAE rap context than they are in the context of British-English and French classical music and speech. Moreover, Patel et al. found that in their two-dimensional rhythm–melody space plot, “a line connecting English and French speech (. . .) would lie at a very similar angle to a line connecting English and French music” (Patel et al., 2006, p. 3042), the difference between the prosody and music vectors being only 14.2°. From the present study's rhythm–melody space plot for East Coast and West Coast AAE speech and East and West Coast rap (Figure 8), a similar observation can be made, with an only slightly larger angle between the vectors for prosody and for rap, namely 20.2°. In light of the Patel et al. study, it is also noteworthy that on their rhythm–melody space plot with all the analyzed composers plotted individually, they found that the British and French groups hardly overlapped:

English and French occupy distinct regions of [the two-dimensional space of rhythm and melody], despite large variation along any single dimension (. . .). This suggests that the *joint* properties of melody and rhythm, not either one alone, are involved in defining national characteristics of music. (Patel et al., 2006, p. 3042)

Correspondingly, the current study found similar results for West Coast AAE and East Coast AAE prosody (see Figure 9A) as well as West Coast and East Coast rap music (see Figure 9B). The fact

that these findings on rhythm and melody corroborated the patterns observed by Patel et al. (2006) indicates that it will be worthwhile to further explore rhythm and melody as well as music and language together in future studies of prosody and music.


To conclude, this study found that West Coast AAE speech prosody is characterized by more rhythmic and more melodic variation than East Coast AAE speech prosody. Moreover, it was found that West Coast rap was marked by more rhythmic variation than East Coast rap, and while no statistically significant difference between the two regarding melodic variation was found, a trend towards more melodic variation on the West Coast was observed. The study's findings appear to suggest that in the context of West Coast and East Coast AAE and rap, variation in regional dialects and regional styles of music follow similar patterns and may be related to each other. From the findings, it also appears that the influence of sociolinguistic factors (in this case: expressing regional identity linguistically) may not be limited to the domain of language, but that it may cross over into the domain of music as well. Future studies should further explore this study's findings, for instance by also looking at other rappers from the East Coast and West Coast, or by looking at more pieces of music and instances of speech from the artists that were examined in the present study. Moreover, it would be very interesting to replicate the study for other regions of hip-hop as well, such as the Midwest (e.g., Chicago or Detroit) and the South (e.g., Atlanta and Houston) in order to get a better idea of regional variation in AAE prosody and rap flows across the United States. Furthermore, future studies could explore if and how the flows of rappers with different mother tongues differ from each other and if this reflects their native language's prosodic characteristics as well (i.e., explore Kautny's (2015) suggestion that, for instance, English and Spanish rap would be rhythmically different due to rhythmic differences between the English and Spanish language).

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors received financial support from the University of Groningen for the research and authorship of this article.

## ORCID iDs

Steven Gilbers  <https://orcid.org/0000-0002-4218-0770>

Wander Lowie  <https://orcid.org/0000-0002-2241-0276>

## Notes

1. Nas and Large Professor hail from Queens, Rakim hails from Long Island, and Divine Styler from Brooklyn. Guru was born in Massachusetts, but his group Gang Starr was based in New York City.
2. Snoop Dogg and RBX are from Long Beach, California, the South Central Cartel and The Pharcyde are from Los Angeles, E-40 is from the California Bay Area, and while The D.O.C. is originally from Dallas, Texas, he made a name for himself on the West Coast gangsta rap scene, as a solo act and through collaborations with N.W.A and Dr. Dre.
3. Throughout this article, we use the terms 'East Coast' and 'West Coast' despite the fact that our subjects are exclusively from (the greater metropolitan areas of) New York City and Los Angeles. The primary reason for using 'East Coast' and 'West Coast' is that in hip-hop culture, these terms are usually used in a *totum pro parte* fashion, meaning that they often refer to just New York City and Los Angeles hip-hop.
4. The researchers were limited with regards to who could be selected due to pragmatic reasons. For instance, the rappers Eazy-E (Los Angeles) and Big L (New York) both easily fit the study's criteria, but could unfortunately not be included because no interviews with them of sufficient audio quality and

length were available for the speech analyses. Similarly, no/not enough a capella rap verses of sufficient audio quality could be obtained for many of the shortlisted artists (e.g., YG, Mack 10, Vince Staples, and Earl Sweatshirt), making them ineligible for analysis as the researchers' aim was to study the speech and rap of the same individuals.

5. In music terminology, a capella refers to vocals unaccompanied by any instrumentation.
6. If a selected verse happened to be longer than 16 bars, only its first 16 were analyzed.
7. Note that while Maschine (Native Instruments, 2015) was used in this study, most professional digital audio workstations (e.g., Cubase, Ableton, Logic Pro, and FL Studio) are capable of performing the basic functions required for this method to work if they support Musical Instrument Digital Interface (MIDI) and especially when controlled with a MIDI controller (e.g., a keyboard or electronic drum pads), although perhaps not as intuitively or efficiently as Maschine can.
8. While a Maschine Mikro Mk 2 controller was used for the present study, any alternative Maschine controller would have worked as well.
9. The rhythmic analyses were conducted by this article's first author, a rapper and hip-hop musician himself with years of studio experience.
10. See Dobrian (2014) for a detailed discussion of how Musical Instrument Digital Interface's temporal resolution works.
11. It should be noted here that the rhythmic variation coefficient of variation (CV) values for speech and rap are not perfectly comparable, as these are based on related but nevertheless distinct variables: vowel duration for speech and note interval duration for rap. This subtle difference between the measurements might have influenced these results, although CV—being dimensionless (Patel et al., 2006)—is a relatively robust variable across domains.
12. It is worth noting here that Musical Instrument Digital Interface (MIDI)-based methods have been employed in previous research, for instance to test how accurately people can synchronize their finger tapping with auditory stimuli (see, for instance, Repp, 1999). As such, the innovation of the presently proposed method does not lie in the fact that it uses MIDI. Rather, it lies in aspects such as the method's loop-based approach, the fact that music of different tempos can be directly compared to each other without a need for any temporal normalization, the multiple applications, analyses and visualizations that it lends itself to, and the fact that it allows for the analysis of artists' actual timing rather than merely the abstract rhythms portrayed in sheet music.

## References

- Abraham, G. (1974). *The tradition of Western music*. Berkeley, CA: University of California Press.
- Adams, K. (2009). On the metrical techniques of flow in rap music. *Music Theory Online*, 15(5). Retrieved from [www.mtosmt.org/issues/mto.09.15.5/mto.09.15.5.adams.html](http://www.mtosmt.org/issues/mto.09.15.5/mto.09.15.5.adams.html)
- Adams, K. (2015). The musical analysis of hip-hop. In J. A. Williams (Ed.), *The Cambridge companion to hip-hop* (pp. 118–134). Cambridge, UK: Cambridge University Press.
- Adobe Inc. (2007). Adobe Audition [Computer program]. Version 3.0. San Jose, CA.
- Alim, H. S. (2004). Hip Hop Nation Language. In E. Finegan & J. R. Rickford (Eds.), *Language in the U.S.A.: Themes for the twenty-first century* (pp. 387–409). Cambridge, UK: Cambridge University Press.
- Asghedom, E., & Jackson, K. (2018). Last time that I checc'd [Recorded by Nipsey Hussle & YG]. On *Victory lap* [CD]. All Money In & Atlantic Records.
- Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer [Computer program]. Version 6.0.21. Retrieved from <http://www.praat.org/>
- Broadus, C., Jr., Young, A., & Arnaud, D. (1993). Murder was the case [Recorded by Snoop Doggy Dogg]. On *Doggystyle* [CD]. "Death Row Records", Interscope Records, Atlantic Records, & Entertainment One.
- Cole, J., Thomas, E. R., Britt, E., & Cogshall, E. L. (2005). Intonational distinctiveness of African American English. Paper presented at New Ways of Analyzing Variation XXXIV, New York, NY.
- Connor, M. E. (2018). *The musical artistry of rap*. Jefferson, MO: McFarland.
- Cook, N. D. (2002). *Tone of voice and mind: The connections between intonation, emotion, cognition and consciousness*. Amsterdam, Netherlands: John Benjamins.

- Dobrian, C. (2014). Timing in MIDI files. May 19, 2014. Retrieved from <https://sites.uci.edu/camp2014/2014/05/19/timing-in-midi-files/>
- Edwards, P. (2013). *How to rap 2: Advanced flow and delivery techniques*. Chicago, IL: Chicago Review Press.
- Flash Music Games (2018). Binary MIDI file to text conversion [Computer program]. Retrieved from <http://flashmusicgames.com/midi/mid2txt.php>
- Foreman, C. (2000). Identification of African-American English from prosodic cues. *Texas Linguistic Forum*, 43, 57–66.
- Forman, M. (2000). “Represent”: Race, space and place in rap music. *Popular Music*, 19(1), 65–90.
- Forman, M. (2002). *The ‘hood comes first: Race, space, and place in rap and hip hop*. Middletown, CT: Wesleyan University Press.
- Gilbers, S. (2018). The linguistic and lyrical development of 2Pac in relation to regional hip-hop identity and conflict. In A. S. Ross & D. J. Rivers (Eds.), *The sociolinguistics of hip-hop as critical conscience: Dissatisfaction and dissent*. Houndmills, UK: Palgrave Macmillan.
- Gilbers, S., Hoeksema, N., De Bot, K., & Lowie, W. (2019). Regional variation in West and East Coast African-American English prosody and rap flows. Retrieved from <https://osf.io/dzc3s/>
- Gooden, S. (2009). Authentically black, bona fide Pittsburgher: A first look at intonation in African American women’s language in Pittsburgh. In S. L. Lanehart (Ed.), *African American women’s language: Discourse, education, and identity* (pp. 142–164). Newcastle-upon-Tyne, UK: Cambridge Scholars.
- Goodwin, M. H., Goodwin, C., & Yaeger-Dror, M. (2002). Multi-modality in girls’ game disputes. *Journal of Pragmatics*, 34, 1621–1649.
- Green, L. J. (2002). *African American English: A linguistic introduction*. Cambridge, UK: Cambridge University Press.
- Hall, R. A., Jr. (1953). Elgar and the intonation of British English. *The Gramophone*, June, 6–7 (reprinted in *Intonation: Selected readings*, pp. 282–285, by D. Bolinger (Ed.), 1972, Hammondsworth, UK: Penguin).
- Hansen, N. C., Sadakata, M., & Pearce, M. (2016). Non-linear changes in the rhythm of European art music: Quantitative support for historical musicology. *Music Perception*, 33(4), 414–431.
- Hart, J. ‘t (1976). Psychoacoustic backgrounds of pitch contour stylization. *I.P.O. Annual Progress Report*, 11, 11–19.
- Hess, M. (2009). Introduction: ‘It’s only right to represent where I’m from’: Local and regional hip hop scenes in the United States. In M. Hess (Ed.), *Hip hop in America: A regional guide* (pp. vii–xxix). Santa Barbara, CA: Greenwood Publishing Group.
- Holliday, N. R. (2016). *Intonational variation, linguistic style and the black/biracial experience* (Doctoral dissertation). New York University, USA.
- Hudson, A. I., & Holbrook, A. (1981). A study of reading fundamental vocal frequency of young black adults. *Journal of Speech and Hearing Research*, 24(2), 197–201.
- Hudson, A. I., & Holbrook, A. (1982). Fundamental frequency characteristics of young black adults: Spontaneous speaking and oral reading. *Journal of Speech and Hearing Research*, 25(1), 25–28.
- Huron, D., & Ollen, J. (2003). Agogic contrast in French and English themes: Further support for Patel and Daniele (2003). *Music Perception*, 21(2), 267–271.
- Jekiel, M. (2014). Comparing rhythm in speech and music: The case of English and Polish. In *Yearbook of the Poznań Linguistic Meeting*, 1, 55–71.
- Jones, N., & Martin, C. (1994). N.Y. state of mind [Recorded by Nas]. On *Illmatic* [CD]. Columbia Records.
- Jun, S.-A., & Foreman, C. (1996). Boundary tones and focus realization in African-American intonation. Paper presented at the 3rd Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, Hawaii, December 2–6, 1996.
- Kautny, O. (2015). Lyrics and flow in rap music. In J. A. Williams (Ed.), *The Cambridge companion to hip-hop* (pp. 101–117). Cambridge, UK: Cambridge University Press.
- Kelley, A. (2011). “A melody? Word.” *The Duke Chronicle*, April 1, 2011. Retrieved from <https://www.dukechronicle.com/article/2011/04/melody-word>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163.



- Krims, A. (2000). *Rap music and the poetics of identity*. Cambridge, UK: Cambridge University Press.
- Lee, C., Brown, L., & Müllensiefen, D. (2017). The musical impact of multicultural London English (MLE) speech rhythm. *Music Perception*, 34(4), 452–481.
- Loman, B. (1975). Prosodic patterns in a negro American dialect. In H. Ringbom, A. Ingberg, R. Norrman, K. Nyholm, R. Westman, & K. Wikberg (Eds.), *Style and text: Studies presented to Nils Erik Enkvist* (pp. 219–242). Stockholm, Sweden: Språkförlaget Skriptor AB.
- Low, E. L., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*, 43(4), 377–401.
- Markman, R. (2013). 20 years later, Snoop Dogg has ‘never’ listened to *Doggystyle*. *MTV News*, November 22, 2013. Retrieved from <http://www.mtv.com/news/1717998/snoop-dogg-doggystyle-20-years-never-listened/>
- McGowan, R. W., & Levitt, A. G. (2011). A comparison of rhythm in English dialects and music. *Music Perception*, 28(3), 307–314.
- McLarty, J. A. (2011). An apparent time analysis of intonation using ex-slaves: Comparing Raleigh, North Carolina AAE and EAE pitch accent types and frequencies (Master’s capstone project). North Carolina State University, USA.
- Mertens, P. (2004). Un outil pour la transcription de la prosodie dans les corpus oraux [A tool for the transcription of prosody in oral corpora]. *Traitement Automatique des Langues*, 45, 109–130. [In French.]
- Morgan, M. (1993). Hip-hop hooray! The linguistic production of identity. Paper presented at the 92nd Annual Meeting of the American Anthropological Association, Washington, DC, November 17–21, 1993.
- Morgan, M. (1998). More than a mood or an attitude: Discourse and verbal genres in African American English. In S. Mufwene (Ed.), *African American English* (pp. 251–280). New York, NY: Routledge.
- Morgan, M. (2001). “Nuthin’ but a G thang”: Grammar and language ideology in hip hop identity. In S. L. Lanehart (Ed.), *Sociocultural and historical contexts of African American English* (pp. 187–209). Amsterdam, Netherlands: John Benjamins.
- Morgan, M. (2002). *Language, discourse and power in African American culture*. Cambridge, UK: Cambridge University Press.
- Native Instruments (2015). Maschine [Computer program]. Version 2.3. Berlin.
- Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, 87(1), B35–45.
- Patel, A. D., Iversen, J. R., & Rosenberg, J. C. (2006). Comparing the rhythm and melody of speech and music: The case of British English and French. *Journal of the Acoustical Society of America*, 119(5), 3034–3047.
- Repp, B. H. (1999). Control of expressive and metronomic timing in pianists. *Journal of Motor Behavior*, 31(2), 145–164.
- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, 12(6), 969–992.
- Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll, & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371–388). New York, NY: Oxford University Press.
- Segalowitz, N., Poulsen, C., & Segalowitz, S. (1999). RT coefficient of variation is differentially sensitive to executive control involvement in an attention-switching task. *Brain & Cognition*, 40(1), 255–258.
- Setaro, S., Kross, J., & Griffin, W. M. Jr. (2014). Episode 65: Rakim. *The Cipher Show*, April 1, 2014. Podcast retrieved from <http://theciphershow.com/episode/65/>
- Schneider, E. W. (1996). Introduction: Research trends in the study of American English. In E. W. Schneider (Ed.), *Focus on the USA* (pp. 1–12). Amsterdam, Netherlands: John Benjamins.
- Tarone, E. E. (1973). Aspects of intonation in Black English. *American Speech*, 48(1), 29–36.
- Temperley, D. (2017). Rhythmic variability in European vocal music. *Music Perception*, 35(2), 193–199.
- The best rappers from Los Angeles. (n.d.) Retrieved from <https://www.ranker.com/list/best-los-angeles-rappers/ranker-hip-hop>
- The best New York rappers. (n.d.) Retrieved from <https://www.ranker.com/list/best-new-york-rappers/ranker-hip-hop>



- Thomas, E. R. (1999). A first look at African-American Vernacular English intonation. Paper presented at New Ways of Analyzing Variation XXVIII, Toronto, Canada, October 16, 1999.
- Thomas, E. R. (2015). Prosodic features of African American English. In J. Bloomquist, L. J. Green, & S. L. Lanehart (Eds.), *The Oxford handbook of African American language* (pp. 420–438). Oxford: Oxford University Press.
- Thomas, E. R., & Carter, P. M. (2006). Rhythm and African American English. *English World-Wide*, 27(1), 331–355.
- Thomas, E. R., Lass, N. J., & Carpenter, J. (2010). Identification of African American speech. In D. R. Preston & N. Niedzielski (Eds.), *A reader in sociophonetics* (pp. 266–285). New York, NY: De Gruyter Mouton.
- Van Handel, L., & Song, T. (2010). The role of meter in compositional style in 19th-century French and German art song. *Journal of New Music Research*, 39(1), 1–11.
- Wallace, C., Franklin, R., Caldwell, B., Eaves, H., III, & Williams, J. (1997). Sky's the limit [Recorded by The Notorious B.I.G. & 112]. On *Life after death* [CD]. Bad Boy Records & Arista Records.
- Wolfram, W. (2007). Sociolinguistic folklore in the study of African American English. *Language and Linguistic Compass*, 1(4), 292–313.
- Wolfram, W., & Kohn, M. E. (2015). Regionality in the development of African American English. In J. Bloomquist, L. J. Green, & S. Lanehart (Ed.), *The Oxford handbook of African American language* (pp. 140–160). Oxford, UK: Oxford University Press.
- Wolfram, W., & Thomas, E. R. (2002). *The development of African American English*. Malden, MA: Blackwell.
- Young, A., Broadus, C. Jr., & Curry, T. L. (1992). Nuthin' but a G thang [Recorded by Dr. Dre & Snoop Doggy Dogg]. On *The chronic* [CD]. Death Row Records, Interscope Records, & Priority Records.

#### Appendix. Information regarding the interviews selected for prosodic analysis.

Subject	Description	Source	Year
Daz Dillinger	Vlad TV interview	<a href="https://www.youtube.com/watch?v=Tc_sjmgRgVA">youtube.com/watch?v=Tc_sjmgRgVA</a>	2015
Dr. Dre	Rhyme & Reason documentary	<a href="https://www.youtube.com/watch?v=DOsLMpphQ_M">youtube.com/watch?v=DOsLMpphQ_M</a>	1997
Ice Cube	Onrust (Dutch TV) interview	<a href="https://www.youtube.com/watch?v=KqFclBhYipM">youtube.com/watch?v=KqFclBhYipM</a>	1991
Kendrick Lamar	GQ interview	<a href="https://www.youtube.com/watch?v=4IPD5PtqMiE">youtube.com/watch?v=4IPD5PtqMiE</a>	2016
Nipsey Hussle	Big Boy's Neighborhood (Real 92.3) radio interview	<a href="https://www.youtube.com/watch?v=IE3rTJi-54k">youtube.com/watch?v=IE3rTJi-54k</a>	2015
Snoop Dogg	The Arsenio Hall Show TV interview	<a href="https://www.youtube.com/watch?v=O24JLNincTM">youtube.com/watch?v=O24JLNincTM</a>	1994
The Game	Big Boy's Neighborhood (Real 92.3) radio interview	<a href="https://www.youtube.com/watch?v=6QqyrHZIbWc">youtube.com/watch?v=6QqyrHZIbWc</a>	2016
Warren G	Ebro in the Morning (Hot 97) radio interview	<a href="https://www.youtube.com/watch?v=0blMD48aR0s">youtube.com/watch?v=0blMD48aR0s</a>	2015
50 Cent	MTV News interview	<a href="https://www.youtube.com/watch?v=3eSpL-qfok">youtube.com/watch?v=3eSpL-qfok</a>	2005
Biggie Smalls	Interview (original source unknown)	<a href="https://www.youtube.com/watch?v=Lv8PdI477ks">youtube.com/watch?v=Lv8PdI477ks</a>	1995
Fabulous	The Breakfast Club (Power 105.1 FM) radio interview	<a href="https://www.youtube.com/watch?v=qxjqtiQrjNk">youtube.com/watch?v=qxjqtiQrjNk</a>	2016
Jay-Z	The Fresh Air (NPR) interview	<a href="https://www.npr.org/2017/06/16/533216823/jay-z-the-fresh-air-interview">npr.org/2017/06/16/533216823/jay-z-the-fresh-air-interview</a>	2010
Mos Def	Live interview (92nd Street Y Plus)	<a href="https://www.youtube.com/watch?v=1kwIulMRlAk">youtube.com/watch?v=1kwIulMRlAk</a>	2014
Nas	BET Presents (BET) interview	<a href="https://www.youtube.com/watch?v=2bCf-60IHMo">youtube.com/watch?v=2bCf-60IHMo</a>	2003
Raekwon	Sway in the Morning (Shade 45) interview	<a href="https://www.youtube.com/watch?v=cxMOXK2vdI4">youtube.com/watch?v=cxMOXK2vdI4</a>	2015
Rakim	The Cipher Show podcast interview	<a href="https://theciphershow.com/episode/65/">theciphershow.com/episode/65/</a>	2014